T.C.

YAŞAR UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

MASTER THESIS

NONSMOOTH ANALYSIS IN SWITCHING CONTROL PROBLEM

Merve ŞENGÜL

Supervisor

Assist. Prof. Dr. Shahlar MAHARRAMOV

İzmir, 2011

T.C.

YAŞAR UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

MASTER THESIS

NONSMOOTH ANALYSIS IN SWITCHING CONTROL PROBLEM

Merve ŞENGÜL

Supervisor

Assist. Prof. Dr. Shahlar MAHARRAMOV

İzmir, 2011

YEMİN METNİ

Yüksek lisans tezi olarak sunduğum “Nonsmooth Analysis in Switching Control Problem” adlı çalışmamın tarafımdan bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım kaynakların kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

25/05/2011

Merve ŞENGÜL

TEŞEKKÜR

Yüksek lisans tezimi hazırlarken bana sabırla rehberlik eden ve desteğini hiçbir zaman eksik etmeyen danışman hocam Yrd. Doç. Dr. Shahlar MAHARRAMOV'a, lisans ve yüksek lisans eğitimimde ders aldığım tüm hocalarıma, hazırlık aşamasında destek olan tüm sevdiklerime ve her zaman yanımda olan canım babam Ömer Faruk ŞENGÜL, canım annem Sebahat ŞENGÜL ve canım abim Kemal ŞENGÜL'e sonsuz teşekkürler.

Merve ŞENGÜL

ABSTRACT

Master Thesis

NONSMOOTH ANALYSIS IN SWITCHING CONTROL PROBLEM

Merve ŞENGÜL

Yaşar University

Institute of Natural and Applied Sciences

In this thesis we study some properties of exhausters, quasidifferential and Frechet superdifferential and their applications to the switching control problem and discrete control problem.

We also consider the necessary optimality condition via exhauster, quasidifferential and Frechet superdifferential for the continuous switching control problem and necessary optimality condition for discrete optimal control problem with nonsmooth data (basic subdifferential).

In this way, we use the knowledge of the nonsmooth analysis. By using the increment formula we obtain necessary optimality conditions for the switching control problem. The minimizing functional satisfying nonsmoothness properties. The obtained optimality condition is an analog of the Pontryagin maximum principle for the switching control problem.

Keywords: Exhauster, Quasidifferential, Frechet Superdifferential, Pontryagin Maximum Principle

CONTEXT

INTRODUCTION

The thesis consists of three sections.

In the first section we consider necessary optimality condition for the switching optimal control problem. The problem in this section is same as the problem that we consider in the last section but in this case minimizing functional satisfying Frechet superdifferential condition.

Switching versions of the maximum principle have been presented in [13, 35, 40] and [48]. A dynamic programming approach for hybrid systems and special issue on hybrid system are discussed in [1, 2]. In [10, 23], a computational method for solving an optimal control problem, governed by a switched dynamical system with time delay and control parameterizations for optimal control of switching system, are developed. The approach is to parameterize the switching instants as a new parameter vector to be optimized. Then, the gradient of the cost function is obtained via solving a number of delay differential equations forward in time. On this basis, the optimal control problem can be solved as a mathematical programming problem. In [24] and [25], discrete switched control problems have been studied. All these articles consider smooth hybrid optimal control problem. The nonsmooth version of the hybrid optimal control problem has not been studied extensively. To our best knowledge, there is only one article which considers the nonsmooth version of the hybrid maximum principle, namely the paper [48]. In this paper, the author obtains the nonsmooth version of the hybrid maximum principle by using "Boltyanskii approximation cone" (By using this method, smooth version of the hybrid maximum principle was obtained by Boltyanskii in [5]). In [48], the author assumes the switching cost and endpoint functionals are nonsmooth. He applies generalized gradients and proves the hybrid maximum principle. Then the author extends this principle for the semidifferentiable switching and endpoint functionals. He also notes that this can be proved by using the Warga's generalized derivative. However, this paper does not consider the hybrid maximum principle using Frechet upper subdifferential. (for the definition of Frechet upper and lower subdifferentials see, for example, [33]).

The second section of the thesis is dedicated to the nonsmooth optimal control problems governed by discrete-time systems with the delays in state variables. Problems of this type arise in variational analysis of delay-differential systems via discrete approximations (cf. [30, 31] and their predecessors for non-delayed systems in [39] and [28, 29]). They are important for many applications, especially to economic modelling, to qualitative and numerical aspects of optimization and control of various hereditary processes, to numerical solutions of control systems with distributed parameters, etc. (see, e.g., [4, 11, 30, 37, 49] and the references there in). Note that delayed discrete systems may be reduced to non-delayed ones of a bigger dimension by a multi-step procedure and that they both can be reduced to finite-dimensional mathematical programming. Nevertheless, optimal control problems of type (P) deserve a special attention in order to obtain results that take into account their particular dynamic structure and the influence of delays on the process of dynamic optimization.

It is well known that, while for continuous-time systems optimal controls satisfy the Pontrjagin maximum principle without restrictive assumptions [36], its discrete analog (the discrete maximum principle) does not generally hold unless a certain convexity is imposed a priori on the control system (see, e.g., [4, 19, 21, 37] and their references). A clear explanation of this phenomenon is given in Section 5.9 of Pshenichnyi's book [38] (the first edition), where it is shown why discrete systems require a convexity assumption for the validity of the maximum principle, while continuous-time systems enjoy it automatically due to the so-called "hidden convexity". The relationships between convexity and the maximum principle are transparent from the viewpoint of nonsmooth analysis due to the special nature of the normal cone to convex sets (cf. [39] and [28]).

The goal of this section is to derive the necessary optimality conditions in the form of the discrete maximum principle for problem (P) and some of its generalizations. Our standing assumption is that $f = f(t,x,y,u)$ is continuous with respect to all variables but t and continuous differentiable with respect to the state variables (x,y) for all $t \in T$ and $u \in U$ near the optimal solution under consideration. We do not assume any smoothness of the cost function φ and derive new versions of the discrete maximum principle with transversality conditions taking into account the nonsmoothness of φ. A striking result obtained in this thesis, new for both delayed and non-delayed systems, is the superdifferential form of the discrete maximum principle, where the transversality condition is expressed in terms of the so-called

Frechet superdifferential. This is a rather surprising result, since it applies to minimization problems for which subdifferential forms of necessary optimality conditions are more conventional. We also obtain the discrete maximum principle for nonsmooth problems with transversality conditions of subdifferential type, which extend known results to the case of delayed systems. We will discuss the relationships between the superdifferential and subdifferential forms of the discrete maximum principle: they are generally independent, while the superdifferential one may be considerably stronger in some situations when it applies.

In last, third, section we consider optimal control for switching system in the case of minimizing functional satisfying quasidifferential and exhauster conditions in the Demyanov and Rubinov sense.

A switched system is a particular kind of hybrid system that consists of several subsystems and a switching law specifying the active subsystem at each time instant. Examples of switched systems can be found in chemical processes, automotive systems, and electrical circuit systems, etc. Recently, optimal control problems of hybrid and switched systems have been attracting researchers from various fields in science and engineering, due to problems significance in the theory and application. The available results in the literature on such problems can be classified into two categories, i.e., theoretical and practical. [35, 6, 48, 10, 24, 25, 26, 7, 5] contain primarily theoretical results. These results extended the classical maximum principle or the dynamic programming approach to such problems. Among them, earliest results which proves a maximum principle for hybrid system with autonomous switchings by Seidman in [46]. More complicated versions of the maximum principle under various additional assumptions are proved by Sussmann in [48] and by Piccoli in [35]. All these article dedicate to the smooth switching optimal control problem (only Sussmann's article [48], it is studied switching system which minimizing functional and constraints are satisfying the generalized differentation). In the last section of the presented thesis the author's aim to establish necessary optimality condition by using exhausters and quasidifferentiable in the sense of Demyanov and Rubinov [14, 15]. We consider minimizing functional is positively homogeneous (p.h). Positively homogeneous (p.h) functions play on outstanding role in Nonsmooth Analysis and Nondifferentiable Optimization since (first-order) optimality conditions are normally expressed in terms of directional derivatives of their generalizations (the Dini and Hadamard upper and lower direcitonal derivatives, the Clarke derivative, the Michael-Penot derivative etc.). All these derivatives are positively homogeneous functions of direction. In the convex case the directional derivative is convex (and p.h), by the Minkovwski duality, optimality conditions can be stated in geometric terms. Attempts to reduce the problem of minimizing an arbitrary function to a sequence of convex problems were undertaken, among others by Pschenichnyi [39], who introduced the notations of upper convex and lower concave approximations and by Clarke [12], who introduced generalized derivatives. Demyanov and Rubinov [14] proposed to consider exhaustive families of upper convex and lower concave approximations. The last section addresses to learn role exhausters and quasidifferentiability in the switching control problem.

1. NECESSARY OPTIMALITY CONDITIONS FOR SWITCHING CONTROL PROBLEMS

1.1 Preliminaries

We recall some definitions from nonsmooth analysis which will be applied to find the superdifferential from the necessary optimality condition for the step discrete system.

Given a nonempty set $\Omega \subset R^n$, consider the associated distance fuction

$$\mathrm{dist}(x;\Omega) = \inf_{\omega\in\Omega} \|x - \omega\|$$

and define Euclidean projector of x onto Ω by

$$\Pi(x;\Omega) := \{\omega \in \Omega \mid \|x-\omega\| = \mathrm{dist}\,(x;\Omega)\}.$$

The set $\Pi(x;\Omega)$ is nonempty for every $x \in R^n$ if the set Ω is closed and bounded. The normal cone in finite dimensional spaces is defined by using the Euclidean projector:

$$\mathrm{N}(\bar{x};\Omega) := \limsup_{x\to\bar{x}} [cone(x - \Pi(x;\Omega))],$$

while the basic subdifferential $\partial\varphi(\bar{x})$ is defined geometrically via the normal cone to the epigraph of φ is a real valued finite function,

$\partial\varphi(\bar{x}) := \{x^* \in \mathrm{R}^n \mid (x^*,-1) \in \mathrm{N}((\bar{x},\varphi(\bar{x})); epi\varphi\}$ and

$epi\varphi := \{(x;\mu) \in \mathrm{R}^{n+1} \mid \mu \geq \varphi(x)\}$ is the epigraph of φ. This nonconvex cone to closed sets and corresponding subdifferential of lower semicontinuous extended real-valued functions were introduced in [33, 32]. Note that this cone is nonconvex (see [25, 33, 32]) and for the locally lipschitz functions convex hull of a subdifferential is a Clarke generalized subdifferential;

$\overline{\varphi}_k(x^0) = co\partial\varphi(x^0)$ (here $\overline{\varphi}_k(x^0)$ is Clarke generalized subdifferential [12, 42]). If φ_k is lower semicontinuous in some neighborhood of x, then its basic subdifferential can be expressed as: $\partial\varphi(x^0) := \limsup_{x\to x^0} \partial\varphi(x)$.

Here,

$$\hat{\partial}\varphi(x^0) := \left\{ x^* \in \mathrm{R}^n : \liminf_{u\to x^0} \frac{\varphi(u) - \varphi(x^0) - \langle x^*, u - x^0 \rangle}{|u - x^0|} \geq 0 \right\}$$

is the Frechet subdifferential. By using plus-minus symmetric constructions, we can write

$$\partial^+\varphi(x^0) := -\partial(-\varphi)(x^0), \ \hat{\partial}^+\varphi(x^0) := -\hat{\partial}(-\varphi)(x^0)$$

which are called basic superdifferential and Frechet superdifferential, respectively. Here

$$\hat{\partial}\varphi^+(x^0) := \left\{ x^* \in \mathrm{R}^n : \limsup_{u\to x^0} \frac{\varphi(u) - \varphi(x^0) - \langle x^*, u - x^0 \rangle}{|u - x^0|} \leq 0 \right\}$$

For a locally Lipschitz function subdifferential and superdifferential may be different. For example, if we take $\varphi(x) = |x|$ on R, then $\partial\varphi(0) = [-1, 1]$, but $\partial^+\varphi(0) = \{-1, 1\}$. If φ is locally Lipschitz continuous at a point x^0 then the strict differentiability of the function φ at x^0 (see [26]) is equivalent to

$\partial\varphi(x^0) = \partial^+\varphi(x^0) = \{\nabla\varphi(x^0)\}$. If $\partial\varphi(x^0) = \hat{\partial}\varphi(x^0)$ then this function is lower regular at x^0. Symmetrically, we can define upper regularity of the function using the superdifferential and Frechet superdifferential. Also, if we extended real-valued function is locally Lipschitz and upper regular at a given point, then its Frechet superdifferential is not empty at this point. Furthermore, it is equal to Clarke generalized subdifferential at this point. In this thesis we will use the following theorem.

