ACADEMIC PRESS

J. Math. Anal. Appl. ●●● (●●●●) ●●●–●●●

Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS

www.academicpress.com

Mixed constrained control problems

Maria do Rosário de Pinho

Faculdade de Engenharia da Universidade do Porto, DEEC/ISR, Rua Dr. Roberto Frias s/n, 4200 465 Porto, Portugal

Received 19 November 2001

Submitted by H. Frankowska

Abstract

Necessary optimality conditions are derived in the form of a weak maximum principle for optimal control problems with mixed state-control equality and inequality constraints. In contrast to previous work these conditions hold when the Jacobian of the active constraints, with respect to the unconstrained control variable, has full rank. A feature of these conditions is that they are stated in terms of a joint Clarke subdifferential. Furthermore the use of the joint subdifferential gives sufficiency for nonsmooth, normal, linear convex problems. The main point of interest is not only the full rank condition assumption but also the nature of the analysis employed in this paper. A key element is the removal of the constraints and application of Ekeland's variational principle.

Keywords: Optimal control; Mixed state and control constraints; Maximum principle

1. Introduction

In this paper we focus on the derivation of necessary conditions of optimality for certain optimal control problems involving mixed state and control constraints in the form of both inequalities and equalities under assumptions which are, in some sense, minimal.

E-mail address: mrpinho@fe.up.pt.

0022-247X/02/$ – see front matter
PII: S0022-247X(02)00351-7

The problem of interest is the following:

$$
\text{(P)}\quad \begin{cases} \text{Minimize } l\big(x(0), x(1)\big) + \int_0^1 L\big(t, x(t), u(t), v(t)\big)\,dt \\ \text{subject to} \\ \dot{x}(t) = f\big(t, x(t), u(t), v(t)\big) & \text{a.e. } t \in [0,1], \\ 0 = b\big(t, x(t), u(t), v(t)\big) & \text{a.e. } t \in [0,1], \\ 0 \geqslant g\big(t, x(t), u(t), v(t)\big) & \text{a.e. } t \in [0,1], \\ v(t) \in V(t) & \text{a.e. } t \in [0,1], \\ x(0) \in C_0, \\ x(1) \in C_1, \end{cases}
$$

where $l:\Re^n \times \Re^n \to \Re$, $L:[0,1]\times\Re^n\times\Re^{k_u}\times\Re^{k_v}\to\Re$, $f:[0,1]\times\Re^n\times\Re^{k_u}\times\Re^{k_v}\to\Re^n$, $b:[0,1]\times\Re^n\times\Re^{k_u}\times\Re^{k_v}\to\Re^{m_b}$, $g:[0,1]\times\Re^n\times\Re^{k_u}\times\Re^{k_v}\to\Re^{m_g}$ are given functions, C_0, $C_1 \subset \Re^n$ given sets and $V:[0,1]\to\Re^{k_v}$ is a given multifunction. We set $k = k_u + k_v$ and $m = m_b + m_g$. Throughout this paper we assume that $k_u \geqslant m$. The control variable comprises two components, u and v. When no distinction between such components is needed we may refer to the control variable simply as $w = (u, v) \in W(t)$, where $W:[0,1]\to\Re^k$ is a given multifunction.

For (P) a process is a pair (x, u, v) comprising measurable control functions u and v and an absolutely continuous function $x \in W^{1,1}([0,1],\Re^n)$ satisfying the constraints of the problem. A process (x, u, v) of (P) is called a *weak local minimizer* if there exists some $\varepsilon > 0$, such that it minimizes the cost over all processes of (P) which satisfy $(x(t), u(t), v(t)) \in T_\varepsilon(t)$ for a.e. $t \in [0,1]$, where $T_\varepsilon(t) = (\bar{x}(t), \bar{u}(t), \bar{v}(t)) + \varepsilon B$, B denotes the closed unit ball and $(\bar{x}(t), \bar{u}(t), \bar{v}(t))$ is a reference process.

Assume that $(\bar{x}(t), \bar{u}(t), \bar{v}(t))$ is an optimal process for (P) and let $\bar{f}(t)$, $\bar{b}(t)$, etc., denote the corresponding function evaluated at $(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$. Given two functions ϕ and φ, $[\phi, \varphi](\cdot)$ denotes the function $(\phi(\cdot), \varphi(\cdot))$.

Necessary conditions in the form of weak maximum principles for (P) have previously been derived assuming full rankness of a given matrix $F(t)$, i.e., assuming that $\det F(t)F(t)^{\mathrm{T}} \geqslant L$ for a.e. $t \in [a,b]$, for some $L > 0$. The matrix F has been considered to be $\Upsilon_1(t) = \nabla_u[\bar{b}, \bar{g}](t)$ in [7] and [9],

$$
\Upsilon_2(t) = \begin{bmatrix} \bar{b}_u(t) \\ \bar{g}_u^{\mathcal{I}_\beta(t)}(t) \end{bmatrix}
$$

in [10], where $\mathcal{I}_\beta(t) = \{i \in \{1,\dots,m_g\}:\ \bar{g}_i(t) \geqslant -\beta\}$ and $\bar{g}_u^{\mathcal{I}_\beta(t)}(t)$ denotes the matrix we obtain after removing from $g_u(t)$ all the rows of index $i \notin \mathcal{I}_\beta(t)$, and

$$
\Upsilon_3(t) = \begin{bmatrix} \bar{b}_u(t) & 0 \\ \bar{g}_u(t) & \operatorname{diag}\{-\bar{g}_i(t)\}_{i\in\{1,\dots,m_g\}} \end{bmatrix}
$$

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.3(1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 3

in [12]. The full rank condition imposed on Υ_1, Υ_2 and Υ_3 are related to each other and they are sufficient for the matrix

$$\Upsilon_4(t) = \begin{bmatrix} \bar{b}_u(t) \\ \bar{g}_u^{\mathcal{I}_a(t)}(t) \end{bmatrix} \tag{1}$$

to be of full rank (see [4]), where

$$\mathcal{I}_a(t) = \left\{ i \in \{1, \ldots, m_g\}: \quad \bar{g}_i(t) = 0 \right\} \tag{2}$$

is the set of active constraints and, as before, $\bar{g}_u^{\mathcal{I}_a(t)}(t)$ denotes the matrix we obtain after removing from $\bar{g}_u(t)$ all the rows of index $i \notin \mathcal{I}_a(t)$.

Matrix $\Upsilon_4(t)$ is of interest. In fact, necessary conditions in the form of weak maximum principles are such that the derivative with respect to u of $\bar{g}_i$, for $i \notin \mathcal{I}_a(t)$, does not take any part in the determination of the multipliers. It is then reasonable to conjecture that necessary conditions of optimality for (P) hold when the full rankness condition is imposed merely on matrix $\Upsilon_4(t)$. In this paper we show that this is indeed the case. A common approach to derive necessary conditions for (P) under a full rankness condition is to associate with (P) a problem with only equality mixed constraints. Such approach is no longer valid when full rank is imposed on $\Upsilon_4(t)$. The key idea behind the proof of our main result is based on a technique explored in [13] for optimal control problems with pure state constraints. It consists on the definition of a sequence of standard optimal control problems in which the constraints are incorporated both in the cost and in the dynamics.

