AN EULER-LAGRANGE INCLUSION FOR OPTIMAL CONTROL PROBLEMS WITH STATE CONSTRAINTS

M. D. R. DE PINHO[1], M. M. A. FERREIRA[1], AND F. A. C. C. FONTES[2]

[1] ISR and DEEC
Faculdade de Engenharia da Universidade do Porto
Rua Dr. Roberto Frias
4200-465 Porto, Portugal.
tel: +351 225 081 617/794
fax: +351 225 081 440
e-mail: {mrpinho,mmf}@fe.up.pt

[2] Departamento de Matemática, CMAT,
Universidade do Minho,
4800-058 Guimarães, Portugal.
tel: +351 253 510 218
fax: +351 253 510 153
e-mail: ffontes@math.uminho.pt

Abstract. New first-order necessary conditions of optimality for control problems with pathwise state constraints are given. These conditions are a variant of a nonsmooth maximum principle which includes a joint subdifferential of the Hamiltonian — a condition called Euler-Lagrange inclusion (ELI). The main novelty of the result provided here is the ability to address state constraints while using an ELI.

The ELI conditions have a number of desirable properties. Namely, they are, in some cases, able to convey more information about minimizers, and for the normal convex problems they are sufficient conditions of optimality. It is shown that these strengths are retained in the presence of state constraints.

key words: Optimal control, maximum principle, state constraints, nonsmooth analysis, Euler-Lagrange inclusion.

Date: July 2, 2001, (First Submitted February 5, 2001. Accepted March 29, 2001).
1991 *Mathematics Subject Classification.* 49K15.

1. Introduction

We focus on the optimal control problem

$$(P) \quad \text{Minimize} \quad \int_0^1 L(t, x(t), u(t))\, dt + g(x(0), x(1)) \tag{1.1}$$

subject to

$$\dot{x}(t) = f\,(t, x(t), u(t)) \quad \text{a.e. } t \in [0,1] \tag{1.2}$$

$$x(0) \in C_0$$

$$x(1) \in C_1$$

$$u(t) \in U(t) \quad \text{a.e. } t \in [0,1]$$

$$h\,(t, x(t)) \leqslant 0 \quad \text{for all } t \in [0,1], \tag{1.3}$$

for which the data comprises functions $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, and $h : [0,1] \times \mathbb{R}^n \to \mathbb{R}$; a multifunction $U : [0,1] \rightrightarrows \mathbb{R}^m$; and sets $C_0, C_1 \subset \mathbb{R}^n$.

The domain of problem (P) is the set of *admissible processes*, namely pairs (x, u) comprising a measurable control function u and a corresponding state trajectory $x \in W^{1,1}([0,1], \mathbb{R}^n)$ which satisfies the constraints of (P) and for which $t \mapsto L(t, x(t), u(t))$ is integrable.

For the problem above we provide a variant of the Maximum Principle involving a "joint" generalized gradient of the Hamiltonian — a condition we call Euler Lagrange Inclusion, ELI for short. The new set of conditions extends previous results in [1] to problems with state constraints.

To illustrate the main ideas behind the new conditions let us consider the nonsmooth Maximum Principle (see e.g. [2]). It asserts that an optimal process $(\bar{x}, \bar{u})$ of (P) satisfies, among others, the conditions

$$-\dot{p}(t) \in \operatorname{co} \partial_x H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad \text{a.e. } t, \tag{1.4}$$

$$\dot{\bar{x}}(t) = \operatorname{co} \partial_p H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad \text{a.e. } t, \tag{1.5}$$

$$H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) = \max_{u \in U(t)} H_\lambda(t, \bar{x}(t), q(t), u) \quad \text{a.e. } t, \tag{1.6}$$

where $\operatorname{co} \partial H_\lambda$ denotes the *Clarke subdifferential* of the Hamiltonian

$$H_\lambda(t, x, p, u) := p \cdot f(t, x, u) - \lambda L(t, x, u).$$

The existence of a function of bounded variation q is assumed. This function q is related to its absolutely continuous "part", p, by the addition of an integral term involving a multiplier η associated with the state constraint as follows:

$$q(t) = p(t) + \int_{[0,t)} \mathrm{d}\eta(s).$$

Under mild hypotheses on the data, which include Lipschitz continuity of L and f with respect to (x, u), a commonly used form of a weak maximum principle is obtained when (1.6) is replaced by

$$\xi(t) \in \operatorname{co} \partial_u H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad \text{a.e. } t, \tag{1.7}$$

and

$$\xi(t) \in \operatorname{co} N_{U(t)}(\bar{u}(t)) \quad \text{a.e. } t. \tag{1.8}$$

The Euler-Lagrange Inclusion type of condition (ELI) for problem (P) we consider in this work involves a joint subdifferential of the Hamiltonian in the (x, q, u) variables

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) \in \operatorname{co} \partial H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad \text{a.e.} \tag{1.9}$$

together with the inclusion (1.8).

It is a well known fact that, for nonsmooth problems, (1.4), (1.5) and (1.7) are not equivalent to (1.9). Inclusion (1.9) can give more information in situations where

$$\operatorname{co} \partial H_\lambda \neq \operatorname{co} \partial_x H_\lambda \times \operatorname{co} \partial_p H_\lambda \times \operatorname{co} \partial_u H_\lambda.$$

Here we generalize existent results on an ELI to optimal control problems with state constraints along the trajectory. The main strengths of such type of conditions are retained in the presence of state constraints. Namely, these conditions can eliminate processes that satisfy the nonsmooth maximum principle and yet are not locally optimal. Furthermore, for convex problems our result, in the normal form, is also a sufficient condition.

Noteworthy, the key idea behind the proof of the necessary conditions of optimality we report in here is very simple: a sequence of optimal control problems is constructed in which the state constraint is replaced by a penalty term that is added to the cost. This is an approach we retrieve from the seminal paper of Vinter and Pappas [3].

2. Preliminaries

Here and throughout, B represents the *closed* unit ball centred at the origin and $|\cdot|$ 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

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

The linear space $W^{1,1}([0,1];\mathbb{R}^p)$ denotes the space of absolutely continuous functions, $L^1([0,1];\mathbb{R}^p)$ the space of integrable functions and $L^\infty([0,1];\mathbb{R}^p)$ the space of essentially bounded functions from $[0,1]$ to $\mathbb{R}^p$.

By $w^*\text{-}\lim_{i\to\infty} m_i$, we mean the *weak** limit of the sequence m_i. Here, m_i is a sequence in $C^*([0,1],\mathbb{R})$, the dual space of the space of continuous functions defined in $[0,1]$ and taking values in $\mathbb{R}$.

