AN ACCESSORY PROBLEM FOR OPTIMAL CONTROL PROBLEMS WITH MIXED CONSTRAINTS

MARIA DO ROSÁRIO DE PINHO* and DALILA FONTES†

* Departamento de Engenharia Electrotecnica e de Comp., Inst. Sistemas e Rob., Faculdade de Eng. Universidade do Porto, Rua dos Bragas, 4050 123 Porto, Portugal, mrpinho@fe.up.pt.

† Faculdade de Economia Universidade do Porto, Rua Dr. Roberto Frias, 4200 464 Porto, Portugal, fontes@fep.up.pt.

Abstract

Nonlinear optimal control problems can be associated with linear quadratic optimal control problems. Such relation are of interest when deriving second order necessary conditions for the original problem in terms of conjugate points.

In this work we consider an optimal control problem with mixed constraints in the form of equalities and inequalities to which we associate a linear quadratic optimal control problem. For the proposed auxiliary problem we prove that the cost is in fact nonnegative. The proof of the main result highlights the relation between nonnegativeness of the cost of the associated problem and the optimality of the solution of the original problem.

Key words: Optimal control, mixed constraints, accessory problem.

1 Introduction

Consider the following optimal control problem (P):

$$\left\{\begin{array}{l} \text{Minimize } l(x(b)) + \int_a^b L(t,x(t),u(t))dt \\ \text{subject to} \\ \begin{array}{rcll} \dot{x}(t) & = & f(t,x(t),u(t)) & \text{a.e.} \\ 0 & = & b(t,x(t),u(t)) & \text{a.e.} \\ 0 & \geq & g(t,x(t),u(t)) & \text{a.e.} \\ x(a) & = & x_a & \\ x(b) & \in & C & \end{array} \end{array}\right.$$

where $l : \mathbb{R}^n \to \mathbb{R}$, $L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$, $f : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$, $b : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{m_b}$, $g : [a,b] \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{m_g}$, and $C \subset \Re^n$ a set defined as

$$C = \{x \in \Re^n : h(x) = 0\},$$

where $h : \Re^n \to \Re^r$ and $r \leq n$.

Throughout this paper we assume that $k \geq m_b + m_g$.

(P) is an optimal control problem involving equality and inequality state-dependent control constraints, also known as mixed state-control constraints.

Following standard procedure applied when deriving second order conditions (see, for example, [8]), we associate with (P) an auxiliary problem, which is to minimise a second variation $J_2(v)$ over all solutions (y,v) of some linearised system. The "auxiliary" problem we consider has been proposed before, for example, in [5], [8] and [2].

Although here we do not derive second order conditions (those can be found, for example, in [2]) we show that if $(\bar{x},\bar{u})$ solves (P), then the proposed linear quadratic problem, called an "accessory problem", has nonnegative cost. In contrast with [2], we highlight the relation between the nonnegativeness of the linear quadratic problem and the optimality of the solution of (P). This is done by building a family of admissible processes of (P) depending on some parameter ϵ.

In [4] and [9] similar results were established when the constraint functions b and g depend only on the control variable u. We generalise their work to cover constraints depending jointly on x and u. Our approach differ from theirs since no admissible directions set is defined. Notably, we apply an Uniform Implicit Function Theorem previously obtained (see [1]) to build a family of admissible processes for (P).

2 Preliminaries

The notation $r \geq 0$ means that each component r_i of $r \in \mathbb{R}^r$ is nonnegative. $\langle \cdot,\cdot \rangle$ denotes the Euclidean scalar product on finite dimensional vector space $\mathbb{R}^k$, $|\cdot| = \sqrt{\langle \cdot,\cdot \rangle}$ the Euclidean norm, and $|\cdot|$ the induced matrix norm on $\mathbb{R}^{m \times k}$. To simplify notation $\bar{\phi}(t)$ will denote the evaluation of a function ϕ at $(t,\bar{x}(t),\bar{u}(t),\bar{v}(t))$ (or $(t,\bar{x}(t),\bar{w}(t))$, whereas ϕ may be L, f, b, g or its derivatives.

The following result, proved in [1], will be of importance in the forthcoming developments, mainly in the proof of our main result.

Theorem 2.1 (Uniform Implicit FunctionTheorem)
Consider a set $T \subset \mathbb{R}^k$, a number $\bar{\varepsilon} > 0$, a family of functions

$$\{\psi_a : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n\}_{a \in T}$$

and a point $(u_0, v_0) \in \mathbb{R}^m \times \mathbb{R}^n$ such that

$$\psi_a(u_0, v_0) = 0$$

for all $a \in T$. Assume that:

(i) *ψ_a is continuously differentiable on $(u_0, v_0) + \bar{\varepsilon} B$ for all $a \in T$.*

(ii) *There exists a monotone increasing function $\theta : (0, \infty) \to (0, \infty)$ with $\theta(s) \downarrow 0$ as $s \downarrow 0$ such that, for all $a \in T$, $(u', v'), (u, v) \in (u_0, v_0) + \bar{\varepsilon} B$,*

$$|\, \nabla \psi_a(u', v') - \nabla \psi_a(u, v) \,| \leq \theta(\,|\, (u', v') - (u, v) \,|\,).$$

