ESAIM: COCV
October 2005, Vol. 11, 614–632
DOI: 10.1051/cocv:2005020

ESAIM: Control, Optimisation and Calculus of Variations

UNMAXIMIZED INCLUSION NECESSARY CONDITIONS FOR NONCONVEX CONSTRAINED OPTIMAL CONTROL PROBLEMS *

Maria do Rosário de Pinho[1], Maria Margarida Ferreira[1] and Fernando Fontes[2]

Abstract. Necessary conditions of optimality in the form of Unmaximized Inclusions (UI) are derived for optimal control problems with state constraints. The conditions presented here generalize earlier optimality conditions to problems that may be nonconvex. The derivation of UI-type conditions in the absence of the convexity assumption is of particular importance when deriving necessary conditions for constrained problems. We illustrate this feature by establishing, as an application, optimality conditions for problems that in addition to state constraints incorporate mixed state-control constraints.

Mathematics Subject Classification. 49K15.

Received January 5, 2004. Revised November 19, 2004.

Introduction

In this paper we derive necessary conditions of optimality involving unmaximized inclusion type conditions for optimal control problems with pure state constraints and we report on some main applications. These subsume and substantially extend the results in [3, 5, 6, 8].

The problem of interest is:

$$
\textbf{(P)} \quad \begin{cases} \text{Minimize } g(x(0),x(1)) + \displaystyle\int_0^1 L(t,x(t),u(t))\mathrm{d}t \\ \text{subject to} \\ \qquad \dot{x}(t) = f(t,x(t),u(t)) & a.e.\ t \in [0,1] \\ h(t,x(t)) \leq 0 & \text{for all } t \in [0,1] \\ \qquad u(t) \in U(t) & a.e.\ t \in [0,1] \\ (x(0),x(1)) \in C. \end{cases}
$$

Keywords and phrases. Optimal control, state constraints, nonsmooth analysis, Euler-Lagrange inclusion.

* *The financial support from Projecto FCT POSI EEI/14156/1998 is gratefully acknowledged. The authors wish to thank Richard B. Vinter for helpful discussions and the anonymous referee for constructive comments.*

[1] ISR and DEEC, Faculdade de Engenharia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal; mrpinho@fe.up.pt; mmf@fe.up.pt

[2] Officina Mathematica, Universidade do Minho, 4800-058 Guimarães, Portugal; ffontes@mct.uminho.pt

Here $g:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$, $L:[0,1]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$, $f:[0,1]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$, $h:[0,1]\times\mathbb{R}^n\to\mathbb{R}$ are given functions, $C\subset\mathbb{R}^n\times\mathbb{R}^n$ a given set and $U:[0,1]\rightrightarrows\mathbb{R}^m$ is a given multifunction.

First order necessary conditions for the optimal control problem (P) when the data may be nonsmooth have undergone continuous development. Such conditions, written in the form of maximum principles, are based on the generalizations of the concept of the "subdifferential" of convex functions, to larger function classes. We refer to these extended optimality conditions as the "nonsmooth" maximum principle.

For (P) the nonsmooth maximum principle (see *e.g.* [12]) asserts that for a local minimizer $(\bar{x},\bar{u})$ there exist a function $p\in W^{1,1}$, a nonnegative measure μ representing an element in C^* and a scalar $\lambda\geq 0$ such that

$$\mu\{[0,1]\}+||p||_{L_\infty}+\lambda>0, \tag{0.1}$$

$$(-\dot{p}(t),\dot{\bar{x}}(t))\in\ \text{co}\ \partial_{x,p}H(t,\bar{x}(t),\bar{u}(t),q(t),\lambda)\quad a.e. \tag{0.2}$$

$$\max_{u\in U(t)}H(t,\bar{x}(t),u,q(t),\lambda)=H(t,\bar{x}(t),\bar{u}(t),q(t),\lambda)\quad a.e. \tag{0.3}$$

$$(p(0),-q(1))\in\lambda\partial g(\bar{x}(0),\bar{x}(1))+N_C(\bar{x}(0),\bar{x}(1)) \tag{0.4}$$

$$\gamma(t)\in\bar{\partial}_x h(t,\bar{x}(t))\quad \mu-a.e. \tag{0.5}$$

$$\text{supp}\{\mu\}\subset\{t\in[0,1]:\ h(t,\bar{x}(t))=0\} \tag{0.6}$$

where H is the *Hamiltonian* for (P),

$$H(t,x,u,p,\lambda)=p\cdot f(t,x,u)-\lambda L(t,x,u) \tag{0.7}$$

and q is defined as

$$q(t)=\begin{cases}p(t)+\displaystyle\int_{[0,t)}\gamma(s)\mu(\mathrm{d}s)\ \ t\in[0,1)\\ p(t)+\displaystyle\int_{[0,1]}\gamma(s)\mu(\mathrm{d}s)\ \ t=1.\end{cases} \tag{0.8}$$

In the above conditions ∂g denotes the *limiting subdifferential* of g with respect to its arguments, N_C denotes the *limiting normal cone* to C, and co $\partial_{x,p}H$ denotes the *Clarke subdifferential* of H with respect to x and p. We also make use of the subdifferential $\bar{\partial}_x h$. These concepts from nonsmooth analysis are defined in the next section. Conditions (0.1) to (0.6) were initially obtained in [2] with the use of Clarke's subdifferential and normal cones in (0.4). The transversality conditions (derived for problems without state constraints) were refined in the present form (as shown in (0.4)) in [9] (see also [10]). The full result as stated above is given in [12].

For standard optimal control problems (when the state constraint is not present) it has been highlighted that the normal form of the nonsmooth maximum principle alluded above fails to provide sufficient conditions for linear-convex problems, in contrast to the analogous maximum principle applicable to problems with differentiable data. In [6] a weak nonsmooth maximum principle is proposed for *standard* optimal control problems which provides, in the normal form, sufficiency to linear-convex nonsmooth problems. It is formulated as *Unmaximized Inclusion* type conditions, (denoted in what follows simply as *UI* and also known in the literature as Euler Lagrange Inclusion-type conditions) involving a joint subdifferential of the Hamiltonian in the (x,p,u) variables. Noteworthy these UI-type conditions have some relevant applications. For example, consider the problem (P) when the constraint $h(t,x(t))\leq 0$ is absent and mixed constraints of the form

$$b(t,x,u)=0,\quad d(t,x,u)\leq 0$$

are added. Such are called optimal control problems with mixed constraints. The UI-type conditions for standard optimal control problems have been used as an intermediate step to establish a strong maximum principle [8] and a weak maximum principle [5] for those problems, under some regularity assumptions on the functions b and d. Of interest is the fact that the weak maximum principle for problems with mixed constraints also involves UI-type conditions [5]. Additionally, UI-type conditions for standard optimal control problems has been successfully used to derive necessary conditions for problems with differential algebraic equations (DAE's) (see [7]).

The main shortcoming of the UI-type conditions reported in [6] and its subsequent applications is the inability to deal with problems involving pure state constraints. A first attempt to extend the result in [6] to cover problem (P) was made in [3]. Here UI-type conditions are derived for (P) by assuming that the velocity set

$$\{(L(t,x,u)), f(t,x,u)),\ u \in U(t)\} \tag{0.9}$$

is convex for all $(t,x) \in [0,1] \times \mathbb{R}^n$. Although these conditions are also sufficient for linear convex problems the convexity assumption (0.9) prevents its application to problems with mixed constraints and problems with DAE's even when the convexity assumption on the velocity set is imposed on the original data. Thus, the generalization in [3] proved to be quite poor.

In this paper we prove UI-type conditions for (P) by not requiring the velocity set to be convex. This alone is a significant improvement on [3]. Moreover, in the absence of the convexity assumption our main result has an important role as an analytical tool in generalizing optimality conditions provided in [5, 7, 8] to cover problems that additionally incorporate pure state constraints. The proof of UI-type conditions we report in here requires techniques quite different and more demanding than those used in [3]. Central to the proof of our main result is the construction of a mini-max problem with convex "velocity set" in which the constraint functional

$$\max_{t\in[0,1]} h(t,x(t))$$

appears in the cost. The application of the result in [3] to this auxiliary problem and the use of relaxation techniques similar to those in the proof of Proposition 9.5.4 in [12], yield multipliers for the maximum principle for the original problem.

