Pure and Applied Functional Analysis
Volume 1, Number 2, 2016, 185–196

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL
© Copyright 2016

ON NECESSARY CONDITIONS FOR IMPLICIT CONTROL SYSTEMS

MARIA DO ROSARIO DE PINHO

Abstract. In this paper we focus on maximum principle for implicit systems. Central to all our discussion is Theorem 6.1 derived in [4]. This result, providing nonsmooth maximum principle for implicit systems has not deserved much attention in the literature, a situation we want to remedy here discussing its applicability and implications. Also, to enlarge its applicability, we also extend Theorem 6.1 in [4] to cover problems with set constrained implicit systems and where only measurability of the data is assumed with respect to the control variable. We do that applying again tools developed in [4].

To highlight the special features of these companion results we turn from nonsmooth to smooth and simple problems. We start by stating the special "smooth" counter part of nonsmooth results. Keeping our attention centred on smooth problems, we then explore the connections between maximum principles for optimal control problems involving semi-explicit Differential Algebraic Equations (DAE's) and our smooth results for specific problems where the implicit system comes in the form of equalities.

1. Introduction

Necessary conditions for optimal control problems with nonsmooth data have been derived for various problems in the last decades. However, not much has been done with respect to problems involving implicit systems. Exceptions can be found in [8] and [12], where the autonomous case is considered, and recently, in [4]. An important feature of such systems is the fact that some may be reduced to dynamic models taking the form of a coupled set of differential and algebraic equations (DAE's), systems widely used in engineering, specially in process systems engineering.

This paper centres on necessary conditions for optimal control problems involving implicit control systems and it is based on Theorem 6.1 in [4]. It turns out that this theorem breaks new ground; in contrast to [12] it covers nonautonomous and nonconvex problems. Moreover, it covers some special cases of problems with DAE's. Noteworthy, necessary conditions for optimal control problems with DAE's systems although the focus of attention (in particular, because they play an important role in the design of computation schemes; see, for example, [7,9,11,12,15,21]).

2010 *Mathematics Subject Classification.* 49K99, 34A09, 49K15.

Key words and phrases. Optimal control, maximum principle, implicit systems, differential and algebraic equations, nonsmooth analysis.

Theorem 6.1 in [4] has received none or little attention in the literature, a situation we aim to remedy here. Our first step, however, is to generalize the nonsmooth maximum principle in Theorem 6.1 to cover problems with set constrained implicit control systems of the form

$$f(t, x(t), \dot{x}(t), u(t)) \in \Phi, \quad u \in U, \tag{1.1}$$

and with less regularity with respect to the control variable. This is done in Section 2. Although we first consider nonsmooth problems, we recur to smooth problems to illustrate some special features of our result in Section 3. There, we also consider the smooth case when Φ in (1.1) reduces to $\{0\}$. The last result of Section 3, Corollary 3.2, plays a crucial role in the discussion of necessary conditions for problems with semi-explicit DAE's in section 4 of the form

$$E\dot{x}(t) - g(t, x(t), u(t)) = 0. \tag{1.2}$$

where E is a $N \times n$ constant matrix.

Our problem of interest is a fixed time optimal control problem involving implicit systems:

$$(P) \qquad \begin{cases} \text{Minimize} \quad l(x(a), x(b)) \\ \text{subject to} \\ \qquad\qquad f(t, x(t), \dot{x}(t), u(t)) \in \Phi \text{ a.e.}, \\ \qquad\qquad u(t) \in U \text{ a.e.}, \\ \qquad\qquad (x(a), x(b)) \in E, \end{cases}$$

where $l : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^N$ and $\Phi \subset \mathbb{R}^N$, $U \subset \mathbb{R}^k$ and $E \subset \mathbb{R}^n \times \mathbb{R}^n$ are all closed sets. Recall that the Mayer form adopted in (P) is not restrictive since problems with an integral cost can be easily reformulated in the above form by well known state augmentation techniques.

Crucial to our forthcoming analysis is the reformulation of the implicit system $f(t, x, \dot{x}, u) \in \Phi$ as

$$\begin{cases} \dot{x}(t) & = v, \\ f(t, x, v, u) & \in \Phi. \end{cases} \tag{1.3}$$

Clearly, and not surprisingly, this reformulation transforms an implicit control systems into a system with mixed state-control constraints.

The introduction of the variable v has implications with respect to the nature of x. Since x is assumed to be an absolutely continuous function, this scheme prevents us from treat some components of the state as measurable functions, a subject that will be discussed later on in section 4.

