Proceedings of the International Congress of Mathematicians
Hyderabad, India, 2010

Optimal Control under State Constraints

Hélène Frankowska*

Abstract

Optimal control under state constraints has brought new mathematical challenges that have led to new techniques and new theories. We survey some recent results related to issues of regularity of optimal trajectories, optimal controls and the value function, and discuss optimal synthesis and necessary optimality conditions. We also show how abstract inverse mapping theorems of set-valued analysis can be applied to study state constrained control systems.

Mathematics Subject Classification (2010). 49K15, 34A60, 47J07, 49N35, 49N60.

Keywords. Optimal control, state constraints, value function, optimal synthesis, normal maximum principle, smoothness of optimal trajectories, regularity of the adjoint variable.

1. Introduction

We consider here finite dimensional control systems by which we mean ordinary differential equations of the form

$$\begin{cases} x'(t) &= f(t,x(t),u(t)), \quad u(t) \in U \quad \text{a.e. in } [0,1], \\ x(0) &= x_0, \end{cases} \tag{1}$$

where U is a complete separable metric space, $f : [0,1] \times \boldsymbol{R}^n \times U \to \boldsymbol{R}^n$, $x_0 \in \boldsymbol{R}^n$, $x(t)$ is called the state of the system, t denotes the time, $x'(\cdot)$ is the derivative of $x(\cdot)$ with respect to time, and the function $u(\cdot)$ has to be chosen so that the corresponding solution $x(\cdot)$ has some desirable properties; in other words, $u(\cdot)$ "controls" the solution $x(\cdot)$ of (1). As a set of controls we choose the set of all Lebesgue measurable functions $u(\cdot) : [0,1] \to U$, while a solution

*Combinatoire & Optimisation, Université Pierre et Marie Curie, case 189, 4 place Jussieu, 75252 Paris Cedex 05, France. E-mail: frankowska@math.jussieu.fr.

$x(\cdot)$ (called here a trajectory of control system) is an absolutely continuous function satisfying $x(0) = x_0$ and $x'(t) = f(t, x(t), u(t))$ almost everywhere in $[0,1]$ (in the sense of Lebesgue measure) for some control $u(\cdot)$. Once we choose a control $u(\cdot)$, system (1) becomes an ordinary differential equation for which conditions for existence, uniqueness and properties of trajectories are classically known. Measurable controls have proven to be well adapted for investigation of existence of solutions to optimal control problems in the general case.

The importance of control systems in mathematics and science is nowadays widely acknowledged. Indeed, not only do they respond to basic issues arising in engineering and social sciences, but they serve as subsumption of previous theories - e.g., the classical calculus of variations - and stimulus to progress in related mathematical fields as well. For example, some fundamental research directions such as weak solutions of nonlinear first (and second) order partial differential equations, set-valued, variational and nonsmooth analysis, have found their inspiration and motivation in control theory and differential games.

The analysis of properties of trajectories of (1) becomes much more challenging to study if the states $x(t)$ are required to belong to a certain region; then we say that the control system (1) is subject to *state constraints*. To be specific, let K be a given closed subset of $\boldsymbol{R}^n$ and consider state constraint of the form

$$x(t) \in K \quad \text{for all} \;\; t \in [0,1]. \tag{2}$$

A trajectory $x(\cdot)$ of (1) satisfying the state constraint (2) is called a *viable* (or feasible) trajectory of the control system. Properties of viable trajectories could be quite different from those of system (1) only.

From now on, we denote by $\mathcal{C}$ the space of continuous functions from $[0,1]$ to $\boldsymbol{R}^n$ with the supremum norm $\|\cdot\|_{\mathcal{C}}$ and by $W^{1,1}([0,1];\boldsymbol{R}^n)$ the space of absolutely continuous functions from $[0,1]$ to $\boldsymbol{R}^n$ with the norm $\|w\|_{W^{1,1}} = \|w\|_{L^1} + \|w'\|_{L^1}$. Let $\mathcal{B}_{\mathcal{C}}$ and $\mathcal{B}_{W^{1,1}}$ denote the closed unit balls in the corresponding spaces.

Consider the set $\mathcal{S}(x_0)$ of all trajectories of (1) and let $\mathcal{S}_K(x_0)$ denote the set of all trajectories of (1), (2). If f is Lipschitz with respect to x with a constant independent of t and u, then the set-valued map $x_0 \rightsquigarrow \mathcal{S}(x_0)$ is Lipschitz continuous in the sense that for some $L \geq 0$, $\mathcal{S}(x_0) \subset \mathcal{S}(y_0) + L|x_0 - y_0|\mathcal{B}_{W^{1,1}}$ for all $x_0,\, y_0 \in \boldsymbol{R}^n$. This is no longer the case for the set-valued map $\mathcal{S}_K(\cdot)$ even for simple sets K and even with $\mathcal{B}_{W^{1,1}}$ replaced by $\mathcal{B}_{\mathcal{C}}$.

Example. $K = \{x = (x_1, x_2) \in \boldsymbol{R}^2 \,|\, x_2 \leq x_1^2\}$, $U = [-1,1] \times \{0\}$. Consider the following control system under a state constraint

$$x'(t) = u(t), \;\; u(t) \in U \;\; \text{a.e. in} \;\; [0,1], \;\; x(t) \in K \;\; \forall\, t \in [0,1].$$

Then $x(t) := (1-t, 0)$ is a viable trajectory with the initial state $x(0) = (1,0)$. Pick any $x_2 \in (0,1]$ and $y(\cdot) \in \mathcal{S}_K((1, x_2))$. Denoting by $|\cdot|$ the Euclidean norm in $\boldsymbol{R}^2$, we get $|x(1) - y(1)| \geq \sqrt{x_2}$ and therefore $\mathcal{S}_K(\cdot) : K \rightsquigarrow \mathcal{C}$ is not Lipschitz on any neighborhood of $(1,0)$. □

• **Existence of viable trajectories.** In general, for some initial conditions, trajectories of (1) satisfying the state constraint (2) may not exist; for instance when for every $u \in U$, $f(0, x_0, u)$ points strictly outside of K and f is continuous. The tangent vectors to the set K turn out to be very instrumental for investigating existence of trajectories of a control system under state constraint.

Denote the distance from $y \in \boldsymbol{R}^n$ to K by $\operatorname{dist}(y; K)$. The contingent cone and the Clarke tangent cone to K at $x \in K$ are defined respectively by

$$\begin{array}{rcl} T_K(x) & := & \{v \in \boldsymbol{R}^n \mid \liminf_{h\to 0+} \frac{1}{h}\operatorname{dist}(x + hv; K) = 0\}; \\ C_K(x) & := & \{v \in \boldsymbol{R}^n \mid \lim_{h\to 0+,\, K\ni y\to x} \frac{1}{h}\operatorname{dist}(y + hv; K) = 0\}; \end{array}$$

and the normal cone to K at x by $N_K(x) := \{p \in \boldsymbol{R}^n \mid \langle p, v\rangle \le 0 \;\; \forall\, v \in C_K(x)\}$.

Set $f(t, x, U) := \bigcup_{u\in U}\{f(t, x, u)\}$ and $W(t, x) := \{u \in U \,|\, f(t, x, u) \in T_K(x)\}$. Observe that $x(\cdot) \in \mathcal{S}_K(x_0)$ if and only if $x(0) = x_0$ and for almost every $t \in [0, 1]$

$$x'(t) \in f(t, x(t), W(t, x(t))) = f(t, x(t), U) \cap T_K(x(t)).$$

In other words $x(\cdot)$ is a trajectory of the following control system

$$x'(t) = f(t, x(t), u(t)), \;\; u(t) \in W(t, x(t)) \quad \text{a.e. in } [0, 1], \;\; x(0) = x_0. \tag{3}$$

The main difficulty in investigating the above system consists in high irregularity of the set-valued map $(t, x) \rightsquigarrow W(t, x)$. For this reason usually one constructs trajectories of (1) satisfying state constraint (2) instead of solving directly (3).

Existence of viable trajectories can be studied using viability theory, which was developed for systems described by differential inclusions. Control systems are a particular case of differential inclusions for set-valued maps $(t, x) \rightsquigarrow f(t, x, U)$.

Let $F : \boldsymbol{R}^n \rightsquigarrow \boldsymbol{R}^n$ be a Marchaud map, i.e. an upper semicontinuous set-valued map with nonempty convex compact values and linear growth and let $x_0 \in K$. Consider the differential inclusion

$$\left\{ \begin{array}{rcll} x'(t) & \in & F(x(t)) & \text{for a.e. } t \ge 0 \\ x(t) & \in & K & \text{for all } t \ge 0 \\ x(0) & = & x_0. & \end{array} \right. \tag{4}$$

A locally absolutely continuous function $x(\cdot) : [0, \infty) \to \boldsymbol{R}^n$ satisfying the above relations is called a viable (in K) trajectory. A necessary and sufficient condition for the existence of a trajectory of (4) for every $x_0 \in K$ is the viability condition

$$F(x) \cap T_K(x) \neq \emptyset \quad \forall\, x \in K. \tag{5}$$

(see for instance [3, Theorems 3.3.2 and 3.3.5]). Similar conditions allowing to handle time dependent constraints $K(t)$ with F depending also measurably on time as well as infinite dimensional control systems and stochastic systems can be found for instance in [3]. Viability theory has numerous applications to control, such as for investigation of uniqueness of solutions to Hamilton-Jacobi-Bellman equations, of invariance of stochastic control systems, of optimal synthesis, etc.

• **Inward pointing condition.** In the example above the mapping $f(x,u) = u$ is Lipschitz, $f(x,U) = [-1,1] \times \{0\}$ is compact and convex and viability condition (5) is satisfied. Thus (5) is not sufficient for the local Lipschitz continuity of $\mathcal{S}_K(\cdot) : K \rightsquigarrow \mathcal{C}$ on K (and on the interior $\mathrm{Int}\, K$ of K) even when f is Lipschitz. For guaranteing such property a stronger tangential condition is needed

$$F(x) \cap \mathrm{Int}\, C_K(x) \neq \emptyset \quad \forall\, x \in \partial K. \tag{6}$$

For K having a smooth boundary, the control system (1) with the time independent f and $F(x) := f(x,U)$, (6) is equivalent to the so called *inward pointing condition*

$$\forall\, x \in \partial K,\ \exists\, u_x \in U \ \text{ such that } \ \langle n_x, f(x,u_x)\rangle < 0 \tag{7}$$

for the outward unit normal n_x to K at x. Condition (7) was introduced in [79] to investigate continuity of the value function of an infinite horizon problem and then to study uniqueness of viscosity solutions to a Hamilton-Jacobi PDE.

• **Inward pointing condition and linearization of control systems.** When K is an intersection of closed sets with smooth boundaries, a generalization of (6) to the time dependent f implies that, under some mild assumptions on f and a transversality assumption on K, for any $n \times n$ matrices $A(t) = (a_{ij}(t)),\ i,j = 1,...,n$ with $a_{ij}(\cdot) \in L^1(0,1)$ and every trajectory/control pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ of (1), (2), there exists a solution $w(\cdot)$ to the following linear control system

$$\begin{cases} w'(t) = A(t)w(t) + v(t), \quad v(t) \in T_{co\, f(t,\bar{x}(t),U)}(f(t,\bar{x}(t),\bar{u}(t))) \ \text{ a.e.} \\ w(0) = 0 \\ w(t) \in \mathrm{Int}\, T_K(\bar{x}(t)) \ \text{ for all } t \in (0,1], \end{cases} \tag{8}$$

where co states for the convex hull (see [11], and [53] for $w(0) = w_0 \in \mathrm{Int}\, T_K(\bar{x}(0))$). When $A(t) = \frac{\partial f}{\partial x}(t,\bar{x}(t),\bar{u}(t))$, the control system in (8) is a linearization of (1) along the trajectory/control pair $(\bar{x}(\cdot),\ \bar{u}(\cdot))$, while the relation $w(t) \in \mathrm{Int}\, T_K(\bar{x}(t))$ can be considered as linearization of the state constraint along $\bar{x}(\cdot)$.

Existence of a solution to (8) is important in various applications. For instance it yields normality of necessary optimality conditions for some optimal control problems. Observe that it resembles a constraint qualification condition in mathematical programming, which guarantees existence of Lagrange multipliers in normal form. In Section 2 existence of a solution to (8) is used to investigate local Lipschitz continuity of $\mathcal{S}_K(\cdot) : \mathrm{Int}\, K \rightsquigarrow \mathcal{C}$ and in Section 4 it is applied to derive normal first order necessary optimality conditions.

• **Value function of the Mayer optimal control problem.** Let $\varphi : \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ be an extended-real-valued function, bounded from below.

