A maximum principle for optimal control problems with mixed constraints

M. D. R. De Pinho  
Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia da Universidade do Porto, 4099 Porto Codex, Portugal

R. B. Vinter  
Centre for Process Systems Engineering and Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London SW7 2BY, UK

And  
H. Zheng  
Department of Business Studies, Edinburgh University, 50 George Street, Edinburgh EH8 9JY, UK

[Received on 15 March 1999; accepted on 10 September 1999]

Necessary conditions in the form of maximum principles are derived for optimal control problems with mixed control and state constraints. Traditionally, necessary conditions for problems with mixed constraints have been proved under hypothesis which include the requirement that the Jacobian of the mixed constraint functional, with respect to the control variable, have full rank. We show that it can be replaced by a weaker ‘interiority’ hypothesis. This refinement broadens the scope of the optimality conditions, to cover some optimal control problems involving differential algebraic constraints, with index greater than unity.

Keywords: optimal control; maximum principle; mixed constraints; differential algebraic equations.

1. Introduction

Consider the optimal control problem

\[
\begin{align*}
\text{(P)} & \quad \text{Minimize } h(x(a), x(b)) \\
& \quad \text{subject to} \\
& \quad \dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
& \quad 0 = b(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
& \quad u(t) \in U(t) \quad \text{a.e. } t \in [a, b] \\
& \quad (x(a), x(b)) \in C
\end{align*}
\]

where \( h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, b : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k, \) \( U(t) \subset \mathbb{R}^m \) for each \( t \in [a, b] \) \( C \subset \mathbb{R}^n \times \mathbb{R}^n. \)
A (Lebesgue) measurable function \( u : [a, b] \to \mathbb{R}^m \) such that \( u(t) \in U(t) \) a.e. \( t \in [a, b] \) is called a control function. A function \( x \in AC([a, b]; \mathbb{R}^p) \) satisfying \( \dot{x}(t) = f(t, x(t), u(t)) \) and \( 0 = b(t, x(t), u(t)) \) a.e. is called a state trajectory (corresponding to \( u \)). A couple \((x, u)\) comprising a state trajectory \( x \) and the control \( u \) with which it is associated is called a process. It is a minimizer if it minimizes the cost over processes satisfying the endpoint constraints. Take a minimizer \((\bar{x}, \bar{u})\).

In this paper we seek necessary conditions of optimality in the form of a maximum principle for \((P)\). A notable feature of the class of problems which we consider is that the differential and algebraic constraints cannot automatically be reduced by state transformations to a more tractable set of pure differential constraints.

By considering \((P)\) as a limiting case as \( \epsilon \to 0 \) of an optimal control problem with additional dynamic constraint \( \epsilon \dot{y} = b \), we are led to expect that a maximum principle applies which asserts the existence of an absolutely continuous function \( p \), a measurable function \( q \) and a scalar \( \lambda \geq 0 \) such that

\[
\|p\| + \lambda = 1, \\
-\dot{p}(t) = p(t)f_x(t, \bar{x}(t), \bar{u}(t)) + q(t)b_x(t, \bar{x}(t), \bar{u}(t)), \\
u \to p(t)f(t, \bar{x}(t), u) + q(t)b(t, \bar{x}(t), u)
\]
is maximized over \( U(t) \) at \( u = \bar{u}(t) \) and

\[(p(a), -p(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \partial h(\bar{x}(a), \bar{x}(b)).\]

Here \( N_C \) and \( \partial h \) denote the limiting normal cone and limiting subdifferential (see the definitions below). For simplicity we treat the case of smooth \( f \) and \( b \).

We focus on the question of whether optimality conditions of this type are valid. \((P)\) is a problem with ‘mixed’ state and control functional equality constraints. Such problems have been addressed by a number of authors (Hestenes, 1966; Stefani & Zezza, 1996, etc). A common hypothesis under which necessary conditions have previously been derived is that the Jacobi matrix \( \nabla_u b(t, \bar{x}(t), \bar{u}(t)) \) has full rank. An exception is to be found in an earlier paper by Schwarzkopf (1976), who briefly sketches, in the case of a scalar state variable, a derivation of necessary conditions when the full rank condition is replaced by the requirement that

\[
\delta B \subset b(t, \bar{x}(t), U(t)) \tag{1.1}
\]

for some \( \delta > 0 \). (Notice that this ‘interiority’ condition does not guarantee that \( \nabla_u b(t, \bar{x}(t), \bar{u}(t)) \) has full rank.) This hypothesis is of interest because (in contrast to the full-rankness condition) it is satisfied in certain cases when \((P)\) is a reformulation of an optimal control problem involving differential algebraic systems of index higher than unity (or indeed of systems for which the index is not even defined).

Different variational techniques are used in Devdaryani & Ledyayev (in press) to derive a maximum principle, under a less restrictive constraint qualification than (1.1), for problems reducible to free endpoint problems. By contrast, we treat general constrained endpoint constraints.

In this paper we provide a simple derivation of necessary conditions for optimal control problems with mixed constraints, under the above interiority hypothesis. Additionally we...
assume that the generalized velocity set

\[ \{(f(t,x,u),b(t,x,u)) : u \in U(t)\} \]

is convex. Unlike the results of Schwarzkopf (1976) we allow nonsmooth data. The nature of the hypotheses is explored in a number of examples. The proof technique is to associate with \((P)\) an ‘auxiliary’ control problem, to which we apply a recent sharpened form of the Euler–Lagrange inclusion, and then show that the resulting conditions can be transformed into conditions of the required nature.

2. Preliminaries

\(|\cdot|\) will denote the Euclidean norm. \(B\) is the closed unit ball. A vector \(a\) is called nonnegative if all its components are nonnegative, written as \(a \geq 0\). Similarly a matrix \(A\) is nonnegative if all its entries are nonnegative, written as \(A \geq 0\).

We shall make use of some properties of \(M\)-matrices, which we list here for convenience.

**Definition 2.1** Let \(A\) be a \(m \times m\) real matrix. \(A\) is an \(M\)-matrix if \(a_{ij} \leq 0, i \neq j\) and the real parts of the eigenvalues of \(A\) are positive.

**Lemma 2.2** (Berman & Plemmons, 1994) The following statements are equivalent:

(i) \(A\) is an \(M\)-matrix.
(ii) There exists \(A^{-1}\) and \(A^{-1} \geq 0\).
(iii) \(a_{ii} > 0\) and \(a_{ii} > \sum_{j \neq i} |a_{ij}|\) for \(i = 1, \ldots, m\).

We shall also require the following constructs from nonsmooth analysis.

