

NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

ADJOINT DIFFERENTIAL INCLUSIONS IN
NECESSARY CONDITIONS FOR THE MINIMAL
TRAJECTORIES OF DIFFERENTIAL INCLUSIONS

Halina Frankowska

August 1984
CP-84-37

Collaborative Papers report work which has not been performed solely at the International Institute for Applied Systems Analysis and which has received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria



PREFACE

This paper extends Pontryagin's maximum principle to differential inclusions and nonsmooth criterion functions, relying on a checkable "surjectivity property" of a "linearized set-valued system" around the optimal trajectory. As an example, Pontryagin's principle is obtained for optimal control problems with constraints on both the initial and the final states.

The research described here was undertaken within the framework of the Dynamics of Macrosystems Feasibility Study in the System and Decision Sciences Program.

ANDRZEJ WIERZBICKI
Chairman
System and Decision
Sciences Program



CONTENTS

Introduction	1
1. Asymptotic Tangent Cone and Asymptotic Differential of a Set-Valued Map	6
2. The Differential Inclusion Problem	12
3. The Abstract Problem	15
4. An Example	21
5. Infinite Horizon Problem	26
REFERENCES	31



ADJOINT DIFFERENTIAL INCLUSIONS
IN NECESSARY CONDITIONS FOR THE MINIMAL
TRAJECTORIES OF DIFFERENTIAL INCLUSIONS

Halina Frankowska

Introduction

Consider the following control system:

$$(0.1) \quad \begin{cases} \dot{x} = f(x, u(t)) \\ u(t) \in U(x) \end{cases}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ and $U : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a set-valued map.

Let $g : \mathbb{R}^{2n} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a given function, and consider the problem of minimizing $g(x(0), x(1))$ over the set K of solutions x to (0.1)

Let z be a minimizer

If U does not depend on x and if the data are smooth enough the Pontriagin maximum principle [15] tells us that for some absolutely continuous function $p : [0, 1] \rightarrow \mathbb{R}^n$ the following holds:

$$(0.2) \quad \begin{cases} -\dot{p}(t) = \left[\frac{\partial f}{\partial x} (z(t), \bar{u}(t)) \right]^* p(t) \\ \langle p(t), f(z(t), \bar{u}(t)) \rangle = \max_{u \in U} \langle p(t), f(z(t), u) \rangle \end{cases}$$

$$(0.3) \quad (-p(0), p(1)) = g'(z(0), z(1))$$

where \bar{u} is the control associated with z and $\left[\frac{\partial f}{\partial x}(z(t), \bar{u}(t))\right]^*$ denotes the transpose of the Jacobian of f with respect to x at $(z(t), \bar{u}(t))$. The case of control problem with constraints bearing on initial and final states can be embedded in the above framework, when g is no longer smooth but just a function taking also infinite values. This is a first motivation to tackle the non-smooth case.

A series of papers took the issue of adapting this result to the case of non-smooth functions by using one or another of the many generalized gradients (see Clarke [8] for instance).

To study the necessary conditions in a more general case we have to consider the set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by

$$F(x) = \{f(x, u) : u \in U(x)\}$$

and an associated differential inclusion

$$(0.4) \quad \dot{x} \in F(x)$$

Under some measurability assumptions on f it can be shown that the solutions to (0.1) and (0.4) coincide. So, to get a characterization of z , we can just study (0.4).

Such an approach to optimal control problem was first proposed by Ważewski in [21] and has been the object of consideration by many authors. See for example, Aubin-Clarke [2], Blagodatskich [4], Clarke [8], Frankowska-Olech [11], Ioffe [12], Lasry-Berliocchi [13], Rockafellar [16].

The question arises naturally how to formulate a maximum principle for an optimal trajectory of such a differential inclusion.

For obtaining results similar to (0.2), (0.3) in the set-valued case we need a notion generalizing the derivative and its transpose to a set-valued map $F : E \rightrightarrows E_1$, where E, E_1 are Banach spaces.

For that purpose we shall adopt the geometric point of view. When F is a smooth function, the graph of its derivative is the tangent space to the graph of the function. In the case of a non-smooth function or a set-valued map F , we need to define a tangent cone to the graph to be able to use the same strategy. Many candidates for the role of tangent cone to a set have been proposed; let me mention the contingent cone, introduced by Bouligand in the early thirties, or the tangent cone introduced by Clarke [5] in 1975.

But whatever the choice of a "tangent cone" $T_K(x)$ to a subset K at a point $x \in K$ is, we can use it to define the derivative of a set-valued map F at a point (x,y) of its graph.

Let $T_{\text{graph}(F)}(x,y)$ be the chosen tangent cone and let us call its negative polar cone the normal cone to K at (x,y) , and denote it by $N_{\text{graph}(F)}(x,y) := T_{\text{graph}(F)}(x,y)^{\bar{}}$. Then the derivative $DF(x,y)$ of F at (x,y) is the set valued map from E to E_1 defined by

$$v \in DF(x,y)(u) \quad \text{iff} \quad (u,v) \in T_{\text{graph}(F)}(x,y)$$

and the co-derivative $DF(x,y)^*$ of F at (x,y) is the set-valued map from E_1^* to E^* defined by

$$r \in DF(x,y)^*(q) \quad \text{iff} \quad (r,-q) \in N_{\text{graph}(F)}(x,y)$$

which can be regarded as the transpose of $DF(x,y)$, (see a survey in Chapter 7 of the book by Aubin-Ekeland [3]).

We define also a generalized gradient of a real-valued function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ at $x \in \text{Dom}(f)$ which we denote by $\partial f(x)$. The necessary conditions then take the following form:

There exists an absolutely continuous function $p : [0,1] \rightarrow \mathbb{R}^n$ satisfying the following conditions:

$$(0.2)' \quad -\dot{p}(t) \in DF(z(t), \dot{z}(t))^*(p(t))$$

$$(0.3)' \quad (-p(0), p(1)) \in \partial g(z(0), z(1))$$

The objective of this paper is twofold. The first one is to derive inclusions (0.2)' and (0.3)' using a suitable concept of tangent cone such that the associated notions of co-derivative and generalized gradient are reasonably small. The second one is related to "calmness assumption" introduced by Clarke (see [6], [7], [8]). We replace it by a "surjectivity assumption" which states that the "linearized problem" around the optimal solution is solvable. This is a checkable assumption: we shall illustrate this point in Section 4 as we apply our approach to optimal control with constraints bearing both on the initial state and the final state. In this example "calmness" seems to be harder to verify. The intermediate tangent cone plays an important role in this paper. This is due namely to the fact that we can "compute" the intermediate tangent cone to the set of solutions to the differential inclusion (0.4) as the cone of solutions w to the "linearized" differential inclusion

$$w'(t) \in dF(z(t), \dot{z}(t))(w(t))$$

where $dF(z(t), \dot{z}(t))$ denotes the intermediate derivative of F at $(z(t), \dot{z}(t))$ (see Frankowska [10]). This is the reason why we cannot avoid using it for solving our type of problem. It enjoys also other interesting properties: in particular, the associated generalized gradient $\partial f(x)$ is smaller than Clarke's generalized gradient and has the following property: If f is Fréchet differentiable at x , then $\partial f(x)$ reduces to $f'(x)$ (whereas we require that f is regularly differentiable for the Clarke generalized gradient to reduce to $f'(x)$).

The choice of a tangent cone is analogous to the choice of an adequate concept of derivative: it depends upon the problem at hand. Let us mention only that the contingent derivative (cf. Aubin-Ekeland [3]) is a generalization of the Gâteaux derivative, the intermediate derivative--a generalization of the Fréchet derivative and the Clarke derivative--a generalization of a continuous Fréchet derivative.