As in [4], here we consider problems with nonsmooth data. Optimality conditions obtained are stated in terms of a "joint" subdifferential $\operatorname{co} \partial_{x,p,u,v} H_\lambda(t, x, u, v, p, q, r)$. This special feature is of interest for linear convex problems. Indeed, nonsmooth weak maximum principles commonly stated in terms of $\operatorname{co} \partial_x H_\lambda \times \operatorname{co} \partial_p H_\lambda \times \operatorname{co} \partial_u H_\lambda \times \operatorname{co} \partial_v H_\lambda$ can fail to provide sufficiency. An example is given in [6]. Here, we show that our necessary conditions, under a normality hypothesis, are sufficient for optimality in the normal nonsmooth linear-convex case.

2. Preliminaries

Here and throughout, B represents the closed unit ball centered at the origin. The notation $r \geqslant 0$ means that each component r_i of $r \in \Re^r$ is nonnegative. $|\cdot| = \sqrt{\langle \cdot, \cdot \rangle}$ denotes the Euclidean norm, and $|\cdot|$ the induced matrix norm on $\Re^{m \times k}$. The linear space $W^{1,1}([0,1]; \Re^p)$ denotes the space of absolutely continuous functions, $L^1([0,1]; \Re^p)$ and $L^\infty([a,b]; \Re^p)$ denote respectively the space of integrable functions and the space of essentially bounded functions from $[0,1]$ to $\Re^p$. Since we assume only measurability of the data with respect to t, a variant of a uniform implicit function theorem, derived in [3], will be essential in our setup. Such theorem asserts that if $\phi(t, x_0(t), u_0(t)) = 0$ almost everywhere, then an implicit function $\varphi(t,u)$ exists and the *same* neighborhood of u_0 can be chosen for all t.

We make use of constructs from nonsmooth analysis of *limiting normal cone to* a closed set A at x, written $N_A(x)$, and *limiting subdifferential* of a semicontinuous function f at x, written $\partial f(x)$. When the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, $\operatorname{co}\partial f(x)$, which may be defined directly, coincides with the *Clarke subdifferential*. The full calculus for limiting subdifferential and limiting normal cones in finite dimensions can be found, for example, in [8,11,13]. Properties of Clarke's subdifferentials (upper semi-continuity, sum rules, etc.), can be found in [1].

Consider a standard optimal control problem

$$
\text{(S)}\quad \begin{cases} \text{Minimize } l\big(x(0),x(1)\big)+\int_0^1 L\big(t,x(t),w(t)\big)\,dt \\ \text{subject to} \\ \dot{x}(t)=f\big(t,x(t),w(t)\big) \quad \text{a.e. } t\in[0,1], \\ w(t)\in W(t) \quad \text{a.e. } t\in[0,1], \\ x(0)\in C_0, \\ rx(1)\in C_1, \end{cases}
$$

where l, L, f, C_0 and C_1 are as defined before for (P) and $W:[0,1]\to\Re^k$ is a given multifunction. Assume that $(\bar{x},\overline{w})$ is a reference process of (S) and $\varepsilon>0$ a parameter. We invoke the following hypotheses on the data of (S):

(H1) $[L,f](\cdot,x,w)$ is measurable for each (x,w) and $[L,f](t,\cdot,\cdot)$ is Lipschitz continuous with Lipschitz constant $k_f(t)$ on $(\bar{x}(t),\overline{w}(t))+\varepsilon B$ for almost every $t\in[0,1]$ and k_f is an L^1-function.

(H2) The cost l is Lipschitz continuous on a neighborhood of $(\bar{x}(0),\bar{x}(1))$ and C_0 and C_1 are closed.

(H3) The multifunction W has Borel measurable graph and, for almost every $t\in[0,1]$, the set $W_\varepsilon(t):=(\overline{w}(t)+\varepsilon B)\cap W(t)$ is closed.

The following Euler–Lagrange inclusion for (S), provided in [2], will also be of importance in our analysis.

Proposition 2.1. *Let* $(\bar{x},\overline{w})$ *denote a weak local minimizer for* (S). *If* (H1)–(H3) *are satisfied and* $H(t,x,p,w)=p\cdot f(t,x,w)-\lambda L(t,x,w)$ *defines the Hamiltonian, then there exist* $\lambda\geqslant 0$, $p\in W^{1,1}([0,1];\Re^n)$ *and* $\zeta\in L^1([0,1];\Re^k)$ *such that, for almost every* $t\in[0,1]$,

$$
\lambda+\|p\|_{L^\infty}=1,
$$

$$
\big(-\dot{p}(t),\dot{\bar{x}}(t),\zeta(t)\big)\in\operatorname{co}\partial H\big(t,\bar{x}(t),p(t),\overline{w}(t)\big),
$$

$$
\zeta(t)\in\operatorname{co}N_{W(t)}\big(\overline{w}(t)\big),
$$

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.5 (1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 5

$$\big(p(0), -p(1)\big) \in N_{C_0 \times C_1}\big(\bar{x}(0), \bar{x}(1)\big) + \lambda \partial l\big(\bar{x}(0), \bar{x}(1)\big),$$

where ∂H denotes the limiting subdifferential in the (x, p, w) variables.

3. Main results

We now concentrate on (P). With reference to a process $(\bar{x}, \bar{u}, \bar{v})$ of (P) and a parameter $\varepsilon > 0$, we invoke hypotheses (H1)–(H3) stated in the previous section (for $\overline{w} = (\bar{u}, \bar{v})$ and $W(t) = \Re^{k_u} \times V(t)$), and the following additional hypotheses:

(H4) $b(\cdot, x, u, v)$ and $g(\cdot, x, u, v)$ are measurable for each (x, u, v). There exists an L^1 function $L_{b,g}$ such that, for almost every $t \in [0, 1]$, $[b, g](t, \cdot, \cdot, \cdot)$ is continuously differentiable and Lipschitz continuous with Lipschitz constant $L_{b,g}(t)$ on $T_\varepsilon(t)$.

(H5) There exists an increasing function $\tilde{\theta} : \Re^+ \to \Re^+$, $\tilde{\theta}(s) \downarrow 0$ as $s \downarrow 0$, such that, for all (x', u', v'), $(x, u, v) \in T_\varepsilon(t)$ and for almost every $t \in [0, 1]$,

$$\big|\nabla_{x,u,v}[b, g](t, x', u', v') - \nabla_{x,u,v}[b, g](t, x, u, v)\big| \leqslant \tilde{\theta}\big(\big|(x', u', v') - (x, u, v)\big|\big).$$

There exists $K_{b,g} > 0$ such that, for almost every $t \in [0, 1]$,

$$\big|\nabla_x\big[\bar{b}, \bar{g}\big](t)\big| + \big|\nabla_{u,v}\big[\bar{b}, \bar{g}\big](t)\big| \leqslant K_{b,g}.$$

(H6) There exists $K > 0$ such that, for almost every $t \in [0, 1]$,

$$\det\big\{\Upsilon_4(t)\Upsilon_4^{\mathrm{T}}(t)\big\} \geqslant K \quad \text{where } \Upsilon_4(t) = \begin{bmatrix} \bar{b}_u(t) \\ \bar{g}_u^{\mathcal{I}_a(t)}(t) \end{bmatrix}.$$

Theorem 3.1. *Let* $(\bar{x}, \bar{u}, \bar{v})$ *be a* weak local minimizer *for* (P). *If, for some* $\varepsilon > 0$, *hypotheses* (H1)–(H6) *are satisfied and*

$$H_\lambda(t, x, p, q, r, u, v) := p \cdot f(t, x, u, v) + q \cdot b(t, x, u, v) + r \cdot g(t, x, u, v) - \lambda L(t, x, u, v)$$

defines the Hamiltonian, then there exist $p \in W^{1,1}([0, 1]; \Re^n)$, $\xi \in L^1([0, 1]; \Re^{k_v})$ *and* $\lambda \geqslant 0$ *such that, for almost every* $t \in [0, 1]$,

(i) $\|p\|_{L_\infty} + \lambda \neq 0$,
(ii) $(-\dot{p}(t), \dot{x}(t), 0, \xi(t)) \in \operatorname{co} \partial_{x,p,u,v} H_\lambda(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t), \bar{v}(t))$,
(iii) $\xi(t) \in \operatorname{co} N_{V(t)}(\bar{v}(t))$,
(iv) $r(t) \cdot g(t, \bar{x}(t), \bar{u}(t)) = 0$ *and* $r(t) \leqslant 0$,
(v) $(p(0), -p(1)) \in N_{C_0}(\bar{x}(0)) \times N_{C_1}(\bar{x}(1)) + \lambda\, \partial l(\bar{x}(0), \bar{x}(1))$.