1. Preliminaries

Here and throughout, B represents the closed unit ball centered at the origin, $|\cdot|$ the Euclidean norm or the induced matrix norm on $\mathbb{R}^{m\times k}$. The *Euclidean distance function* with respect to $A \subset \mathbb{R}^k$ is

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

We make use of the following concepts from nonsmooth analysis. A vector $p \in \mathbb{R}^k$ is a *limiting normal to* a set A of $\mathbb{R}^k$ at a point $\tilde{x}$ in A if there exist $p_i \to p$ and $x_i \to \tilde{x}$, $x_i \in A$ and a sequence of positive scalars $\{M_i\}$, such that, for each $i \in \mathbb{N}$,

$$p_i \cdot (x - x_i) \leqslant M_i|x - x_i|^2 \qquad \text{for all } x \in A.$$

The *limiting normal cone to* A at x, written $N_A(x)$, is the set of all limiting normals to A at x. 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

$$\partial f(x) := \big\{\zeta \ :\ (\zeta, -1) \in N_{\text{epi}\,\{f\}}(x, f(x))\big\},$$

where *epi* $\{f\} = \big\{(x,\eta) \ :\ \eta \geqslant f(x)\big\}$ denotes the epigragh set.

The above concepts of limiting normal cone and limiting subdifferential were first introduced in [9]. The full calculus for these constructions in finite dimensions is described in [10, 11].

In the case that the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, co $\partial f(x)$, coincides with the *Clarke subdifferential*, which may be defined directly. Properties of Clarke subdifferentials (upper semi-continuity, sum rules, etc.), are described in [2].

We also make use of the subdifferential $\bar{\partial}_x h$ defined as

$$\bar{\partial}_x h(t,x) := \text{co}\left\{\lim \xi_i : \xi_i = \nabla_x h(t_i,x_i)\,\forall i \text{ and } (t_i,x_i) \xrightarrow{h} (t,x)\right\}. \tag{1.10}$$

For an optimal control problem, such as (P) (or for problem (Q) defined below), a (Lebesgue) measurable function $u : [0,1] \to \mathbb{R}^m$ such that $u(t) \in U(t)$ for almost every $t \in [0,1]$ is called a *control function.* An absolutely continuous function x ($x \in W^{1,1}$) satisfying $\dot{x}(t) = f(t,x(t),u(t))$ *a.e.* is called a *state trajectory* (corresponding to u). A pair (x,u) comprising a state trajectory x and a control u with which it is associated is called a *process.* A process is called *admissible* if it satisfies all problem constraints, namely end points constraint $(x(0),x(1)) \in C$, the state constraint $h(t,x(t)) \leq 0$ for all $t \in [0,1]$(and in the case of problem (Q), also the mixed state-control constraint $0 = b(t,x(t),u(t))$ *a.e.* $t \in [0,1]$).

An admissible process $(\bar{x},\bar{u})$ is a *local minimizer* for this problem if there exists a parameter $\bar{\delta} > 0$ such that

$$g(x(0),x(1)) + \int_0^1 L(t,x(t),u(t))\mathrm{d}t \geq g(\bar{x}(0),\bar{x}(1)) + \int_0^1 L(t,\bar{x}(t),\bar{u}(t))\mathrm{d}t \tag{1.11}$$

for all admissible processes $(x(t),u(t))$ which satisfy

$$| x(t) - \bar{x}(t) | \leq \bar{\delta}, \tag{1.12}$$

and it is a *weak local minimizer* if it satisfies (1.11) over admissible processes satisfying

$$| x(t) - \bar{x}(t) | \leq \bar{\delta}, \qquad | u(t) - \bar{u}(t) | \leq \bar{\delta} \quad a.e. \tag{1.13}$$

It is worth mentioning that local minimizers as defined above correspond to *s*trong local minimizers in the problem reformulation *via* differential inclusions.

Let us consider the optimal control problem (P). The following hypotheses, which make reference to a parameter $\delta > 0$ and a reference process $(\bar{x},\bar{u})$, are imposed:

H1 The function $t \to [L,f](t,x,u) := (L(t,x,u), f(t,x,u))$ is Lebesgue measurable for each pair (x,u) and there exists a function K in L^1 such that

$$\left|[L,f](t,x,u) - [L,f](t,x',u')\right| \leqslant K(t)[|x-x'|^2 + |u-u'|^2]^{1/2}$$

for $x,x' \in \bar{x}(t) + \delta B$, and $u,u' \in \bar{u}(t) + \delta B$ *a.e.* $t \in [0,1]$.

H2 The multifunction U has Borel measurable graph and

$$U_\delta(t) := (\bar{u}(t) + \delta B) \cap U(t)$$

is closed for almost all $t \in [0,1]$.

H3 The endpoint constraint set C is closed and g is locally Lipschitz in a neighbourhood of $(\bar{x}(0),\bar{x}(1))$.

H4 For $x \in \bar{x}(t) + \delta B$ the function $t \to h(t,x)$ is continuous and there exists a scalar $K_h > 0$ such that the function $x \to h(t,x)$ is Lipschitz of rank K_h for all $t \in [0,1]$.

Assume additionally that the *convexity assumption* (**CH**) holds:

CH The velocity set

$$\{(L(t,x,u), f(t,x,u)) : u \in U(t)\}$$

is convex for all $(t,x) \in [0,1] \times \mathbb{R}^n$.

The following unmaximized inclusion for optimal control problems, a simple variation of the result provided in [3], will be used.

Proposition 1.1. *Let $(\bar{x}, \bar{u})$ be a weak local minimizer for problem (P). Assume that H1–H4 and CH are satisfied. Then, there exists an absolutely continuous function $p : [0, 1] \to \mathbb{R}^n$, integrable functions $\xi : [0, 1] \to \mathbb{R}^m$ and $\gamma : [0, 1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0, 1], \mathbb{R})$, and a scalar $\lambda \geqslant 0$ such that*

$$\mu\{[0,1]\} + ||p||_{L_\infty} + \lambda > 0, \tag{1.14}$$

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) \in \operatorname{co} \partial H(t, \bar{x}(t), q(t), \bar{u}(t), \lambda) \quad a.e.\ t \in [0, 1], \tag{1.15}$$

$$\xi(t) \in \operatorname{co} N_{U(t)}(\bar{u}(t)) \quad a.e.\ t \in [0, 1], \tag{1.16}$$

$$(p(0), -q(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial g(\bar{x}(0), \bar{x}(1)), \tag{1.17}$$

$$\gamma(t) \in \bar{\partial}_x h(t, \bar{x}(t)) \quad \mu\text{-}a.e., \tag{1.18}$$

$$\operatorname{supp}\{\mu\} \subset \{t \in [0, 1] : h\,(t, \bar{x}(t)) = 0\}\,, \tag{1.19}$$

where H and q are defined as in (0.7) and (0.8).

Proof. This is simple variation of a result provided in [3] for which the endpoint constraints are separate, of the type $x(0) \in C_0$ and $x(1) \in C_1$, with $C_0, C_1 \subset \mathbb{R}^n$. Problem (P) is easily converted into one with separate endpoint constraints by considering an additional state $y \in \mathbb{R}^n$ with dynamics $\dot{y}(t) = 0$ and modifying the endpoint constraints to:

$$(x(0), y(0)) \in C,$$

$$(x(1), y(1)) \in \{(x, y) \in \mathbb{R}^{2n} : x = y\}.$$

The application of the UI conditions of [3] to the resulting problem and simple manipulations yield the desired result. □

2. Main result

In this section we state the main result of this paper, namely a weak maximum principle for problem (P) covering cases when the velocity set is nonconvex.

The following theorem mainly states that the assertions of Proposition 1.1 remain valid when the convexity assumption CH is dropped. We also sharpen the necessary conditions by replacing the subdifferential $\bar{\partial}_x h$ (as defined in (1.10)) by a possibly more refined one $\partial_x^> h$, defined as

$$\partial_x^> h(t, x) = \operatorname{co} \left\{ \xi : \exists\, (t_i, x_i) \xrightarrow{h} (t, x) : \ h(t_i, x_i) > 0\, \forall i,\ \nabla_x h(t_i, x_i) \to \xi \ \right\}. \tag{2.20}$$

Theorem 2.1. *Let $(\bar{x}, \bar{u})$ be a weak local minimizer for problem (P). Assume that, for some $\delta > 0$, H1–H4 are satisfied. Then there exist an absolutely continuous function $p : [0, 1] \to \mathbb{R}^n$, integrable functions $\xi : [0, 1] \to \mathbb{R}^m$ and $\gamma : [0, 1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0, 1], \mathbb{R})$, and a scalar $\lambda \geqslant 0$ such that conditions (1.14)–(1.19) of Proposition 1.1 hold with $\partial_x^> h(t, x)$ replacing $\bar{\partial}_x h(t, \bar{x}(t))$.*

3. Applications

The removal of the convexity assumption on (0.9) is of particular importance when deriving necessary conditions for problems that in addition to state constraints incorporate mixed state-control constraints.