In general, the intermediate tangent cone is not convex. In many applications the convexity is required. Our results can be

formulated with different convex subcones of the intermediate tangent cone (one among possible candidates is the Clarke tangent cone, which is always convex and contained in the intermediate tangent cone). To fix the ideas we shall choose one particular subcone, the asymptotic tangent cone, which contains the Clarke cone and coinciding with the intermediate cone when the latter is convex.

The reader used to Clarke's notion of tangency may replace the notions of asymptotic derivative, co-derivative and gradient in Theorem 2.3 by those obtained through Clarke's definition, to get the same kind of results.

In Section 4 we give an example of a problem with initial and end point constraints and study the surjectivity assumption in this case. In this example "calmness" seems to be harder to verify.

Our results can be applied also to the study of the "generalized Bolza problem", exactly in the same way as it was done by Clarke in [6]. Under some "reasonable" assumptions the generalized Bolza problem can be written in differential inclusion form (cf. [6]). Then the necessary conditions from Theorem 2.3 can be expressed in terms of the generalized Euler-Lagrange equation for the Lagrangian.

We devote the first section to a presentation of the asymptotic tangent cone. Section 2 deals with the necessary conditions satisfied by an optimal solution to a differential inclusion problem. We state the main result and begin the proof, which reduces this problem to an abstract optimization problem. This general problem is then studied in the third section. In the fourth section we give an example of the application of the main theorem. In the fifth section, we apply the method to a non-convex infinite-horizon problem, and extend to this case results of Aubin-Clarke [2].

The author wishes to thank Ivar Ekeland for the many suggestions and advice which made the presentation of this paper much clearer.

1. Asymptotic Tangent Cone and Asymptotic Differential of a Set-Valued Map

Let E be a Banach space. We denote by $\overset{\circ}{B}$ the open unit ball in E and by $\langle \cdot, \cdot \rangle$ the canonical bilinear form on $E^* \times E$.

Consider a subset $K \subset E$ and a point $x \in K$. There exist in the literature different notions of "tangent cones" to K at x . We recall in particular the definitions of the contingent cone

$$(1.1) \quad T_K(x) := \bigcap_{\substack{\varepsilon > 0 \\ \delta > 0}} \bigcup_{h \in]0, \delta[} \left[\frac{1}{h}(K-x) + \varepsilon \overset{\circ}{B} \right]$$

(see Aubin-Ekeland [3], Chapter 7).

the tangent cone in the sense of Clarke

$$(1.2) \quad C_K(x) := \bigcap_{\substack{\varepsilon > 0 \\ \rho > 0}} \bigcup_{\substack{\delta > 0 \\ h \in]0, \delta[}} \bigcap_{\substack{x' \in B(x, \rho) \cap K \\ h \in]0, \delta[}} \left[\frac{1}{h}(K-x') + \varepsilon \overset{\circ}{B} \right]$$

(see Clarke [8], Aubin-Ekeland [3], Chapter 7)

and the intermediate tangent cone

$$(1.3) \quad I_K(x) := \bigcap_{\varepsilon > 0} \bigcup_{\delta > 0} \bigcap_{h \in]0, \delta[} \left[\frac{1}{h}(K-x) + \varepsilon \overset{\circ}{B} \right]$$

(see Ursescu [19]).

The relations (1.1), (1.2), (1.3) can be written in terms of the Kuratowski $\lim \sup$ and $\lim \inf$ in the following way:

$$(1.1)' \quad T_K(x) = \limsup_{h \rightarrow 0_+} \frac{1}{h}(K-x)$$

$$(1.2)' \quad C_K(x) = \liminf_{\substack{x' \rightarrow x \\ h \rightarrow 0_+}} \frac{1}{h}(K-x')$$

$$(1.3)' \quad I_K(x) = \liminf_{h \rightarrow 0_+} \frac{1}{h}(K-x)$$

All the above sets are closed cones satisfying $C_K(x) \subset I_K(x) \subset T_K(x)$. Moreover, $C_K(x)$ is convex. For further properties of $C_K(x)$, $T_K(x)$ see [3], [5], [8], [17]. The cone $I_K(x)$ is less known. It can be also characterized by using the distance function.

(1.4) Proposition: Let $d_K(y)$ denote the distance of $y \in E$ to K . Then

$$I_K(x) = \{v \in E : \lim_{h \rightarrow 0^+} \frac{d_K(x + hv)}{h} = 0\}$$

Consider a function $f : E \rightarrow R \cup \{+\infty\}$ and let $\text{epi}(f)$ denote the epigraph of f . As an important example we shall study the set $I_{\text{epi}(f)}(x, f(x))$. We recall first:

(1.5) Definition: For $Q : R \times E \rightarrow R \cup \{+\infty\}$ set

$$\limsup_{h \rightarrow 0^+} \inf_{u' \rightarrow u} Q(h, u') : = \sup_{\varepsilon > 0} \inf_{\delta > 0} \sup_{h \in]0, \delta[} \inf_{u' \in u + \varepsilon B} Q(h, u')$$

(see Rockafellar [17]).

Let us introduce the following

(1.6) Definition: For $f : E \rightarrow R \cup \{+\infty\}$, $x \in \text{Dom}(f)$

$$i_+ f(x)(u) : = \limsup_{h \rightarrow 0^+} \inf_{u' \rightarrow u} \frac{f(x + hu') - f(x)}{h}$$

(1.7) Proposition: Let f, x be as in (1.6); then

$$I_{\text{epi}(f)}(x, f(x)) = \text{epi}(i_+ f(x))$$

Proof: Let $K = \text{epi}(f)$ and $(u, v) \in \text{epi}(i_+ f(x))$. Then for all $\varepsilon > 0$ and all small $h > 0$ there exist $u_h \in u + \varepsilon B$

such that $f(x + hu_h) - f(x) \leq h i_+ f(x)(u) + h\varepsilon \leq h(v + \varepsilon)$. It implies that for all small $h > 0$ $(x, f(x)) + h(u_h, v + \varepsilon) \in K$. Thus by (1.4) $(u, v) \in I_K(x, f(x))$. Conversely, if $(u, v) \in I_K(x, f(x))$ then by (1.3) for all $\varepsilon > 0$ there exist $\delta > 0$ such that for any $h \in]0, \delta[$ we have $(u, v) \in \frac{1}{h}[K - (x, f(x))] + \varepsilon \overset{a}{B}$. This means $(x, f(x)) + h(u, v) \in K + \varepsilon \overset{a}{B}$ and hence

$$\inf_{u' \in B(u, \varepsilon)} \frac{f(x + hu') - f(x)}{h} \leq v + \varepsilon. \text{ Therefore,}$$

$$i_+ f(x)(u) \leq v \quad \text{and} \quad (u, v) \in \text{epi}(i_+ f(x)).$$

The function $i_+ f(x)(\cdot)$ is lower semicontinuous and positively homogeneous.

In the study of some non-smooth problems we are often led to deal with convex tangent cones and convex functions. We shall now define one of them, which is the one we shall be using subsequently:

(1.8) Definition: The asymptotic tangent cone to a subset K at $x \in K$ is given by

$$I_K^a(x) = \{u \in I_K(x) : u + I_K(x) \subset I_K(x)\}$$

Its negative polar, given by

$$N_K^a(x) = \{p \in E^* : \langle p, v \rangle \leq 0 \text{ for all } v \in I_K^a(x)\}$$

is called the asymptotic normal cone to K at x .

(1.9) Remark: $I_K^a(x)$ is closed convex cone. One can easily verify the following relation

$$C_K(x) \subset I_K^a(x) \subset I_K(x) \subset T_K(x)$$

The cones $I_K(x)$ and $I_K^a(x)$ have been also used in [14].

As it is done in [3] we can define now the derivative to a set-valued map F from E to a Banach space E_1 .

