
Differential Inclusion Approach for Mixed Constrained Problems Revisited

REPORT FEUP/DEEC/MRP1/2013

Maria do Rosário DE PINHO, Igor KORNIENKO

1

2013

¹The support of the European Union Seventh Framework Programme [FP7-PEOPLE-2010-ITN] under grant agreement n. 64735-SADCO is greatly acknowledged. This work was also supported by FEDER funds through COMPETE and by Portuguese funds through the Portuguese Foundation for Science and Technology, within projects PTDC/EEA-CRO/116014/2009 and PTDC/EEI-AUT/1450/ 2012.

Abstract

Properties of control systems described by differential inclusions are well established in the literature. Of special relevance to optimal control problems are properties concerning measurability, convexity, compactness of trajectories and Lipschitz continuity of the multifunctions mapping defining the differential inclusion of interest. In this work we concentrate on dynamic control systems coupled with mixed state-control constraints. We characterize a class of such systems that can be described by an appropriate differential inclusion so as exhibit “good” properties of the multifunction. We also illustrate the importance of our findings by treating some applications scenarios.

1 Introduction

Control systems described in terms by differential inclusions have been extensively studied in the literature (see, e.g., [1, 2, 5, 7, 12, 13, 16, 17] to name but a few). Differential inclusions appear in control theory when dynamical systems are expressed as

$$\dot{x}(t) \in F(t, x(t)), \quad (1)$$

where $t \in I \subset \mathbb{R}$, $x \in \mathbb{R}^n$ and F is a multifunction (or set-valued mapping) with closed values in \mathbb{R}^n . Such systems make it possible to study in a uniform way a large number of control problems (in this respect see for example [5]). Indeed, differential inclusions have proved to be a useful framework for optimal control problems. They are convenient to state conditions under which existence of solution is ensured and an useful tool to derive optimality conditions.

It is commonly accepted that differential inclusions are a "natural framework" to study dynamical systems with mixed state- control constraints (see [17], pp. 38). Such approach has been used for example in [7], [10] and, recently, in [8] and [9]. The usefulness of differential inclusions for optimal control problems requires that differential inclusion have certain properties. In particular, it is essential to establish under which conditions the trajectories of control systems described in terms of ordinary differential equations are the functions satisfying (1). In this respect many questions arise as those on the measurability of the multifunction defining the differential inclusion (so existence of measurable selections is guaranteed), compactness of trajectories, convexity properties (two subjects relevant for the existence of solution to optimal control problems), etc. Although such aspects are clearly and concisely treated in the literature for control systems of the form

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)) & \text{a.e. } t \in [a, b], \\ u(t) \in U(t) & \text{a.e. } t \in [a, b], \end{cases}$$

(see for example Chapter 2 in [17]), the same cannot be said when control systems are coupled with mixed constraints. The system of interest, herein denoted as (Σ) , comprises a differential equation

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

and mixed constraints

$$(x(t), u(t)) \in S(t) \quad \text{a.e. } t \in [a, b], \quad (2)$$

together with boundary conditions

$$(x(a), x(b)) \in E. \quad (3)$$

The data comprises a fixed interval $[a, b]$, a function $f : [a, b] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^n$, a multifunction $S : [a, b] \rightarrow \mathbb{R}^n \times \mathbb{R}^k$ and a set $E \subset \mathbb{R}^n \times \mathbb{R}^n$. For such system, a pair (x, u) comprising an absolutely continuous function x (the state trajectory) and a measurable function u (the control), will be called throughout an *feasible process* if it satisfies all the constraints of the above system.

Our aim is to gather together properties of multifunctions relevant to optimal control problems involving control systems of the form (Σ) . To highlight the required properties, while keeping exposition as simple as possible, we work under somewhat strong assumptions.

Conditions under which the state trajectories for (Σ) coincide with the trajectories of a certain differential inclusion

$$\dot{x}(t) \in F_m(t, x(t)) \quad \text{a.e. } t \in [a, b],$$

(where F_m is a multifunction to be defined shortly) satisfying the boundaries constraints (3) will be central in our analysis.

We shall pay particular attention to the case where

$$S(t) := \{(x, u) \in \mathbb{R}^n \times U : g(t, x(t), u(t)) \leq 0\}, \quad (4)$$

where $U \subset \mathbb{R}^k$ and $g : [a, b] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^m$. However, we do not limit our discussion to this case.

We emphasize that the contribution of this paper does not reside on the novelties of the results (since most of them appear to be known) but rather on gathering them together. A remarkable fact is that our work highlights the importance of a *bounded slope* condition imposed on the mixed constraints in the vein [7].

This paper is organized in the following way. In section II we introduce a definition and an important result due to Clarke (in [7]) that will be essential in our setting. Before engaging in our main theme we briefly report

on two optimal control problems where inequality mixed constraint appear naturally. Next, in section IV, we present the main assumptions. Our main results appear in Section V. In the last section we present a sketch of proof of our last (and probably the most important) theorem.

Notations: If $g \in \mathbb{R}^m$, the inequality $g \leq 0$ is interpreted component-wise. Define $\mathbb{R}_-^m = \{\xi \in \mathbb{R}^m : \xi \leq 0\}$ and likewise for \mathbb{R}_+^m . The closed ball centred at x with radius δ is denoted $\bar{B}(x, \delta)$ and likewise for the open ball regardless of the dimension of the underlying space. On the other hand, \bar{B} and B denote the closed and open unit ball centred at the origin. Also $|\cdot|$ is the Euclidean norm or the induced matrix norm on $\mathbb{R}^{p \times q}$.

