

# ON OPTIMALITY CONDITIONS FOR OPTIMAL CONTROL PROBLEMS WITH MIXED CONSTRAINTS

Maria do Rosário de Pinho\*, Achim Ilchmann†

\* Departamento de Engenharia Electrotecnica e de Comp., Inst. Sistemas e Rob., Faculdade de Eng. Universidade do Porto, Rua Dr. Roberto Frias 4200 465, Porto, Portugal,  
fax: ++351 22 508 1443  
e-mail: mrpinho@fe.up.pt

† Institute of Mathematics, Technical University Ilmenau, Weimarer Straße 25, 98693 Ilmenau, FRG  
e-mail: ilchmann@mathematik.tu-ilmenau.de

**Keywords:** Optimal control, mixed constraints, maximum principle, nonsmooth analysis, weak minimizers.

## Abstract

An optimal control problem for time-varying, nonlinear differential equations with state-dependent control constraints is considered, the data may be nonsmooth. A weak maximum principle for problems with equality mixed constraints and pointwise set constraints imposed only on some components of the control variable is presented. This result is then applied to derive another weak maximum principle for problems with both equality and inequality mixed constraints. The essential assumption is a full rankness condition imposed on the constraints. Additionally, a lemma shows how the results are related to previous contributions.

## 1 Introduction

Optimality conditions for control problems with mixed state-control constraints have been the focus of attention for a long time. In particular, the subject of necessary conditions in the form of maximum principles have been addressed by a number of authors; see [7], [11], [13], [15], to name but a few. Weak maximum principles, which apply to *weak local solutions*, covering problems with possibly nonsmooth data, have been considered in [3] and, in a more general setting, in [13].

Various recent results, including those of the present paper, can be captured as special cases of the following optimal control problem (**P**) with *mixed constraints*, also known as state-dependent control constraints:

$$\left\{ \begin{array}{l} \text{Minimize } l(x(a), x(b)) \\ \text{subject to} \\ \dot{x}(t) = f(t, x(t), u(t), v(t)) \quad a.a. \quad t \in [a, b] \\ 0 = b(t, x(t), u(t), v(t)) \quad a.a. \quad t \in [a, b] \\ 0 \geq g(t, x(t), u(t), v(t)) \quad a.a. \quad t \in [a, b] \\ (u(t), v(t)) \in U(t) \times V(t) \quad a.a. \quad t \in [a, b] \\ (x(a), x(b)) \in C \end{array} \right.$$

where  $l : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \rightarrow \mathbb{R}^n$ ,  $b : [a, b] \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \rightarrow \mathbb{R}^{m_b}$ ,  $g : [a, b] \times \mathbb{R}^n \times \mathbb{R}^{k_u} \times \mathbb{R}^{k_v} \rightarrow \mathbb{R}^{m_g}$ ,  $U : [a, b] \rightarrow \mathbb{R}^{k_u}$  and  $V : [a, b] \rightarrow \mathbb{R}^{k_v}$  are given multifunctions and  $C \subset \mathbb{R}^n \times \mathbb{R}^n$  a given set. Throughout this paper we assume that  $k_u + k_v \geq m_b + m_g$ .

We seek optimality necessary conditions in the form of a weak maximum principle which apply to various special cases of problem (**P**). We are particularly interested in generalizing classical results (see [7] and [11]) to cover problems with possibly nonsmooth data.

Usually one has  $m_b \geq 1$ ,  $m_g \geq 1$  and, for all  $t \in [a, b]$ ,  $U(t) \subset \mathbb{R}^{k_u}$  or  $V(t) \subset \mathbb{R}^{k_v}$ . However, we allow for  $m_b, m_g = 0$ ,  $U(t) = \mathbb{R}^{k_u}$ , or  $V(t) = \mathbb{R}^{k_v}$  to signify the case where there are no explicit equality or inequality state-control constraints or no pointwise set constraints on some components of the control variable.

Weak local solutions are defined as follows.

**Definition 1.1** A process  $(\bar{x}, \bar{u}, \bar{v})$  of (**P**), i.e., a triple of an absolutely continuous function  $\bar{x} : [a, b] \rightarrow \mathbb{R}^n$  and Lebesgue measurable functions  $\bar{u} : [a, b] \rightarrow \mathbb{R}^{k_u}$ ,  $\bar{v} : [a, b] \rightarrow \mathbb{R}^{k_v}$  satisfying the constraints of (**P**), is called a *weak local minimizer* if, and only if, there exists some  $\varepsilon > 0$ , such that it minimizes the cost over all processes  $(x, u, v)$  of (**P**) which satisfy

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

where

$$T_\varepsilon(t) = (\bar{x}(t) + \varepsilon \bar{B}) \times (U(t) \cap (\bar{u}(t) + \varepsilon \bar{B})) \times (V(t) \cap (\bar{v}(t) + \varepsilon \bar{B})), \quad (1.1)$$

and  $\bar{B}$  denotes the closed unit ball.  $\square$

Necessary conditions have previously been derived assuming that a certain matrix  $F(t)$  has full rank in the sense that  $\det F(t)F(t)^T \geq L$  for a.a.  $t \in [a, b]$ , for some  $L > 0$ .