(iii) *$\nabla_v \psi_a(u_0, v_0)$ is nonsingular for each $a \in T$ and there exists $c > 0$ such that, for all $a \in T$,*

$$|\, [\nabla_v \psi_a(u_0, v_0)]^{-1} \,| \leq c.$$

Then there exist $\delta \geq 0$ and a family of continuously differentiable functions

$$\{\phi_a : u_0 + \delta B \to v_0 + \bar{\varepsilon} B\}_{a \in T},$$

which are Lipschitz continuous with common Lipschitz constant k, such that, for all $a \in T$,

$$\begin{aligned} v_0 &= \phi_a(u_0) \\ \psi_a(u, \phi_a(u)) &= 0 \ \text{ for all } u \in u_0 + \delta B \\ \nabla_u \phi_a(u_0) &= -[\nabla_v \psi_a(u_0, v_0)]^{-1} \nabla_u \psi_a(u_0, v_0). \end{aligned}$$

The numbers δ and k depend on θ, c and $\bar{\varepsilon}$ only.

Furthermore, if T is a Borel set and $a \mapsto \psi_a(u, v)$ is a Borel measurable function for each

$$(u, v) \in (u_0, v_0) + \bar{\varepsilon} B,$$

then $a \mapsto \phi_a(u)$ is a Borel measurable function for each $u \in u_0 + \delta B$.

A *process* $(\bar{x}, \bar{u})$ of (P), i.e., a pair of an absolutely continuous function $\bar{x} : [a, b] \longrightarrow \mathbb{R}^n$ and measurable function $\bar{u} : [a, b] \longrightarrow \mathbb{R}^k$ satisfying the constraints of (P), is called a *weak local minimizer* if, and only if, there exists some $\bar{\varepsilon} > 0$, such that it minimizes the cost over all processes (x, u) of (P) which satisfy

$$(x(t), u(t)) \in T_{\bar{\varepsilon}}(t), \qquad \text{for a.a. } t \in [a, b],$$

where $T_{\bar{\varepsilon}}(t) = \big(\bar{x}(t) + \bar{\varepsilon}\bar{B}\big) \times \big(\bar{u}(t) + \bar{\varepsilon}\bar{B}\big)$. Here $\bar{B}$ denotes the closed unit ball.

Since we have inequality constraints in (P) we introduce the set $\mathcal{I}_a(t)$ of indexes of the *active constraints* as

$$\mathcal{I}_a(t) = \big\{ i \in \{1, \ldots, m_g\} \ \big| \ g_i(t, \bar{x}(t), \bar{u}(t)) = 0 \big\}.$$

and we denote its cardinal by $q_a(t)$.

Let

$$g_u^{\mathcal{I}_a(t)}\big(t, \bar{x}(t), \bar{u}(t)\big) \in \mathbb{R}^{q_a(t) \times k}$$

(if $q_a(t) = 0$, then the latter holds vacuously) denote the matrix we obtain after removing from $\bar{g}_u(t)$ all the rows of index $i \notin \mathcal{I}_a(t)$.

We shall invoke the following hypotheses on (P), which refers to some process $(\bar{x}, \bar{u})$ and parameter $\bar{\varepsilon} > 0$:

H1. $L(\cdot, x, u)$, $f(\cdot, x, u)$, $b(\cdot, x, u)$, are measurable for each (x, u).

For almost every $t \in [a, b]$, $f(t, \cdot, \cdot)$, $L(t, \cdot, \cdot)$, $b(t, \cdot, \cdot)$ and $g(t, \cdot, \cdot)$ are twice continuously differentiable on $(\bar{x}(t), \bar{u}(t)) + \bar{\varepsilon} B$.

$f(t, \cdot, \cdot)$, $L(t, \cdot, \cdot)$, $b(t, \cdot, \cdot)$ and $g(t, \cdot, \cdot)$ and their derivatives are essentially bounded at $(t, \bar{x}(t), \bar{u}(t))$.

H2. There exists a monotone increasing function $\theta : (0, \infty) \to (0, \infty)$ with $\theta(s) \downarrow 0$ as $s \downarrow 0$ such that, for all $t \in [a, b]$, $(x, u), (x', u') \in (\bar{x}(t), \bar{u}(t)) + \bar{\varepsilon} B$,

$$\begin{aligned} |\, \nabla_{x,u} \varphi(t, x, u) - {} & \nabla_{x,u} \varphi(t, x', u') \,| \leq \\ & \theta(\,|\, (x', u') - (x, u) \,|\,) \end{aligned}$$

where

$$\varphi(t, x, u) = (L(t, x, u), f(t, x, u), b(t, x, u), g(t, x, u))^T.$$

H3. There exists $K > 0$ such that, for almost all $t \in [a, b]$,

$$\det \ \big\{ \Upsilon(t) \Upsilon(t)^T \big\} \geq K$$

where

$$\Upsilon(t) = \begin{bmatrix} b_u(t, \bar{x}(t), \bar{u}(t)) \\ g_u^{I_a(t)}(t, \bar{x}(t), \bar{u}(t)) \end{bmatrix}$$

H4. h an l are C^2 on $\{x : |\, x - \bar{x}(b) \,| < \bar{\varepsilon}\}$ and the matrix $h'(\bar{x}(b))$ has full rank.