If $\Omega \subset \mathbb{R}^p$ and $F : \Omega \rightarrow \mathbb{R}^q$ is a multifunction (or set-valued mapping), then the graph of F is defined as

$$\text{Gr } F := \{(x, y) \in \Omega \times \mathbb{R}^q : y \in F(x)\}.$$

We say that a set $S \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ is $\mathcal{L} \times \mathcal{B}$ -measurable when we refer to measurability relative to the σ -field generated by the products of Lebesgue measurable subsets in \mathbb{R} and Borel measurable subsets in $\mathbb{R}^n \times \mathbb{R}^m$.

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

Take $A \subset \mathbb{R}^n$ to be a closed set with and consider $x^* \in A$. Also let $f : \mathbb{R}^k \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function. With respect to f , $x^* \in \mathbb{R}^k$ will be such that $f(x^*) < +\infty$. Concerning nonsmooth analysis we use the following notation: $N_A^P(x^*)$ is the *proximal normal cone* to A at x^* , $N_A^L(x^*)$ is the *limiting normal cone* to A at x^* , $N_A^C(x^*)$ is the *Clarke normal cone* to A at x^* , $\partial^L f(x^*)$ is *limiting subdifferential* of f at x^* and $\partial^C f(x^*)$ is *(Clarke) subdifferential* of f at x^* . If f is Lipschitz continuous near x^* , the convex hull of the limiting subdifferential, $\text{co } \partial^L f(x^*) = \partial^C f(x^*)$.

2 Auxiliary Result

Before proceeding we state an adaptation of Theorem 3.5.2 in [7] that will be important in the forthcoming analysis.

Consider a multifunction $\Gamma : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^k$. For each $t \in [a, b]$, consider the multifunction $x \rightarrow \Gamma(t, x)$ and suppose that the graph $G(t)$ of this multifunction is closed. Suppose that $u^* \in \Gamma(t, x^*)$ and that the following condition holds:

[BS'] There exist $\varepsilon > 0$, $R > 0$ and $\mathcal{K} > 0$ such that, for almost all t ,

$$x \in B(x^*, \varepsilon), \quad u \in B(u^*, R), \quad (\alpha, \beta) \in N_{G(t)}^P(x, u) \implies |\alpha| \leq \mathcal{K}|\beta|.$$

The following theorem asserts that a multifunction satisfying [BS'] is pseudo-Lipschitz.

Theorem 2.1 (adaptation of Theorem 3.5.2 in [7]) *Let $x \rightarrow \Gamma(t, x)$ satisfy [BS'] near $(x^*, u^*) \in G(t)$. Then for any $\xi \in]0, 1[$ and any $x_1, x_2 \in B(x^*, \bar{\varepsilon})$ the following holds*

$$\Gamma(t, x_1) \cap \bar{B}(u^*, (1 - \xi)R) \subset \Gamma(t, x_2) + \mathcal{K}|x_1 - x_2|\bar{B},$$

where $\bar{\varepsilon} = \min\{\varepsilon, \xi R/3\mathcal{K}\}$.

3 Main Assumptions

Mixed constraints, also known as state dependent control constraints, can be written in the general form (see [8]) as (2). We associate with S the multifunction $S_m : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^k$ defined as

$$S_m(t, x) = \{u \in \mathbb{R}^k : (x, u) \in S(t)\}.$$

For each $t \in [a, b]$ the set $S(t)$ is the graph of $x \rightarrow S_m(t, x)$, that is,

$$(x, u) \in S(t) \iff u \in S_m(t, x).$$

Let us also introduce two more multifunctions $F : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m$ and $F_m : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}$ defined as

$$F(t, x) : = \{(f(t, x, u), g(t, x, u)) : u \in U\}, \quad (5)$$

$$F_m(t, x) : = \{f(t, x, u) : u \in S_m(t, x)\} \quad (6)$$

The multifunction (5) will be of importance when we concentrate on S as in (4). Consider now an absolutely continuous function $x^* : [a, b] \rightarrow \mathbb{R}^n$ such that

$$\dot{x}^*(t) \in F_m(t, x^*(t)) \quad \text{a.e.} \quad (7)$$

Take $\varepsilon > 0$ and define

$$X(t) := x^*(t) + \varepsilon \bar{B} \quad \text{and} \quad S_\varepsilon^*(t) := S(t) \cap ((x^*(t) + \varepsilon B) \times \mathbb{R}^k). \quad (8)$$

We now state several assumptions that will be use in the forthcoming analysis. Let $\phi : [a, b] \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^p$ be a general function (ϕ may then be replaced by f or g).

- (B1) The function $t \rightarrow \phi(t, x, u)$ is \mathcal{L} -measurable for all $(x, u) \in \mathbb{R}^n \times \mathbb{R}^k$.
- (B2) The multifunction S is \mathcal{L} -measurable and, for each $t \in [a, b]$, $S(t)$ is closed.
- (B3) The set E is closed.
- (B4) For each $t \in [a, b]$ and $x \in X(t)$, there exists $u \in \mathbb{R}^k$ such that $(x, u) \in S(t)$. Furthermore, each $t \in [a, b]$ there exists a constant σ such that

$$(x, u) \in S(t) \implies |u| < \sigma.$$

- (BS) There exists a constant $\mathcal{K} > 0$ such that, for almost all $t \in [a, b]$ and all $(x, u) \in S_\varepsilon^*(t)$,

$$(\alpha, \beta) \in N_{S_\varepsilon^*(t)}^P(x, u) \implies |\alpha| \leq \mathcal{K}|\beta|.$$

- (CA) For all $t \in [a, b]$ and $x \in X(t)$, $F_m(t, x)$ is convex.