Furthermore, there exist an integrable function K_Q and a constant $C_Q > 0$ such that

$$\left|\left(q(t), r(t)\right)\right| \leqslant K_Q(t)\left|p(t)\right| + C_Q\left|p(1)\right| \quad \textit{for a.e. } t \in [0,1].$$

The novelty of the theorem is that, unlike previously proved results (see [4,7, 9,10,12]) the conclusions hold when the full rank condition is imposed merely on matrix Υ_4. Recall that, as mentioned in the Introduction, the full rank of matrix Υ_4 does not necessarily imply the full rank of other matrices used in the existing literature. Additionally the Theorem 3.1 provides necessary conditions of optimality for problem (P) with data possibly nonsmooth and they are stated in terms of a "joint" subdifferential (see (ii)).

Necessary conditions of optimality in terms of a "joint" subdifferential like those of Theorem 3.1 derived for standard optimal control problems and optimal control problems with state constraints (see [2] and [5], respectively) are also sufficient conditions for linear-convex problems in the normal form. This is also the case for linear-convex problems with mixed constrained problems as we show next. By linear-convex problems we mean problem (P) where

$$f\left(t, x(t), u(t), v(t)\right) = A(t)x(t) + B(t)u(t) + C(t)v(t),$$
$$b\left(t, x(t), u(t), v(t)\right) = D(t)x(t) + E(t)u(t) + F(t)v(t),$$
$$g\left(t, x(t), u(t), v(t)\right) = G(t)x(t) + J(t)u(t) + K(t)v(t),$$

and the following hypotheses hold:

(HC1) C_0 and C_1 are convex, $V(t)$ is convex for a.e. $t \in [0,1]$.
(HC2) The function $t \to L(\cdot, x, u, v)$ is measurable and l and $(x,u,v) \to L(t,x,u,v)$ are convex.
(HC2) The functions A, D, G, E, F, J and K are integrable and B and C are measurable. We denote such problem by (LC)

Proposition 3.1. *Let $(\bar{x}, \bar{u}, \bar{v})$ be a process for problem (LC). Assume that* (HC1)–(HC3) *are in force and that $(\bar{x}, \bar{u}, \bar{v})$ is a normal extremal in the sense that there exist $p \in W^{1,1}$, ζ, q, $r \in L^1$ such that the conditions*

$$\left(-\dot{p}(t), \dot{\bar{x}}(t), 0, \zeta(t)\right) \in \operatorname{co} \partial H_\lambda\left(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t), \bar{v}(t)\right) \quad \textit{a.e.}, \tag{3}$$

$$\zeta(t) \in \operatorname{co} N_{V(t)}\left(\bar{v}(t)\right) \quad \textit{a.e.}, \tag{4}$$

$$r(t) \leqslant 0 \quad \textit{and} \quad r(t)\cdot\left(G(t)\bar{x}(t) + J(t)\bar{u}(t) + K(t)\bar{v}(t)\right) = 0 \quad \textit{a.e.}, \tag{5}$$

$$\left(p(0), -p(1)\right) \in N_{C_0\times C_1}\left(\bar{x}(0), \bar{x}(1)\right) \tag{6}$$

are satisfied for $\lambda = 1$. Then $(\bar{x}, \bar{u}, \bar{v})$ is a weak local minimizer.

In the Proposition above, subdifferentials and normal cones are understood in the sense of convex analysis.

Proof. Here we follow the approach of the proof of Proposition 4.1 in [2]. Details are omitted.

Let (x,u,v) be an arbitrary admissible process and the pair $(\bar{x},\bar{u},\bar{v})$ be a normal extremal for (LC). We compute the difference of the cost between (x,u,v) and $(\bar{x},\bar{u},\bar{v})$. In doing so the following remarks will be of help. By definition of a process $t\to L(t,x(t),u(t),v(t))$, $t\to B(t)u(t)$ and $t\to C(t)v(t)$ are integrable. By (5) and since (x,u,v) is a process for (LC) we have

$$r\cdot\big(G(x-\bar{x})+J(u-\bar{u})+K(v-\bar{v})\big)\geqslant 0\quad \text{a.e.} \tag{7}$$

From (6), there exists a $\sigma\in N_{C_0\times C_1}(\bar{x}(0),\bar{x}(1))$ such that

$$\big(p(0),-p(1)\big)-\sigma\in\partial l\big(\bar{x}(0),\bar{x}(1)\big). \tag{8}$$

Finally, recall also that for a generic convex function f, we have

$$f(x)-f(\bar{x})-\zeta\cdot(x-\bar{x})\geqslant 0,\quad \forall\zeta\in\partial f(\bar{x}). \tag{9}$$

Taking into account (3)–(6) and (7), the difference of the cost between (x,u,v) and $(\bar{x},\bar{u},\bar{v})$

$$\begin{aligned}\Delta=&\int_0^1\Big(L\big(t,x(t),u(t),v(t)\big)-L\big(t,\bar{x}(t),\bar{u}(t),\bar{v}(t)\big)\Big)\,dt\\&+l\big(x(0),x(1)\big)-l\big(\bar{x}(0),\bar{x}(1)\big)\end{aligned}$$

is computed as in the proof of Proposition 4.1 in [2] leading to

$$\begin{aligned}\Delta&\geqslant l\big(x(0),x(1)\big)-l\big(\bar{x}(0),\bar{x}(1)\big)+\int_0^1\frac{d}{dt}p\cdot(x-\bar{x})\,dt\\&=l\big(x(0),x(1)\big)-l\big(\bar{x}(0),\bar{x}(1)\big)\\&\quad-\big(p(0),-p(1)\big)\cdot\big(x(0)-\bar{x}(0),x(1)-\bar{x}(1)\big)\\&\quad+\sigma\cdot\big(x(0)-\bar{x}(0),x(1)-\bar{x}(1)\big)-\sigma\cdot\big(x(0)-\bar{x}(0),x(1)-\bar{x}(1)\big)\end{aligned}$$

(8) and (9) ensure that this last expression is nonnegative, proving the proposition. □