Define the Hamiltonian:

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

Under these set of hypotheses a Weak Maximum Principle, proved in [3], asserts the existence of of an absolutely continuous function p, two L^∞-functions

q and r, a vector μ and a scalar $\lambda \geq 0$ such that, for almost all $t \in [a,b]$,

$$
\begin{array}{llcl}
\textbf{(i)} & 0 & \neq & \| p \|_{L_\infty} + \lambda, \\
\textbf{(ii)} & -\dot{p}(t) & = & H_x(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t)), \\
\textbf{(iii)} & 0 & = & H_u(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t)), \\
\textbf{(iv)} & 0 & = & r(t) g(t, \bar{x}(t), \bar{u}(t)) \text{ and } r(t) \leq 0, \\
\textbf{(v)} & -p(b) & = & h'(\bar{x}(b))^T \mu + \lambda \nabla l(\bar{x}(b)).
\end{array}
$$

An admissible process for (P), (x,u), is called an *extremal* if it satisfies H1-H4 and there exist multipliers, i.e., functions p, q and r, $\mu \in \mathbb{R}^r$, and a scalar $\lambda \geq 0$ such that conclusions (i) through (v) above hold.

For any extremal (x,u) we define the set of multipliers as $\Lambda = (p(\cdot), q(\cdot), r(\cdot), \lambda, \mu)$. An extremal is called a *normal* extremal if there exists one set of multipliers satisfying $\lambda = 1$. An extremal (x,u) of (P) is *strongly normal* if the only solution of the system:

$$
\left\{
\begin{array}{rcl}
-\dot{p}(t) & = & H_x(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t)) \\
0 & = & H_u(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t)) \\
0 & = & r(t) \cdot g(t, \bar{x}(t), \bar{u}(t)) \\
-p(b) & = & h'(\bar{x}(b))^T \mu
\end{array}
\right.
$$

is $p(\cdot) \equiv 0$. It can also be deduced from the Weak Maximum Principle in [3] that if $p(\cdot) \equiv 0$, then $r(t) = q(t) = 0$ for almost all $t \in [a,b]$.

Throughout the remainder of this paper we assume that $(\bar{x}, \bar{u})$ is a strong normal solution of (P) and that the set of multipliers Λ associated with $(\bar{x}, \bar{u})$ is a *singleton*. Denote the multipliers as p, q, r, $\lambda = 1$ and μ.

3 Main Result

We now associate with (P) the following linear quadratic problem (AC):

$$
\left\{
\begin{array}{l}
\text{Minimise} \\
\qquad J_2(v) := \frac{1}{2} \left\{ l_2(y(b)) + \displaystyle\int_a^b L_2(t, y(t), v(t)) dt \right\} \\
\text{subject to} \\
\begin{array}{rcll}
\dot{y}(t) & = & \bar{f}_x(t) y(t) + \bar{f}_u(t) v(t) & \text{a.e} \\
0 & = & \bar{b}_x(t) y(t) + \bar{b}_u(t) v(t) & \text{a.e} \\
0 & = & \bar{g}_x^{I_a(t)}(t) y(t) + \bar{g}_u^{I_a(t)}(t) v(t) & \text{a.e} \\
y(a) & = & 0 & \\
My(b) & = & 0 &
\end{array}
\end{array}
\right.
$$

where $M = h'(\bar{x}(b))$,

$$l_2(y(b)) = y^T(b) \Gamma y(b),$$

$$\Gamma = \left\{ l''(\bar{x}(b)) + \left[h''(\bar{x}(b))^T \mu \right] \right\}$$

and

$$L_2(t, y(t), v(t)) = -y(t)^T \bar{H}_{xx}(t) y(t) - 2y(t)^T \bar{H}_{xu}(t) v(t) - v(t)^T \bar{H}_{uu}(t) v(t)$$

For simplicity of exposition, we define (AC) with only equality constraints of the form

$$0 = \bar{g}_x^{I_a(t)}(t) y(t) + \bar{g}_u^{I_a(t)}(t) v(t),$$

although another, possibly more general, "accessory problem" can be considered, as proposed in [5], by separating $\bar{g}_x^{I_a(t)}(t) y(t) + \bar{g}_u^{I_a(t)}(t) v(t)$ in both inequality and equality constraints in the following way. Define two subsets of $I_a(t)$:

$$
\begin{array}{rcl}
I_0(t) & = & \{ i \in I_a(t) \colon\ r_i(t) = 0 \} \\
I_<(t) & = & \{ i \in I_a(t) \colon\ r_i(t) < 0 \}
\end{array}
$$

and consider

$$
\begin{array}{rcl}
0 & = & \bar{g}_x^{I_<(t)}(t) y(t) + \bar{g}_u^{I_<(t)}(t) v(t) \\
0 & \geq & \bar{g}_x^{I_0(t)}(t) y(t) + \bar{g}_u^{I_0(t)}(t) v(t)
\end{array} .
$$

Our approach can be generalized to cover this other problem and it will be the focus of some future work.