As we have mentioned before, we apply a smooth version of our results to problems with implicit control systems of the form (1.2) in our last section. We consider three cases: when E is of full row rank, when E is of full column rank and when E is not of full rank. In the first case, we show that the adjoint equation can be written in the form of the initial implicit system. In the two latter cases, however, (1.2) is rewritten as a DAE's in the semi-explicit form and application of necessary conditions is not possible unless lack of full rankness is somehow compensated as, for example when DAE's is of *index one*. The necessary conditions we obtain for DAE's of index one coincide with known results in the literature ([7,9,11,12,15,21]),

but they differ in so far as they are obtained without appealing to implicit function theorems.

To keep the exposition short and simple, we do not present the statement of the necessary condition when g in (1.2) is nonsmooth. Those can nevertheless be easily derived from our Theorem 2.1 or Theorem 2.2 or Lemma 3.1 below (choice would depend on the assumptions) yielding new nonsmooth necessary conditions for DAE's. Preliminaries of this sort can be found in [14].

Notation: If g is a vector, $g \in \mathbb{R}^m$, the inequality $g \leq 0$ is interpreted component wise.

We will denote by $\mathbb{B}$ the closed unit ball centred at the origin regardless of the dimension of the underlying space. Also $|\cdot|$ is the Euclidean norm or the induced matrix norm on $\mathbb{R}^{p\times q}$.

Take any $A \subset \mathbb{R}^n$. Then the *Euclidean distance function* with respect to A is defined as

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

Consider now a function $h : [a,b] \to \mathbb{R}^p$. We say that $h \in W^{1,1}([a,b];\mathbb{R}^p)$ if and only if it is absolutely continuous; in $h \in L^1([a,b];\mathbb{R}^p)$ iff h is integrable; and in $h \in L^\infty([a,b];\mathbb{R}^p)$ iff it is essentially bounded. The norm of $L^1([a,b];\mathbb{R}^p)$ is denoted by $\|\cdot\|_1$ and the norm of $L^\infty([a,b];\mathbb{R}^p)$ is $\|\cdot\|_\infty$.

We make use of concepts from nonsmooth analysis. Thorough discussion of basic concepts of nonsmooth analysis can be found ,for example, in [2], [6], [22], [20] and [17]. Here we introduce only the notation of some concepts used throughout this paper.

Let $A \subset \mathbb{R}^n$ to be a closed set with and consider $x_* \in A$. The *limiting normal cone to A at x_** (also known as *Mordukhovich* normal cone) is denoted by $N_A^L(x_*)$ while the *Clarke* normal cone is $N_A^C(x_*)$.

Take a lower semicontinuous function $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x_* \in \mathbb{R}^k$ where $f(x_*) < +\infty$. Then the *limiting subdifferential*, also known as *Mordukhovich subdifferential*, of f at x_* is denoted by $\partial^L f(_*)$. Recall that when the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, co $\partial^L f(x)$, coincides with the *(Clarke) subdifferential*, denoted here by $\partial^C f(x)$.

2. Main Results

In this section we present two variants nonsmooth maximum principles for (P) of different nature. The first one, denoted here simply as the nonsmooth maximum principle, is closed related to Theorem 6.1 in [4]. The second result is a *hybrid* nonsmooth maximum principle in line with Theorem 3.2 also in [4]. Their difference lies in the assumptions; while the function f is assumed locally Lipschitz continuous with respect to u for the first Theorem, in the second case only measurability of f with respect to u is imposed. Both results hold for *strong local minimizers* for (P), whose definition we present next.

A pair (x,u), comprising an absolutely continuous function x and a measurable function u is an admissible process for (P), if it satisfies all the constraints of the problem.We say that (x_*,u_*) is a *strong local minimizer* for (P) if it is an admissible process for (P) minimizing the cost $J(x,u) := l(x(a),x(b))$ over all admissible

processes (x, u) such that

$$|x(t) - x_*(t)| \leq \varepsilon$$

for some $\varepsilon > 0$.