Theorem 1.1.1. ([33]) Let $\varphi : X \to \overline{R}$ be a proper function . Assume that φ is finite at a point $\overline{x}$. Then for every $x^* \in \hat{\partial}\varphi(\overline{x})$ there is a function $s: X \to R$ with $s(\overline{x}) = \varphi(\overline{x})$ and whenever $x \in X$ such that s(.) is Frechet differentiable at $\overline{x}$ with $\nabla s(\overline{x}) = x^*$.

1.2 Problem formulation

We consider the following optimization problem

$$\dot{x}_K(t) = f_K(x_K(t), u_K(t), t), \quad t \in [t_{K-1}, t_K], \quad K = 1, 2, ..., N \tag{1.1}$$

$$x_1(t_0) = x_0 \tag{1.2}$$

$$F_K(x_N(t_N), t_N) = 0, \quad K = 1, 2, ..., N \tag{1.3}$$

$$x_{K+1}(t_K) = M_K(x_K(t_K), t_K), \quad K = 1, 2, ..., N-1 \tag{1.4}$$

$$\begin{aligned} &\min S(u_1, ..., u_N, t_1, ..., t_N) \\ &= \sum_{K=1}^{N} \varphi_K(x_K(t_K)) + \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} L(x_K, u_K, t)dt \end{aligned} \tag{1.5}$$

Here $f_K : R \times R^n \times R^r \to R^n$, M_K and F_K are continuous, at least continuously partially differentiable vector-valued functions with respect to their variables, $L : R^n \times R^r \times R \to R$ is continuous and have continuous partial derivative with respect to their variables, $M_K : R^n \times R \to R$ and $\varphi_K(.)$ are given differentiable functions, $u_K(t) : R \to U_K \subset R^r$ are controls. The sets U_K are assumed to be nonempty and bounded. Here (1.4) are switching conditions. It is required to find the control $u_1, u_2, ..., u_K$, switching points $t_1, t_2, ..., t_{N-1}$ and the end point t_N (here t_N is not fixed) with corresponding state $x_1, x_2, ..., x_N$ satisfying (1.1)-(1.4) so that the fuction S(.) in (1.5) is minimized. We will derive necessary conditions for smooth and nonsmooth version of these problems (in the case of smooth and nonsmooth cost functionals).

Denote:

$$\theta = (t_1, t_2, ..., t_N),\ x(t) = (x_1(t), x_2(t), ..., x_N(t)),$$

$$u(t) = (u_1(t), u_2(t), ..., u_N(t)).$$

Our aim is to find tuple $(x(t), u(t), \theta)$ which solves problem (1.1)-(1.5). Such tuple will be called optimal control for the problem (1.1)-(1.5).

Theorem 1.2.1. Let the $(x(t), u(t), \theta)$ be an optimal control for Problem (1.1)-(1.5). Then there are vector functions $p_K(t),\ K = 1, 2, ..., N$ such that following conditions hold.

1) State equation.

$$\dot{x}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial p_K}$$
$$t \in [t_{K-1}, t_K],\quad K = 1, 2, ..., N$$

2) Costate equation.

$$\dot{p}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial x_K}$$
$$t \in [t_{K-1}, t_K],\quad K = 1, 2, ..., N$$

3) At t_N, the function $p_N(.)$ satisfies

$$p_N(t_N) = \frac{\partial \varphi_N(x_N(t_N))}{\partial x_N} + \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial x_N}$$

4) Necessary conditions

$$\max_{u_K \in U_K} H_K(x_K^0, u_K(t), p_K(t), t) = H_K(x_K^0, u_K^0, p_K(t)t), \quad K = 1, 2, ..., N$$
$$t \in [t_{K-1}, t_K]$$

5) Necessary conditions at the switching points

$$p_K(t_K) = \frac{\partial \varphi_K(t_K)}{\partial x_K} - p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial x_K}, \quad K = 1, 2, ..., N-1$$

$$\left(\sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial t_N}\right)\delta_{L,N}$$
$$-\left(\sum_{K=1}^{N-1} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial t_K}\right) \times \left(1 - \delta_{L,N}\right) = 0.$$

$$\textit{Here } \delta_{L,N} = \begin{cases} 1, & L = N \\ 0, & L \neq N \end{cases}, \mathrm{L} = 1, 2, ..., N,$$
$$H_K(x_K, u_K, p_K, t) = L_K(x_K, u_K, p_K, t) + p_K^T . f_K(x_K, u_K, p_K, t)$$
and λ_K, $K = 1, 2, ..., N$ *are vectors.*

Proof. We use Lagrange multipliers to adjoint the state and conjugate equations in the theorem. Then, by using Lagrange multipliers rule, we can write

$$S' = \sum_{K=1}^{N} \varphi_K(x_K(t_K)) + \sum_{K=1}^{N} \lambda_K F_K(x_N(t_N), t_N)$$
$$+ \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} (L(x_K, u_K, t) + p_K^T(t)(f_K(x_K, u_K, t) - \dot{x}_K(t)))dt$$

By determining

$$H_K(x_K, p_K, u_K, t) = L_K(x_K, u_K, t) + p_K(t) f_K(x_K, u_K, t) \text{ for } t \in [t_{K-1}, t_K]$$

we have:

$$S^{'} = \sum_{K=1}^{N} \varphi_K(x_K(t_K)) + \sum_{K=1}^{N} \lambda_K F_K(x_N(t_N), t_N)$$
$$+ \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} (H_K(x_K, p_K, u_K, t) + p_K^T x_K(t))dt$$

From the calculus of variations, we can obtain that the first variation of $\delta S^{'}$ as:

$$\delta S^{'} = \sum_{K=1}^{N} \frac{\partial \varphi_K(x_K(t_K))}{\partial x_K} \delta x_K(t_K) + \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial x_N} \delta x_N(t_N)$$
$$+ \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_K)}{\partial t_N} \delta t_N + \sum_{K=1}^{N} \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial x_K} \delta x_K(t)$$
$$+ \sum_{K=1}^{N} \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial u_K} \delta u_K + \sum_{K=1}^{N} \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial p_K} \delta p_K$$
$$- \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} p_K(t) \dot{x}_K(t) dt$$

The latter term in previous equation can be computed as follows;

$$\sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} p_K(t) dx_K(t) = \sum_{K=1}^{N} p_K(t_K) x_K(t_K) - (p_K(t_{K-1}) x_K(t_{K-1}))$$
$$- \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} \dot{p}_K(t) \delta x_K dt$$
$$= \sum_{K=1}^{N} p_K(t_K) \delta x_K(t_K) - \sum_{K=1}^{N-1} p_{K+1}(t_K) \delta x_{K+1}(t_K)$$
$$- p_1(t_0) \delta x_1(t_0) - \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} \dot{p}_K(t) \delta x_K dt$$

Here we have taken into account:

$$\sum_{K=1}^{N} p_K(t_{K-1}) x_K(t_{K-1}) = \sum_{K=1}^{N-1} p_{K+1}(t_K) x_{K+1}(t_K) + p_1(t_0) x_1(t_0)$$

Since $\delta x_1(t_0)=0$ using (1.4) we get

$$\sum_{K=1}^{N} p_{K+1}(t_K)\left(\frac{\partial M_K(x_K(t_K),t_K)}{\partial x_K}\delta x_K(t_K)+\frac{\partial M_K(x_K(t_K),t_K)}{\partial t_K}\delta t_K\right)$$
$$=\sum_{K=1}^{N-1} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial x_K}\delta x_K(t_K)$$
$$+\sum_{K=1}^{N-1} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial t_K}\delta t_K$$

Then, first variation has the following form;

$$\delta S'=\sum_{K=1}^{N-1}\frac{\partial s_K(x_K(t_K))}{\partial x_K}\delta x_K(t_K)+\frac{\partial s(x_N(t_N))}{\partial x_N}\delta x_N(t_N)$$
$$+\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N),t_N)}{\partial x_N}\delta x_N(t_N)+\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N),t_N)}{\partial t_N}\delta t_N$$
$$+\sum_{K=1}^{N}\frac{\partial H_K(u_K,x_K,p_K,t)}{\partial u_K}\delta u_K+\sum_{K=1}^{N}\frac{\partial H_K(u_K,x_K,p_K,t)}{\partial p_K}\delta p_K$$
$$-\sum_{K=1}^{N-1} p_K(t_K)\delta x_K(t_K)-p_N(t_N)\delta x_N(t_N)$$
$$-\sum_{K=1}^{N-1} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial x_K}\delta x_K(t_K)$$
$$-\sum_{K=1}^{N-1} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial t_K}\delta t_K-\sum_{k=1}^{N} p_k(t)\delta x_k(t_k)$$
$$=\sum_{K=1}^{N-1}\left(\frac{\partial s_K(x_K(t_K))}{\partial x_K}-p_K(t_K)-p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial x_K}\right)\delta x_K(t_K)$$
$$+\left(\frac{\partial s_N(x_N(t_N))}{\partial x_N}+\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N),t_N)}{\partial x_N}-p_N(t_N)\right)\delta x_N(t_N)$$
$$\sum_{L=1}^{N}\left[\left(\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N),t_N)}{\partial t_N}\right)\delta_{L,N}-\left(\sum_{K=1}^{N-1}p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K),t_K)}{\partial t_K}\right)\left(1-\delta_{L,N}\right)\right]\delta t_N$$
$$+\sum_{K=1}^{N}\left(\frac{\partial H_K(x_K,p_K,u_K,t)}{\partial x_K}-\dot{p}_K(t)\right)\delta x_K+\sum_{K=1}^{N}\frac{\partial H_K(x_K,p_K,u_K,t)}{\partial u_K}\delta u_K$$
$$+\sum_{K=1}^{N}\left(\frac{\partial H_K(x_K,p_K,u_K,t)}{\partial p_K}-\dot{p}_K(t)\right)\delta p_K$$

The latter sum is known because

$$\frac{\partial H_K(x_K, u_K, p_K, t)}{\partial p_K} = \dot{x}_K(t).$$

According to a necessary condition for an optimal solution $\delta S' = 0$. Setting to zero coefficients of the independent increments $\delta x_N(t_N), \delta x_K(t_K), \delta x_K, \delta u_K$, and δp_K yields the necessary optimality condition in the following form

$$\dot{x}_K(t) = \frac{\partial H_K(u_K, x_K, p_K, t)}{\partial p_K}$$

$$\dot{p}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial x_K}$$

$$\frac{\partial H_K(u_K, x_K, p_K, t)}{\partial u_K} = 0$$

$$\frac{\partial \varphi_K(t_K)}{\partial x_K} - p_K(t_K) - p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial x_K} = 0, \quad K = 1, 2, ..., N-1$$

$$\frac{\partial \varphi_N(t_N)}{\partial x_N} + \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial x_N} - p_N(t_N) = 0$$

$$\frac{\partial \varphi_K(t_K)}{\partial x_K} - p_K(t_K) - p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial x_K} = 0, \quad K = 1, 2, ..., N-1$$

$$\left(\sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial t_N}\right)\delta_{L,N} - \left(\sum_{K=1}^{N} p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial t_K}\right)\left(1 - \delta_{L,N}\right) = 0.$$

This completes the proof.

Let us now assume that the objective function $\varphi_K(.)$ is Frechet upper subdifferentiable (superdifferentiable) at the point $\bar{x}_K(t_K)$. Then one can prove the following theorem for the nonsmooth version of problem (1.1)-(1.5).