Theorem 3.1 *Assume that $(\bar{x}, \bar{u})$ is an optimal solution of (P) and that it is strongly normal. Then, for any admissible solution (y,v) of (AC), we have*

$$J_2(v) \geq 0.$$

4 Proof of the Main Result

Set $\delta \in L^\infty([a,b]; \Re^{m_g})$ to be

$$
\delta_i(t) = \left\{
\begin{array}{lll}
1 & \text{if} & i \in I_a(t) \\
0 & \text{if} & i \notin I_a(t)
\end{array}
\right. .
$$

Define two diagonal matrices $\Delta(t), \Delta'(t) \in M_{m_g \times m_g}$ as

$$
\begin{array}{rcl}
\Delta(t) & = & \text{diag } \{ \delta_1(t), \ldots, \delta_{m_g}(t) \}, \\
\Delta'(t) & = & I - \Delta(t).
\end{array}
$$

Consider now a $L^\infty([a,b]; \Re^{m_g})$-function ω and two systems, (S)

$$
\left\{
\begin{array}{rcl}
\dot{y}(t) & = & \bar{f}_x(t) y(t) + \bar{f}_u(t) v(t) \\
0 & = & \bar{b}_x(t) y(t) + \bar{b}_u(t) v(t) \\
0 & = & \bar{g}_x^{I_a(t)}(t) y(t) + \bar{g}_u^{I_a(t)}(t) v(t) \\
y(a) & = & 0 \\
My(b) & = & 0
\end{array}
\right.
$$

and (S^*):

$$\begin{cases} \dot{y}(t) &= \bar{f}_x(t)y(t) + \bar{f}_u(t)v(t) \\ 0 &= \bar{b}_x(t)y(t) + \bar{b}_u(t)v(t) \\ 0 &= \Delta(t)\bar{g}_x(t)y(t) + \Delta(t)\bar{g}_u(t)v(t) + \\ & \qquad \Delta'(t)\omega(t) \\ y(a) &= 0 \\ My(b) &= 0 \end{cases}$$

Observe that any solution of (S) is an admissible solution for problem (AC) and that $(y \equiv 0, v \equiv 0)$ is a solution of (S).

As it can easily be seen, $(\tilde{y}, \tilde{v})$ is a solution of (S) if and only if there exists an L^∞ function $\tilde{\omega}$ such that $(\tilde{y}, \tilde{v}, \tilde{\omega})$ is a solution of (S^*).

If $(\tilde{y}, \tilde{v}, \tilde{\omega})$ is a solution of (S^*), then $\tilde{\omega}_i(t) = 0$ for any $i \notin I_a(t)$.

Set $w = (v, \omega)$. Defining

$$A(t) = \bar{f}_x(t), \quad B(t) = [\bar{f}_u(t) \;\; 0],$$

$$C(t) = \begin{bmatrix} \bar{b}_x(t) \\ \Delta(t)\bar{g}_x(t) \end{bmatrix},$$

and

$$D(t) = \begin{bmatrix} \bar{b}_u(t) & 0 \\ \Delta(t)\bar{g}_u(t) & \Delta'(t) \end{bmatrix},$$

we can rewrite (S^*) as

$$\begin{cases} \dot{y}(t) &= A(t)y(t) + B(t)w(t) \quad \text{a.e.} \\ 0 &= C(t)y(t) + D(t)w(t) \quad \text{a.e.} \\ y(a) &= 0 \\ My(b) &= 0 \end{cases}.$$

Let (y, w) be a solution of (S^*). Since, by H3, $D(t)$ is of full rank, we set

$$\begin{aligned} D^{\#}(t) &= D(t)^T \left(D(t)D(t)^T\right)^{-1} \\ \bar{A}(t) &= A(t) - B(t)D^{\#}(t)C(t) \\ \bar{B}(t) &= B(t)\left(I - D^{\#}(t)D(t)\right) \\ \Pi_1(t) &= \left(I - D^{\#}(t)D(t)\right). \end{aligned}$$

Define the system $(\hat{S})$:

$$\begin{cases} \dot{y}(t) &= \bar{A}(t)y(t) + \bar{B}(t)\zeta(t) \\ y(a) &= 0 \\ My(b) &= 0 \end{cases}$$

The following lemma relates the solutions of (S^*) and $(\hat{S})$. The proof is an easy task that we omit here.

Lemma 4.1 *If (y, w), is a solution of (S^*), then there exists an $\zeta \in L^\infty$ such that*

$$\Pi_1(t)(\zeta(t) - w(t)) = 0$$

and (y, ζ) is a solution of $(\hat{S})$.