As mentioned in the introduction, the UI-type conditions for standard optimal problems proved in [6] were successfully applied to derive necessary conditions for constrained optimal control problems under certain regularity assumptions. Noteworthy, the weak maximum principle involving UI-type conditions in [6] was used to derive not only UI-type conditions for other problems but also *strong* forms of the maximum principle (applied to local minimizers). For optimal control problems involving differential algebraic equations (DAE's) of index one (see [1] for the definition of index for DAE's) a strong maximum principle and a weak maximum principle involving UI-type conditions are proved in [7]. For optimal control with mixed constraints a strong maximum principle is derived in [8] assuming a certain "interiority " assumption on the mixed constraints and, in [5], UI-type conditions are validated assuming the full rankness of the derivatives with respect to the control of the function defining the mixed constraints. In [5,8] some differentiability with respect to the state and/or the control variable of the functions defining the mixed constraints is imposed. As mentioned in [8] optimal control problems involving DAE's of index higher than one can sometimes be reformulated as problems with mixed constraints. Systems involving DAE's are of interest since there are many applications for such dynamic models in process systems engineering, robotics, etc. See [1,7] in that respect.

The derivation of the results in [5,7,8] follows a simple and common approach. Different regularity assumptions imposed on the data of the problem under consideration permit the association of an auxiliary problem with the original problem. In all the three cases the auxiliary problem is a standard optimal control problem to which the UI-type conditions obtained in [6] apply.

It is then reasonable to hope that UI-type conditions for problem (P) can also be successfully applied to generalize the necessary conditions of optimality in [5,7,8] to cover problems involving additionally pure state constraints. Following the approach used in [5,7,8] the idea would be of associating with the original problem an auxiliary problem of the form

$$(P_{\text{aux}}) \quad \begin{cases} \text{Minimize} & g(x(0),x(1)) + \displaystyle\int_0^1 \tilde{L}(t,x(t),w(t))\mathrm{d}t \\ \text{subject to} & \\ \quad \dot{x}(t) = \phi(t,x(t),w(t)) & a.e. \\ \quad 0 \geq h(t,x(t)) & \text{for all } t \\ \quad w(t) \in W(t) & a.e. \\ \quad (x(0),x(1)) \in C, & \end{cases}$$

where w is the new control variable taking values in the control set $W(t)$.

Unfortunately it turns out that, even when convexity is imposed on the original data, Proposition 1.1 cannot be applied to (P_{aux}). This is due to the fact that the velocity set of (P_{aux})

$$\{(\tilde{L}(t,x(t),w(t)),\phi(t,x(t),w(t))) : w \in W(t)\}$$

fails to be convex. Theorem 2.1, by validating the conclusions of Proposition 1.1 *in the absence of convexity assumption* is then the true successor of the result in [6], since it can be used as an intermediate step to derive necessary conditions of optimality for problems involving not only mixed constraints or DAE's but also pure state constraints.

In this section we illustrate this feature by deriving a maximum principle for a problem with both state constraints and mixed constraints. The problem of interest is:

$$\textbf{(Q)} \quad \begin{cases} \text{Minimize } g(x(0),x(1)) + \displaystyle\int_0^1 L(t,x(t),u(t))\mathrm{d}t \\ \text{subject to} \\ \quad \dot{x}(t) = f(t,x(t),u(t)) & a.e.\ t \in [0,1] \\ \quad 0 = b(t,x(t),u(t)) & a.e.\ t \in [0,1] \\ \quad h(t,x(t)) \leq 0 & \text{for all } t \in [0,1] \\ \quad u(t) \in U(t) & a.e.\ t \in [0,1] \\ \quad (x(0),x(1)) \in C, & \end{cases}$$

in which $b : [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$, and $m \geq k$.

A weak maximum principle for a particular case of problem (Q) in the form of UI-type conditions is reported in [4]. A set of hypotheses comprising a full rank condition permits the construction of an auxiliary problem to which Theorem 2.1 applies.

Next we derive a strong maximum principle for (Q) (not involving UI-type conditions) using Theorem 2.1 as an intermediate step. In this respect a new set of assumptions, which make reference to a parameter $\delta > 0$ and a reference process $(\bar{x}, \bar{u})$, are imposed on the data of (Q):

HQ1 $\big(L(\cdot, x, \cdot), f(\cdot, x, \cdot), b(\cdot, x, \cdot)\big)$ are $\mathcal{L} \times \mathcal{B}$ measurable for each x.

HQ2 There exists an integrable function $K_{f,L}$ such that, for almost every $t \in [0,1]$, $(L(t, \cdot, u), f(t, \cdot, u))$ is Lipschitz continuous on $\bar{x}(t) + \delta B$ for $u \in U(t)$ with Lipschitz constant $K_{f,L}(t)$.

HQ3 $b(t, \cdot, u)$ is a continuously differentiable function with Lipschitz constant K_b on $\bar{x}(t) + \delta B$ for all $u \in U(t)$ and almost all $t \in [0,1]$.

HQ4 There exists a $\eta > 0$ such that

$$\eta B \subset b(t, \bar{x}(t), U(t)) \quad \text{for } a.e. \ t \in [0,1].$$

HQ5 Graph of U is a Borel measurable set.

In addition we assume that

HQ6 The set

$$\{(\ell, v) : \ell \geqslant L(t, x, u), v = (f(t, x, u), b(t, x, u)), u \in U(t)\}$$

is **convex** for each $t \in [0,1]$.

Observe that HQ1, HQ2 and HQ5 are less restrictive hypotheses than their counterparts H1, H2. HQ3 is relevant for the definition of the multiplier ρ associated with b. A special feature of the above hypotheses is the *interiority* assumption HQ4. For details we refer the reader to [8]. Also HQ6 is essential as illustrated through example in [8].

It is of interest to note that Theorem 2.1 holds for weak local minimizers but the following theorem is a necessary condition in the form of strong maximum principle which holds for local minimizers.

Theorem 3.1 (maximum principle for problem (Q))**.** *Let $(\bar{x}, \bar{u})$ be a local minimizer for (Q). Assume that, for some $\delta > 0$, hypotheses HQ1-HQ6, H3 and H4 are satisfied. Let the Hamiltonian be*

$$H_Q(t, x, p, \rho, u, \lambda) \ = p \cdot f(t, x, u) + \rho \cdot b(t, x, u) - \lambda L(t, x, u).$$

Then there exist $p \in W^{1,1}([0,1]; \mathbb{R}^n)$, $\rho \in L^1([0,1]; \mathbb{R}^k)$ $\gamma \in L^1([0,1]; \mathbb{R}^n)$, a nonnegative Radon measure $\mu \in C^*([0,1], \mathbb{R})$, and $\lambda \geq 0$ such that, for almost every $t \in [0,1]$,

(i) $\mu\{[0,1]\} + ||p||_{L_\infty} + \lambda > 0,$

(ii) $-\dot{p}(t) \in \operatorname{co} \partial_x H_Q(t, \bar{x}(t), q(t), \rho(t), \bar{u}(t), \lambda), \quad a.e.$

(iii) $H_Q(t, \bar{x}(t), q(t), \rho(t), \bar{u}(t), \lambda) = \max_{u \in U(t)} H_Q(t, \bar{x}(t), q(t), \rho(t), u, \lambda), \quad a.e.$

(iv) $(p(0), -q(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \, \partial g(\bar{x}(0), \bar{x}(1)),$

(v) $\gamma(t) \in \partial_x^{>} h(t, \bar{x}(t))$ *μ-a.e.,*

(vi) $\operatorname{supp}\{\mu\} \subset \{t \in [0,1] : h(t, \bar{x}(t)) = 0\},$

where q is as in (0.8).

As in [8] this theorem can be easily generalized to cover inequality mixed constraints.

Proof. Here we give an outline of the proof. Many steps are omitted since they consist on repeating the same arguments used in [8]. Mainly the definition of the auxiliary problem (R) below follows *exactly* the steps in [8]. The pure state constraint $0 \geq h(t, x(t))$ does not play any role in Step 0 and Step 1 below.

We prove the theorem when $L \equiv 0$. This restriction can be removed by state augmentation techniques. Details are omitted.

Step 1. We verify the conclusions of the Theorem under the hypothesis that

H* There exists a $c_f \in L^1$ and a scalar c_b such that

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

for all $u \in U(t)$ and almost every $t \in [0, 1]$.

This condition can be removed by reasoning along the lines of the proof of Theorem 5.1.2. of [2].