We make use of the following concepts from nonsmooth analysis. A vector $p\in\mathbb{R}^k$ is a *limiting normal to* a closed set A of $\mathbb{R}^k$ at a point x in A if there exist $p_i\to p$, $x_i\to x$, 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

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

where $epi\,\{f\} = \big\{(x,\eta) \;:\; \eta\geqslant f(x)\big\}$ denotes the epigragh set . In the case that the function f is Lipschitz continuous near x, the convex hull of the limiting subdifferential, $\mathrm{co}\,\partial f(x)$, coincides with the *Clarke subdifferential*, which may be defined directly. Properties and the calculus for these various constructions may be found in [4, 5, 6, 2].

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

$$\bar\partial h(t,x) := \mathrm{co}\,\{\lim\xi_i \;:\; \xi_i\in\partial_x h(t_i,x_i),(t_i,x_i)\mapsto(t,x)\}.$$

We say that an admissible process $(\bar x,\bar u)$ is a *local minimizer* of (P) if there exists $\varepsilon>0$ such that

$$\int_0^1 L(t,\bar x(t),\bar u(t))dt + g(\bar x(0),\bar x(1)) \leqslant \int_0^1 L(t,x(t),u(t))dt + g(x(0),x(1))$$

over all admissible processes (x,u) satisfying

$$x(t)\in\bar x(t)+\varepsilon B,\quad u(t)\in\bar u(t)+\varepsilon B.$$

Let us consider the optimal control problem (P) when *no state constraints* are present and suppose that $(\bar x,\bar u)$ is a local minimizer in the sense defined above. The following hypotheses, which make reference to a parameter $\varepsilon>0$, are imposed:

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

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

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

H2: The multifunction U has Borel measurable graph and

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

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

H3: The endpoint constraint sets C_0 and C_1 are closed and g is locally Lipschitz around $(\bar{x}(0), \bar{x}(1))$.

The following Euler Lagrange Inclusion for optimal control problems, provided in [1], will be of importance in our analysis.

Proposition 2.1. *Let $(\bar{x}, \bar{u})$ be a local minimizer for (P) when* no state constraints *are present. Assume that H1–H3 are satisfied and*

$$H_\lambda(t,x,p,u) \;=\; p \cdot f(t,x,u) - \lambda L(t,x,u)$$

defines the Hamiltonian. Then there exist $\lambda \geqslant 0$, $p \in W^{1,1}([0,1];\mathbb{R}^n)$ and $\zeta \in L^1([0,1];\mathbb{R}^m)$ such that, for almost all $t \in [0,1]$,

$$\begin{aligned} \lambda + \| p \|_{L^\infty} &\neq 0 \\ (-\dot{p}(t), \dot{\bar{x}}(t), \zeta(t)) &\in co\, \partial H_\lambda(t, \bar{x}(t), p(t), \bar{u}(t)) \\ \zeta(t) &\in co\, N_{U(t)}(\bar{u}(t)) \\ (p(0), -p(1)) &\in N_{C_0 \times C_1}(\bar{x}(0), \bar{x}(1)) + \lambda \partial g(\bar{x}(0), \bar{x}(1)). \end{aligned}$$

3. Main Results

We first establish the validity of necessary conditions of optimality in the form of an Euler-Lagrange Inclusion for optimal control problems with state constraints. For such problem we consider two additional hypotheses:

H4: For $x \in \bar{x}(t) + \varepsilon B$ 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 $t \in [0,1]$.

H5: The extended velocity set

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

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

Theorem 3.1. *Let $(\bar{x}, \bar{u})$ be a local minimizer for problem (P). Assume that H1–H5 are satisfied and*

$$H_\lambda(t,x,p,u) := p \cdot f(t,x,u) - \lambda L(t,x,u)$$

defines the Hamiltonian. 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

$$\mu\{[0,1]\} + ||p||_{L_\infty} + \lambda > 0, \tag{3.10}$$

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) \in \operatorname{co} \partial H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad a.e.\ t \in [0,1], \tag{3.11}$$

$$\xi(t) \in \operatorname{co} N_{U(t)}(\bar{u}(t)) \quad a.e.\ t \in [0,1], \tag{3.12}$$

$$(p(0), -q(1)) \in N_{C_0}(\bar{x}(0)) \times N_{C_1}(\bar{x}(1)) + \lambda \partial g(\bar{x}(0), \bar{x}(1)), \tag{3.13}$$

$$\gamma(t) \in \bar{\partial} h(t, \bar{x}(t)) \quad \mu\text{-}a.e.\ , \tag{3.14}$$

$$\operatorname{supp}\{\mu\} \subset \{t \in [0,1] : h\,(t, \bar{x}(t)) = 0\}\,, \tag{3.15}$$

where

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

To see that the optimality conditions of Theorem 3.1 are distinct from the maximum principle we study an example based on one given in [1]. The optimal control problem of interest is the following

$$(P1) \quad \text{Minimize} \quad \int_0^1 \{w_1|x - u_1| + w_2|x - u_2| + x\}\, dt$$

subject to

$$\dot{x}(t) = 4w_1u_1 + 4w_2u_2 \qquad \text{a.e. } t \in [0,1]$$

$$x(0) = 0$$

$$u_1(t), u_2(t) \in [-1,1] \quad \text{a.e. } t \in [0,1]$$

$$(w_1(t), w_2(t)) \in \{(w_1, w_2) : w_1 \geq 0, w_2 \geq 0, \; w_1 + w_2 = 1\} \qquad \text{a.e. } t \in [0,1]$$

$$x(t) \geqslant -1 \qquad \text{for all } t \in [0,1],$$

Here $h(t,x) = -x - 1$.

The process $(\bar{x} \equiv 0, \bar{u}_1 = \bar{u}_2 = \bar{w}_2 \equiv 0, \bar{w}_1 = 1)$ is admissible and has cost zero but is it is not optimal since, for any $\alpha \in (0, \frac{1}{4}]$, the process $x_\alpha(t) = -4\alpha t, u_{1\alpha}(t) = -\alpha$, $u_{2\alpha}(t) = w_{2\alpha}(t) = 0, w_{1\alpha}(t) = 1$, which satisfies the state constraint, has cost -0.75α. Yet, for the process ($\bar{x} \equiv 0, \bar{u}_1 = \bar{u}_2 = \bar{w}_2 \equiv 0, \bar{w}_1 = 1$) the standard nonsmooth maximum principle is satisfied in normal form, with $\lambda = 1$ and all the other multipliers null. On the other hand, the conclusions of Theorem 3.1 are not satisfied for such process. In fact, if Theorem 3.1 were to apply, we would need the existence of an absolutely continuous function p such that $-\dot{p} = -\lambda(1 + \sigma_1)$, $p(1) = 0$ and $4p + \lambda\sigma_1 = 0$, where $\sigma_1 \in [-1, 1]$. It is easy to check that no such p exists.