If (y, ζ) is a solution of $(\hat{S})$, then (y, w), where

$$w(t) = \Pi_1(t)\zeta(t) - D^{\#}(t)C(t)y(t),$$

is a solution of (S^*).

Let $\Phi(t, a)$ be the transition matrix of

$$\dot{y}(t) = \bar{A}(t)y(t). \tag{4.1}$$

The reachable set of $(\hat{S})$ is

$$\mathcal{R}_a(b) = \left\{ \int_a^t \Phi(b,a)\Phi^{-1}(s,a)B(s)\Pi_1(t)\zeta(s)ds : \zeta \in L^\infty \right\}$$

Definition 4.2 *Let M be a full rank $r \times n$ real matrix. The system $(\hat{S})$ is M-controllable if*

$$M\mathcal{R}_a(b) = \Re^r.$$

By Lemma 4.1, M-controllability of $(\hat{S})$ is equivalent to the M-controllability of $(\tilde{S})$ (when considering the controllability we ignore the end point constraint $My(b) = 0$).

Proposition 4.2 of [8] asserts that $(\bar{x}, \bar{u})$ is strongly normal on $[a, b]$ if and only if $(\hat{S})$ is M-controllable.

We now focus on the system

$$\begin{cases} \dot{\psi}(t) &= A(t)\psi(t) + B(t)\xi(t) + c(t) \\ 0 &= C(t)\psi(t) + D(t)\xi(t) + d(t) \\ M\psi(b) + e &= 0, \end{cases} \tag{4.2}$$

where c and d are two essentially bounded functions and $e \in \Re^n$.

Lemma 4.3 *For any $c, d \in L^\infty$ and any $e \in \Re^n$, the system (4.2) has a solution.*

Proof. The system (4.1) has an unique solution given by $\psi(t) = \Phi(t, a)\psi_0$. Let $\mathcal{L} \in \Re^r$ and set $\psi_0 = \Phi^{-1}(b, a)M^T\left(MM^T\right)^{-1}\mathcal{L}$. Then the solution $\bar{\psi}$ of (4.1) is such that

$$M\bar{\psi}(b) = \mathcal{L}.$$

Premultiplying the algebraic equation of (4.2) by $D^{\#}(t)$, we obtain

$$-D^{\#}(t)C(t)\psi(t) - D^{\#}(t)d(t) = D^{\#}(t)D(t)\xi(t).$$

Subtracting ξ to both sides of this equality and replacing ξ in the differential equation of (4.2) we deduce that y is a solution of

$$\dot{\psi}(t) = \bar{A}(t)\psi(t) + \bar{B}(t)\xi(t) + r(t) \tag{4.3}$$

where $r(t) = c(t) - B(t)D^{\#}(t)d(t)$.

Denote by $(\tilde{\psi}, \tilde{\xi})$ a solution of (4.3).

Set $\lambda = M\tilde{\psi}(b)$ and $\mathcal{L} = -\lambda - e$.

Let $\bar{\psi}$ be the solution of (4.1) $(M\bar{\psi}(b) = \mathcal{L})$.

Then, (ψ, ξ), where $\psi(t) = \tilde{\psi}(t) + \bar{\psi}(t)$ and $\xi(t) = -D^{\#}(t)C(t)\psi(t) - \Pi_1(t)\tilde{\xi}(t) - D^{\#}(t)d(t)$ is a solution of (4.3) and

$$M\psi(b) = M\tilde{\psi}(b) + M\bar{\psi}(b) = \lambda - \lambda - e = -e$$

completing the proof. ■

Let $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_r$ be a linearly independent set of $\Re^n$-vectors and let $e_1, e_2, \ldots, e_r$ be such that $Me_i = \mathcal{L}_i$ for $i = 1, \ldots, r$. Set

$$\varphi(t) = [\bar{b}_x(t), \bar{g}_x^{I_a(t)}(t)].$$

For each $i = 1, \ldots, r$ consider the system

$$(S_i) \begin{cases} \dot{y}(t) &= A(t)y(t) + \bar{f}_u(t)v(t) \\ 0 &= \varphi(t)y(t) + \Upsilon(t)v(t) \\ y(a) &= 0 \\ y(b) &= e_i \end{cases} \quad (4.4)$$

The hypotheses H1 - H4 , the strongly normal assumption of $(\bar{x}, \bar{u})$ on $[a, b]$ and Proposition 4.2 of ([8]) ensure that the above system is M-controllable on $[a, b]$. It follows that there exists a solution of (y_i, v_i) of (S_i), for each $i = 1, \ldots, r$.

We now focus on the proof of Theorem 3.1.

Proof of Theorem 3.1. Let $\bar{\varepsilon} > 0$ be the parameter in hypotheses H1-H4.

Let (y, w), where $w = (v, \omega)$, be a solution of (S^*). Recall that (y, v) is an admissible solution of (AC).

Let $z \in \Re^{m_g}$ be the vector $z = (1, 1, \ldots, 1)^T$.

Let (y_i, v_i) be a solution of (4.4) for each $i = 1, \ldots, r$.