Step 2. The *interiority* condition HQ4 and a measurable selection theorem allow us to choose control functions $\{u_1, \dots, u_k\}$ and $\{v_1, \dots, v_k\}$ such that the matrices

$$B_1(t, x) := (\Delta b(t, x, u_1(t)), \dots, \Delta b(t, x, u_k(t))),$$
$$B_2(t, x) := (\Delta b(t, x, v_1(t)), \dots, \Delta b(t, x, v_k(t))),$$

where $\Delta b(t, x, u) = b(t, x, u) - b(t, x, \bar{u}(t))$, satisfy, for all $x \in \bar{x}(t) + \varepsilon' B$ (for some ε') and almost every $t \in [0, 1]$,

$$B_1^{-1}(t, x) \geq 0, \qquad B_2^{-1}(t, x) \leq 0.$$

Furthermore there exists a constant m_b such that

$$| B_1^{-1}(t, \bar{x}(t)) | \leq m_b, \quad | B_2^{-1}(t, \bar{x}(t)) | \leq m_b.$$

Next a finite collection of control functions $\{w_i\}_{i=1}^M$ (which includes $\bar{u}$) is chosen. Let β be a measurable function such that $\beta(t) \in S$ where

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

Set $\gamma^1, \gamma^2, \alpha^1, \alpha^2 : [0, 1] \times \mathbb{R}^n \times S \to \mathbb{R}^k$ as

$$\gamma^1(t, x, \beta) := \max \left\{ -b(t, x, \bar{u}(t)) - \sum_{l=1}^M \beta_l \Delta b(t, x, w_l(t)), 0 \right\},$$
$$\gamma^2(t, x, \beta) := \min \left\{ -b(t, x, \bar{u}(t)) - \sum_{l=1}^M \beta_l \Delta b(t, x, w_l(t)), 0 \right\},$$

where the max and min are taken componentwise, and

$$\alpha^1(t, x, \beta) = B_1^{-1}(t, x)\gamma^1(t, x, \beta), \quad \alpha^2(t, x, \beta) = B_2^{-1}(t, x)\gamma^2(t, x, \beta).$$

One may prove that $\alpha^1(t, x, \beta) \geq 0$, $\alpha^2(t, x, \beta) \geq 0$ and

$$b(t, x, \bar{u}(t)) + B_1(t, x)\alpha^1(t, x, \beta) + B_2(t, x)\alpha^2(t, x, \beta) + \sum_{l=1}^M \beta_l \Delta b(t, x, w_l(t)) = 0 \tag{3.21}$$

where $(t, x, \beta) \in [0,1] \times \mathbb{R}^n \times S$. Finally define the matrices

$$F_1(t,x) := (\Delta f(t,x,u_1(t)), \dots \Delta f(t,x,u_k(t))),$$

$$F_2(t,x) := (\Delta f(t,x,v_1(t)), \dots, \Delta f(t,x,v_k(t))),$$

and a function

$$\phi(t,x,\beta) = f(t,x,\bar{u}(t)) + F_1(t,x)\alpha^1(t,x,\beta) + F_2(t,x)\alpha^2(t,x,\beta) + \sum_{l=1}^{M} \beta_l \Delta f(t,x,w_l(t)).$$

It turns out that $\phi(\cdot,x,\beta)$ is measurable and that $\phi(t,\cdot,\cdot)$ is Lipschitz continuous near $(\bar{x}(t),0)$.

Step 3. For some $\zeta > 0$ consider the following problem

$$(R) \quad \begin{cases} \text{Minimize} & g(x(0),x(1)) \\ \text{subject to} & \\ \quad \dot{x}(t) = \phi(t,x(t),\beta(t)) & a.e. \\ \quad h(t,x(t)) \leq 0 & \text{for all } t \\ \quad \beta(t) \in S_\zeta & a.e. \\ \quad (x(0),x(1)) \in C & \end{cases}$$

where $S_\zeta = S \cap (\zeta B)$. The convexity assumption HQ6 together with (3.21) guarantee that $(\bar{x}, \bar{\beta} \equiv 0)$ is a weak local minimizer of (R) for a sufficiently small ζ (see [8]).

The Hamiltonian for (R) is $H_R(t,x,p,\beta) = p \cdot \phi(t,x,\beta)$. Problem (R) satisfies the conditions under which Theorem 2.1 is applicable. It asserts the existence of $\lambda \geq 0$, an absolutely continuous function $p : [0,1] \to \mathbb{R}^n$, integrable functions $\pi : [0,1] \to \mathbb{R}^k$ and $\gamma : [0,1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0,1];\mathbb{R})$ such that

$$\mu\{[0,1]\} + ||p||_{L_\infty} + \lambda = 1, \tag{3.22}$$

$$(-\dot{p}(t), \dot{\bar{x}}(t), \pi(t)) \in \operatorname{co} \partial H_R(t, \bar{x}(t), q(t), 0) \quad a.e., \tag{3.23}$$

$$\pi(t) \in \operatorname{co} N_{S_\zeta}(0) \quad a.e.\ t \in [0,1], \tag{3.24}$$

$$(p(0), -q(1)) \in N_C(\bar{x}(0), \bar{x}(1)) + \lambda \partial g(\bar{x}(0), \bar{x}(1)), \tag{3.25}$$

$$\gamma(t) \in \partial_x^> h(t, \bar{x}(t)) \quad \mu\text{-}a.e., \tag{3.26}$$

$$\operatorname{supp}\{\mu\} \subset \{t \in [0,1] : h(t, \bar{x}(t)) = 0\}, \tag{3.27}$$

where q is defined as in (0.8).

Step 4. Estimating the subdifferentials of ϕ we may deduce that for almost every $t \in [0,1]$

$$\begin{aligned}(-\dot{p}(t), \pi(t)) \in\ & \{\operatorname{co}\partial_x\, q \cdot f(t, \bar{x}(t), \bar{u}(t))\} \\ & \times\{q \cdot \Delta f(t, \bar{x}(t), w_1(t)), \dots, q \cdot \Delta f(t, \bar{x}(t), w_M(t))\} \\ & +\{\rho \cdot \nabla_x b(t, \bar{x}(t), \bar{u}(t))\} \\ & \times\{\rho \cdot \Delta b(t, \bar{x}(t), w_1(t)), \dots, \rho \cdot \Delta b(t, \bar{x}(t), w_M(t))\}\end{aligned} \tag{3.28}$$

where, for some L^∞ functions $\mathcal{M}$ and $\mathcal{N}$,

$$\rho(t) := -\mathcal{M}(t)\big(B_1^{-1}\big)^T(t) F_1^T(t) q(t) - \mathcal{N}(t)\big(B_2^{-1}\big)^T(t) F_2^T(t) q(t)$$

and F_1, F_2, B_1^{-1} and B_2^{-1} are evaluated at $(t, \bar{x}(t))$. Introducing the Hamiltonian

$$H(t, x, p, \rho, u) := p \cdot f(t, x, u) + \rho \cdot b(t, x, u)$$

and also noting that $N_{S_\varsigma}(0) = \{(\rho_1, \ldots, \rho_M) : \rho_i \leq 0\}$ we obtain from (3.28) that

$$-\dot{p}(t) \in \operatorname{co} \partial_x H(t, \bar{x}(t), q(t), \rho(t), \bar{u}(t)) \tag{3.29}$$

and

$$q \cdot \Delta f(t, \bar{x}, w_i) + \rho \Delta b(t, \bar{x}, w_i) \leq 0$$

for $i = 1, \ldots, M$, that is,

$$Max_{u \in \hat{U}(t)} \{H(t, \bar{x}(t), q(t), \rho(t), u)\} = H(t, \bar{x}(t), q(t), \rho(t), \bar{u}(t)) \ a.e.$$

in which $\hat{U}(t) = \bigcup_i^M \{w_i(t)\}$.

Step 5. The final step of the proof is to show that this last relationship remains true when we replace $\hat{U}(t)$ by $U(t)$ and the removal of the interim hypothesis H*. Details can be found in [2, 8]. □

4. Proof of the Theorem 3.1

Once again, we prove the theorem when $L \equiv 0$. This restriction can be removed by state augmentation techniques. Details are omitted. The Hamiltonian considered is then simply

$$H(t, x, p, u) := p \cdot f(t, x, u).$$

We start by noting some additional properties that the data of problem (P) satisfy and that will be of use later.

Let $\varepsilon = \min\{\bar{\delta}, \delta\}$, where $\bar{\delta} > 0$ is as in (1.13). By H1 and H2 the following conditions are satisfied:

AH1: there exists an integrable function ϕ such that for almost all $t \in [0, 1]$

$$| f(t, x, u) | \leq \phi(t)$$

for all $(x, u) \in (\bar{x}(t), \bar{u}(t)) + \varepsilon B$.

AH2: the set $f(t, x, U_\varepsilon(t))$ is compact for all $x \in \bar{x}(t) + \varepsilon B$.