(1.10) Definition: The asymptotic derivative of F at $(x,y) \in \text{graph}(F)$ is a set-valued map $D_a F(x,y) : E \rightrightarrows E_1$ defined by

$$v \in D_a F(x,y)(u) \quad \text{iff} \quad (u,v) \in I_{\text{graph}(F)}^a(x,y)$$

(1.11) Definition: The asymptotic co-derivative of F at $(x,y) \in \text{graph}(F)$ is a set-valued map $D_a F(x,y)^* : E_1^* \rightrightarrows E^*$ defined by

$$r \in D_a F(x,y)^*(q) \quad \text{iff} \quad (r,-q) \in N_{\text{graph}(F)}^a(x,y)$$

for all $q \in E_1^*$. Equivalently:

$$\langle r,u \rangle - \langle q,v \rangle \leq 0 \quad \text{for all} \quad v \in D_a F(x,y)(u)$$

Let $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$, $x \in \text{Dom}(f)$. Define $F(y) = f(y) + \mathbb{R}_+$ for all $y \in E$, i.e. $\text{graph}(F) = \text{epi}(f)$.

(1.12) Definition: The subset

$$\partial_a f(x) = D_a F(x,f(x))^*(1)$$

is called the asymptotic gradient of f at x .

We recall that f is regularly Gâteaux differentiable at $x \in \text{Dom}(f)$ if it has the Gâteaux derivative $f'(x) \in E^*$ and for all $u \in E$

$$\lim_{\substack{u' \rightarrow u \\ h \rightarrow 0_+}} \frac{f(x+hu') - f(x)}{h} = \langle f'(x), u \rangle$$

Observe that if f is as above then by (1.7), $\partial_a f(x)$ is single-valued and

$$\partial_a f(x) = \{f'(x)\}.$$

Remark: The asymptotic gradient is well-defined for Fréchet-differentiable functions. Recall that Clarke's generalized gradient may not be defined for such functions: they have to be regularly differentiable.

(1.13) Definition: For all $u \in E$ set

$$i_+^a f(x)(u) := \inf \{ r : r \in D_a F(x, f(x))(u) \}$$

We obtain from (1.8), (1.11)

(1.14) Proposition: $i_+^a f(x)(u) = \sup_v [i_+ f(x)(u+v) - i_+ f(x)(v)]$

$$\partial_a f(x) = \{ q \in E^* : \langle q, u \rangle \leq i_+^a f(x)(u) \text{ for all } u \in E \}$$

The following proposition is similar to one from [18] concerning the subgradients of convex functions.

(1.15) Proposition: If $f, g : E \rightarrow \mathbb{R} \cup \{+\infty\}$, $f \leq g$ and $x \in E$ is such that $f(x) = g(x) < +\infty$. Then

$$\partial_a f(x) \subset \partial_a g(x)$$

Let W, H, T be Banach spaces and $W \subset H$, $L \in L(W, E)$, $\gamma \in L(W, T)$ be continuous linear operators and let $f : W \rightarrow \mathbb{R} \cup \{+\infty\}$; $g : T \rightarrow \mathbb{R} \cup \{+\infty\}$ be given functions, $F : H \rightarrow E$ be a set-valued map. We denote by K the set of all solutions to the inclusion $Lx \in F(x)$.

(1.16) Lemma: Assume that $z \in K$ provides a finite minimum to the problem:

$$\text{minimize } \{ f(x) + g(\gamma x) : x \in W, Lx \in F(x) \}$$

If γ has a continuous right inverse and f is locally Lipschitzian at z then for all $w \in I_K(z)$ we have

$$i_+ f(z)(w) + i_+ g(\gamma z)(\gamma w) \geq 0 \quad \sim$$

Proof: If $w \in I_K(z)$ then for all $h > 0$ there exists $w_h \in W$ such that $z + hw_h \in K$, $\lim_{h \rightarrow 0+} w_h = w$

Since z minimizes $f + g \circ \gamma$ on K we have

$$i_+(f + g \circ \gamma)(z)(w) \geq 0$$

But
$$i_+[f + g \circ \gamma](z)(w) \leq \limsup_{h \rightarrow 0+} \sup_{w' \rightarrow w} [f(z + hw') -$$

$$- f(z)]/h + \limsup_{h \rightarrow 0+} \inf_{w' \rightarrow w} [g(\gamma z + h\gamma w') - g(\gamma z)]/h$$

and since γ has a continuous right inverse we also have

$$\limsup_{h \rightarrow 0+} \inf_{w' \rightarrow w} [g(\gamma z + h\gamma w') - g(\gamma z)]/h$$

$$= \limsup_{h \rightarrow 0+} \inf_{t \rightarrow \gamma w} [g(\gamma z + ht) - g(\gamma z)]/h = i_+g(\gamma z)(\gamma w)$$

It implies that

$$(1.17) \quad \limsup_{\substack{w' \rightarrow w \\ h \rightarrow 0+}} \frac{f(z + hw') - f(z)}{h} + i_+g(\gamma z)(\gamma w) \geq 0$$

By Lipschitz continuity of f we also have

$$(1.18) \quad i_+f(z)(w) = \limsup_{\substack{w' \rightarrow w \\ h \rightarrow 0+}} \frac{f(z + hw') - f(z)}{h}$$

Adding (1.17) and (1.18) we finally get

$$i_+f(z)(w) + i_+g(\gamma z)(\gamma w) \geq 0$$

Since w is arbitrary, the proof follows.

2. The Differential Inclusion Problem

Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a set-valued map of closed graph. Consider the differential inclusion

$$(2.1) \quad \dot{x} \in F(x)$$

An absolutely continuous (a.c.) function $x : [0,1] \rightarrow \mathbb{R}^n$ is a solution of (2.1) iff

$$\dot{x}(t) \in F(x(t)) \quad \text{a.e.}$$

Let K denote the set of all solutions of (2.1); $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be a Lipschitzian function; $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. Then for some $c > 0$

$$|\varphi(r)| \leq c(1+|r|) \quad \text{for all } r \in \mathbb{R}^n$$

and the functional defined by $f(x) = \int_0^1 \varphi(x(t)) dt$ is finite and Lipschitzian on L^1 . Consider the following problem

$$(2.2) \quad \text{minimize } \{g(x(0), x(1)) + \int_0^1 \varphi(x(t)) dt : x \in K\}$$

(2.3) Theorem: Assume the minimum in (2.2) is finite, and $z \in K$ is a minimizer. Assume that F is Lipschitzian in some neighborhood of $z([0,1])$, for the Hausdorff metric. If the following "surjectivity" assumption holds: for some $p > 1$ and all $u, e \in L^p$ there exists a solution $w \in W^{1,p}([0,1], \mathbb{R}^n)$ of

$$i) \quad \dot{w}(t) \in D_a F(z(t), \dot{z}(t))(w(t) + u(t)) + e(t) \quad \text{a.e.}$$

$$ii) \quad i_+^a g(z(0), z(1))(w(0), w(1)) < \infty$$

Then there exists a function $q \in W^{1,p^*}(0,1), \mathbb{R}^n$ (where $\frac{1}{p} + \frac{1}{p^*} = 1$) such that

$$-\dot{q}(t) \in \partial_a \varphi(z(t)) + D_a F(z(t), \dot{z}(t)) * (q(t)) \quad \text{a.e.}$$

$$(-q(0), q(1)) \in \partial_a g(z(0), z(1))$$

We shall prove the above theorem in several steps.