Many such rankness conditions, together with conditions enforcing continuity of the data with respect to  $t$ , permit the application of classical Implicit Function theorems, thereby allowing the removal of the state-dependent control constraints.

An exception is to be found in a paper by Pales and Zeidan [13]. They prove a multiplier rule for an abstract nonsmooth

problem with mixed and pure constraints, and then derive necessary conditions. In the absence of pure state constraints, the weak maximum principle obtained in this way is validated with the full rankness imposed on

$$\Upsilon(t) = \begin{bmatrix} b_u(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) & 0 \\ g_u(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) & G(t) \end{bmatrix}, \quad (1.2)$$

where

$$G(t) = \text{diag} \{-g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))\}_{i \in \{1, \dots, m_g\}}.$$

In the present paper we present, first, a weak maximum principle for optimal control problems with equality mixed constraints ( $m_g = 0$ ) and pointwise set constraints imposed only on *some components* of the control variable ( $U(t) = \mathbb{R}^{k_u}$ ,  $k_v \geq 1$ ), derived in [4]. The novelty of such weak maximum principle is that it extends a result in [3] to treat problems for which some pointwise set constraints on the control are also present. Different to previous work, it assumes only measurability of the data with respect to  $t$ . Thus, a sharpened variant of the Implicit Function Theorem, a *Uniform* Implicit Function Theorem previously obtained in [6], must be used. Secondly, we provide a weak maximum principle for the particular case that  $U(t) = \mathbb{R}^{k_u}$  in  $(P)$  under the full rankness condition of the matrix (1.2) as in [13]. By contrast, we consider pointwise set constraints in some components of the control variable  $w = (u, v)$  and relax some of the smoothness assumptions on the dynamics.

Finally, a technical lemma is given and used to compare the different full rankness conditions of the literature. In the presence of both inequality and equality state-control constraints, a restriction on the generality of our results is the assumption that  $U(t) = \mathbb{R}^{k_u}$ . Nonetheless, these problems are of interest when higher order conditions are considered.

A notable feature of the proofs in this paper is that they provide a simple and transparent derivation of necessary conditions, which might also be worth knowing for classical second order conditions.

## 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  $\mathbb{R}^{m \times k}$ . We set

$$\mathbb{R}_{\geq 0}^m = \{x \in \mathbb{R}^m \mid x_i \geq 0 \text{ for } i = 1, \dots, m\}$$

and  $\mathbb{R}_{\leq 0}^m$  accordingly. The *Euclidean distance function* with respect to  $A \subset \mathbb{R}^k$  is

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

We will often refer to the control variable as being  $w$  whenever we do not want to distinguish between components. In that case, a weak local minimizer will be denote by  $(\bar{x}, \bar{w})$ , the

control set will be  $W(t) \subset \mathbb{R}^k$ ,  $k = k_u + k_v$ , and the set (1.1) will be written as

$$T_\varepsilon(t) = (\bar{x}(t) + \varepsilon \bar{B}) \times (W(t) \cap (\bar{w}(t) + \varepsilon \bar{B})).$$

The linear space  $W^{1,1}([a, b]; \mathbb{R}^p)$  denotes the space of absolutely continuous functions,  $L^1([a, b]; \mathbb{R}^p)$  the space of integrable functions and  $L^\infty([a, b]; \mathbb{R}^p)$  the space of essentially bounded functions from  $[a, b]$  to  $\mathbb{R}^p$ , respectively.

We will make use of concepts from nonsmooth analysis. Of foremost importance are those of *limiting normal cone* to a set  $A$  at  $x$ , written  $N_A(x)$  and the *limiting subdifferential* of a lower semi continuous function  $f$  at  $x$ , written  $\partial f(x)$ . Such concepts were first introduced in [9]. The full calculus for these constructions in finite dimensions are described in [10] and [14].

In the case that the function  $f$  is Lipschitz continuous near  $x$ , the convex hull of the limiting subdifferential,  $\text{co } \partial f(x)$ , coincides with the (*Clarke*) *generalized gradient*, which may be defined directly. Properties of generalized gradients (upper semi-continuity, sum rules, etc.) are described in [1].

## 3 Main Results

To simplify notation,  $\bar{\phi}(t)$  will denote the evaluation of a function  $\phi$  at  $(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ , where  $\phi$  may be  $f$ ,  $b$ ,  $g$  or its derivatives.