Indeed, ϕ, as defined in AH1, can be chosen to be $K(t)\varepsilon + | \dot{\bar{x}}(t) |$.

The proof breaks into steps. First, we derive necessary conditions of optimality for the following "minimax" optimal control problem in which the constraint functional $\max_{t \in [0,1]} h(t, x(t))$ appears in the cost:

$$(R) \begin{cases} \text{Minimize } \tilde{g}(x(0), x(1), \max_{t \in [0,1]} h(t, x(t))) \\ \text{over } x \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \dot{x}(t) \ = \ f(t, x(t), u(t)) \quad a.e., \\ u(t) \ \in \ U_\varepsilon(t) \quad a.e., \\ (x(0), x(1)) \ \in \ C_0 \times \mathbb{R}^n, \end{cases}$$

in which (besides the previously specified data), $C_0 \subset \mathbb{R}^n$ is a given closed set and $\tilde{g} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is a given function.

Proposition 4.1. *Let $(\bar{x}, \bar{u})$ be a weak local minimizer for (R). Assume that the data for problem (R) satisfies hypotheses H1 – H4 and G below:*

G: *$\tilde{g}$ is Lipschitz continuous on a neighbourhood of*

$$(\bar{x}(0), \bar{x}(1), \max_{t\in[0,1]} h(t, \bar{x}(t)))$$

and $\tilde{g}$ is monotone in the z variable, in the sense that

$$z' \geq z \qquad \textit{implies} \qquad \tilde{g}(y, x, z') \geq \tilde{g}(y, x, z),$$

for each $(y, x) \in \mathbb{R}^n \times \mathbb{R}^n$.

Then there exists an absolutely continuous function $p : [0,1] \to \mathbb{R}^n$, integrable functions $\xi : [0,1] \to \mathbb{R}^m$ and $\gamma : [0,1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0,1], \mathbb{R})$ such that

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) \in \operatorname{co} \partial H(t, \bar{x}(t), q(t), \bar{u}(t)) \quad a.e., \tag{4.30}$$

$$\xi(t) \in \operatorname{co} N_{U(t)}(\bar{u}(t)) \quad a.e.\ t \in [0,1], \tag{4.31}$$

$$\left(p(0), -q(1), \int_{[0,1]} \mu(\mathrm{d}s)\right) \in \partial\tilde{g}(\bar{x}(0), \bar{x}(1), \max_{t\in[0,1]} h(t, \bar{x}(t))) + N_{C_0}(\bar{x}(0)) \times \{0, 0\} \tag{4.32}$$

$$\operatorname{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = \max_{s\in[0,1]} h(s, \bar{x}(s))\}, \tag{4.33}$$

$$\gamma(t) \in \bar{\partial}_x h(t, \bar{x}(t)) \quad \mu\text{- } a.e. \tag{4.34}$$

Here

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} \gamma(s)\mu(\mathrm{d}s) & \textit{if } t \in [0,1) \\ p(1) + \int_{[0,1]} \gamma(s)\mu(\mathrm{d}s) & \textit{if } t = 1. \end{cases}$$

Proof. By adjusting $\varepsilon > 0$ we can arrange that $(\bar{x}, \bar{u})$ is minimizing with respect to processes (x, u) satisfying the constraints of (R) and also $\|x - \bar{x}\|_{L^\infty} \leq \varepsilon$ and $\|u - \bar{u}\|_{L^\infty} \leq \varepsilon$. Define

$$Q := \{x \in W^{1,1} : x(0) \in C_0,\ \dot{x}(t) \in f(t, x(t), U_\varepsilon(t))\}.$$

By the Generalized Filippov Selection Theorem (see *e.g.* [12], Th. 2.3.13), $\bar{x}$ is a minimizer for the problem

$$\begin{cases} \text{Minimize } \tilde{g}(x(0), x(1), \max_{t\in[0,1]} h(t, x(t))) \\ \text{over arcs } x \in Q \text{ satisfying } \|x - \bar{x}\|_{L^\infty} < \varepsilon. \end{cases}$$

In view of the Relaxation Theorem (see *e.g.* [12], Th. 2.7.2), any arc x in the set

$$Q_r := \{x \in W^{1,1} : x(0) \in C_0, \dot{x}(t) \in \operatorname{co} f(t, x(t), U_\varepsilon(t))\}$$

which satisfies $||x - \bar{x}||_{L^\infty} < \varepsilon$ can be approximated by an arc y in Q satisfying $\|y - \bar{x}\|_{L^\infty} < \varepsilon$. The continuity of the mapping

$$x \to \tilde{g}(x(0), x(1), \max_{t\in[0,1]} h(t, x(t)))$$

on a neighbourhood of $\bar{x}$ (with respect to the supremum norm topology) asserts that $\bar{x}$ is a minimizer for the optimization problem

$$\begin{cases} \text{Minimize } \tilde{g}(x(0), x(1), \max_{t\in[0,1]} h(t, x(t))) \\ \text{over } x \in Q_r \text{ and } \|x - \bar{x}\|_{L^\infty} < \varepsilon. \end{cases}$$

Set $\bar{z} := \max_{t\in[0,1]} h(t,\bar{x}(t))$. By the Generalized Filippov Selection Theorem and Carathéodory's Theorem, and in view of the monotonicity property of $\tilde{g}$,

$$\{\bar{x}, \bar{y} \equiv \tilde{g}(\bar{x}(0),\bar{x}(1),\bar{z}), \bar{z}, (\bar{u}_0,\ldots,\bar{u}_n) \equiv (\bar{u},\ldots,\bar{u}), (\lambda_0,\lambda_1,\ldots,\lambda_n) \equiv (1,0,\ldots,0)\}$$

is a local minimizer for the optimization problem

$$\begin{cases} \text{Minimize } y(1) \\ \text{over } x \in W^{1,1},\ y \in W^{1,1},\ z \in W^{1,1} \\ \text{and measurable functions } u_0,\ldots,u_n,\ \lambda_0,\ldots,\lambda_n \text{ satisfying} \\ \dot{x}(t) \ = \ \sum_i \lambda_i(t) f(t,x(t),u_i(t)),\ \dot{y}(t) \ = \ 0,\ \dot{z}(t) \ = \ 0 \quad a.e., \\ (\lambda_0(t),\ldots,\lambda_n(t)) \in \Lambda,\ u_i(t) \in U_\varepsilon(t),\ i=0,\ldots,n \quad a.e., \\ h(t,x(t)) - z(t) \ \le \ 0 \quad \text{for all } t\in[0,1], \\ (x(0),x(1),z(0),y(0)) \ \in \ \text{epi}\{\tilde{g} + \Psi_{C_0\times\mathbb{R}^n\times\mathbb{R}}\}. \end{cases}$$

Here

$$\Lambda \ := \ \{\lambda'_0,\ldots,\lambda'_n : \lambda'_i \ge 0 \text{ for } i=0,\ldots,n \text{ and } \sum_i \lambda'_i = 1\},$$

Ψ_A is the indicator function of a set A and $(\lambda_0,\ldots,\lambda_n)$, $(u_0,\ldots,u_n)$ are regarded as control variables. Since the velocity set is convex, this is a problem to which Proposition 1.1 is applicable. We deduce existence of absolutely continuous functions $p_1 : [0,1] \to \mathbb{R}^n$, $p_2 : [0,1] \to \mathbb{R}$, $p_3 : [0,1] \to \mathbb{R}$, integrable functions $\xi_j : [0,1] \to \mathbb{R}^m$, $j=0,1,...,n$, $\eta : [0,1] \to \mathbb{R}^{n+1}$, and $\gamma : [0,1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0,1],\mathbb{R})$, such that

$$\mu\{[0,1]\} + ||(p_1,p_2,p_3)||_{L_\infty} + \lambda > 0, \tag{4.35}$$

$$(-\dot{p}_1(t), -\dot{p}_2(t), -\dot{p}_3(t), \dot{\bar{x}}(t), \dot{\bar{y}}(t), \dot{\bar{z}}(t), \xi_0(t),\ldots,\xi_n(t)), \eta(t)) \in \tag{4.36}$$
$$\operatorname{co}\partial\tilde{H}(t,\bar{x}(t),\bar{y}(t),\bar{z}(t),q_1(t),p_2(t),q_3(t),(\bar{u}(t),\ldots,\bar{u}(t)),(1,0,\ldots,0)) \quad a.e.$$

$$\xi_i(t) \in \operatorname{co} N_{U_\varepsilon(t)}(\bar{u}(t)) \quad a.e.\ t\in[0,1], \quad i\in\{0,\ldots,n\} \tag{4.37}$$

$$\eta(t) \in \operatorname{co} N_\Lambda(1,0,\ldots,0) \quad a.e.\ t\in[0,1], \tag{4.38}$$