Theorem 1.2.2. (Nonsmooth version) Let objective function $\varphi_K(.)$ is Frechet upper subdifferentiable (superdifferentiable) at the point $\bar{x}_K(\bar{t}_K)$ *and* $(\bar{u}(t), \bar{x}(t), \bar{\theta})$ be an optimal solution to the control problem (1.1)-(1.5). Then, every collection of Frechet upper subgradients (supergradients) $x_K^* \in \hat{\partial}^+ \varphi_K(\bar{x}_K(\bar{t}_K)),\ K = 1, 2, ..., N$ conditions in Theorem 1.2.1. hold with the corresponding trajectory $p_K(.)$ of the conjugate system, the condition (3) and condition (5) in Theorem 1.2.1. replacing by following conditions:

$$p_K(\bar{t}_K) = x_K^* - p_{K+1}(\bar{t}_K)\frac{\partial M_K(\bar{x}_K(\bar{t}_K), \bar{t}_K)}{\partial x_K}, \quad K = 1, 2, ..., N-1$$

$$p_N(\bar{t}_N) = x_N^* + \sum_{K=1}^{N} \frac{\partial F_K(\bar{x}_N(\bar{t}_N), \bar{t}_N)}{\partial x_N}$$

$$\left(\sum_{K=1}^{N} \frac{\partial F_K(\bar{x}_N(\bar{t}_N), \bar{t}_N)}{\partial t_N}\right)\delta_{L,N} - \left(\sum_{K=1}^{N-1} p_{K+1}(\bar{t}_K)\frac{\partial M_K(\bar{x}_K(\bar{t}_K), \bar{t}_K)}{\partial t_K}\right)\left(1-\delta_{L,N}\right) = 0,$$

$$\text{here } L = 1, 2, ..., N, \quad \bar{\theta} = (\bar{t}_1, \bar{t}_2, ..., \bar{t}_N),$$

$$\bar{x}(t) = (\bar{x}_1(t), \bar{x}_2(t), ..., \bar{x}_N(t)), \quad \bar{u}(t) = (\bar{u}_1(t), \bar{u}_2(t), ..., \bar{u}_N(t)).$$

Proof. Take any arbitrary set of Frechet upper subgradients $x_K^* \in \hat{\partial}^+ \varphi_K(\bar{x}_K(\bar{t}_K)),\ K = 1, 2, ..., N$ and employ the smooth variational description of $-x_K^*$ from assertion (i) of Theorem 1 (see [33]). As a result, we find functions $s_K : X \to R$ for $K = 1, 2, ..., N$ satisfying the relations

$$s_K(\bar{x}_K(\bar{t}_K)) = \varphi_K(\bar{x}_K(\bar{t}_K)), \quad s_K(x_K(t)) = \varphi_K(x(t))$$

in some neighborhood of $\bar{x}_K(t_K)$ and such that each of them Frechet differentiable at $\bar{x}_K(\bar{t}_K)$ with $\nabla s_K(\bar{x}_K(\bar{t}_K)) = x_K^*,\ \ K = 1, 2, ..., N$.

By using construction of these functions we easily deduce that the process $(\bar{u}(.), \bar{x}(.), \bar{\theta})$ is an optimal solution to the following control problem:

$$\min S(u_1, ..., u_N, t_1, ..., t_N) = \sum_{K=1}^{N} s_K(x_K(t_K)) + \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} L(x_K, u_K, t)dt$$

subject to conditions (1.1)-(1.4). The initial data of the latter optimal control problem satisfy all the assumptions of Theorem 1.2.1. Thus, applying the above maximum principle to the problem (1.1)-(1.5) and taking into account that

$$\nabla s_K(\bar{x}_K(\bar{t}_K)) = x_K^*, \quad K = 1, 2, ..., N$$

we complete the proof of the theorem.

Lemma 1.2.3. Let $\varphi : R \to \bar{R}$ be locally Lipschitz continuous at $\bar{x}$ and upper regular at this point. Then Frechet superdifferential is not empty at this point and coinside with the Clarke subdifferential at this point, $0 \neq \hat{\partial}^+ \varphi(\bar{x}) = \bar{\partial}\varphi(\bar{x})$.

Proof. The nonemptiness of $\hat{\partial}^+ \varphi(\bar{x})$ directly follows from $\partial\varphi(\bar{x}) \neq 0$ for locally Lipschitzian functions and the definition of upper regularity. Due to $\bar{\partial}\varphi(\bar{x}) = co\,\varphi(\bar{x})$, any locally Lipschitz function is lower regular at $\bar{x}$ if and only if $\hat{\partial}\varphi(\bar{x}) = \bar{\partial}\varphi(\bar{x})$. Hence, the upper regularity of φ at $\bar{x}$ and the plus-minus symmetry of the generalized gradient imply that $\hat{\partial}^+ \varphi(\bar{x}) = -\hat{\partial}(-\varphi)(\bar{x}) = -\bar{\partial}(-\varphi)(\bar{x}) = \bar{\partial}\varphi(\bar{x})$ which completes the proof.

Corollary 1.2.4. Let $\{\bar{u}_K(.), \bar{x}_K(.), \bar{\theta}\}$ be an optimal solution of the control problem (1.1)-(1.5) and assume that $\varphi_K(.)$ is locally Lipschitz and upper regular at $\bar{x}_K(t_K)$. Then, for any Clarke generalized gradient $x_K^* \in \bar{\partial}\varphi_K(\bar{x}_K(t_K))$ the maximum principle and transversality condition is satisfied.

The proof follows from Theorem 1.2.2. and Lemma 1.2.3.

1.3 Necessary conditions for cost uniformly upper subdifferentiable functionals

In this section we present uniformly upper subdifferential form of the main problem.

Definition 1.3.1. (Uniform upper subdifferentiability). A function $\varphi : R^n \to \overline{R}$ is uniformly upper subdifferentiable at a point $\overline{x}$, if it is finite at this point and there exists a neighborhood V of $\overline{x}$ such that for every $x \in V$ there exists $x^* \in R^n$ with the following property: Given any $\varepsilon \rangle 0$, there exists $\eta \rangle 0$ for which

$$\varphi(v) - \varphi(x) - \left\langle x^*, v - x \right\rangle \leq \varepsilon \|v - x\|$$

whenever $v \in V$ with $\|v - x\| \leq \eta$. It is easy to check that the class of uniformly upper subdifferentiable functions include continuously differentiable functions and concave continuous functions, and also are closed with respect to taking the minimum over compact sets.

It is well known that a function uniformly upper subdifferentiable in some neighborhood of a given point is upper regular, Lipschitz continuous at this point (see [32], Proposition 3.2). Then:

Corollary 1.3.2. Let $\left\{\overline{u}_K(.), \overline{x}_K(.), \overline{\theta}\right\}$ be an optimal solution to Problem (1.1)-(1.5). Assume that φ_K is uniformly upper subdifferentiable in some neighborhood of the point $\overline{x}_K(t_K)$. Then for every upper subgradient $x_K^* \in \hat{\partial}^+ \varphi_K(\overline{x}_K(t_K))$, $K = 1, 2, ..., N$ the maximum condition, transversality conditions and necessary conditions in the switching points are satisfied in Theorem 1.2.2.

Proof. Let φ_K be uniformly upper subdifferentiable in some neighborhood of the point $\bar{x}_K(t_K)$. Then by using Proposition 3.2 ([32]) we can say φ_K is upper regular at $\bar{x}_K$ and Lipschitz continuous at this point. Then, by using Corollary 1.2.4. and Theorem 1.2.2., we can write that, for every upper subgradient $x_K^* \in \partial\varphi_K(\bar{x}_K(t_K))$ where $K = 1, 2, ..., N$ the maximum condition, the transversality condition and necessary conditions at the switching points are satisfied in Theorem 1.2.1.

2. DISCRETE MAXIMUM PRINCIPLE FOR NONSMOOTH OPTIMAL CONTROL PROBLEMS WITH DELAYS

Our notation is basically Standard (see, e.g., [41]).

$$\operatorname*{Limsup}_{x \to \bar{x}} F(x) := \left\{ y \in \mathrm{R}^m \middle| \exists \text{ sequences } x_k \to \bar{x} \text{ and } y_K \to y \text{ with } y_k \in F(x_k) \text{ for all } k \in \mathrm{N} \right\}$$

denotes the Painleve-Kuratowski upper (outer) limit for a set-valued mapping $F : \mathrm{R}^{\mathrm{n}} \underset{\to}{\to} \mathrm{R}^{\mathrm{m}}$ as $x \to \bar{x}$. The expressions

$$\mathrm{cl}\Omega \ , \mathrm{co}\Omega \ , \text{and } \mathrm{cone}\Omega := \left\{ ax \middle| a \rangle 0, \ x \in \Omega \right\}$$

stand for the closure, convex hull, and conic hull of a set Ω, respectively. The notation $x \xrightarrow{\varphi} \bar{x}$ with $\varphi(x) \to \varphi(\bar{x})$

2.1 Tools of nonsmooth anaysis

In this section we review several constructions of nonsmooth analysis and their properties needed in what follows. For more information we refer the reader to [12, 28, 41].

Let Ω be a nonempty set in R^{n}, and let

$$\Pi(x;\Omega) := \left\{ w \in \mathrm{cl}\Omega \text{ with } |x - w| = \mathrm{dist}(x;\Omega) \right\}$$

be the Euclidean projector of x to the closure of Ω. The basic normal cone [3] to Ω at $\bar{x} \in \mathrm{cl}\Omega$ is defined by

$$N(\bar{x};\Omega) := \operatorname*{Lim\,sup}_{x \to \bar{x}} \left[\mathrm{cone}\left(x - \Pi\left(x;\Omega \right) \right) \right]. \tag{2.1}$$

This cone if often nonconvex, and its convex closure agrees with the Clarke normal cone [35].

Given an extended-real-valued function $\varphi : \mathrm{R}^n \to \overline{\mathrm{R}} := [-\infty, \infty]$ finite at $\bar{x}$, we define its basic subdifferential [28] by

$$\partial\varphi(\bar{x}) := \left\{x^* \in \mathrm{R}^n \middle| (x^*, -1) \in N((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi)\right\}, \tag{2.2}$$

Where $\mathrm{epi}\,\varphi := \left\{(x, \mu) \in \mathrm{R}^{n+1} \middle| \mu \geq \varphi(x)\right\}$ stands for the epigraph of φ. If φ is locally Lipschitzian around $\bar{x}$, then $\partial\varphi(\bar{x})$ is a nonempty compact set satisfying

$$(x^*, -\lambda) \in N((\bar{x}, \varphi(\bar{x})); \mathrm{epi}\,\varphi) \Leftrightarrow \lambda \geq 0,\ x^* \in \lambda\partial\varphi(\bar{x}). \tag{2.3}$$

One always has $\bar{\partial}\varphi(\bar{x}) = \mathrm{co}\,\partial\varphi(\bar{x})$ for the Clarke generalized gradient of locally Lipschitzian functions [12]. Note the latter construction, in contrast to (2.2), possesses the classical plus-minus symmetry $\bar{\partial}(-\varphi)(\bar{x}) = -\bar{\partial}\varphi(\bar{x})$. If φ is lower semicontinuous around $\bar{x}$, then the basic subdifferential (2.2) admits the representation

$$\partial\varphi(\bar{x}) = \operatorname*{Lim\,sup}_{x \xrightarrow{\varphi} \bar{x}} \hat{\partial}\varphi(x)$$

in terms of the so-called Frechet subdifferential of φ at x defined by

$$\hat{\partial}\varphi(x) := \left\{x^* \in \mathrm{R}^n \middle| \liminf_{u \to x} \frac{\varphi(u) - \varphi(x) - \langle x^*, u - x\rangle}{|u - x|} \geq 0\right\} \tag{2.4}$$

The symmetric constructions

$$\partial^+\varphi(\bar{x}) := -\partial(-\varphi)(\bar{x}),\ \hat{\partial}^+\varphi(\bar{x}) := -\hat{\partial}(-\varphi)(\bar{x}) \tag{2.5}$$

to (2.2) and (2.4) are called, respectively, tha basic superdifferential and the Frechet superdifferential af φ at $\bar{x}$. Note that

$$\hat{\partial}^{+}\varphi(\bar{x}) := \left\{ x^{*} \in \mathrm{R}^{\mathrm{n}} \middle| \limsup_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^{*}, u - \bar{x} \rangle}{|x - \bar{x}|} \leq 0 \right\} \tag{2.6}$$

and that both $\hat{\partial}\varphi(\bar{x})$ and $\hat{\partial}^{+}\varphi(\bar{x})$ are nonempty simultaneously if and only if φ is Frechet differentiable at $\bar{x}$, in which case they both reduce to the classical (Frechet) derivative of φ at this point:

$$\hat{\partial}\varphi(\bar{x}) = \hat{\partial}^{+}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\} \tag{2.7}$$