4. Sufficiency for Convex Problems

The example of the previous section is a nonsmooth, convex problem. Nonsmoothness is crucial here. Indeed, the classical smooth maximum principle provides necessary conditions which are sufficient for convex problems, even with state constraints (see [7, 8]). However, nonsmooth versions of the maximum principle fail to provide a sufficient condition for convex problems as illustrated.

We now show that the normal form of the ELI condition of Theorem 3.1 provides, for convex problems, a sufficient condition for optimality.

In this respect, let us consider the following convex optimal control problem with state constraints:

$$(CP) \quad \text{Minimize} \quad \int_0^1 L(t, x(t), u(t))\, dt + g(x(0), x(1))$$

$$\text{subject to}$$

$$\begin{aligned}
&\dot{x}(t) = A(t)x(t) + B(t)u(t) && \text{a.e. } t \in [0,1]\\
&x(0) \in C_0\\
&x(1) \in C_1\\
&u(t) \in U(t) && \text{a.e. } t \in [0,1]\\
&D(t)x(t) \leqslant 0 && \text{for all } t \in [0,1].
\end{aligned}$$

We assume that the sets C_0, C_1 are convex; the multifunction $U(t)$ is convex for a.e. t in $[0, 1]$, the functions g and $(x, u) \to L(t, x, u)$ are convex, the function $A : [0, 1] \to \mathbb{R}^{n\times n}$ is integrable, the function $B : [0, 1] \to \mathbb{R}^{n\times m}$ is measurable, and the function $D : [0, 1] \to \mathbb{R}^{1\times n}$ is continuous.

Definition 4.1. *A process $(\bar{x}, \bar{u})$ is a normal extremal if it is an admissible process and satisfies conditions (3.10)–(3.15) with $\lambda = 1$.*

When we apply the above definition to problem (CP), the normal cones and subdifferentials involved can be understood in the sense of convex analysis due to the convexity properties of the data.

Proposition 4.1. *If the pair $(\bar{x}, \bar{u})$ is a normal extremal for problem (CP), then it is a minimizer.*

Proof. Let (x, u) be an arbitrary admissible process and the pair $(\bar{x}, \bar{u})$ be a normal extremal for the convex problem (CP). We will show that

$$\int_0^1 [L(t, x(t), u(t)) - L(t, \bar{x}(t), \bar{u}(t))] \, dt + g(x(0), x(1)) - g(\bar{x}(0), \bar{x}(1)) \geqslant 0.$$

Since $(\bar{x}, \bar{u})$ is a normal extremal there exist an absolutely continuous function $p : [0,1] \to \mathbb{R}^n$, an integrable function $\xi : [0,1] \to \mathbb{R}^n$, a nonnegative Radon measure $\mu \in C^*([0,1], \mathbb{R})$ and vectors $(\zeta_0, \zeta_1) \in \mathbb{R}^n \times \mathbb{R}^n$, such that

$$(\dot{p}(t) + q(t)A(t), -\xi(t) + q(t)B(t)) \in \partial L(t, \bar{x}(t), \bar{u}(t)) \quad \text{a.e. }, \tag{4.16}$$

$$\xi(t) \cdot (u - \bar{u}(t)) \leqslant 0 \quad \forall u \in U(t), \text{ a.e. }, \tag{4.17}$$

$$(p(0), -q(1)) - (\zeta_0, \zeta_1) \in \partial g(\bar{x}(0), \bar{x}(1)), \tag{4.18}$$

$$(\zeta_0, \zeta_1) \cdot [(x_0, x_1) - (\bar{x}(0), \bar{x}(1))] \leqslant 0 \quad \forall (x_0, x_1) \in C_0 \times C_1, \tag{4.19}$$

$$\text{supp}\{\mu\} \subset \{t \in [0,1] : D(t)\bar{x}(t) = 0\}, \tag{4.20}$$

where

$$q(t) = \begin{cases} p(t) + \int_{[0,t)} D(s)\mu(ds) & t \in [0,1) \\ p(t) + \int_{[0,1]} D(s)\mu(ds) & t = 1. \end{cases}$$

By definition of a process $t \to L(t, x(t), u(t))$ and $t \to B(t)u(t)$ are integrable.

Computing the difference of the cost between (x,u) and $(\bar{x},\bar{u})$ we get

$$\int_0^1 [L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))]\, dt + g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1))$$

$$= \int_0^1 \big[L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))\big]\, dt + g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1))$$

$$+ \int_0^1 q(t)\big[\dot{x}(t) - A(t)x(t) - B(t)u(t) - (\dot{\bar{x}}(t) - A(t)\bar{x}(t) - B(t)\bar{u}(t))\big]\, dt$$

$$= \int_0^1 \big[L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))\big]\, dt + g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1))$$

$$+ \int_0^1 \big[(q(t) - p(t))\cdot(\dot{x}(t) - \dot{\bar{x}}(t)) - q(t)A(t)(x(t) - \bar{x}(t))$$

$$-q(t)B(t)(u(t) - \bar{u}(t))\big]\, dt$$

$$+p(1)[x(1) - \bar{x}(1)] - p(0)[x(0) - \bar{x}(0)] - \int_0^1 \dot{p}(t)[x(t) - \bar{x}(t)]dt.$$

In the last step, we perform an integration by parts of $\int_0^1 p\cdot(\dot{x} - \dot{\bar{x}})$.

Rearranging and after some more addition and subtraction of terms we obtain

$$\int_0^1 \big[L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))\big]\, dt + g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1))$$

$$= \int_0^1 \big[L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))\big]dt$$

$$- \int_0^1 \Big[[\dot{p}(t) + q(t)A(t)]\,[x(t) - \bar{x}(t)] + [q(t)B(t) - \xi(t)]\,[u(t) - \bar{u}(t)]\Big]\, dt$$

$$+g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1))$$

$$-[(p(0), -q(1)) - (\zeta_0,\zeta_1)]\cdot[(x(0),x(1)) - (\bar{x}(0),\bar{x}(1))]$$

$$+ \int_0^1 [q(t) - p(t)]\,[\dot{x}(t) - \dot{\bar{x}}(t)]\, dt - [q(1) - p(1)]\,[x(1) - \bar{x}(1)]$$

$$- \int_0^1 \xi(t)\,[u(t) - \bar{u}(t)]\, dt - (\zeta_0,\zeta_1)\cdot[(x(0),x(1)) - (\bar{x}(0),\bar{x}(1))]$$

From (4.17) and (4.19), we can deduce that the two terms in the last line are nonnegative. By the properties of a generic convex function f,

$$f(x) - f(\bar{x}) - \zeta \cdot (x - \bar{x}) \geq 0, \quad \forall \zeta \in \partial f(\bar{x}).$$