Proof: First we introduce the following notations.

$$E = L^P((0,1), R^n)$$

$$W = W^{1,P}((0,1), R^n)$$

$$T = R^n \times R^n$$

$$\gamma \in L(W, T) \text{ is a "trace operator", } \gamma(x) = (x(0), x(1))$$

$$L \in L(W, E) \text{ is the operator of differentiation, } Lx = \dot{x}$$

L^* be its transpose.

$$F : E \rightarrow E \text{ be defined by } F(x) = \{y \in E : y(t) \in F(x(t)) \text{ a.e.}\}$$

If $q \in W^{1,P^*}((0,1), R^n)$ then integration by parts gives

$$(2.4) \quad \langle L^*q, w \rangle = \langle (-q(0), q(1)), (w(0), w(1)) \rangle - \int_0^1 \dot{q}w$$

In the new notations, z solves the following problem

$$(2.5) \quad \text{minimize } \{f(x) + g(\gamma x) : x \in W, Lx \in F(x)\}.$$

$$(2.6) \quad \underline{\text{Lemma:}} \text{ For all } w \in W \text{ satisfying}$$

$$(2.7) \quad (w(t), \dot{w}(t)) \in I_{\text{graph } F}(z(t), \dot{z}(t)) \text{ a.e. in } [0,1]$$

we have $w \in I_K(x)$ and

$$i_+ f(z)(w) + i_+ g(\gamma z)(\gamma w) \geq 0$$

Proof: We proved in Frankowska [10] that if $w \in W^{1,1}(0,1)$ satisfies 2.7 then $w \in I_K(z)$. This and Lemma 1.16 complete the proof.

(2.8) Remark: Lemma 2.6 can be viewed as a necessary condition for z to be a minimum. In order to obtain an "adjoint" necessary condition we shall use a separation theorem. At this point we need to use convexity properties. For that purpose, we use a convex sub-cone of $I_K(x)$. We can choose Clarke's tangent cone, but it may be

important to use a larger cone, the asymptotic tangent cone and its related concepts. For this reason, we shall replace

$I_{\text{graph}F}(z(t), \dot{z}(t))$, $i_+^a f(z)$, $i_+^a g(\gamma z)$ by

$I_{\text{graph}F}^a(z(t), \dot{z}(t))$, $i_+^a f(z)$ and $i_+^a g(\gamma z)$ respectively.

(2.9) Lemma: The cone

$$C = \{ (x, y) \in E \times E : y(t) \in D_a F(z(t), \dot{z}(t))(x(t)) \}$$

is closed, convex and $(1 \times L)^{-1} C \subset I_K(z)$

Proof: If $(x_n, y_n) \in C$ and $\lim_{n \rightarrow \infty} x_n = x$, $\lim_{n \rightarrow \infty} y_n = y$ in E then

$(x_n, y_n)(t) \rightarrow (x, y)(t)$ a.e.. Since $I_{\text{graph}(F)}^a(z(t), \dot{z}(t))$ is closed and convex cone C has the same properties. The second claim follows from Lemma 2.6. Let C^- be the negative polar cone to C .

(2.10) Lemma: If an a.c. function $q \in W^{1, P^*}((0, 1), \mathbb{R}^n)$ satisfies

$$(-\dot{q}, -q) \in \partial_a f(z) \times \{0\} + C^-$$

$$(-q(0), q(1)) \in \partial_a g(\gamma z)$$

then q also satisfies the requirement of Theorem 2.3.

Proof: Let $\xi \in \partial_a f(z)$ be such that $(-\dot{q} - \xi, -q) \in C^-$. Assume for a moment that

$$\sup_{\tau \in I_{\text{graph}(F)}^a(z(t), \dot{z}(t))} \langle (-\dot{q}(t) - \xi(t), -q(t)), \tau \rangle = +\infty$$

on a set $U \subset [0, 1]$ of positive measure. Let

$$Q(t) = \{ \tau \in I_{\text{graph}(F)}^a(z(t), \dot{z}(t)) : \langle (-\dot{q}-\xi, -q)(t), \tau \rangle > 0 \} \cap B$$

One can easily verify that the map $t \rightarrow I_{\text{graph}(F)}^a(z(t), \dot{z}(t))$ is measurable. Therefore also the map $t \rightarrow Q(t)$ is measurable. Thus there exists a measurable selection $\sigma(t) \in Q(t)$ on U . Let

$$\bar{\sigma}(t) = \begin{cases} \sigma(t) & \text{for } t \in U \\ 0 & \text{otherwise} \end{cases}$$

Then $\bar{\sigma} \in C$ and $\langle (-\dot{q}-\xi, q), \bar{\sigma} \rangle = \int_0^1 \langle (-\dot{q}-\xi, -q)(t), \bar{\sigma}(t) \rangle dt > 0$ which contradicts the definition of C .

Thus the proof of Theorem 2.3 will be complete if we prove that the assumption of Lemma 2.10 is verified.

This will be shown in the next section where an abstract problem is treated.

3. The Abstract Problem

Consider reflexive Banach spaces W, H, E, T , where $W \subset H$ and continuous linear operators

$$L \in L(W, E) \quad , \quad \gamma \in L(W, T)$$

We are supposing here that the injection $i : W \rightarrow H$ is continuous and that the

"trace property" γ has a continuous right inverse and the kernel W_0 of γ is dense in H

holds true.

We denote by i_0 the restriction of i to W_0 . Let L_0 be the restriction of L to W_0 and L_0^* denotes its transpose. Define

$$E_0^* = \{ p \in E^* : L_0^* p \in H^* \}$$

Thus L_0^* maps E_0^* to H^* .

For the problem considered in Section 2 we have $H = E$,
 $E_0^* = W^{1,p}((0,1), \mathbb{R}^n)$ and $L_0^* q = -\dot{q}$ on E_0^*

Equipped with the graph norm E_0^* is a Banach space. If the "trace property" holds then we have the following abstract Green formula ([1]) which corresponds to integration by part from Section 2:

There is a unique operator $\beta^* \in L(E_0^*, T^*)$ such that for all $u \in W$, $p \in E_0^*$

$$\langle L_0^* p, u \rangle - \langle p, Lu \rangle = \langle \beta^* p, \gamma u \rangle$$

Let $F : H \rightrightarrows E$ be a set-valued map and

$$f : H \rightarrow \mathbb{R} \cup \{+\infty\}$$

$$g : T \rightarrow \mathbb{R} \cup \{+\infty\}$$

be given. Consider the problem

$$\text{minimize } \{ f(x) + g(\gamma x) : x \in W , Lx \in F(x) \}$$

which contains problem (2.5) as a particular case. Let C be a closed convex cone, C^- be its negative polar.

Consider closed convex process

$$G : H \rightrightarrows E , \quad G^* : E^* \rightrightarrows H^*$$

defined by

$$p \in G(q) \quad \text{iff} \quad (q, p) \in C$$

$$r \in G^*(q) \quad \text{iff} \quad (r, -q) \in C^-$$

We denote by K the set of all solutions to the inclusion $Lx \in F(x)$.

(3.1) Theorem: Assume that $z \in K$ provides a finite minimum to the problem (3.1), and that f is locally Lipschitzian at z . Further assume that

$$(1 \times L)^{-1}C \subset I_K(x)$$

and the following surjectivity assumption holds true: for all $u, e \in H \times E$, there exists a solution $w \in W$ to the problem

$$(i) \quad Lw \in G(w+u) + e$$

$$(ii) \quad \gamma w \in \text{Dom } i_+^a g(\gamma z)$$

Then there exists $q \in E_0^*$ such that

$$L_0^* q \in \partial_a f(z) + G^*(q)$$

$$-\beta^* q \in \partial_a g(\gamma z)$$

Proof: Note that the assumptions of Theorem 3.1 imply the assumptions of Lemma 1.16. Therefore, z is also a solution to problem

$$(3.2) \quad i_+^a f(z)(w) + i_+^a g(\gamma z)(\gamma w) \geq 0 \quad \text{for all } (w, Lw) \in C$$

Set

$$\pi(w) := i_+^a f(z)(w) \quad , \quad \psi(t) := i_+^a g(\gamma z)(t)$$

and

$$\Pi := \{p \in H^* : \langle p, w \rangle \leq \pi(w) \text{ for all } w \in H\} = \partial_a f(z)$$

$$\Psi := \{q \in T^* : \langle q, t \rangle \leq \psi(t) \text{ for all } t \in T\} = \partial_a g(\gamma z)$$

The functions π, ψ are lower semicontinuous, convex and positively homogeneous. The sets Π, Ψ are closed and convex.