**Definition 2.3** Given a closed set \(A \subset \mathbb{R}^k\) and a point \(x \in A\), the limiting normal cone to \(A\) at \(x\), written \(N_A(x)\), is the set of all vectors \(p \in \mathbb{R}^k\) satisfying the following condition: there exist \(p_i \to p\) and \(x_i \to x\) in \(A\) such that, for each \(i\)

\[ p_i \cdot (x - x_i) \leq o(|x - x_i|) \text{ for all } x \in A, \]

in which \(o(\alpha)/\alpha \to 0\) as \(\alpha \downarrow 0\).

**Definition 2.4** Given a lower semicontinuous function \(f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}\) and a point \(x \in \mathbb{R}^k\) such that \(f(x) < +\infty\), the limiting subdifferential of \(f\) at \(x\), written \(\partial f(x)\), is the set

\[ \partial f(x) := \{\xi : (\xi, -1) \in \text{epi}(f)(x, f(x))\} \]

in which \(\text{epi}(f)\) denotes the epigraph set \(\{(x, \alpha) \in \mathbb{R}^k \times \mathbb{R} : \alpha \geq f(x)\}\).

Properties of limiting normal cones, limiting subdifferentials and associated calculus rules can be found in Clarke (1983), Loewen (1993) and Mordukhovich (1994).

Finally, we state an earlier derived weak maximum principle for optimal control problems (see de Pinho & Vinter, 1995) which will be quoted in the proofs.
THEOREM 2.5 Take $\epsilon > 0$. Let $(\tilde{x}, \tilde{u})$ be a minimizer for the problem

\[
\begin{aligned}
\text{Minimize } & h(x(a), x(b)) \\
\text{subject to } & \\
\dot{x}(t) = f(t, x(t), u(t)) & \text{ a.e. } t \in [a, b] \\
u(t) \in U(t) & \text{ a.e. } t \in [a, b] \\
(x(a), x(b)) & \in C
\end{aligned}
\]

over elements $(x, u)$ comprising an absolutely continuous function $x$ and a measurable function $u$ satisfying the constraints of the above problem and for which

\[|x(t) - \bar{x}(t)| \leq \epsilon \quad \text{and} \quad |u(t) - \bar{u}(t)| \leq \epsilon \quad \text{a.e.}\]

Assume that

(i) $f(\cdot, x, u)$ is measurable, and there exists $k \in L^1$ such that

\[|f(t, x, u) - f(t, x', u')| \leq k(t)|(x, u) - (x', u')|\]

for all $x, x' \in \bar{x}(t) + \epsilon B$, $u, u' \in \bar{u}(t) + \epsilon B$ and almost all $t \in [a, b]$.

(ii) Graph $U$ is a Borel measurable set and

\[(\bar{u}(t) + \epsilon B) \cap U(t)\]

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

(iii) $h$ is locally Lipschitz continuous and $C$ is closed.

There exists $\lambda \geq 0$, $p(\cdot) \in AC([a, b]; \mathbb{R}^n)$ and $\zeta(\cdot) \in L^1([a, b]; \mathbb{R}^n)$ such that

\[\lambda + \|p(\cdot)\|_{L^\infty} \neq 0\]

\[(-\dot{p}(t), \dot{x}(t), \zeta(t)) \in \partial p(t) \cdot f(t, \tilde{x}(t), \tilde{u}(t)) \quad \text{a.e.}\]

\[\zeta(t) \in \overline{CN}_{U(t)}(\tilde{u}(t)) \quad \text{a.e.}\]

\[(p(a), -p(b)) \in NC(\tilde{x}(a), \tilde{u}(b)) + \lambda \partial h(\tilde{x}(a), \tilde{x}(b))\]

where $\partial p \cdot f(t, x, u)$ denotes the limiting subdifferential of $(x, p, u) \rightarrow p \cdot f(t, x, u)$ for fixed $t$.

3. Main results

We shall invoke the following hypotheses on the data for $(P)$, which make reference to some process $(\tilde{x}(\cdot), \tilde{u}(\cdot))$ and parameter $\epsilon > 0$:

H1. $f(\cdot, x, \cdot)$ and $b(\cdot, x, \cdot)$ are $\mathcal{L} \times \mathcal{B}$ measurable for each $x$.

H2. There exists an integrable function $K_F(\cdot)$ such that, for almost every $t \in [a, b]$, $f(t, \cdot, u)$ is Lipschitz continuous on $\tilde{x}(t) + \epsilon B$ for $u \in U(t)$.

H3. There exists a constant $K_b > 0$ such that, for almost every $t \in [a, b]$, $b(t, \cdot, u)$ is a continuously differentiable function with Lipschitz constant $K_b$ on $\tilde{x}(t) + \epsilon B$ for all $u \in U(t)$ and almost every $t \in [a, b]$. 
H4. There exists a positive number δ such that

\[ δB \subset b(t, \ddot{x}(t), U(t)) \]

for almost every \( t \in [a, b] \).

H5. \( h \) is Lipschitz continuous on a neighbourhood of \((\ddot{x}(a), \ddot{x}(b))\) and \( C \) is closed.

H6. Graph \( U \) is a Borel measurable set.

H7. \([(f(t, x, u), b(t, x, u)) : u \in U(t)\)] is convex for each \( t \in [a, b] \).

Define the Hamiltonian:

\[ H(t, x, p, q, u) = p \cdot f(t, x, u) + q \cdot b(t, x, u). \]

THEOREM 3.1 (Maximum Principle) Let \((\ddot{x}(\cdot), \ddot{u}(\cdot))\) be a minimizing process of \((P)\).

Assume that, for some \( \epsilon > 0 \), hypotheses H1–H7 are satisfied. Then there exist \( p \in AC([a, b]; \mathbb{R}^n), q(\cdot) \in L^1([a, b]; \mathbb{R}^m) \) and \( \lambda \geq 0 \) such that:

(i) \( \|p\|_\infty + \lambda = 1 \).

(ii) \(-\dot{p}(t) \in \text{co} \, \partial_t H(t, \ddot{x}(t), p(t), q(t), \ddot{u}(t)) \) a.e.

(iii) \( u \to H(t, \ddot{x}(t), p(t), q(t), u) \) is maximized over \( U(t) \) at \( u = \ddot{u}(t) \) a.e.

(iv) \((p(a), -\dot{p}(b)) \in N_C(\ddot{x}(a), \ddot{x}(b)) + \lambda \partial h(\ddot{x}(a), \ddot{x}(b)).\)

We now consider a number of examples which provide some insights into the nature of the optimal control problems to which the above maximum principle is applicable.