Consider the Mayer problem

$$\text{minimize}\,\{\varphi(x(1)) \mid x(\cdot) \in \mathcal{S}_K(x_0)\}\,. \tag{9}$$

A trajectory $\bar{x}(\cdot) \in \mathcal{S}_K(x_0)$ is called optimal if $\varphi(\bar{x}(1)) = \min_{x(\cdot)\in\mathcal{S}_K(x_0)} \varphi(x(1)) < +\infty$. Let $t_0 \in [0,1]$, $y_0 \in K$ and consider the control system

$$x'(t) = f(t, x(t), u(t)), \quad u(t) \in U \;\text{ a.e. in }\; [t_0, 1], \quad x(t_0) = y_0. \tag{10}$$

The value function $V : [0,1] \times \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ associated to (9) is defined by

$$V(t_0, y_0) = \inf\{\varphi(x(1)) \mid x(\cdot) \;\text{ is a trajectory of }\; (10),\; x([t_0, 1]) \subset K\}, \tag{11}$$

(we adopt the convention that the infimum *inf* over an empty set is equal to $+\infty$).

Value functions arising in various optimal control problems have been extensively used since their introduction by Bellman and Isaacs in the fifties. In general, even for smooth f, φ and in the absence of state constraints, the value function of Mayer's problem may be not differentiable. Its lack of differentiability is related to the multiplicity of optimal trajectories (see [19]). This may be also explained by the shocks of characteristics of the associated Hamilton-Jacobi equation, see [25, 26, 55]. Conversely, as it was shown in [16] - [18], the absence of shocks guarantees smoothness of the value function. The Hamiltonian $H : [0,1] \times \boldsymbol{R}^n \times \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ of the Mayer problem is defined by

$$H(t, x, p) = \sup_{u \in U} \langle p, f(t, x, u) \rangle.$$

Under appropriate assumptions, V is a unique solution in a generalized sense to the Hamilton-Jacobi equation

$$-\frac{\partial V}{\partial t} + H\left(t, x, -\frac{\partial V}{\partial x}\right) = 0, \quad V(1, x) = \varphi(x), \quad (t, x) \in [0,1] \times K. \tag{12}$$

It is well known that (12) does not have smooth solutions and for this reason the notion of solution was extended to non differentiable functions. The most popular are continuous viscosity solutions using superdifferentials and subdifferentials instead of gradients for defining super and subsolutions to (12). See for instance [6, 32, 33, 46] and the references contained therein.

The Hamiltonian H defined above is convex with respect to the last variable. When $K = \boldsymbol{R}^n$ (no state constraint) this actually allowed to get a simpler definition of lower semicontinuous solution involving only subdifferentials and equalities (see [7] for an approach based on PDE arguments and [50, 51, 58] for the one based on viability theory). An extension to systems under state constraints is given in [57].

The value function is also an important tool for investigating optimality conditions. For instance it follows from [19] that in the absence of state constraints, sufficient conditions for optimality for the Mayer problem can be expressed using extremal points of the generalized gradients of the value function and that the optimal trajectories are unique at points of differentiability of the value function whenever H is smooth enough. Furthermore, the adjoint variable in the maximum principle (discussed below) satisfies some additional relations involving superdifferentials of the value function. These additional relations make the maximum principle not only a necessary but also a sufficient optimality condition (see [19] and [81]). When φ and f are smooth enough, the value function is semiconcave (see [24] for a nice collection of results on semiconcave functions and applications of semiconcavity to problems of optimal control or [19] for both a proof of semiconcavity of the value function of Mayer's problem and sufficient optimality conditions). In the presence of state constraint, in general, V is not semiconcave even for smooth f, φ. Observe that if $\varphi(\cdot)$ is locally Lipschitz, then local Lipschitz continuity of $\mathcal{S}_K(\cdot) : K \rightsquigarrow \mathcal{C}$ yields local Lipschitz continuity of the value function.

• **Maximum principle.** Assume f differentiable with respect to x and φ differentiable. Let $\bar{x}(\cdot)$ be optimal for problem (9) and let $\bar{u}(\cdot)$ be a corresponding control. Then, under some technical assumptions, the celebrated maximum principle under state constraint holds true (see [43] and also [60] for an earlier version): there exist $\lambda \in \{0, 1\}$, an absolutely continuous mapping $p(\cdot) : [0, 1] \to \boldsymbol{R}^n$ and a mapping $\psi(\cdot) : [0, 1] \to \boldsymbol{R}^n$ of bounded total variation satisfying

(i) the adjoint equation (where * states for the transposition)

$$-p'(t) = \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^*(p(t) + \psi(t)) \ \text{ a.e. in } \ [0, 1], \tag{13}$$

(ii) the maximum principle

$$\langle p(t)+\psi(t), f(t, \bar{x}(t), \bar{u}(t))\rangle = \max_{u \in U}(\langle p(t)+\psi(t), f(t, \bar{x}(t), u)\rangle \ \text{ a.e. in } \ [0, 1] \tag{14}$$

and the transversality condition $-p(1) - \psi(1) = \lambda \nabla \varphi(\bar{x}(1))$. Furthermore $\psi(\cdot)$ is linked to the state constraint in the following way : there exist a positive (scalar) Radon measure μ on $[0, 1]$ and a Borel measurable $\nu(\cdot) : [0, 1] \to \boldsymbol{R}^n$ satisfying

$$\nu(s) \in N_K(\bar{x}(s)) \cap B \quad \mu - a.e., \tag{15}$$

$$\psi(t) = \int_{[0,t]} \nu(s) d\mu(s) \quad \forall\, t \in (0, 1]. \tag{16}$$

See the monograph [82] for different forms of maximum principle under state constraints and [72] for some historical comments on the maximum principle.

The above necessary optimality condition is said to be normal if $\lambda = 1$. It was shown in [11] and [53] that a generalized inward pointing condition yields normality of the maximum principle for a class of state constraints with non smooth boundaries. Normality is very useful for investigation of Lipschitz continuity of optimal trajectories (with respect to the time), to avoid the Lavrentieff phenomenon, see [21, 41, 54, 59, 62, 67, 78]. Let us underline that regularity of optimal trajectories is important in discrete approximations and hence for numerical solutions.

• **Adjoint state and gradient of the value function.** It is well known that if $K = \boldsymbol{R}^n$ and if the value function is differentiable, then $-p(t) = \nabla_x V(t, \bar{x}(t))$ for all t. In [30], for $K = \boldsymbol{R}^n$, this relation was extended to a locally Lipschitz value function using generalized gradient instead of gradient. It follows from [27] that in the presence of a state constraint if $V(0, \cdot)$ is locally Lipschitz at $x_0 \in \operatorname{Int} K$, then $p(\cdot)$ in the maximum principle satisfies $-p(0) \in \lambda \partial_x V(0, x_0)$, where $\partial_x V(0, x_0)$ denotes the generalized gradient of $V(0, \cdot)$ at x_0. For K with a smooth boundary, the relation $-p(t) - \psi(t) \in \partial_x V(t, \bar{x}(t))$ for a.e. $t \in [0, 1]$ was recently obtained in [14], using a slightly different notion of generalized gradient on the boundary of K.

• **Regularity of optimal trajectories.** The mapping $\psi(\cdot)$ in the maximum principle may be very irregular and have an infinite number of jumps (see [71] for a relevant example in $\boldsymbol{R}^n$ with $n \geq 3$). For this reason optimal controls may also be highly discontinuous with respect to the time. However for some classes of nonlinear constrained optimal control problems of Bolza type (see (34) in Section 4 below) this is no longer the case. This was observed first in [62] for linear problems with convex cost and convex state constraints and extended in [67] to constrained control systems that are nonlinear with respect to the state. Further generalizations to systems affine with respect to control under nonlinear state constraints were obtained in [59]. In [53] it was shown that for the Bolza optimal control problem, whose Hamiltonian has a coercive gradient in the directions normal to constraint, $\psi(\cdot)$ is continuous on $(0, 1)$. This helps to investigate the continuity of optimal controls. Moreover, under a uniform coercivity assumption in the directions normal to constraint, $\psi(\cdot)$ becomes absolutely continuous on $(0, 1)$, implying in turn that optimal trajectories have absolutely continuous derivatives. For some classes of control systems this allows to get absolutely continuous and even Lipschitz continuous optimal controls.

• **Outline.** In the next section we discuss the local Lipschitz continuity of $\mathcal{S}_K(\cdot)$ and in Section 3 the local Lipschitz continuity of the value function and optimal synthesis. Section 4 relates the adjoint state $p(0)$ of the maximum principle to the generalized gradient of the value function. Finally, Section 5 is devoted to smoothness of $\psi(\cdot)$ in the maximum principle and regularity of optimal trajectories and controls for the Bolza optimal control problem.

2. Lipschitz Dependence of Viable Trajectories on Initial States and Inverse Mapping Theorems

By $B(x_0,\varepsilon)$ (or $B_X(x_0,\varepsilon)$) we denote the closed ball in a metric space X of center $x_0 \in X$ and radius $\varepsilon > 0$ and by B or B_Y the closed unit ball centered at zero in a Banach space Y. The Euclidean norm in $\boldsymbol{R}^n$ is denoted by $|\cdot|$.

Let (X, d_X) be a metric space, Y be a Banach space and $G : X \rightsquigarrow Y$ be a set-valued map. G is said to be locally Lipschitz, if it has nonempty values and for every $x_0 \in X$ there exist $\varepsilon > 0,\ L \geq 0$ such that $G(x_1) \subset G(x_2) + L|x_1 - x_2|B_Y$ for all $x_1,\, x_2 \in B_X(x_0,\varepsilon)$. The graph of G is defined by $\mathrm{Graph}\,(G) := \{(x,y) \mid y \in G(x)\}$.

Consider a set-valued map $F : [0,1] \times \boldsymbol{R}^n \rightsquigarrow \boldsymbol{R}^n$, a closed set $K \subset \boldsymbol{R}^n$, $x_0 \in K$ and the differential inclusion

$$\left\{ \begin{array}{lcl} x'(t) & \in & F(t,x(t)) \quad \text{for a.e. } \ t \in [0,1], \\ x(0) & = & x_0. \end{array} \right. \tag{17}$$

It is worth to underline that if the mapping f from the introduction is measurable with respect to t and continuous with respect to x, u, then the set of trajectories of control system (1) coincides with the set of trajectories of differential inclusion (17) for $F(t,x) = f(t,x,U)$, see for instance [5, Theorem 8.2.10]. Define

$$\mathcal{S}(x_0) := \left\{ x(\cdot) \in W^{1,1}([0,1];\boldsymbol{R}^n) \ \mid \ x(\cdot) \ \text{ satisfies } \ (17) \right\},$$

$$\mathcal{S}_K(x_0) := \left\{ x(\cdot) \in \mathcal{S}(x_0) \ \mid \ x(t) \in K \ \text{ for all } t \in [0,1] \right\}.$$

We say that $\mathcal{S}_K(\cdot)$ is locally $\mathcal{C}-$Lipschitz (respectively $W^{1,1}-$Lipschitz) on a subset $\mathcal{D} \subset K$ if it is locally Lipschitz as a set-valued map from $\mathcal{D}$ into the space $\mathcal{C}$ (respectively into the space $W^{1,1}([0,1];\boldsymbol{R}^n)$).

Theorem 2.1. *Assume there exists $\gamma > 0$ such that $\sup_{v\in F(t,x)} |v| \leq \gamma(1+|x|)$ and $F(t,x)$ is nonempty and closed for all $(t,x) \in [0,1] \times \boldsymbol{R}^n$, that F is locally Lipschitz and that the "inward pointing condition"*

$$F(t,x) \cap \mathrm{Int}\, C_K(x) \neq \emptyset \quad \forall\, x \in \partial K, \ \ \forall\, t \in [0,1] \tag{18}$$

holds true. Then the set-valued map $\mathcal{S}_K(\cdot)$ is locally $\mathcal{C}-$Lipschitz on K.

The above theorem is an extension of a result due to Filippov [45] to systems under state constraints. In the absence of a state constraint a stronger conclusion holds true : $\mathcal{S}(\cdot)$ is locally $W^{1,1}-$Lipschitz under less restrictive assumptions, for instance F may be unbounded and measurably dependent on time. Furthermore [45] provides also estimates of the $W^{1,1}-$distance from an arbitrary $x(\cdot) \in W^{1,1}([0,1];\boldsymbol{R}^n)$ to the set $\mathcal{S}(x_0) \subset W^{1,1}([0,1];\boldsymbol{R}^n)$.

There exist several approaches dealing with the question of Lipschitz continuity of $\mathcal{S}_K(\cdot)$. We briefly describe some of them.