(3.4) Lemma: Let $\pi : H \rightarrow R \cup \{+\infty\}$, $\psi : T \rightarrow R \cup \{+\infty\}$ be lower semicontinuous, convex, positively homogeneous functions and $C \subset H \times E$ be a closed convex cone, and the sets Π, Ψ be defined as in (3.3). Assume that the following set

$$A := i^* \Pi + \gamma^* \Psi + \{i^* r - L^* q \mid (r, -q) \in C^-\}$$

is closed in W^* . Then the following statements are equivalent:

(1) $\pi(w) + \psi(\gamma w) \geq 0$ for all $w \in W$ of $Lw \in G(w)$

(2) There is $q \in E_0^*$ such that:

$$L_0^* q \in \Pi + G^*(q)$$

$$-\beta^* q \in \Psi$$

Proof: Assume (1) holds. We claim that there is $q \in E^*$ satisfying

$$(3.5) \quad 0 \in i^* \Pi + \gamma^* \Psi + i^* G^*(q) - L^* q$$

Indeed, assume that it does not hold. By reflexivity of W and the separation theorem there is $w \in W$ such that for all $\alpha \in \Pi$, $(r, -q) \in C^-$, $\alpha' \in \Psi$

$$\langle i^* \alpha, w \rangle + \langle \gamma^* \alpha', w \rangle + \langle i^* r, w \rangle - \langle L^* q, w \rangle \leq \rho < 0$$

where $\rho < 0$ is fixed. Hence

$$(3.6) \quad \langle \alpha, w \rangle + \langle \alpha', \gamma w \rangle + \langle r, w \rangle - \langle q, Lw \rangle \leq \rho < 0$$

for all $\alpha \in \Pi$, $\alpha' \in \Psi$, $(r, -q) \in C^-$. Since C^- is a cone it implies $(w, Lw) \in C$ or $Lw \in G(w)$. By (1) $\pi(w) + \psi(\gamma w) \geq 0$. On the other hand setting $r = 0$, $q = 0$ in (3.6) we get:

$$\pi(w) + \psi(\gamma w) = \sup_{\alpha \in \Pi} \langle \alpha, w \rangle + \sup_{\alpha' \in \Psi} \langle \alpha', \gamma w \rangle \leq \rho < 0$$

The obtained contradiction proves (3.5). Let $q \in E^*$, $\alpha \in \Pi$, $\alpha' \in \Psi$, $r \in G^*(q)$ be such that

$$(3.7) \quad i^* \alpha + \gamma^* \alpha' + i^* r - L^* q = 0$$

Thus for all $w \in W_0$ we have $\langle \alpha, w \rangle + \langle r, w \rangle - \langle q, L_0 w \rangle = 0$. It implies

$$(3.8) \quad L_0^* q = i_0^* \alpha + i_0^* r$$

Since $i_0^* : H^* \rightarrow W_0^*$ is the canonical injection and W_0 is dense in H it implies $L_0^* q \in H^*$ and thus that $q \in E_0^*$. By applying now (3.7) to any $w \in W$ using Green formula and (3.8) we obtain

$$\begin{aligned} 0 &= \langle \alpha, w \rangle + \langle \alpha', \gamma w \rangle + \langle r, w \rangle - \langle q, Lw \rangle = \langle \alpha, w \rangle + \langle \alpha', \gamma w \rangle \\ &\quad + \langle r, w \rangle + \langle \beta^* q, \gamma w \rangle - \langle L_0^* q, w \rangle = \langle \alpha' + \beta^* q, \gamma w \rangle \end{aligned}$$

Since $\gamma(W) = T$ it implies $\alpha' + \beta^* q = 0$ or (2).

To prove the converse assume (2) holds. Then there is $\alpha \in \Pi$, $\alpha' \in \Psi$, $(r, -q) \in C^-$ such that $q \in E_0^*$ and $L_0^* q = \alpha + r$, $-\beta^* q = \alpha'$. By Green formula

$$\alpha + r = \gamma^* \beta^* q + L^* q = L^* q - \gamma^* \alpha'$$

and

$$\alpha + \gamma^* \alpha' = L^* q - r$$

Assume $w \in W$ is such that $Lw \in G(w)$. Then $\pi(w) + \psi(\gamma w) \geq \langle \alpha, w \rangle + \langle \alpha', \gamma w \rangle = \langle \alpha + \gamma^* \alpha', w \rangle = \langle L^* q - r, w \rangle = - \langle (r, -q), (w, Lw) \rangle \geq 0$.

Which proves (1).

(3.9) Lemma: Under all assumptions of Lemma 3.4 assume that for all $(u, v, e) \in H \times H \times E$ there is $w \in W$ solving the problem

$$Lw \in G(w+u) + e, \quad w \in \text{Dom}(\pi) - v, \quad \gamma w \in \text{Dom}(\psi)$$

then A is closed in W^* .

Proof: Let $a_n = i^* \alpha_n + \gamma^* \alpha'_n + i^* r_n - L^* q_n$, where $\alpha_n \in \Pi$, $\alpha'_n \in \Psi$, $(r_n, -q_n) \in C^-$, $n = 1, 2, \dots$. Assume $\lim_{n \rightarrow \infty} a_n = a$ in W^* . First we shall prove that $\{(\alpha_n, r_n, -q_n)\}$ is bounded. Since H and E are reflexive, it is enough to show that for all $(u, v, e) \in H \times H \times E$

$$(3.10) \quad \sup_{n \geq 1} (\langle \alpha_n, v \rangle + \langle r_n, u \rangle + \langle q_n, e \rangle) < \infty$$

(any weakly bounded set is bounded). Fix $(u, v, e) \in H \times H \times E$ and let $w \in W$ be such that

$$Lw \in G(w+u) + e, \quad w \in \text{Dom}(\pi) - v, \quad \gamma w \in \text{Dom}(\psi)$$

(it exists by assumptions). Then for some $y \in G(w+u)$:

$$\begin{aligned} \langle \alpha_n, v \rangle + \langle r_n, u \rangle + \langle q_n, e \rangle &= \langle \alpha_n, v \rangle + \langle r_n, u \rangle + \langle q_n, Lw - y \rangle \\ &= \langle \alpha_n, v \rangle + \langle r_n, u \rangle - \langle q_n, y \rangle + \langle L^* q_n, w \rangle = \langle \alpha_n, v \rangle + \langle r_n, u \rangle \\ &\quad - \langle q_n, y \rangle + \langle i^* \alpha_n + \gamma^* \alpha'_n + i^* r_n - a_n, w \rangle = \langle \alpha_n, v + w \rangle + \langle \alpha'_n, \gamma w \rangle + \\ &\quad + \langle (r_n, -q_n), (u + w, y) \rangle - \langle a_n, w \rangle \leq \pi(v + w) + \psi(\gamma w) - \langle a_n, w \rangle \end{aligned}$$

Moreover, $\{\langle a_n, w \rangle\}$ is bounded which implies (3.10). Consequently, $\{\|\alpha_n\|\}$, $\{\|r_n\|\}$, $\{\|q_n\|\}$ are bounded. By reflexivity we may assume that $\{\alpha_n\}$, $\{r_n\}$, $\{q_n\}$ are weakly convergent to some α, r, q respectively. Because Π, C^- are closed and convex by Mazur Lemma [10], $\alpha \in \Pi$, $(r, -q) \in C^-$. Let σ be the right inverse of γ . Then $\alpha'_n = \sigma^* \gamma^* \alpha'_n = \sigma^* (a_n - i^* \alpha_n - i^* r_n + L^* q_n)$. Since σ^* is continuous by the previous part we obtain that α'_n is weakly convergent to $\alpha' = \sigma^* (a - i^* \alpha - i^* r + L^* q)$ and $\alpha' \in \Psi$. Thus:

$$a = i^* \alpha + \gamma^* \alpha' + i^* r - L^* q$$

which proves the theorem.