In contrast, the basic subdifferential and superdifferential are simultaneously nonempty for every locally Lipschitzian function; they may be essentially different, e.g., for $\varphi(x) = |x|$ on R when $\partial\varphi(0) = [-1,1]$ and $\partial^{+}\varphi(0) = \{-1,1\}$. Note also that if φ is Lipschitz continuous around $\bar{x}$, then

$$\partial\varphi(\bar{x}) = \partial^{+}\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\} \tag{2.8}$$

if and only if φ is strictly differentiable at $\bar{x}$, i.e.,

$$\lim_{\substack{x \to \bar{x} \\ x' \to \bar{x}}} \frac{\varphi(x) - \varphi(x') - \langle \nabla\varphi(\bar{x}), x - x' \rangle}{|x - x'|} = 0$$

which happens, in particular, when φ is continuously differentiable around $\bar{x}$. The singleton relations (2.8) may be violated if φ is just differentiable but not strictly differentiable at x. For example, if $\varphi(x) = x^{2}\sin(1/x)$ for $x \neq 0$ with $\varphi(0) = 0$, then

$$\partial\varphi(0) = \partial^{+}\varphi(0) = [-1,1] \text{ while } \hat{\partial}\varphi(0) = \hat{\partial}^{+}\varphi(0) = \{0\}$$

Recall [3] that φ is lower regular at $\bar{x}$ if $\partial\varphi(\bar{x}) = \hat{\partial}\varphi(\bar{x})$. This happens, in particular, when φ is either strictly differentiable at $\bar{x}$ or convex. Moreover, lower regularity holds for the class of weakly convex functions [34], which includes both smooth and convex functions and is closed with respect to taking the maximum over compact sets. Note that the latter class is a subclass of quasidifferentiable functions is the sense of Pshenichnyi [38].

A large class of lower regular functions (in somewhat stronger sense) has been studied in [41] under the name of amenability. It was shown there that the class of amenable functions enjoys a fairly rich calculus and includes a large core of functions frequently encountered in finite-dimensional minimization.

Symmetrically, φ is upper regular at $\bar{x}$ if $\partial^{+}\varphi(\bar{x}) = \hat{\partial}^{+}\varphi(\bar{x})$. It follows from (2.5) that this property is equivalent to the lower regularity of $-\varphi$ at $\bar{x}$. Thus all the facts about subdifferentials and lower regularity relative to minimization can be symmetrically transferred to superdifferentials and upper regularity relative to maximization. The point is that in the next section we are going to apply superdifferentials and upper regularity relative to maximization problems. The following proposition is useful in this respect.

Proposition 2.1.1. Let $\varphi : \mathrm{R}^{n} \to \overline{\mathrm{R}}$ be Lipschitz continuous around $\bar{x}$ and upper regular at this point. Then $0 \neq \hat{\partial}^{+}\varphi(\bar{x}) = \bar{\partial}\varphi(\bar{x})$.

Proof. The nonemptiness of $\hat{\partial}^{+}\varphi(\bar{x})$ follows directy from $\partial\varphi(\bar{\mathrm{x}}) \neq 0$ for locally Lipschitzian functions and the definition of upper regularity. Due to $\bar{\partial}\varphi(\bar{x}) = \mathrm{co}\,\partial\varphi(\bar{x})$, any local Lipschitzian function is lower regular at $\bar{x}$ if and only if $\hat{\partial}\varphi(\bar{x}) = \bar{\partial}\varphi(\bar{x})$. Hence the upper regularity of φ at $\bar{x}$ and the plus-minus symmetry of the generalized gradient imply that

$$\hat{\partial}^{+}\varphi(\bar{x}) = -\hat{\partial}(-\varphi)(\bar{x}) = -\bar{\partial}(-\varphi)(\bar{x}) = \bar{\partial}\varphi(\bar{x})$$

which ends the proof of the proposition.

Note that all the assumptions of Proposition 2.1.1. hold for concave functions continuous around $\bar{x}$.

2.2 Superdifferential form of the discrete maximum principle

The following problem (P) of the Mayer type is considered as the basic model:

minimize $J(x,u) := \varphi(x(t_1))$ (i)

over discrete control processes $\{x(.), u(.)\}$ satisfying

$$x(t+h) = x(t) + hf(t, x(t), x(t-\tau), u(t)) , \qquad x(t_0) \equiv x_0 \in R^n \tag{ii}$$

$$u(t) \in U \, , \; t \in T := \{t_0, t_0 + h, ..., t_1 - h\}, \tag{iii}$$

$$x(t) = c(t) \, , \; t \in T_0 := \{t_0 - \tau, t_0 - \tau + h, ..., t_0 - h\}, \tag{iv}$$

where $h \rangle 0$ is a discrete stepsize, $\tau = Nh$ is a time delay with some $N \in \mathrm{N} := \{1, 2, ...\}, U,$ is a compact set describing constraints on control values in (iii), and c(.) is a given function describing the initial "delay" condition (iv) for the delayed system (ii).

In this section we first study the discrete optimal control problem (P) defined in (i)-(iv) and then consider its multiple delay generalization. Let $\{x(.), u(.)\}$ be a feasible process to (P), and let $\{\overline{x}(.), \overline{u}(.)\}$ be an optimal process to this problem. For convenience sake we introduce the following notation:

$$\xi(t) := (x(t), x(t-\tau)), \quad \overline{\xi}(t) := (\overline{x}(t), \overline{x}(t-\tau)),$$
$$f(\mathrm{t}, \xi, \mathrm{u}) := f(t, x(t), x(t-\tau), u(t)), \quad f(\mathrm{t}, \overline{\xi}, \mathrm{u}) := f(t, \overline{x}(t), \overline{x}(t-\tau), u(t)),$$
$$f(\mathrm{t}+\tau, \xi, \mathrm{u}) := f(t+\tau, x(t+\tau), x(t), u(t+\tau)),$$
$$\Delta x(t) := x(t) - \overline{x}(t), \quad \Delta f(t) := f(\mathrm{t}, \xi, \mathrm{u}) - f(\mathrm{t}, \overline{\xi}, \overline{\mathrm{u}}),$$
$$\Delta_u f(t) = f(\mathrm{t}, \overline{\xi}, \mathrm{u}) - f(\mathrm{t}, \overline{\xi}, \overline{\mathrm{u}}).$$

Using this notation, we define the adjoint system

$$\begin{aligned} p(t) = p(t+h) + h\frac{\partial f^*}{\partial x}(t,\bar{\xi},\bar{u})p(t+h) \\ + h\frac{\partial f^*}{\partial y}(t+\tau,\bar{\xi},\bar{u})p(t+\tau+h), t \in T \end{aligned} \tag{2.9}$$

to (2.2) along the optimal process $\{\bar{x}(.),\bar{u}(.)\}$. Consider the Hamilton-Pontryagin function

$$H(t, p(t+h), \xi(t), u(t)) := \langle p(t+h), f(t,\xi(t),u(t)) \rangle, \tag{2.10}$$

which allows us to rewrite the adjoint system (2.9) in the simplified form

$$p(t) = p(t+h) + h\left[\frac{\partial H}{\partial x}(t) + \frac{\partial H}{\partial y}(t+\tau)\right]$$

with $H(t) := H(t, p(t+h), \bar{\xi}(t), \bar{u}(t))$. Form the set

$$\Lambda(\bar{u}(t)) := \left\{u \in U \middle| f(t,\bar{\xi},u) \in \sigma(f(t,\bar{\xi},\bar{u}); f(t,\bar{\xi},U))\right\}. \tag{2.11}$$

where $\sigma(q;Q)$ denotes the star-neighborhood of $q \in Q$ relative to Q

$$\sigma(q;Q) := \left\{a \in Q \middle| \exists \varepsilon_k \downarrow 0 \text{ such that } q + \varepsilon_k(a-q) \in Q \text{for all} k \in \mathrm{N}\right\} \tag{2.12}$$

It easily follows from (2.11) and (2.12) that $\Lambda(\bar{u}(t)) = U$ if the set $f(t,\bar{\xi},U)$ is convex. The following theorem establishes a new superdifferential form of the discrete maximum principle for both delayed and non-delayed systems.

Theorem 2.2.1. Let $\{\bar{x}(.), \bar{u}(.)\}$ be an optimal process to (P). Assume that $\varphi : \mathrm{R}^n \to \overline{\mathrm{R}}$ is finite at $\bar{x}(t_1)$ and that $\hat{\partial}^+ \varphi(\bar{x}(t_1)) \neq 0$. Then for any $x^* \in \hat{\partial}^+ \varphi(\bar{x}(t_1))$ one has the discrete maximum principle

$$\begin{aligned} & H(t, p(t+h), \bar{x}(t), \bar{x}(t-\tau), \bar{u}(t)) \\ & = \max_{u \in \Lambda(\bar{u}(t))} H(t, p(t+h), \bar{x}(t), \bar{x}(t-\tau), u(t)), \quad t \in T, \end{aligned} \tag{2.13}$$

where p(.) is an adjoint trajectory satisfying (2.9) and the transversality conditions

$$p(t_1) = -x^*, \quad p(t) = 0 \ \text{ for } t \rangle t_1. \tag{2.14}$$

The maximum condition (2.13) is global over all $u \in U$ if the set $f(t, \bar{\xi}, U)$ is convex.

Proof. Take an arbitrary $x^* \in \hat{\partial}^+ \varphi(\bar{x}(t_1))$. It follows from (2.6) that

$$\varphi(x) - \varphi(\bar{x}(t_1)) \leq \langle x^*, x - \bar{x}(t_1) \rangle + o(\|x - \bar{x}(t_1)\|) \tag{2.15}$$

for all x sufficiently close to $\bar{x}(t_1)$. Put $p(t_1) := -x^*$ and derive from (2.15) and (i) that

$$J(x,u) - J(\bar{x}, \bar{u}) = -\langle p(t_1), \Delta x(t_1) \rangle + o(\|\Delta x(t_1)\|) \tag{2.16}$$

for all feasible process $\{x(.), u(.)\}$ to (P) such that $x(t_1)$ is sufficiently close to $\bar{x}(t_1)$. One always has the identity

$$\begin{aligned} \langle p(t_1), \Delta x(t_1) \rangle &= \sum_{t=t_0}^{t_1 - h} \langle p(t+h) - p(t), \Delta x(t) \rangle \\ &+ \sum_{t=t_0}^{t_1 - h} \langle p(t+h), \Delta x(t+h) - \Delta x(t) \rangle \end{aligned} \tag{2.17}$$

Due to (ii) we get the representation

$$\Delta x(t+h)-\Delta x(t)=h\Delta f(t)$$
$$=h\left[\Delta_u f(t)+\frac{\partial f}{\partial x}(t,\overline{\xi},\overline{u})\Delta x(t)+\frac{\partial f}{\partial y}(t,\overline{\xi},\overline{u})\Delta x(t-\tau)+\eta(t)\right],$$

where the remainder $\eta(t)$ is computed by

$$\eta(t)=\left(\frac{\partial f}{\partial x}(t,\overline{\xi},u)-\frac{\partial f}{\partial x}(t,\overline{\xi},\overline{u})\right)\Delta x(t)+\left(\frac{\partial f}{\partial y}(t,\overline{\xi},u)-\frac{\partial f}{\partial y}(t,\overline{\xi},\overline{u})\right)\Delta x(t-\tau)$$
$$+o\left(\left|\Delta x(t)\right|\right)+o\left(\left|\Delta x(t-\tau)\right|\right).$$

This allows us to present the second sum in (2.17) as

$$\sum_{t=t_0}^{t_1-h}\left\langle p(t+h,\Delta x(t+h)-\Delta x(t)\right\rangle$$
$$=h\sum_{t=t_0}^{t_1-h}\left\langle p(t+h),\Delta_u f(t)+\frac{\partial f}{\partial x}(t,\overline{\xi},\overline{u})\Delta x(t)+\frac{\partial f}{\partial y}(t,\overline{\xi},\overline{u})\Delta x(t-\tau)+\eta(t)\right\rangle.$$

Using the equalities

$$\Delta x(t)=0 \text{ for } t\le t_0\,,\ \ p(t+h)=0 \text{ for } t\ge t_1$$

and shifting the summation above, we have

$$\sum_{t=t_0}^{t_1-h}\left\langle p(t+h,\frac{\partial f}{\partial y}(t,\overline{\xi},\overline{u})\Delta x(t-\tau)\right\rangle$$
$$=\sum_{t=t_0}^{t_1-h}\left\langle p(t+\tau+h),,\frac{\partial f}{\partial y}(t+\tau,\overline{\xi},\overline{u})\Delta x(t)\right\rangle \tag{2.18}$$

Finally, substituting (2.9), (2.17), and (2.18) into (2.16), we obtain

$$
\begin{aligned}
J(x,u)-J(\overline{x},\overline{u}) = -h\sum_{t=t_0}^{t_1-h}\Delta_u H(t) \\
-h\sum_{t=t_0}^{t_1-h}\langle p(t+h),\eta(t)\rangle + o(|\Delta x(t_1)|) \geq 0
\end{aligned}
\tag{2.19}
$$

with $\Delta_u H(t) := H(t, p(t+h), \overline{\xi}(t), u(t)) - H(t, p(t+h), \overline{\xi}(t), \overline{u}(t))$ where $\Delta x(t_1)$ is sufficiently small.