• The **first** one was initiated in [79] for the time independent control systems when the boundary of K is C^2 and extended in [23] to Hilbert spaces. It is based on the local Lipschitz continuity of $\mathcal{S}(\cdot)$ and on a modification of controls in a suitable way to satisfy state constraints.

Actually the very same approach can be used to prove Theorem 2.1. More precisely fix $r > 0$ and $x_0 \in K$. Then there exists $L_r \geq 0$ such that for any $x_1,\, x_2 \in K \cap B(x_0, r)$ and every $y(\cdot) \in \mathcal{S}_K(x_1)$ we can find $\tilde{z}(\cdot) \in \mathcal{S}(x_2)$ satisfying $\|y - \tilde{z}\|_{\mathcal{C}} \leq L_r|x_1 - x_2|$. Then to prove Theorem 2.1 it is sufficient to find $z(\cdot) \in \mathcal{S}_K(x_2)$ verifying $\|z - \tilde{z}\|_{\mathcal{C}} \leq \bar{c}|x_1 - x_2|$ for a constant $\bar{c} \geq 0$ depending only on (the magnitude of) $r + |x_0|$.

To construct such $z(\cdot)$, assume that $\tilde{z}(\cdot) \notin \mathcal{S}_K(x_2)$ and define $t_0 = \inf\{t \mid \tilde{z}(t) \notin K\}$. By the inward pointing condition, it can be shown that for some $v_0 \in F(t_0, z(t_0)) \cap \operatorname{Int} C_K(z(t_0)), \varepsilon > 0$ and $\tau > t_0$ there exists a trajectory $z(\cdot)$ of the differential inclusion $z'(s) \in F(s, z(s))$ a.e. in $[t_0, \tau]$ such that

$$z(s) \in z(t_0) + (s - t_0)v_0 + B\left(0, \varepsilon(s - t_0)\right) \subset \operatorname{Int} K \quad \forall\, s \in (t_0, \tau].$$

Filippov's theorem from [45] and the local Lipschitz continuity of $F(\cdot,\cdot)$ imply the existence of a trajectory $x(\cdot)$ to the differential inclusion $x'(s) \in F(s, x(s))$ a.e. in $[\tau, 1]$ such that $|x'(s) - \tilde{z}'(s - \tau + t_0)| \leq c(\tau - t_0)$ for all $s \in [\tau, 1]$ and $x(\tau) = z(\tau)$, where the constant $c \geq 0$ depends only on $|x_0| + r$. It follows that for some $t_1 > \tau$, $x([\tau, t_1[) \subset \operatorname{Int} K$ and either $t_1 = 1$ or $x(t_1) \in \partial K$ and $t_1 < 1$. Denote by $z(\cdot)$ the restriction of $x(\cdot)$ to $[\tau, t_1]$. By assumptions of Theorem 2.1 it is possible to choose $\beta > 0$ and $\alpha > 0$ depending only on $|x_0| + r$ in such way that for some $\tau > t_0$ satisfying $\tau - t_0 \leq \beta|x_1 - x_2|$ we have $t_1 \geq \min\{\tau + \alpha, 1\}$ for sufficiently small $|x_1 - x_2|$.

Repeating the described process (a finite number of times) we construct $z(\cdot)$ on $[0, 1]$ as required. This approach uses a time shift in the definition of $z(\cdot)$ on $[t_0 + \tau, t_1]$ (which is not convenient when applied to some questions arising in differential games, where players have to adapt to each other strategies without knowing the future, i.e. using non anticipative controls).

• The **second** approach uses the so called neighbouring feasible trajectories theorems. These theorems provide a sufficient condition for the existence of $L_r \geq 0$ depending only on $|x_0| + r$ such that for any $y_0 \in K \cap B(x_0, r)$ and $x(\cdot) \in \mathcal{S}(y_0)$ we can find $\bar{x}(\cdot) \in \mathcal{S}_K(y_0)$ satisfying $\|x - \bar{x}\|_{\mathcal{C}} \leq L_r \max_{s\in[0,1]} \operatorname{dist}(x(s); K)$. This approach was initiated in [47] for differential inclusions under a much stronger inward pointing condition. Constructions proposed in proofs of these theorems are still "anticipative". Neighbouring feasible trajectories theorems imply the local $\mathcal{C}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$. We refer to [8] for the most recent neighbouring feasible trajectories theorem in the space $W^{1,1}([0,1]; \boldsymbol{R}^n)$ for F depending measurably on time and K having a smooth boundary.

When the boundary of K is non smooth and F is discontinuous with respect to the time, neighbouring feasible trajectories theorems are no longer valid neither in $W^{1,1}([0,1];\boldsymbol{R}^n)$ nor even in $\mathcal{C}$. Some counterexamples are proposed in [8] with a state independent F and K being a convex cone in $\boldsymbol{R}^2$.

• The **third** approach was initiated in [9] for control systems of the form (1) with bounded f and when the boundary of K is smooth. It was further assumed that the sets $f(t,x,U)$ are convex and closed. Then an extension to unbounded set-valued maps was proposed in [10], where, instead of Lipschitz continuity of $\mathcal{S}_K(\cdot)$, its pseudo-Lipschitz continuity was investigated. The advantage of this construction is of formulating the problem into the viability theory framework and therefore proceeding in a non anticipative way. Let $y(\cdot) \in \mathcal{S}_K(x_1)$ and let $u(\cdot)$ be a control corresponding to $y(\cdot)$. Set $r(t,x) := \operatorname{dist}\big(f(t,x,u(t)); f(t,x,U) \cap T_K(x)\big)$,

$$G(t,x) := f(t,x,U) \cap B(f(t,x,u(t)), r(t,x))$$

and consider the differential inclusion

$$\begin{cases} z'(t) \in G(t,z(t)) \\ z(0) = x_2 \in K. \end{cases}$$

By the measurable viability theorem from [56] it has a viable trajectory $z(\cdot) \in \mathcal{S}_K(x_2)$. An analysis of $z(\cdot)$ yields estimates $\|y-z\|_{W^{1,1}} \leq L_r|x_1 - x_2|$ for a constant $L_r \geq 0$ depending only on $|x_0| + r$.

• To **summarize**, the first construction allows us to prove the local $\mathcal{C}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$. When the boundary of K is sufficiently smooth, the second and third approaches imply $W^{1,1}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$ even when F is only measurable in time. Still counterexamples to neighbouring feasible trajectories theorems do exist when K is an intersection of sets with smooth boundaries and F is discontinuous in time. The third approach provides a non anticipative construction when in addition the sets $f(t,x,U)$ are convex and closed.

• We propose now an **inverse mapping theorem** approach to $\mathcal{C}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$ on $\operatorname{Int} K$ for constraints with possibly nonsmooth boundary and f measurable with respect to the time.

Recall that for a subset $Q \subset \boldsymbol{R}^n$ with nonempty boundary ∂Q the oriented distance to ∂Q is the function $d_Q(\cdot) : \boldsymbol{R}^n \to \boldsymbol{R}$ defined by

$$d_Q(x) := \operatorname{dist}(x;Q) - \operatorname{dist}(x;\boldsymbol{R}^n \setminus Q).$$

See [37, 38] for interesting and detailed discussions about relations between smoothness of the oriented distance function and regularity of the boundary of Q. We write $d_Q \in \mathcal{C}^{1,1}_{loc}$ if for any $r > 0$ there exists $\varepsilon > 0$ such that the gradient of $d_Q(\cdot)$ is Lipschitz continuous on $(\partial Q \cap B(0,r)) + B(0,\varepsilon)$.

In Theorem 2.2 below we suppose that K is an intersection of closed sets having smooth boundaries

$$\begin{cases} \emptyset \neq K = \cap_{j=1}^{m} K_j \\ \text{for some closed subsets } K_j \subset \boldsymbol{R}^n \text{ such that } d_j \in \mathcal{C}^{1,1}_{loc} \quad \forall\, j = 1, ..., m \ ; \\ 0 \notin co\{\nabla d_j(x) \mid j \in J(x)\} \quad \forall\, x \in \partial K, \end{cases} \tag{19}$$

where $d_j = d_{K_j}$ and $J(x) = \{j \,|\, x \in \partial K_j\}$. Then $T_{K_j}(x) = \{v \mid \langle \nabla d_j(x), v\rangle \leq 0\}$ for every $j \in J(x)$, $T_K(x) = \bigcap_{j=1}^{m} T_{K_j}(x)$ and $T_K(x) = C_K(x)$.

Can we expect the local $\mathcal{C}-$Lipschitz continuity of the set-valued map $\mathcal{S}_K(\cdot)$ on the interior of K when f is only measurable in time and the inward pointing condition holds true ? A positive answer is provided in [13] on the basis of an inverse mapping theorem of set-valued analysis. Lipschitz-like properties of inverse maps in an abstract setting were studied by many authors, see for instance [2, 5, 42, 64] and the references contained therein.

Consider the Banach space $\mathcal{C}_0 := \{x(\cdot) \in \mathcal{C} \,|\, x(0) = 0\}$ with the norm $\|\cdot\|_{\mathcal{C}}$ and define for every $y_0 \in K$ the set

$$\mathcal{K}(y_0) := \{x(\cdot) \in \mathcal{C} \mid x(0) = y_0,\ x(t) \in K \quad \forall\, t \in [0,1]\}\,.$$

Then $\mathcal{K}(y_0)$ is a complete metric space with the metric induced by the $\mathcal{C}$-norm. We associate with every $y_0 \in K$ the set-valued map $G_{y_0} : \mathcal{K}(y_0) \rightsquigarrow \mathcal{C}_0$ defined by $G_{y_0}(y(\cdot)) = \mathcal{S}(y_0) - y(\cdot)$ and consider the problem: find $x(\cdot; y_0) \in \mathcal{K}(y_0)$ such that $0 \in \mathcal{S}(y_0) - x(\cdot; y_0)$. Observe that $x(\cdot; y_0) \in \mathcal{S}_K(y_0)$. Thus $\mathcal{C}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$ may be studied by investigation of Lipschitz behaviour of the set-valued map $G_{y_0}^{-1}(0)$ with respect to the parameter y_0. This question is related to set-valued implicit function theorems (see [42] for a very clear exposition of this topic).

Let U be a complete separable metric space and $f : [0,1] \times \boldsymbol{R}^n \times U \to \boldsymbol{R}^n$. In the theorem below we impose the following assumptions on f

$$\begin{cases} f \text{ is Lebesgue measurable in } t \text{ and continuous in } u\,; \\ \forall\, r > 0,\, \exists\, k_r \geq 0 \ \text{ such that } f(t, \cdot, u) \text{ is } k_r - \text{Lipschitz on } B(0,r)\ \forall\, t,\, u\,; \\ \exists\, \gamma > 0 \ \text{ such that } \ \sup_{u \in U} |f(t,x,u)| \leq \gamma(1 + |x|) \ \ \forall\, t,\, x\,; \\ f(t,x,U) \text{ is compact } \ \forall\, t,\, x \end{cases} \tag{20}$$

and the inward pointing condition

$$\begin{cases} \forall\, r > 0,\, \exists\, \rho_r > 0 \ \text{ such that } \ \forall\, x \in \partial K \cap B(0,r),\ \forall\, t \in [0,1], \\ \exists\, v_{t,x} \in co\, f(t,x,U) \text{ satisfying } \ \langle \nabla d_j(x), v_{t,x}\rangle \leq -\rho_r \quad \forall\, j \in J(x). \end{cases} \tag{21}$$

Theorem 2.2 ([13]). *Assume (19) - (21) and that f is differentiable with respect to x. Then the set-valued map $\mathcal{S}_K(\cdot)$ is locally $\mathcal{C}-$Lipschitz on* Int K.

To simplify the discussion of methodology for proving Theorem 2.2 via set-valued inverse mapping theorems we consider only the case when in addition

$$f(t,x,U) \text{ is convex for all } t,\, x.$$

The relaxation theorem and some variational arguments allow to remove this assumption. Then $\mathrm{Graph}(G_{y_0})$ is closed for every $y_0 \in K$. Assumptions (19)-(21), convexity of $f(t,x,U)$ and the measurable viability theorem from [56] imply that $\mathcal{S}_K(x_0) \neq \emptyset$ for all $x_0 \in K$. Furthermore, for every $x_0 \in K$ and $r > 0$ there exists $c(x_0,r) \geq 0$ such that for any $x_1,\, x_2 \in K \cap B(x_0,r)$ and all $\bar{z}(\cdot) \in \mathcal{S}(x_1)$, $\bar{y}(\cdot) \in \mathcal{K}(x_1)$, we can find $z(\cdot) \in \mathcal{S}(x_2)$ and $y(\cdot) \in \mathcal{K}(x_2)$ satisfying $\|z(\cdot) - \bar{z}(\cdot)\|_{\mathcal{C}} + \|y(\cdot) - \bar{y}(\cdot)\|_{\mathcal{C}} \leq c(x_0,r)|x_1 - x_2|$.

