Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Optimal Control Problems with Nonsmooth Mixed Constraints

M.d.R. de Pinho and G. N. Silva

Abstract— **In this paper we present a weak maximum principle for optimal control problems involving mixed constraints and pointwise set control constraints. Notably such result holds for problems with possibly nonsmooth mixed constraints. Although the setback of such result resides on a convexity assumption on the "extended velocity set", we show that if the number of mixed constraints is one, such convexity assumption may be removed when an interiority assumption holds.**

I. INTRODUCTION

The interest in constrained optimal control problems has witnessed a significant growth in areas like robotics, economics and process systems engineering. In this respect optimal control problems with mixed constraints are of particular importance. One area of application of optimality conditions for such problems is the control of devices modelled by differential algebraic equations (DAE systems). DAE models are nowadays of interest in mechanics and economics and widespread in chemical process engineering ([8], [22], [16]).

As the number of applications increases so does the need to broaden the scope of optimality conditions to cover larger classes of problems ([25], [6], [24]). Necessary conditions in the form of maximum principles for optimal control problems with mixed constraints have been addressed by a number of authors; see for example [13], [18], [19], [12], [15], to name but a few. Maximum principles covering problems with nonsmooth dynamics and some smoothness imposed on mixed constraints have also been considered in [19], [8], [9] and recently in [6] and [10]. To the best of our knowledge the derivation of necessary conditions for nonsmooth mixed constraints remains a largely unexplored area (an exception may be found in [14] where autonomous problems are considered), a surprising fact taking into account the fast development of nonsmooth methods for optimal control since the publication of the seminal book [2].

In this paper we investigate the possibilities of deriving nonsmooth maximum principles for optimal control problems with nonsmooth inequality mixed constraints. Our main result is a weak nonsmooth maximum principle. Although it is obtained under a convexity assumption, we show how, for some special cases, such assumption may be weakened. The problem of interest is (P)

$$\begin{cases} \text{Minimize } l(x(0),x(1)) \\ \text{subject to} \\ \qquad \dot{x}(t) = f(t,x(t),u(t),v(t)) \quad \text{a.e.} \quad t\in[0,1] \\ \qquad 0 \geqslant g(t,x(t),u(t),v(t)) \quad \text{a.e.} \quad t\in[0,1] \\ \qquad v(t) \in V(t) \quad \text{a.e.} \quad t\in[0,1] \\ \qquad (x(0),x(1)) \in C. \end{cases}$$

Here the given functions $f\colon[0,1]\times\mathbb{R}^n\times\mathbb{R}^{k_u}\times\mathbb{R}^{k_v}\to\mathbb{R}^n$, $g\colon[0,1]\times\mathbb{R}^n\times\mathbb{R}^{k_u}\times\mathbb{R}^{k_v}\to\mathbb{R}^m$, and multifunction $V\colon[0,1]\rightrightarrows\mathbb{R}^{k_v}$ describe the system dynamics and control constraints, while the given set $C\subset\mathbb{R}^n\times\mathbb{R}^n$ and function $l\colon\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ specify the endpoint constraints and costs.

A *process* is a triple (x,u,v) comprising a function $x\in W^{1,1}([0,1];\mathbb{R}^n)$ and measurable functions $u\colon[0,1]\to\mathbb{R}^{k_u}$ and $v\colon[0,1]\to\mathbb{R}^{k_v}$. An *admissible process for* (P) is a process satisfying the constraints. Here $W^{1,1}(T;\mathbb{R}^n)$ denotes the space of absolutely continuous functions mapping T to $\mathbb{R}^n$.

An admissible process $(\bar{x},\bar{u},\bar{v})$ is a *minimizer* (also known as *weak minimizer*) for (P) if there exists $\delta'>0$ such that

$$l(\bar{x}(0),\bar{x}(1))\leqslant l(x(0),x(1))$$

holds for all admissible processes (x,u,v) satisfying the following conditions for almost every $t\in[0,1]$:

$$\begin{gathered}|x(t)-\bar{x}(t)|\leqslant\delta', \\ |u(t)-\bar{u}(t)|\leqslant\delta',\ |v(t)-\bar{v}(t)|\leqslant\delta'.\end{gathered} \tag{1}$$

To simplify the exposition, we shall sometimes refer to the control as $w=(u,v)\in\mathbb{R}^k$, where $k=k_u+k_v$, taking values in $w(t)\in W(t)=\mathbb{R}^{k_u}\times V(t)$.

II. PRELIMINARIES

For g in $\mathbb{R}^m$, inequalities like $g\leqslant 0$ are interpreted componentwise. We focus on a particular process $(\bar{x},\bar{w})$, and write $\bar{\phi}(t)$ instead of $\phi(t,\bar{x}(t),\bar{w}(t))$ for both $\phi=f$ and $\phi=g$.