$$(p_1(0), -q_1(1), p_3(0), p_2(0)) \ \in \ \partial\Psi_{\text{epi}\{\tilde{g}+\Psi_{C_0\times\mathbb{R}^n\times\mathbb{R}}\}}(\bar{x}(0),\bar{x}(1),\bar{z},\bar{y}), \tag{4.39}$$

$$-p_2(1) = \lambda, \tag{4.40}$$

$$-q_3(1) = 0, \tag{4.41}$$

$$\text{supp}\{\mu\} \subset \{t : h(t,\bar{x}(t)) = \bar{z}\},$$
$$\gamma(t) \ \in \ \bar{\partial}_x h(t,\bar{x}(t)) \quad \mu\text{ - } a.e., \tag{4.42}$$

where

$$\tilde{H}(t,x,y,z,p_1,p_2,p_3,(u_0,\ldots,u_n),(\lambda_1,\ldots,\lambda_n)) = (p_1,p_2,p_3)\cdot\left(\sum_i \lambda_i(t) f(t,x,u_i),0,0\right),$$

$$q_1(t) = \begin{cases} p_1(t) + \int_{[0,t)} \gamma(s)\mu(\mathrm{d}s) \ t\in[0,1) \\ p_1(t) + \int_{[0,1]} \gamma(s)\mu(\mathrm{d}s) \ t = 1, \end{cases}$$

and

$$q_3(t) = \begin{cases} p_3(t) - \int_{[0,t)} \mu(\mathrm{d}s) \;\; t \in [0,1) \\ p_3(t) - \int_{[0,1]} \mu(\mathrm{d}s) \;\; t = 1. \end{cases}$$

Since $\tilde{H}$ is independent of y and z we have

$$-\dot{p}_2 = 0 \text{ and } -\dot{p}_3 = 0.$$

Estimating the Clarke's subdifferential of

$$\hat{H}(t, x, p_1, (u_0, \ldots, u_n), (\lambda_1, \ldots, \lambda_n)) = p_1 \cdot \sum_i \lambda_i f(t, x, u_i),$$

at the point $(\bar{x}(t), q_1(t), (\bar{u}(t), \ldots, \bar{u}(t)), (1, 0, \ldots, 0))$, and appealing to the sum and product rules (Ths. 5.4.1 and 5.4.2 of [12]) we are led to

$$\operatorname{co} \partial \hat{H} \subset \sum_{i=0}^{n} \operatorname{co} \partial_{x, p_1, (u_0, \ldots, u_n), (\lambda_0, \ldots, \lambda_n)} [\lambda_i \, p_1 \cdot f(t, x, u_i)].$$

For each $i \in \{0, \ldots, n\}$, the product rule (Th. 5.4.2 in [12]) allows us to deduce that

$$\operatorname{co} \partial_{x, p_1, (u_0, \ldots, u_n), (\lambda_0, \ldots, \lambda_n)} [\lambda_i \, p_1 \cdot f(t, x, u_i)] \subset \Big\{ \big(\lambda_i \mu_i, \lambda_i f(t, x, u_i), (0, \ldots, \overbrace{\lambda_i \nu_i}^{i\text{th position}}, \ldots, 0), (0, \ldots, \overbrace{p_1 \cdot f(t, x, u_i)}^{i\text{th position}}, \ldots, 0)\big) : (\mu_i, \nu_i) \in \operatorname{co} \partial_{x,u} \, p_1 \cdot f(t, x, u_i) \Big\}.$$

Relationship (4.36) therefore yields

$$\begin{aligned} &(-\dot{p}_1(t), \dot{\bar{x}}(t), \xi_0(t)) \in \operatorname{co} \partial_{x, p_1, u} [q_1(t) \cdot f(t, \bar{x}(t), \bar{u}(t))], \\ &\eta(t) = \big(q_1(t) \cdot f(t, \bar{x}(t), \bar{u}(t)), \ldots, q_1(t) \cdot f(t, \bar{x}(t), \bar{u}(t))\big), \\ &\xi_1(t) \equiv \ldots \equiv \xi_n(t) \equiv 0. \end{aligned} \tag{4.43}$$

Observe that the multiplier η defined above satisfies (4.38).

Conditions (4.39) to (4.41) imply (using the definition of limiting subdifferential)

$$(p_1(0), -q_1(1), \int_{[0,1]} \mu(\mathrm{d}s)) \;\in\; \lambda \partial \tilde{g}(\bar{x}(0), \bar{x}(1), \bar{z}) + N_{C_0}(\bar{x}(0)) \times \{(0,0)\}. \tag{4.44}$$

From the above relationships, we deduce that $(p_1, \mu, \lambda) \neq 0$. But in fact $\lambda > 0$. This is because, if $\lambda = 0$, then by (4.44) $\mu = 0$ and by (4.40)–(4.41) $p_2 = p_3 = 0$. Now, again by (4.44) $p_1(1) = 0$ and by (4.43) $p_1 \equiv 0$, an impossibility. By scaling the multipliers we can arrange that $\lambda = 1$. Finally we can identify p_1 and ξ_0 with the p and ξ from the proposition we want to prove. Reviewing our findings, we see that all the assertions of the proposition have been proved. □

The next stage is to get the necessary conditions as mentioned in Theorem 2.1. We use Proposition 4.1 but we allow a general end-point constraint. We shall also sharpen the necessary conditions, replacing the subdifferential $\bar{\partial}_x h$ by the more refined subdifferential $\partial_x^{>} h$.

Consider the set

$$W := \{(x,u,e) : (x,u) \text{ satisfies } \dot{x}(t) = f(t,x(t),u(t)), \tag{4.45}$$
$$u(t) \in U_\varepsilon(t) \ a.e.,\ e \in \mathbb{R}^n,\ (x(0),e) \in C \text{ and } \|x-\bar{x}\|_{L^\infty} \le \varepsilon\} \tag{4.46}$$

and define $d_W : W \times W \to \mathbb{R}$

$$d_W((x,u,e),(x',u',e')) \;=\; |x(0)-x'(0)| + |e-e'| + \int_0^1 |u(t)-u'(t)|\,\mathrm{d}t.$$

Choose $\varepsilon_i \downarrow 0$ and, for each i, define the function

$$\tilde{g}_i(x,y,x',y',z) \;:=\; \max\{g(x,y) - g(\bar{x}(0),\bar{x}(1)) + \varepsilon_i^2, z, |x'-y'|\}.$$

Now consider the optimization problem

$$\text{Minimize } \{\tilde{g}_i(x(0),e,x(1),e, \max_{t\in[0,1]} h(t,x(t))) : (x,u,e) \in W\}.$$

We omit the straightforward verification of the following properties of the metric space (W, d_W):

(i) d_W defines a metric on W, and (W, d_W) is a complete metric space;
(ii) if $(x_i,u_i,e_i) \to (x,u,e)$ in (W,d_W) then $\|x_i - x\|_{L^\infty} \to 0$;
(iii) the function
$$(x,u,e) \to \tilde{g}_i(x(0),e,x(1),e, \max_{t\in[0,1]} h(t,x(t)))$$
is continuous on (W, d_W).

Notice that
$$\tilde{g}_i(\bar{x}(0),\bar{x}(1),\bar{x}(1),\bar{x}(1), \max_{t\in[0,1]} h(t,\bar{x}(t))) \;=\; \varepsilon_i^2.$$

Since $\tilde{g}_i$ is non-negative valued it follows that $(\bar{x},\bar{u},\bar{x}(1))$ is an "ε_i^2-minimizer" for the above minimization problem.

According to Ekeland's Theorem, there exists a sequence $\{(x_i,u_i,e_i)\}$ in W such that for each i,

$$\tilde{g}_i(x_i(0),e_i,x_i(1),e_i, \max_{t\in[0,1]} h(t,x_i(t))) \le g_i(x(0),e,x(1),e, \max_{t\in[0,1]} h(t,x(t))) + \varepsilon_i d_W((x,u,e),(x_i,u_i,e_i)) \tag{4.47}$$

for all $(x,u,e) \in W$ and also
$$d_W((x_i,u_i,e_i),(\bar{x},\bar{u},\bar{x}(1))) \;\le\; \varepsilon_i. \tag{4.48}$$
The condition (4.48) implies that $e_i \to \bar{x}(1)$ and u_i converges to $\bar{u}$ in the L^1 norm. We can arrange by subsequence extraction that u_i converges to $\bar{u}$ almost everywhere. From the properties of f it follows that $x_i \to \bar{x}$ uniformly.

Define the arc $y_i \equiv e_i$. Then $y_i \to \bar{x}(1)$ uniformly.