Let us prove that (2.19) implies that, $\Delta_u H(t) \leq 0$ for any $t \in T$ and $u \in \Lambda(\overline{u}(t))$, which is equivalent to the discrete maximum principle (2.13). Assuming the contrary, we find

$$
\theta \in T \text{ and } u \in \Lambda(\overline{u}(\theta)) \;\; \Delta_u H(\theta) := a \rangle 0 . \tag{2.20}
$$

By definitions (2.11) and (2.12), there are sequences
$\varepsilon_k \downarrow 0$ and $u_k \in U$ such that

$$
f(\theta,\overline{\xi},\overline{u}) + \varepsilon_k f(\theta,\overline{\xi},u) - f(\theta,\overline{\xi},\overline{u}) := f(\theta,\overline{\xi},u_k) \in f(\theta,\overline{\xi},U),
$$

which is equivalent to

$$
\Delta_{u_k} f(\theta) := f(\theta,\overline{\xi},u_k) - f(\theta,\overline{\xi},\overline{u}) = \varepsilon_k (f(\theta,\overline{\xi},u) - f(\theta,\overline{\xi},\overline{u})) := \varepsilon_k \Delta_u f(\theta).
$$

Now let us consider needle variations of the optimal control defined as

$$
v_k(t) = \begin{cases} u_k & \text{if } \; t = 0, \\ \overline{u}(t) & \text{if } \; t \in T / \{\theta\}, \end{cases}
$$

which are feasible to (P) for all $k \in$ N, and let $\Delta_k x(t)$ be the corresponding perturbations of the optimal trajectory generated by $v_k(t)$. One can see that

$$\Delta_k x(t) = 0 \text{ for } t = t_0,\dots,\theta \text{ and } |\Delta_k x(t)| = O(\varepsilon_k) \text{ for } t = \theta + h,\dots,t_1 .$$

This implies that

$$\left(\frac{\partial f}{\partial x}(t,\bar{\xi},v_k) - \frac{\partial f}{\partial x}(t,\bar{\xi},\bar{u})\right)\Delta_k x(t)$$
$$+\left(\frac{\partial f}{\partial y}(t,\bar{\xi},v_k) - \frac{\partial f}{\partial y}(t,\bar{\xi},\bar{u})\right)\Delta_k x(t-\tau) = 0\,,\ t \in T$$

and that $\eta_k(t) = o(\varepsilon_k), \mathrm{k} \in \mathrm{N}$, for the corresponding remainders $\eta_k(.)$ defined above. Hence

$$J(x_k, v_k) - J(\bar{x},\bar{u}) = -h\Delta_{u_k} H(\theta) - h\sum_{t=t_0}^{t_1-h} \langle p(t+h), \eta_k(t)\rangle = -\varepsilon_k ha + o(\varepsilon_k) \langle 0$$

for all large $\mathrm{k} \in \mathrm{N}$ due to (2.20). Since $x_k(t_1) \to \bar{x}(t_1)$ as $k \to \infty$, this contradicts (2.19) and completes the proof of the theorem.

Let us present two important corollaries of Theorem 2.2.1. The first one assumes that φ is (Frechet) differentiable at the point $\bar{x}(t_1)$. Note that it may not be strictly differentiable (and hence not upper regular) at this point as for the function $\varphi(x) = x^2 \sin(1/x)$ for $x \neq 0$ with $\varphi(0) = 0$ (see definitions in Section 2). If φ is continuously differentiable around $\bar{x}(t_1)$ and $f = f(t,x,u)$ in (ii), then this result and its proof go back to the discrete maximum principle for non-delayed systems established in [19, Chapter IX].

Corollary 2.2.2. Let $\{\bar{x}(.),\bar{u}(.)\}$ be an optimal process to (P), where φ is assumed to be differentiable at $\bar{x}(t_1)$. Then one has the discrete maximum principle (2.13) with p(.) satisfying (2.9) and

$$p(t_1) = -\nabla\varphi(\bar{x}(t_1))\,,\ p(t) = 0 \text{ for } t \rangle t_1 \tag{2.21}$$

Proof. It follows from Theorem 2.2.1. due to the second relation in (2.7), which ensures that (2.14) reduces to (2.21).

The next corollary provides a striking result for upper regular and Lipschitz continuous cost function φ. In this case the discrete maxımum principle holds with the transversality condition $p(t_1) = -x^*$ given by any vector x^* from the generalized gradient $\overline{\partial}\varphi(\overline{x}(t_1))$, while conventional results ensure such conditions only for some subgradient.

Corollary 2.2.3. Let $\{\overline{x}(.), \overline{u}(.)\}$ be an optimal process to (P), where φ is assumed to be Lipschitz continuous around $\overline{x}(t_1)$ and upper regular at this point. Then for every vector $x^* \in \overline{\partial}\varphi(\overline{x}(t_1)) \neq 0$ one has the maximum principle (2.13) with p(.) satisfying (2.9) and (2.14).

Proof. Follows from Theorem 2.2.1 and proposition 2.1.1.

Now let us consider an extension (P_1) of problem (P) to the case of multiple delays: minimize (i) over discrete control processes $\{x(.), u(.)\}$ satisfying the system

$$\begin{aligned} x(t+h) &= x(t) \\ &+ hf(t, x(t), x(t-\tau_1), ..., x(t-\tau_m), u(t)), x(t_0) = x_0 \in \mathrm{R}^n \end{aligned} \tag{2.22}$$

with many delays $\tau_i = \mathrm{N}_i h$ for $N_i \in \mathrm{N}$ and $i = 1, ..., m$ subject to constraints (iii) and (iv), where $f = f(t, x, x_1, ..., x_m, u)$ satisfies our standing assumption and where the initial interval T_0 is correspondingly modified.

Denote $\overline{\xi}(t) := (\overline{x}(t), \overline{x}(t-\tau_1), ..., \overline{x}(t-\tau_m))$ and define p(.) satisfying (2.14) and the adjoint system

$$\begin{aligned} p(t) &= p(t+h) + h\frac{\partial f^*}{\partial x}(t, \overline{\xi}, \overline{u})p(t+h) \\ &+ h\sum_{i=1}^{m}\frac{\partial f^*}{\partial x}(t+\tau_i, \overline{\xi}, \overline{u})p(t+\tau_i+h) \end{aligned} \tag{2.23}$$

for $t \in T$, which can be rewritten in the Hamiltonian form

$$p(t) = p(t+h) + h\frac{\partial H}{\partial x}(t) + h\sum_{i=1}^{m}\frac{\partial H}{\partial x_i}(t+\tau_i)$$

in terms of (2.10) with $H(t) := H(t, p(t+h), \bar{\xi}(t), \bar{u}(t))$. The proof of the following theorem is similar to the basic case of Theorem 2.2.1. and can be omitted.

Theorem 2.2.4. Let $\{\bar{x}(.), \bar{u}(.)\}$ be an optimal process to (P_1) with $\hat{\partial}^{+}\varphi(\bar{x}(t_1)) \neq 0$. Then for any $x^* \in \hat{\partial}^{+}\varphi(\bar{x}(t_1))$ one has the discrete maximum principle

$$\begin{aligned} &H(t, p(t+h), \bar{\xi}(t), \bar{u}(t)) \\ &= \max_{u \in \Lambda(\bar{u}(t))} H(t, p(t+h), \bar{\xi}(t), u) \text{ for all } t \in T \end{aligned} \tag{2.24}$$

where p(.) is an adjoint trajectory satisfying (2.14) and (2.23).

Of course, we have the corollaries of Theorem 2.2.4. similar to the above ones for Theorem 2.2.1. Let us obtain another corollary of Theorem 2.2.4. for a counterpart (P_2) of the optimal control problem (P) involving discrete systems of neutral type

$$\begin{aligned} &x(t+h) = x(t) \\ &\quad + hf\left(t, x(t), x(t-\tau), \frac{x(t-\tau+h) - x(t-\tau)}{h}, u(t)\right), \quad t \in T \end{aligned} \tag{2.25}$$

where $\frac{x(t-\tau+h) - x(t-\tau)}{h}$ can be treated as an analog of the delayed derivative $\dot{x}(t-\tau)$ under the time discretization and where $f = f(t, x, y, z, u)$ satisfies our standing assumption.

Given an optimal process $\{\bar{x}(.), \bar{u}(.)\}$ to (P_2), we put

$$\bar{\xi}(t) := \left(\bar{x}(t), \bar{x}(t-\tau), \frac{\bar{x}(t-\tau+h) - \bar{x}(t-\tau)}{h}\right), \quad t \in T, \tag{2.26}$$

and define the adjoint discrete neutral-type system

$$
\begin{aligned}
p(t) = p(t+h) &+ h\frac{\partial f^*}{\partial x}(t,\bar{\xi},\bar{u})p(t+h) \\
&+ h\frac{\partial f^*}{\partial y}(t+\tau,\bar{\xi},\bar{u})p(t+\tau+h) \\
&+ \frac{\partial f^*}{\partial z}(t+\tau-h,\bar{\xi},\bar{u})p(t+\tau) \\
&- \frac{\partial f^*}{\partial z}(t+\tau,\bar{\xi},u)p(t+\tau+h)\ ,\ \ t \in T
\end{aligned}
\tag{2.27}
$$

Corollary 2.2.5. Let $\{\bar{x}(.),\bar{u}(.)\}$ be an optimal process to (P_2) with $\hat{\partial}^+\varphi(\bar{x}(t_1)) \neq 0$. Then for any $x^* \in \hat{\partial}^+\varphi(\bar{x}(t_1))$ one has the discrete maxımum principle (2.24), where $\bar{\xi}(.)$ is defined in (2.26) and where p(.) is an adjoint trajectory satisfying (2.14) and (2.27).

Proof. Observe that the neutral system (2.25) can be easily reduced to (2.22) with two delays. Thus this corollary follows from Theorem 2.2.4. via simple calculations.

A drawback of the superdifferential form of the discrete maximum principle established above is that the Frechet superdifferential may be empty for nice functions important in nonsmooth minimization, e.g., for convex functions that are not differentiable at the minimum points. In the next section we derive results on the discrete maximum principle that cover delayed problems of type (P) with general nonsmooth cost functions φ. Results of the latter subdifferential type are applicable to a broad class of nonsmooth problems, but they may not be as sharp as the superdifferential form of Theorem 2.2.1. when it applies.

2.3 Discrete maximum principle in terms of basic normals and subgradients

In this section of the thesis, we present nonsmooth versions of the discrete maximum principle for the delayed problem (P) in (i)-(iv) with transversality conditions expressed in terms of basic normals and subgradients defined in Section 2.1. The corresponding modifications for problems (P_1) and (P_2) can be made similarly to Section 2.2.

Theorem 2.3.1. Let $\{\bar{x}(.),\bar{u}(.)\}$ be an optimal process to (P), and let $\bar{x} := \bar{x}(t_1)$. Assume that the set $f(t,x,y,U)$ is convex around $(\bar{x}(t),\bar{x}(t-\tau))$ for all $t \in T$. Then one has the following assertions.