2.1. **Nonsmooth Maximum Principles for** (P)**.** Define the sets

$$S(t) := \{(x, v, u) : (f(t, x, v, u), u) \in \Phi \times U\}. \tag{2.1}$$

$$S_*^\epsilon(t) := \{(x, v, u) \in S(t) : \ |x - x_*(t)| \leq \varepsilon\} \tag{2.2}$$

and

$$S(t, u) := \{(x, v) : f(t, x, v, u) \in \Phi\}. \tag{2.3}$$

$$S_*^\epsilon(t, u) := \{(x, v) \in S(t, u) : \ |x - x_*(t)| \leq \varepsilon\} \tag{2.4}$$

The following *basic hypotheses* are imposed throughout: the function l is locally Lipschitz, $(t, (x, v, u)) \to f(t, (x, v, u))$ is $\mathcal{L} \times \mathcal{B}$ measurable [1], the set $S(t)$ is closed, the graph of $t \to S(t)$ is $\mathcal{L} \times \mathcal{B}$ measurable and the set U is compact and Φ and E are closed sets.

Consider also the following assumptions.

L1* There exists a constant k_f such that, for almost every $t \in [a, b]$, for every (x_i, v_i, u_i) with $|x_i - x_*(t)| \leq \varepsilon$, we have

$$|f(t, x_1, v_1, u_1) - f(t, x_2, v_2, u_2)| \leq k_f[|x_1 - x_2| + |v_1 - v_2| + |u_1 - u_2|].$$

CQ1 There exists constant M such that, for almost every $t \in [a, b]$, all $(x, v, u) \in S_*^\varepsilon(t)$ and all $(\lambda, \mu) \in N_\Phi^L(f(t, x, v, u)) \times N_U^L(u)$, we have

$$(\alpha, \beta_1, \beta_2 - \mu) \in \partial_{x,v,u}^L \langle \lambda, f(t, x, v, u)\rangle \implies |\lambda| \leq M|(\beta_1, \beta_2)|.$$

For our *hybrid* nonsmooth maximum principle, **L1*** and **CQ1** are replaced by the assumptions stated next.

L2* There exists a constant k_f such that, for almost every $t \in [a, b]$, for every (x_i, v_i) in a neighborhood of $S_*^\varepsilon(t, u)$, $(i = 1, 2)$, we have

$$|f(t, x_1, v_1, u_1) - f(t, x_2, v_2, u_2)| \leq k_f[|x_1 - x_2| + |v_1 - v_2|].$$

CQ2 For each $u \in U$, the set $S(t, u)$ is closed and there exists a constant M such that, for almost every $t \in [a, b]$, all $u \in U$, all $(x, v) \in S_*^\varepsilon(t, u)$ and all $\lambda \in N_\Phi^L(f(t, x, v, u))$, we have

$$(\alpha, \beta) \in \partial_{x,v}^L \langle \lambda, f(t, x, v, u)\rangle \implies |\lambda| \leq M|\beta|.$$

Our first result is a simple adaptation of Theorem 6.1 in [4]; it holds under assumptions that, although stronger than those appearing in [4], are nevertheless of interest for applications.

[1] relative to the σ-field generated by the product of Lebesgue measurable subsets in $\mathbb{R}$ and Borel measurable subsets in $R^n \times \mathbb{R}^m \times \mathbb{R}^k$

Theorem 2.1. *Let (x_*, u_*) be a strong local minimizer for (P). Assume that the basic assumptions,* **L1*** *and* **CQ1** *are satisfied. Then there exist $p \in W^{1,1}([a,b];\mathbb{R}^n)$ and a scalar $\lambda_0 \geq 0$ such that:*

$$||p||_\infty + \lambda_0 > 0, \tag{2.5}$$

$$(p(a), -p(b)) \in N_E^L(x_*(a), x_*(b)) + \lambda_0 \partial^L l(x_*(a), x_*(b)), \tag{2.6}$$

for almost every $t \in [a,b]$

$$(-\dot{p}(t), 0, 0) \in \partial^C_{x,v,u}\langle p(t), \dot{x}_*(t)\rangle - N^C_{S(t)}(x_*(t), \dot{x}_*(t), u_*(t)) \tag{2.7}$$

and, for all (v,u) such that $u \in U$ and $f(t, x_*(t), v, u) \in \Phi$, we have

$$\langle p(t), v\rangle \leq \langle p(t), \dot{x}_*(t)\rangle. \tag{2.8}$$

Observe that Theorem 6.1 in [4] holds when $\Psi = \{0\}$. However, the tools in [4] permit its extension to closed sets Ψ.