(3.11) Proof of Theorem 3.1: We apply Lemmas 3.4 and 3.9 to problem (3.2). Then we obtain the existence of $q \in E_0^*$ satisfying

$$L_0^* q \in \partial_a f(z) + G^*(q)$$

$$-g^* q \in \partial_a g(\gamma z)$$

(3.12) Proof of Theorem 2.3: Assumptions of Theorem 2.3, Lemma 2.9, and Theorem 3.1 imply that the assumption of Lemma 2.10 is verified. This concludes the proof of Theorem 2.3.

4. An Example

Let U be a compact topological space, and let a continuous function $f : \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$ be given. Consider a nonlinear control system:

$$(4.1) \quad \begin{cases} \dot{x} = f(x, u(t)) \\ u(t) \in U \text{ is measurable} \end{cases}$$

We denote by K the set of all solutions of (4.1). Let two subsets C_0, C_1 of \mathbb{R}^n and a Lipschitzian function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be given. We shall study the following problem:

$$(4.2) \quad \text{minimize } \left\{ \int_0^1 \varphi(x(t)) dt : x \in K, x(0) \in C_0, x(1) \in C_1 \right\}$$

Assume a trajectory-control pair (z, \bar{u}) solves (4.2). We associate with (4.1) a linear control system

$$(4.3) \quad \begin{cases} \dot{w} \in \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w + I_{f(z(t), U)}^a(\dot{z}(t)) \\ w(0) \in I_{C_0}^a(z(0)) \end{cases}$$

Let $R(1)$ denote the reachable set of (4.3) at time 1. One can verify that it is a convex cone.

(4.4) Theorem: Assume there exists an open neighborhood V of $z([0,1])$ such that $\frac{\partial f}{\partial x}$ is continuous on $V \times U$ and for almost all $t \in [0,1]$, the set-valued map $Q : \text{graph} F \rightarrow U$ defined by

$$Q(x,w) = \{u \in U : f(x,u) = w\}$$

is lower semicontinuous at $(z(t), \dot{z}(t))$. If the following surjectivity assumption holds true:

$$(4.5) \quad I_{C_1}^a(z(1)) - R(1) = \mathbb{R}^n$$

then for all $p > 1$ there exists $q \in W^{1,p}(0,1)$ such that

$$\dot{q}(t) \in \partial_a \varphi(z(t)) - \frac{\partial f}{\partial x}(z(t), \bar{u}(t)) * q(t)$$

$$\max \{ \langle q(t), y \rangle : y \in I_{f(z(t), U)}^a(\dot{z}(t)) \} = 0$$

$$-q(0) \in N_{C_0}^a(z(0)) ; q(1) \in N_{C_1}^a(z(1))$$

Proof: Set $F(x) = \{f(x,u) : u \in U\}$. By a Filippov theorem, the set of solutions K coincides with the set of all solutions of the differential inclusion.

$$\dot{x} \in F(x)$$

Moreover, the graph of F is closed and F is Lipschitzian on V . Define $g : \mathbb{R}^{2n} \rightarrow \mathbb{R} \cup \{+\infty\}$ by

$$g(x,y) = \begin{cases} 0 & \text{if } x \in C_0, y \in C_1 \\ +\infty & \text{otherwise} \end{cases}$$

Thus z solves the problem

$$\text{minimize } \{ g(x(0), x(1)) + \int_0^1 \varphi(x(t)) dt : \dot{x}(t) \in F(x(t)) \}$$

We shall apply Theorem 2.3. For this we need to compute $D_a F(z(t), \dot{z}(t))$ and verify the surjectivity assumption.

Step 1: We claim that for almost all $t \in [0, 1]$

$$(4.6) \quad \left\{ \left(w, \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w + I_{f(z(t), U)}(\dot{z}(t)) \right) : w \in \mathbb{R}^n \right\} = \\ = I_{\text{graph } F}(z(t), \dot{z}(t))$$

Indeed, if $(w, s) \in I_{\text{graph } F}(z(t), \dot{z}(t))$ then for all $h > 0$ there exist w_h, s_h such that $\lim_{h \rightarrow 0^+} (w_h, s_h) = (w, s)$ and

$$(z(t) + hw_h, \dot{z}(t) + hs_h) \in \text{graph } F$$

Let $u_h \in U$ be such that $\lim_{h \rightarrow 0^+} u_h = \bar{u}(t)$ and $\dot{z}(t) + hs_h = f(z(t) + hw_h, u_h)$. (It exists for almost all t by the lower semicontinuity assumption.) Then since $\frac{\partial f}{\partial x}$ is continuous and U is compact we have

$$s = \lim_{h \rightarrow 0^+} \frac{f(z(t) + hw_h, u_h) - \dot{z}(t)}{h} = \lim_{h \rightarrow 0^+} \frac{f(z(t), u_h) - \dot{z}(t)}{h} \\ + \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w. \quad \text{It implies that}$$

$$s - \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w = \lim_{h \rightarrow 0^+} \frac{f(z(t), u_h) - \dot{z}(t)}{h} \in$$

$$\in I_{f(z(t), U)}(\dot{z}(t)). \quad \text{Hence}$$

$$(w, s) \in \left\{ \left(w, \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w + I_{f(z(t), U)}(\dot{z}(t)) \right) : w \in \mathbb{R}^n \right\}$$

Because (w, s) is an arbitrary point in $I_{\text{graph } F}(z(t), \dot{z}(t))$, we proved that $I_{\text{graph } F}(z(t), \dot{z}(t))$ is contained in the right-hand side of the above inclusion. To prove the equality of (4.6), pick up any point \bar{x} in $I_{f(z(t), U)}(\dot{z}(t))$ and let $u_h \in U$ be such

that $\lim_{h \rightarrow 0^+} u_h = \bar{u}(t)$

$$\lim_{h \rightarrow 0^+} \frac{f(z(t), u_h) - \dot{z}(t)}{h} = r$$

(It exists for almost all $t \in [0, 1]$.) Then for all $w \in \mathbb{R}^n$ we have

$$\begin{aligned} & \lim_{h \rightarrow 0^+} \frac{f(z(t) + hw, u_h) - \dot{z}(t)}{h} = \\ & = \lim_{h \rightarrow 0^+} \frac{f(z(t), u_h) - \dot{z}(t)}{h} + \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w \\ & = r + \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w, \text{ which achieves the proof of (4.6).} \end{aligned}$$

Step 2: It follows from Step 1 that

$$D_a F(z(t), \dot{z}(t))(w) = \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w + I_{f(z(t), U)}^a(\dot{z}(t))$$

Fix any $p > 1$.