Let  $\mathcal{I}_a(t)$  be the set of indexes of the *active constraints*, i.e.,

$$\mathcal{I}_a(t) = \{i \in \{1, \dots, m_g\} \mid g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0\}$$

and its complement, the set of indexes of the *inactive constraints*,

$$\mathcal{I}_c(t) = \{1, \dots, m_g\} \setminus \mathcal{I}_a(t).$$

$q_a(t)$  denotes the cardinal of  $\mathcal{I}_a(t)$  and  $q_c(t)$  be the cardinal of  $\mathcal{I}_c(t)$ . Let

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

(if  $q_a(t) = 0$ , then the latter holds vacuously) denote the matrix we obtain after removing from  $g_u(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$  all the rows of index  $i \in \mathcal{I}_c(t)$ .

We shall invoke the following hypotheses on  $(P)$  which make reference to a process  $(\bar{x}, \bar{u}, \bar{v})$  of  $(P)$  and some scalar  $\varepsilon > 0$ :

- (H1)  $f(\cdot, x, u, v)$  is measurable for each  $(x, u, v)$  and  $f(t, \cdot, \cdot, \cdot)$  is Lipschitz continuous with Lipschitz constant  $k_f(t)$  on  $T_\varepsilon(t)$  for almost all  $t \in [a, b]$ , and  $k_f$  is an  $L^1$ -function.
- (H2) The cost  $l$  is Lipschitz continuous on a neighborhood of  $(\bar{x}(a), \bar{x}(b))$  and  $C$  is closed.
- (H3) *graph*  $U(\cdot)$  is Borel measurable and  $U(t) \cap (\bar{u}(t) + \varepsilon \bar{B})$  is closed for almost all  $t \in [a, b]$ .

**(H4)**  $b(\cdot, x, u, v)$  and  $g(\cdot, x, u, v)$  are measurable for each  $(x, u, v)$ . For almost all  $t \in [a, b]$ ,  $b(t, \cdot, \cdot, \cdot)$  and  $g(t, \cdot, \cdot, \cdot)$  are continuously differentiable functions on  $(\bar{x}(t), \bar{u}(t), \bar{v}(t)) + \varepsilon \bar{B}$ .

There exists an increasing function  $\tilde{\theta} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ ,  $\tilde{\theta}(s) \downarrow 0$  as  $s \downarrow 0$ , such that, for all  $(x', u', v'), (x, u, v) \in (\bar{x}(t), \bar{u}(t), \bar{v}(t)) + \varepsilon \bar{B}$  and for almost all  $t \in [a, b]$ ,

$$\begin{aligned} & |\nabla_{x,u,v}[b, g](t, x', u', v') - \nabla_{x,u,v}[b, g](t, x, u, v)| \\ & \leq \tilde{\theta} \left( \left| (x', u', v') - (x, u, v) \right| \right). \end{aligned}$$

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

$$\begin{aligned} & |\nabla_x[b, g](t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| \\ & + |\nabla_u[b, g](t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| \\ & + |\nabla_v[b, g](t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| \leq K_{b,g}. \end{aligned}$$

**(H5)** There exists  $K > 0$  such that, for almost all  $t \in [a, b]$ ,

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

where  $\Upsilon(t)$  is defined in (1.2).

**Theorem 3.1** (Weak Maximum Principle for (P) without inequality constraints)

Let  $(\bar{x}, \bar{u}, \bar{v})$  be a weak local minimizer for problem  $(\mathbf{P}_=)$

$$\left\{ \begin{array}{l} \text{Minimize } l(x(a), x(b)) \\ \text{subject to} \\ \quad \dot{x}(t) = f(t, x(t), u(t), v(t)) \quad \text{a.a. } t \in [a, b] \\ \quad 0 = b(t, x(t), u(t), v(t)) \quad \text{a.a. } t \in [a, b] \\ \quad (u(t), v(t)) \in \mathbb{R}^{k_u} \times V(t) \quad \text{a.a. } t \in [a, b] \\ \quad (x(a), x(b)) \in C. \end{array} \right.$$

Assume that, for some  $\varepsilon > 0$ , hypotheses (H1)-(H5) are satisfied. Note that the matrix  $\Upsilon(t)$  in (H5) simplifies to  $\Upsilon(t) = b_u(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$  and the Hamiltonian is

$$H(t, x, p, q, u, v) = \langle p, f(t, x, u, v) \rangle + \langle q, b(t, x, u, v) \rangle.$$

Then there exist  $p \in W^{1,1}([a, b]; \mathbb{R}^n)$ ,  $\zeta \in L^1([a, b]; \mathbb{R}^{k_v})$  and  $\lambda \geq 0$  such that, for almost all  $t \in [a, b]$ ,

- (i)  $\|p\|_{L^\infty} + \lambda \neq 0$
- (ii)  $(-\dot{p}(t), 0, \zeta(t)) \in \text{co } \partial_{x,u,v} H(t, \bar{x}(t), p(t), q(t), \bar{u}(t), \bar{v}(t))$ ,
- (iii)  $\zeta(t) \in \text{co } N_{V(t)}(\bar{v}(t))$ ,
- (iv)  $(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial l(\bar{x}(a), \bar{x}(b))$ .