A special feature of Theorem 2.1 is the Lipschitz behavior of f with respect to the control, an assumption not enforced to obtain other necessary conditions available in the literature. This situation can be partially fixed appealing to Theorem 3.2 in [4]. This yields our second result:

Theorem 2.2. *Let (x_*, u_*) be a local minimum for problem (P). Assume that the basic assumptions as well as* **L2*** *and* **CQ2** *are satisfied. Then there exist $p \in W^{1,1}([a,b];\mathbb{R}^n)$ and a scalar $\lambda_0 \geq 0$ such that conditions* (2.5) *and* (2.6) *in Theorem 2.1 are satisfied together with:*

$$(-\dot{p}(t), 0) \in \partial^C_{x,v}\langle p(t), \dot{x}_*(t)\rangle - N^C_{S(t,u_*(t))}(x_*(t), \dot{x}_*(t)) \quad a.e. \tag{2.9}$$

and, for all $u \in U$ and $(x_*(t), v) \in S(t,u)$ for a.e. t,

$$\langle p(t), v\rangle \leq \langle p(t), \dot{x}_*(t)\rangle. \tag{2.10}$$

Proof. Rewrite (P) in the following form

$$\begin{cases} \text{Minimize} \quad l(x(a), x(b)) \\ \text{subject to} \\ \qquad \dot{x}(t) = v(t) \quad \text{a.e.} \\ \qquad (x(t), v(t)) \in S(t, u(t)) \quad \text{a.e.} \\ \qquad u(t) \in U \quad \text{a.e.} \\ \qquad (x(a), x(b)) \in E. \end{cases}$$

Application of Theorem 3.2 in [4] yields the required conditions. □

Clearly, Theorem 3.2 in [4] originates Theorem 2.2, when applied to our problem. The division on the control into two components,one constrained and another unconstrained, is not new; we refer the reader [13] and references within in this respect.

The applicability of both Theorems is shadowed by the presence of the normal cone of $S(t)$ or $S(t, u^*(t))$ in the adjoint inclusions.

3. Smooth Case

We now explore the implications of the above theorems to smooth problems. First, however, let us consider an intermediate case when all the assumptions of Theorem 2.1 are enforced and, additionally, $(x,v,u) \to f(t,x,v,u)$ is *strict differentiable* at $(x_*(t), \dot{x}_*(t), u_*(t))$ for almost every t and that both sets Φ and U are such that their limiting normal cones coincide to their Clarke normal cone (i.e, when Φ and U are *regular* in the sense of Clarke, [2]). Under these additional hypotheses the adjoint inclusion in Theorem 2.1 can be written in an explicit multiplier form as we see next.

Suppose that, in addition to the hypotheses of Theorem 2.1, $(x,v,u) \to f(t,x,v,u)$ is *strict differentiable* at $(x_*(t), \dot{x}_*(t), u_*(t))$ and

$$\left\{\begin{array}{rcl} N^L_\Phi(f(x_*(t), \dot{x}_*(t), u_*(t))) & = & N^C_\Phi(f(t, x_*(t), \dot{x}_*(t), u_*(t))), \\ N^L_U(u_*(t)) & = & N^C_U(u_*(t)) \end{array}\right. \tag{3.1}$$

(It is well known that when Φ and U are convex, then (3.1) holds.) In this situation Proposition 4.1 in [4] asserts the existence of measurable functions $\mu : [a,b] \to \mathbb{R}^k$ and $\lambda : [a,b] \to \mathbb{R}^N$, where $\mu(t) \in N^C_U(u_*(t))$ and $\lambda(t) \in N^C_\Phi(f(t, x_*(t), \dot{x}_*(t), u_*(t)))$ a.e., such that (2.7) reads

$$(-\dot{p}(t), 0, \mu(t)) \in \partial^C_{x,v,u}\langle p(t), \dot{x}_*(t)\rangle - \partial^C_{x,v,u}\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle. \tag{3.2}$$

Appealing to the properties of Clarke subdifferential for strict differentiablity (see [2]) we have

$$\partial^C_{x,v,u}\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle = \nabla_{x,v,u}\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle$$

and $\partial^C_{x,v,u}\langle p(t), \dot{x}_*(t)\rangle = (0, p(t), 0)$. It follows that

$$\left\{\begin{array}{rcl} \dot{p}(t) & = & \nabla_x\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle, \\ p(t) & = & \nabla_v\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle, \\ -\mu(t) & = & \nabla_u\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle, \end{array}\right.$$

where $\lambda(t) \in N^C_\Phi(f(t, x_*(t), \dot{x}_*(t), u_*(t)))$ and $\mu(t) \in N^C_U(u^*(t))$. We summarize our findings in the following Corollary:

Lemma 3.1. *Let (x_*, u_*) be a local minimum for problem (P). Assume that the assumptions of Theorem 2.1 hold and that $(x,v,u) \to f(t,x,v,u)$ is strict differentiable at $(x_*(t), \dot{x}_*(t), u_*(t))$ a.e. and* (3.1) *holds almost everywhere. Then there exist $p \in W^{1,1}([a,b];\mathbb{R}^n)$, measurable functions λ and μ, where $\mu(t) \in N^C_U(u_*(t))$ and $\lambda(t) \in N^C_\Phi(f(t, x_*(t), \dot{x}_*(t), u_*(t)))$ a.e., and a scalar $\lambda_0 \geq 0$ satisfying* (2.5), (2.6), (2.8) *and*

$$(\dot{p}(t), p(t), -\mu(t)) = \nabla_{x,v,u}\langle \lambda(t), f(t, x_*(t), \dot{x}_*(t), u_*(t))\rangle \ \ a.e.. \tag{3.3}$$

Furthermore,

$$|\lambda(t)| \leq M|p(t)| \ \ a.e. \tag{3.4}$$

where M is the constant in **CQ1** .

Lemma 3.1 covers the *strict differentiability* version of Theorem 2.1. To avoid being repetitive, we leave out an analogous result one can easily obtain appealing now to Theorem 2.2.

Lemma 3.1 requires strict differentiability simply along the optimal solution. Clearly, it holds for (P) when f is C^1, the set U convex and $\Phi = \{0\}$. To further illustrate the implications of the above necessary conditions we now focus on the smooth case (f in C^1) with $\Phi = \{0\}$.

In this scenario, it is a simple matter to see that **CQ1** is equivalent to the *full row rankness* of matrix $f_v(t, x, v, u)$ around the optimal solution (clearly, an implicit condition is $n \geq N$), a condition we can write as

$$\begin{cases} \text{for all } \lambda \in \mathbb{R}^N, \text{ all } (x, v, u) \in S_*^\varepsilon(t): \\ \nabla_v \langle \lambda, f(t, x, v, u) \rangle = 0 \implies \lambda = 0. \end{cases} \tag{3.5}$$

We now turn to (3.3). To do so let us set

$$f_v^+(t, x_*(t), \dot{x}_*(t), u_*(t)) = \Big(\nabla_v f(t)(\nabla_v f(t)^T\Big)^{-1} \nabla_v f(t),$$

where $\nabla_v f(t) = \nabla_v f(t, x_*(t), \dot{x}_*(t), u_*(t))$. This matrix is the left inverse of $\big(\nabla_v f(t, x_*(t), \dot{x}_*(t), u_*(t))\big)^T$. Then, the multiplier λ in (3.3) reduces to

$$\lambda(t) = f_v^+(t, x_*(t), \dot{x}_*(t), u_*(t))p(t) \tag{3.6}$$

and we get

$$\dot{p}(t) = (\nabla_x f(t, x_*(t), \dot{x}_*(t), u_*(t)))^T f_v^+(t, x_*(t), \dot{x}_*(t), u_*(t))p(t) \tag{3.7}$$

$$-\mu(t) = (\nabla_u f(t, x_*(t), \dot{x}_*(t), u_*(t)))^T f_v^+(t, x_*(t), \dot{x}_*(t), u_*(t))p(t), \tag{3.8}$$

where, as before, $\mu \in N_U^C(u_*(t))$.

The maximum principle for the smooth version of (P) is then given by (3.7), (3.8), with $\mu \in N_U^C(u_*(t))$, together with (2.5), (2.6) and (2.8). The adjoint equation (3.7) is of a different nature. A possible alternative way to obtain such conditions for smooth problem would be the use of implicit function theorems similar to what is done in [7].

For completeness we summarize our findings below.

Corollary 3.2. *Let (x_*, u_*) be a local minimum for problem (P) where $\Phi = \{0\}$. Assume the basic assumptions. Assume also that the function f is C^1, (3.5) holds and the set U is compact and satisfies the condition*

$$N_U^L(u) = N_U^C(u). \tag{3.9}$$

Then there exist $p \in W^{1,1}([a, b]; \mathbb{R}^n)$, a measurable function μ, where $\mu(t) \in N_U^C(u_*(t))$ a.e., and a scalar $\lambda_0 \geq 0$ satisfying (2.5), (2.6), (2.8), (3.7) and (3.8).