Here and throughout, $\mathbb{B}$ represents the closed unit ball centered at the origin regardless of the dimension of the underlying space and $|\cdot|$ the Euclidean norm or the induced matrix norm on $\mathbb{R}^{p\times q}$. For each t in $[0,1]$ and some $\delta>0$, we define

$$T_\delta(t)=\bar{x}(t)+\delta\mathbb{B}=\{y\in\mathbb{R}^n\ :\ |y-\bar{x}(t)|\leqslant\delta\}\,. \tag{2}$$

The financial support from ISR- Research Unit 147, is gratefully acknowledged

Maria do Rosário de Pinho is with ISR and DEEC, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias , 4200-465 Porto, Portugal mrpinho@fe.up.pt

Geraldo Nunes Silva is with DCCE, IBILCE, Universidade Estadual Paulista, UNESP, São José do Rio Preto, SP, Brazil gsilva@ibilce.unesp.br

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

Likewise we set

$$V_\delta(t) = V(t) \cap \big(\bar{v}(t) + \delta\mathbb{B}\big), \tag{3}$$
$$U_\delta(t) = \bar{u}(t) + \delta\mathbb{B}, \tag{4}$$
$$W_\delta(t) = U_\delta(t) \times V_\delta(t). \tag{5}$$

The *Euclidean distance function* with respect to a given set $A \subset \mathbb{R}^k$ is

$$d_A \colon \mathbb{R}^k \to \mathbb{R}, \qquad y \mapsto d_A(y) = \inf\,\{|y-x| \,:\, x \in A\}\,.$$

A function $h\colon [0,1] \to \mathbb{R}^p$ lies in $W^{1,1}([0,1];\mathbb{R}^p)$ if and only if it is absolutely continuous; in $L^1([0,1];\mathbb{R}^p)$ iff it is integrable; and in $L^\infty([0,1];\mathbb{R}^p)$ iff it is essentially bounded. The norm of $L^1([0,1];\mathbb{R}^p)$ is denoted by $\|\cdot\|_1$ and the norm of $L^\infty([0,1];\mathbb{R}^p)$ is $\|\cdot\|_\infty$.

We make use of standard concepts from nonsmooth analysis. Let $A \subset \mathbb{R}^k$ be a closed set with $\bar{x} \in A$. The *limiting normal cone to* A *at* $\bar{x}$ is denoted by $N_A(\bar{x})$.

Given a lower semicontinuous function $f\colon \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $\bar{x} \in \mathbb{R}^k$ where $f(\bar{x}) < +\infty$, $\partial f(\bar{x})$ denotes the *limiting subdifferential* of f at $\bar{x}$. When the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, co $\partial f(x)$, coincides with the *(Clarke) subdifferential*. Properties of Clarke's subdifferentials (upper semi-continuity, sum rules, etc.), can be found in [2].

For details on such nonsmooth analysis concepts, see [2], [21], [3], [25] and [17].

III. Auxiliary Results

A weak nonsmooth Maximum Principle given by Proposition 6.1 in [7] will play a crucial rule in our analysis. It provides unmaximized Hamiltonian Inclusion (UHI) for standard optimal control problems of the form

$$(S)\quad \begin{cases} \text{Minimize } l(x(1)) + \displaystyle\int_0^1 G(t,x(t),u(t))\,dt \\ \text{subject to} \\ \dot{x}(t) \;=\; f(t,x(t),u(t)) \quad \text{a.e.} \quad t \in [0,1] \\ u(t) \;\in\; U(t) \quad \text{a.e.} \quad t \in [0,1] \\ x(0) \;\in\; C_0, \end{cases}$$

where $U\colon [0,1] \rightrightarrows \mathbb{R}^k$ is a given multifunction.

For the above problem consider the following assumptions which make reference to a reference process $(\bar{x},\bar{u})$ and a parameter $\delta > 0$:

(A1) For each $(x,u) \in \mathbb{R}^n \times \mathbb{R}^k$, the function $t \to (f(t,x,u), G(t,x,u))$ is Lebesgue measurable, and there exist a function $L \in L^1$ such that both $\phi = f$ and $\phi = G$ obey this inequality for almost every t in $[0,1]$:

$$|\phi(t,x,u) - \phi(t,y,w)| \leqslant L(t)\,|(x,u)-(y,w)|$$

for all $x,y \in T_\delta(t),\ u,w \in U(t)$.

(A2) The multifunction U has Borel measurable graph. The set $U_\delta(t)$, defined in (4), is closed for almost every $t \in [0,1]$.

(A3) The endpoint constraint set C_0 is closed; the cost function l is locally Lipschitz in a neighbourhood of $\bar{x}(1)$.