Furthermore, there exists an  $M > 0$  such that

$$\left| (\Upsilon(t) \Upsilon(t)^T)^{-1} \right| \leq M \quad \text{a.a. } t \in [a, b], \quad (3.3)$$

and

$$|q(t)| \leq k_f(t) M K_{b,g} |p(t)| \quad \text{a.a. } t \in [a, b]. \quad (3.4)$$

□

We omit the proof of Theorem 3.1. Its main step consists in rewriting the constraints so that a Uniform Implicit Function theorem, proved in [6], applies. Observe that a sharpened version of an Implicit Function theorem is needed, since we work under the assumption that the data is merely measurable with respect to  $t$ . Then we associate with  $(\mathbf{P}_=)$  an ‘‘auxiliary problem’’. This problem is a standard optimal control problem, but we must apply the nonsmooth Maximum Principle Proposition, derived in [5], to yield the required conditions for (P). Details can be found in [4].

We now turn to optimal control problems with mixed state-control constraints in the form of equalities and inequalities and we assume that  $U(t) = \mathbb{R}^{k_u}$ .

**Theorem 3.2** (Weak Maximum Principle for (P) with  $U(t) = \mathbb{R}^{k_u}$ )

Let  $(\bar{x}, \bar{u}, \bar{v})$  be a weak local minimizer for (P) with  $U(t) = \mathbb{R}^{k_u}$ . Assume that, for some  $\varepsilon > 0$ , hypotheses (H1)-(H5) are satisfied and that  $g(\cdot, \bar{x}(\cdot), \bar{u}(\cdot), \bar{v}(\cdot))$  is bounded on  $[a, b]$ . Define the Hamiltonian to be

$$\begin{aligned} H(t, x, p, q, r, u, v) = \\ \langle p, f(t, x, u, v) \rangle + \langle q, b(t, x, u, v) \rangle + \langle r, g(t, x, u, v) \rangle. \end{aligned}$$

Then there exist  $p \in W^{1,1}([a, b]; \mathbb{R}^n)$ ,  $\eta(t) \in L^1([a, b]; \mathbb{R}^{k_v})$  and  $\lambda \geq 0$  such that, for almost all  $t \in [a, b]$ ,

- (i)  $\|p\|_{L^\infty} + \lambda \neq 0$ ,
- (ii)  $(-\dot{p}(t), 0, \eta(t)) \in \text{co } \partial_{x,u,v} H(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t), \bar{v}(t))$ ,
- (iii)  $\eta(t) \in \text{co } N_{V(t)}(\bar{v}(t))$ ,
- (iv)  $\langle r(t), g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) \rangle = 0$  and  $r(t) \leq 0$ ,
- (v)  $(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial l(\bar{x}(a), \bar{x}(b))$ .

Furthermore, there exists an  $M > 0$  such that

$$|(q(t), r(t))| \leq k_f(t) M K_{b,g} |p(t)| \quad \text{a.a. } t \in [a, b]. \quad (3.5)$$

□

The proof of Theorem 3.2 is provided in Section 4.

The following lemma and remark will be used in clarifying the meaning of (H5) and other assumptions made in the literature.

**Lemma 3.3** Let  $k, n \in \mathbb{N}$  such that, for all  $t \in [a, b]$ ,  $m(t), q(t), l(t) \in \mathbb{N}_0$ ,  $m(t) + q(t) = n \leq k$ , and consider

$$J(t) := \begin{bmatrix} A(t) & 0 & 0 \\ N(t) & 0 & B(t) \end{bmatrix} \in \mathbb{R}^{n \times (k+l(t)+q(t))},$$

with

$$A(t) \in \mathbb{R}^{m(t) \times k}, \quad N(t) \in \mathbb{R}^{q(t) \times k}, \quad B(t) \in \mathbb{R}^{q(t) \times q(t)}.$$

Then the following conditions

- (i)  $\exists c_A > 0 : \det A(t)A(t)^T \geq c_A \quad a.a. t \in [a, b]$
- (ii)  $\exists c_B > 0 : \det B(t)B(t)^T \geq c_B \quad a.a. t \in [a, b]$
- (iii)  $\exists c > 0 : \det J(t)J(t)^T \geq c \quad a.a. t \in [a, b]$

are related as follows

- (i), (ii)  $\wedge N \in L^\infty \implies$  (iii),
- (iii)  $\implies$  (i),
- (iii)  $\stackrel{i.g.}{\not\implies}$  (ii). □

The proof of Lemma 3.3 is provided in Section 4.