Remark. The above Corollary can be easily adapted to cover situations when $\Phi = \{x : x \leq 0\}$ or $\Phi = \{0\} \times \{x : x \leq 0\}$. In these cases, the smooth counterparts of **CQ1** can be easily obtained (in this respect we refer to reader to [4]).

4. Special Cases

We consider some problems involving DAE's. Throughout we consider the assumptions under which Corollary 3.2 is valid and

$$f(t, x, \dot{x}, u) = E\dot{x} - g(t, x, u), \tag{4.1}$$

where E is a $N \times n$ matrix with $\text{rank}(E) = r$. This incorporates three different situations. Matrix E may be

Case (A): of full row rank ($N \leq n$ and $r = N$);
Case (B): of column full rank ($N > n$ and $r = n$);
Case (C): $r < \min\{n, N\}$.

Here we dwell on the smooth cases. Remarkably, however, Theorems 2.1 or 2.2 or Lemma 3.1 are of importance because they provide necessary conditions for the DAE's problems (4.1) (and some more general one, indeed) when the function g is nonsmooth. To keep our analysis simple, we do not state such results; they can be easily obtained with the tools developed here.

In what follows, we consider the data smooth and the matrix E appearing in the system (4.1) in all the three cases **(A)**–**(C)** but under some simple forms. For simplicity of exposition we assume the matrix E to be constant. The case where E is dependent on t, if its required properties are assumed to hold for almost every t, can be treated analogously.

4.1. **Necessary conditions for a case (A).** In case **(A)**, we consider matrix E in (4.1) to be of the form

$$E = \begin{bmatrix} E_a & 0 \end{bmatrix},$$

where E_a is a $N \times N$ nonsingular matrix. Considering x partitioned as $x = (y, z)$, with $y \in \mathbb{R}^N$ and $z \in \mathbb{R}^{n-N}$, the equation (4.1) reduces to the ODE

$$\dot{y}(t) = E_a^{-1} g(t, y, z, u). \tag{4.2}$$

An important feature of (4.2) is the presence of the z, a component of state variable x, not associated with a differential equation. If we were to apply known necessary conditions to (P) involving (4.2), we would question the role of z: should it be a control or a state? In situation where (4.1) reduces to (4.2), we may have to consider z to be a "state". In this case, we could reformulate system $E_a\dot{y}(t) - g(t, y, z, u) = 0$ to the form

$$\begin{cases} \dot{y}(t) &= v_a(t), \\ \dot{z}(t) &= v_b(t), \\ 0 &= E_a v_a(t) - g(t, y(t), z(t), u(t)). \end{cases}$$

Here $v = (v_a, v_b)$ plays the role of an unconstrained control. However, the Jacobian of $E_a v_a - g(t, y, z, u)$ with respect to v would not be of full rank and the application of Corollary 3.2 would be compromised, unless the lack of full rankness were compensated identifying the control v with v_a and some additional components of u.

To apply Corollary 3.2 to our general system $E_a\dot{y}(t) - g(t, y, z, u) = 0$, we consider the system

$$\begin{cases} \dot{y}(t) &= v_a(t), \\ 0 &= E_a v_a(t) - g(t, y(t), z(t), u(t)). \end{cases}$$

Here the control variable is (v_a, z, u), where both v_a and z are unconstrained. Moreover, we need to assume that the cost function $l(x(a), x(b))$ depends only on y and the constraint $(x(a), x(b)) \in E$ reduces to $(y(a), y(b)) \in E$. This yields the following result.

Corollary 4.1. *Consider x partitioned as $(y, z) \in \mathbb{R}^N \times \mathbb{R}^{n-N}$. Let (y_*, z_*, u_*) be a local minimum for (P) when*

$$f(t, y, z, \dot{y}, \dot{z}, u) = E_a\dot{y}(t) - g(t, y, z, u),$$

where g is a C^1 function, E_a is a $N \times N$ nonsingular matrix, U is compact and satisfies (3.9)*, $l(x(a), x(b)) = l(y(a), y(b))$, $(x(a), x(b)) \in E$ reduces to $(y(a), y(b)) \in E$. Then there exist $p \in W^{1,1}([a, b]; \mathbb{R}^N)$, a measurable function $\mu : [a, b] \to \mathbb{R}^k$, with $\mu(t) \in N_U^C(u_*(t))$ a.e., and a scalar $\lambda_0 \geq 0$ such that:*

$$||p||_\infty + \lambda_0 > 0, \tag{4.3}$$

$$(p(a), -p(b)) \in N_E^L(y_*(a), y_*(b)) + \lambda_0 \partial^L l(y_*(a), y_*(b)), \tag{4.4}$$