The surjectivity assumption of Theorem 2.3 has the following form: for all $u, e \in L^p(0, 1)$ there exists a solution $w \in W^{1, p}([0, 1]; \mathbb{R}^n)$ of

$$\begin{aligned} \dot{w}(t) & \in \frac{\partial f}{\partial x}(z(t), \bar{u}(t))(w(t) + u(t)) + I_{f(z(t), U)}^a(\dot{z}(t)) + e(t) \\ w(0) & \in I_{C_0}^a(u(0)) ; w(1) \in I_{C_1}^a(z(1)) \end{aligned}$$

Let $X(t)$ be the matrix (fundamental solution) satisfying

$$\dot{X}(t) = \frac{\partial f}{\partial x}(z(t), \bar{u}(t))X(t) \quad \text{a.e.}$$

$$X(0) = \text{Id}$$

Then the surjectivity condition has the following form: for all $v \in L^p(0, 1)$ there exists $w \in W^{1, p}([0, 1]; \mathbb{R}^n)$ satisfying

$$(4.7) \quad \begin{cases} \dot{w}(t) \in \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w(t) + I_{f(z(t), U)}^a(\dot{z}(t)) + X(t)X^{-1}(1)v(t) \\ w(0) \in I_{C_0}^a(z(0)) \end{cases}$$

and

$$(4.8) \quad w(1) \in I_{C_1}^a(z(1))$$

The reachable set $R'(1)$ of inclusion (4.7) at time 1 is equal to

$$R(1) + \int_0^1 v(t)dt$$

Condition (4.8) implies that

$$R'(1) \cap I_{C_1}^a(z(1)) \neq \emptyset$$

or equivalently that

$$\int_0^1 v(t)dt \in I_{C_1}^a(z(1)) - R(1)$$

Since v is an arbitrary function in $L^P(0,1)$ we proved that the surjectivity assumption is equivalent to

$$I_{C_1}^a(z(1)) - R(1) = \mathbb{R}^n$$

which ends the proof of Theorem 4.4.

(4.9) Remark: Observe that when there are no constraints on the final state (i.e. $C_1 = \mathbb{R}^n$), then assumption (4.5) is automatically satisfied (because $I_{C_1}^a(z(1)) = \mathbb{R}^n$). This happens whenever $z(1)$ belongs to the interior of C_1 . In this case Theorem 4.4 reduces to a non-smooth version of the Pontriagin principle.

We also observe that in Theorem 4.4 we may assume less regularity on f : instead of assuming that $\frac{\partial f}{\partial x}$ is continuous on $V \times U$

it is enough to suppose that for some $L > 0$ f is L -Lipschitzean in the first variable on a neighborhood of $z([0,1])$ and for almost all $t \in [0,1]$

$$\lim_{\substack{h \rightarrow 0^+ \\ u \rightarrow \bar{u}(t)}} \frac{f(z(t) + hw, u) - f(z(t), u)}{h} = \frac{\partial f}{\partial x}(z(t), \bar{u}(t))w$$

Then the same conclusions hold true.

5. Infinite Horizon Problem

Let U be a compact subset in \mathbb{R}^m , A be a $n \times n$ matrix, B be a $n \times m$ matrix, $x_0 \in \mathbb{R}^n$, $\delta > 0$ and a locally Lipschitzean function $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ be given. Consider the following problem:

$$(5.1) \quad \text{minimize} \quad \int_0^\infty \varphi(x(t), u(t)) e^{-\delta t} dt$$

over the trajectory-control pairs (x, u) satisfying

$$(5.2) \quad \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0 \\ u(t) \in U \quad \text{is measurable} \end{cases}$$

This problem was studied in Aubin-Clarke [2] when U is convex, and by many other authors.

The abstract theorems of Section 3 can be applied as well to this new problem, but we would prefer to have more precise results. So we shall study this problem through the same framework but applying the main ideas step by step.

We posit the following growth assumption on φ :

$$\left\{ \begin{array}{l} \text{there are numbers } c, p \geq 1 \text{ such that for every } (x, u) \text{ and} \\ \xi \in \partial_a \varphi(x, u) : \\ |\xi| \leq c (1 + |(x, u)|^{p-1}) \end{array} \right.$$

It easily implies that

$$(5.3) \quad |\varphi(x,u)| \leq |\varphi(0,0)| + c(|x,u| + |x,u|^p)$$

Thus if $u \in L_m^p = L^p(0, \infty; \mathbb{R}^m, e^{-\delta t} dt)$, $x \in L_n^p$ then the integral in (5.1) is finite.

Let λ be the maximum of real parts of the eigenvalue of A .

(5.4) Theorem: Under the above assumptions, assume (z, \bar{u}) solves the considered problem and $\delta > p\lambda$. Then there exists an a.c. function $q : [0, \infty) \rightarrow \mathbb{R}^n$ and measurable functions ξ_1, ξ_2 such that

$$(i) \quad -\dot{q}(t) = A^* q(t) - e^{-\delta t} \xi_1(t) \quad \text{a.e.}$$

$$(ii) \quad (\xi_1(t), \xi_2(t)) \in \partial_a \varphi(z(t), \bar{u}(t))$$

$$(iii) \quad \max \{ \langle q(t), Bw \rangle - e^{-\delta t} \langle \xi_2(t), w \rangle : w \in I_U^\infty(\bar{u}(t)) \} = 0$$

$$(iv) \quad \int_0^\infty e^{(p^*-1)\delta t} |q(t)|^{p^*} dt < \infty ; \int_0^\infty e^{(p^*-1)\delta t} |\dot{q}(t)|^{p^*} dt < \infty$$

where $p^* > 1$ is defined by $\frac{1}{p} + \frac{1}{p^*} = 1$ provided $p > 1$.

If $p = 1$ we have instead: $e^{\delta t} |q(t)|$, $e^{\delta t} |\dot{q}(t)|$ are bounded.

$$(v) \quad \lim_{t \rightarrow \infty} e^{(p^*-1)\delta t} |q(t)|^{p^*} = 0 \quad \text{if } p > 1 ; \text{ if } p = 1$$

then we have instead: $e^{\delta t} |q(t)|$ tends to a finite limit as t goes to $+\infty$.

Proof: It is not restrictive to assume that $x_0 = 0$. For any $u \in L_m^p$ the solution x to (5.2) is given by

$$x(t) = \int_0^t e^{A(t-\tau)} B u(\tau) d\tau$$

and belongs to $W_\delta^{1,p} = \{w \in H^1(0, \infty; \mathbb{R}^n, e^{-\delta t} dt) : w \in L_n^p, \dot{w} \in L_n^p\}$

(see [2, Lemma 3.1]). For all $u \in L_m^p$ set

$$f(u) = \int_0^\infty e^{-\delta t} \varphi \left(\int_0^t e^{A(t-\tau)} B u(\tau) d\tau, u(t) \right) dt$$

Then \bar{u} minimizes f over all $u \in L_m^p$ satisfying $u(t) \in U$.

The following result is analogous to Lemma 1.16.

(5.5) Lemma: If $u \in C = \{u \in L_m^p : u(t) \in I_U^\infty(\bar{u}(t))\}$ then

$$(5.6) \quad \int_0^\infty e^{-\delta t} i_+^a \varphi(z(t), \bar{u}(t)) \left(\int_0^t e^{A(t-\tau)} B u(\tau) d\tau, u(t) \right) dt \geq 0$$

Proof: We introduce the following notations

$$\psi(t, s, v) := i_+^a \varphi(z(t), \bar{u}(t))(s, v)$$

$$Lu(t) := \int_0^t e^{A(t-\tau)} B u(\tau) d\tau \quad \text{for } u \in L_m^p$$

The growth condition implies easily that for $u \in L_m^p$ the function $t \rightarrow \psi(t, Lu(t), u(t))$ belongs to $L^1(0, \infty; \mathbb{R}^{n \times m}; e^{-\delta t} dt)$. Thus the integral (5.6) is finite. Hence it is enough to show that for all bounded $u \in C$

$$\int_0^\infty e^{-\delta t} \psi(t, Lu(t), u(t)) dt \geq 0$$

Fix any such u and let $h_k > 0$ be a sequence converging to zero. We can find a sequence of measurable uniformly bounded functions u_k such that $\bar{u}(t) + h_k u_k(t) \in U$ and $\lim_{k \rightarrow \infty} u_k(t) = \bar{u}(t)$ for all $t \geq 0$. Let $x_k(t) = Lu_k(t)$. By the growth condition

$$(5.7) \quad \begin{aligned} & |\varphi(z(t) + h_k x_k(t), \bar{u}(t) + h_k u_k(t)) - \varphi(z(t), \bar{u}(t))| \leq \\ & \leq c(1 + (|z(t), \bar{u}(t)| + |x_k(t), u_k(t)|)^{p-1}) h_k |(x_k, u_k)(t)| \\ & \leq M(1 + |(z(t), \bar{u}(t))|^{p-1}) h_k \quad \text{for some } M > 0 \end{aligned}$$