The difference is then simplified to

$$\begin{aligned}
&\int_0^1 \big[L(t, x(t), u(t)) - L(t, \bar{x}(t), \bar{u}(t))\big]\, dt + g(x(0), x(1)) - g(\bar{x}(0), \bar{x}(1)) \\
&\quad \geq \int_0^1 [q(t) - p(t)]\,[\dot{x}(t) - \dot{\bar{x}}(t)]\, dt - [q(1) - p(1)][x(1) - \bar{x}(1)] \\
&\quad = \int_0^1 \int_{[0,t)} D(s)\mu(ds)[\dot{x}(t) - \dot{\bar{x}}(t)]\, dt - \int_{[0,1]} D(s)\mu(ds)(x(1) - \bar{x}(1)).
\end{aligned}$$

Now consider a sequence $\{m_i\}$ of measures which are absolutely continuous with respect to the Lebesgue measure and such that

$$w^*\text{-}\lim_{i\to\infty} m_i = \mu.$$

Then

$$\begin{aligned}
&\int_0^1 \int_{[0,t)} D(s)\mu(ds)[\dot{x}(t) - \dot{\bar{x}}(t)]\, dt - \int_{[0,1]} D(s)\mu(ds)[x(1) - \bar{x}(1)] \\
&\quad = \lim_{i\to\infty} \int_0^1 \int_{[0,t)} D(s)m_i(s)\, ds[\dot{x}(t) - \dot{\bar{x}}(t)]\, dt \\
&\qquad\qquad - \int_{[0,1]} D(s)m_i(s)\, ds[x(1) - \bar{x}(1)] \\
&\quad = \lim_{i\to\infty} - \int_0^1 D(t)m_i(t)[x(t) - \bar{x}(t)]\, dt \\
&\quad = - \int_{[0,1]} D(t)[x(t) - \bar{x}(t)]\, \mu(dt)
\end{aligned}$$

We claim that

$$D(t)[x(t) - \bar{x}(t)] \leqslant 0, \ \mu\text{-a.e.}$$

In fact, by (4.20) $D(t)\bar{x}(t) = 0$, μ-a.e.. The admissibility of x yields the required inequality.

We conclude from the above that

$$\int_0^1 \big[L(t,x(t),u(t)) - L(t,\bar{x}(t),\bar{u}(t))\big]\,dt + g(x(0),x(1)) - g(\bar{x}(0),\bar{x}(1)) \geqslant 0.$$

proving the proposition. □

5. Proof of the Main Result

We now prove Theorem 3.1. The main step consists in constructing a sequence of optimal control problems where the state constraint is replaced by a penalty term added to the cost. Proposition 2.1 applies to each problem of the sequence.

We note for future use that the following conditions are automatically satisfied:

$$|f(t,\bar{x}(t),u) - f(t,\bar{x}(t),\bar{u}(t))| + |L(t,\bar{x}(t),u) - L(t,\bar{x}(t),\bar{u}(t))| \leqslant K(t)|u-\bar{u}(t)|$$

for all $u \in U_\varepsilon(t)$ a.e. $t \in [0,1]$ and there exists an integrable function c such that

$$|f(t,\bar{x}(t),u)| + |L(t,\bar{x}(t),u)| \leqslant c(t)$$

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

We proceed in several steps.

Step 1: We define a sequence of problems (P_i), without explicit state constraints.

Consider a sequence of positive numbers $\{k_i\}_{i\in\mathbb{N}}$ and the family of problems

$$(P_i) \quad \text{Minimize} \quad \int_0^1 L(t,x(t),u(t))\,dt + g(x(0),x(1)) + k_i \int_0^1 h^+(t,x(t))\,dt$$

subject to

$$\begin{aligned} &\dot{x}(t) = f(t,x(t),u(t)) && \text{a.e. } t \in [0,1] \\ &(x(0),x(1)) \in C_0 \times C_1 && \\ &u(t) \in U_\varepsilon(t) && \text{a.e. } t \in [0,1] \\ &x(t) \in \bar{x}(t) + \varepsilon B && \text{for all } t \in [0,1], \end{aligned}$$

where

$$h^+(t,x) := \max\{0, h(t,x)\}.$$

Notice that the state constraint $h(t,x(t)) \leqslant 0$ in (P) is incorporated in the cost function of each P_i. This component of the cost function acts as a penalty for the non fulfillment of the state constraint.

Define

$$\hat{H}_\lambda(t,x,p,u) = p \cdot f(t,x,u) - \lambda L(t,x,u) - \lambda k_i h^+(t,x).$$

For the above family of problems we assume the interim hypothesis

IH:

$$\lim_{k_i \to \infty} \inf\{P_i\} = \inf\{P\}$$

We first prove the theorem under this hypothesis. In the final stages of the proof we show that IH follows from H5.

Step 2: We consider a sequence $\{k_i\}$, $k_i > 0$, tending to infinity and a corresponding sequence of optimization problems to which Ekeland's Theorem applies.

Define the set V as the set of pairs (u,s) such that

$$\begin{aligned} V: \ = \ & \{(u,s) : \exists x_u \text{ such that } \dot{x}_u(t) = f(t,x_u(t),u(t)),\ x_u(0) = s, \\ & s \in C_0,\ x_u(1) \in C_1, u(t) \in U_\varepsilon(t) \text{ and } x_u(t) \in \bar{x}(t) + \varepsilon B\} \end{aligned}$$

Set

$$\delta(u,v) = \| u - v \|_{L^1}$$

and provide V with the metric

$$\Delta((u,s),(w,s')) = \|s - s'\| + \delta(u,w).$$

Consider now the sequence of optimization problems

$$\begin{aligned} (R_i) \quad & \text{Minimize} \quad J_i(u,s) \\ & \text{subject to} \\ & \qquad (u,s) \in V \end{aligned}$$

where

$$J_i(u,s) = \int_0^1 L(t,x_u,u)dt + g(x_u(0),x_u(1)) + \int_0^1 k_i h^+(t,x_u)dt.$$

Observe that $(\bar{x},\bar{u})$ is an admissible solution of (P_i) for any i. Set

$$\varepsilon_i = J_i(\bar{u},\bar{x}(0)) - \inf\{J_i(u,s) : (u,s) \in V\}.$$

Then $\varepsilon_i \geqslant 0$ and, since $h(t,\bar{x}(t)) \leqslant 0$ for all t in [0,1], we get

$$\varepsilon_i = \inf\{P\} - \inf\{P_i\}$$

It follows from the interim hypothesis that

$$\varepsilon_i \to 0. \tag{5.21}$$

The set V is a complete metric space, the functions J_i are continuous, and $(\bar{u}, \bar{x}(0))$ is an ε_i-solution for (R_i).

We are now in a position to apply Ekeland's Variational Principle [4] to (R_i), which asserts the existence of a sequence $\{(v_i, s_i)\}$ in V such that

$$\Delta((v_i, s_i), (\bar{u}, \bar{x}(0)) \leqslant \sqrt{\varepsilon_i}, \tag{5.22}$$

and

$$\bar{J}_i(v_i, s_i) \leqslant \bar{J}_i(u, s) \text{ for all } \ (u, s) \in V \tag{5.23}$$

where

$$\bar{J}_i(u, s) = J_i(u, s) + \sqrt{\varepsilon_i}\delta(u, v_i).$$

Let x_i be the admissible trajectory x_{v_i}.