$$\dot{p}(t) = -(\nabla_y g(t, y_*(t), z_*(t), u_*(t)))^T \left(E_a^T\right)^{-1} p(t) \ \ a.e. \tag{4.5}$$

$$0 = -(\nabla_z g(t, y_*(t), z_*(t), u_*(t)))^T \left(E_a^T\right)^{-1} p(t) \ \ a.e. \tag{4.6}$$

$$-\mu(t) = -(\nabla_u g(t, y_*(t), z_*(t), u_*(t)))^T \left(E_a^T\right)^{-1} p(t) \ \ a.e. \tag{4.7}$$

and, for all $u \in U$,

$$\langle p(t), E_a^{-1} g(t, y_*(t), z_*(t), u)\rangle \leq \langle p(t), E_a^{-1} g(t, y_*(t), z_*(t), u_*(t))\rangle. \tag{4.8}$$

Indeed, and as we would expect, these necessary conditions coincide with well known necessary conditions when the ODE considered is (4.2). It is however worth mentioning that the equations (4.5) and (4.6) come in the form of (4.1) setting $q(t) = E_a^T p(t)$ and recalling that $x = (y, z)$ we get

$$E^T \dot{q}(t) + (\nabla_x g(t, x_*(t), u_*(t)))^T q(t) = 0.$$

In this way, we get adjoint equations which themselves come in the form of the implicit system.

4.2. **Necessary conditions for a case (B).** Turning now to case **(B)**, let us consider $E = \begin{bmatrix} E_b \\ 0 \end{bmatrix}$, where E_b is a $n \times n$ nonsingular matrix and

$$g(t, x, u) = \begin{bmatrix} g_d(t, x, u) \\ g_a(t, x, u) \end{bmatrix}.$$

It follows that (4.1) reduces to

$$\begin{cases} \dot{x}(t) &= E_b^{-1} g_d(t, x, u), \\ 0 &= g_a(t, x, u). \end{cases} \tag{4.9}$$

Clearly, this is a DAE system, that is, (4.9) comprises an ordinary differential equation (ODE) coupled with an algebraic equation. However, the variable x is not partitioned as in the previous case and so there is no "fast" state variable. Thus the *full column rank* of E in (4.1) yields the system (4.9) with *mixed constraints* in the form of equality. Application of Corollary 3.2 is possible if the lack of full rankness is compensated with the derivatives of g_a with respect to some constrained components of u (see [4]). It is worth mentioning that the conditions under which Corollary 3.2 holds require that assumption

4.3. **Necessary conditions for a case (C).** We finally turn to case **(C)**. We make no assumptions on how N and n are related and we consider

$$E = \begin{bmatrix} E_c & 0 \\ 0 & 0 \end{bmatrix},$$

where E_c is a $r \times r$ nonsingular matrix, where $r < \min\{n, N\}$. Under such circumstances we consider x partitioned as

$$x = (y, z) \text{ and } g(t, y, z, u) = \begin{bmatrix} g_d(t, y, z, u) \\ g_a(t, y, z, u) \end{bmatrix},$$

where $y \in \mathbb{R}^r$, $g_d(t, y, z, u) \in \mathbb{R}^r$, $g_a(t, y, z, u) \in \mathbb{R}^{n-r}$ and $z \in \mathbb{R}^{n-r}$. Thus (4.1) now reads as the following DAE's

$$\begin{cases} \dot{y}(t) &= E_c^{-1} g_d(t, y, z, u), \\ 0 &= g_a(t, y, z, u). \end{cases} \tag{4.10}$$

In contrast with (4.9), we now have a "slow" state y (or "differential" state) and a "fast" state z (or "algebraic" state).

It is worth mentioning that when the end point constraints and the cost function do not depend on z, then z may be seen as an unconstrained control (see [15]). We can then define the control to be $w = (z, u)$. If, furthermore, $\nabla_z g_a(t, y, z, u)$ is invertible, then (4.10) is an *index one* DAE's.

5. Conclusions

We extended the results in [4] to cover nonsmooth optimal control problem problems with implicit constraints expressed as set constraints. We also considered the case where the control is partitioned into two components, one constrained and another component unconstrained, as in [13], when only measurability with respect to the constrained control component is assumed. Our approach is based on (1.3). For our problem (P) with smooth data, necessary conditions have been derived appealing to implicit function theorem (see, for example, [7], [21] [11], [19], and [14]). Corollary 3.2 yields the same set of necessary conditions but avoids the calculation of implicit functions.