Theorem 2.2 is deduced from a result similar to the classical implicit function theorem. Indeed the underlying idea is to show that for any $y_0 \in \mathrm{Int}\, K$ and any $\bar{x}(\cdot) \in \mathcal{S}_K(y_0) \subset \mathcal{K}(y_0)$ the "derivative" of G_{y_0} at $\bar{x}$ is surjective. However $\mathcal{K}(y_0)$ being a metric space and G_{y_0} being a set-valued map, derivatives have to be replaced by *set-valued variations* and surjectivity by an *uniform covering property* of variations. Furthermore, Lipschitz continuity of the inverse is replaced by *pseudo-Lipschitz continuity* (also called Aubin continuity in [42]), because "surjectivity" at $\bar{x}$ implies Lipschitz-like behaviour of the inverse only in a neighborhood of $(0,\bar{x})$. In such framework a general inverse mapping theorem from [49] can be applied to deduce the local $\mathcal{C}-$Lipschitz continuity of $\mathcal{S}_K(\cdot)$ on $\mathrm{Int}\, K$.

Definition 2.3 ([49]). Let $\Phi : X \rightsquigarrow Y$ be a set-valued map from a metric space X to a Banach space Y. The variation of Φ at $(\bar{x},\bar{y}) \in \mathrm{Graph}(\Phi)$ is the closed subset of Y defined by

$$\Phi^{(1)}(\bar{x},\bar{y}) := \mathrm{Limsup}_{h\to 0+} \frac{\Phi(B(\bar{x},h)) - \bar{y}}{h}.$$

In the above Limsup stands for the Painlevé-Kuratowski upper limit of sets (see for instance [65, 5]). In other words $v \in \Phi^{(1)}(\bar{x},\bar{y})$ whenever there exist $h_i > 0$ converging to 0 and $v_i \in Y$ converging to v such that $\bar{y} + h_i v_i \in \Phi(B(\bar{x},h_i))$.

Let (Y,d_Y) be a metric space. The Hausdorff semidistance between two subsets A and C of Y is defined by

$$e(A;C) := \sup_{a\in A} \mathrm{dist}_Y(a;C) \in \boldsymbol{R}_+ \cup \{+\infty\},$$

where we set $e(A;C) = +\infty$ if one of the subsets A, C is empty.

Definition 2.4 ([2]). Let $\Psi : X \rightsquigarrow Y$ be a set-valued map from a metric space (X,d_X) to a metric space (Y,d_Y). Ψ is called pseudo-Lipschitz at $(\bar{\zeta},\bar{\xi}) \in \mathrm{Graph}\,(\Psi)$ if there exist $L > 0$ and $\eta > 0$ such that

$$e(\Psi(\zeta) \cap B_Y(\bar{\xi},\eta);\, \Psi(\zeta')) \leq L d_X(\zeta,\zeta'),\;\; \forall\, \zeta,\, \zeta' \in B_X(\bar{\zeta},\eta).$$

The main assumption of the inverse mapping theorem [49, Theorem 6.1] is a uniform covering property of variations. In terms of our setting, denoting by $\mathcal{B}_{\mathcal{C}_0}$ the closed unit ball in $\mathcal{C}_0$, the uniform covering property at $x_0 \in \mathrm{Int}\, K$ means that for some $\rho > 0$, $\varepsilon_0 > 0$ and $\varepsilon > 0$,

$$\rho \mathcal{B}_{\mathcal{C}_0} \subset G^{(1)}_{y_0}(y(\cdot), z(\cdot) - y(\cdot)) \tag{22}$$

for all $x_1, y_0 \in B(x_0, \varepsilon_0) \subset K$, $\bar{x}(\cdot) \in \mathcal{S}_K(x_1)$, $y(\cdot) \in \mathcal{K}(y_0)$ and $z(\cdot) \in \mathcal{S}(y_0)$ satisfying $||\bar{x}(\cdot) - y(\cdot)||_{\mathcal{C}} + ||\bar{x}(\cdot) - z(\cdot)||_{\mathcal{C}} \le \varepsilon$. If such uniform covering condition holds true, then the inverse set-valued map $G^{-1}_{y_0}$ defined by

$$G^{-1}_{y_0}(\zeta) = \{\xi \mid \zeta \in G_{y_0}(\xi)\}$$

is pseudo-Lipschitz on a neighbourhood of $(z(\cdot) - y(\cdot), y(\cdot))$ for all $z(\cdot)$, $y(\cdot)$ as above. Furthermore L and η of Definition 2.4 do not depend on $\bar{x}(\cdot)$, $y(\cdot)$ and $z(\cdot)$. That is for some $c \ge 0$, $\eta_0 > 0$, $\eta > 0$ and all $x_1, y_0 \in B(x_0, \eta_0) \subset K$, $\bar{x}(\cdot) \in \mathcal{S}_K(x_1)$, $y(\cdot) \in \mathcal{K}(y_0)$ and $z(\cdot) \in \mathcal{S}(y_0)$ satisfying $||\bar{x}(\cdot) - y(\cdot)||_{\mathcal{C}} + ||\bar{x}(\cdot) - z(\cdot)||_{\mathcal{C}} \le \eta$ the following inequality holds

$$\mathrm{dist}_{\mathcal{C}}(y(\cdot); G^{-1}_{y_0}(0)) \le c\|z(\cdot) - y(\cdot)\|_{\mathcal{C}}.$$

This implies Theorem 2.2, because it is enough to choose $z(\cdot) \in \mathcal{S}(x_2)$ and $y(\cdot) \in \mathcal{K}(x_2)$ satisfying $\|z(\cdot) - \bar{x}(\cdot)\|_{\mathcal{C}} + \|y(\cdot) - \bar{x}(\cdot)\|_{\mathcal{C}} \le c(x_0, \eta_0)|x_1 - x_2|$.

The following lemma allows to check the covering property.

Lemma 2.5 ([13]). *Assume (19) - (21), that f is differentiable with respect to x and $x_0 \in \mathrm{Int}\, K$. Then for some $\alpha > 0$, $\varepsilon > 0$ and $\sigma > 0$, for any $x_1 \in B(x_0, \varepsilon) \subset K$ and $\bar{x}(\cdot) \in \mathcal{S}_K(x_1)$ with $\max_{t\in[0,1]} d_K(\bar{x}(t)) > -\alpha$, the following holds true: if $y_0 \in B(x_0, \varepsilon)$, $y(\cdot) \in \mathcal{K}(y_0)$, $z(\cdot) \in \mathcal{S}(y_0)$ are such that $||y(\cdot) - \bar{x}(\cdot)||_{\mathcal{C}} + ||z(\cdot) - \bar{x}(\cdot)||_{\mathcal{C}} \le \varepsilon$, then for any control $u(\cdot)$ satisfying $z'(t) = f(t, z(t), u(t))$ a.e., there exist $\delta > 0$, a measurable selection $v(t) \in co\, f(t, z(t), U)$ a.e. in $[0, 1]$ and a solution $\overline{w}(\cdot)$ to the linear system*

$$\begin{cases} w'(t) = \frac{\partial f}{\partial x}(t, z(t), u(t))w(t) + r(t)(v(t) - z'(t)), \quad r(t) \ge 0 \ \text{ a.e. in } [0, 1] \\ w(0) = 0, \end{cases} \tag{23}$$

such that $||\overline{w}(\cdot)||_{\mathcal{C}} \le \frac{1}{2}$,

$$\max_{t\in[0,\delta]} d_K(y(t)) < 0 \ \text{ and } \ \langle \nabla d_j(y(t)), \overline{w}(t)\rangle \le -\sigma \quad \forall\, t \in (\delta, 1],\ j \in J(y(t)). \tag{24}$$

Observe that (23) is a linear control system with non negative scalar controls. To check that variations of G_{y_0} do have a uniform covering property, consider α, ε, σ as in Lemma 2.5. We may assume that $\sigma < 1$ and $\varepsilon < \frac{\alpha}{2}$. Let $x_1,\ y_0 \in B(x_0, \varepsilon)$, $\bar{x}(\cdot) \in \mathcal{S}_K(x_1)$, $y(\cdot) \in \mathcal{K}(y_0)$ and $z(\cdot) \in \mathcal{S}(y_0)$ be such that $||y(\cdot) - \bar{x}(\cdot)||_{\mathcal{C}} + ||z(\cdot) - \bar{x}(\cdot)||_{\mathcal{C}} \le \varepsilon$. If $\max_{t\in[0,1]} d_K(\bar{x}(t)) \le -\alpha$, then

$y(\cdot) + h\mathcal{B}_{\mathcal{C}_0} \subset \mathcal{K}(y_0)$ for all sufficiently small $h > 0$. Hence $z(\cdot) - y(\cdot) + h\mathcal{B}_{\mathcal{C}_0} \subset G_{y_0}(B(y(\cdot), h))$ and therefore $\mathcal{B}_{\mathcal{C}_0} \subset G_{y_0}^{(1)}(y(\cdot), z(\cdot) - y(\cdot))$. Consider next the case $\max_{t\in[0,1]} d_K(\bar{x}(t)) > -\alpha$ and let $\overline{w}(\cdot)$ be as in Lemma 2.5.

By the variational equation of control theory (see for instance [48]), there exist $\overline{w}_h(\cdot)$ converging uniformly to $\overline{w}(\cdot)$ as $h \to 0+$ such that $z(\cdot) + h\overline{w}_h(\cdot) \in \mathcal{S}(y_0)$. Let $w(\cdot) \in \mathcal{C}_0$ be such that $\|w(\cdot)\|_{\mathcal{C}} \le \frac{\sigma}{2}$. From (24) we deduce that for all small $h > 0$, $y(\cdot) + h(\overline{w}_h(\cdot) - w(\cdot)) \in \mathcal{K}(y_0) \cap \mathcal{B}_{\mathcal{C}}(y(\cdot), h)$. Therefore for all small $h > 0$,

$$z(\cdot) - y(\cdot) + hw(\cdot) = z(\cdot) + h\overline{w}_h(\cdot) - (y(\cdot) + h(\overline{w}_h(\cdot) - w(\cdot))) \in G_{y_0}(\mathcal{B}_{\mathcal{C}}(y(\cdot), h)),$$

implying that $w(\cdot) \in G_{y_0}^{(1)}(y(\cdot), z(\cdot) - y(\cdot))$. Thus (22) holds true with ρ replaced by $\frac{\sigma}{2}$. Therefore variations do have the announced uniform covering property.

3. Value Function and Optimal Synthesis

Let $\varphi : \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ be an extended-real-valued lower semicontinuous, bounded from below function and f, U, K be as in the introduction. Consider the Mayer optimal control problem

$$\text{minimize}\,\{\varphi(x(1)) \mid x(\cdot) \in \mathcal{S}_K(x_0)\} \tag{25}$$

and let $V : [0, 1] \times \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ be the value function associated to it by (11). Then $\bar{x}(\cdot) \in \mathcal{S}_K(x_0)$ is optimal for the Mayer problem if and only if $V(\cdot, \bar{x}(\cdot)) \equiv const \neq +\infty$. Therefore if $K = \boldsymbol{R}^n$ and $V \in C^1$, then, by (12), the set-valued map $\Lambda : [0, 1] \times \boldsymbol{R}^n \rightsquigarrow U$ given by

$$\Lambda(t, x) := \{u \in U \mid H(t, x, -V'_x(t, x)) = \langle -V'_x(t, x), f(t, x, u)\rangle\}$$

can be seen as an optimal synthesis for the optimal control problem (25). Indeed a trajectory/control pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ is optimal for (25) if and only if $\bar{u}(t) \in \Lambda(t, \bar{x}(t))$ a.e. Thus the set of optimal trajectories coincides with the set of trajectories of

$$x'(t) \in f(t, x(t), \Lambda(t, x(t))) \quad \text{a.e. in } [0, 1], \quad x(0) = x_0. \tag{26}$$

If V is not differentiable, but f is sufficiently smooth with respect to x, then it is still possible to express the optimal synthesis using superdifferentials of the value function. Recall [33] that the superdifferential of a function $g : \boldsymbol{R}^n \to \boldsymbol{R}$ at x is a closed convex, possibly empty, subset of $\boldsymbol{R}^n$ defined by

$$\partial_+ g(x) = \left\{ p \in \boldsymbol{R}^n \mid \limsup_{y\to x} \frac{g(y) - g(x) - \langle p, y - x\rangle}{|y - x|} \le 0 \right\}.$$

Assume next that f is time independent, that $f(x, U)$ is closed and convex for every x and define for all $t \in (0, 1)$ and $x \in \boldsymbol{R}^n$

$$\Lambda(t, x) := \{u \in U \mid p_t + \langle p_x, f(x, u)\rangle = 0 \ \ \forall\, (p_t, p_x) \in \partial_+ V(t, x)\}. \tag{27}$$

If $\partial_+ V(t,x) = \emptyset$, then we set $\Lambda(t,x) = \emptyset$. Observe that $f(x, \Lambda(t,x))$ is closed and convex (possibly empty) for all $t \in (0,1)$ and $x \in \boldsymbol{R}^n$. From [52, Theorems 4.1 and 4.3], it follows that for smooth enough f and φ, the differential inclusion (26) with $\Lambda(t,x)$ defined by (27) characterizes all optimal trajectories. There is no analogue of this result when f is only Lipschitz continuous. Also, in general, the set-valued map $(t,x) \rightsquigarrow f(x, \Lambda(t,x))$ is not upper semicontinuous.