Since, by (5.22), v_i converges to $\bar{u}$ in L^1, there is a subsequence which converges almost everywhere to $\bar{u}$. We also have $s_i \to \bar{x}(0)$. Without relabelling, we can conclude by [9, Lemma 9] that

$$x_i \to \bar{x} \qquad \text{uniformly.} \tag{5.24}$$

By discarding initial terms in the sequence if necessary, we can ensure that $x_i(t) \in \bar{x}(t) + \left(\frac{1}{2}\right) \varepsilon B$, for all i. Clearly then, and from (5.23), we deduce that $((x_i, 0), v_i)$ is a local minimizer for the following control problem $(\bar{P}_i)$:

$$(\bar{P}_i) \quad \text{Minimize} \quad \int_0^1 L(t, x(t), u(t))\, dt + g(x(0), x(1))$$

$$+k_i \int_0^1 h^+(t, x(t)) + \sqrt{\varepsilon_i} y(1)$$

subject to

$$\dot{x}(t) = f\,(t, x(t), u(t)) \quad \text{a.e. } t \in [0,1]$$

$$\dot{y}(t) = \|u(t) - v_i(t)\| \quad \text{a.e. } t \in [0,1]$$

$$x(0) \in C_0$$

$$y(0) = 0$$

$$x(1) \in C_1$$

$$u(t) \in U_\varepsilon(t) \qquad \text{a.e. } t \in [0,1].$$

Step 3: Necessary conditions of optimality for problem $(\bar{P}_i)$ are given in the following lemma.

Lemma 5.1. *There exists an absolutely continuous function $p_i : [0,1] \to \Re^n$, integrable functions $\rho_i, \xi_i : [0,1] \to \Re^m$, a measure $\pi_i \in C^*([0,1];\mathbb{R}^n)$, a nonnegative measure $\mu_i \in C^*([0,1];\mathbb{R})$ having support in $\{t \in [0,1] : \; h(t, x_i(t)) = 0\}$, measurable functions Φ_i, Λ_i*

$$\begin{aligned} \Phi_i(t) &= (\alpha_i^1(t), 0, \alpha_i^3(t)) \\ \Lambda_i(t) &= (\beta_i^1(t), 0, \beta_i^3(t)), \end{aligned}$$

a Borel measurable function Γ_i

$$\Gamma_i(t) = (\gamma_i^1(t), 0, 0),$$

a nonnegative number λ_i, and $b_i \in \mathbb{R}^n$ such that

$$\Phi_i(t) \in co\, \partial_{x,p,u} f(t, x_i(t), v_i(t)) \quad a.e. \tag{5.25}$$

$$\Lambda_i(t) \in co\, \partial_{x,p,u} L(t, x_i(t), v_i(t)) \quad a.e. \tag{5.26}$$

$$\Gamma_i(t) \in co\, \partial_{x,p,u} h(t, x_i(t)) \quad \mu_i \; a.e. \tag{5.27}$$

and

$$\Psi_i(t) = (\psi_i^1(t), \psi_i^2(t), \psi_i^3(t)) \in \operatorname{co} \partial_{x,p,u} H_{\lambda_i}(t, x_i(t), p_i(t), v_i(t))$$

where,

$$\begin{aligned} \psi_i^1(t) &= p_i(t)\alpha_i^1(t) - \lambda_i \beta_i^1(t) \\ \psi_i^2(t) &= f(t, x_i(t), v_i(t)) \\ \psi_i^3(t) &= p_i(t)\alpha_i^3(t) - \lambda_i \beta_i^3(t) \end{aligned}$$

and

$$\xi_i(t) = \psi_i^3(t) - \lambda_i \sqrt{\varepsilon_i}\rho_i(t) \tag{5.28}$$

$$\xi_i(t) \in co\, N_{U_\varepsilon(t)}(v_i(t)) \text{ a.e.} \tag{5.29}$$

$$-\int_B d\pi_i(t) = \int_B \psi_i^1(t)\, dt - \int_B \gamma_i^1(t)\, d\mu_i(t) \tag{5.30}$$

for all Borel sets $B \in [0,1]$,

$$p_i(t) = b_i + \int_{[0,t)} d\pi_i(\tau) \quad t \in (0,1] \tag{5.31}$$

$$\left(b_i, -b_i - \int_{[0,1]} d\pi_i(t)\right) \in N_{C_0 \times C_1}(x_i(0), x_i(1)) + \lambda_i \partial g(x_i(0), x_i(1)) \tag{5.32}$$

$$|b_i| + |\lambda_i| + |\mu_i| = 1 \tag{5.33}$$

Proof. The conditions under which Proposition 2.1 is validated are satisfied by the data of $(\bar{P}_i)$. Applying the proposition, we deduce the existence of an absolutely continuous function $(p_i, r_i) : [0,1] \to \mathbb{R}^{n+1}$, an integrable function $\xi_i : [0,1] \to \mathbb{R}^n$, and a scalar $\lambda_i \geqslant 0$ such that

$$||(p_i, r_i)||_{L_\infty} + \lambda_i > 0, \tag{5.34}$$

$$(-\dot{p}_i(t), -\dot{r}_i(t), \dot{x}_i(t), 0, \xi_i(t)) \in \tag{5.35}$$

$$\text{co}\, \partial_{x,y,p,r,u} \tilde{H}_{\lambda_i}(t, x_i(t), 0, p_i(t), r_i(t), v_i(t)) \quad \text{a.e. } t \in [0,1],$$

$$\xi_i(t) \in \text{co}\, N_{U_\varepsilon(t)}(v_i(t)), \tag{5.36}$$

and

$$\begin{aligned} (p_i(0), r_i(0), -p_i(1), -r_i(1)) \in & N_{\tilde{C}}(x_i(0), 0, x_i(1), 0) \\ & + \lambda_i(0,0,0,\sqrt{\varepsilon_i}) \\ & + \lambda_i \partial \tilde{g}(x_i(0), 0, x_i(1), 0). \end{aligned} \tag{5.37}$$

where

$$\tilde{C} = C_0 \times \{0\} \times C_1 \times \mathbb{R},$$

the function $\tilde{g} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \to \Re$ is defined as

$$\tilde{g}(x_0, y_0, x_1, y_1) := g(x_0, x_1)$$

and

$$\tilde{H}_\lambda(t, x, y, p, r, u) := p \cdot f(t, x, u) + r\|u - v_i\| - \lambda L(t, x, u) - \lambda k_i h^+(t, x).$$

It follows from (5.35) and (5.37) that r_i is constant and equal to $r_i(t) = -\lambda_i \varepsilon_i^{\frac{1}{2}}$, for every $t \in [0, 1]$.