It is our belief that our results for smooth problem could possibly be obtained using an approach in line with that developed in [13](if this has been done, it is unknown to us). Noteworthy, this would enable the addition of pure state constraints which are considered in [13] but not in [4]. Nevertheless, it is fair to expect that in the near future the recently developments in [1] may be used together with those in [4] to allow for state constraints to be included for nonsmooth problem in

the form of (P), given rise to generalizations of both Theorems 2.1 and 2.2. The introduction of state constraints is of importance when *higher index* DAE systems are considered. These will be the focus of future research.

Acknowledgments

The author thanks Hélène Frankowska for calling her attention to implict systems and her paper [8]. We also thank the anonymous referee for various suggestions and comments that greatly helped improved this paper.

References

[1] P. Bettiol, A. Boccia and R. Vinter, *Stratified Necessary Conditions for Differential Inclusions with State Constraints*, SIAM J. Control Optim. **51** (2013), 3903–3917.
[2] F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.
[3] F. Clarke, *Necessary Conditions in Dynamic Optimization*, Mem. Amer. Math. Society, 2005.
[4] F. Clarke, MdR de Pinho, *Optimal control problems with mixed constraints*, SIAM J. Control Optim. **48** (2010), 4500–4524.
[5] F. Clarke, Y. Ledyaev, MdR de Pinho, *An extension of the Schwarzkopf multiplier rule in optimal control*, SIAM J. Control Optim. **49** (2011), 599–610.
[6] F. Clarke, Y. Ledyaev, R. J. Stern, P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer-Verlag, New York, 1998.
[7] M. de Pinho and R. Vinter, *Necessary conditions for optimal control problems involving nonlinear differential algebraic equations*, J. Math. Anal. Appl. **21** (1997), 493–516.
[8] H. Frankowska, *Contingent Cones to Reachable Sets of Control systems*, SIAM J. Control Optim. **27** (1989), 170–198.
[9] M. Gerdts, *Local minimum principle for optimal control problems subject to differential-algebraic equations of index two*, J. Optim. Theory Appl. **130** (2006), 441–460.
[10] M. Gerdts, *ptimal Control of Ordinary Differential Equations and Differential-Algebraic Equations*, Habilitation Thesis, Department of Mathematics University of Bayreuth, Germany, 2006.
[11] M. Gerdts, *Optimal Control of Odes and Daes*, Walter de Gruyter, 2012.
[12] E. N. Devdaryani, Y. S. Ledyaev, *Maximum principle for implicit control systems*, Appl. Math. Optim. **40** (1996), 79–103.
[13] A. V. Dmitruk, *Maximum principle for the general optimal control problem with phase and regular mixed constraints.* Comput. Math. Model. **4** (1993), 364–377.
[14] I. Kornienko, M. Gerdts, M. de Pinho, *New version of necessary conditions for optimal control problems with differential algebraic equations*, in: Proceedings of MTNS 2012, Melbourne, Australia, July 2012.
[15] P. Kunkel and M. Volker, *Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index*, Math. Control Signals Syst. **22** (2008) 227–269.
[16] P. Loewen, *Optimal Control via Nonsmooth Analysis*, CRM, American Mathematical Society, 1993.
[17] B. Mordukhovich, *Variational Analysis and Generalized Differentiation. Basic Theory*, Fundamental Principles of Mathematical Sciences, vol. 330, Springer-Verlag, Berlin, 2006.
[18] R. Pytlak, *Optimal control of differential-algebraic equations of higher index, Part 1: First-order approximations*, J Optim Theory Appl. **134** (2007), 61–75.
[19] R. Pytlak, *Optimal control of differential-algebraic equations of higher index, Part 2: Necessary optimality conditions*, J. Optim. Theory Appl. **134** (2007), 77–90 .
[20] R. T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. vol. 317 Springer Verlag, Berlin, 1988.
[21] T. Roubicek and M. Valasek, *Optimal control of causal differential-algebraic systems*, J. Math. Anal. Appl. **269** (2002), 616–641.
[22] R. Vinter, *Optimal Control*, 2000, Birkhäuser, Boston.

Manuscript received 18 December 2015
revised 1 March 2016

MdR de Pinho
Faculdade de Engenharia da Universidade do Porto, DEEC, SYSTEC, Rua Dr. Roberto Frias, s/n, 4200–465 Porto, Portugal
E-mail address: mrpinho@fe.up.pt