EXAMPLE 1 One area of application of optimality conditions for optimal control problems with mixed constraints is to the control of devices modelled by differential algebraic systems of equations (DAE systems). DAE models are widespread in the chemical processing industry. Such problems can take the form

\[
(P) \quad \begin{align*}
\text{Minimize } & h(x(a), x(b)) \\
\text{subject to } & \dot{x}(t) = \ddot{f}(x(t), y(t), v(t)) \quad \text{a.e. } t \in [a, b] \\
& 0 = \ddot{b}(x(t), y(t)) \quad \text{a.e. } t \in [a, b] \\
& v(t) \in V(t) \quad \text{a.e. } t \in [a, b] \\
& y(t) \in A(t) \quad \text{a.e. } t \in [a, b] \\
& (x(a), x(b)) \in C,
\end{align*}
\]

where \( h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \ddot{f} : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}^n, \ddot{b} : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}_l, V(t) \subset \mathbb{R}_m, A(t) \subset \mathbb{R}_l, C \subset \mathbb{R}^n \times \mathbb{R}_m.\)

In this problem, the dynamic constraint \((\ddot{x}, 0) = (\ddot{f}, \ddot{b})\) describes the evolution of the ‘differential’ and ‘algebraic’ variables \( x \) and \( y \), respectively.

In certain cases, a DAE model can be reduced to a system of ordinary differential equations by repeated differentiation. The number of differentiations required to accomplish this task is referred to as the index of the DAE model.

Regarding the derivation of necessary conditions, the most studied case is that when the index is unity. Here, it is assumed that

\[ \nabla_y \ddot{b}(x, y) \text{ has full rank, for all } (x, y). \]
One differentiation of $t \to \hat{b}(x(t), y(t))$ then yields a system of differential equations for $(x, y)$. (See de Pinho & Vinter, 1995).

The analysis in this paper provides optimality conditions for DAE optimal control problems, involving DAE dynamic constraints for which the index is, possibly, greater than unity. Consider the case of $(P)'$ in which $n = l = 1, m = 1$ and

$$h(x_0, x_1) = x_1,$$
$$\tilde{f}(x, y, v) = d(x)y^3 + v,$$
$$\tilde{b} = y^3,$$
$$C = \{0\} \times \mathbb{R}^n.$$

Here $d(\cdot) : \mathbb{R} \to \mathbb{R}$ is some Lipschitz continuous function, $V(t) = [0, 1]$ and $A(t) = [-1, +1].$

The minimizer is clearly $(\hat{x}, \hat{y}, \hat{v}) = (0, 0, 0).$

This problem can be fitted to the framework of problem $(P)$. To do this, we regard $(y, v)$ as the control variable $u = (u_1, u_2)$ and set $f(t, x, u) = \tilde{f}(x, u_1, u_2), b(t, x, u) = \tilde{b}(x, u_1).$ The data for this problem satisfy all the hypotheses of Theorem 3.1 (including the convexity and interiority conditions H4 and H7), yet the DAE dynamic constraint has infinite index (in the sense that the constraint cannot be reduced to a system of first order differential equations by a finite number of differentiations).

The purpose of the next two examples is to illustrate that neither the convexity nor the interiority hypotheses can be dispensed with.

**Example 2** Consider the case of $(P)$ in which $n = k = 1, m = 2$ and

$$h(x_0, x_1) = x_1,$$
$$f(x, y, v) = u_1,$$
$$b = x - u_2^2,$$
$$U(t) = [-1, +1] \times [-1, +1],$$
$$C = \{0\} \times \mathbb{R}.$$

$(\hat{x}, \hat{u}_1, \hat{u}_2) = (0, 0, 0)$ is a minimizer for this problem. If the assertions of Theorem 3.1 were valid, there would exist multipliers $(p(\cdot), q(\cdot), \lambda)$ such that

$$-\dot{p}(t) = q(t)$$
$$0 = \max_{(u_1, u_2) \in U(t)} \{p \cdot u_1 - q \cdot u_2^2\},$$
$$-p(1) = \lambda.$$

The only solution is $(p, q, \lambda) \equiv (0, 0, 0).$ This is not possible, in view of the nontriviality property of the multipliers. It follows that the maximum principle, as stated above, is not valid for this problem. In this example, the interiority hypothesis H4 is violated.
EXAMPLE 3 Consider finally the case of \((P)\) in which \(n = k = 1, m = 2\) and
\[
\begin{align*}
h(x_0, x_1) &= -x_1, \\
f(x, y, v) &= u_1^2, \\
b &= u_1 + u_2, \\
U(t) &= [-1, +1] \times [0, +1], \\
C &= \{0\} \times \mathbb{R}.
\end{align*}
\]
\((\tilde{x}, \tilde{u}_1, \tilde{u}_2) = (0, 0, 0)\) is a minimizer for this problem. If the assertions of Theorem 3.1 were valid, there would exist multipliers \((p(\cdot), q(\cdot), \lambda)\) such that
\[
\begin{align*}
\hat{p}(t) &= 0, \\
0 &= \max_{(u_1, u_2) \in U(t)} \{ p \cdot u_1^2 + q \cdot (u_1 + u_2) \}, \\
p(1) &= \lambda.
\end{align*}
\]

If \(\lambda = 0\), then \(p(t) \equiv 0\). This is not possible, by nontriviality of the multipliers. Suppose, however, that \(\lambda > 0\). By scaling, we can arrange that \(\lambda = 1\). It follows that \(p(t) = q(t) = 1\) and consequently
\[
0 \geq u_1^2 - u_1 - u_2
\]
for all \((u_1, u_2) \in U(t)\). This is impossible. Indeed, if \(u_1 = -\frac{1}{8}\) and \(u_2 = 0\), we are led to the contradiction \(-\frac{1}{8} + \frac{1}{8} = \frac{3}{8} > 0\). The maximum principle, as stated above, does not apply to this problem. The pathological aspect of this problem is that the convexity hypothesis \(H7\) is violated.

Before proving the Theorem 3.1 we briefly discuss a problem where, in addition to the equality constraint \(b(t, x, a) = 0\), a ‘unilateral’ constraint \(g(t, x, a) \leq 0\) is imposed:

\[
\begin{cases}
\text{Minimize } h(x(a), x(b)) \\
\text{subject to} \\
\dot{x}(t) = f(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
0 = b(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
0 \geq g(t, x(t), u(t)) \quad \text{a.e. } t \in [a, b] \\
u(t) \in u(t) \quad \text{a.e. } t \in [a, b] \\
(x(a), x(b)) \in C,
\end{cases}
\]

where \(f, h, b\) and \(U\) are defined as before and \(g : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^l\).