Since $f(\bar{u}) = \min \{f(u) : u \in L_m^p, u(t) \in I_U^\infty(\bar{u}(t))\}$ we have

$$\limsup_{k \rightarrow \infty} \frac{1}{h_k} [f(\bar{u} + h_k \bar{u}_k) - f(\bar{u})] \geq 0 .$$

Because of (5.7) we can use Fatou Lemma. Hence

$$\begin{aligned} 0 &\leq \int_0^\infty e^{-\delta t} \limsup_{k \rightarrow \infty} \frac{1}{h_k} [\varphi(z(t) + h_k x_k(t), \bar{u}(t) + h_k u_k(t)) \\ &\quad - \varphi(z(t), \bar{u}(t))] dt \leq \int_0^\infty e^{-\delta t} \psi(t, Lu(t), u(t)) dt \end{aligned}$$

Let $(\xi_1(t), \xi_2(t)) \in \partial_a \varphi(z(t), u(t))$ be such that

$$i_+^a \varphi(z(t), \bar{u}(t)) = \langle (\xi_1(t), \xi_2(t)), \varphi(L\bar{u}(t), \bar{u}(t)) \rangle$$

Then the measurability of $\partial_a \varphi(z(\cdot), \bar{u}(\cdot))$ and $(L\bar{u}(\cdot), \bar{u}(\cdot))$ and the boundedness of $\partial_a \varphi(z(t), \bar{u}(t))$ yield $\xi_1, \xi_2 \in L^\infty$, and for all $u \in V$ we have

$$(5.8) \quad \int_0^\infty \langle e^{-\delta t} \xi_1(t), Lu(t) \rangle + e^{-\delta t} \langle \xi_2(t), u(t) \rangle dt \geq 0$$

Let $r > 0$ be so small that $\lambda p + r < \delta$ and let $\|\xi_1\|$ be the norm of ξ_1 in $L_n^{p^*}(0, \infty; \mathbb{R}^n, e^{-\delta t} dt)$ where $\frac{1}{p} + \frac{1}{p^*} = 1$. By the Hölder inequality and since $|e^{A^* \tau}| \leq e^{\lambda \tau}$ we obtain

$$\begin{aligned} &\int_0^\infty e^{rt} \left| \int_0^t e^{A^* \tau} \xi_1(\tau) e^{-\delta \tau} d\tau \right|^p dt \leq \int_0^\infty e^{rt} \left(\int_0^t e^{\lambda \tau} |\xi_1(\tau)| e^{-\delta \tau} d\tau \right)^p dt \\ &\leq \int_0^\infty e^{rt} \int_0^t e^{(\lambda p - \delta) \tau} d\tau \|\xi_1\|^p \leq c \int_0^\infty e^{(\lambda p + r - \delta)t} dt < \infty \end{aligned}$$

and

$$\int_0^\infty |e^{A^* \tau} \xi_1(\tau) e^{-\delta \tau}|^p d\tau < \infty$$

Therefore

$$\lim_{t \rightarrow \infty} \int_0^t e^{A^* \tau} \xi_1(\tau) e^{-\delta \tau} d\tau = 0$$

Let $u \in C$ be such that $\int_0^\infty e^{-A\tau} Bu(\tau) d\tau$ exists. Then integrating by part we have

$$\begin{aligned} \int_0^\infty \langle e^{-\delta t} \xi_1(t), Lu(t) \rangle dt &= \int_0^\infty \langle e^{A^*t} e^{-\delta t} \xi_1(t), \\ \int_0^t e^{-A\tau} Bu(\tau) d\tau \rangle &= - \int_0^\infty \langle \int_0^t e^{A^*\tau} e^{-\delta\tau} \xi_1(\tau) d\tau, e^{-At} Bu(t) \rangle dt \\ &= - \int_0^\infty \langle \int_0^t e^{-A^*(t-\tau)} e^{-\delta\tau} \xi_1(\tau) d\tau, Bu(t) \rangle dt \end{aligned}$$

Let $q(t) = \int_0^t e^{-A^*(t-\tau)} e^{-\delta\tau} \xi_1(\tau) d\tau$. Then q satisfies (i). Moreover by (5.8)

$$\int_0^\infty \langle q(t), Bu(t) \rangle - e^{-\delta t} \langle \xi_2(t), u(t) \rangle dt \leq 0$$

whenever $\int_0^\infty e^{-A\tau} Bu(\tau) d\tau$ exists. It implies (iii). The relations (iv), (v) follow as in [2]. So the proof is complete.

REFERENCES

- [1] Aubin, J-P. 1979. Applied Functional Analysis, Wiley Interscience.
- [2] Aubin, J-P., F.H. Clarke. 1979. Shadow prices and duality for a class of optimal control problems, SIAM J. of Control, 17 no.5, pp.567-586.
- [3] Aubin, J-P., I. Ekeland. 1984. Applied Nonlinear Analysis, Wiley Interscience,
- [4] Blagodatskich, V. 1979. On the theory of sufficient conditions for optimality, Proceedings of Steklov Institute of Mathematics, Issue 3, pp.81-90.
- [5] Clarke, F.H. 1975. Generalized Gradient and Applications. Trans. Amer. Math. Soc., 205, pp.247-262.
- [6] Clarke, F.H. 1976. The generalized problem of Bolza. SIAM J. of Control, 14, pp.682-699.
- [7] Clarke, F.H. 1976. Optimal solutions to differential inclusions. J. Opt. Theory Appl. vol 19, no.3, pp.469-478.
- [8] Clarke, F.H. 1983. "Optimization and Non-smooth Analysis", Wiley Interscience.
- [9] Ekeland, I., and R. Temam. 1974. "Analyse convexe et problèmes variationnels". Dunod, Paris.
- [10] Frankowska, H. 1984. Contrôlabilité locale et propriétés des semi-groupes de correspondances, CRASt.... (Detailed version to appear).

- [11] Frankowska, H., C. Olech. 1982. Boundary solutions to differential inclusions, *J. Diff. Eqs.* 44, pp.156-165.
- [12] Ioffe, A. 1981. Non-smooth analysis: differential calculus of nondifferentiable mappings, *Trans. Amer. Math. Soc.*, 266 (1), pp.1-56.
- [13] Lasry, J.M., and H. Berliocchi. 1973. Principe de Pontriagin pour des systèmes régis par une équation différentielle multivoque, *CRAS, Paris*, vol. 277, pp.1103-1105.
- [14] Penot, J.P., and P. Terpolilli. 1983. Cônes tangents et singularités, *CRAS, Paris*, vol. 296, pp.721-724.
- [15] Pontriagin, L., V. Boltyanskii, V. Gamkrelidze, E. Mischenko. 1962. "The mathematical theory of optimal process", Wiley Interscience Publishers, New York.
- [16] Rockafellar, R.T. 1975. Existence theorems for general control problems of Bolza and Lagrange. *Adv. in Math.* 15 pp.312-323.
- [17] Rockafellar, R.T. 1980. Generalized directional derivatives and subgradients of non-convex functions. *Canad. J. Math.* 32, pp.257-280.
- [18] Rockafellar, R.T. 1970. "Convex analysis", 1970. Princeton University Press, Princeton, New Jersey.
- [19] Ursescu, C. 1982. Tangent set's calculus and necessary conditions for extremality. *SIAM J. of Control*, 20(4), pp.563-574.
- [20] Wagner, D.H. 1977. Survey of measurable selection theorems, *SIAM J. of Control*, 15, pp.859-903.
- [21] Ważewski, T. 1964. On an optimal control problem, *Proc. Conference "Differential equations and their applications"*, Prague, pp.229-242.