4. Proof of Theorem 3.1

Recall the definition of the set of active constraints $\mathcal{I}_a(t)$ (see (2)) and let $\mathcal{I}_c(t)=\{i\in\{1,\dots,m_g\}\colon\ \bar{g}_i(t)<0\}$ be its complement. Set $q_a(t)$ to be the cardinal of $\mathcal{I}_a(t)$ and $q_c(t)$ the cardinal of $\mathcal{I}_c(t)$. Note that $q_a(t)+q_c(t)=m_g$. For $i\in\{1,\dots,m_g\}$, define $\delta_i:[0,1]\to\Re$ to be

$$\delta_i(t)=\begin{cases}1 & \text{if } i\in\mathcal{I}_a(t),\\ 0 & \text{if } i\in\mathcal{I}_c(t),\end{cases} \tag{10}$$

and consider the following matrices

$$\Delta(t)=\operatorname{diag}\{\delta_i(t)\}_{i=1}^{m_g},\qquad \Delta'(t)=I-\Delta(t),$$
$$\overline{\Gamma}_x(t)=\begin{bmatrix}\bar b_x(t)\\ \Delta(t)\bar g_x(t)\end{bmatrix},\qquad \overline{\Gamma}_u(t)=\begin{bmatrix}\bar b_u(t)\\ \Delta(t)\bar g_u(t)\end{bmatrix},$$
$$\overline{\Gamma}_v(t)=\begin{bmatrix}\bar b_v(t)\\ \Delta(t)\bar g_v(t)\end{bmatrix},\qquad \Gamma_\alpha(t)=\begin{bmatrix}0\\ \Delta'(t)\end{bmatrix},$$
$$\Gamma_\beta(t)=\begin{bmatrix}0\\ \Delta(t)\end{bmatrix},\qquad D(t)=\begin{bmatrix}1 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix},$$

where $D(t)\in\Re^{q_c(t)\times q_c(t)}$.

Let $\bar\beta,\bar\alpha:[0,1]\to\Re^{m_g}$ be defined componentwise as $\bar\beta_i(t)=-\bar g_i(t)$ and $\bar\alpha_i(t)=0$. The functions $\bar\beta$ and $\bar\alpha$ are measurable. Let α,β be measurable functions. Take $(\bar x,\bar u,\bar v)$ to be the process in Theorem 3.1 and $\varepsilon>0$ the parameter. We prove the theorem in the case $L\equiv 0$. This restriction is lifted by the use of well known augmentation techniques and an appeal to standard estimation on limiting subdifferentials; details are omitted.

We start the proof which breaks into steps. The proof consists on two major steps. Firstly, an uniform implicit function d is determined by applying an uniform implicit function theorem previously proved in [3] to the active mixed constraints. This function allows us to define a sequence of optimal control problems to which Ekeland's variational principle applies.

Step 1. We apply a uniform implicit function theorem to a function $\mu:[0,1]\times\Re^n\times\Re^k\times\Re^{m_b}\times\Re^{m_g}\times\Re^m\to\Re^m$ in order to obtain an "uniform" implicit function d.

Let μ be defined as

$$\begin{aligned}&\mu\big(t,\xi,(u,v),\alpha,\beta,\eta\big)\\ &\quad=b\big(t,\bar x(t)+\xi,\bar u(t)+u+\overline{\Gamma}_u(t)^{\mathrm T}\eta,\bar v(t)+v\big)\\ &\qquad\Delta(t)\big[g\big(t,\bar x(t)+\xi,\bar u(t)+u+\overline{\Gamma}_u(t)^{\mathrm T}\eta,\bar v(t)+v\big)+\bar\beta(t)+\beta\big]\\ &\qquad+\Phi(t,\alpha,\eta),\end{aligned} \tag{11}$$

where $\Phi(t,\alpha,\eta) = \Delta'(t)(\alpha + \Gamma_\alpha(t)^{\mathrm{T}}\eta)$. μ satisfies the conditions under which Theorem 3.2 in [3] applies. Observe that if the components of g are permute in such a way that the active constraints come first, we have

$$\overline{\Gamma}(t) = \frac{\partial\mu}{\partial\eta}(t,0,0,0,0,0) = \begin{bmatrix} \Upsilon_4(t)\Upsilon_4(t)^{\mathrm{T}} & 0 \\ 0 & D(t)D(t)^{\mathrm{T}} \end{bmatrix}, \tag{12}$$

where $\Upsilon_4(t)$ is as defined in (H6). It follows that there exists a constant $C > 0$ such that

$$\big|\big[\overline{\Gamma}(t)\big]^{-1}\big| \leqslant C \tag{13}$$

for almost every $t \in [0,1]$.

Theorem 3.2 in [3] ensures the existence of $\sigma \in (0,\varepsilon)$, $\delta \in (0,\varepsilon)$ and a mapping $d:[0,1] \times (\sigma B) \times (\sigma B) \times (\sigma B) \times (\sigma B) \times (\sigma B) \to \delta B$ such that $d(\cdot,\xi,u,v,\alpha,\beta)$ is a measurable function for fixed (ξ,u,v,α,β), the functions $\{d(t,\cdot,\cdot,\cdot,\cdot,\cdot):\ t \in [0,1]\}$ are Lipschitz continuous with common Lipschitz constant, $d(t,\cdot,\cdot,\cdot,\cdot,\cdot)$ is continuously differentiable for fixed t, $d(t,0,0,0,0,0) = 0$ for almost every $t \in [0,1]$ and, for all $(\xi,u,v,\alpha,\beta) \in \sigma B \times \sigma B \times \sigma B \times \sigma B \times \sigma B$ and for almost every $t \in [0,1]$,

$$\mu\big(t,(\xi,u,v,\alpha,\beta),d(t,\xi,u,v,\alpha,\beta)\big) = 0, \tag{14}$$

$$\nabla_{\xi,u,v,\alpha,\beta} d(t,0,0,0,0,0) = -\big[\overline{\Gamma}(t)\big]^{-1}\big(\overline{\Gamma}_x(t),\overline{\Gamma}_u(t),\overline{\Gamma}_v(t),\overline{\Gamma}_\alpha(t),\overline{\Gamma}_\beta(t)\big). \tag{15}$$

Choose $\sigma_1, \delta_1 > 0$ such that

$$\begin{cases} \sigma_1 \in \big(0,\min\{\sigma,\varepsilon/2\}\big), \\ \delta_1 \in \big(0,\min\{\delta,\varepsilon/2\}\big), \\ \sigma_1 + K_{b,g}\delta_1 \in (0,\varepsilon/2), \end{cases} \tag{16}$$

where $K_{b,g}$ is given by (H5).

Step 2. We now define an optimization problem to which Ekeland's theorem applies.

Define the functions

$$g^+(t,x,u,v) = \max\big\{0,\ g_1(t,x,u,v),\ldots,g_{m_g}(t,x,u,v)\big\},$$
$$f_1(t,x,u,v,\alpha,\beta) = f\big(t,x,u+\overline{\Gamma}_u(t)^{\mathrm{T}}\tilde{d}(t),v\big),$$
$$f_2(t,x,u,v,\alpha,\beta) = g^+\big(t,x,u+\overline{\Gamma}_u(t)^{\mathrm{T}}\tilde{d}(t),v\big),$$

where $\tilde{d}(t) = d(t,x-\bar{x}(t),u-\bar{u}(t),v-\bar{v}(t),\alpha-\bar{\alpha}(t),\beta-\bar{\beta}(t))$. Define also the sets $\mathcal{B} = \{\xi \in \Re^{m_g}:\ \xi_i \geqslant 0,\ i = \{i,\ldots,m_g\}\}$, $\mathcal{B}_{\sigma_1}(t) = \mathcal{B} \cap (\bar{\beta}(t) + \sigma_1 B)$, $V_{\sigma_1}(t) = V(t) \cap (\bar{v}(t) + \sigma_1 B)$.