**Remark 3.4** In what follows, we assume that the data satisfies (H4).

- (a) Using the abbreviations introduced at the beginning of Section 3, we set

$$A(t) = \begin{bmatrix} \bar{b}_u(t) \\ \bar{g}_u^{\mathcal{I}_a(t)}(t) \end{bmatrix} \in \mathbb{R}^{(m_g + q_a(t)) \times k_u},$$

$$N(t) = \bar{g}_u^{\mathcal{I}_c(t)}(t) \in \mathbb{R}^{q_c(t) \times k_u},$$

$$B(t) = \text{diag} \{-\bar{g}_i(t)\}_{i \in \mathcal{I}_c(t)} \in \mathbb{R}^{q_c(t) \times q_c(t)},$$

and  $m_g = q_a(t) + q_c(t)$ . In above, the functions are evaluated at  $(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$ .

If the components of  $g$  are permuted in such a way that the active constraints come first, then  $\Upsilon(t)$  as defined in (1.2) and  $J(t)$  as in Lemma 3.3 coincide and we have

$$\Upsilon(t) = J(t) = \begin{bmatrix} \bar{b}_u(t) & 0 & 0 \\ \bar{g}_u^{\mathcal{I}_a(t)}(t) & 0 & 0 \\ \bar{g}_u^{\mathcal{I}_c(t)}(t) & 0 & \text{diag} \{-\bar{g}_i(t)\}_{i \in \mathcal{I}_c(t)} \end{bmatrix}$$

where

$$\Upsilon(t) \in \mathbb{R}^{(m_b + m_g) \times (k_u + m_g)}.$$

Note that the permutation of the components of  $g$  depends on  $t$  but does not change a full rankness condition. Suppose there exists some  $L > 0$  such that

$$\det \{ \nabla_u [b, g](t) \nabla_u [b, g](t)^T \} \geq L \quad a.a. t \in [a, b],$$

where

$$[b, g](t) = (b(\bar{x}(t), \bar{u}(t), \bar{v}(t)), g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))).$$

Then

$$\nabla_u [b, g](t) = \begin{bmatrix} A(t) \\ N(t) \end{bmatrix}$$

and an application of Weyl's Theorem (see e.g. [8]) yields condition (iii) of Lemma 3.3, or equivalently (H5).

(b) A sufficient condition for (H5) to hold is given in (a) above. Conversely, if (H5), or equivalently condition (iii) of Lemma 3.3, is satisfied, then condition (i) of Lemma 3.3 holds true. However, there are systems for which (i) of Lemma 3.3 holds but (iii) does not. Since, by (H4),  $N \in L^\infty$ , candidates are systems, for which (iii) is not valid. In [4] an optimal control problem for which condition (i) of Lemma 3.3 holds but not (iii) is given. The interest of such problem resides on the simple observation that the data of the problem does not satisfy (H5) but, nevertheless, Theorem 3.2 provides a nontrivial set of multipliers associated with the optimal solution of it. This does not come as a surprise. In fact, since the components of the multiplier  $r$  are zero when  $i \in \mathcal{I}_c(t)$ , the derivative with respect to  $u$  of the corresponding components of  $g$  do not take any part in the determination of the multipliers.

## 4 Proofs

**Proof of Lemma 3.3:** Set  $\mathcal{M} \subset [a, b]$  to be the largest set such that (i)-(iii) hold for all  $t \in \mathcal{M}$ .

To simplify the notation, define for sequences  $\{t_i\}_{i \in \mathbb{N}} \subset \mathcal{M}$  and  $\{(x_i, y_i)\}_{i \in \mathbb{N}} \subset \mathbb{R}^{m(t_i) \times q(t_i)}$ ,

$$\varphi_i := |A(t_i)^T x_i + N(t_i) y_i|^2 + |B(t_i)^T y_i|^2.$$

**“(i), (ii) &  $N \in L^\infty \implies$  (iii)”:** Seeking a contradiction, suppose there exists a sequence  $\{t_i\}_{i \in \mathbb{N}} \subset \mathcal{M}$  and

$$\{(x_i, y_i)\}_{i \in \mathbb{N}} \subset \mathbb{R}^{m(t_i) \times q(t_i)} \quad (4.6)$$

with

$$|(x_i^T, y_i^T)| = 1 \quad \text{for all } i \in \mathbb{N}, \quad (4.7)$$

such that

$$\lim_{i \rightarrow \infty} \varphi_i = 0 \quad (4.8)$$

Suppose there are only finitely many  $q(t_i) \geq 0$ . If necessary, extract a subsequence (we do not relabel)  $\{t_i\}_{i \in \mathbb{N}}$  so that (4.6), (4.7) and (4.8) hold and  $q(t_i) = 0$  for all  $i \in \mathbb{N}$ . It follows that

$$\lim_{i \rightarrow \infty} \varphi_i = \lim_{i \rightarrow \infty} |A(t_i)^T x_i|^2 = 0 \quad \text{and} \quad |x_i| = 1 \quad \text{for all } i \in \mathbb{N}.$$

This violates (i) and yields a contradiction.