(a)Let φ be lower semicontinuous around $\bar{x}$. Then there is a nonzero vector $(x^*,\lambda) \in \mathrm{R}^{n+1}$ such that $\lambda \geq 0$

$(x^*,-\lambda) \in N((\bar{x},\varphi(\bar{x}));\mathrm{epi}\varphi)$, and the discrete maximum principle

$$\begin{aligned} &H(t,p(t+h),\bar{x}(t),\bar{x}(t-\tau),\bar{u}(t)) \\ &= \max_{u \in U} H(t,p(t+h),\bar{x}(t),\bar{x}(t-\tau),u)\ ,\ \ t \in T \end{aligned} \tag{2.28}$$

holds with p(.) satisfying (2.9) and (2.14).

(b) Let φ be Lipschitz continuous around $\bar{x}$. Then there is $x^* \in \partial\varphi(\bar{x})$ such that (2.28) holds with p(.) satisfying (2.9) and (2.14).

Proof. We will proceed similarly to the non-delayed case using the method of metric approximation (cf. [28, Section 11]). This method allows us to approximate the original nonsmooth problem by a family of smooth discrete problems with delays and then arrive at the desired conclusions by a limiting procedure involving the corresponding results.

Let us first prove assertion (a). Taking a parameter $\gamma \in \mathrm{R}$, we consider a parametric family of the following optimal control pronlems (P_γ) for delayed discrete systems with the distance-type cost functional:

$$\text{Minimize } J_\gamma(x,u) := \text{dist}((x(t_1),\gamma);\text{epi}\varphi) + \sum_{t=t_0}^{t_1} \left| x(t) - \bar{x}(t) \right|^2$$

over control processes $\{x(.),u(.)\}$ subject to constraints (ii)-(iv).

Let $\bar{\gamma} := \varphi(\bar{x}(t_1))$, and let $\{\bar{x}_\gamma(.),\bar{u}_\gamma(.)\}$ be optimal processes to (P_γ) that obviously exist by the classical Weierstrass theorem due to the standing assumptions. It follows from the structure of (P_γ) and the optimality of $\{\bar{x}(.),\bar{u}(.)\}$ in the original problem (P) that $\bar{x}_\gamma(t) \to \bar{x}(t)$ as $\gamma \to \bar{\gamma}$ for all $t \in T \cup \{T_1\}$. Moreover,

$$m_\gamma := \text{dist}((\bar{x}_\gamma(t_1),\gamma);\text{epi}\gamma) \rangle\, 0 \text{ whenever } \gamma \,\langle\, \bar{\gamma}. \qquad (2.29)$$

The latter allows us to conclude that, for any $\gamma \,\langle\, \bar{\gamma}$, the process $\{\bar{x}_\gamma(.),\bar{u}_\gamma(.)\}$ is optimal to the smooth problem $(\overline{P}_\gamma)$ of minimizing the functional

$$\bar{J}_\gamma(x,u) := \left(\left| x(t_1) - x_\gamma \right|^2 + \left| \gamma - w_\gamma \right|^2 \right)^{1/2} + \sum_{t=t_0}^{t_1} \left| x(t) - \bar{x}(t) \right|^2$$

subject to (ii)-(iv), where (x_γ, w_γ) is an arbitrary element of the Euclidean projector $\Pi((\bar{x}_\gamma(t_1),\gamma);\text{epi}\varphi)$ of $(\bar{x}_\gamma(t_1),\gamma)$ to the closed set epi φ. Introducing an additional state variable $x_{n+1}(t)$ by

$$x_{n+1}(t+h) = x_{n+1}(t) + \left| x(t) - \bar{x}(t) \right|^2, \; x_{n+1}(t_0) = 0 \qquad (2.30)$$

we rewrite problem $(\overline{P}_\gamma)$ in the equivalent form of minimizing the Mayer-type functional

$$\begin{aligned} \bar{J}_\gamma(x,x_{n+1},u) := & \left(\left| x(t_1) - x_\gamma \right|^2 + \left| \gamma - w_\gamma \right|^2 \right)^{1/2} \\ & + x_{n+1}(t_1) + \left| x(t_1) - \bar{x}(t_1) \right|^2 \end{aligned} \qquad (2.31)$$

over $\{x(.), x_{n+1}(.), u(.)\}$ satisfying (ii)-(iv) and (2.30).

Denote $\bar{\xi}_\gamma(t) := (\bar{x}_\gamma(t), \bar{x}_\gamma(t-\tau))$ and observe that the sets $f(t, \bar{\xi}_\gamma(t), U)$ are convex for all $t \in T$ while the cost function in (2.31) is differentiable at $(\bar{x}_\gamma(t_1), \bar{x}_{n+1}(t_1))$, where $\bar{x}_{n+1}(.)$ is generated by $\bar{x}_\gamma(.)$ in (2.30). Now applying Corollary 2.2.4. to problem (P_γ) as $\gamma \langle \bar{\gamma}$ and taking into account the structure of the cost function (2.31), we arrive at the discrete maximum principle

$$H(t, p_\gamma(t+h), \bar{\xi}_\gamma(t), \bar{u}_\gamma(t)) = \max_{u \in U} H(t, p_\gamma(t+h), \bar{\xi}_\gamma(t), u), \; t \in T \;,$$

where $p_\gamma(.)$ satisfies the adjoint system (2.9) along $\{\bar{x}_\gamma(.), \bar{u}_\gamma(.)\}$ with the transversality conditions

$$p_\gamma(t_1) = -\frac{\bar{x}_\gamma(t_1) - x_\gamma}{m_\gamma} - 2(\bar{x}_\gamma(t_1) - \bar{x}(t_1)), \; p_\gamma(t) = 0 \text{ for } t \rangle t_1 \;,$$

where $m_\gamma \rangle 0$ is given in (2.29), and where

$$\left(\frac{|\bar{x}_\gamma(t_1) - x_\gamma|}{m_\gamma} \right)^2 + \left(\frac{|\gamma - w_\gamma|}{m_\gamma} \right)^2 = 1$$

Passing to the limit as $\gamma \uparrow \bar{\gamma}$ in the above relations and using the constructions of the basic normal cone (2.1), we arrive at all the conclusions of (a).

To justify (b) when φ is Lipschitz continuous around $\bar{x}(t_1)$, we observe that in this case one has $x^* \in \lambda \partial \varphi(\bar{x}(t_1))$ from (a) and (2.3). The latter implies that $\lambda \neq 0$, which yields (b) by normalization.

Let us compare the superdifferential and subdifferential forms of the discrete maximum principle from Theorems 2.2.1. and 2.3.1., respectively. As mentioned above, Theorem 2.3.1. is applicable to a broad class of nonsmooth problems (P) while Theorem 2.2.1. requires that $\hat{\partial}^+ \varphi(\bar{x}(t_1)) \neq 0$, which excludes many nonsmooth functions. On the other hand, the superdifferential form has essential advantages for special classes of cost functions.

First we observe that Theorem 2.2.1. implies the gradient form (2.21) of transversality when φ is just differentiable at $\bar{x}(t_1)$ (it may even not be Lipschitz continuous around this point) while Teorem 2.3.1. ensures (2.21) only when φ is strictly differentiable at $\bar{x}(t_1)$ (see (2.8) and the related discussion in Section 2). The most striking difference between subdifferential and superdifferential transversality conditions appears in the case of upper regular and locally Lipschitzian cost functions. In this case Theorem 2.3.1. provides the discrete maximum principle generated by some subgradient $x^* \in \overline{\partial}\varphi(\bar{x}(t_1)) \subset \overline{\partial}\varphi(x(t_1))$ in (2.14) while Corollary 2.2.5. ensures it for every $x^* \in \overline{\partial}\varphi(\bar{x}(t_1))$. This is a big difference!

3. OPTIMALITY CONDITIONS VIA EXHAUSTERS AND QUASIDIFFERENTIABILITY IN SWITCHING CONTROL PROBLEM

3.1 Some knowledge about quasidifferential and exhausters

Let us begin with basic constructions of directional derivative (or its generalization) used in the sequel. We refer the reader to the book by Demyanov and Rubinov [14, 15] and articles Roshchina [43, 44, 45], Demyanov and Roshchina [16, 17, 18]. Let $f : X \to R, X \in R^n$ be an open set. The function $f_H^{\uparrow}$ ($f_H^{\downarrow}$) is called Hadamard upper (lower) derivative of the function f at the point x in the direction g if there exist limit

$$f_H^{\uparrow}(x,g) := \limsup_{[\alpha,g']\to[+0,g]} \frac{1}{\alpha}\left[f(x+\alpha g') - f(x)\right] \tag{3.1}$$

$$\left(f_H^{\downarrow}(x,g) := \liminf_{[\alpha,g']\to[+0,g]} \frac{1}{\alpha}\left[f(x+\alpha g') - f(x)\right]\right) \tag{3.2}$$

Note that limit in (3.1, 3.2) always exist, but are not necessarily finite. This derivative is positively homogeneous function of direction. The Gateaux upper (lower) subdifferential of the function f at a point $x_0 \in X$ can be defined as follows

$$\partial_G^+ f(x_0) = \left\{ v \in R^n \,\middle|\, \limsup_{t\downarrow 0} \frac{f(x_0+tg) - f(x_0)}{t} \le (v,g), \forall g \in R^n \right\}$$

$$\left(\partial_G^- f(x_0) = \left\{ v \in R^n \,\middle|\, \liminf_{t\downarrow 0} \frac{f(x_0+tg) - f(x_0)}{t} \ge (v,g), \forall g \in R^n \right\}\right)$$

The set

$$\hat{\partial}^+ f(x_0) = \left\{ v \in R^n \,\middle|\, \limsup_{x\to x_0} \frac{f(x) - f(x_0) - \langle v, x-x_0\rangle}{\|x-x_0\|} \le 0 \right\}$$

$$\left(\hat{\partial} f(x_0) = \left\{ v \in R^n \,\middle|\, \limsup_{x\to x_0} \frac{f(x) - f(x_0) - \langle v, x-x_0\rangle}{\|x-x_0\|} \le 0 \right\}\right)$$

is called, respectively, upper (lower) Frechet subdifferential of the function f at the point x_0.

It is known that, if f is a quasidifferentiable function then [15] its directional derivative at a point x is represented as

$$f'(x,g) = \max_{v\in\underline{\partial} f(x)}(v,g) + \min_{w\in\overline{\partial} f(x)}(w,g)$$

where $\underline{\partial} f(x), \overline{\partial} f(x) \subset R^n$ are convex compact sets. From last relation, easly we can reduce that

$$f'(x,g) = \min_{w\in\overline{\partial} f(x)} \max_{v\in w+\underline{\partial} f(x)}(v,g) = \max_{v\in\underline{\partial} f(x)} \min_{w\in v+\overline{\partial} f(x)}(v,g)$$

It means that for the function $h(g) = f'(x,g)$ the upper and lower exhausters can be describe following way

$$E^* = \left\{C = w + \underline{\partial} f(x) \middle| w \in \overline{\partial} f(x)\right\}$$
$$E_* = \left\{C = v + \overline{\partial} f(x) \middle| v \in \underline{\partial} f(x)\right\}$$

It follows that the directional derivative of a quasidifferentiable function (as a function of direction) is positively homogeneous and quasidifferentiable.