- (LC) There exist constants k_x^ϕ and k_u^ϕ such that, for almost all $t \in [a, b]$ and all $(x_i, u_i) \in S_\varepsilon^*(t)$ ($i = 1, 2$), we have

$$|\phi(t, x_1, u_1) - \phi(t, x_2, u_2)| \leq k_x^\phi |x_1 - x_2| + k_u^\phi |u_1 - u_2|.$$

Some of the above assumptions could (in some situations) be stated in weaker forms. Such strengthening of the assumptions include the second part of (B4) and also the second part of nature of the constant in (BS) and (LC). Indeed, the parameters \mathcal{K} in (BS) and k_x^ϕ and k_u^ϕ in (B1) are constants while in [8] such parameters are considered to be merely measurable functions. These strengthening of the hypotheses however allow us to avoid some technical details in the forthcoming analysis.

4 On $S(t)$ defined by (4)

Let us analyse the case where the set $S(t)$ has a specific structure given by (4). First we introduce additional assumptions on U and F .

- (ICU) The set U is compact and for each $x \in \mathbb{R}^n$ there exists a $u \in U$ such that $g(t, x, u) \leq 0$.
- (IMC) There exists a constant M such that, for almost every t , all $(x, u) \in S_\varepsilon^*(t)$, $\eta \in N_U^L(u)$, $\gamma \in \mathbb{R}_+^m$ with $\langle \gamma, g(t, x, u) \rangle = 0$, we have

$$(\alpha, \beta - \eta) \in \partial_{x,u}^L \langle \gamma, g(t, x, u) \rangle \implies |\gamma| \leq M|\beta|.$$

- (ICC) For all $t \in [a, b]$ and $x \in X(t)$, $F(t, x)$ is convex.

Next we shall relate these assumptions with previous ones imposed on $S(t)$ and F_m .

Lemma 4.1 *Consider $S(t)$ as defined by (4). Assume that g satisfies (B1) and (LC) and that (ICU) holds. Then S is a Lebesgue measurable multifunction and for each t , $S(t)$ is closed.*

Remark: For S as defined by (4) Lemma 4.1 states conditions on g and U implying that (B2) and (B4) hold.

Proof. For each $t \in [a, b]$, $S(t)$ is nonempty by (ICU). By (LC) we know that g is a Carathéodory function. Then Proposition 14.33 in [15] asserts that $S(t)$ is a closed set for each t and $t \rightarrow S(t)$ is Lebesgue measurable.

Now we investigate the relation between (IMC) and (BS). We first characterize $(\alpha, \beta) \in N_{S(t)}^L(x, u)$.

Lemma 4.2 Consider $S(t)$ as defined by (4) and assume that g satisfies (B1) and (LC) and that (ICU) and (IMC) hold. Then for almost every $t \in [a, b]$, for all $(x, u) \in S_\varepsilon^*(t)$ and all $(\alpha, \beta) \in N_{S_\varepsilon^*(t)}^L(x, u)$, there exists an $\gamma \geq 0$ with $\langle \gamma, g(t, x, u) \rangle = 0$ such that

$$(\alpha, \beta) \in \partial_{(x,u)}^L \langle \gamma, g(t, x, u) \rangle + \{0\} \times N_U^L(u). \quad (9)$$

Proof. (Proof of Lemma 4.2) Let $t \in [a, b]$ be such that (IMC) and (LC) hold. Let $\varphi(x, u) = g(t, x, u)$ and set

$$C_1(t) = \varphi^{-1}(\mathbb{R}_-^m) \quad \text{and} \quad C_2(t) = X(t) \times U.$$

Now take any

$$(x, u) \in S(t), \quad x \in x^*(t) + \varepsilon B \quad \text{and} \quad (\alpha, \beta) \in N_{S(t)}^L(x, u).$$

Observe that $x \in X(t)$ but it is not on the boundary of X . This will be of importance in what follow since (IMC) will be invoked.

Our next task is to characterize $N_{C_1(t)}^L(\varphi(x, u))$ in terms of $\partial_{(x,u)}^L \langle \gamma, \varphi(x, u) \rangle$. Corollary 10.50 in [15] is essential here. To do so, we first claim that if $\gamma \in N_{\mathbb{R}_-^m}^L(\varphi(x, u))$ such that $(0, 0) \in \partial_{(x,u)}^L \langle \gamma, \varphi(x, u) \rangle$, then $\gamma = 0$. Take any such γ . Since $\gamma \in N_{\mathbb{R}_-^m}^L(\varphi(x, u))$, we have

$$\langle \gamma, \varphi(x, u) \rangle = 0, \quad \gamma \geq 0.$$

Since (IMC) holds, we deduce from $(0, 0) \in \partial_{(x,u)}^L \langle \gamma, \varphi(x, u) \rangle$ and the fact that $0 \in N_U^L(u)$, that $|\gamma| \leq 0$. It follows that $\gamma = 0$.