Assume that \((\tilde{x}(\cdot), \tilde{u}(\cdot))\) is an optimal solution of \((Q)\) and that basic hypotheses \(H1\)–\(H3\) and \(H5\)–\(H7\) are satisfied when \(b\) is replaced by \(\tilde{b} : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^L (L = k + l)\) defined by
\[
\tilde{b}(t, x, u) := (b(t, x, u), g(t, x, u)).
\]

Furthermore, assume that the following hypothesis is also verified:
H4′. There exists a positive scalar \( \delta \) such that:

\[
\delta B \subset \{(b(t, \bar{x}(t), U(t)), g(t, \bar{x}(t), U(t)) + v) : v \in \mathbb{R}^I, v \geq 0\}
\]

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

Define the Hamiltonian

\[
H_Q(t, x, p, q, r, u) = p \cdot f(t, x, u) + q \cdot b(t, x, u) + r \cdot g(t, x, u).
\]

For problem (Q) the following result holds:

**Corollary 3.2** Let \((\bar{x}(\cdot), \bar{\bar{u}}(\cdot))\) be a minimizing process for problem (Q). Assume that hypotheses H1–H3, H5–H7 (when \( b \) is replaced by \( \bar{b} \)) together with H4′ are satisfied. Then there exist an absolutely continuous function \( p \), functions \( q \in L^1([a, b]; \mathbb{R}^I) \) and \( r \in L^1([a, b]; \mathbb{R}^I) \) and a scalar \( \lambda \geq 0 \) such that:

(i) \( \|p\|_{\infty} + \lambda = 1 \).

(ii) \( -\bar{p}(t) \in \text{co}_\mathcal{H}_Q(t, \bar{x}(t), p(t), q(t), r(t)) \text{ a.e.} \)

(iii) \( w \to H_Q(t, \bar{x}(t), p(t), q(t), r(t), u) \) is maximized over \( U(t) \) at \( u = \bar{\bar{u}}(t) \text{ a.e.} \)

(iv) \( (p(a), -q(b)) \in N_C(\bar{x}(a), \bar{x}(b)) + \lambda \delta h(\bar{x}(a), \bar{x}(b)). \)

(v) \( r(t) \leq 0, \quad r(t) \cdot g(t, \bar{x}(t), \bar{\bar{u}}(t)) = 0 \text{ a.e. } t \in [a, b]. \)

**Proof of Corollary 3.2.** We start by pointing out that the introduction of a new control \( v \) allows us to rewrite (Q) as

\[
(Q') \quad \begin{align*}
\text{Minimize} & \quad h(x(a), x(b)) \\
\text{subject to} & \quad \begin{cases}
\dot{x}(t) = f(t, x(t), u(t)) & \text{a.e. } t \in [a, b] \\
0 = b(t, x(t), u(t)) & \text{a.e. } t \in [a, b] \\
0 = g(t, x(t), u(t)) + v(t) & \text{a.e. } t \in [a, b] \\
(u(t), v(t)) \in U(t) \times \{v \in \mathbb{R}^I : v \geq 0\} & \text{a.e. } t \in [a, b] \\
(x(a), x(b)) \in C.
\end{cases}
\end{align*}
\]

This new problem is in the form of \( P \) where \( w = (u, v) \) is the new control variable and \( W(t) = U(t) \times \{v \in \mathbb{R}^I : v \geq 0\} \) the control set. Obviously \((\bar{x}, \bar{\bar{u}}, \bar{v})\) is feasible for (Q') where

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

It is easy to see that, in fact, \((\bar{x}, \bar{\bar{u}}, \bar{v})\) is optimal for (Q').

Additionally observe that H4′ implies that

\[
B_{\delta}(0, 0) \subset \bar{b}(t, \bar{x}(t), W(t)) \text{ for almost all } t \in [a, b],
\]

where \( \bar{b}(t, x, w) = (b(t, x, u), g(t, x, u) + v). \)

We now apply Theorem 3.1 to (Q') to obtain functions \( p, q \) and \( r \) and a scalar \( \lambda \geq 0 \) satisfying (i), (ii) and (iv) of the theorem and

\[
p(t) \cdot f(t, \bar{x}, \bar{\bar{u}}) = \max_{(u,v) \in W(t)} \{ p \cdot f(t, \bar{x}, u) + q \cdot b(t, \bar{x}, u) + r \cdot g(t, \bar{x}, u) + r \cdot v \} \text{ a.e. (3.2)}
\]
Let us fix $v = \tilde{v}$ in (3.2). As $\tilde{v} = -g(t, \tilde{x}, \tilde{u})$, it follows that

$$p(t) \cdot f(t, \tilde{x}, \tilde{u}) + r \cdot g(t, \tilde{x}(t), \tilde{u}(t)) = \max_{u \in U(t)} \{ p \cdot f(t, \tilde{x}, u) + q \cdot b(t, \tilde{x}, u) + r \cdot g(t, \tilde{x}(t), u) \} \quad \text{a.e.}$$

(3.3)

i.e. we recover (iii).

Going back to (3.2) and fixing now $u = \tilde{u}$ we obtain

$$0 = \max_{v \geq 0} \{ r \cdot (g(t, \tilde{x}, \tilde{u}) + v) \}. \quad (3.4)$$

If $g(t, \tilde{x}(t), \tilde{u}(t)) < 0$, then (3.4) implies $r = 0$. If $g(t, \tilde{x}(t), \tilde{u}(t)) = 0$, then (3.4) implies $r(t) \leq 0$. We then deduce the complementary slackness condition

$$r(t) \leq 0, \quad r(t) \cdot g(t, \tilde{x}(t), \tilde{u}(t)) = 0 \quad \text{a.e. } t \in [a, b].$$

The proof is complete.

4. Proof of Theorem 3.1

The proof technique is to associate with (P) an ‘auxiliary’ control problem, to which we apply an Euler–Lagrange inclusion previously obtained (see Theorem 2.5), and then show that the resulting conditions can be transformed into conditions of the required nature. The auxiliary problem is a ‘nonsmooth’ optimal control problem.

Proof. We verify the assertions of the theorem initially under the additional hypothesis

(H$^\circ$). There exists a function $c_f \in L^1$ and a number $c_b$ such that

$$|f(t, \tilde{x}(t), u)| \leq c_f(t) \quad \text{and} \quad |b(t, \tilde{x}(t), u)| \leq c_b,$$

for all $u \in U(t)$ and almost every $t \in [a, b]$.

This hypothesis will be eliminated in the final stages of the proof.