If $h(g)$ is upper semicontinuous in g, then [14, 18] $h(g)$ can be expressed as

$$h(g) = \inf_{C\in E^*} \max_{v\in C}(v,g)$$

and if $h(g)$ is lower semicontinuous in g, then $h(g)$ can be written as in the form

$$h(g) = \sup_{C\in E_*} \min_{v\in C}(v,g)$$

In [9], Castellani proved that if h is Lipschitz then h can be written in the forms

$$h(g) = \min_{C \in E^*} \max_{v \in C} (v, g)$$

and

$$h(g) = \max_{C \in E_*} \min_{v \in C} (v, g)$$

The pair $E = [E^*, E_*]$ of families of totally bounded, convex compact sets a biexhausters, E^* being an upper exhausters and E_* a lower one. In [17, theorem 3.3] and in [18, theorem 2], the authors wrote and proved relationship between upper exhausters and Frechet lower subdifferential. They also wrote about relationship between lower exhauster and Frechet upper subdifferential and remark that this relationships can be prove easly by using symmetrical construction. For the continence of our future work in current article, let us prove this relationship. It is clear that Frechet upper subdifferential can be Express with the Hadamard upper derivative following way [17, lemma 3.2]

$$\partial_F^+ f(x_0) = \partial_F^+ f_H^{\uparrow}(x_0, 0_n).$$

Then:

Theorem 3.1.1. Let E_* be lower exhausters of the positively homogeneous function $h: R^n \to R$. Then $\bigcap_{C \subset E_*} C = \hat{\partial}^+ h(0_n)$, where $\hat{\partial}^+ h$ is the Frechet upper subdifferential of the h at 0_n and for the positively homogeneous function $h: R^n \to R$ the Frechet superdifferential at the point zero follows,

$$\hat{\partial}^+ h(0_n) = \{ v \in R^n \,|\, h(x) - (v, x) \le 0 \} \tag{3.3}$$

Proof. Take any $v_0 \in \bigcap_{C \subset E_*} C$. Then from the definition of lower exhausters that,

$v_0(x) \geq h(x),\ \forall x \in R^n \Rightarrow \bigcap_{C \subset E_*} C \subset \hat{\partial}^+ h(0_n)$

Consider now any $v_0 \in \hat{\partial}^+ h(0_n) \Rightarrow$

$$v_0(x) \geq h(x). \tag{3.4}$$

Let us consider $v_0 \notin \bigcap_{C \subset E_*} C$. Then there exist $C_0 \in E^*$ where $v_0 \notin C_0$.

Then by separation theorem there exist $x_0 \in R^n$ such that

$$(x_0, v_0) \leq \max_{v \in C_0}(x_0, v) \leq h(x)$$

It is conducts (3.3) and $v_0 \in C$ for every $C \in E^*$ and due to arbitrary. This means that $v_0 \in \bigcap_{C \subset E_*} C$. It is complete proof of the theorem.

Lemma 3.1.2. The Frechet upper (lower) and Gateaux upper (lower) subdifferentials of a positively homogeneous function at zero coincide.

Proof. Let $h: R^n \to R$ be a positively homogeneous function. It is not difficult to observe that every $g \in R^n$ and every $t \rangle 0$

$$\frac{h(0_n + t_g) - h(0_n)}{t} = \frac{th(g)}{t} = h(g)$$

Hence, the Gateaux lower subdifferential of h at 0_n take the form

$$\partial_G^+ h(0_n) = \left\{ v \in R^n \middle| h(g) \leq (v, g),\ \forall g \in R^n \right\}$$

which coincide with the representation of the Frechet upper subdifferentials of the positively homogeneous function (see [22], Proposition 1.9).

3.2 Problem formulation and necessary optimality principle

We consider the following optimization problem:

$$\dot{x}_K(t) = f_K(x_K(t), u_K(t), t), \quad t \in [t_{K-1}, t_K], \quad K = 1, 2, ..., N \tag{3.5}$$

$$x_1(t_0) = x_0 \tag{3.6}$$

$$F_K(x_N(t_N), t_N) = 0, \quad K = 1, 2, ..., N \tag{3.7}$$

$$\mathrm{x}_{\mathrm{K}+1}(t_K) = M_K(x_K(t_K), t_K), \; K = 1, 2, ..., N-1 \tag{3.8}$$

$$\min . S(u_1, ..., u_N, t_1, ..., t_N) = \sum_{K=1}^{N} \varphi_K(x_K(t_K)) + \sum_{K=1}^{N} \int_{t_{K-1}}^{t_K} L(x_K, u_K, t) dt \tag{3.9}$$

Remark 3.2.1. We consider the problem (3.5)-(3.9) in the first section (the problem (1.1)-(1.5)) but in this section we extend this result in the case of minimizing functional satisfies quasidifferential and exhauster conditions in the Demyanov and Rubinov sense.

Here $f_K : R \times R^n \times R^r \to R^n$, M_K and F_K are continuous, at least continuously partially differentiable vector-valued functions with respect to their variables, $L : R^n \times R^r \times R \to R$ is continuous and have continuous partial derivative with respect to their variables, $\varphi_K(.)$ has Frechet upper subdifferentiable (superdifferentiable) at a point $\bar{x}_K(t_K)$ and positively homogeneous functional, $u_K(t) : R \to U_K \subset R^r$ are controls. The sets U_K are assumed to be nonempty and bounded. Here (3.8) are switching conditions. It is required to find the control $u_1, u_2, ..., u_N$, switching points $t_1, t_2, ..., t_{N-1}$ and the end point t_N (here $t_1, t_2, ..., t_N$ are not fixed) with corresponding state $x_1, x_2, ..., x_N$ satisfying (3.5)-(3.9) so that the function $S(u_1, ... u_N, t_1, ..., t_N)$ in (3.9) is minimized. We will derive necessary conditions for nonsmooth version of these problems (by using exhausters and quasidifferentiable in the Demyanov sense).

Denote:

$$\theta = (t_1, t_2, ..., t_N),\ x(t) = (x_1(t), x_2(t), ..., x_N(t)),\ u(t) = (u_1(t), u_2(t), ..., u_N(t)).$$

Our aim is to find tuple $(x(t), u(t), \theta)$ which solves problem (3.5)-(3.9). Such tuple will be called optimal control for the problem (3.5)-(3.9). At first we assume that $\varphi_K(.)$ is Hadamard upper differentiable at the point $\bar{x}_K(t_K)$ to the zero direction. Then, $\varphi_K(.)$ is uppersemicontinuous [47] and it has exhaustive family of lower concave approximations of $\varphi_K(.)$ [15, theorem 3]. Then:

Theorem 3.2.2. (Necessary optimality condition in terms of lower exhauster)
Let $(\bar{u}_K(.), \bar{x}_K(.), \bar{\theta})$ be an optimal solution to the control problem (3.5)-(3.9). Then, every collections of the element from intersection of the subsets of the lower exhauster of the functional $\varphi_K(\bar{x}_K(t_K)),\ x_K^* \in \bigcap_{C_K \in E_{*,K}} C_K,\ K = 1, 2, ..., N$, there exist vector functions $p_K(t),\ K = 1, 2, ..., N$ which one has following necessary optimality condition hold:

1) State equation.

$$\dot{x}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial p_K}$$

2) Costate equation.

$$\dot{p}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial x_K}$$

3) At the switching points, $t_1, t_2, ..., t_{N-1}$

$$x_K^* - p_K(t_K) - p_{K+1}(t_K)\frac{\partial M_K(x_K(t_K), t_K)}{\partial x_K} = 0,\ K = 1, 2, ..., N-1$$

4) Stationarity condition

$$\frac{\partial H_K(x_K, p_K, u_K, t)}{\partial u_K} = 0,\ K = 1, 2, ..., N,\ t \in [t_{K-1}, t_K]$$

5) At the end point t_N

$$p_N(t_N) = x_N^* + \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial x_N}$$

$$\left(\sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial t_N}\right)\delta_{L,N} - \left(\sum_{K=1}^{N-1} p_{K+1}(t_K) \frac{\partial M_K(x_K(t_K), t_K)}{\partial t_K}\right)(1 - \delta_{L,N}) = 0,$$

$$L = 1, 2, ..., N \ \ here\ \ \delta_{L,N} = \begin{cases} 1, & L = N \\ 0, & L \neq N \end{cases},$$

$$H_K(x_K, u_K, p_K, t) = L(x_K, u_K, p_K, t) + p_K^T f_K(x_K, u_K, p_K, t),$$

where $E_{*,K}$ is lower exhauster of the functional $\varphi_K(x_K(t_K))$ and $\lambda_K,\ K = 1, 2, ..., N$ are the vectors, $p_K(.)$ is defined by the conditions (3.2) and (3.3) in the theorem, later.

Proof. To prove the theorem, take any elements from intersection of the subset of the exhauster, $x_K^* \in \bigcap_{C_K \in E_{*,K}} C_K,\ K = 1, 2, ..., N$. Then by using theorem 3.1 we can write that $x_K^* \in \hat{\partial}^+ \varphi_K(\bar{x}_K(t_K))$. Then, apply the variational description from theorem 1.88 ((i)) in [33] to the subgradients $-x_K^* \in \hat{\partial}^+(-\varphi_K(\bar{x}_K(t_K)))$. In this way we find functions $s_K : X \to R$ for K=1, 2, …, N satisfying the relations $s_K(\bar{x}_K(t_K)) = \varphi_K(\bar{x}_K(t_K))$ and $s_K(x_K(t)) \geq \varphi_K(x_K(t))$ in some neighborhood of $\bar{x}_K(t_K)$ and such that each $s_K(.)$ is continuously differentiable at $\bar{x}_K(t_K)$ with $\nabla s_K(\bar{x}_K(t_K)) = x_K^*,\ K = 1, 2, ..., N$. It is easy to check that $\bar{x}_K(.)$ is a local solution to the following optimization problem of type (3.5)-(3.9) but with cost continuously differentiable around $\bar{x}_K(.)$. This means that, we deduce the optimal control problem (3.5)-(3.9) with the nonsmooth cost functional to the smooth cost functional data

$$\min .S(u_1,...,u_N,t_1,...,t_N)=\sum_{K=1}^{N}s_K(x_K(t_K))+\sum_{K=1}^{N}\int_{t_{K-1}}^{t_K}L(x_K,u_K,t)dt$$

taking into account that

$$\nabla s_K(\overline{x}_K(t_K))=x_K^*,\ K=1,2,...,N$$

Then, by using Lagrange multipliers rule and by using results which described in first section where we calculated first variation of the minimizing functional. We can obtain first variation of the minimizing functional in the following form;

$$\begin{aligned}
\delta S' &= \sum_{K=1}^{N-1}\left(\frac{\partial s_K(x_k(t_K))}{\partial x_K}-p_K(t_K)-p_{K+1}(t_K)\frac{\partial M_K(x_k(t_K),t_K)}{\partial x_K}\right)\delta x_K(t_K)\\
&+\left(\frac{\partial s_N(x_N(t_N))}{\partial x_N}+\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N))}{\partial x_N}-p_N(t_N)\right)\delta x_N(t_N)\\
&+\sum_{L=1}^{N}\left[\left(\sum_{K=1}^{N}\lambda_K\frac{\partial F_K(x_N(t_N))}{\partial t_N}\right)\delta_{L,N}-\left(\sum_{K=1}^{N-1}p_{K+1}(t_K)\frac{\partial M_K(x_k(t_K),t_K)}{\partial t_K}\right)(1-\delta_{L,N})\right]\delta t_N\\
&+\sum_{K=1}^{N}\left(\frac{\partial H_K(x_k,p_K,u_K,t)}{\partial x_K}-p_K(t)\right)\delta x_K+\sum_{K=1}^{N}\frac{\partial H_K(x_k,p_K,u_K,t)}{\partial u_K}\delta u_K\\
&+\sum_{K=1}^{N}\left(\frac{\partial H_K(x_k,p_K,u_K,t)}{\partial p_K}-p_K(t)\right)\delta p_K
\end{aligned}$$

The latter sum is known because

$$\frac{\partial H_K(x_K,u_K,p_K,t)}{\partial p_K}=\dot{x}_K(t)$$

According to a necessary condition for an optimal solution $\delta J'=0$. Setting to zero coefficients of the independent increments $\delta x_N(t_N),\delta x_K(t_K),\delta x_K,\delta u_K$ and δp_K and taking into account that

$$\nabla s_K(\overline{x}_K(t_K))=x_K^*,\quad K=1,2,...,N$$

yields the necessary optimality condition in the following form

$$\dot{x}_K(t) = \frac{\partial H_K(u_K, x_K, p_K, t)}{\partial p_K}, \quad K = 1, 2, ..., N$$
$$\dot{p}_K(t) = \frac{\partial H_K(x_K, u_K, p_K, t)}{\partial x_K}, \quad K = 1, 2, ..., N$$
$$\frac{\partial H_K(u_K, x_K, p_K, t)}{\partial u_K} = 0, \quad K = 1, 2, ..., N$$
$$x_K^* - p_K(t_K) - p_{K+1}(t_K) = \frac{\partial M_K(x_K(t_K), t_K)}{\partial x_K} = 0, \quad K = 1, 2, ..., N-1$$
$$x_N^* + \sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial x_N} - p_N(t_N) = 0$$
$$\left(\sum_{K=1}^{N} \lambda_K \frac{\partial F_K(x_N(t_N), t_N)}{\partial t_N} \right) \delta_{L,N}$$
$$- \left(\sum_{K=1}^{N} p_{K+1}(t_K) = \frac{\partial M_K(x_K(t_K), t_K)}{\partial t_K} \right) \left(1 - \delta_{L,N}\right) = 0.$$

This completes the proof of the theorem.