Then Corollary 10.50 in [15] holds yielding

$$N_{C_1(t)}^L(x, u) \subset \bigcup \left\{ \partial_{(x,u)}^L \langle \gamma, \varphi(x, u) \rangle : \gamma \in N_D^L(\varphi(x, u)) \right\}. \quad (10)$$

This means that there exists a $\gamma \geq 0$, $\langle \gamma, \varphi(x, u) \rangle = 0$ such that if $(v_1, v_2) \in N_{C_1(t)}^L(x, u)$, then $(v_1, v_2) \in \partial^L \langle \gamma, \varphi(x, u) \rangle$.

Our next step is to prove that $N_{C_1(t)}^L(x, u)$ and $N_{C_2(t)}^L(x, u)$ are transversal in (x, u) , i.e.,

$$(\xi, \zeta) \in -N_{C_1(t)}^L(x, u) \cap N_{C_2(t)}^L(x, u) \implies (\xi, \zeta) = (0, 0). \quad (11)$$

Since $N_{C_2(t)}^L(x, u) = N_{X(t)}^L(x, u) \times N_U^L(u)$, $x \in \text{int}X(t)$, we have $\zeta \in N_U^L(u)$ and $\xi = 0$. By (10) we have, for some γ with the required properties,

$$(0, -\zeta) \in \partial^L \langle \gamma, \varphi(x, u) \rangle.$$

Invoking (IMC) with $\alpha = 0$, $\beta = 0$ and $\eta = \zeta$, we deduce that $\gamma = 0$. But then $(0, \zeta) = (0, 0)$, proving (11). Then we can apply Theorem 6.42 in [15] to conclude that

$$N_{S_\varepsilon^*(t)}(x, u) \subset N_{C_1(t)}^L(x, u) + N_{C_2(t)}^L(x, u).$$

It follows from the above that (9) holds, proving the Lemma. \square

As a consequence of Lemma 4.2 we relate (IMC) and (BS).

Corollary 4.3 Under the assumptions of Lemma 4.2, (BS) holds.

Proof. Take any $(\alpha, \beta) \in N_{S(t)}^P(x, u)$. Since $S_\varepsilon^*(t) \subset S(t)$ we have $N_{S(t)}^P(x, u) \subset N_{S_\varepsilon^*(t)}^P(x, u)$. On the other hand, we also have $N_{S_\varepsilon^*(t)}^P(x, u) \subset N_{S_\varepsilon^*(t)}^L(x, u)$. Thus $(\alpha, \beta) \in N_{S_\varepsilon^*(t)}^L(x, u)$ and it follows from Lemma 4.2 and (IMC) that for $\gamma \geq 0$ with $\langle \gamma, g(t, x, u) \rangle = 0$,

$$\eta \in N_U^L(u), \quad (\alpha, \beta - \eta) \in \partial_{(x,u)}^L \langle \gamma, g(t, x, u) \rangle \implies |\gamma| \leq M|\beta|.$$

By (LC) we also conclude that $(x, u) \rightarrow \langle \gamma, g(t, x, u) \rangle$ is Lipschitz continuous with constant $|\gamma| \max\{k_x^g, k_u^g\}$. We then conclude that

$$|\alpha| \leq |(\alpha, \beta - \eta)| \leq \max\{k_x^g, k_u^g\} |\gamma| \leq \max\{k_x^g, k_u^g\} M |\beta|.$$

But then (BS) holds with $\mathcal{K} = \max\{k_x^g, k_u^g\} M$. \square

5 Convexity of $F_m(t, x)$

Let us now concentrate on (CA) and (ICC). Although for a general $S(t)$, (CA) may be difficult to check, the same is not true when $S(t)$ is as defined by (4). In this case, (ICC) is a condition easier to verify and, as we will show soon, implies (CA). However, the opposite implication does not hold in general. Given the structure of $S(t)$ one may be tempted to think that other easier verifiable condition would involve the multifunctions

$$\begin{aligned} F^f(t, y) &= \{f(t, y, u) : u \in U\}, \\ G^g(t, y) &= \{g(t, y, u) : u \in U\}. \end{aligned}$$

Next we investigate the existence of any relations between convexity properties of F , F_m , F^f and G^g when $S(t)$ is defined by inequalities mixed constraints (4).

Lemma 5.1 *Consider any $t \in [a, b]$ and $x \in X(t)$ such that $S_m(t, x) \neq \emptyset$ where $S(t)$ is defined as*

$$S(t) := \{(x, u) \in \mathbb{R}^n \times U : g(t, x(t), u(t)) \leq 0\}.$$

The following relations hold:

1. $F(t, x)$ convex $\implies F_m(t, x)$ convex, but the opposite implication does not hold.
2. $F(t, x)$ convex $\implies F^f(t, x)$ and $G^g(t, x)$ are convex, but the opposite implication does not hold.
3. The convexity of $F_m(t, x)$ does not imply the convexity of $F^f(t, x)$ and $G^g(t, x)$ and the opposite implication does not hold.

Proof.

1. $F(t, x)$ **convex** $\implies F_m(t, x)$ **convex**.

Take any $v_1, v_2 \in F_m(t, x)$. Then there exist $u_1, u_2 \in U$ such that $v_1 = f(t, x, u_1)$, $v_2 = f(t, x, u_2)$, $g(t, x, u_1) \leq 0$ and $g(t, x, u_2) \leq 0$. Set $z_i = g(t, x, u_i)$, $i = 1, 2$. We have $(v_i, z_i) \in F(t, x)$, $i = 1, 2$. Since $F(t, x)$ is convex, for any $\beta \in [0, 1]$, there exists $u \in U$ such that $(v, z) = \beta(v_1, z_1) + (1 - \beta)(v_2, z_2) = (f(t, x, u), g(t, x, u))$. But $z = \beta z_1 + (1 - \beta)z_2 = g(t, x, u) \leq 0$. Thus $v \in F_m(t, x)$ proving convexity of $F_m(t, x)$.