We can express the minimization property (4.47) as follows: $(x_i, y_i, w_i \equiv 0, u_i)$ is a local minimizer for the optimal control problem

$$(\mathbf{R}_i) \quad \begin{cases} \text{Minimize } \tilde{g}_i(x(0),y(0),x(1),y(1),\max_{t\in[0,1]} h(t,x(t))) \\ \qquad +\varepsilon_i[|x(0)-x_i(0)| + |y(0)-y_i(0)| + w(1)] \\ \text{over } x,y,w \in W^{1,1} \text{ and measurable functions } u \text{ satisfying} \\ \dot{x}(t) \;=\; f(t,x(t),u(t)),\ \dot{y}(t) \;=\; 0,\ \dot{w}(t) \;=\; |u(t)-u_i(t)| \quad a.e., \\ u(t) \in U_\varepsilon(t) \quad a.e., \\ ((x(0),y(0),w(0)) \in C \times \{0\}. \end{cases}$$

Notice that the cost function of (R_i) satisfies assumption G of Proposition 4.1, thus this is an example of the optimal control problem to which Proposition 4.1 applies. We deduce existence of $p_i \in W^{1,1}$, $d_i \in \mathbb{R}^n$, $r_i \in \mathbb{R}$, nonnegative Radon measure $\mu_i \in C^*$ and integrable functions γ_i, ξ_i satisfying

$$(-\dot{p}_i(t), -\dot{d}_i(t), -\dot{r}_i(t), \dot{x}_i(t), \dot{y}_i(t), 0, \xi_i(t)) \in \operatorname{co} \partial H_i(t, x_i(t), y_i(t), 0, q_i(t), d_i(t), r_i(t), u_i(t)) \quad a.e.\, t \in [0,1]$$
$$\xi_i(t) \in \operatorname{co} N_{U_\varepsilon(t)}(u_i(t)) \quad a.e.\, t \in [0,1] \quad (4.49)$$

$$(p_i(0), d_i(0), r_i(0), -q_i(1), -d_i(1), -r_i(1), \int_{[0,1]} \mu_i(\mathrm{d}t)) \in \; N_{C\times\{0\}} \times \{0,0,0,0\} + \partial \{\tilde{g}_i(x, y, x', y', z)$$
$$+\varepsilon_i[|x - x_i(0)| + |y - y_i(0)| + w]\} \quad (4.50)$$

where the last gradient is calculated at the point

$$(x_i(0), y_i, w_i(0) = 0, x_i(1), y_i, w_i(1) = 0, z_i = \max\{h(t, x_i(t))\})$$

$$\operatorname{supp}\{\mu_i\} \;\subset\; \{t : h(t, x_i(t)) = \max_{s\in[0,1]} \{h(s, x_i(s))\}\},$$
$$\gamma_i(t) \;\in\; \bar{\partial}_x h(t, x_i(t)) \;\mu-a.e.$$

The Hamiltonian H_i is defined as

$$H_i(t, x, y, w, p, d, r, u) := p \cdot f(t, x, u) + r|u - u_i(t)|,$$

and in the above relationships $q_i := p_i + \int \gamma_i \mu_i(\mathrm{d}s)$. To derive these conditions, we have identified p, d, and r as the adjoint variables associated with the x, y and w variables respectively.

Observe that

$$\begin{aligned}\operatorname{co} \partial H_i(t, x, y, w, p, d, r, u) &\subset \operatorname{co} \partial\, [p \cdot f(t, x, u)] + \operatorname{co} \partial\, [r|u - u_i(t)|] \\ &= \{(a, 0, 0, b, 0, 0, c) : (a, b, c) \in \operatorname{co} \partial_{x,p,u} p \cdot f(t, x, u)\} \\ &\quad + \{(0, 0, 0, 0, 0, |u - u_i(t)|, re) : e \in \operatorname{co} \partial_u |u - u_i(t)|\}.\end{aligned}$$

We have

$$\begin{aligned}&\operatorname{co} \partial H_i(t, x_i(t), y_i(t), 0, q_i(t), d_i(t), r_i(t), u_i(t)) \subset \\ &\quad \{(a, 0, 0, b, 0, 0, c + r_i\beta_i) : ||\beta_i|| \le 1 \text{ and } (a, b, c) \in \operatorname{co} \partial H(t, x_i(t), q_i(t), u_i(t))\}\end{aligned}$$

So, $\dot{d}_i(t) = 0$ and $\dot{r}_i(t) = 0$. Let us say then $d_i(t) \equiv d_i$ and $r_i(t) \equiv r_i$, constants. We also have that

$$(-\dot{p}_i(t), \dot{x}_i(t), \xi_i(t)) \in \operatorname{co} \partial H(t, x_i(t), q_i(t), u_i(t)) + (0, 0, r_i\beta_i) \quad (4.51)$$

with $||\beta_i|| \le 1$.

From condition (4.50) we deduce that

$$\begin{aligned}&(p_i(0), d_i, r_i, -q_i(1), -d_i, -r_i, \int_{[0,1]} \mu_i(\mathrm{d}t)) \in N_C \times \mathbb{R}^n \times \{(0,0,0,0)\} \\ &\quad + \{(a, b, 0, c, d, 0, e) : (a, b, c, d, e) \in \partial \tilde{g}_i(x_i(0), y_i(0), x_i(1), y_i(1), \max\{h(t, x_i(t))\}) \\ &\quad + \varepsilon_i\, [B \times B \times \{0\} \times \{(0,0)\} \times \{1\} \times \{0\}]\end{aligned}$$

which implies that

$$(p_i(0), d_i, -q_i(1), -d_i, \int_{[0,1]} \mu_i(\mathrm{d}t)) \in N_C \times \{(0,0,0)\}$$
$$+ \partial \tilde{g}_i(x_i(0), y_i(0), x_i(1), y_i(1), \max\{h(t, x_i(t))\}) + \varepsilon_i (B \times B) \times \{(0,0,0)\} \quad (4.52)$$

and

$$r_i = -\varepsilon_i.$$

From condition (4.52) we deduce that $\{\|\mu_i\|_{T.V}\}$, $\{d_i\}$ and $\{p_i(1)\}$ are bounded sequences. By (4.51) $\{p_i\}$ is uniformly bounded and $\{\dot{p}_i\}$ and $\{\xi_i\}$ are uniformly integrably bounded. We deduce that, following a subsequence extraction,

$$p_i \to p \text{ uniformly}, \quad d_i \to d, \quad \xi_i \to \xi \text{ in the } L^1 \text{ norm},$$

and

$$\mu_i \to \mu, \ \gamma_i \mu_i(\mathrm{d}t) \to \gamma \mu(\mathrm{d}t) \text{ weakly}^*,$$

for some $p \in W^{1,1}$, $d \in \mathbb{R}^n$, $\xi \in L^1$, nonnegative Radon measure $\mu \in C^*$ and some Borel measurable function γ, as $i \to \infty$. Furthermore

$$\mathrm{supp}\{\mu\} \ \subset \ \{t : h(t, \bar{x}(t)) = \max_{s \in [0,1]} h(s, \bar{x}(s))\}$$

and

$$\gamma(t) \in \bar{\partial}_x h(t, \bar{x}(t)) \ \mu - a.e.$$

Write $q = p + \int \gamma \mu(\mathrm{d}s)$.

By subsequence extraction we can have $\{\xi_i\}$ converging to ξ almost everywhere.

A convergence analysis along the lines of the proof of Theorem 3.1 in [3] and an appeal to the upper semi continuity properties of limiting subdifferentials and normal cones allow us to pass to the limit in relationships (4.51) and (4.49). There results

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) \in \mathrm{co}\, \partial H(t, \bar{x}(t), q(t), \bar{u}(t)) \quad a.e. \ t \in [0,1],$$

and

$$\xi(t) \in \mathrm{co}\, N_{U_\varepsilon(t)}(\bar{u}(t)) \quad a.e. \ t \in [0,1].$$

Let us now analyse the transversality condition (4.52). An important implication is that

$$\tilde{g}_i(x_i(0), y_i(0), x_i(1), y_i(1), \max_{s \in [0,1]} h(s, x_i(s))) \ > \ 0 \quad (4.53)$$

for all i sufficiently large. This is because, if $\tilde{g}_i = 0$ for some sufficiently large i,

$$x_i(1) \ = \ y_i(1) \ = \ y_i(0), \ (x_i(0), x_i(1)) \ \in \ C, \ \max_{s \in [0,1]} h(s, x_i(s)) \ \le \ 0,$$

$$\|x_i - \bar{x}\|_{L^\infty} \ \le \ \varepsilon, \ \|u_i - \bar{u}\|_{L^\infty} \ \le \ \varepsilon$$

and

$$g(x_i(0), x_i(1)) \ \le \ g(\bar{x}(0), \bar{x}(1)) - \varepsilon_i^2,$$

in violation of the optimality of $(\bar{x}, \bar{u})$ (see the definition of $\tilde{g}_i$).