Step 1 Define $e_i = (e_{i1}, e_{i2}, \ldots, e_{ik}) \in \mathbb{R}^k$ as

$$e_{ij} = \begin{cases} \delta_1 & \text{if } i = j \\ -\delta_1/k & \text{if } i \neq j \end{cases}$$

where $\delta_1 \in (0, \delta)$ ($\delta$ given by H4). Then by H4 and a measurable selection theorem, there exist control functions $\{u_1(), \ldots, u_k()\}$ and $\{v_1(), \ldots, v_k()\}$ satisfying

$$b(t, \tilde{x}(t), u_i(t)) = e_i,$$

$$b(t, \tilde{x}(t), v_i(t)) = -e_i.$$

Fix $i$. For almost every $t \in [a, b]$ we have

$$b_1(t, x, u_i(t)) \geq b_1(t, \tilde{x}(t), u_i(t)) - K_b|x - \tilde{x}(t)| = \delta_1 - K_b|x - \tilde{x}(t)|,$$

$$b_1(t, x, \tilde{u}(t)) \leq K_b|x - \tilde{x}(t)|.$$
and for \( j \neq i \)

\[
b_j(t, x, u_i(t)) \leq b_j(t, \tilde{x}(t), u_i(t)) + K_b|x - \tilde{x}(t)|
\]

\[
= -\delta_1/k + K_b|x - \tilde{x}(t)|,
\]

\[
b_j(t, x, \bar{u}(t)) \geq -K_b|x - \bar{x}(t)|.
\]

Define

\[
\Delta b(t, x, u) := b(t, x, u) - b(t, x, \bar{u}(t)).
\]

Setting \( \epsilon_1 = \min \left\{ \frac{\delta_1}{2K_b}, \epsilon \right\} \) we deduce that for all \( x \in \tilde{x}(t) + \epsilon_1 B \), almost every \( t \in [a, b] \) we have

\[
\Delta b(t, x, u_i(t)) \geq \delta_1 - K_b|x - \tilde{x}(t)| - K_b|x - \bar{x}(t)|
\]

\[
= \delta_1 - 2K_b|x - \bar{x}(t)| > 0. \tag{4.1}
\]

and, for \( j \neq i \),

\[
\Delta b_j(t, x, u_i(t)) \leq -\delta_1/k + K_b|x - \tilde{x}(t)| + K_b|x - \bar{x}(t)|
\]

\[
= -\delta_1/k + 2K_b|x - \bar{x}(t)| < 0. \tag{4.2}
\]

Similarly, we can show that for all \( x \in \tilde{x}(t) + \epsilon_1 B \) and almost every \( t \in [a, b] \) we obtain

\[
\Delta b_j(t, x, u_i(t)) > -\delta_1/k - 2K_b|x - \tilde{x}(t)|, \tag{4.3}
\]

and also that

\[
\Delta b_i(t, x, v_i(t)) < 0,
\]

\[
\Delta b_j(t, x, v_i(t)) > 0 \quad \text{if} \ j \neq i.
\]

Define the \( k \times k \) matrices:

\[
B_1(t, x) := (\Delta b(t, x, u_1(t)), \ldots, \Delta b(t, x, u_k(t)))
\]

and

\[
B_2(t, x) := (\Delta b(t, x, v_1(t)), \ldots, \Delta b(t, x, v_k(t))).
\]

**Lemma 4.1** There exists an \( \epsilon' > 0 \) such that for all \( x \in \tilde{x}(t) + \epsilon'B \) and almost every \( t \in [a, b] \) both matrices \( B_1(t, x) \) and \( -B_2(t, x) \) are \( M \)-matrices.

**Proof.** We prove that \( B_1(t, x) \) is an \( M \)-matrix. From (4.1), (4.2) and (4.3) it follows that for \( \epsilon' \in (0, \epsilon_1) \), \( x \in \tilde{x}(t) + \epsilon'B \) and for almost all \( t \in [a, b] \),

\[
\sum_{j \neq i} |\Delta b_j(t, x, u_i(t))| \leq \frac{\delta_1(k - 1)}{k} + 2K_b(k - 1)|x - \tilde{x}(t)|
\]

\[
< \delta_1 - 2K_b|x - \tilde{x}(t)|
\]

\[
\leq \Delta b_i(t, x, u_i(t)).
\]

This follows from Lemma 2.2 that for all \( x \in \tilde{x}(t) + \epsilon'B \) and almost all \( t \in [a, b] \), \( B_1(t, x) \) is an \( M \)-matrix. Similar arguments show that \( -B_2(t, x) \) is an \( M \)-matrix. This completes the proof. \( \square \)
Lemma 4.1 and Lemma 2.2 assure us that \( B_1^{-1}(t, x) \geq 0 \) and \( B_2^{-1}(t, x) \leq 0 \) for all \( x \in \bar{x}(t) + \epsilon'B \) and almost all \( t \in [a, b] \).

It can be easily checked that there exists a scalar \( m_\beta > 0 \), which depends on \( \delta \) and \( m \), such that

\[
|B_1^{-1}(t, \bar{x}(t))| \leq m_\beta \quad \text{and} \quad |B_2^{-1}(t, \bar{x}(t))| \leq m_\beta.
\]

**STEP 2** We now choose a finite collection of controls functions \( \{w_i(\cdot)\}_{i=1}^M \) (which includes \( \bar{u}(\cdot) \)). Let \( \beta(t) = (\beta_1(t), \ldots, \beta_M(t)) \) be any measurable function taking values in \( S \) where

\[
S = \left\{ (\beta_1, \ldots, \beta_M) \in \mathbb{R}^M : \sum_{i=1}^M \beta_i \leq 1, \quad \beta_i \geq 0, \quad i = 1, \ldots, M \right\}.
\]

Define the vector-valued functions \( \gamma^1, \gamma^2 : [a, b] \times \mathbb{R}^n \to \mathbb{R}^k \) as

\[
\gamma^1(t, x, \beta) := \max \left\{ -b(t, x, \bar{u}(t)) - \sum_{i=1}^M \beta_i \Delta b(t, x, w_i(t)), 0 \right\},
\]

\[
\gamma^2(t, x, \beta) := \min \left\{ -b(t, x, \bar{u}(t)) - \sum_{i=1}^M \beta_i \Delta b(t, x, w_i(t)), 0 \right\},
\]

where the max and min are taken componentwise. Observe that \( \gamma^1(t, x, \beta) \geq 0 \) and \( \gamma^2(t, x, \beta) \leq 0 \) for all \( (t, x, \beta) \in [a, b] \times \mathbb{R}^n \times S \). The functions \( \gamma^i(\cdot, x, \cdot) \) are \( \mathcal{L} \)-measurable, \( \gamma^i(\cdot, \cdot, \cdot) \) are Lipschitz continuous on \( (\bar{x}(t) + \epsilon'B) \times S \), for \( i = 1, 2 \), and \( \gamma^1(t, \bar{x}(t), 0) = \gamma^2(t, \bar{x}(t), 0) = 0 \) for almost every \( t \in [a, b] \).