Let $\tilde{g}(t, x, u) = (b(t, x, u), g(t, x, u))$.

Let (ψ, ξ) be a solution of (4.2) where

$$2c(t) = y^T(t)\bar{f}_{xx}(t)y(t) + 2y^T(t)\bar{f}_{xu}(t)v(t) + v^T(t)\bar{f}_{uu}(t)v(t)$$

$$2d(t) = y^T(t)\bar{\tilde{g}}_{xx}(t)y(t) + 2y^T(t)\bar{\tilde{g}}_{ux}(t)v(t) + v^T(t)\bar{\tilde{g}}_{uu}(t)v(t),$$

and $e = \frac{1}{2}y^T(b)h''(\bar{x}(b))y(b)$.

Let $\gamma = (\gamma_1, \gamma_2) \in \Re^{n+m}$, where $m = m_b + m_g$.

Let $\beta = (\beta_1, \ldots, \beta_r) \in \Re^r$ and $\epsilon \in \mathbb{R}$.

Define the matrices

$$L_1(t) = [L_{11}(t) \;\; 0] \quad \text{and} \quad L_2(t) = [0 \;\; D^T(t)]$$

where

$$L_{11}(t) = -\int_a^t A(s)L_{11}(s)ds + I.$$

Set $x_0(\epsilon, \gamma, \alpha) = \bar{x}(a)$,

$$x(t, \epsilon, \gamma, \beta) = \bar{x}(t) + \epsilon y(t) + \epsilon^2\psi(t) + L_1(t)\gamma + \sum_{i=1}^r \beta_i y_i(t) \quad (4.5)$$

and

$$u(t, \epsilon, \gamma) = \bar{u}(t) + \epsilon v(t) + \epsilon^2\xi(t) + L_2(t)\gamma + \sum_{i=1}^r \beta_i v_i(t). \quad (4.6)$$

Consider the function

$$F(t, \epsilon, \gamma, \beta) = (F_1(t, \epsilon, \gamma, \beta), F_2(t, \epsilon, \gamma, \beta), F_3(\epsilon, \gamma, \beta))$$

where

$$F_1(t, \epsilon, \gamma, \beta) = x(t, \epsilon, \gamma, \beta) - x_0(\epsilon, \gamma, \beta) - \int_a^t f(s, x(s, \epsilon, \gamma, \beta), u(s, \epsilon, \gamma, \beta))ds,$$

$$F_2(t, \epsilon, \gamma, \beta) = b(t, x(t, \epsilon, \gamma, \beta), u(t, \epsilon, \gamma, \beta)),$$

$$F_3(t, \epsilon, \gamma, \beta) = \Delta(t)g(t, x(t, \epsilon, \gamma, \beta), u(t, \epsilon, \gamma, \beta)) + \Delta'(t)\Big(\gamma_2 + \epsilon\omega(t) + \epsilon^4\rho(t, \epsilon, \gamma, \beta)\Big)$$

with

$$\rho(t, \epsilon, \gamma, \beta) = g(s, x(s, \epsilon, \gamma, \beta), u(s, \epsilon, \gamma, \beta)) - \bar{g}(t) + \bar{\varepsilon}z.$$

We have $F(t, 0, 0, 0) = 0$.

Differentiating F with respect to ϵ, we obtain

$$\frac{\partial F}{\partial \epsilon}(t, 0, 0, 0) = \begin{pmatrix} y(t) - \int_a^t \Big(A(s)y(s) + B(s)v(s)\Big)ds \\ C(t)y(t) + D(t)w(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Differentiating F with respect to γ, we get

$$\frac{\partial F}{\partial \gamma}(t, 0, 0, 0) = \begin{bmatrix} I & 0 \\ C(t)L_{11}(t) & D(t)D^T(t) + \begin{bmatrix} 0 \\ \Delta'(t) \end{bmatrix} \end{bmatrix}.$$

We deduce from H3 that $\frac{\partial F}{\partial \gamma}(t, 0, 0, 0)$ is nonsingular. By H1, it can be proved that $\frac{\partial F}{\partial \gamma}(t, 0, 0, 0)$ satisfies condition (iii) of Theorem 2.1. This, together with H2, allows the application of Theorem 2.1 which asserts the existence of a $\delta > 0$ and a function

$$\gamma : J \times (-\delta, \delta) \times (-\delta, \delta)^r \to \bar{\varepsilon}B,$$

where $J \subset [a, b]$ is a set of full measure and B is the unitary open ball in $\mathbb{R}^{n+m}$, such that $\gamma(t, \cdot, \cdot)$ is C^2, $\gamma(t, 0, 0) = 0$ and

$$F(t, \epsilon, \gamma(t, \epsilon, \alpha), \beta) = 0 \quad (4.7)$$

for almost all $t \in [a, b]$ and all $\epsilon \in (-\delta, \delta)$.