Proposition 3.1: Assume that $(\bar{x},\bar{u})$ is a local minimizer for (S) and assume that there exits a scalar $\delta > 0$ such that A1 – A3 are satisfied. Set

$$H(t,x,p,u) = p \cdot f(t,x,u) - L(t,x,u).$$

Then there exist an absolutely continuous function $p\colon [0,1] \to \mathbb{R}^n$ and an integrable function $\xi\colon [0,1] \to \mathbb{R}^k$ such that

$$(-\dot{p}(t), \xi(t)) \in \\ \operatorname{co}\partial_{x,u} H(t,\bar{x}(t),p(t),\bar{u}(t)) \quad \text{a.e. } t \in [0,1],$$

$$\xi(t) \in \beta(t)\operatorname{co}\partial d_{U_\delta(t)}(\bar{u}(t)) \quad \text{a.e. } t \in [0,1],$$

$$p(0) \in N_{C_0}(\bar{x}(0) \\ -p(1)) \in \partial l(\bar{x}(1)).$$

where $\beta \in L^1$ depends only on δ, L, K_f $(K_f(t) = |\dot{\bar{x}}(t)|)$ and the Lipschitz constant of l (defined in A1–A3).

IV. Main Result

Next we focus on (P).

A. Hypotheses

Define the function

$$g^+(t,x,w) = \max\{0, g_1(t,x,w), \ldots, g_m(t,x,w)\}\,.$$

The following two sets of hypotheses on the data of (P), which make reference to a parameter $\delta > 0$ and a reference process $(\bar{x},\bar{u},\bar{v})$, will be important:

(H1) For each $(x,u,v) \in \mathbb{R}^n \times \mathbb{R}^k$, the function $t \to (f(t,x,u,v), g(t,x,u,v))$ is Lebesgue measurable. Also, there exists a function $L \in L^1$ such that both $\phi = f$ and $\phi = g$ obey this inequality for almost every t in $[0,1]$:

$$|\phi(t,x,u,v) - \phi(t,x',u',v')| \leqslant \\ L(t)\,|(x,u,v)-(x',u',v')|$$

for all $x,x' \in T_\delta(t),\ (u,v),(u',v') \in \mathbb{R}^k$.

(H2) The endpoint constraint set C is closed; the cost function l is locally Lipschitz in a neighbourhood of $(\bar{x}(0),\bar{x}(1))$.

(H3) Both $K_f(t) := |f(t,\bar{x}(t),\bar{u}(t),\bar{v}(t))|$ and $K_g(t) := |g(t,\bar{x}(t),\bar{w}(t),\bar{v}(t))|$ are integrable on $[0,1]$.

(H4) There exist a constant $K_1 > 0$ and a function $h \in L^\infty([0,1];\mathbb{R}^k)$, with $|h(t)| = 1$ a.e., such that the following condition is satisfied for almost every $t \in [0,1]$, all $(x,u,v) \in T_\delta(t) \times U_\delta(t) \times V_\delta(t)$, all $j \in \{1,\ldots,m\}$ and all vectors $(\gamma^j,\psi^j,\phi^j) \in$ co $\partial_{x,u,v} g_j(t,x,u,v)$:

$$(\psi^j,\phi^j) \cdot h(t) \geqslant K_1.$$

(H5) The multifunction V has Borel measurable graph. The set $V_\delta(t)$ as defined in (3), is closed for almost every $t \in [0,1]$.

Consider additionally the following convexity assumption:

(CC) For almost every $t \in [0,1]$, each of the following sets is convex:

$$V^+(t,x) = \{(f(t,x,u,v), g^+(t,x,u,v)+s) \\ u \in U_\delta(t), v \in V_\delta(t),\ s \geqslant 0\}.$$

Hypothesis H4, essential in our setup, is a nonsmooth version of regularity assumptions on the mixed constraints. Assuming smoothness of the function g, H4 coincides with the well known positive linear independence of the gradients $\nabla_v g_i$. In this respect we refer the reader to [12], [20], [6] and [10]).