Consider a sequence of positive scalars $\{\varepsilon_k\}_{k\in\mathbb{N}}$, such that $\lim_{k\to\infty}\varepsilon_k = 0$ and set

$$\Psi_k(x,y,x',y',z) = \max\big\{l(x,y) - l\big(\bar{x}(0),\bar{x}(1)\big) + \varepsilon_k^2, z + |x'-y'|\big\}.$$

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.10 (1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 10

Take W to be the set of all measurable functions (u, v, α, β) and all vectors $(a, b, c, e, h) \in \Re^n \times \Re^n \times \Re^n \times \Re^n \times \Re$ such that, for almost every $t \in [0, 1]$, $(u(t), \alpha(t)) \in (\bar{u}(t) + \sigma_1 B) \times \sigma_1 B$, $(v(t), \beta(t)) \in V_{\sigma_1}(t) \times \mathcal{B}_{\sigma_1}(t)$, $(a, b) \in C_0 \times C_1$ and for which there exist absolutely continuous functions x, y and z such that

$$\begin{aligned} &\dot{x}(t) = f_1(t, x, u, v, \beta, \alpha) && \text{a.e.,} \\ &\dot{y}(t) = 0 && \text{a.e.,} \\ &\dot{z}(t) = f_2(t, x, u, v, \beta, \alpha) && \text{a.e.,} \\ &\big(x(t), y(t), z(t)\big) \in \big(\bar{x}(t), \bar{x}(1), 0\big) + \sigma_1 B && \text{a.e.,} \\ &\big(x(0), y(0), z(0)\big) = (a, b, 0), \\ &\big(x(1), y(1), z(1)\big) = (c, e, h). \end{aligned}$$

To simplify the notation set $E = (a, b, c, e, h) \in \Re^{4n+1}$. Let

$$|E - E'| = |a - a'| + |b - b'| + |c - c'| + |e - e'| + |h - h'|$$

and

$$\begin{aligned} &\nu\big((u, v, \alpha, \beta), (u', v', \alpha', \beta')\big) \\ &\quad = \int_0^1 \big|u(t) - u'(t)\big|\, dt + \int_0^1 \big|v(t) - v'(t)\big|\, dt + \int_0^1 \big|\alpha(t) - \alpha'(t)\big|\, dt \\ &\qquad + \int_0^1 \big|\beta(t) - \beta'(t)\big|\, dt. \end{aligned}$$

Define $d_W : W \times W \to \Re$,

$$\begin{aligned} &d_W\big((u, v, \beta, \alpha, E), (u', v', \beta', \alpha', E')\big) \\ &\quad = \nu\big((u, v, \alpha, \beta), (u', v', \alpha', \beta')\big) + |E - E'|. \end{aligned}$$

Consider the sequence of optimization problems

$$(\mathrm{R_k}) \quad \begin{cases} \text{Minimize } J_k(u, v, \alpha, \beta, E) \\ \text{subject to } (u, v, \alpha, \beta, E) \in W, \end{cases}$$

where $J_k(u, v, \alpha, \beta, E) = \Psi_k(x(0), y(0), x(1), y(1), z(1))$. Observe that d_W defines a metric in W and, with respect to this metric, the set W is a complete metric space and the function $(u, v, \alpha, \beta, E) \to J_k(u, v, \alpha, \beta, E)$ is continuous on (W, d_W).

Let $\overline{E} = (\bar{x}(0), \bar{x}(1), \bar{x}(1), \bar{x}(1), 0)$. For all $k \in \mathbb{N}$, $J_k(u, v, \alpha, \beta, E) \geqslant 0$ and $J_k(\bar{u}, \bar{v}, \bar{\alpha}, \bar{\beta}, \overline{E}) = \varepsilon_k^2$. It follows that $(\bar{u}, \bar{v}, \bar{\alpha}, \bar{\beta}, \overline{E})$ is an "ε_k^2-minimizer" for $(\mathrm{R_k})$. According to Ekeland's variational principle (see [13]), there exists a sequence $(u_k, v_k, \alpha_k, \beta_k, E_k) \in W$ such that, for each $k \in \mathbb{N}$,

$$d_W\big((u_k, v_k, \alpha_k, \beta_k, E_k), \big(\bar{u}, \bar{v}, \bar{\alpha}, \bar{\beta}, \overline{E}\big)\big) \leqslant \varepsilon_k \tag{17}$$

and $(u_k, v_k, \alpha_k, \beta_k, E_k)$ minimizes the perturbed cost function $J_k(u, v, \alpha, \beta, E) + \varepsilon_k d_W((u_k, v_k, \alpha_k, \beta_k, E_k), (u, v, \alpha, \beta, E))$ for all $(u, v, \alpha, \beta, E) \in W$.

Step 3. Rewriting the conclusions of Ekeland's theorem in control theoretic terms we obtain a sequence of standard optimal control problems.

Write (x_k, y_k, z_k) the trajectory corresponding to $(u_k, v_k, \alpha_k, \beta_k, E_k)$. For each $k \in \mathbb{N}$, the process $(x_k, y_k, z_k, w_1 \equiv 0, w_2 \equiv 0, w_3 \equiv 0, w_4 \equiv 0, u_k, v_k, \alpha_k, \beta_k)$ solves the control problem $(\mathrm{C_k})$:

$$\begin{cases} \text{Minimize} \\ \Psi_k\big(x(0), y(0), x(1), y(1), z(1)\big) + \varepsilon_k\big|x(0) - x_k(0)\big| + \varepsilon_k\big|x(1) - x_k(1)\big| \\ \quad + \varepsilon_k\big|y(0) - y_k(0)\big| + \varepsilon_k\big|y(1) - y_k(1)\big| + \varepsilon_k\big|z(1) - z_k(1)\big| \\ \quad + \varepsilon_k w_1(1) + \varepsilon_k w_2(1) + \varepsilon_k w_3(1) + \varepsilon_k w_4(1) \\ \text{subject to} \\ \dot{x}(t) = f_1(t, x, u, \alpha, \beta), \qquad \dot{y}(t) = 0, \\ \dot{z}(t) = f_2(t, x, u, \alpha, \beta), \\ \dot{w}_1(t) = \big|u(t) - u_k(t)\big|, \qquad \dot{w}_2(t) = \big|v(t) - v_k(t)\big|, \\ \dot{w}_3(t) = \big|\alpha(t) - \alpha_k(t)\big|, \qquad \dot{w}_4(t) = \big|\beta(t) - \beta_k(t)\big|, \\ \big(x(t), y(t), z(t)\big) \in \big(\bar{x}(t), \bar{x}(1), 0\big) + \sigma_1 B, \\ \big(u(t), v(t), \alpha(t), \beta(t)\big) \in \big(\bar{u}(t) + \sigma_1 B\big) \times V_{\sigma_1}(t) \times \sigma_1 B \times \mathcal{B}_{\sigma_1}(t), \\ \big(x(0), y(0), z(0)\big) \in C_0 \times C_1 \times \{0\}, \\ \big(w_1(0), w_2(0), w_3(0), w_4(0)\big) = (0, 0, 0, 0), \end{cases}$$

where all the equalities and inclusions but the last two are to be understood in an *almost everywhere* sense.