Differentiating (4.7) with respect to ϵ we conclude that

$$\frac{\partial \gamma}{\partial \epsilon}(t,0,0)=0.$$

For all $i=1,\dots,r$, we also have

$$\frac{\partial F}{\partial \beta_i}(t,0,0,0)=0,$$

asserting that $\frac{\partial \gamma}{\partial \beta_i}(t,0,0)=0$.

By definition of ψ and ξ,

$$\frac{\partial^2 F}{\partial \epsilon^2}(t,0,0,0)=0$$

and we deduce that $\frac{\partial^2 \gamma}{\partial \epsilon^2}(t,0,0)=0$.

Replacing γ by $\gamma(t,\epsilon,\beta)$ in (4.5) and (4.6), we have two functions

$$\begin{array}{rcl}\hat{u}(t,\epsilon,\beta) & = & u(t,\epsilon,\gamma(t,\epsilon,\beta),\beta)) \\ \hat{x}(t,\epsilon,\beta) & = & x(t,\epsilon,\gamma(t,\epsilon,\beta),\beta))\end{array}$$

defined on $J\times(-\delta,\delta)\times(-\delta,\delta)^r$, where $J\subset[a,b]$ is a set of full measure. These functions have the regularity properties of $\gamma(t,\epsilon,\beta)$. It follows from the above that $\hat{u}(t,0,0)=\bar{u}(t)$, $\hat{x}(t,0,0)=\bar{x}(t)$ and

$$\left\{\begin{array}{rclrcl}\frac{\partial \hat{u}}{\partial \epsilon}(t,0,0) & = & v(t), & \frac{\partial \hat{x}}{\partial \epsilon}(t,0,0) & = & y(t), \\ \frac{\partial \hat{u}}{\partial \beta_i}(t,0,0) & = & v_i(t), & \frac{\partial \hat{x}}{\partial \beta_i}(t,0,0) & = & y_i(t), \\ \frac{\partial^2 \hat{u}}{\partial \epsilon^2}(t,0,0) & = & 2\xi(t), & \frac{\partial^2 \hat{x}}{\partial \epsilon^2}(t,0,0) & = & 2\psi(t),\end{array}\right.$$

for almost all $t\in[a,b]$.

Define now the map $G:(-\delta,\delta)\times(-\delta,\delta)^r\to\Re^r$ as

$$G(t,\epsilon,\beta)=h(\hat{x}(b,\epsilon,\beta)).$$

Then $G(0,0)=0$ and

$$\frac{\partial G}{\partial \beta_j}(0,0)=Me_i=\mathcal{L}_i.$$

Since $\{\mathcal{L}_1,\dots,\mathcal{L}_r\}$ is a linearly independent set of vectors, the matrix $\frac{\partial G}{\partial \beta}(0,0)$ is nonsingular. The classic implicit function theorem asserts the existence of a $0<\delta_0\le\delta$ and a C^2 function β defined on $(-\delta_0,\delta_0)$, with $\beta(0)=0$ and such that

$$h(\hat{x}(b,\epsilon,\beta(\epsilon))=0 \tag{4.8}$$

for all $\epsilon\in(-\delta_0,\delta_0)$. Differentiating (4.8) twice we can further conclude that $\beta_i'(0)=0$ for $i=1,\dots,r$ and $\beta_i''(0)=0$, for $i=1,\dots,r$.

Let $\bar{\delta}=\min\{1,\delta_0\}$. We now define on $J\times(-\bar{\delta},\bar{\delta})$, where, again, $J\subset[a,b]$ is of full measure,

$$\begin{array}{rcl}\tilde{u}(t,\epsilon) & = & \hat{u}(t,\epsilon,\beta(\epsilon)) \\ \tilde{x}(t,\epsilon) & = & \hat{x}(t,\epsilon,\beta(\epsilon))\end{array}$$

We conclude from the above that

$$F(t,\epsilon,\gamma(t,\epsilon,\beta(\epsilon)),\beta(\epsilon))=0 \tag{4.9}$$

for almost all $t\in[a,b]$ and all $\epsilon\in(-\bar{\delta},\bar{\delta})$.

Observe that $\tilde{\gamma}(t,\epsilon)=\gamma(t,\epsilon,\beta(\epsilon))$ is a function defined on $J\times(-\bar{\delta},\bar{\delta})$ and taking values in $\bar{\varepsilon}B$.

It follows from (4.9) and the definition of F that, for almost all $t\in[a,b]$, all $\epsilon\in(-\bar{\delta},\bar{\delta})$

$$\begin{array}{rcll}\dot{\tilde{x}}(t,\epsilon) & = & f(t,\tilde{x}(t,\epsilon),\tilde{u}(t,\epsilon)) & \\ b(t,\tilde{x}(t,\epsilon),\tilde{u}(t,\epsilon)) & = & 0 & \\ g_i(t,\tilde{x}(t,\epsilon),\tilde{u}(t,\epsilon)) & \le & 0 & \forall\, i\in I_a(t) \\ \tilde{x}(a,\epsilon) & = & \bar{x}(a) & \\ h(\tilde{x}(b,\epsilon)) & = & 0 & \end{array}$$

and, for all $i\notin I_a(t)$,

$$\begin{array}{rl}\epsilon^4\big(g_i(t,\tilde{x}(t,\epsilon),\tilde{u}(t,\epsilon)) & -g_i(t,\bar{x}(t),\bar{u}(t))\big)= \\ & -\tilde{\gamma}_2^i(t,\epsilon)-\epsilon^4\bar{\varepsilon}\le \\ & \bar{\varepsilon}(1-\epsilon^4)\le 0\end{array}$$

This means that, for all $\epsilon\in(-\bar{\delta},\bar{\delta})$, $(\tilde{x}(\cdot,\epsilon),\tilde{u}(\cdot,\epsilon))$ is an admissible solution of (P).