Directional derivatives of V seem to be better adapted to express "synthesis equations" for optimal trajectories to encompass problems with Lipschitz dynamics and state constraints. For all $t \in [0,1)$, $x \in K$ such that $V(t,x) \neq +\infty$ and all $\bar{v} \in \boldsymbol{R}^n$ the contingent derivative of V at (t,x) in the direction $(1,\bar{v})$ is defined by

$$D_\uparrow V(t,x)(1,\bar{v}) := \liminf_{h \to 0+,\ v \to \bar{v}} \frac{V(t+h, x+hv) - V(t,x)}{h}.$$

We associate to it the set

$$G(t,x) = \{v \in f(t,x,U) \mid D_\uparrow V(t,x)(1,v) \leq 0\}.$$

The proof of the next result is immediate.

Proposition 3.1. *Assume that V is locally Lipschitz on $[0,1] \times K$. Then $x(\cdot) \in \mathcal{S}_K(x_0)$ is optimal for the Mayer problem if and only if*

$$x'(t) \in G(t, x(t)) \text{ a.e. in } [0,1], \quad x(0) = x_0. \tag{28}$$

A refinement of the results of the previous section allows to deduce the following two theorems about the local Lipschitz continuity of the value function.

Theorem 3.2. *Assume that φ is locally Lipschitz, that for every $r > 0$ there exists $L_r > 0$ such that $f(\cdot,\cdot,u)$ is L_r−Lipschitz on $[0,1] \times B(0,r)$ for all $u \in U$, that f is continuous with respect to u, that the sets $f(t,x,U)$ are closed and for some $\gamma > 0$, $\sup_{u \in U} |f(t,x,u)| \leq \gamma(1+|x|)$ for all $(t,x) \in [0,1] \times \boldsymbol{R}^n$. If for every $t \in [0,1]$ and $x \in \partial K$, $f(t,x,U) \cap \operatorname{Int} C_K(x) \neq \emptyset$, then V is locally Lipschitz on $[0,1] \times K$.*

Theorem 3.3. *Assume (19) - (21), that f is differentiable with respect to x and that φ is locally Lipschitz. Then V is locally Lipschitz on $[0,1] \times \operatorname{Int} K$.*

Theorem 3.2 and Proposition 3.1 allow to characterize all optimal trajectories of the optimal control problem (25) as trajectories of the differential inclusion (28) when the inward pointing condition is satisfied. The differential inclusion (28) is not simple to handle because, in general, the set-valued map G neither has convex values nor it is upper semicontinuous. If $K = \boldsymbol{R}^n$ and V is semiconcave (see [19] for sufficient conditions for such regularity of V), then the set-valued map $(t,x) \rightsquigarrow G(t,x)$ is upper semicontinuous. However, in general, for control systems under state constraints V is not semiconcave.

In **conclusion**, the optimal synthesis problem presents the same difficulty than investigation of control systems under state constraints mentioned in the introduction - it leads to control systems (differential inclusions) having highly irregular right-hand sides.

An **alternative** way to characterize optimal trajectories is to consider an extended constrained control system under an extended state constraint that we now describe. Below we denote by $epi(V)$ the epigraph of V defined by

$$epi(V) = \{(t, x, r) \in [0, 1] \times K \times \boldsymbol{R} \mid r \geq V(t, x)\}.$$

Assume that $V(0, x_0) < +\infty$ and consider the following viability problem

$$\begin{cases} s'(t) &=& 1, \;\; s(0) = 0 \\ x'(t) &=& f(t, x(t), u(t)), \;\; u(t) \in U \text{ a.e. in } [0, 1], \;\; x(0) = x_0 \\ z'(t) &=& 0, \;\; z(0) = V(0, x_0) \\ (s(t), x(t), z(t)) &\in& epi(V) \;\; \text{for all} \;\; t \in [0, 1]. \end{cases} \tag{29}$$

Then a trajectory $\bar{x}(\cdot)$ of (1), (2) is optimal for the Mayer problem (25) if and only if for some real-valued absolutely continuous functions $s(\cdot)$ and $z(\cdot)$ defined on $[0, 1]$, the triple $(s(\cdot), \bar{x}(\cdot), z(\cdot))$ satisfies (29).

Observe that if (20) holds true and $f(t, x, U)$ is convex for every $(t, x) \in [0, 1] \times \boldsymbol{R}^n$, then V is lower semicontinuous and therefore $epi(V)$ is a closed set. The viability problem (29) is a new control system under a state constraint where two very simply evolving variables (s, z) were added. Such transformations, introduced in [3], now became standard in various applications of viability theory.

It is worth to underline that algorithms for solving (12) approximate numerically the value function and optimal controls, but not super/subdifferentials. Note that once a viable trajectory of (29) has been found, an optimal control can be associated to it by a measurable selection theorem.

Regular optimal synthesis should not be expected for general nonlinear control systems. However a locally Lipschitz continuous approximate optimal synthesis can be derived via non smooth analysis techniques, see [29].

4. Value Function and Maximum Principle

Let $\varphi : \boldsymbol{R}^n \to \boldsymbol{R}$ be locally Lipschitz and f, U, K be as in the introduction. Consider again the Mayer minimization problem (25). This section illustrates how Lemma 2.5, Theorem 3.3 and arguments of convex analysis can be applied to derive a normal first order necessary optimality condition involving the generalized gradient of $V(0, \cdot)$ at $x_0 \in \operatorname{Int} K$. The result we state below is by no means the most general. Notations $\partial\varphi(x_0)$ and $\partial_x V(0, x_0)$ stand respectively for the Clarke generalized gradient of φ and of $V(0, \cdot)$ at x_0.

Let $NBV([0,1];\boldsymbol{R}^n)$ (Normalized Bounded Variations) denote the space of functions $\psi : [0,1] \to \boldsymbol{R}^n$ of bounded variation on $[0,1]$, continuous from the right on $(0,1)$ and such that $\psi(0) = 0$. The norm of $\psi(\cdot) \in NBV([0,1];\boldsymbol{R}^n)$, $\|\psi\|_{TV}$, is the total variation of $\psi(\cdot)$ on $[0,1]$.

Theorem 4.1 ([27]). *Assume (19) - (21), that f is differentiable with respect to x and let $x_0 \in \operatorname{Int} K$. If $\bar{x}(\cdot) \in \mathcal{S}_K(x_0)$ is optimal for problem (25) and $\bar{u}(\cdot)$ is a control corresponding to $\bar{x}(\cdot)$, then there exist $\psi(\cdot) \in NBV([0,1];\boldsymbol{R}^n)$ and $p(\cdot) \in W^{1,1}([0,1];\boldsymbol{R}^n)$ satisfying the adjoint equation (13), the maximum principle (14), the transversality condition $-p(1) - \psi(1) \in \partial\varphi(\bar{x}(1))$ and linked to the value function by the inclusion*

$$-p(0) \in \partial_x V(0, x_0). \tag{30}$$

Furthermore $\psi(\cdot)$ satisfies (16) for a positive (scalar) Radon measure μ on $[0,1]$ and a Borel measurable $\nu(\cdot) : [0,1] \to \boldsymbol{R}^n$ as in (15).

It follows from [14], that for a state constraint K with smooth boundary and under more general assumptions on f, $-p(t) - \psi(t) \in \partial_x V(t, \bar{x}(t))$ for a.e. $t \in [0,1]$ satisfying $\bar{x}(t) \in \operatorname{Int} K$ and also that a related inclusion holds true for a.e. $t \in [0,1]$ such that $\bar{x}(t) \in \partial K$.

Sketch of proof of Theorem 4.1. By Theorem 3.3, V is locally Lipschitz on $[0,1] \times \operatorname{Int} K$. Consider the linearized control system

$$w'(t) = \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))w(t) + v(t), \quad v(t) \in T_{cof(t,\bar{x}(t),U)}(f(t, \bar{x}(t), \bar{u}(t))) \text{ a.e.} \tag{31}$$

and define the convex sets

$$\mathcal{S}^L = \left\{ w(\cdot) \in W^{1,1}([0,1];\boldsymbol{R}^n) \mid w(\cdot) \text{ is a trajectory of (31)} \right\},$$

$$\mathcal{K}^L = \{ w(\cdot) \in \mathcal{C} \mid w(t) \in C_K(\bar{x}(t)) \ \ \forall\, t \in [0,1] \}.$$

Then $\operatorname{Int} \mathcal{K}^L = \{ w(\cdot) \in \mathcal{C} \mid w(t) \in \operatorname{Int} C_K(\bar{x}(t)) \ \forall\, t \in [0,1] \}$. As in Lemma 2.5, it can be shown that for every $\theta \in \boldsymbol{R}^n$ there exists $w(\cdot) \in \mathcal{S}^L \cap \operatorname{Int} \mathcal{K}^L$ such that $w(0) = \theta$. Recall that the Clarke directional derivative of a locally Lipschitz function $g : \boldsymbol{R}^n \to \boldsymbol{R}$ at $y_0 \in \boldsymbol{R}^n$ in the direction θ is defined by

$$g^0(y_0)(\theta) = \limsup_{y \to y_0,\ h \to 0+} \frac{g(y + h\theta) - g(y)}{h}.$$

Let $V_x^0(0, x_0)(\theta)$ be defined as above for $g(\cdot) = V(0, \cdot)$, $y_0 = x_0$ and let $w(\cdot) \in \mathcal{S}^L \cap \operatorname{Int} \mathcal{K}^L$. Consider any sequence $h_i \to 0+$. By the variational inclusion from [48] there exist $x_i(\cdot) \in \mathcal{S}(x_0 + h_i w(0))$ such that $\frac{x_i(\cdot) - \bar{x}(\cdot)}{h_i}$ converges uniformly to $w(\cdot)$ when $i \to \infty$. Then for all large i, $x_i(\cdot) \in \mathcal{S}_K(x_0 + h_i w(0))$. Since V is nondecreasing along viable trajectories, it follows that $\varphi(x_i(1)) \geq V(0, x_0 + h_i w(0))$. Therefore, from the optimality of $\bar{x}(\cdot)$ we deduce that $\varphi^0(\bar{x}(1))(w(1)) +$

$V_x^0(0,x_0)(-w(0)) \geq 0$. Denoting by $\mathrm{cl}\,\mathcal{S}^L$ the closure of $\mathcal{S}^L$ in the space $\mathcal{C}$, we obtain

$$\varphi^0(\bar{x}(1))(w(1)) + V_x^0(0,x_0)(-w(0)) \geq 0 \quad \forall\, w(\cdot) \in \mathrm{cl}\,\mathcal{S}^L \cap \mathrm{Int}\,\mathcal{K}^L. \tag{32}$$

Define the linear operator $\gamma : \mathcal{C} \to \boldsymbol{R}^n \times \boldsymbol{R}^n$ by $\gamma(x(\cdot)) = (x(0), x(1))$ for all $x(\cdot) \in \mathcal{C}$. For a subset $\mathcal{E} \subset \mathcal{C}$, let $[\mathcal{E}]^+$ denote its positive polar cone. Inequality (32) implies that for some $a \in -\partial_x V(0,x_0)$ and $b \in \partial\varphi(\bar{x}(1))$ we have $\gamma^*(a,b) \in [\mathrm{cl}\,\mathcal{S}^L \cap \mathrm{Int}\,\mathcal{K}^L]^+ = [\mathrm{cl}\,\mathcal{S}^L]^+ + [\mathcal{K}^L]^+$. Hence for some $\beta \in [\mathcal{K}^L]^+$

$$\gamma^*(a,b) - \beta \in [\mathrm{cl}\,\mathcal{S}^L]^+. \tag{33}$$

Using [74] it can be deduced that there exists $\psi(\cdot) \in NBV([0,1];\boldsymbol{R}^n)$ satisfying (16) for a positive (scalar) Radon measure μ on $[0,1]$ and a Borel measurable $\nu(\cdot) : [0,1] \to \boldsymbol{R}^n$ as in (15) such that for every $x(\cdot) \in \mathcal{C}$, $\beta(x) = \int_0^1 x(t)d\psi(t)$ (the Stieltjes integral) and $\|\beta\| = \|\psi\|_{TV}$ (see [27] for details). Observe that $\mathcal{S}^L$ is the set of trajectories of a linear control system without state constraint. A direct analysis of (33) allows to conclude that $b = -p(1) - \psi(1)$ and $a = p(0)$ for some $p(\cdot)$ as in (13), (14). □

Remark 4.2. The derived necessary optimality condition is normal. Assumption (21) yields $\mathcal{S}^L \cap \mathrm{Int}\,\mathcal{K}^L \neq \emptyset$. Without assuming (21) this intersection may be empty. Still a necessary optimality condition can be obtained by applying the separation theorem to the convex sets $\mathcal{S}^L$ and $\mathrm{Int}\,\mathcal{K}^L$. The necessary condition is then abnormal ($\lambda = 0$) and $p(0) = 0$ (see [27, Proof of Theorem 3.4], where a similar result was derived for a differential inclusion under state and end point constraints).