Since $\varepsilon_k \to 0$, we have from (17) that $(x_k, y_k, z_k) \to (\bar{x}, \bar{x}(1), 0)$ uniformly. By discarding initial terms of the sequence, if necessary, we have $(x_k(t), y_k(t), z_k(t)) \in (\bar{x}(t), \bar{x}(1), 0) + (\sigma_1/2)B$ for all k. Then $(x_k, y_k, z_k, w_1 \equiv 0, w_2 \equiv 0, w_3 \equiv 0, w_4 \equiv 0, u_k, v_k, \alpha_k, \beta_k)$ is a weak local minimizer of a variant of $(\mathrm{C_k})$ obtained by dropping the state constraint "$(x(t), y(t), z(t)) \in (\bar{x}(t), \bar{x}(1), 0) + \sigma_1 B$."

Step 4. We obtain necessary conditions to problem $(\mathrm{C_k})$.

Lemma 4.1. *Let* $(x_k, y_k, z_k, 0, 0, 0, 0, u_k, v_k, \alpha_k, \beta_k)$ *be a weak local minimizer for problem* $(\mathrm{C_k})$. *Set* $H(t, x, p, r, u, v, \alpha, \beta) = p \cdot f_1(t, x, u, v, \alpha, \beta) + r \cdot f_2(t, x, u, v, \alpha, \beta)$. *Then there exist scalars* λ_k, η_k *and* r_k, *vectors* $q_k, e_k \in \Re^n$, *integrable functions* $\xi_k : [0, 1] \to \Re^{k_v}$, $\zeta_k : [0, 1] \to \Re^{m_g}$ *and an absolutely continuous function* $p_k \in W^{1,1}$ *such that*:

(a) $\lambda_k + \|p_k\|_\infty + |q_k| + |r_k| + 4\lambda_k\varepsilon_k = 1$,

(b) $\lambda_k \geqslant 0, \qquad \eta_k \in [0, 1], \qquad |e_k| = 1,$

(c) $$\big(-p_k(1)-\lambda_k(1-\eta_k)e_k, -q_k+\lambda_k(1-\eta_k)e_k, -r_k-\lambda_k(1-\eta_k)\big) \in \lambda_k\varepsilon_k(B\times B\times B),$$

(d) $$\big(p_k(0), q_k\big) \in N_{C_0}\big(x_k(0)\big)\times N_{C_1}(y_k)+\lambda_k\eta_k\partial l\big(x_k(0), y_k(0)\big) + \lambda_k\varepsilon_k(B\times B),$$

(e) $$\zeta_k(t)\in \operatorname{co} N_{\mathcal{B}_{\sigma_1(t)}}\big(\beta_k(t)\big), \qquad \xi_k(t)\in \operatorname{co} N_{V_{\sigma_1(t)}}\big(v_k(t)\big) \quad a.e.,$$

(f) $$\big(-\dot p_k(t), \dot x_k(t), \dot z_k(t), 0, \xi_k(t), 0, \zeta_k(t)\big) \in \operatorname{co}\partial H\big(t, x_k(t), p_k(t), r_k, u_k(t), v_k(t), \alpha_k(t), \beta_k(t)\big) + \lambda_k\varepsilon_k\big(\{0\}\times\{0\}\times\{0\}\times B\times B\times B\times B\big) \quad a.e.$$

Proof. We only give a sketch of the proof. For details see [6]. It is a simple matter to verify that the hypotheses are satisfied under which Proposition 2.1 applies to $(\mathrm{C_k})$. The set of necessary conditions we obtain by applying Proposition 2.1 to $(\mathrm{C_k})$ can be rewritten as in the lemma by observing the following:

(i) The Hamiltonian for $(\mathrm{C_k})$ is

$$\begin{aligned} &h(t,x,y,z,w_1,w_2,w_3,w_4,p,q,r,\pi_1,\pi_2,\pi_3,\pi_4,u,v,\alpha,\beta)\\ &\quad = H(t,x,p,q,r,u,v,\alpha,\beta)+q\cdot 0+\pi_1|u-u_k|+\pi_2|v-v_k|\\ &\qquad +\pi_3|\alpha-\alpha_k|+\pi_4|\beta-\beta_k|. \end{aligned}$$

(ii) For i sufficiently large, it can easily be shown that

$$\Psi_k\big(x_k(0), y_k, x_k(1), y_k, z_k(1)\big) > 0.$$

(iii) The max rule guarantees that there exists $\eta_k\in[0,1]$ such that

$$\begin{aligned} &\partial\Psi_k\big(x_k(0), y_k, x_k(1), y_k, z_k(1)\big)\\ &\quad \subset \eta_k\partial l\big(x_k(0), y_k(0)\big)\times\{0,0,0,0,0,0,0,0,0,0,0,0\}\\ &\qquad +(1-\eta_k)\{0,0,0,0,0,0,0\}\times\big\{(e,-e)\colon e\in\Re^n, |e|=1\big\}\\ &\qquad\quad \times\{1,0,0,0,0\}. \qquad \square \end{aligned}$$

Step 5. We now consider $\varepsilon_k\to 0$ and we take limits to obtain necessary conditions for (P).

Recall that $(x_k, y_k, z_k)\to(\bar x, \bar x(1), 0)$ uniformly. Since $(u_k, v_k, \alpha_k, \beta_k)\to(\bar u, \bar v, 0, \bar\beta)$ strongly in L^1 we can arrange by subsequence extraction that $(u_k, v_k, \alpha_k, \beta_k)\to(\bar u, \bar v, 0, \bar\beta)$ almost everywhere. It also follows that $\tilde d(t)\to 0$, where

$$\begin{aligned} \tilde d_k(t) = d\big(&t, x_k(t)-\bar x(t), u_k(t)-\bar u(t), v_k(t)-\bar v(t), \alpha_k(t)-\bar\alpha(t),\\ &\beta_k(t)-\bar\beta(t)\big) \end{aligned}$$

and, consequently, $\tilde u_k(t) = u_k(t)+\overline{\Gamma}_u(t)^{\mathrm T}\tilde d_k(t)\to\bar u(t)$ almost everywhere.

The sequences $\{e_k\}$ and $\{\eta_k\}$ are uniformly bounded by (b) of the Lemma 4.1. Since $\varepsilon_k \to 0$, we conclude from (a) of the Lemma 4.1 that λ_k is also uniformly bounded. We can therefore arrange, again by subsequence extraction if necessary, that $e_k \to e$, $\lambda_k \to \lambda$ and $\eta_k \to \eta$, where $|e| = 1$, $\lambda \geqslant 0$ and $\eta \in [0, 1]$.

Now the sequences $\{p_k\}$, $\{t \to \int_0^t \xi_k\, ds\}$ and $\{t \to \int_0^t \zeta_k\, ds\}$ are equicontinuous and uniformly bounded and $\{\dot{p}_k\}$ is uniformly integrable bounded. Standard compactness arguments and an appeal to Dunford–Pettis criterion for L^1 compactness ensure that, by further extraction of subsequences if necessary, $p_k \to p$ uniformly, $\int_0^t \zeta_k\, ds \to \int_0^t \zeta\, ds$ uniformly, $\int_0^t \xi_k\, ds \to \int_0^t \xi\, ds$ uniformly for some $p \in W^{1,1}$ and $\xi,\ \zeta \in L^1$, and $\dot{p}_k \to \dot{p}$, $\xi_k \to \xi$ and $\zeta_k \to \zeta$ weakly in L^1.