If $F_m(t, x)$ convex $\not\Rightarrow F(t, x)$ convex.

Take $U = [-1, 1]$, $f(t, x, u) = u$ and $g(t, x, u) = -(u + 1)^2$. Then for any x , the set $F_m(t, x) = [-1, 1]$ is convex. However

$$F(t, x) = \{(u, -(u + 1)^2) : u \in U\}$$

is not.

2. $F(t, y)$ **convex** $\implies F^f(t, y)$ and $G^g(t, y)$ **are convex**.

Fix y and take any $v_1, v_2 \in F(t, x)$. Then there exist $u_1, u_2 \in U$ such that $v_1 = f(t, x, u_1)$ and $v_2 = f(t, x, u_2)$. Set $z_1 = g(t, x, u_1)$ and $z_2 = g(t, x, u_2)$. Then, for any $\beta \in [0, 1]$ $(v, z) = \beta(v_1, z_1) + (1 - \beta)(v_2, z_2)$ is such that $(v, z) \in F(t, x)$, i.e, there exists $u \in U$ such that $(v, z) = (f(t, x, u), g(t, x, u))$. It follows that $v \in F^f(t, x)$ and $z \in G^g(t, x)$ proving convexity of $F^f(t, x)$ and $G^g(t, x)$.

If $F^f(t, x)$ and $G^g(t, x)$ convex $\not\Rightarrow F(t, x)$ convex.

To see this it is enough to define $U = [-1, 1]$, $f(t, x, u) = u^2$, and $g(t, x, u) = u$. Then $F^f(t, x) = [0, 1]$, $G^g(t, x) = [-1, 1]$ and $F(t, x) = \{(u^2, u) : u \in [-1, 1]\}$.

3. **If $F_m(t, x)$ convex $\not\Rightarrow F^f(t, x)$ and $G^g(t, x)$ convex.**

Take $U = [-1, 1]$, $f(t, x, u) = u$ and $g(t, x, u) = (-u, u^3 - u)$. Then $S_m(t, x) = [0, 1]$ and consequently $F_m(t, x) = [0, 1]$ is convex. On the other hand, although $F^f(t, x) = [0, 1]$ is convex, we do not have convexity of $G^g(t, x) = \{(-u, u^3 - u) : u \in [-1, 1]\}$.

If $F^f(t, x)$ and $G^g(t, x)$ are convex $\not\Rightarrow F_m(t, x)$ convex.

Take $U = [-1, 1]$, $f(t, x, u) = u$ and $g(t, x, u) = -u^2 + 1/4$. Then $F^f(t, x) = [-1, 1]$ and $G^g(t, x) = [-3/4, 1/4]$ are both convex. However, $F_m(t, x) = [-1, -1/2] \cup [1/2, 1]$ is not convex.

We summarize our findings:

$F^{f,g}(t, y)$	convex	$\begin{array}{c} \implies \\ \not\Leftarrow \end{array}$	$F(t, y), G(t, y)$	convex
\Downarrow	\Uparrow		\Uparrow	\Downarrow
$F_m(t, y)$	convex	\iff	$F_m(t, y)$	convex

6 Properties of the Multifunctions

Here we establish important properties of the multifunctions S and F_m . We first focus on a general S and later on we turn to (4).

Lemma 6.1 *Assume that (B2) and (B4) hold and that f satisfies (B1) and (LC). Then*

1. *For each $t \in [a, b]$ and each $x \in X(t)$, $S_m(t, x)$ and $F_m(t, x)$ are nonempty and compact.*
2. *The multifunction F_m is $\mathcal{L} \times \mathcal{B}$ -measurable.*
3. *The graph of $(t, x) \rightarrow S_m(t, x)$ is a $\mathcal{L} \times \mathcal{B}$ -measurable set.*
4. *For almost all $t \in [a, b]$ and all $x(t) \in X(t)$ there exists an integrable function c such that for all $\gamma(t) \in F_m(t, x(t))$ we have $|\gamma(t)| \leq c(t)$.*

Proof. The first part of the Lemma follows from (B4) and (LC). For each $t \in [a, b]$ and each $x \in X(t)$, (B4) guarantees that the sets $S_m(t, x)$ and $F_m(t, x)$ are nonempty. The set $S_m(t, x)$ is compact since it is closed by (B2) and bounded by (B4). Taking into account that $u \rightarrow f(t, x, u)$ is continuous by (LC), we get the compactness of the set $F_m(t, x)$.