B. Weak maximum principle

In the context of problem (P), the unmaximized Hamiltonian is the function $H\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \to \mathbb{R}$ defined by

$$H(t,x,p,r,u,v) := p \cdot f(t,x,u,v) + r \cdot g(t,x,u,v). \quad (6)$$

It appears throughout our main result, which describes properties enjoyed by every local minimizer for (P):

Theorem 4.1: Let $(\bar{x},\bar{u},\bar{v})$ be a minimizer for problem (P). Assume H1–H5 and CC. Then there exist an absolutely continuous function $p\colon [0,1] \to \mathbb{R}^n$, integrable functions $\xi\colon [0,1] \to \mathbb{R}^{k_v}$ and $r\colon [0,1] \to \mathbb{R}^m$, and a scalar $\lambda \geqslant 0$ such that

$$\|p\|_\infty + \lambda > 0, \quad (7)$$

$$(-\dot{p}(t), 0, \xi(t)) \in \\ \mathrm{co}\, \partial_{x,u,v} H(t,\bar{x}(t),p(t),r(t),\bar{u}(t),\bar{v}(t)) \text{ a.e. } t, \quad (8)$$

$$\xi(t) \in \beta(t)\, \mathrm{co}\, \partial d_{V_\delta}(\bar{v}(t)) \text{ a.e. } t, \quad (9)$$

$$r(t)\cdot g(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)) = 0 \text{ and } r(t) \leqslant 0 \text{ a.e. } t, \quad (10)$$

$$(p(0), -p(1)) \in N_C(\bar{x}(0),\bar{x}(1)) + \lambda \partial l(\bar{x}(0),\bar{x}(1)), \quad (11)$$

where $\beta \in L^1$ depends only on δ, L, K_f ($K_f(t) = |\dot{\bar{x}}(t)|$) and the Lipschitz constant of l.

The main setback to the application of the above theorem is that hypothesis CC is quite restrictive. The removal of such hypothesis will be the focus of future research. However, as we will see next, in the special case of (P) when the inequality mixed constraint is reduced to one, this assumption can easily be removed.

C. Scalar valued Mixed Constraints

We now consider (P) when the number of inequality mixed constraints is one (i.e., $g : [0,1] \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \to \mathbb{R}$). Consider the following condition:

(INT) For almost every t in $[0,1]$, we have

$$\{w \in W_\delta(t) \,:\, g(t,x,w) \leqslant 0\} \neq \emptyset,$$

for all $x \in T_\delta(t)$.

Theorem 4.2: Let $(\bar{x},\bar{w})$ be a minimizer for problem (P) when $m = 1$ (i.e., $g : [0,1] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$). Assume H1–H5 and INT. Then there exist an absolutely continuous function $p\colon [0,1] \to \mathbb{R}^n$, integrable functions $r\colon [0,1] \to \mathbb{R}^m$, $\xi\colon [0,1] \to \mathbb{R}^{k_v}$ and a scalar $\lambda \geqslant 0$ such that (7)–(11) are satisfied.

V. Proof of Theorem 4.2

We provide an outline of the proof. Details are omitted.

The proof breaks in three steps. We first prove the theorem under the interim hypotheses

(IH) For almost every $t \in [0,1]$, each of the following sets is convex:

$$F(t,x) = \{(f(t,x,w), g(t,x,w)) :\ w \in W_\delta(t)\}$$

for all $x \in T_\delta(t)$.

(ECS) $l(x(0),x(1)) = l(x(1))$ and $C = C_0 \times \mathbb{R}^n$ where $C_0 \subset \mathbb{R}^n$ is a closed set.

Step 1: *Show that IH imply CC.*

Take any $w, w' \in W_\delta(t)$ and any $\alpha \in [0,1]$. Hypothesis IH asserts the existence of $\widetilde{w} \in W_\delta(t)$ such that

$$f(t,x,\widetilde{w}) = \alpha f(t,x,w) + (1-\alpha) f(t,x,w'), \quad (12)$$

$$g(t,x,\widetilde{w}) = \alpha g(t,x,w) + (1-\alpha) g(t,x,w'). \quad (13)$$

Now take any $s, s' \geqslant 0$. We claim that there exists a $\widehat{s} \geqslant 0$ and $\widehat{w} \in W_\delta(t)$ such that

$$f(t,x,\widehat{w}) = \alpha f(t,x,w) + (1-\alpha) f(t,x,w'), \quad (14)$$

$$g_0^+(t,x,\widehat{w}) + \widehat{s} = \\ \alpha g_0^+(t,x,w) + (1-\alpha) g_0^+(t,x,w') + \bar{s}, \quad (15)$$

where $\bar{s} = \alpha s + (1-\alpha)s'$. Consider three cases:

1. If $g(t,x,w), g(t,x,w') \geqslant 0$, then $g(t,x,\widetilde{w}) \geqslant 0$, $g_0^+(t,x,w) = g(t,x,w)$, $g_0^+(t,x,w') = g(t,x,w')$, $g_0^+(t,x,\widetilde{w}) = g(t,x,\widetilde{w})$. By (12) and (13) we deduce that taking $\widehat{w} = \widetilde{w}$ and $\widehat{s} = \bar{s}$, then (14) and (15) hold.
2. Consider now the case when $g(t,x,w), g(t,x,w') \leqslant 0$. It follows from (13) that $g(t,x,\widetilde{w}) \leqslant 0$. Since $g_0^+(t,x,w) = 0$, $g_0^+(t,x,w') = 0$, $g_0^+(t,x,\widetilde{w}) = 0$ we deduce that (14) and (15) hold with $\widehat{w} = \widetilde{w}$ and $\widehat{s} = \bar{s}$.
3. Suppose that $g(t,x,w) < 0$ and $g(t,x,w') \geqslant 0$. Then $g_0^+(t,x,w) = 0$ and $g_0^+(t,x,w') = g(t,x,w')$. Set

$$A = g(t,x,\widetilde{w}), \\ B = \alpha g_0^+(t,x,w) + (1-\alpha) g_0^+(t,x,w'),$$

where $\widetilde{w}$ is such that (12) and (13) hold. Clearly we have

$$B = (1-\alpha)g(t,x,w') \geqslant 0.$$

In view of (13) two different things may happen.

(i) Suppose that we have $A \geqslant 0$. Then $g_0^+(t,x,\widetilde{w}) = A$. Since $B \geqslant 0$ and

$$A = \alpha g(t,x,w) + (1-\alpha)g(t,x,w'),$$

we have, by (13), $0 \leqslant A < B$. Set $\sigma = B - A$. Then

$$B = A + \sigma = g_0^+(t,x,\widetilde{w}) + \sigma.$$

Take $\widehat{w} = \widetilde{w}$ and $\widehat{s} = \sigma + \bar{s} \geqslant 0$. Then (14) and (15) hold.

(ii) Suppose that now $A < 0$. Then $g_0^+(t,x,\widetilde{w}) = 0$. Set $\sigma = B$. Take $\widehat{w} = \widetilde{w}$ and $\widehat{s} = \sigma + \bar{s} \geqslant 0$. Then (14) holds and, since

$$g_0^+(t,x,\widehat{w}) + \widehat{s} = \alpha g_0^+(t,x,w) + (1-\alpha)g_0^+(t,x,w') + \bar{s},$$

(15) also holds.

4. The case when $g(t,x,w) \geqslant 0$ and $g(t,x,w') < 0$ is treated exactly like case 3.

Notice that the fact that g is a scalar valued function was essential.

We conclude that CC holds. Applying Theorem 4.1 we deduce that the conclusions of Theorem 4.2 are satisfied under IH.

Step 2: *Removal of IH.*

By adjusting $\delta > 0$ we can arrange that $(\bar{x}, \bar{w})$ is a minimizer over all admissible processes (x, w) for (P) such that $\|x - \bar{x}\|_\infty \leqslant \delta$ and $\|w - \bar{w}\|_\infty \leqslant \delta$.

Define

$$\widetilde{F}(t,x,y) = \{(f(t,x,w), g(t,x,w) + s) : w \in W_\delta(t), s \geqslant 0\} \bigcap (\mathbb{R}^n \times \{0\}).$$

Under our assumptions $\widetilde{F}(t,x,y)$ is nonempty for each $(t,xy,)$ in the set

$$\Omega = \{(t,x,y) \in [0,1] \times \mathbb{R}^n \times \mathbb{R} \,:\, (x,y) \in (\bar{x}(t),0) + \delta\mathbb{B}\}.$$

Write

$$R := \left\{x \in W^{1,1} \,:\, x(0) \in C_0,\ (\dot{x}(t),0) \in \widetilde{F}(t,x(t),0)\right\}.$$

By the Generalized Filippov Selection Theorem [25, Thm.2.3.13]), $\bar{x}$ is a minimizer for the problem

$$\begin{cases} \text{Minimize } l(x(1)) \\ \text{over arcs } x \text{ in } R \text{ satisfying } \|x-\bar{x}\|_{L^\infty} < \delta. \end{cases}$$

A straightforward modification of the proof of the Relaxation Theorem (see, e.g., [25, Thm.2.7.2]) implies that any arc x in the set

$$R_r := \left\{x \in W^{1,1} \,:\, x(0) \in C_0, (\dot{x}(t),0) \in \operatorname{co} F(t,x(t).0)\right\}$$

which satisfies $\|x - \bar{x}\|_{L^\infty} < \delta$ can be approximated by an arc z in R satisfying $\|z - \bar{x}\|_{L^\infty} < \delta$. The continuity of the mapping

$$x \to l(x(0), x(1))$$

on a neighborhood of $\bar{x}$ implies that $\bar{x}$ is a minimizer for the optimization problem

$$\begin{cases} \text{Minimize } l(x(1)) \\ \text{over arcs } x \in R_r \text{ satisfying } \|x-\bar{x}\|_{L^\infty} < \delta. \end{cases}$$