• **Maximum principle of the Bolza problem.** In the next section we shall use the maximum principle of a Bolza optimal control problem under state and end point constraints that we recall now. For every $x(\cdot) \in \mathcal{S}_K(x_0)$ let us denote by $\mathcal{U}(x(\cdot))$ the set of all controls corresponding to $x(\cdot)$, that is $u(\cdot) \in \mathcal{U}(x(\cdot))$ if and only if $u : [0,1] \to U$ is Lebesgue measurable and $x'(t) = f(t,x(t),u(t))$ a.e.

Let $K_1 \subset \boldsymbol{R}^n$ and $\ell : [0,1] \times \boldsymbol{R}^n \times U \to \boldsymbol{R}$. Consider the minimization problem

$$\min\{\varphi(x(1)) + \int_0^1 \ell(s,x(s),u(s))ds \,|\, x(\cdot) \in \mathcal{S}_K(x_0), u(\cdot) \in \mathcal{U}(x(\cdot)), x(1) \in K_1\}. \tag{34}$$

Denote by $M(n \times n)$ the set of $n \times n$ matrices and for every $\lambda \geq 0$, define the Hamiltonian $H_\lambda : [0,1] \times \boldsymbol{R}^n \times \boldsymbol{R}^n \to \boldsymbol{R} \cup \{+\infty\}$ of the Bolza problem by

$$H_\lambda(t,x,p) = \sup_{u \in U}(\langle p, f(t,x,u)\rangle - \lambda\ell(t,x,u)).$$

Definition 4.3. A trajectory/control pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ of (1), (2) with $\bar{x}(1) \in K_1$ satisfies the maximum principle (of problem (34)) if there exist $\lambda \in \{0,1\}$, $\psi(\cdot) \in NBV([0,1]; \boldsymbol{R}^n)$ and $p(\cdot) \in W^{1,1}([0,1]; \boldsymbol{R}^n)$ not vanishing simultaneously such that for some $\pi_1 \in \boldsymbol{R}^n$ and some integrable $A(\cdot) : [0,1] \to M(n \times n)$, $\pi(\cdot) : [0,1] \to \boldsymbol{R}^n$

$$-p'(t) = A(t)^*(p(t) + \psi(t)) - \lambda\pi(t) \;\text{ a.e. in } [0,1], \tag{35}$$

$$-p(1) - \psi(1) \in \lambda\pi_1 + N_{K_1}(\bar{x}(1)), \tag{36}$$

$$\langle p(t) + \psi(t), f(t, \bar{x}(t), \bar{u}(t))\rangle - \lambda\ell(t, \bar{x}(t), \bar{u}(t)) = H_\lambda(t, \bar{x}(t), p(t) + \psi(t)) \;\text{ a.e.} \tag{37}$$

and (16) holds true for a positive (scalar) Radon measure μ on $[0,1]$ and a Borel measurable $\nu(\cdot) : [0,1] \to \boldsymbol{R}^n$ satisfying (15). The maximum principle is called normal if $\lambda = 1$.

Remark 4.4. If there exist $\varepsilon > 0$ and $k(\cdot) \in L^1(0,1)$ such that $f(t, \cdot, u)$ and $\ell(t, \cdot, u)$ are $k(t)-$Lipschitz on $B(\bar{x}(t), \varepsilon)$ for a.e. $t \in [0,1]$ and all $u \in U$, then under some mild regularity assumptions on f and ℓ, every optimal trajectory/control pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ satisfies the maximum principle of Definition 4.3 for some $A(t) \in \partial_x f(t, \bar{x}(t), \bar{u}(t))$ (generalized Jacobian of $f(t, \cdot, \bar{u}(t))$ at $\bar{x}(t)$), $\pi(t) \in \partial_x \ell(t, \bar{x}(t), \bar{u}(t))$ (generalized gradient of $\ell(t, \cdot, \bar{u}(t))$ at $\bar{x}(t)$), $\pi_1 \in \partial\varphi(\bar{x}(1))$ and some $(\lambda, p, \psi) \neq 0$ (see [82]).

Furthermore, some other maximum principles that differ only in the adjoint equation (having the Hamiltonian or the Euler-Lagrange form) can be rewritten with the adjoint equation like (35). For instance, if $p(\cdot) \in W^{1,1}([0,1]; \boldsymbol{R}^n)$ and $k(\cdot) \in L^1([0,1]; \boldsymbol{R}_+)$, are such that $|p'(t)| \le k(t)|p(t) + \psi(t)| + \lambda k(t)$ a.e., then it is not difficult to find $A(t)$ and $\pi(t)$ not necessarily related to the generalized Jacobian of $f(t, \cdot, \bar{u}(t))$ and to the generalized gradient of $\ell(t, \cdot, \bar{u}(t))$ such that $A(\cdot)$ and $\pi(\cdot)$ are integrable and (35) holds true. This is particularly useful for deducing normality for other forms of the maximum principle.

• **Normality of the maximum principle.** We provide next a sufficient condition for normality of the maximum principle of the Bolza problem when $x_0 \in \operatorname{Int} K$ under the following *outward* pointing condition

$$\begin{cases} \forall\, r > 0,\ \exists\, \sigma_r > 0 \quad \text{such that} \quad \forall\, t \in [0,1],\ \ \forall\, x \in \partial K \cap B(0,r), \\ \exists\, v_{t,x} \in co\, f(t,x,U) \text{ satisfying } \langle \nabla d_j(x), v_{t,x}\rangle \ge \sigma_r \quad \forall\, j \in J(x). \end{cases} \tag{38}$$

In Proposition 4.5 below, (38) can be assumed only for $x = \bar{x}(t)$ and all $t \in [0,1]$.

Consider a trajectory/control pair $(\bar{x}(\cdot), \bar{u}(\cdot))$ of (1), (2) satisfying the maximum principle (of problem (34)) for some (λ, p, ψ), π_1, $\pi(\cdot)$ and $A(\cdot)$. The reachable set from zero at time $t \in [0,1]$ of the linear control system

$$w'(t) = A(t)w(t) + v(t), \;\; v(t) \in T_{cof(t,\bar{x}(t),U)}(f(t, \bar{x}(t), \bar{u}(t))) \;\text{ a.e.} \tag{39}$$

is a convex cone in $\boldsymbol{R}^n$ defined by

$$\mathcal{R}(t) = \{w(t) \mid w(\cdot) \text{ is a trajectory of (39) on } [0,t],\ w(0) = 0\}.$$

Proposition 4.5. *Assume (19), (20), (38). Let* $x_0 \in \operatorname{Int} K$ *and let a trajectory/control pair* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *of (1), (2) with* $\bar{x}(1) \in K_1$ *satisfy the maximum principle for some* $(\lambda, p, \psi) \neq 0$ *and* $A(\cdot)$*. Further assume that* $\bar{x}([0,1]) \cap \partial K \neq \emptyset$*,* $\operatorname{Int} C_{K_1}(\bar{x}(1)) \cap \operatorname{Int} T_K(\bar{x}(1)) \neq \emptyset$ *and for* $t_0 = \min\{t \in [0,1] \mid \bar{x}(t) \in \partial K\}$ *we have* $\operatorname{Int} T_K(\bar{x}(t_0)) \subset \mathcal{R}(t_0)$*. Then* $\lambda = 1$*.*

Proof. As in [53], normality will follow if there exists a solution $w(\cdot)$ to (8) satisfying $0 \neq w(1) \in \operatorname{Int} C_{K_1}(\bar{x}(1))$. Fix any $0 \neq w_1 \in \operatorname{Int} C_{K_1}(\bar{x}(1)) \cap \operatorname{Int} T_K(\bar{x}(1))$. Combining constructions of [53, Corollary 6.4] and [11, Proof of Theorem 3.2] (made backward in time) we obtain a solution $w(\cdot)$ to (39) defined on $[t_0, 1]$ and satisfying $w(1) = w_1$ and $w(t) \in \operatorname{Int} T_K(\bar{x}(t))$ for all $t \in [t_0, 1]$. Because $\operatorname{Int} T_K(\bar{x}(t_0)) \subset \mathcal{R}(t_0)$, $w(\cdot)$ can be extended on the time interval $[0, t_0]$ by a solution to (39) in such way that $w(0) = 0$. Since $\bar{x}([0, t_0)) \subset \operatorname{Int} K$, the proof is complete. □

When the end point is free, i.e. $K_1 = \boldsymbol{R}^n$, some sufficient *inward* pointing conditions for normality can be found in [11, 53, 54]. In [73] for a free end point optimal control problem the normal maximum principle was derived by the penalization of a state constraint satisfying the inward pointing condition.

5. Regularity of Optimal Trajectories and Controls

Consider again the Bolza problem (34). We discuss here regularity (with respect to the time) of trajectories and controls satisfying the normal maximum principle.

Let the Hamiltonian H_1 be defined as in the previous section. Recall that $H_1(t, x, \cdot)$ is convex and for every $q \in \boldsymbol{R}^n$ and $u \in U$ satisfying $H_1(t, x, q) = \langle q, f(t, x, u)\rangle - \ell(t, x, u)$, we have $f(t, x, u) \in \partial_p H_1(t, x, q)$, where $\partial_p H_1(t, x, q)$ denotes the subdifferential of $H_1(t, x, \cdot)$ at q. In Proposition 5.1 below we consider only Lipschitz continuous optimal trajectories. A sufficient condition for the existence of a Lipschitz continuous optimal trajectory for the Bolza problem can be found for instance in [21, 54]. We also impose some global assumptions on H_1. However most of them can be localized for $H_1(t, \bar{x}(t), \cdot)$, where $\bar{x}(\cdot)$ is a trajectory of the control system (1), (2) under investigation. Define

$$\mathcal{M} := \{(t, x, f(t, x, u), r) \mid t \in [0,1],\ x \in K,\ u \in U,\ r \geq \ell(t, x, u)\}.$$

Proposition 5.1 ([53]). *Assume (19), that* $\mathcal{M}$ *is closed, that* H_1 *is continuous and that* $H_1(t, x, \cdot)$ *is differentiable for all* $(t, x) \in [0,1] \times K$*. Let* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *satisfy the normal maximum principle for some* $p(\cdot)$*,* $\psi(\cdot)$*. If* $\bar{x}(\cdot)$ *is Lipschitz, then* $\bar{x}(\cdot) \in C^1([0,1])$*, the mapping* $(0,1) \ni t \mapsto \frac{\partial H_1}{\partial p}(t, \bar{x}(t), p(t) + \psi(t))$ *is continuous and* $\bar{x}'(t) = \frac{\partial H_1}{\partial p}(t, \bar{x}(t), p(t) + \psi(t))$ *for every* $t \in (0,1)$*. Furthermore,* $\psi(\cdot)$ *is continuous on* $(0,1)$ *provided for every* $t \in (0,1)$*,* $x \in \partial K$ *and* $p,\ q \in \boldsymbol{R}^n$

the following implication holds true

$$\left(\begin{array}{l} p-q \in N_K(x) \\ H_1(t,x,p) = H_1(t,x,q) \\ \frac{\partial H_1}{\partial p}(t,x,p) = \frac{\partial H_1}{\partial p}(t,x,q) \end{array}\right) \quad \Longrightarrow \quad p = q. \tag{40}$$

Observe that (40) is satisfied, in particular, when $\frac{\partial H_1}{\partial p}(t,x,\cdot)$ is strictly monotone in the directions normal to K at every $x \in \partial K$, i.e. when for every $t \in (0,1)$ and all $p \neq q \in \boldsymbol{R}^n$ satisfying $p - q \in N_K(x)$ we have

$$\left\langle \frac{\partial H_1}{\partial p}(t,x,p) - \frac{\partial H_1}{\partial p}(t,x,q), p-q \right\rangle > 0.$$

If for all t, x, the Hamiltonian $H_1(t,x,\cdot)$ is twice differentiable and has strictly positive second derivative, then the last inequality is satisfied for all $p \neq q \in \boldsymbol{R}^n$.