Define

$$z_i \ = \ \max_{s \in [0,1]} h(s, x_i(s)).$$

We verify the following estimate for $\partial\tilde{g}_i$:

$$\partial\tilde{g}_i(x_i(0), y_i(0), x_i(1), y_i(1), z_i) \;\subset \big\{(a, b, e, -e, c) \,\in\, \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} : \tag{4.54}$$

$$\exists\tilde{\lambda} \geq 0 \text{ such that } \tilde{\lambda} + |e| = 1 \text{ and}$$

$$(a, b, c) \;\in\; \tilde{\lambda}\,\partial\max\{g(x, y) - g(x_i(0), y_i(0)) + \varepsilon_i^2, z\}|_{(x_i(0),y_i(0),z_i)}. \tag{4.55}$$

There are two cases to consider:

Case a. $x_i(1) = y_i(1)$. In this case (by (4.53))

$$\tilde{g}_i(x, y, x', y', z) \;=\; \max\{g(x, y) - g(x_i(0), y_i(0)) + \varepsilon_i^2, z\}$$

near $(x_i(0), y_i(0), x_i(1), y_i(1), z_i)$. Consequently (4.55) is true with $e = 0$.

Case b. $x_i(1) \neq y_i(1)$. In this case, from the Chain Rule,

$$\partial|x - y|\;|_{(x_i(1),y_i(1))} \in \{(e, -e) : |e| = 1\}.$$

Inclusion (4.55) then follows from the Max Rule for limiting subdifferentials.

Again by the Max Rule,

$$\partial\max\{g(x, y) - g(x_i(0), y_i(0)) + \varepsilon_i^2, z\}|_{(x_i(0),y_i(0),z_i)} \;\subset\; \{(\alpha\partial g(x_i(0), y_i(0)), (1 - \alpha)) : \alpha \in [0, 1]\}, \tag{4.56}$$

and we deduce from (4.52) and (4.55) that there exist $\tilde{\lambda}_i \geq 0$ and $\alpha_i \in [0, 1]$ such that

$$(p_i(0), d_i, -q_i(1), -d_i, \int_{[0,1]} \mu_i(\mathrm{d}t)) = (c_1, c_2, 0, 0, 0) + (a, b, e, -e, c) + \varepsilon_i(b_1, b_2, 0, 0, 0)$$

where

$$\begin{gathered}(c_1, c_2) \in N_C(x_i(0), y_i(0)),\\ |e| + \tilde{\lambda}_i = 1,\\ (a, b, c) = \tilde{\lambda}_i(\alpha_i g_1, 1 - \alpha_i), \quad g_1 \in \partial g(x_i(0), y_i(0)),\\ b_1, b_2 \in B.\end{gathered}$$

This implies that

$$\begin{cases} p_i(0) = c_1 + a + \varepsilon_i b_1\\ d_i = c_2 + b + \varepsilon_i b_2\\ -q_i(1) = e\\ -d_i = -e\\ \displaystyle\int_{[0,1]} \mu_i(\mathrm{d}t) = c\end{cases}$$

and from this we get

$$\begin{cases} d_i = -q_i(1)\\ \tilde{\lambda}_i + |q_i(1)| = 1\\ ||\mu_i||_{TV} = c = \tilde{\lambda}_i(1 - \alpha_i)\\ (p_i(0), -q_i(1)) \in N_C(x_i(0), y_i(0)) + \alpha_i\tilde{\lambda}_i\partial g(x_i(0), y_i(0)) + \varepsilon_i(b_1, b_2).\end{cases}$$

Observe also that

$$\mu_i = 0 \text{ if } z_i \leq 0,$$

by (4.52), since $z_i \leq 0$ implies

$$\tilde{g}_i(x,y,x',y',z) := \max\{g(x,y) - g(\bar{x}(0),\bar{x}(1)) + \varepsilon_i^2, |x'-y'|\}$$

for (x,y,x',y',z) near $(x_i(0),y_i(0),x_i(1),y_i(1),z_i)$ which in turns implies $\alpha_i = 1$, and consequently $\|\mu_i\|_{T.V.} = 0$.

Now choose $\lambda_i = \alpha_i\tilde{\lambda}_i$. It follows from the above relations that

$$1 = \tilde{\lambda}_i + |q_i(1)| = \tilde{\lambda}_i\alpha_i + \tilde{\lambda}_i(1-\alpha_i) + |q_i(1)| = \lambda_i + \|\mu_i\|_{T.V.} + |q_i(1)|. \tag{4.57}$$

Along a subsequence $\lambda_i \to \lambda$, for some $\lambda \geq 0$.

In the limit as $i \to \infty$ we obtain from the relations above the conditions

$$(p(0), -q(1)) \in \lambda\partial g(\bar{x}(0),\bar{x}(1)) + N_C(\bar{x}(0),\bar{x}(1))$$

and

$$\lambda + \|\mu\|_{T.V.} + |q(1)| = 1.$$

Surveying these relationships we see that the proposition is proved, except that $\bar{\partial}_x h$ replaces $\partial_x^{>} h$. To attend to this remaining detail we must examine two possible outcomes of the sequence construction we have undertaken.

(A): $\max_{t\in[0,1]} h(t,x_i(t)) > 0$ for at most a finite number of i's. For this possibility $\mu_i = 0$ for all i sufficiently large, by (4.52) and (4.53). Then the preceding convergence analysis gives $\mu = 0$ and $\gamma \in \partial_x^{>} h(t,\bar{x}(t))$ μ - *a.e.*, trivially.

(B): $\max_{t\in[0,1]} h(t,x_i(t)) > 0$ for an infinite number of i's. Now we can arrange, by a further subsequence extraction, that

$$\max_{t\in[0,1]} h(t,x_i(t)) > 0 \quad \text{for all } i. \tag{4.58}$$

We have

$$\text{supp}\{\mu_i\} \subset \{t : h(t,x_i(t)) = \max_{s\in[0,1]} h(s,x_i(s))\}$$

and, since the inequality (4.58) is strict, the condition $\gamma_i \in \bar{\partial}_x h(t,x_i(t))$ can be replaced by the more precise relationship

$$\gamma_i(t) \in \partial_x^{>} h(t,x_i(t)) \quad \mu_i\text{ - } a.e.$$

The fact that we can arrange, in the limit, that

$$\gamma(t) \in \partial_x^{>} h(t,\bar{x}(t))$$

now follows from Proposition 9.2.1 in [12].

References

[1] K.E. Brenen, S.L. Campbell and L.R. Petzold, *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations. Classics Appl. Math.* SIAM, Philadelphia (1996).

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis.* Wiley, New York (1983). Reprinted as Vol. 5 of *Classics Appl. Math.* SIAM, Philadelphia (1990).

[3] M.d.R. de Pinho, M.M.A. Ferreira and F.A.C.C. Fontes, An Euler-Lagrange inclusion for optimal control problems with state constraints. *J. Dynam. Control Syst.* **8** (2002) 23–45.

[4] M.d.R. de Pinho, M.M.A. Ferreira and F.A.C.C. Fontes, Necessary conditions in Euler-Lagrange inclusion form for constrained nonconvex optimal control problems, in *Proc. of the 10th Mediterranean Conference on Control and Automation.* Lisbon, Portugal (2002).

[5] M.d.R. de Pinho and A. Ilchmann, Weak maximum principle for optimal control problems with mixed constraints. *Nonlinear Anal. Theory Appl.* **48** (2002) 1179–1196.

[6] M.d.R. de Pinho and R.B. Vinter, An Euler-Lagrange inclusion for optimal control problems. *IEEE Trans. Aut. Control* **40** (1995) 1191–1198.

[7] M.d.R. de Pinho and R.B. Vinter, Necessary conditions for optimal control problems involving nonlinear differential algebraic equations. *J. Math. Anal. Appl.* **212** (1997) 493–516.

[8] M.d.R. de Pinho, R.B. Vinter and H. Zheng, A maximum principle for optimal control problems with mixed constraints. *IMA J. Math. Control Inform.* **18** (2001) 189–205.

[9] B.S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.* **40** (1976) 960–969.

[10] B.S. Mordukhovich, *Approximation Methods in Problems of Optimization and Control*. Nakua, Moscow; the 2nd edition to appear in Wiley-Interscience (1988).

[11] R.T. Rockafellar and B. Wets, *Variational Analysis*. Springer, Berlin (1998).

[12] R.B. Vinter, *Optimal Control*. Birkhauser, Boston (2000).