By the Generalized Filippov Selection Theorem and Carathéodory's Theorem,

$$\{\bar{x}, (\bar{w}_0, \ldots, \bar{w}_M) \equiv (\bar{w}, \ldots, \bar{w}), (\lambda_0, \lambda_1, \ldots, \lambda_M) \equiv (1, 0, \ldots, 0)\}$$

is a minimizer for the optimal control problem (B) defined as

$$\begin{cases} \text{Minimize } l(x(1)) \\ \text{over } x \in W^{1,1} \text{and measurable functions } \lambda_0, \ldots, \lambda_M, \\ w_0, \ldots, w_M \ \text{ satisfying} \\ \dot{x}(t) \ = \displaystyle\sum_i \lambda_i(t) f(t,x(t),w_i(t)), \ \text{a.e.}, \\ 0 \ \geqslant \displaystyle\sum_i \lambda_i(t) g(t,x(t),w_i(t)), \quad \text{a.e.} \\ (\lambda_0(t), \ldots, \lambda_M(t)) \in \Lambda, \\ w_i(t) \in W_\delta(t),\ i = 0, \ldots, M \quad \text{a.e.} \\ x(0) \ \in C. \end{cases}$$

Here

$$\Lambda \ := \ \left\{\lambda'_0, \ldots, \lambda'_M : \lambda'_i \geqslant 0 \text{ for } i = 0, \ldots, M \text{ and } \sum_i^M \lambda'_i = 1\right\},$$

and $(\lambda_0, \ldots, \lambda_M)$, $(w_0, \ldots, w_M)$ are regarded as control variables.

Rewriting the conclusions of Theorem 4.1 when applied to problem (C) gives necessary conditions for our problem (P) when ESC holds.

Step 3: *Validation of the result obtained in Step 2 when hypothesis ECS is removed.*

Step 3. a): We show that the Theorem holds for a problem in the form of (P) but allowing end point constraints of the form $(x(0), x(1)) \in C_0 \times C_1$ (we still consider $l = l(x(1))$).

Let D denote the set of pairs (w, s) such that $w \colon [0,1] \to \mathbb{R}^k$ is a measurable function and $s \in C_1$ for which there exist absolutely continuous functions (x, y) such that

$$\begin{cases} \dot{x}(t) &= & f(t,x(t),w(t)), \quad \text{a.e.}\, t \\ \dot{y}(t) &= & 0, \text{a.e.}\, t \\ 0 &\geqslant & g(t,x(t),w(t)), \quad \text{a.e.}\, t \\ w(t) &\in & W_\delta(t), \quad \text{a.e.}\, t \\ (x(t),y(t) &\in & T_\delta(t) \times T_\delta(t) \quad \text{for all } t \\ (x(0),y(0)) &\in & C_0 \times C_1 \\ x(1) &= & s \end{cases}$$

We provide D with the metric

$$\Delta((w,w'),(s,s')) = \int_0^1 |w(t) - w'(t)|\, dt + |s - s'|,$$

Choose a sequence ε_i such that $\varepsilon_i \downarrow 0$ and $\sum \varepsilon_i < +\infty$, and, for each i, define the function

$$l_i(x,y) = \max\left\{l(y) - l(\bar{x}(1)) + \varepsilon_i^2,\, |x-y|\right\}.$$

Consider

$$(R_i) \quad \begin{cases} \text{Minimize} & l_i(x(1), y(1)) \\ \text{subject to} & (w,s) \in D. \end{cases}$$

Since $(\bar{w}, \bar{x}(1)) \in D$, D is nonempty. It is a simple matter to check that (D, Δ) is a complete metric space on which the functional $l_i \colon D \to \mathbb{R}$ is continuous.

Notice that $l_i(\bar{x}(1), \bar{x}(1)) = \varepsilon_i^2$. Since $l_i \geqslant 0$, it follows that $(\bar{w}, \bar{x}(1))$ is a "ε_i^2-minimizer for (R_i). Then we apply Ekeland's Variational Principle (Theorem 3.3.1 in [25]). Rewriting the conclusions in control theoretical terms we obtain a sequence of perturbed problems to which the necessary conditions obtained in the previous step hold. Taking limits we obtain the required conclusions.