If there exist infinitely many  $i \in \mathbb{N}$  with  $q(t_i) \geq 1$ , then extract a subsequence (again we do not relabel)  $\{t_i\}_{i \in \mathbb{N}}$  such that  $q(t_i) \geq 1$  for all  $i \in \mathbb{N}$ . If  $\{|y_i|\}_{i \in \mathbb{N}}$  is bounded away from 0, then, by (ii) we have  $\lim_{i \rightarrow \infty} y_i = 0$ , contradicting the boundedness of  $\{|y_i|\}_{i \in \mathbb{N}}$  away from zero. If, however,  $\{|y_i|\}_{i \in \mathbb{N}}$  is *not* bounded away from 0, then we may extract a further subsequence  $\{t_i\}_{i \in \mathbb{N}}$  such that  $\lim_{i \rightarrow \infty} y_i = 0$ . In this case the sequence defined by (4.6)-(4.8) gives  $\lim_{i \rightarrow \infty} A(t_i)^T x_i + N(t_i) y_i = 0$ , and, since  $N$  is bounded along the subsequence, we deduce that  $\lim_{i \rightarrow \infty} A(t_i)^T x_i = 0$ , whence, by (i),  $\lim_{i \rightarrow \infty} x_i = 0$ . This contradicts  $|(x_i^T, y_i^T)| = 1$  for all  $i \in \mathbb{N}$ , and thus the first part of the proof is complete.

“(iii)  $\implies$  (i)”: Once again, seeking a contradiction, suppose that there exists a sequence  $\{t_i\}_{i \in \mathbb{N}} \subset \mathcal{M}$  such that

$$\lim_{i \rightarrow \infty} A(t_i)^T x_i = 0 \quad \text{and} \quad |x_i| = 1 \quad \text{for all } i \in \mathbb{N}.$$

Then, for

$$(x_i^T, y_i^T) = \begin{cases} (x_i^T, 0), & \text{if } q(t_i) \geq 1 \\ x_i^T, & \text{if } q(t_i) = 0 \end{cases}$$

it follows that

$$\lim_{i \rightarrow \infty} \varphi_i = 0 \quad \text{and} \quad |(x_i^T, y_i^T)| = 1.$$

This contradicts (iii) and completes the second part of the proof.

“(iii)  $\not\Rightarrow$  (ii)”: <sup>i.g.</sup> The constant matrices  $A = [0, 1]$ ,  $N = [1, 0]$  and  $B = 0$  satisfy

$$\begin{bmatrix} A & 0 & 0 \\ N & 0 & B \end{bmatrix} \begin{bmatrix} A & 0 & 0 \\ N & 0 & B \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and hence (iii) is certainly satisfied but not (ii).  $\square$

Now we are in a position to prove Theorem 3.2 by applying Theorem 3.1 to a suitable auxiliary problem.

**Proof of Theorem 3.2:** In order to apply Theorem 3.1, we define an auxiliary optimal control problem.

Let  $\alpha, \zeta : [a, b] \rightarrow \mathbb{R}^{m_g}$  be measurable functions. Define two matrices

$$E(t) = \text{diag} \{-g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))\}_{i \in \{1, \dots, m_g\}},$$

$$Z(t) = \text{diag} \{z_1(t), \dots, z_{m_g}(t)\},$$

where

$$z_i(t) = \begin{cases} 1 & \text{if } i \in \mathcal{I}_a(t) \\ 0 & \text{if } i \in \mathcal{I}_c(t). \end{cases}$$

For  $\varepsilon \in (0, 1)$  and a weak local minimizer  $(\bar{x}, \bar{u}, \bar{v})$  for (P) with  $U(t) = \mathbb{R}^{k_u}$ , we consider the optimal control problem ( $\mathbf{P}_{\text{aux}}$ )

$$\begin{cases} \text{Minimize } l(x(a), x(b)) \\ \text{subject to} \\ \dot{x}(t) = f(t, x, u, v) \\ 0 = b(t, x, u, v) \\ 0 = g(t, x, u, v) + E(t)\alpha(t) + Z(t)\zeta(t) \\ (u(t), v(t), \alpha(t), \zeta(t)) \in \mathbb{R}^{k_u} \times V(t) \times \mathbb{R}^{m_g} \times \mathbb{R}_{\geq 0}^{m_g} \\ (x(a), x(b)) \in C, \end{cases}$$

where all the equalities and inclusions but the last one hold for almost all  $t \in [a, b]$ . We proceed in several steps.