Recall that

$$H_\lambda(t, x, p, u) = p \cdot f(t, x, u) - \lambda L(t, x, u).$$

By the sum rule [4] we get

$$\operatorname{co} \partial_{x,y,p,r,u} \tilde{H}_{\lambda_i}(t, x, y, p, r, u)$$

$$\subseteq \operatorname{co} \partial_{x,y,p,r,u} H_{\lambda_i}(t, x, p, u) + \operatorname{co} \partial_{x,y,p,r,u} [r\|u - v_i\|] - \operatorname{co} \partial_{x,y,p,r,u} \lambda_i k_i h^+(t, x)$$

$$= \operatorname{co} \partial_{x,y,p,r,u} H_{\lambda_i}(t, x, p, u) + \{0\} \times \{0\} \times \{0\} \times \{\|u - v_i\|\} \times r \operatorname{co} \partial_u \|u - v_i\| -$$

$$\operatorname{co} \partial_{x,y,p,r,u} \lambda_i k_i h^+(t, x)$$

The condition (5.35) can now be written as

$$\begin{aligned} (-\dot{p}_i(t), \dot{x}_i(t), \xi_i(t)) \in \operatorname{co} \partial_{x,p,u} H_{\lambda_i}(t, x_i(t), p_i(t), v_i(t)) + \\ \{0\} \times \{0\} \times \{-\lambda_i \sqrt{\varepsilon_i} \rho_i\} - \operatorname{co} \partial_{x,p,u} \lambda_i k_i h^+(t, x_i(t)), \end{aligned} \tag{5.38}$$

where ρ_i belongs to the closed unit ball B in $\Re^m$.

The max rule implies that

$$\operatorname{co} \partial_x h^+(t, s) \subset \begin{cases} \{\gamma\sigma : \ \gamma \in \operatorname{co} \partial_x h(t, s), \ \sigma \in [0, 1]\} & \text{for } h(t, s) = 0 \\ \{0\} & \text{for } h(t, s) < 0 \end{cases}$$

Define the multifunctions

$$\Sigma_i(t) = \begin{cases} [0, 1] & \text{for } h(t, x_i(t)) = 0 \\ \{0\} & \text{otherwise,} \end{cases}$$

$$\Omega_i(t) = \operatorname{co} \partial_{x,p,u} H_{\lambda_i}(t, x_i(t), p_i(t), v_i(t)) \times B \times \operatorname{co} \partial_{x,p,u} h(t, x_i(t)) \times \Sigma_i(t).$$

Given $\eta = (\Psi, \rho, \Gamma, \sigma)$, define

$$\mathcal{H}_i(t, \eta) = \Psi - \lambda_i \sqrt{\varepsilon_i}(0, 0, \rho) - \lambda_i k_i \Gamma \sigma.$$

It now follows from (5.38) that

$$(-\dot{p}_i(t), \dot{x}_i(t), \xi_i(t)) \in \{\mathcal{H}_i(t, \eta) : \eta \in \Omega_i(t)\}, \text{ a.e.}$$

Observe that the function $\mathcal{H}_i(t, \eta)$ is measurable in t and continuous in η and that $\Omega_i(t)$ takes compact values and is measurable.

Appealing to a measurable selection theorem we deduce existence of measurable functions $\eta_i = (\Psi_i, \rho_i, \Gamma_i, \sigma_i)$ such that

$$\eta_i(t) \in \Omega_i(t)$$

and

$$(-\dot{p}_i(t), \dot{x}_i(t), \xi_i(t)) = \mathcal{H}_i(t, \eta_i(t)) \quad \text{a.e. } t \in [0, 1]. \tag{5.39}$$

The sum rule and the properties of H_{λ_i} allow us to express Ψ_i as in the lemma.

At this stage it is convenient to introduce the nonnegative measure μ_i in $C^*([0, 1]; \mathbb{R})$ such that

$$\int_B d\mu_i(t) = \int_B \lambda_i k_i \sigma_i(t) dt,$$

for any Borel set B. Then, in terms of $d\mu_i$, we can write

$$\int_{[0,t)} \lambda_i k_i \Gamma_i(t) \sigma_i(t) dt = \int_{[0,t)} \Gamma_i(t) d\mu_i(t).$$

If we consider π_i to be the measure defined by

$$d\pi_i(t) = \dot{p}_i(t) dt$$

the conclusions of Lemma 5.1 follow from (5.34) (we normalize the multipliers to get the final form (5.33)), (5.36), (5.37) and (5.39). □

Step 4: We now consider $k_i \to \infty$ and we obtain the necessary conditions for the original problem P.

Standard arguments involving the Gronwall's inequality assert the existence of a constant K_1 such that

$$|\pi_i| \leqslant K_1.$$

It follows from (5.31) and (5.33) that

$$|p_i(t)| \leqslant 1 + K_1$$

for all $t \in [0, 1]$ and all i. By extraction of subsequence, if necessary, we obtain

$$\pi_i \to \pi \qquad \text{weakly}^*$$

for some measure π. The weak*-lower semicontinuity of the dual norm allows us to conclude that

$$|\pi| \leqslant K_1.$$

By (5.33) we may also arrange that

$$\mu_i \to \mu \ \text{ weakly}^*, \quad b_i \to b \quad \text{and} \quad \lambda_i \to \lambda \tag{5.40}$$

for some μ, b and λ. We deduce that $|\mu_i| \to |\mu|$. By (5.33), we further deduce that

$$|b| + |\lambda| + |\mu| = 1.$$

The convergence of b_i to b together with Lemma 4.3 of [3] guarantee that, for some subsequence,

$$p_i(t) \to q(t) \qquad a.e.,$$

where q is a function of bounded variation given by

$$q(t) = b + \int_{[0,t)} d\pi,$$

and

$$b_i + \int_{[0,1]} d\pi_i \to b + \int_{[0,1]} d\pi \tag{5.41}$$

By H1, H4, (5.25) and (5.26), we have for some integrable function k

$$|\Phi_i(t)|, |\Lambda_i(t)| \leqslant k(t), \tag{5.42}$$

and

$$|\xi_i(t) + \lambda_i \sqrt{\varepsilon_i} \rho_i(t)| \leqslant k(t), \tag{5.43}$$

for i sufficiently large. Extracting a subsequence, if necessary, we may arrange that

$$\Phi_i \to \Phi, \quad \Lambda_i \to \Lambda, \quad \Psi_i \to \Psi \tag{5.44}$$

for some $\Phi, \Lambda, \Psi \in L^1$.