Step 3. b): We now consider problem (P) with cost $l(x(0), x(1))$ and end point constraints $(x(0), x(1)) \in C_0 \times C_1$. For such problem we consider the reformulated problem in which the underlying time interval is $[-1, 1]$.

$$\begin{cases} \text{Minimize } l(z(1), x(1)) \\ \text{subject to} \\ \quad \dot{z}(t) = 0,\ t \in [-1, 0), \\ \quad \dot{z}(t) = 0,\ t \in [0, 1], \\ \quad \dot{x}(t) = 0,\ t \in [-1, 0), \\ \quad \dot{x}(t) = f(t, x(t), w(t)), \\ \quad t \in [0,1] \\ \quad 0 \geqslant \widetilde{g}(t, x(t), w(t)),\ t \in [-1, 0), \\ \quad 0 \geqslant g(t, x(t), w(t)),\ t \in [0, 1] \\ \quad w(t) \in (-1, \dots, -1) + 1/2\mathbb{B}, \quad t \in [-1, 0), \\ \quad w(t) \in W_\delta(t), \quad t \in [0, 1] \\ \quad (z(-1), x(-1)) \in C_1' \\ \quad (z(1), x(1)) \in C_0 \times C_1. \end{cases}$$

where $\widetilde{g}(t,x,w) = Dw$, D is a $m \times k$ constant matrix such that each column vector has all entries equal 0 except one that is equal to 1 and $C_1' = \{(a,b) \in R^n \times \mathbb{R}^n : a = b\}$. Applying the result as obtained in Step 3. a) we get the required necessary conditions.

Step 3. c): Finally we generalize the Theorem to cover the case when $(x(0), x(1)) \in C$ for a closed subset $C \subset \mathbb{R}^n \times \mathbb{R}^n$.

A standard state-augmentation trick converts problem (P) as stated into a problem with separated endpoint constraints. It suffices to introduce an additional state $y \in \mathbb{R}^n$ with dynamics $\dot{y}(t) = 0$ and to impose the modified endpoint constraints

$$(x(0), y(0)) \in C,$$
$$(x(1), y(1)) \in \left\{(x,y) \in \mathbb{R}^{2n} : x = y\right\}.$$

The results already obtained apply to the augmented problem, and the stated result for (P) is easily extracted from them.

The proof is complete.

VI. SKETCH OF THE PROOF OF MAIN RESULT

We now present a brief sketch of the proof of Theorem 4.1. For details see [11].

Notice that the local minimality of $(\bar{x}, \bar{u})$ provides some $\delta' > 0$ as described in the paragraph containing (1). By reducing this constant if necessary, we can also rely on the properties in CC.

Step 1 First the Theorem is proved for problem Q:

$$\begin{cases} \text{Minimize } l(x(1)) \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), w(t)) \text{ a.e. } t \\ \quad 0 \geqslant g(t, x(t), w(t)) \text{ a.e. } t \\ \quad w(t) \in W(t) \text{ a.e. } t \\ \quad x(0) \in C_0. \end{cases}$$

Problem (Q) is a special case of (P) in which $C = C_0 \times \mathbb{R}^n$ and $l(x_0, x_1) = l(x_1)$.

For each fixed $i \in \mathbb{N}$, consider also the optimal control problem

$$(\mathrm{Q_i}) \begin{cases} \text{Minimize } l(x(1)) + i\displaystyle\int_0^1 g^+(t, x(t), w(t))\, dt \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), w(t)) \text{ a.e. } t \\ \quad w(t) \in W_\delta(t) \text{ a.e. } t \\ \quad x(0) \in C_0. \end{cases}$$

This differs from (Q) by explicitly localizing the control constraint around $\bar{w}$, and by shifting the mixed constraints into the objective function. The integral

$$i \int_0^1 g^+(t, x(t), w(t))\, dt.$$

penalizes violation of the mixed constraints. We temporarily assume that

(IH2) $\liminf_{i \to \infty} (Q_i) = \inf(Q)$.

holds. In the final stage of this first step we show that CC implies IH2.

Let E denote the set of pairs (u, s), where $s \in \mathbb{R}^n$ and $u \colon [0,1] \to \mathbb{R}^k$ is measurable, for which there exist absolutely continuous functions x such that

$$\begin{cases} \dot{x}(t) = f(t, x(t), w(t)) \text{ a.e. } t \\ w(t) \in W_\delta(t) \text{ a.e. } t \\ x(t) \in T_\delta(t) \text{ a.e. } t \\ x(0) \in C_0 \\ x(1) = s \end{cases}$$

We provide E with the metric

$$\Delta((w,s),(w',s')) = \int_0^1 |w(t) - w'(t)|\, dt + |s - s'|,$$

and define $J_i: W \to \mathbb{R}$ using the arc x mentioned above:

$$J_i(u,s) = l(x(1)) + i\int_0^1 g^+(t, x(t), w(t))\,dt.$$

Moreover, problem (Q_i) above is closely related to the abstract problem

$$(R_i) \quad \begin{cases} \text{Minimize} & J_i(w,s) \\ \text{subject to} & (w,s) \in E. \end{cases}$$

Clearly $(\bar{w}, \bar{x}(1))$ is admissible for (R_i), with

$$J_i(\bar{w}, \bar{x}(1)) = l(\bar{x}(1)) = \inf P$$

since $g^+(t, \bar{x}(t), \bar{w}(t)) = 0$ for almost every $t \in [0,1]$. Application of Ekeland's Variational Principle to (R_i) leads to a sequence of perturbed optimal control problems to which Proposition 3.1 applies. Taking limits we obtain necessary conditions for (Q). Finally we show that CC implies IH2.