It now follows from the above and from (a), (c) and (d) of Lemma 4.1 that $r_k \to r = -\lambda(1-\eta)$, $q_k \to q = \lambda(1-\eta)e$, $p(1) = -\lambda(1-\eta)e$,

$$\big(p(0), -p(1)\big) \in N_{C_0}\big(\bar{x}(0)\big) \times N_{C_1}\big(\bar{x}(1)\big) + \lambda\eta\partial l\big(\bar{x}(0), \bar{x}(1)\big), \tag{18}$$

and $\lambda + \|p\|_\infty + 2|p(1)| = 1$ (observe that $p(1) = -q$). Since $|p(1)| = \lambda(1-\eta)$, we get $\lambda\eta + 3|p(1)| + \|p\|_\infty = 1$, a condition which ensures that $\tilde{\lambda} + \|p\|_\infty \neq 0$, where $\tilde{\lambda} = \lambda\eta \geqslant 0$.

The properties of the limiting normal cone and limiting subdifferential and an application of Theorem 3.1.7 of [1] allow us to pass to the limit in relationships (e) and (f) of Lemma 4.1. There results

$$\begin{aligned} &\big(-\dot{p}(t), \dot{\bar{x}}(t), 0, 0, \xi(t), 0, \zeta(t)\big) \\ &\quad \in \operatorname{co} \partial H\big(t, \bar{x}(t), p(t), r, \bar{u}(t), \bar{v}(t), 0, \bar{\beta}(t)\big) \end{aligned} \tag{19}$$

and

$$\zeta(t) \in \operatorname{co} N_{\mathcal{B}_{\sigma_1}(t)}\big(\bar{\beta}(t)\big), \qquad \xi(t) \in \operatorname{co} N_{V_{\sigma_1}(t)}\big(\bar{v}(t)\big). \tag{20}$$

We conclude that there exist scalars $\tilde{\lambda} \geqslant 0$ and $r \leqslant 0$ and functions $p \in W^{1,1}$, $\zeta, \xi \in L^1$ such that

(A′) $\tilde{\lambda} + \|p\|_\infty \neq 0$,
(B′) $(-\dot{p}(t), \dot{\bar{x}}(t), 0, 0, \xi(t), 0, \zeta(t)) \in \operatorname{co} \partial H(t, \bar{x}(t), p(t), r, \bar{u}(t), \bar{v}(t), 0, \bar{\beta}(t))$ a.e.,
(C′) $\zeta(t) \in \operatorname{co} N_{\mathcal{B}_{\sigma_1}(t)}(\bar{\beta}(t))$ a.e.,
(D′) $\xi(t) \in \operatorname{co} N_{V_{\sigma_1}(t)}(\bar{v}(t))$ a.e.,
(E′) $(p(0), -p(1)) \in N_{C_0}(\bar{x}(0)) \times N_{C_1}(\bar{x}(1)) + \tilde{\lambda}\partial l(\bar{x}(0), \bar{x}(1))$,

where $H(t, x, p, r, u, v, \alpha, \beta) = p \cdot f_1(t, x, u, v, \alpha, \beta) + r f_2(t, x, u, v, \alpha, \beta)$.

Step 6. Finally we rewrite relationships (A′)–(E′) in the required form.
Observe that

$$N_{\mathcal{B}_{\sigma_1}}\big(\bar{\beta}\big) = \big\{\theta \in \Re^{m_g} \colon\ \theta_i = 0 \text{ if } i \in \mathcal{I}_c(t),\ \theta_i \leqslant 0 \text{ if } i \in \mathcal{I}_a(t)\big\}.$$

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.14 (1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 14

Then, from (C′) and the definition of $\bar\beta$ we deduce that

$$\zeta_i(t)=0 \quad \text{if } i\in\mathcal{I}_c(t), \qquad \zeta_i(t)\leqslant 0 \quad \text{if } i\in\mathcal{I}_a(t). \tag{21}$$

We deduce from the max rule (applied to f_2), the chain rule (see [1]) and the differentiable properties of d the following estimation for $\operatorname{co}\partial H$:

$$\begin{aligned}
&\operatorname{co}\partial H(t,x,p,r,u,v,\alpha,\beta)\\
&\quad\subset\Big\{\Big(\mu-\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_x^{\mathrm T}(t)+r\gamma(t)\cdot g_x(t)\\
&\qquad\quad -r\big(\gamma(t)\cdot g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_x^{\mathrm T}(t),\\
&\qquad\quad \nu,0,\rho-\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_u^{\mathrm T}(t)+r\gamma(t)\cdot g_u(t)\\
&\qquad\quad -r\big(\gamma(t)\cdot g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_u^{\mathrm T}(t),\\
&\qquad\quad \varrho-\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_v^{\mathrm T}(t)+r\gamma(t)\cdot g_v(t)\\
&\qquad\quad -r\big(\gamma(t)\cdot g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_v^{\mathrm T}(t),\\
&\qquad\quad -\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_\alpha^{\mathrm T}(t)-r\big(\gamma(t)\cdot g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_\alpha^{\mathrm T}(t),\\
&\qquad\quad -\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_\beta^{\mathrm T}(t)-r\big(\gamma(t)\cdot g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_\beta^{\mathrm T}(t)\Big):\\
&\qquad\quad (\mu,\nu,0,\rho,\varrho)\in\operatorname{co}\partial p\cdot f,\ \gamma=(\gamma_1,\ldots,\gamma_{m_g})\text{ measurable},\\
&\qquad\quad \gamma_i(t)\in[0,1],\ \gamma_i(t)\geqslant 0\text{ if }\bar g_i(t)=0,\ \gamma_i(t)=0\text{ if }\bar g_i(t)\leqslant 0\Big\}
\end{aligned}$$

in which f, g_i, g_x, etc., are evaluated at

$$\big(t,x,u+\bar u(t)+\overline{\Gamma}_u^{\mathrm T}(t)d\big(t,x-\bar x(t),u-\bar u(t),v-\bar v(t),\alpha-\bar\alpha,\beta-\bar\beta(t)\big)\big).$$

Appealing to an appropriate selection theorem, we deduce existence of measurable functions

$$\big(\mu(t),\nu(t),0,\rho(t),\varrho(t)\big)\in\operatorname{co}\partial_{x,p,r,u,v}p\cdot f \quad \text{a.e.}, \tag{22}$$

$$\gamma=\big\{(\gamma_1,\ldots,\gamma_{m_g}):\ \gamma_i(t)\in[0,1]\big\}\quad\text{and}$$

$$\begin{cases}\gamma_i(t)\geqslant 0 & \text{if } \bar g_i(t)=0,\\ \gamma_i(t)=0 & \text{if } \bar g_i(t)\leqslant 0\end{cases}\quad \text{a.e.}, \tag{23}$$

$$\zeta(t)\in\operatorname{co}N_{\mathcal{B}_{\sigma_1}(t)}\big(\bar\beta(t)\big)\quad\text{a.e.}, \tag{24}$$

$$\xi(t)\in\operatorname{co}N_{V_{\sigma_1}(t)}\big(\bar v(t)\big)\quad\text{a.e.}, \tag{25}$$

such that

$$\begin{aligned}
&\big(-\dot p(t),\dot x(t),0,0,\xi(t),0,\zeta(t)\big)\\
&\quad=\Big(\mu-\rho\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_x^{\mathrm T}(t)+r\gamma(t)\cdot\bar g_x(t)\\
&\qquad\quad -r\big(\gamma(t)\cdot\bar g_u(t)\big)\overline{\Gamma}_u^{\mathrm T}(t)\overline{\Gamma}^{-1}(t)\overline{\Gamma}_x^{\mathrm T}(t),
\end{aligned}$$