Appealing to the upper semicontinuity of limiting subdifferential we can write:

$$\Phi \in \operatorname{co} \partial_{x,p,u} f(t, \bar{x}(t), \bar{u}(t)) \quad a.e.,$$

$$\Lambda \in \operatorname{co} \partial_{x,p,u} L(t, \bar{x}(t), \bar{u}(t)) \quad a.e.$$

and

$$\Psi \in \operatorname{co} \partial_{x,p,u} H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \quad a.e.$$

Since $\lambda_i \to \lambda$, $\rho_i(t) \in B$ and $\varepsilon_i \to 0$, we may extract a subsequence, if necessary, such that

$$\xi_i \to \xi \tag{5.45}$$

with respect to the weak L^1 topology, for some $\xi \in L^1$. By the properties of the limiting normal cone and since $v_i(t) \to \bar{u}(t)$ almost everywhere, we conclude from (5.29), (5.43) and (5.45) that

$$\xi(t) \in \operatorname{co} N_{U_\varepsilon(t)}(\bar{u}(t)) \quad \text{a.e.} \tag{5.46}$$

Since $\partial_x h(t,x) \subset \bar{\partial}_x h(t,x)$ and $\bar{\partial}_x h(t,x)$ is of closed graph for any i, using Lemma 4.5 of [3] we deduce from (5.27) the existence of a Borel measurable, μ-integrable function $\Gamma = (\gamma^1, 0, 0)$ such that

$$\gamma^1(t) \in \bar{\partial}_x h(t, \bar{x}(t)) \quad \mu - \text{a.e.}\ .$$

Taking into account the fact that the multifunctions $s \to N_{C_0}(s)$ and $s \to N_{C_1}(s)$ have closed graph and by the properties of the subdifferential ∂g, we also deduce that (5.32) holds with the i's deleted.

We now turn to the support of μ. Seeking a contradiction let us assume that

$$\mu\big(\{t : \ h(t, \bar{x}(t) < 0\}\big) > 0.$$

The σ additivity of the measure implies the existence of an $e > 0$ such that $\mu(E) > 0$ for $E = \{t : \ h(t, \bar{x}(t)) < -e\}$. By (5.24) and H4,

$$\{t : \ h(t, \bar{x}(t)) < -e\} \subset \{t : \ h(t, x_i(t)) \leqslant -e/2\}$$

for sufficiently large i. But $\{t : \ h(t, x_i(t)) \leqslant -e/2\}$ is contained in the complement of the support of μ_i, whence $\mu_i(E) = 0$. Since h is continuous with respect to t, we conclude that E is an open set. Recall that μ_i converges weakly* to μ. Thus $\mu(E) = 0$, a contradiction. We have just shown that

$$\mu\big(\{t : \ h(t, \bar{x}(t)) < 0\}\big) = 0. \tag{5.47}$$

Lemma 4.3 of ([3]) allow us to write equation (5.30) with i's deleted and we have

$$\begin{aligned} -q(t) + b &= \int_{[0,t)} \psi^1(t)\, dt - \int_{[0,t)} \gamma^1(t)\, d\mu(t) \\ &= \int_{[0,t)} \big(q(t)\alpha^1(t) - \lambda\beta^1(t)\big)\ dt - \int_{[0,t)} \gamma^1(t)\, d\mu(t) \end{aligned}$$

Defining the function

$$p(t) := q(t) - \int_{[0,t)} \gamma_1(s) d\mu(s)$$

which is absolutely continuous, we obtain,

$$-\dot{p}(t) = q(t)\alpha^1(t) - \lambda\beta^1(t).$$

We also have

$$\begin{aligned} \xi(t) &= q(t)\alpha^3(t) - \lambda\beta^3(t) \\ \dot{\bar{x}}(t) &= f(t, \bar{x}(t), \bar{u}(t)). \end{aligned}$$

Then we can write

$$(-\dot{p}(t), \dot{\bar{x}}(t), \xi(t)) = \Psi(t) \in \text{co}\, \partial_{x,p,u} H_\lambda(t, \bar{x}(t), q(t), \bar{u}(t)) \ a.e. \tag{5.48}$$

The necessary conditions of Theorem 3.1 are then established.

Step 5: Finally, we show that H5 implies the interim hypothesis IH:

$$\lim_{k_i \to \infty} \inf\{P_i\} = \inf\{P\}.$$

Observe that this was essential to get (5.21).

Let be $k_i \to \infty$ and choose a corresponding sequence of admissible processes (x_i, u_i) for P_i such that

$$\int_0^1 L(t, x_i(t), u_i(t))\, dt + g(x_i(0), x_i(1)) + k_i \int_0^1 h^+(t, x_i(t))\, dt \leqslant \inf\{P_i\} + \frac{1}{k_i} \tag{5.49}$$

Define the multifunction

$$F(t,x) = \{(v_1, v_2) \in \mathbb{R} \times \mathbb{R}^n : (v_1, v_2) = (L(t,x,u), f(t,x,u)),\ u \in U_\varepsilon(t)\}.$$

By H1, H2 and H5, this multifunction is closed, convex, measurably Lipschitz and integrably bounded, (see definitions in [4]), on

$$\Omega = \{(t,x) \in \mathbb{R} \times \mathbb{R}^n : t \in [0,1],\ x \in \bar{x}(t) + \varepsilon B\}$$

Applying [4, Theorem 3.1.7] we deduce the existence of a subsequence of $\{(z_i, x_i)\}$, where $z_i(t) = \int_0^t L(s, x_i(s), u_i(s))ds$, which converges uniformly to an arc (z, x) such that

$$(\dot{z}(t), \dot{x}(t)) \in F(t, x(t))$$

and $x(0) \in C_0$, $z(0) = 0$. An appropriate measurable selection theorem asserts the existence of a measurable function u such that $u(t) \in U_\varepsilon(t)$ and

$$\dot{z}(t) = L(t, x(t), u(t)) \quad a.e.$$

$$\dot{x}(t) = f(t, x(t), u(t)) \quad a.e.$$

Furthermore,

$$z(1) = \int_0^1 L(t, x(t), u(t))\, dt.$$

Since

$$x_i \to x$$

uniformly and C_1 is closed we further deduce that

$$x(1) \in C_1.$$

As $k_i \to \infty$, $\int_0^1 L(t, x_i(t), u_i(t))\, dt + g(x_i(0), x_i(1)) + k_i \int_0^1 h^+(t, x_i(t))\, dt$ is bounded, and $t \mapsto h^+(t, x(t))$ is continuous (by H4) we have

$$\int_0^1 h^+(t, x(t))\, dt = \lim_{i\to 0} \int_0^1 h^+(t, x_i(t))\, dt = 0. \tag{5.50}$$