We now turn to 2. of the Lemma. Take any open set $A \subset \mathbb{R}^n$. We want to prove that

$$\{(t, x) \in [a, b] \times \mathbb{R}^n : F_m(t, x) \cap A \neq \emptyset\}$$

is $\mathcal{L} \times \mathcal{B}$ -measurable. Since, by (B1) and (LC), $t \rightarrow f(t, x, u)$ is measurable for each (x, u) and $(x, u) \rightarrow f(t, x, u)$ is continuous for each t , Proposition 2.3.6 in [17] asserts that f is an $\mathcal{L} \times \mathcal{B}$ -measurable function. It follows that the set

$$f^{-1}(A) = \{(t, x, u) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^k : f(t, x, u) \in A\}$$

is $\mathcal{L} \times \mathcal{B}$ measurable. On the other hand, the multifunction $t \rightarrow S(t)$ is \mathcal{L} -measurable and closed valued by (B4). It follows from, for example, Theorem 2.3.7 in [17], that its graph

$$\Upsilon := \{(t, x, u) \in [a, b] \times \mathbb{R}^n \times \mathbb{R}^k : (x, u) \in S(t)\}, \quad (12)$$

is a $\mathcal{L} \times \mathcal{B}$ measurable set. Consequently, $f^{-1}(A) \cap \Upsilon$ is a $\mathcal{L} \times \mathcal{B}$ -measurable set. Taking into account that (B4) guarantees that for any

$$(t, x) \in \{(t, x) \in [a, b] \times \mathbb{R}^n : F_m(t, x) \cap A \neq \emptyset\}$$

there exists a $u \in \mathbb{R}^k$ such that

$$(t, x, u) \in f^{-1}(A) \cap \Upsilon$$

we conclude the $\mathcal{L} \times \mathcal{B}$ measurability of F_m .

Statement 3. of the Lemma follows from the $\mathcal{L} \times \mathcal{B}$ measurability of the set (12) and the fact that $(x, u) \in S(t)$ is equivalent to $u \in S_m(t, x)$.

It remains to prove 4. Take $t \in [a, b]$ such that $\dot{x}^*(t) \in F_m(t, x^*(t))$ (see (7)). Let u^* be such that $u^* \in S_m(t, x^*(t))$ and $\dot{x}^*(t) = f(t, x^*(t), u^*(t))$. Take x such that $x \in X(t)$. Since by (B4) we have $F_m(t, x) \neq \emptyset$, take any $\gamma \in F_m(t, x)$. By definition of F_m there exists a $u \in S_m(t, x)$ such that $\gamma = f(t, x, u)$. Appealing to (LC) we now have

$$|\gamma| \leq |f(t, x^*(t), u^*(t))| + 2k_x^f \varepsilon + 2k_u^f \sigma = |\dot{x}^*(t)| + 2k_x^f \varepsilon + 2k_u^f \sigma.$$

Set $c(t) = |\dot{x}^*(t)| + 2k_x^f \varepsilon + 2k_u^f \sigma$. Observe that upper bound does not depend on the choice of x or u and it holds for almost every t . Since \dot{x}^* is an integrable function we conclude that $c \in L^1$ proving our claim. \square

Remark: It is important to emphasize that to prove 4. of Lemma 6.1 the fact that our choice of x^* satisfying (7) is essential to assert the integrability of c . A choice of any absolutely continuous function as x^* , not necessarily satisfying (7), would not have been enough to guarantee 4. However, if had we assumed the function f to be uniformly bounded, then we would get 4. with c a constant. \square

We now investigate Lipschitz properties of $x \rightarrow S_m(t, x)$ and $x \rightarrow F_m(t, x)$ for each t . In this respect, (BS) is essential as we shall see. Indeed, conditions (B1), (B2), (B4) and (LC) by themselves, are not enough to guarantee lower semi-continuity of $x \rightarrow S_m(t, x)$ or $x \rightarrow F_m(t, x)$, let alone Lipschitz continuity, as the following example shows.

Example 6.2 Let us fix $t \in [a, b]$ (the interval $[a, b]$ here has no relevance) and set $S(t) = \{(x, u) \in \mathbb{R}^n \times \mathbb{R} : u \in [-1, 1], u|x| \leq 1\}$. Since, for each t ,

$$S_m(t, x) = \begin{cases} [-1, 1] & \text{if } x = 0, \\ [-1, 0] & \text{if } x \neq 0, \end{cases}$$

we have $F_m(t, x) = \{x + u : u \in S_m(t, x)\}$. It is a simple matter to see that (B1), (B2), (B4) hold and that $f(x, u) = x + u$ satisfies (LC). However, both F_m and S_m fail to be lower semi-continuous. To see that consider any sequence $\{x_i\}$ such that $x_i \neq 0$ and $x_i \rightarrow 0$. Then $1/2 \in S_m(t, 0)$ and $1/2 \in F_m(t, 0)$. But there is no convergent sequence $\{u_i\}$ with limit equal to $1/2$, since $u_i \leq 0$. Consequently, there is no sequence $\gamma_i \in F_m(t, x_i)$ converging $1/2$.

Assumption (BS) excludes this example from our context. Indeed, for any t , we have $(1, 0) \in N_{S(t)}^P(0, 1/2)$ and for any constant \mathcal{K} we have $1 > 0 \cdot \mathcal{K}$. So (BS) is not satisfied. \square

Remark: It is worth mentioning, for future reference, that the set $S(t)$ in the above example can be defined as in (4) where $U = [-1, 1]$ and $g(x, u) = u|x|$.

Under our conditions an appeal to Theorem 2.1 guarantees that $x \rightarrow S_m(t, x)$ is not merely pseudo-Lipschitz, it is in fact Lipschitz continuous as we show next.