Note that if $\frac{\partial H_1}{\partial p}$ is continuous, then every $\bar{x}(\cdot)$ satisfying the normal maximum principle (for some control $\bar{u}(\cdot)$) is Lipschitz continuous. The next theorem provides a sufficient condition for the absolute continuity of the mapping $\psi(\cdot)$ on $(0,1)$.

Theorem 5.2 ([53]). *Assume (19), that $\mathcal{M}$ is closed, H_1 is continuous, $H_1(t,x,\cdot)$ is differentiable for all $(t,x) \in [0,1] \times K$ and $\frac{\partial H_1}{\partial p}$ is locally Lipschitz on $[0,1] \times K \times \boldsymbol{R}^n$. Further assume that for every $t \in (0,1)$, $x \in \partial K$ and p, $q \in \boldsymbol{R}^n$ the implication (40) holds true and that for every $r > 0$ there exist $k_r > 0$ and $\bar{\varepsilon} > 0$ such that for all $t \in [0,1]$, $x \in \partial K \cap B(0,r)$ and p, $q \in B(0,r)$ we have*

$$(p-q \in N_K(x) \cap B(0,\bar{\varepsilon})) \implies \left\langle \frac{\partial H_1}{\partial p}(t,x,p) - \frac{\partial H_1}{\partial p}(t,x,q), p-q \right\rangle \geq k_r |p-q|^2.$$

If $(\bar{x}(\cdot), \bar{u}(\cdot))$ satisfy the normal maximum principle for some $p(\cdot)$, $\psi(\cdot)$, then $\psi(\cdot)$ is absolutely continuous on $(0,1)$ and $\bar{x}'(\cdot)$ is absolutely continuous on $[0,1]$.

Moreover, if $p(\cdot)$ is Lipschitz, then $\psi(\cdot)$ is Lipschitz on $(0,1)$ and $\bar{x}'(\cdot)$ is Lipschitz on $[0,1]$.

The coercivity assumption of the above theorem is automatically satisfied for all p, $q \in B(0,r)$ if $H_1(t,x,\cdot)$ is twice differentiable, $\frac{\partial^2 H_1}{\partial p^2}$ is continuous and $\frac{\partial^2 H_1}{\partial p^2}(t,x,\cdot) > 0$ for all $(t,x) \in [0,1] \times \partial K$. Observe that $p(\cdot)$ is Lipschitz whenever $A(\cdot)$ and $\pi(\cdot)$ of the maximum principle are essentially bounded.

The proof of the above theorem relies on an induction argument developed in [62] for linear control systems, convex Lagrangian ℓ and convex state constraints. Some sufficient conditions for Hölder continuity of derivatives of optimal trajectories can be found in [12] and [78].

Corollary 5.3. *Under all the assumptions of Theorem 5.2, suppose that* H_1 *is locally Lipschitz. If* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *satisfy the normal maximum principle for some* $p(\cdot)$, $\psi(\cdot)$, *then there exists an absolutely continuous function* $\phi : [0,1] \to \boldsymbol{R}$ *such that* $\phi(t) = \ell(t, \bar{x}(t), \bar{u}(t))$ *a.e. Moreover if* $p(\cdot)$ *is Lipschitz, then* $\phi(\cdot)$ *is Lipschitz.*

By Theorem 5.2, $\phi(\cdot)$ defined by $\phi(t) = \langle p(t)+\psi(t), \bar{x}'(t)\rangle - H_1(t, \bar{x}(t), p(t)+\psi(t))$ for $t \in (0,1)$ and $\phi(0) = \phi(0+)$, $\phi(1) = \phi(1-)$ is absolutely continuous. Furthermore, by (37), $\phi(t) = \ell(t, \bar{x}(t), \bar{u}(t))$ a.e. in $[0,1]$ implying the above Corollary.

Regularity of $\psi(\cdot)$, $p(\cdot)$ and $\bar{x}'(\cdot)$ helps to study regularity of optimal controls with respect to the time.

Proposition 5.4. *Assume that* U *is a closed convex subset of* $\boldsymbol{R}^m$, *that* f, ℓ *are defined on* $[0,1] \times \boldsymbol{R}^n \times \boldsymbol{R}^m$ *and are continuous, and that* $f(t,x,\cdot)$, $\ell(t,x,\cdot)$ *are differentiable for all* $(t,x) \in [0,1] \times K$. *Define* $\mathcal{H} : [0,1] \times \boldsymbol{R}^n \times \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}$ *by*

$$\mathcal{H}(t,x,p,u) := \langle p, f(t,x,u)\rangle - \ell(t,x,u).$$

If for some $\Phi : \boldsymbol{R}_+ \times \boldsymbol{R}_+ \to \boldsymbol{R}$ *satisfying* $\lim_{k\to+\infty} \frac{\Phi(k,r)}{k} = +\infty$ *for every* $r > 0$,

$$\begin{cases} i) \;\; \ell(t,x,u) \geq \Phi(|f(t,x,u)|, r) \;\forall\, (t,x) \in [0,1] \times (K \cap B(0,r)),\; r > 0,\; u \in U; \\ ii) \;\; \lim_{|u|\to+\infty} \left(\inf_{(t,x)\in[0,1]\times(K\cap B(0,r))} \ell(t,x,u)\right) = +\infty \quad \forall\, r > 0; \\ iii) \;\; \forall\, (t,x,p) \in [0,1] \times K \times \boldsymbol{R}^n,\;\; \forall\, u_1,\; u_2 \in U \quad \text{with} \quad u_1 \neq u_2, \\ \qquad \left\langle \frac{\partial \mathcal{H}}{\partial u}(t,x,p,u_1) - \frac{\partial \mathcal{H}}{\partial u}(t,x,p,u_2), u_2 - u_1 \right\rangle > 0, \end{cases}$$

then for every $(t,x,p) \in [0,1] \times K \times \boldsymbol{R}^n$, *there exists a unique* $v(t,x,p) \in U$ *such that* $H_1(t,x,p) = \mathcal{H}(t,x,p,v(t,x,p))$. *Furthermore* $H_1(\cdot,\cdot,\cdot)$ *and* $v(\cdot,\cdot,\cdot)$ *are continuous.*

Corollary 5.5. *Under all the assumptions of Proposition 5.4, let* $(\bar{x}(\cdot), \bar{u}(\cdot))$ *satisfy the normal maximum principle for some* $p(\cdot)$, $\psi(\cdot)$. *If* $\psi(\cdot)$ *is continuous on* $(0,1)$, *then there exists a continuous mapping* $u_0(\cdot) : [0,1] \to U$ *such that* $u_0(\cdot) = \bar{u}(\cdot)$ *a.e. in* $[0,1]$. *Consequently* $\bar{x}(\cdot)$ *is Lipschitz.*

Furthermore, if $v(\cdot,\cdot,\cdot)$ *is locally Lipschitz on* $[0,1] \times K \times \boldsymbol{R}^n$ *and* $\psi(\cdot)$ *is absolutely continuous on* $(0,1)$, *then* $u_0(\cdot)$ *is absolutely continuous and if* $p(\cdot)$ *is Lipschitz and* $\psi(\cdot)$ *is Lipschitz on* $(0,1)$, *then* $u_0(\cdot)$ *is Lipschitz.*

Proof. By Proposition 5.4, for every $t \in (0,1)$ there exists a unique $u_0(t) := v(t, \bar{x}(t), p(t) + \psi(t)) \in U$ with $\mathcal{H}(t, \bar{x}(t), p(t) + \psi(t), u_0(t)) = H_1(t, \bar{x}(t), p(t) + \psi(t))$. Let $\widetilde{\psi}(\cdot) \in \mathcal{C}$ be such that $\widetilde{\psi}(\cdot) = \psi(\cdot)$ on $(0,1)$. Then $u_0(\cdot) := v(\cdot, \bar{x}(\cdot), p(\cdot) + \widetilde{\psi}(\cdot))$ is continuous on $[0,1]$. By (37), $u_0(t) = \bar{u}(t)$ for a.e. $t \in [0,1]$. The remaining statements follow from the very definition of $u_0(\cdot)$. □

For f affine with respect to controls, the local Lipschitz continuity of $v(\cdot,\cdot,\cdot)$ follows, for instance, from the assumptions of [78]. For such control systems this question is related to Lipschitz continuity of a conjugate function. Assume that $f(t,x,u) = a(t,x) + g(t,x)u$ for some continuous $a : [0,1] \times \boldsymbol{R}^n \to \boldsymbol{R}^n$ and $g : [0,1] \times \boldsymbol{R}^n \to M(n \times m)$, where $u \in \boldsymbol{R}^m$ and let $\ell : [0,1] \times \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}$ be continuous and convex with respect to the last variable. Consider a closed convex subset $U \subset \boldsymbol{R}^m$ and assume that $\ell(t,x,\cdot)$ is differentiable and for all $u_1,\, u_2 \in U$ with $u_1 \neq u_2$

$$\left\langle \frac{\partial \ell}{\partial u}(t,x,u_1) - \frac{\partial \ell}{\partial u}(t,x,u_2), u_1 - u_2 \right\rangle > 0.$$

Then all the conclusions of Proposition 5.4 hold true whenever there exists $\theta : \boldsymbol{R}_+ \times \boldsymbol{R}_+ \to \boldsymbol{R}$ such that for every $r > 0$, $\lim_{k \to +\infty} \frac{\theta(k,r)}{k} = +\infty$ and $\ell(t,x,u) \geq \theta(|u|, r)$ for all $(t,x) \in [0,1] \times (K \cap B(0,r))$ and $u \in U$. Let $\iota_U(\cdot)$ be the indicator function of U and denote by $\ell^F(t,x,\cdot)$ the Fenchel conjugate of $\ell(t,x,\cdot) + \iota_U(\cdot)$. Then, by the uniqueness of $v(t,x,p)$, the function $y \mapsto \ell^F(t,x,y)$ is differentiable and $v(t,x,p) = \frac{\partial \ell^F}{\partial y}(t,x,g(t,x)^* p)$. If $\frac{\partial \ell^F}{\partial y}(\cdot,\cdot,\cdot)$ is locally Lipschitz and $g(\cdot,\cdot)$ is locally Lipschitz, then also $v(\cdot,\cdot,\cdot)$ is locally Lipschitz.

When the mapping

$$[0,1] \times K \times \boldsymbol{R}^n \ni (t,x,q) \rightsquigarrow \Upsilon(t,x,q) := \{u \in U \mid \mathcal{H}(t,x,q,u) = H_1(t,x,q)\}$$

is multivalued, then it may happen that several controls give rise to the same trajectory and an optimal control may be discontinuous. If Υ enjoys some regularity properties, then, taking its selections, it is still possible to deduce the existence of regular optimal controls from the regularity of $p(\cdot) + \psi(\cdot)$ and $\bar{x}(\cdot)$. In general however we can not expect Lipschitz and even continuous optimal controls in the nonlinear case even under assumptions like in Theorem 5.2 guaranteing Lipschitz continuity of derivatives of optimal trajectories.

Acknowledgement

The author is grateful to Piernicola Bettiol, Piermarco Cannarsa, Pierre Cardaliaguet, Asen Dontchev and Sylvain Sorin for many constructive remarks concerning the preliminary version of this manuscript.

References

[1] Arutyunov A. V. & Aseev S. M. (1997) *Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints*, SIAM J. Control Optim., **35**, 930–952.

[2] Aubin J.-P. (1984) *Lipschitz behavior of solutions to convex minimization problems,* Mathematics of Oper. Res., **9**, 87–111.

[3] Aubin J.-P. (1991) VIABILITY THEORY, Birkhäuser, Boston.

[4] Aubin J.-P. & Cellina A. (1984) Differential Inclusions, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin.

[5] Aubin J.-P. & Frankowska H. (1990) Set-Valued Analysis, Birkhäuser, Boston.

[6] Bardi M. & Capuzzo Dolcetta I. (1997) Optimal Control and Viscosity Solutions of Hamilton - Jacobi - Bellman Equations, Birkhäuser, Boston.

[7] Barron E. N. & Jensen R. (1990) *Semicontinuous viscosity solutions of Hamilton-Jacobi equations with convex Hamiltonians*, Comm. Partial Diff. Eq., **15**, 1713–1742.