Suppose that there exists a $\bar{t} \in [0, 1]$ such that

$$h(\bar{t}, x(\bar{t})) > 0.$$

Then there exists an interval $(\bar{t} - \delta, \bar{t} + \delta)$ (or $[\bar{t}, \bar{t} + \delta)$ if $\bar{t} = 0$, or $(\bar{t} - \delta, \bar{t}]$ if $\bar{t} = 1$), for some $\delta > 0$ such that

$$h(t, x(t)) > 0$$

and

$$\int_{\bar{t}-\delta}^{\bar{t}+\delta} h^+(t, x(t)) dt > 0$$

contradicting (5.50). This means that (x, u) is an admissible process of (P) such that

$$x(t) \in \bar{x}(t) + \varepsilon B.$$

It follows that

$$\int_0^1 L(t, x(t), u(t))\, dt + g(x(0), x(1)) \geqslant \int_0^1 L(t, \bar{x}(t), \bar{u}(t))\, dt + g(\bar{x}(0), \bar{x}(1)).$$

From (5.49)

$$\begin{aligned}
\liminf\{Pi\} &\geqslant \lim\left[\int_0^1 L(t,x_i(t),u_i(t))\,dt + g(x_i(0),x_i(1)) + k_i\int_0^1 h^+(t,x_i(t))\,dt\right] - 0\\
&\geqslant \lim\int_0^1 L(t,x_i(t),u_i(t))\,dt + g(x_i(0),x_i(1))\\
&= \int_0^1 L(t,x(t),u(t))\,dt + g(x(0),x(1))\\
&\geqslant \inf\{P\}
\end{aligned}$$

As $(\bar{x}(t),\bar{u}(t))$ is always an admissible process for problem (P_i) we also have,

$$\inf\{P_i\} \leqslant \inf\{P\}$$

and the same inequality is valid in the limit.

We conclude that

$$\liminf\{Pi\} = \inf\{P\},$$

proving the theorem.

6. Remarks

To get necessary conditions for optimal control problems with state constraints it is usually required that the state constraint function is merely upper semicontinuous as a function of t. Within this framework several results involving state constraints may be deduced as it can be verified in the literature. However, the approach we use in the present work does not allow such framework. In fact, when $t \to h(t,x)$ is assumed to be only upper-semicontinuous we cannot guarantee that

$$\lim_{i\to\infty}\inf\{P_i\} = \inf\{P\}.$$

The following example illustrates this statement.

Consider the following problem.

$$(P2) \quad \text{Minimize} \quad \int_0^1 -x(t)\,dt$$

subject to

$$\begin{aligned} &\dot{x}(t) = u(t) && \text{a.e. } t \in [0,1] \\ &x(0) \in [-2, 0] && \\ &u(t) \in [-1, 0] && \text{a.e. } t \in [0,1] \\ &h\left(t, x(t)\right) \leqslant 0 && \text{for all } t \in [0,1], \end{aligned}$$

where, for $\mathcal{A} = \{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$,

$$h(t,x) = \begin{cases} x & \text{if } t \in [0,1] \setminus \mathcal{A} \\ 1 + x & \text{if } t \in \mathcal{A} \cap [0,1]. \end{cases}$$

In this case we have

$$\begin{aligned} \{(t,x) : h(t,x) \leqslant 0\} = &\{(t,x) : x \leqslant 0 \text{ and } t \in [0,1] \setminus \mathcal{A}\} \\ &\cup \{(t,x) : x \leqslant -1 \text{ and } t \in \mathcal{A} \cap [0,1]\}. \end{aligned}$$

The data of this problem satisfies H1–H3 and H5 but $t \to h(t,x)$ fails to be continuous. Indeed, $x \to h(t,x)$ is Lipschitz continuous for each $t \in [0,1]$ and $t \to h(t,x)$ is only upper-semicontinuous. As a function of t, h is continuous on $[0,1] \setminus \mathcal{A}$ and for $\bar{t} \in \mathcal{A} \cap [0,1]$, $h(\bar{t}, x) = 1 + x \geqslant \limsup_{t \to \bar{t}} h(t,x)$.

It is easy to see that if a trajectory x of problem $(P2)$ is such that $h(t, x(t)) \leqslant 0$, for all $t \in [0,1]$, then it must satisfy $x(0) \leqslant -1$. In view of the cost function it is desirable to have $x(t)$ as "large" as possible. But the admissible control values are non-positive.

The optimal solution of problem $(P2)$ is then

$$\bar{x}(t) \equiv -1, \quad \bar{u}(t) \equiv 0,$$

and $\inf\{P\} = 1$.

Consider $\varepsilon \in (0,1)$ and the ε-neighborhood of $\bar{x}$. The sequence of problems (P_i) is now:

$$(P_i) \quad \text{Minimize} \quad \int_0^1 -x(t)\,dt + k_i \int_0^1 h^+(t,x(t))\,dt$$

subject to

$$\begin{aligned} &\dot{x}(t) = u(t) && \text{a.e. } t \in [0,1] \\ &x(0) \in [-2,0] && \\ &u(t) \in [-1,0] && \text{a.e. } t \in [0,1] \end{aligned}$$

Take the constant sequence of arcs $x_i(t) \equiv -1 + \varepsilon/4$, $t \in [0,1]$. Then,

$$h(t,x_i(t)) = h(t,-1+\varepsilon/4) = \begin{cases} -1+\varepsilon/4 & t \in [0,1] \setminus \mathcal{A} \\ \varepsilon/4 & t \in \mathcal{A} \cap [0,1] \end{cases}$$

and

$$h^+(t,x_i(t)) = \begin{cases} 0 & t \in [0,1] \setminus \mathcal{A} \\ \varepsilon/4 & t \in \mathcal{A} \cap [0,1] \end{cases}$$

Since each trajectory x_i is admissible for problem (P_i) we have

$$\inf\{P_i\} \leqslant \int_0^1 (1-\varepsilon/4)\,dt + k_i \int_0^1 h^+(t,x_i(t))\,dt = 1 - \varepsilon/4 < 1$$

As we have seen before that $\inf\{P\} = 1$. So

$$\inf\{P_i\} < \inf\{P\} - \varepsilon/4.$$

Acknowledgements

The authors wish to thank Richard B. Vinter for helpful discussions and suggestions. The financial support from Projecto FCT PRAXIS C/EEI/14156/1998 is gratefully acknowledged.

References

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

[2] R. B. Vinter. *Optimal Control*. Birkhouser, 2000.

[3] R. B. Vinter and G. Pappas. A maximum principle for nonsmooth optimal-control problems with state constraints. *Journal of Mathematical Analysis and Applications*, 89:212–232, 1982.

[4] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983. Republished as vol.5 of Classics in Applied Mathematics, SIAM, 1990.

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

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

[7] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2):181–218, 1995.

[8] L. W. Neustadt. *Optimization, A Theory of Necessary Conditions*. Princeton University Press, 1976.

[9] F. H. Clarke. The maximum principle under minimal hypotheses. *SIAM Journal of Control and Optimization*, 14(6):1078–1091, 1976.