Define the vector-valued functions \( \alpha^1, \alpha^2 : [a, b] \times \mathbb{R}^n \times S \to \mathbb{R}^k \) as

\[
\alpha^1(t, x, \beta) := B_1^{-1}(t, x)\gamma^1(t, x, \beta),
\]

\[
\alpha^2(t, x, \beta) := B_2^{-1}(t, x)\gamma^2(t, x, \beta).
\]

Because \( \gamma^1(t, x, \beta) \geq 0, B_1^{-1}(t, x) \geq 0, \gamma^2(t, x, \beta) \leq 0 \) and \( B_2^{-1}(t, x) \leq 0 \) we have, for all \( x \in \bar{x}(t) + \epsilon'B \), all \( \beta \in S \) and almost every \( t \in [a, b] \):

\[
\alpha^1(t, x, \beta) \geq 0,
\]

\[
\alpha^2(t, x, \beta) \geq 0.
\]

By definition of \( \alpha^1, \alpha^2 \) and \( \gamma^1, \gamma^2 \) we have

\[
b(t, x, \bar{u}(t)) + B_1(t, x)\alpha^1(t, x, \beta) + B_2(t, x)\alpha^2(t, x, \beta)
\]

\[
+ \sum_{i=1}^M \beta_i \Delta b(t, x, w_i(t)) \equiv 0.
\]

Finally we define the \( n \times k \) matrices:

\[
F_1(t, x) := (\Delta f(t, x, u_1(t)), \ldots, \Delta f(t, x, u_k(t))),
\]

\[
F_2(t, x) := (\Delta f(t, x, v_1(t)), \ldots, \Delta f(t, x, v_k(t))).
\]
and a function

\[ \phi(t, x, \beta) := f(t, x, \tilde{u}(t)) + F_1(t, x)\alpha^1(t, x, \beta) + F_2(t, x)\alpha^2(t, x, \beta) \]

\[ + \sum_{l=1}^{M} \beta_l \Delta f(t, x, w_l(t)). \]

Obviously \( \phi(\cdot, x, \beta) \) is \( \mathcal{L} \) measurable and \( \phi(t, \cdot, \cdot) \) is Lipschitz continuous near \((\tilde{x}(t), 0)\).

**STEP 3** For some \( \eta > 0 \), set \( S_\eta := S \cap (\eta B) \) and consider the following problem:

\[
\begin{aligned}
\text{Minimize } & h(x(a), x(b)) \\
\text{subject to } & \left\{ \begin{array}{l}
\dot{x}(t) = \phi(t, x(t), \beta(t)) \quad \text{a.e. } t \in [a, b] \\
\beta(t) \in S_\eta \quad \text{a.e. } t \in [a, b] \\
(x(a), x(b)) \in C.
\end{array} \right.
\end{aligned}
\]

Observe that \((\tilde{x}, \tilde{\beta} \equiv 0)\) is a feasible process for \((R)\). Suppose there exists another feasible process \((x, \beta)\) for \((R)\) with lesser cost, i.e. \( h(x(a), x(b)) < h(\tilde{x}(a), \tilde{x}(b)) \). By H7 and (4.12) we deduce that

\[(\tilde{x}, 0) \in \{ (f(t, x, u), b(t, x, u)) : u \in U(t) \}. \]

Measurable selection theorems assert the existence of a control function \( \tilde{u} \) such that \( \dot{x} = f(t, x, \tilde{u}) \) and \( \theta = b(t, x, \tilde{u}) \). So \( (x, \tilde{u}) \) is a feasible process for \((P)\), contradicting the optimality of \((\tilde{x}, \tilde{u})\). So \((\tilde{x}, \tilde{\beta} \equiv 0)\) is an optimal solution for \((R)\).

The Hamiltonian for \((R)\) is

\[ H_R(t, x, p, \beta) := p \cdot \phi(t, x, \beta) \]

\[ = p \cdot f(t, x, \tilde{u}) + p \cdot F_1(t, x)\alpha^1(t, x, \beta) + p \cdot F_2(t, x)\alpha^2(t, x, \beta) \]

\[ + p \cdot \sum_{l=1}^{M} \beta_l \Delta f(t, x, w_l(t)). \]

Problem \((R)\) satisfies the conditions under which Theorem 2.5 applies. This result asserts the existence of \( \lambda \geq 0, p(\cdot) \in A \subset ([a, b]; \mathbb{R}^n) \) and \( \zeta(\cdot) \in L^1([a, b]; \mathbb{R}^M) \) such that

\[ \lambda + \| p(\cdot) \|_{L^\infty} = 1, \quad (4.13) \]

\[ (-\dot{p}(t), \tilde{x}(t), \zeta(t)) \in \text{co } \partial_{x, p, \beta} H_R(t, \tilde{x}(t), p(t), 0) \quad \text{a.e.} \quad (4.14) \]

\[ \zeta(t) \in \mathcal{C} N_{S_\eta} (0) \quad \text{a.e.} \quad (4.15) \]

\[ (p(a), -p(b)) \in \mathcal{C} N_{C} (\tilde{x}(a), \tilde{x}(b)) + \lambda \partial h(\tilde{x}(a), \tilde{x}(b)). \quad (4.16) \]

We deduce from the max rule (Clarke, 1983) and a selection theorem the following estimates for the generalized Jacobians:

\[ \partial_{x, \beta} \gamma^1(t, \tilde{x}, 0) \subset \{ -\mathcal{M}(t)(\nabla b(t, \tilde{x}, \tilde{u}), \Delta b(t, \tilde{x}, w_1), \ldots, \Delta b(t, \tilde{x}, w_M) ) : \]

\[ \mathcal{M}(t) = \text{diag}(\mu_1(t), \ldots, \mu_k(t)), \quad \mu_i(t) \in [0, 1] \text{ measurable} \]
and

\[ \partial_x p^2(t, \bar{x}, 0) \subset [-\mathcal{N}(t)(\nabla_x b(t, \bar{x}, \bar{u}), \Delta b(t, \bar{x}, w_1), \ldots, \Delta b(t, \bar{x}, w_M)) : \]
\[ \mathcal{N}(t) = \text{diag}(v_1(t), \ldots, v_k(t)), \quad v_i(t) \in [0, 1] \text{ measurable}. \]