Step 2: We extend the results obtained in the previous step to cover the more general problem (P). This is accomplished by repeating the approach used in Step 3 of the proof of Theorem 4.2.

VII. ACKNOWLEDGEMENTS

Maria do Rosário de Pinho would like to thank the Department of Computer Science and Statistics and IBILCE at the Universidade Estadual Pulista (UNESP) São José do Rio Preto, São Paulo, Brazil for the invitation and very warm welcome and stay that allowed to discuss the ideas that resulted in this piece of work with Geraldo Nunes Silva.

REFERENCES

[1] A. V. Arutyunov, *Optimality Conditions. Abnormal and Degenerate Problems*, Mathematics and its Applications 526, Kluwer Academic Publishers, Dordrecht, 2000.
[2] F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.
[3] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Graduate Texts in Mathematics, vol. 178, Springer-Verla, New York, 1998.
[4] M. R. de Pinho and R. B. Vinter, An Euler-Lagrange inclusion for optimal control problems, *IEEE Trans. Automat. Control*, vol 40, 1995, pp 1191-1198.
[5] M. R. de Pinho and A. Ilchmann, Weak maximum principle for optimal control problems with mixed constraints, *Nonlinear Anal.*, vol 48, 2002, pp 1179-1196.
[6] F. Clarke, The maximum principle in optimal control, then and now, *Control Cybernet.*, vol. 34, 2005, pp 709-722.
[7] M. R. de Pinho and R. B. Vinter, An Euler-Lagrange inclusion for Optimal Control Problems, *IEEE Trans. Automat. Control*, vol 40, 1995, pp 1191-1198.
[8] M. R. de Pinho, R. B. Vinter, and H. Zheng, A maximum principle for optimal control problems with mixed constraints, *IMA J. Math. Control Inform.*, vol 18, 2001, pp 189-205.
[9] M. d. R. de Pinho , Mixed constrained control problems, *J. Math. Anal. Appl.*, vol. 278, 2003, pp 293-307.
[10] M.d.R. de Pinho and J. F. Rosenblueth, Necessary Conditions for Constrained Problems under Mangasarian-Fromowitz Conditions*SIAM J. Control Optim.*, vol 47, 2008, pp 535-552.
[11] M.d.R. de Pinho, P. Loewen and G. N. Silva, A weak maximum principle for optimal control problems with nonsmooth mixed constraints, *submitted to Set Valued Analysis*, 2008.
[12] A. V. Dmitruk, Maximum principle for the general optimal control problem with phase and regular mixed constraints, *Comput. Math. Model.*, vol 4, 1993, pp 364-377.
[13] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1966.
[14] E.N. Devdaryani and Y. S. Ledyaev , Maximum principle for implicit control systems, *Appl. Math. Optim.*, vol. 40, 1996, pp 79-103.
[15] A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs 180, American Mathematical Society, Providence, Rhode Island, 1998.
[16] Boris S. Mordukhovich and Lianwen Wang, Optimal control of delay systems with differential and algebraic dynamic constraints, *ESAIM, COCV*, vol 11, 2005, pp 285-309.
[17] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I. Applications*, Fundamental Principles of Mathematical Sciences 330, Springer-Verlag, Berlin, 2006.
[18] L. W. Neustadt, *Optimization, A Theory of Necessary Conditions*, Princeton University Press, New Jersey, 1976.
[19] Z. Páles and V. Zeidan, First- and second-order necessary conditions for control problems with constraints, *Trans. Amer. Math. Soc.*, vol. 346, 1994, pp 421–453.
[20] Z. Páles and V. Zeidan, Optimal control problems with set-valued control and state constraints, *SIAM J. Optim.*, vol 14, 2003, pp 334-358.
[21] R. T. Rockafellar and B. Wets,*Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, 1998.
[22] Tomás Roubícek and Michael Valásek, Optimal control of causal differential-algebraic systems, *J. Math. Anal. Appl.*, vol 269, 2002, pp 616-641.
[23] G. Stefani and P. L. Zezza, Optimality conditions for a constrained control problem, *SIAM J. Control Optim.*, vol 34, 1996, pp 635-659.
[24] Hector J. Sussmann, A strong version of the Maximum Principle under weak hypothesis, in Proceedings of the 33th IEEE Conference on Decision and Control, Orlando, Florida, 1994.
[25] Richard Vinter, *Optimal Control*, Systems Control Found. Appl., Birkhäuser, Boston, 2000.