Define the function

$$J(\epsilon)=l(\tilde{x}(b,\epsilon))+\int_a^b L(t,\tilde{x}(t,\epsilon),\tilde{u}(t,\epsilon))dt$$

on $(-\bar{\delta},\bar{\delta})$. The minimum of J is attained at $\epsilon=0$. Since

$$\begin{array}{rcl}J'(0) & = & \nabla(\bar{x}(b))^T y(b)+ \\ & & \int_a^b(\bar{L}_x(t)y(t)+\bar{L}_u(t)v(t))dt \\ & = & -p^T(b)y(b)+\int_a^b \frac{d}{dt}(p^T(t)y(t))dt \\ & = & -p^T(b)y(b)+p^T(b)y(b)-p^T(a)y(a) \\ & = & 0\end{array}$$

we must have:

$$\begin{array}{rcl}J''(0) & = & y^T(b)l''(\bar{x}(b))y(b)+ \\ & & \int_a^b(y^T(t)\bar{L}_{xx}(t)y(t)+ \\ & & 2y^T(t)\bar{L}_{xu}(t)v(t)+v^T(t)\bar{L}_{uu}(t)v(t))dt \\ & \ge & 0\end{array}$$

Suppressing the t for simplicity, observe that

$$\begin{array}{rcl}\bar{L}_{xx} & = & -\bar{H}_{xx}+p\bar{f}_{xx}+q\bar{b}_{xx}+r\bar{g}_{xx} \\ \bar{L}_{xu} & = & -\bar{H}_{xu}+p\bar{f}_{xu}+q\bar{b}_{xu}+r\bar{g}_{xu} \\ \bar{L}_{xx} & = & -\bar{H}_{uu}+p\bar{f}_{uu}+q\bar{b}_{uu}+r\bar{g}_{uu}.\end{array}$$

We deduce that

$$
\begin{aligned}
J''(0) &= y^T(b)l''(\bar{x}(b))y(b)- \\
& \int_a^b \left(y^T(t)\bar{H}_{xx}(t)y(t) + 2y^T(t)\bar{H}_{xu}(t)v(t)+\right. \\
& \left. v^T(t)\bar{H}_{uu}(t)v(t)\right) dt \\
&= 2J_2(v) \geq 0
\end{aligned}
$$

proving the theorem. ■

Acknowledgement: M. d. R. de Pinho would like to thank the support of Fundação para a Ciência e Tecnologia, Proj. PRAXIS/P/EEI/14156/1998, Portugal.

References

[1] M.d.R. de Pinho and R.B. Vinter, Necessary conditions for optimal control problems involving nonlinear differential algebraic conditions, *J. Math. Anal. and Applications*, 212, 1997, 493-516.

[2] M.d.R. de Pinho and H. Zheng, Conjugate points for optimal control problems with state-dependent control constraints, *Systems Modelling and Optimization* (London: Chapman & Hall/LCR, 1999, Ed. M. P. Polis et all), 162-170.

[3] M.d.R. de Pinho and A. Ilchmann, Weak maximum principle for optimal control problems with mixed constraints, submitted to *Nonlinear Analysis, Theo. Applic.*, 1999.

[4] P. Loewen and H. Zheng, Generalized conjugate points for optimal control problems, *Nonlinear Analysis, Theo. Applic.*, 22(6), 1994, 771-791.

[5] N. Osmolovskii, Second Order conditions for a weak local minimum in an optimal control problem (necessity, sufficiency)., *Dokl. Akad. Nauk SSSR*, 333, 1975, 259-262 (English transl. in *Soviet Math. Dokl, 16, ,1975, 1480-1484.*

[6] G. Stefani G. and P.L. Zezza, Optimality conditions for constrained control problems, *SIAM J. Control and Optim.*, 34(2), 1996, 635-659.

[7] G. Stefani and P.L. Zezza, Constrained regular LQ-control problems, *SIAM J. Control and Optim.*, 35(3), 1997, 879-900.

[8] V. Zeidan, The Riccati equation for optimal control problems with mixed state-control constraints: Necessity and Sufficiency, *SIAM J. Control and Optimization*, 32, 1994, 1297-1321.

[9] V. Zeidan, Admissible directions and generalized coupled points for optimal control problems, *Nonlinear Analysis, Theo. Applic.*, 26(3), 1996, 479-507.