We then conclude that

\[ (-\dot{p}(t), \zeta(t)) \in \{ \text{co } \partial_x p \cdot f(t, \bar{x}(t), \bar{u}(t)) \}
\times \{ p \cdot \Delta f(t, \bar{x}(t), w_1(t)), \ldots, p \cdot \Delta f(t, \bar{x}(t), w_M(t)) \}
\times \{ q \cdot \nabla_x b(t, \bar{x}(t), \bar{u}(t)) \}
\times \{ q \cdot \Delta b(t, \bar{x}(t), w_1(t)), \ldots, q \cdot \Delta b(t, \bar{x}(t), w_M(t)) \} \] (4.17)

for almost all \( t \in [a, b] \), where (for some \( L^\infty \) functions \( M \) and \( \mathcal{N} \))

\[ q(t) := -M(t)(B_1^{-1})^T(t)F(t)p(t) - \mathcal{N}(t)(B_2^{-1})^T(t)F_2^T(t)p(t) \] (4.18)

and \( F_1, F_2, B_1^{-1} \) and \( B_2^{-1} \) are evaluated at \( (t, \bar{x}(t)) \). In view of (4.4) and \((H^*)\), we deduce the existence of an integrable function \( K_q \) such that

\[ |q(t)| \leq K_q(t)|p(t)|. \] (4.19)

If we write

\[ H(t, x, p, q, u) := p \cdot f(t, x, u) + q \cdot b(t, x, u) \]

and also note that

\[ \zeta(t) \in NS_q(0) = \{(\eta_1, \ldots, \eta_M) \in \mathbb{R}^M : \eta_i \leq 0, \quad i \in \{1, \ldots, M\} \}, \]

we obtain from (4.17) that

\[ -\dot{p}(t) \in \text{co } \partial_x H(t, \bar{x}(t), p(t), q(t), \bar{u}(t)) \] (4.20)

and

\[ p \cdot \Delta f(t, \bar{x}, w_i) + q \cdot \Delta b(t, \bar{x}, w_i) \leq 0 \quad \text{for } \quad i = 1, \ldots, M, \] (4.21)

that is,

\[ \text{Max}_{u \in U(t)} \{ H(\bar{x}(t), p(t), q(t), u) \} = H(\bar{x}(t), p(t), q(t), \bar{u}(t)) \quad \text{a.e.} \]

in which

\[ \hat{U}(t) = \bigcup_{i=1}^{M} [u_i(t)]. \]

**STEP 4** Our next task is to show that this last relationship remains true when we replace \( \hat{U}(t) \) by \( U(t) \). This is achieved by reasoning along the lines of the proof of Theorem 5.1.1 of Clarke (1983).
It is convenient to introduce the following scaled version $\tilde{q}(\cdot)$ of $q(\cdot)$:

$$\tilde{q}(t) = K_q(t)^{-1}q(t).$$

(4.22)

In view of (4.20), (4.13) and the Lipschitz continuity and differentiability hypotheses H2 and H3, we deduce that $\|p\|_{L^\infty}$ is bounded by a constant (which does not depend on our choice of $w_i$s). In view of (4.21) and (4.19) we can find a constant $L$ (again independent of the $w_i$) such that $\|r\|_{L^\infty} \leq L$, where

$$r(\cdot) := (p(\cdot), \tilde{q}(\cdot)).$$

Let $\{Z_j\}_{j=1}^\infty$ be an increasing family of finite subsets of $LB$ such that $Z_j + j^{-1}B \supset LB$ for each $j$. Fix $j$ and consider $z \in Z_j$. Appealing to an appropriate selection theorem we may find a measurable function $u_z(\cdot)$, $u_z(t) \in U(t)$ a.e., such that

$$\sup_{u \in U(t)} \{z \cdot \beta(t, u)\} \leq z \cdot \beta(t, u_z(t)) + j^{-1} \text{ a.e.}$$

(4.23)

Here

$$\beta(t, u) := (f(t, \tilde{x}(t), u), c_j(t)\bar{b}(t, \tilde{x}(t), u)).$$

For each $j$ and $t \in [a, b]$, define the measurable multifunction $U_j(t) = \{u_z(t): z \in Z_j\}$.

Let us now apply the results of the earlier analysis for the choice of ‘finitely generated’ control constraint set $\tilde{U}(t) = U_j(t)$. This tells us that there exist $p_j \in AC$, $\tilde{q}_j \in L^\infty$ and $\lambda_j \geq 0$ such that

$$-\dot{p}_j(t) \in \text{co } \partial_x p_j \cdot f + [c_j \tilde{q}_j \cdot \nabla_x b], \text{ a.e.,}$$

(4.24)

$$(p_j(a), -p_j(b)) \in N_C(\tilde{x}(a), \tilde{x}(b)) + \lambda_j \partial \bar{h}(\tilde{x}(a), \tilde{x}(b)),$$

(4.25)

$$(p_j(t), \tilde{q}_j(t)) \cdot \beta(t, u) \leq (p_j(t), \tilde{q}_j(t)) \cdot \beta(t, \tilde{u}(t)) \text{ for all } u \in U_j(t), \text{ a.e.},$$

(4.26)

$$\lambda_j + \|p_j(\cdot)\|_{L^\infty} = 1.$$

(4.27)

Now $\{\|p_j(\cdot)\|_{L^\infty}\}$, $\{\|\tilde{q}_j(\cdot)\|_{L^\infty}\}$ and $\{\lambda_j\}$ are uniformly bounded sequences and $\{\dot{p}_j\}$ is a uniformly integrably bounded sequence. Following extraction of subsequences we have then that

$$p_j(\cdot) \to p(\cdot) \text{ uniformly,} \quad \tilde{q}_j(\cdot) \to \tilde{q}(\cdot) \text{ weakly}^* \text{ in } L^\infty$$

for some $p \in AC$ and $\tilde{q} \in L^\infty$, and $\{\dot{p}_j\}$ has a weak limit in $L^1$. A straightforward modification of the proof of Clarke (1983, Theorem 3.1.7) justifies passing to the limit in (4.24) to obtain

$$-\dot{p}(t) \in \text{co } \partial_x p \cdot f + q \cdot \nabla_x b \quad \text{ a.e.,}$$

where $q := K_q \tilde{q}$ and in which the generalized gradient is evaluated at $(t, \tilde{x}(t), p(t), q(t))$. We deduce from (4.25) and the upper semicontinuity property of the limiting normal cones that

$$(p(a), -p(b)) \in N_C(\tilde{x}(a), \tilde{x}(b)) + \lambda \partial \bar{h}(\tilde{x}(a), \tilde{x}(b)).$$

Also, from (4.27),

$$\lambda + \|p(\cdot)\|_{L^\infty} = 1.$$
We turn now to the ‘maximization of the Hamiltonian’ condition. Let \( \tilde{u}(\cdot) \) be an arbitrary control function.