Lemma 6.3 Assume that (B2), (B4) and (BS) hold. Then there exist constants k_S and ε such that, for almost every t ,

$$x, x' \in x^*(t) + \varepsilon B \implies S_m(t, x) \subset S_m(t, x') + k_S|x - x'|\bar{B}.$$

Proof. Recall that $S(t)$ is the graph of $x \rightarrow S_m(t, x)$ and, by (B2), it is a closed set. Now take t such that (BS) holds. Consider any measurable function u^* such that $u^*(t) \in S_m(t, x^*(t))$ for almost every t . By (B3) we know that for any $u \in S_m(t, x)$ we have $|u| \leq \sigma$. So, for almost every t , any x and u such that $x \in B(x^*(t), \varepsilon)$ (the same ε defining the closed set $X(t)$) and $(x, u) \in S(t)$ we have $u \in B(u^*(t), R)$ with $R = 2\sigma$. Taking into account (BS) we deduce that (BS') holds. We can then apply Theorem 2.1, where $\Gamma(t, x) = S_m(t, x)$ and $G(t) = S(t)$, with $\xi = 1/2$. Observing that, by (B4), for $\varepsilon = \min\{\varepsilon, \frac{\sigma}{3\mathcal{K}}\}$ and any $x \in B(x^*(t), \varepsilon)$ we have

$$S_m(t, x) \cap \bar{B}(u^*(t), (1 - \xi)R) = S_m(t, x) \cap \bar{B}(u^*(t), \sigma) = S_m(t, x)$$

we get our result. \square

As an immediate conclusion from the above Lemma we get the following Corollary.

Corollary 6.4 Assume that (B2), (B4) and (BS) hold and that f satisfies (B1) and (LC). Then there exist constant k_{F_m} such that, for almost every t ,

$$x, x' \in x^*(t) + B(0, \varepsilon) \implies F_m(t, x) \subset F_m(t, x') + k_{F_m}|x - x'|\bar{B}.$$

Proof. Take any $x, x' \in x^*(t) + B(0, \varepsilon)$. Take any $\gamma \in F_m(t, x)$ and $\gamma' \in F_m(t, x')$. Let u and u' be such that $(x, u) \in S_m(t, x)$, $(x', u') \in S_m(t, x')$, $\gamma = f(t, x, u)$ and $\gamma' = f(t, x', u')$. By (LC) and Lemma 6.3 we have

$$\begin{aligned} |f(t, x, u) - f(t, x', u')| &\leq k_x^f|x - x'| + k_u^f|u - u'| \\ &\leq k_x^f|x - x'| + k_u^f\mathcal{K}|x - x'| \\ &= (k_x^f + k_u^f\mathcal{K})|x - x'| \end{aligned}$$

and our result follows with $k_{F_m} = k_x^f + k_u^f\mathcal{K}$. \square

7 Differential Inclusion with Mixed Constraints

We are now in position to invoke Chapter 2 in [17] to obtain several relevant properties for our multifunction F_m and its trajectories.

The observation that the multifunction $X : [a, b] \rightarrow \mathbb{R}^n$ is closed and bounded together with Lemma 6.1 and Theorem 2.5.3 in [17] allow us to deduce the following:

Theorem 7.1 *Assume that (CA) and the conditions under which Lemma 6.1 hold. Take any sequence $\{x_i\}$, $x_i \in W^{1,1}([a, b]; \mathbb{R}^n)$ such that*

$$\text{Gr } x_i \subset \text{Gr } X, \quad \dot{x}_i(t) \in F_m(t, x_i(t)) \text{ a.a. } t \in [a, b], \quad x_i(0) \in X(0).$$

Then there exists a subsequence (we do not relabel) such that

$$x_i \rightarrow x \text{ uniformly} \quad \text{and} \quad \dot{x}_i \rightarrow \dot{x} \text{ weakly in } L^1$$

for some $x \in W^{1,1}([a, b]; \mathbb{R}^n)$ such that $\dot{x}(t) \in F_m(t, x(t))$ a.a. $t \in [a, b]$.

We say that an absolutely continuous function x is a feasible trajectory of F^- if $x(t) \in X(t)$ for all $t \in [a, b]$ and $\dot{x}(t) \in F_m(t, x(t))$ for almost all $t \in [a, b]$. We denote the set of all F_m -feasible trajectories associated with E to be

$$\mathcal{R}_{[a,b]}^*(E) := \{x \in C([a, b]; \mathbb{R}^n) : x \text{ is an } F_m \text{ trajectory and } (x(a), x(b)) \in E\}.$$

Recall that $X(t)$ is closed.

Lemma 7.2 *Assume that (B2) - (B4) as well as (BS) and (CA) hold. Assume also that f satisfies (B1) and (LC). Then $\mathcal{R}_{[a,b]}^*(E)$ is compact with respect to the supremum norm topology.*

Proof. Taking into account Lemma 6.1 and Corollary 6.4, it is a simple matter to see that Theorem 2.6.1 in [17] holds. Our result follows. \square

Now define $\mathcal{S}_{[a,b]}^*(E)$ to be the set of all absolutely continuous functions x associated with a control $u : [a, b] \rightarrow U$ such that $x(t) \in X(t)$ for all $t \in [a, b]$ and

$$\begin{cases} \dot{x}(t) &= f(t, x(t), u(t)) \text{ a.e.,} \\ 0 &\geq g(t, x(t), u(t)) \text{ a.e.,} \\ u(t) &\in U \text{ a.e.,} \\ (x(a), x(b)) &\in E. \end{cases}$$

Consider the S_m , F_m and the corresponding $\mathcal{R}_{[a,b]}^*(E)$ when $S(t)$ satisfies (4), i.e., when

$$S(t) := \{(x, u) \in \mathbb{R}^n \times U : g(t, x(t), u(t)) \leq 0\}.$$

Theorem 7.3 *Consider $S(t)$ as (4) and assume that f and g satisfy (B1) and (LC) and that (ICU), (IMC) and (CA) hold. Then $x \in \mathcal{S}_{[a,b]}^*(E)$ if and only if $x \in \mathcal{R}_{[a,b]}^*(E)$.*