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.15 (1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 15

$$\begin{aligned}&\nu, 0, \rho-\rho \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_u^{\mathrm{T}}(t)+r \gamma(t) \cdot \bar{g}_u(t)\\&-r\big(\gamma(t) \cdot \bar{g}_u(t)\big) \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_u^{\mathrm{T}}(t),\\&\varrho-\rho \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_v^{\mathrm{T}}(t)+r \gamma(t) \cdot \bar{g}_v(t)\\&-r\big(\gamma(t) \cdot \bar{g}_u(t)\big) \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_v^{\mathrm{T}}(t),\\&-\rho \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_\alpha^{\mathrm{T}}(t)-r\big(\gamma(t) \cdot \bar{g}_u(t)\big) \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_\alpha^{\mathrm{T}}(t),\\&-\rho \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_\beta^{\mathrm{T}}(t)-r\big(\gamma(t) \cdot \bar{g}_u(t)\big) \bar{\Gamma}_u^{\mathrm{T}}(t) \bar{\Gamma}^{-1}(t) \bar{\Gamma}_\beta^{\mathrm{T}}(t)\Big)\end{aligned} \tag{26}$$

in which f, g, g_x, etc. are evaluated at $(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ (observe that $d(t, 0, 0, 0, 0, 0) = 0$). Under the hypotheses μ, ν, ρ, ϱ, γ and ζ are all integrable functions.

Define $\widetilde{Q}(t) = -\bar{\Gamma}_u^{\mathrm{T}}(t)\bar{\Gamma}^{-1}(t) \in \Re^{k\times m}$ and $\tilde{r}(t) = r\gamma(t) \in \Re^{m_g}$. Then $\tilde{r} \leqslant 0$. Take $\widehat{Q}$ to be the L^1 function:

$$\widehat{Q}(t) = \big(\hat{q}_1(t), \hat{q}_2(t)\big) = \rho(t)\widetilde{Q}(t) + \tilde{r}(t)\bar{g}_u(t)\widetilde{Q}(t).$$

In terms of $\widehat{Q}$ and $\tilde{r}$, (26) becomes

$$\begin{aligned}&\big(-\dot{p}(t), \dot{x}(t), 0, 0, \xi(t), 0, \zeta(t)\big)\\&\quad=\Big(\mu+\widehat{Q}(t)\bar{\Gamma}_x^{\mathrm{T}}(t)+\tilde{r}(t)\bar{g}_x(t), \nu, 0, \rho+\widehat{Q}(t)\bar{\Gamma}_u^{\mathrm{T}}(t)+\tilde{r}(t)\bar{g}_u(t),\\&\qquad\varrho+\widehat{Q}(t)\bar{\Gamma}_v^{\mathrm{T}}(t)+\tilde{r}(t)\bar{g}_v(t), \widehat{Q}(t)\bar{\Gamma}_\alpha^{\mathrm{T}}(t), \widehat{Q}(t)\bar{\Gamma}_\beta^{\mathrm{T}}(t)\Big).\end{aligned} \tag{27}$$

Since $r \leqslant 0$, by definition of γ we conclude that $\tilde{r}_i(t) = 0$ if $i \in \mathcal{I}_c(t)$ and $\tilde{r}_i(t) \leqslant 0$ if $i \in \mathcal{I}_a(t)$. Furthermore, by definition of $\bar{\Gamma}_\alpha^{\mathrm{T}}(t)$, $\bar{\Gamma}_\beta^{\mathrm{T}}(t)$ and by (21) and (27) we deduce that $\hat{q}_{2_i}(t) = 0$ if $i \in \mathcal{I}_c(t)$ and $\hat{q}_{2_i}(t) \leqslant 0$ if $i \in \mathcal{I}_a(t)$. Set $q(t) = \hat{q}_1(t)$ and $r(t) = \tilde{r}(t) + \hat{q}_2(t)$. Then

$$\begin{aligned}&r_i(t) = 0 \quad \text{if } g_i\big(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)\big) < 0,\\&r_i(t) \leqslant 0 \quad \text{if } g_i\big(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)\big) = 0.\end{aligned} \tag{28}$$

Since b and g are smooth we conclude from (27) that

$$\begin{aligned}&\big(-\dot{p}(t), \dot{x}(t), 0, \xi(t)\big)\\&\quad\in \operatorname{co} \partial_{x,p,u,v}\big\{p \cdot f(t,x,u,v) + q \cdot b(t,x,u,v) + r \cdot g(t,x,u,v)\big\}\end{aligned} \tag{29}$$

in which the subdifferential is evaluated at $(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$. $\tilde{\lambda}$, p, q and r obey the state relationships; (A′), (D′) and (E′) are respectively (i), (iii) and (v) of the theorem with $\lambda = \tilde{\lambda}$, (28) yields (iv) and (29) coincides with (ii) when $L \equiv 0$. As for the final assertion,

$$\big|\big(q(t), r(t)\big)\big| \leqslant k_f(t) K_{b,g} C \big|p(t)\big| + \big(1 + K_{b,g}^2 C\big)\big|p(1)\big|$$

this follows from the definition of q, r, (H1), (H4), (13) and (22). The proof is complete.

S0022-247X(02)00351-7/FLA AID:8171 Vol.●●●(●●●) Yjmaa8171 P.16 (1-16)
ELSGMLTM(JMAA):mNY1s 2002/08/30 Prn:11/09/2002; 15:38 by:R.M. p. 16

Acknowledgment

The financial support from Projecto FCT PRAXIS C/EEI/14156/1998 is gratefully acknowledged.

References

[1] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[2] M.d.R. de Pinho, R.B. Vinter, An Euler–Lagrange inclusion for optimal control problems, IEEE Trans. Automat. Control AC-40 (1995) 1191–1198.
[3] M.d.R. de Pinho, R.B. Vinter, Necessary conditions for optimal control problems involving nonlinear differential algebraic equations, J. Math. Anal. Appl. 212 (1997) 493–516.
[4] M.d.R. de Pinho, A. Ilchmann, Weak maximum principle for optimal control problems with mixed constraints, Nonlinear Anal. 48 (2002) 1179–1196.
[5] M.d.R. de Pinho, M.M.A. Ferreira, F.A.C.C. Fontes, An Euler–Lagrange inclusion for optimal control problems with state constraints, J. Dynam. Control Systems 8 (2002) 23–45.
[6] M.d.R. de Pinho, A report on optimality conditions for optimal control problems with mixed constraints, Internal Report, FEUP, ISR, Porto, Portugal, 2001.
[7] M.R. Hestenes, Calculus of Variations and Optimal Control Theory, Wiley, New York, 1966.
[8] B.S. Mordukhovich, Generalized differential calculus for nonsmooth and set-valued mappings, J. Math. Anal. Appl. 183 (1994) 250–288.
[9] L.W. Neustadt, Optimization, A Theory of Necessary Conditions, Princeton University Press, Princeton, NJ, 1976.
[10] N.P. Osmolovskii, Second order conditions for a weak local minimum in an optimal control problem (necessity and sufficiency), Soviet Math. Dokl. 16 (1975) 1480–1484.
[11] R.T. Rockafellar, B. Wets, Variational Analysis, Springer-Verlag, Berlin, 1998.
[12] Z. Pales, V. Zeidan, First- and second-order necessary conditions for control problems with constraints, Trans. Amer. Math. Soc. 346 (1994) 421–453.
[13] R.B. Vinter, Optimal Control, Birkhäuser, Boston, 2002.