Fix \( j \). For almost every \( t \in [a, b] \), \( z \in Z_j \) can be chosen such that \( |(p_j(t), \tilde{q}_j(t)) - z| \leq j^{-1} \). It follows that, for almost every \( t \),

\[
(p_j(t), \tilde{q}_j(t)) \cdot \beta(t, \tilde{u}(t)) \leq z \cdot \beta(t, \tilde{u}(t)) + j^{-1}(c_f(t) + c_b c_f(t)) \\
\leq z \cdot \beta(t, u_z(t)) + j^{-1}(1 + c_f(t) + c_b c_f(t))
\]

by (4.23)

\[
\leq (p_j(t), \tilde{q}_j(t)) \cdot \beta(t, u_z(t)) + j^{-1}(1 + 2[c_f(t) + c_b c_f(t)]) \\
\leq (p_j(t), \tilde{q}_j(t)) \cdot \beta(t, \tilde{u}(t)) + j^{-1}(1 + 2[c_f(t) + c_f(t)c_b])
\]

by (4.26). We conclude that

\[
\int_a^b \{ p_j \cdot [f(\tilde{u}) - f(\tilde{\bar{u}})] + \tilde{q}_j \cdot c_f[b(\tilde{u}) - b(\tilde{\bar{u}})] \} \, dt \geq -\alpha_j,
\]

in which \( f(\tilde{u}) \) denotes \( f(t, \tilde{x}(t), \tilde{u}(t)) \), etc., and

\[
\alpha_j := j^{-1}[b - a] + 2\|c_f\|_{L^1} + 2c_g(\|c_f\|_{L^1}).
\]

Because \( p_j(\cdot) \to p(\cdot) \) uniformly, \( \tilde{q}_j(\cdot) \to \tilde{q}(\cdot) \) weakly* in \( L^\infty \) and \( q := k_\phi \tilde{q} \), we obtain in the limit

\[
\int_a^b \{ p \cdot f(\tilde{u}) + q \cdot b(\tilde{u}) \} = [p \cdot f(\tilde{u}) + q \cdot b(\tilde{u})] \} \, dt \geq 0. \tag{4.28}
\]

This inequality, which is valid for any control \( \tilde{u} \), is an ‘integral’ form of the maximization of the Hamiltonian condition.

STEP 5 A version of the maximum principle has been proved under the additional hypothesis (H*). Suppose now that (H*) is possibly not satisfied. We set up another problem for which (H*) is satisfied, which involves restricting \( U(\cdot) \) and convexifying the resulting velocity set. (This last modification is carried out to ensure continuing satisfaction of hypotheses H1–H7). The special case of the maximum principle already proved is applied to this problem, leading to the desired optimality condition for the original problem. Take

\[
\tilde{U}(t) := \{ u \in U(t): |f(t, \tilde{x}(t), u) - f(t, \tilde{x}(t), \tilde{u}(t))| \leq 1 \} \\
\text{and} \quad |b(t, \tilde{x}(t), u)| \leq 1,
\]

and consider a variant of (P) in which we replace the ‘differential-algebraic’ constraint by

\[
(\dot{x}, 0) = \left( \sum_{i=0}^{n+k} \alpha_i f(t, x, v_i), \sum_{i=0}^{n+k} \alpha_i b(t, x, v_i) \right),
\]
in which \( \alpha = (\alpha_0, \ldots, \alpha_{n+k}) \) and \( v = (v_0, \ldots, v_{n+k}) \) are regarded as control variables, subject to the constraints

\[
\alpha(t) \in P^{n+k} \quad \text{and} \quad v(t) \in \tilde{U}(t) \times \cdots \times \tilde{U}(t).
\]

Here,

\[
P^{n+k} := \left\{ (\alpha_0', \ldots, \alpha_{n+k}') : \sum_{i=0}^{n+k} \alpha_i' = 1, \alpha_i' \geq 0, \text{ for all } i \right\}.
\]

The new problem satisfies H1–H7, and also (H*). Besides,

\[
(\bar{\alpha}, \alpha(t) \equiv (1, 0, \ldots, 0), v(t) = (\bar{o}(t), \ldots, \bar{o}(t)))
\]

is a minimizer.

The preceding analysis provides elements \( p \in AC, q \in L^1 \) and \( \lambda \geq 0 \), satisfying (4.24), (4.25) and (4.27). Also, the ‘integral’ condition (4.28) is satisfied for every \( \tilde{u} \) in \( \tilde{U}(t) \). We show that this last condition implies

\[
\max_{u \in U(t)} \{ p \cdot f(u) + q \cdot b(u) \} = p(t) \cdot f(\tilde{u}(t)) + q \cdot b(\tilde{u}(t)) \quad \text{a.e.} \quad (4.29)
\]

(where \( f(u) = f(t, \bar{x}(t), u) \), etc.). If this were not true, we would be able to find a control \( u'(...) \in U(...) \) such that

\[
p(t) \cdot \left\{ f(u'(t)) - f(\tilde{u}(t)) \right\} + q(t) \cdot \left\{ b(u'(t)) - b(\tilde{u}(t)) \right\} \geq 0 \quad \text{a.e.},
\]

and the inequality would be strict on a subset of positive measure. Define the measurable function \( \gamma([a, b] \rightarrow [0, 1]) \)

\[
\gamma(t) := (1 + |f(u'(t)) - f(\tilde{u}(t))| + |b(u'(t)) - b(\tilde{u}(t))|)^{-1}.
\]

Then

\[
p(t) \cdot \left\{ f(\tilde{u}(t)) + \gamma(t)(f(u'(t)) - f(\tilde{u}(t))) \right\} + q(t) \cdot \left\{ b(\tilde{u}(t)) + \gamma(t)(b(u'(t)) - b(\tilde{u}(t))) \right\} - [p(t) \cdot f(\tilde{u}(t)) + q(t) \cdot b(\tilde{u}(t))] \geq 0,
\]

and the inequality is, once again, strict on a subset of positive measure. Appealing to the convexity of \( [f(t, x, u), b(t, x, u)]: u \in U(t) \) and a measurable selection theorem, we can find \( \tilde{u}(\cdot) \in U(\cdot) \) such that

\[
f(\tilde{u}(t)) = f(\tilde{u}(t)) + \gamma(t)(f(u'(t)) - f(\tilde{u}(t))) \quad \text{a.e.},
\]

\[
b(\tilde{u}(t)) = b(\tilde{u}(t)) + \gamma(t)(b(u'(t)) - b(\tilde{u}(t))) \quad \text{a.e.}
\]

It follows that

\[
\int_a^b \left( [p(t) \cdot f(\tilde{u}(t)) + q(t) \cdot b(\tilde{u}(t))] - [p(t) \cdot f(\tilde{u}(t)) + q(t) \cdot b(\tilde{u}(t))] \right) \, dt > 0.
\]

But for our choice of \( \gamma, \tilde{u} \) is actually a member of \( \tilde{U}(t) \), which means that (4.28) is satisfied. (4.29) must therefore be true. The theorem is proved. \( \square \)
Acknowledgement

M. d. R. d. P. was supported by F. C. T. Project 2438/95, Portugal.

REFERENCES