[8] Bettiol P., Bressan A. & Vinter R. B. (2009) *On trajectories satisfying a state constraint: $W^{1,1}$-estimates and counter-examples*, Preprint.

[9] Bettiol P., Cardaliaguet P. & Quincampoix M. (2006) *Zero-sum state constrained differential games: Existence of value for Bolza problem*, International Journal of Game Theory, **34**, 495–527.

[10] Bettiol P. & Frankowska H. (2007) *Regularity of solution maps of differential inclusions for systems under state constraints*, Set-Valued Analysis, **15**, 21–45.

[11] Bettiol P. & Frankowska H. (2007) *Normality of the maximum principle for non convex constrained Bolza problems*, J. Diff. Eqs., **243**, 256–269.

[12] Bettiol P. & Frankowska H. (2008) *Hölder continuity of adjoint states and optimal controls for state constrained problems*, Applied Math. and Optim., **57**, 125–147.

[13] Bettiol P. & Frankowska H. (2010) *On the solution map to control systems with multiple state constraints*, Preprint.

[14] Bettiol P. & Vinter R. B. (2010) *Sensitivity interpretations of the co-state variable for optimal control problems with state constraints*, SIAM J. Control Optim., **48**, 3297–3317.

[15] Bonnans J. F. & Hermant A. (2009) *Second-order analysis for optimal control problems with pure and mixed state constraints*, Annales de l'I.H.P. - Nonlinear Analysis, **26**, 561–598.

[16] Byrnes Ch. (1998) *The solution of nonlinear Lagrange and Bolza problems via Riccati partial differential equations*, J. Math. Systems, Estimation and Control, **8**, 1–54.

[17] Byrnes Ch. & Frankowska H. (1992) *Unicité des solutions optimales et absence de chocs pour les équations d'Hamilton-Jacobi-Bellman et de Riccati,* Comptes-Rendus de l'Académie des Sciences, **315**, Série 1, Paris, 427–431.

[18] Byrnes Ch. & Frankowska H. (1998) *Uniqueness of optimal trajectories and the nonexistence of shocks for Hamilton-Jacobi-Bellman and Riccati partial differential equations*, Lecture Notes in Nonlinear Analysis, **2**, J. Schauder Center for Nonlinear Studies, 89–112.

[19] Cannarsa P. & Frankowska H. (1991) *Some characterizations of optimal trajectories in control theory*, SIAM J. Control Optim., **29**, 1322–1347.

[20] Cannarsa P., Frankowska H. & Marchini E. (2009) *Existence and Lipschitz regularity of solutions to Bolza problems in optimal control*, Trans. Amer. Math. Soc., **361**, 4491–4517.

[21] Cannarsa P., Frankowska H. & Marchini E. (2009) *On Bolza optimal control problems with constraints*, Discrete and Continuous Dynamical Systems, **11**, 629–653.

[22] Cannarsa P., Frankowska H. & Sinestrari C. (2000) *Optimality conditions and synthesis for the minimum time problem*, J. Set-Valued Analysis, **8**, 127–148.

[23] Cannarsa P., Gozzi F. & Soner H. M. (1991) *A boundary value problem for Hamilton-Jacobi equations in Hilbert spaces*, Applied Math. and Optim., **24**, 197–220.

[24] Cannarsa P. & Sinestrari C. (2004) SEMICONCAVE FUNCTIONS, HAMILTON-JACOBI EQUATIONS, AND OPTIMAL CONTROL, Birkhäuser, Boston.

[25] Caroff N. & Frankowska H. (1992) *Optimality and characteristics of Hamilton-Jacobi-Bellman equations,* International Series of Numerical Mathematics, **107**, Birkhäuser Verlag, Basel, 169–180.

[26] Caroff N. & Frankowska H. (1996) *Conjugate points and shocks in nonlinear optimal control*, Trans. Amer. Math. Soc., **348**, 3133–3153.

[27] Cernea A. & Frankowska H. (2006) *A connection between the maximum principle and dynamic programming for constrained control problems*, SIAM J. Control Optim., **44**, 673–703.

[28] Clarke F. (1983) OPTIMIZATION AND NONSMOOTH ANALYSIS, Wiley-Interscience, New York.

[29] Clarke F., Rifford L. & Stern R. J. (2002) *Feedback in state constrained optimal control*, ESAIM Control Optim. Calc. Var., **7**, 97–133.

[30] Clarke F. & Vinter R. B. (1987) *The relationship between the maximum principle and dynamic programming*, SIAM J. Control Optim., **25**, 1291–1311.

[31] Clarke F. & Vinter R. B. (1990) *Regularity properties of optimal controls*, SIAM J. Control Optim., **28**, 980–997.

[32] Crandall M. G. & Lions P. L. (1983) *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **277**, 1–42.

[33] Crandall M. G., Evans C. L. & Lions P. L. (1984) *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **282**, 487–502.

[34] Chryssochoos I. & Vinter R. B. (2003) *Optimal control problems on manifolds: a dynamic programming approach*, J. Math. Anal. Appl., **287**, 118–140.

[35] Dal Maso G. & Frankowska H. (2003) *Autonomous integral functionals with discontinuous nonconvex integrands: Lipschitz regularity of minimizers, DuBois-Reymond necessary conditions, and Hamilton-Jacobi equations*, Appl. Math. Optim., **48**, 39–66.

[36] Da Prato G. & Frankowska H. (2004) *Invariance of stochastic control systems with deterministic arguments*, J. Diff. Eqs., **200**, 18–52.

[37] Delfour M. & Zolesio J.-P. (1994) *Shape analysis via oriented distance functions*, J. Func. Analysis, **123**, 129–201.

[38] Delfour M. & Zolesio J.-P. (2004) *Oriented distance function and its evolution equation for initial sets with thin boundary*, SIAM J. Control Optim., **42**, 2286–2304.

[39] Dikusar V. V. & Milyutin A. A. (1989) QUALITATIVE AND NUMERICAL METHODS IN THE MAXIMUM PRINCIPLE, Nauka.

[40] Dontchev A. L. & Hager W. W. (1993) *Lipschitzian stability in nonlinear control and optimization*, SIAM J. Control Optim., **31**, 569–603.

[41] Dontchev A. L. & Hager W. W. (1998) *A new approach to Lipschitz continuity in state constrained optimal control,* Systems and Control Letters, **35**, 137–143.

[42] Dontchev A. L. & Rockafellar R. T. (2009) IMPLICIT FUNCTIONS AND SOLUTION MAPPINGS, Springer Mathematics Monographs, Springer, Dodrecht.

[43] Dubovitskii A. Y. & Milyutin A. A. (1965) *Extremal problems with constraints*, USSR Comput. Math. and Math. Physiscs, **5**, 1–80.

[44] Evans L. C. (1998) PARTIAL DIFFERENTIAL EQUATIONS, American Mathematical Society, Providence.

[45] Filippov A. F. (1967) *Classical solutions of differential equations with multivalued right-hand side*, SIAM J. Control Optim., **5**, 609–621.

[46] Fleming W. H. & Soner H. M. (2005) CONTROLLED MARKOV PROCESSES AND VISCOSITY SOLUTIONS, 2nd edition, Springer-Verlag.

[47] Forcellini F. & Rampazzo F. (1999) *On non-convex differential inclusions whose state is constrained in the closure of an open set*, *J. Differential Integral Equations*, **12**, 471–497.

[48] Frankowska H. (1987) *The maximum principle for an optimal solution to a differential inclusion with end points constraints*, SIAM J. Control Optim., **25**, 145–157.

[49] Frankowska H. (1990) *Some inverse mapping theorems*, Annales de l'I.H.P. - Nonlinear Analysis, **7**, 183–234.

[50] Frankowska H. (1991) *Lower semicontinuous solutions to Hamilton-Jacobi-Bellman equations*, Proceedings of 30th CDC IEEE Conference, Brighton, December 11–13.

[51] Frankowska H. (1993) *Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equation*, SIAM J. Control Optim., **31**, 257–272.

[52] Frankowska H. (2005) *Optimal synthesis via superdifferentials of value function*, Control and Cybernetics, **34**, 787–803.

[53] Frankowska H. (2006) *Regularity of minimizers and of adjoint states for optimal control problems under state constraints*, J. Convex Analysis, **13**, 299–328.

[54] Frankowska H. & Marchini E. (2006) *Existence and Lipschitzianity of optimal trajectories for the Bolza optimal control problem*, Calculus of Variations and PDE, **27**, 467–492.

[55] Frankowska H. & Ochal A. (2005) *On singularities of value function for Bolza optimal control problem*, JMAA, **306**, 714–729.

[56] Frankowska H. & Plaskacz S. (1996) *A measurable upper semicontinuous viability theorem for tubes*, J. of Nonlinear Analysis, TMA, **26**, 565–582.

[57] Frankowska H. & Plaskacz S. (2000) *Semicontinuous solutions of Hamilton-Jacobi-Bellman equations with degenerate state constraints*, J. Math. Anal. Appl., **251**, 818–838.

[58] Frankowska H., Plaskacz S. & Rzezuchowski T. (1995) *Measurable viability theorems and Hamilton-Jacobi-Bellman equation*, J. Diff. Eqs., **116**, 265–305.

[59] Galbraith G. & Vinter R. B. (2003) *Lipschitz continuity of optimal controls for state constrained problems*, SIAM J. Control Optim., **42**, 1727–1744.

[60] Gamkrelidze R. V. (1960) *Optimal processes with bounded phase coordinates*, Izv. Akad. Nauk, USSR, Sec. Mat., **24**, 315–356.

[61] Gavriel C. & Vinter R. B. (2010) *Regularity of minimizers for second order variational problems*, Preprint.

[62] Hager W. W. (1979) *Lipschitz continuity for constrained processes,* SIAM J. Control Optim., **17**, 321–338.

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

[64] Ioffe A. D. (2000) *Metric regularity and subdifferential calculus*, Math. Surveys, **55**, 501–558.

[65] Kuratowski K. (1966) TOPOLOGY, Academic Press/PWN.

[66] Loewen P. & Rockafellar R. T. (1991) *The adjoint arc in nonsmooth optimization*, Trans. Amer. Math. Soc., **325**, 39–72.

[67] Malanowski K. M. (1978) *On the regularity of solutions to optimal control problems for systems linear with respect to control variable*, Arch. Auto i Telemech., **23**, 227–241.

[68] Malanowski K. M. (2003) *On normality of Lagrange multipliers for state constrained optimal control problems*, Optimization, **52**, 75–91.

[69] Malanowski K. M. (2007) *Sufficient optimality conditions in stability analysis for state-constrained optimal control*, Appl. Math. Optim., **55**, 255–271.

[70] Malanowski K. M. (2007) *Stability analysis for nonlinear optimal control problems subject to state constraints*, SIAM J. Optim., **18**, 926–945.

[71] Milyutin A. A. (2000) *On a certain family of optimal control problems with phase constraint*, J. Math. Sc., **100**, 2564–2571.

[72] Pesch H. J. & Plail M. (2009) *The maximum principle of optimal control: a history of ingenious ideas and missed opportunities*, Control and Cybernetics, **38**.

[73] Rampazzo F. & Vinter R. B. (1999) *A theorem on existence of neighbouring trajectories satisfying a state constraint, with applications to optimal control*, IMA J. Math. Control Inform., **16**, 335–351.

[74] Rockafellar R. T. (1971) *Integrals which are convex functionals*, Pacific J. Math., **39**, 439–469.

[75] Rockafellar R. T. & Wets R. J.-B. (1997) VARIATIONAL ANALYSIS, Springer-Verlag, Berlin.

[76] Sarychev A. & Torres D. F. M. (2000) *Lipschitzian regularity of the minimizers for optimal control problems with control-affine dynamics*, Appl. Math. Optim., **41**, 237–254.

[77] Schättler H. (2006) *Local fields of extremals for optimal control problems with state constraints of relative degree 1*, J. Dyn. Control Syst., **12**, 563–599.

[78] Shvartsman I. A. & Vinter R. B. (2006) *Regularity properties of optimal controls for problems with time-varying state and control constraints,* Nonlinear Analysis TMA, **65**, 448–474.

[79] Soner H. M. (1986) *Optimal control with state-space constraints,* SIAM J. Control Optim., **24**, 552–561.

[80] Subbotin A. I. (1995) Generalized solutions of first order PDEs: the dynamical optimization perspective, Birkhäuser, Boston, Basel, Berlin.

[81] Subbotina N. N. (1989) *The maximum principle and the superdifferential of the value function*, Problems Control Inform. Theory, **18**, 151–160.

[82] Vinter R. B. (2000) Optimal control, Birkhäuser, Boston, Basel, Berlin.