Theorem 3.2.3. (Necessary optimality conditions for switching optimal control system in terms of Quasidifferentiability) Let the minimization functional $\varphi_K(.)$ to be positively homogeneous, quasidifferentiable at a point $\bar{x}_K(.)$ and $(\bar{u}_K(.), \bar{x}_K(.), \bar{\theta})$ be an optimal solution to the control problem (3.5)-(3.9). Then there exist vector functions $p_K(t),\ K = 1, 2, ..., N$ and there exist convex compact and bounded set $M(\varphi_K(.))$, which for any elements $x_K^* \in M(\varphi_K(.))$ the necessary optimality conditions 1)-5) in the theorem 3.2.2. are satisfied.

Proof. Let minimization functional $\varphi_K(.)$ to be positively homogeneous, quasidifferentiable at a point $\bar{x}_K(\bar{t}_K)$. Then there exist totally bounded lower exhausters $E_{*,K}$ for the $\varphi_K(.)$ ([17] theorem 4). Let us make substitution $M(\varphi_K(.)) = E_{*,K}$, take any element $x_K^* \in M(\varphi_K(.))$. Then $x_K^* \in E_{*,K}$ also and if we follow the prove description and result in theorem 3.2.2., we can prove the theorem 3.2.3.

If we use relationship between Gateaux upper subdifferential and Dini upper derivative[15, lemma3.6], put substitution $h_K(g) = \varphi^+_{K,H}(\bar{x}_K(\bar{t}_K), g)$ then we can write following corollary. (here $\varphi^+_{K,H}(\bar{x}_K(\bar{t}_K), g)$ is Hadamard upper derivative of the minimizing functional $\varphi_K(.)$ in the direction g)

Corollary 3.2.4. Let the minimization functional $\varphi_K(.)$ to be positively homogeneous, Dini upper differentiable at a point $\bar{x}_K(.)$ and $(\bar{u}_K(.), \bar{x}_K(.), \bar{\theta})$ be an optimal solution to the control problem (3.5)-(3.9). Then for any elements $x^*_K \in \partial^+_G h_K(0_n)$ there exist vector functions $p_K(t),\ K = 1, 2, ..., N$ such that the necessary optimality conditions 1)-5) in the theorem 3.2.2. hold.

Proof. Let take any element $x^*_K \in \partial^+_G h_K(0_n)$. Then by using lemma 3.8 in [17], we can write $x^*_K \in \partial^+_F h_K(0_n)$. Next, if we use lemma 3.2 in [17], then we can put $x^*_K \in \partial^+_F \varphi_K(\bar{x}_K(\bar{t}_K))$. At least, if we follow the theorem 3.1.1. (relationship between upper Frechet subdifferential and exhausters) and the theorem 3.2.2. (necessary optimality condition in terms of exhausters), we can prove the result of the corollary 3.2.4.

CONCLUSION

In thesis, results for nonsmooth optimal control of switching systems are reported. The method takes advantage of the special structure of nonlinear optimal switching control systems with smooth and nonsmooth minimizing functional. Application of necessary optimality condition to the switching optimal control problem is also reported. A further research topic can be carried on the development of methods to search optimality conditions for the nonsmooth switching optimal control problem for the differential and discrete inclusion, nonsmooth optimal switching control problem with delay and neutral type.

We also investigated necessary optimality condition for discrete system in the nonsmooth case. It is first time obtained optimality condition for given problem.

In thesis we tried to get necessary optimality conditions for the switching optimal control problem in terms of exhausters and quasidifferentiable in the Demyanov sense. By using necessary results about relationship Frechet upper subdifferential, Quasidifferentiability and exhausters which was obtained by Demyanov and Roshchina [15], Roshchina [18], and by using results connection Gateaux subdifferentiable and Dini derivative which obtained by Demyanov and Roshchina in [15], it is obtained necessary condition for switching control problem. It is first time studied application quasidifferentiability and exhausters in the switching optimal control problem. But there are some open problem, like Clarke and Penot subdifferentiable in the switching optimal control problem, reduction of the exhausters which it will be help for us to get more constructive optimality condition for the switching control problem.

REFERENCES

1) Antsaklis, P. J. and Nerode, A. (1998), eds., Special issue on hybrid system, IEEE Trans. Automat. Control 43, no, 4.
2) Bensoussan, A. and Menaldi, J. (1997). Hybrid control and dynamic programming, Dyna. Continuous. Discrete, Impul. System., 3, 395-442.
3) Bensoussan, A. and Menaldi, J. (apr.1997). “Hybrid Control and dynamic programming” Dyna. Continuous. Discrete, Impul. System., vol.43, pp.475-482.
4) Boltyanskii, V. G. (1973). Optimal Control of Discrete Systems [in Russian], Nauka, Moscow.
5) Boltyanskii, V. (2004). The maximum principle for variable structure systems, International Journal of Control, 77, 1445-1451.
6) Branicky, M. S., Borkar, V. S. and Mitter, S. K. (1998). A unified framework for hybrid control: model and optimal control theory. IEEE Transactions on Automatic Control, 43(1):31-45.
7) Capuzzo Dolcetta and Evans, L. C. (1984). Optimal switching for ordinary differential equations. SIAM Journal of Control and Optimization, 22(1):143-161.
8) Caravello, R. M. and Piccoli, B. (2002). Hybrid Necessary Principle, preprint SSSA 71, AF.
9) Castellani, M. A. (2000). Dual characterization for proper positively homogeneous functions, J. Global Optim. 16, pp. 393-400.
10) Changzhi, W. and Kok Teo, Rui Li, Zhao, Y. (2006). Optimal control of switching system with time delay. Applied Mathematics Letters 19, 1062-1067.
11) Chukwu, E. N. (1992). Stability and Time-Optimal Control of Hereditary Systems, Academic Pres, Boston.
12) Clarke, F. H. (1983) “Optimization and Nonsmooth Analysis,” Canadian Mathematical Society Series of Monographs and Advanced Texts. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.
13) D’Apice, C. and Garavello, M., Manzo, R., Piccoli, B. (2003). Hybrid optimal control: case study of a car with gears, International Journal of Control 76, 1272-1284.

14) Demyanov, V. F. and Rubinov, A. M. (1982). Elements of Quasidifferential Calculus (In Russian) in Nonsmooth Problems of Optimization Theory and Control, Ch.1, V. F. Demyanov, ed., Leningrad University Pres, Leningrad, pp. 5-127.

15) Demyanov, V. F. and Rubinov, A. M. (1995). Constructive Nonsmooth Analysis. Frankfurt a/M:Verlag Peter Lang.

16) Demyanov, V. F. and Roshchina, V. A. (2005). Constrained optimality conditions in terms of proper and adjoint exhausters, Appl. Comput. Math. 4(2), pp. 25-35, Azerbaijan National Academy of Sciences, Baku.

17) Demyanov, V. F. and Roshchina, V. (2008). Exhausters and subdifferentials in nonsmooth analysis, Optimization, Vol.57, no.1, pp. 41-56.

18) Demyanov, V. F. and Roshchina V. (Mar 2008). Exhausters, optimality conditions and related problems. J. Global Optim., Vol. 40 No: 1-3, 71-83.

19) Gabasov, R. and Kirillova, F. M. (1976). Qualitative Theory of Optimal Process, Marcel Dekker, New York.

20) Gokbayrak, K. and Cassandras, C. G. (2000). A Hieraarchical decomposition method for optimal control of hybrid system, in Proc. 39th IEEE Conf. Decision Control, pp.

21) Haklin, H. (1966). “A maximum principle of the Pontryagin type for systems described by nonlinear difference equations.” J. SIAM Control, 4, 90-112.

22) Kruger, A. Ya. (2003). On Frechet subdifferential. Optimization and related topics, 3., J.Math.Sci.(N.Y).116(3), pp. 3325-3358.

23) Li, R., Teo, K. L., Wong, K. H. and Duan, G. R. (2006). Control parameterization enhancing transform for optimal control of switched systems, Mathematics and Computer Modelling, 43, 1393-1403.

24) Magerramov, Sh. F. and Mansimov, K. B. (2001). Optimization of a class of discrete step control systems, (Russian, English) Comput. Math. Phys., 41, 334-339. Translation from Zh. Vychisl. Math. Fiz., 41, No.3, 360-366.

25) Maharramov, Sh. F. (2008). Optimality Condition of a Nonsmooth Switching Control System, Automatic control and Computer Science, vol. 42. no.2, pp.94-101.

26) Maharramov, Sh. F. (2010). Necessary Optimality Conditions for Switching Control System, Accepted for American Institute for Mathematical Science, Journal of Industrial and Manegament Optimization, 6, february, 10 pages.

27) Maharramov, Sh. F. with Sengul, M. and Gurbuz, E. (2010). The proceedings of 24th Mini Conference, On continuous optimization and information-based technologies in the financial sectors, some properties of weak subdifferential, p.60-65.

28) Mordukhovich, B. S. (1988). Approximation Methods in Problems of Optimization and Control [in Russian], Nauka, Moscow.

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

30) Mordukhovich, B. S. (2000). "Optimal control of difference, differential and differential- difference inclusions," J. Math. Sci., 100, 2323-2632.

31) Mordukhovich, B. S. and Trubnik, R. (2001). "Stability of discrete approximations and necessary optimality conditions for delay-differential inclusions," Ann. Oper. Res., 101, 149-170.

32) Mordukhovich, B. S. and Shvartsman, I. (2004). The maximum principle in constrained optimal control, SIAM J. Control Optim., 43, 1037-1062 (electronic).

33) Mordukhovich, B. S. (2005). "Variational Anaysis and Generalized Differentiation," Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330, 584 pp., Springer-Verlag, Berlin.

34) Nurminskii, E. A. (1979). Numerical Methods for Solutions of Deterministic and Stochastic Minimax Problems [in Russian], Naukova Dumka, Kiev.

35) Piccoli, B. (1999). "Necessary conditions for hybrid optimization", in Proc. 38th IEEE Conf. Decision and Control, pp. 410-415.

36) Pontryagin, A. S., Boltyanskii, V. G., Gamkrelidze, R. V. and Mishchenko, E. F. (1962). The Mathematical Theory of Optimal Processes, Wiley, New York.

37) Propoi, A. I. (1973). Elements of the Theory of Optimal Discrete Processes [in Russian], Nauka, Moscow.

38) Pshenichnyi, B. N. (1971). Necessary Conditions for an Extremum, Marcel Dekker, New York.

39) Pshenichniyi, B. N. (1980). Convex Analysis and Extremal Problems (In Russian), Nauka Publisher, Moscow.

40) Rachev, S. and Ruschendorf, L. (1998). "Mass Trasporation Problems," New York: Springer-Verlag, vol. I, Theory. Probability and its Applications.

41) Rockafellar, R. T. and Wets, R. J-B. (1998). "Variational Analysis", Springer, Berlin.

42) Rockafellar, R. T. and Wets, R. J-B. (1998). "Variational Analysis," Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), 317, Springer-Verlag, Berlin.

43) Roshchina, V. (2007). Relationships between upper exhausters and the basic subdifferential in variational analysis, J.Math.Anal.Appl.334, pp. 261-262.

44) Roshchina, V. A. (Feb 2008). Reducing Exhausters. J. Optim. Theory Appl., Vol.136, No.2, 261-273.

45) Roshchina, V. A. (2008). On Conditions for the Minimality of Exhausters. J. Convex Anal. Vol. 15, No. 4.

46) Seidman, I. T. (december 1987). Optimal control for switching systems. In Proceedings of the 21 stannual Conference on information Sciences and systems, pages 485-489, the John Hopkins University, Baltimore.

47) Shapiro, A. (1990). On concepts of directional differentiability, J.Optim. theory Appl.66(3), pp. 477-487.

48) Sussmann, H. J. (1999). "A maximum principle for hybrid optimal control problems" in Proc.38th IEEE Conf. Decision and Control, pp.425-430.

49) Warga, J. (1972). Optimal Control of Differential and Functional Equations, Acad. Pres, New York.

50) Witsenhausen, H. S. (April 1966). A class of hybrid-state continuous-time dynamic systems. IEEE Transactions on Automatic Control, 11(2):161-167.