**Step 1:** We show that  $(\bar{x}, \bar{u}, \bar{v}, \bar{\alpha}, \bar{\zeta})$  is a weak local minimizer for ( $\mathbf{P}_{\text{aux}}$ ) where

$$\bar{\zeta}_i(t) = -g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))$$

and

$$\bar{\alpha}_i(t) = 1 \quad \text{for } i = 1, \dots, m_g.$$

Suppose that there exists an admissible solution  $(\tilde{x}, \tilde{u}, \tilde{v}, \tilde{\alpha}, \tilde{\zeta})$  for ( $\mathbf{P}_{\text{aux}}$ ) with lesser cost. For any  $i \in \mathcal{I}_a(t)$ , we have  $g_i(t, \tilde{x}(t), \tilde{u}(t), \tilde{v}(t)) + \tilde{\zeta}_i(t) = 0$ , and since  $\tilde{\zeta}_i(t) \geq 0$ , it follows that  $g_i(t, \tilde{x}(t), \tilde{u}(t), \tilde{v}(t)) \leq 0$ . Taking into account that any  $\tilde{\alpha}$  is such that  $\tilde{\alpha}(t) \in \bar{\alpha}(t) + \varepsilon \bar{B}$  and that  $\varepsilon < 1$ , we also deduce that  $g_i(t, \tilde{x}(t), \tilde{u}(t), \tilde{v}(t)) \leq 0$ , for all  $i \in \mathcal{I}_c(t)$ . This means that  $(\tilde{x}, \tilde{u}, \tilde{v})$  is a solution for (P) with lesser cost, contradicting the optimality of  $(\bar{x}, \bar{u}, \bar{v})$ .

**Step 2:** We prove that the data of ( $\mathbf{P}_{\text{aux}}$ ) satisfies the hypotheses under which Theorem 3.1 is applicable, i.e., (H1)-(H5) for

$$\begin{aligned} & \psi(t, x(t), u(t), v(t), \alpha(t), \zeta(t)) = \\ & \left( \begin{array}{c} b(t, x, u, v) \\ g(t, x, u, v) + E(t)\alpha(t) + Z(t)\zeta(t) \end{array} \right) \in \mathbb{R}^{m_b + m_g}, \end{aligned}$$

$(u, \alpha)$  and  $(v, \zeta)$  taking the role of  $b, u$  and  $v$ , respectively, are satisfied.

(H1)-(H3) are immediate. To see (H5) note that the derivative of  $\bar{\psi}$  with respect to  $u$  and  $\alpha$  is, in terms of (1.2),

$$\bar{\psi}_{u, \alpha}(t) = \psi_{u, \alpha}(t, \bar{x}(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t), \bar{\zeta}(t)) = \Upsilon(t).$$

It remains to show (H4). (H4) consists of two parts. We omit the details of the first part, see [4]. We only prove that the uniform bound in (H4). Note that, for almost all  $t \in [a, b]$ ,

$$\begin{aligned} & |\nabla_x \psi(t, \bar{x}(t), \bar{u}(t), \bar{\alpha}(t), \bar{v}(t), \bar{\zeta}(t))| \\ & + |\nabla_{u, \alpha} \psi(t, \bar{x}(t), \bar{u}(t), \bar{\alpha}(t), \bar{v}(t), \bar{\zeta}(t))| \\ & + |\nabla_{v, \zeta} \psi(t, \bar{x}(t), \bar{u}(t), \bar{\alpha}(t), \bar{v}(t), \bar{\zeta}(t))| \\ & \leq |\nabla_x [b, g](t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| + |\nabla_u [b, g](t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| \\ & + |E(t)| + |\nabla_v [b, g]g(t, \bar{x}(t), \bar{u}(t), \bar{v}(t))| + |Z(t)| \\ & \leq K_\psi, \end{aligned}$$

where the existence of  $K_\psi > 0$  is due to the fact that  $|Z(t)| = 1$ ,  $t \mapsto g(t, \bar{x}(t), \bar{u}(t))$  is uniformly bounded by assumption which yields uniform boundedness of  $E(t)$ . This completes the proof of (H4).

**Step 3:** Finally we apply Theorem 3.1 to prove (i)-(iv).

By Step 2, Theorem 3.1 is applicable to ( $\mathbf{P}_{\text{aux}}$ ) and therefore there exist  $p \in W^{1,1}([a, b]; \mathbb{R}^n)$ ,  $(\eta^1, \eta^2) \in L^1([a, b]; \mathbb{R}^{k_v} \times \mathbb{R}^{m_g})$  and  $\lambda \geq 0$  such that, for almost all  $t \in [a, b]$ ,

$$(-\dot{p}(t), 0, \eta^1(t), 0, \eta^2(t)) \in \quad (4.9)$$

$$\begin{aligned} & \text{co } \partial_{x, u, v, \alpha, \zeta} H(t, \bar{x}(t), p(t), q(t), r(t), \bar{u}(t), \bar{v}(t), \bar{\alpha}(t), \bar{\zeta}(t)) \\ & (\eta^1(t), \eta^2(t)) \in \text{co } N_{V(t)}(\bar{v}(t)) \times \text{co } N_{\mathcal{Z}}(\bar{\zeta}(t)) \quad (4.10) \end{aligned}$$