Proof. The implication

$$x \in \mathcal{S}_{[a,b]}^*(E) \implies x \in \mathcal{R}_{[a,b]}^*(E)$$

is trivial. To see that the opposite implication holds take $x \in \mathcal{R}_{[a,b]}^*(E)$. Set $w(t) = \dot{x}(t)$ and $m(t, u) = f(t, x(t), u)$. Consider now the multifunction

$$\Lambda(t) = S(t, x(t)) = \{u \in U : g(t, x(t), u) \leq 0\}.$$

Let $\gamma(t, u) = g(t, x(t), u)$. By Proposition 2.3.4 in [17] $(t, u) \rightarrow \gamma(t, u)$ and $(t, u) \rightarrow m(t, u)$ are $\mathcal{L} \times \mathcal{B}$ measurable. By Proposition 14.33 in [15] Λ is measurable. Then by Theorem 2.3.7 in [17] the graph of Λ is a $\mathcal{L} \times \mathcal{B}$ measurable set. By definition $w(t) \in \{m(t, u) : u \in \Lambda(t)\}$ for almost every $t \in [a, b]$. Theorem 2.3.13 in [17] asserts that there exists a measurable function $u : [a, b] \rightarrow \mathbb{R}^k$ such that

$$w(t) \in \Lambda(t) \text{ a.e.} \quad \text{and} \quad w(t) = m(t, u(t)) \text{ a.e.}$$

It follows that $x \in \mathcal{S}_{[a,b]}^*(E)$, completing our proof.

This Lemma has particularly nice properties when we consider optimal control problems of the form

$$(Q) \quad \left\{ \begin{array}{ll} \text{Minimize } l(x(b)) & \\ \text{subject to} & \\ \quad \dot{x}(t) = f(t, x(t), u(t)) & \text{a.a. } t \in [a, b] \\ \quad u(t) \in U & \text{a.a. } t \in [a, b] \\ \quad g(t, x(t), u(t)) \leq 0 & \text{a.a. } t \in [a, b] \\ \quad (x(a), x(b)) \in \{x_a\} \times E_b & \end{array} \right.$$

where $l : \mathbb{R}^n \rightarrow \mathbb{R}$ is a locally Lipschitz function and E_b a closed set (it may be \mathbb{R}^n in particular). Under our assumptions, Proposition 2.6.2 in [17] asserts that if there exists an feasible pair (x^*, u^*) then (Q) has a minimizer. It is also of interest to apply relaxation results on (Q) when the assumption [C] is removed.

Bibliography

- [1] Aubin J. P., Cellina, A. 1984. *Differential Inclusions: Set-Valued Maps and Viability Theory* Springer-Verlag.
- [2] Aubin, J. P., Frankowska, H. 1990. *Set-Valued Analysis*,. Birkhäuser, Boston.
- [3] Bokanowski O., Cristiani E., Laurent-Varin J., Zidani H. 2012. *Hamilton-Jacobi-Bellman approach for the climbing problem for multi-stage launchers*. Proceedings of MTNS 2012, July 2012, Melbourne, Australia.
- [4] Clarke, F. 1976. *The maximum principle under minimal hypotheses*. SIAM J. Control Optim. 14, pp. 1078-1091.
- [5] Clarke F. 1983. *Optimization and Nonsmooth Analysis*. John Wiley, New York.
- [6] Clarke F, Ledyaev Yu. S., Stern R. J., Wolenski P. R. 1998. *Nonsmooth Analysis and Control Theory*. Springer-Verlag, New York.
- [7] Clarke F. 2005. *Necessary conditions in dynamic optimization*. Mem. Amer. Math. Soc.
- [8] Clarke F, de Pinho M. d. R. 2010. *Optimal control problems with mixed constraints*. SIAM J. Control Optim. 48, pp. 4500–4524.
- [9] Clarke F, Ledyaev Y., de Pinho M. d. R. 2011. *An extension of the schwartzkopf multiplier rule in optimal control*. SIAM J. Control Optim. 49, pp. 599–610.
- [10] Devdaryani E. N., Ledyaev Y. S. 1999. *Maximum principle for implicit control systems*, Appl. Math. Optim., 40, 79–103.
- [11] Kornienko I., de Pinho, MdR, 2013. *Properties of control systems with mixed constraints in the form of inequalities*. Internal Report, ISR, DEEC, FEUP, 2013.
- [12] Loewen P, Rockafellar R. T., 1994. *Optimal Control of Unbounded Differential Inclusions* SIAM J. Control Optim. Vol. 32, No. 2, pp. 442–470.
- [13] Mordukhovich B. 2006. *Variational analysis and generalized differentiation. Basic Theory*. Fundamental Principles of Mathematical Sciences 330, Springer-Verlag, Berlin.
- [14] Miller Neilan, R. L., Lenhart S. 2010. *An Introduction to Optimal Control with an Application in Disease Modeling*. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Volume 75, pp. 67–81.
- [15] Rockafellar R. T. and Wets B. 1998. *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin.
- [16] Smirnov, G. V. 2002 *Introduction to the Theory of Differential Inclusions*. American Mathematical Soc.
- [17] Vinter R. 2000. *Optimal Control*. Birkhäuser, Boston.