$$(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial l(\bar{x}(a), \bar{x}(b)), \quad (4.11)$$

where  $\mathcal{Z} = \mathbb{R}_{\geq 0}^{m_g}$  and

$$\begin{aligned} & H(t, x, p, q, r, u, v, \alpha, \zeta) \\ & = \langle p, f(t, x, u, v) \rangle + \langle q, b(t, x, u, v) \rangle \\ & + \left\langle r, g(t, x, u, v) + E(t)\alpha(t) + Z(t)\zeta(t) \right\rangle. \end{aligned}$$

Observe that

$$N_{\mathcal{Z}}(\bar{\zeta}(t)) = \mathbb{R}_{\leq 0}^{m_g}.$$

Since  $\bar{\zeta}_i(t) = -g_i(t, \bar{x}(t), \bar{u}(t)) = 0$  if  $i \in \mathcal{I}_a(t)$ , and  $\bar{\zeta}_i(t) > 0$  if  $i \in \mathcal{I}_c(t)$ , we deduce from (4.9) and (4.10) that

$$r_i(t) = 0, \quad \text{if } g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) < 0,$$

and

$$r_i(t) \leq 0, \quad \text{if } g_i(t, \bar{x}(t), \bar{u}(t), \bar{v}(t)) = 0.$$

Hence  $\lambda, \eta = \eta_1, p, q$  and  $r$  satisfy (ii)-(v). Note that  $\bar{\psi}_{u,\alpha}(t)$  and  $(q(t), r(t))$  take the role of  $\Upsilon(t)$  and  $q(t)$  in Theorem 3.1, respectively. Now by Step 2 and (H4) there exists an  $M > 0$  such that (3.3) holds. Applying (3.4) yields (3.5). This completes the proof.  $\square$

**Acknowledgments:** The support to Maria do Rosário de Pinho by Fundação para a Ciência e Tecnologia, PBICT/CEG/2438/95, Portugal is greatly acknowledged as well as numerous discussions with Prof. Vera Zeidan. This work was done while Achim Ilchmann was on study leave at the Faculdade de Eng. Universidade do Porto; the hospitality of the Instituto de Sistemas e Robótica, and the support by Fundação para a Ciência e Tecnologia, Praxis XXI, Portugal (BCC/20279/99), are gratefully acknowledged.

## References

- [1] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, (1983).
- [2] E.N. Devdaryani and Y.S. Ledyayev, Maximum principle for implicit control systems, *Appl. Math. Optim.*, **40**, 79–103 (1996).
- [3] M.d.R. de Pinho, On the weak maximum principle for optimal control for problems with state-dependent control constraints, *Nonlinear Analysis, Theo. Applic.*, **30**, Proc. 2nd World Congress of Nonlinear Analysts, 2481–2488 (1997).
- [4] M.d.R. de Pinho and A. Ilchmann, Weak Maximum Principle for optimal control problems with mixed constraints, to appear in *Nonlinear Analysis*, (2001).
- [5] M.d.R. de Pinho and R.B. Vinter, An Euler Lagrange inclusion for optimal control problems, *IEEE Trans. Aut. Control*, AC-**40**, 1191–1198 (1995).
- [6] M.d.R. de Pinho and R.B. Vinter, Necessary conditions for optimal control problems involving nonlinear differential algebraic equations, *J. Math. Anal. and Applications*, **212**, 493–516 (1997).
- [7] M.R. Hestenes, *Calculus of Variations and Optimal Control Theory*, Wiley, New York, (1966).
- [8] R.A. Horn and C.R. Johnson *Matrix Analysis*, Cambridge University Press, Cambridge (1985).
- [9] B.S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, *Journal of Appl. Math. Mech.*, **40**, 960–969 (1976).
- [10] B.S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow (1988); 2nd edition to be published by Wiley-Interscience.
- [11] L.W. Neustadt, *Optimization, A Theory of Necessary Conditions*, Princeton University Press, New Jersey, (1976).
- [12] N. P. Osmolovskii, Second order conditions for a weak local minimum in an optimal control problem (necessity and sufficiency), *Soviet Math. Dokl.*, **16**, 1480–1484, (1975).
- [13] Z. Pales and V. Zeidan, First- and second order necessary conditions for control problems with constraints, *Transc. American Mathematical Society*, **346**, 421–453 (1994).
- [14] R.T. Rockafellar and B. Wets, *Variational Analysis*, Springer Verlag, Berlin, (1998).
- [15] G. Stefani and P.L. Zezza, Optimality conditions for constrained control problems, *SIAM J. Control and Optim.*, **34**, 635–659 (1996).