Serdica Math. J. **39** (2013), 233–270

Serdica
Mathematical Journal

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

A SECOND-ORDER MAXIMUM PRINCIPLE IN OPTIMAL CONTROL UNDER STATE CONSTRAINTS*

Hélène Frankowska, Daniel Hoehener, Daniela Tonon

Communicated by N. Ribarska

Dedicated to Asen Dontchev and Vladimir Veliov on the occasion of their birthdays

Abstract. A second-order variational inclusion for control systems under state constraints is derived and applied to investigate necessary optimality conditions for the Mayer optimal control problem. A new pointwise condition verified by the adjoint state of the maximum principle is obtained as well as a second-order necessary optimality condition in the integral form. Finally, a new sufficient condition for normality of the maximum principle is proposed. Some extensions to the Mayer optimization problem involving a differential inclusion under state constraints are also provided.

1. Introduction. This paper is devoted to second-order necessary optimality conditions for control problems in the presence of pure state constraints

2010 *Mathematics Subject Classification:* 49K15, 49K21, 34A60, 34K35.
Key words: optimal control, second-order necessary optimality conditions, second-order tangents.
* Partial support by the European Commission (FP7-PEOPLE-2010-ITN, Grant Agreement no. 264735-SADCO) is gratefully acknowledged.

together with endpoint constraints. Let us consider first an abstract optimization problem

$$\min_{c \in C} \phi(c), \tag{1}$$

where C is a subset of a Banach space X and $\phi : X \to \mathbb{R}$ is a twice Fréchet differentiable function. Assume that $\bar{c} \in C$ is an optimal solution, i.e.

$$\phi(\bar{c}) = \min_{c \in C} \phi(c),$$

and denote by ϕ' and ϕ'' the first and second-order derivatives of ϕ. If $\bar{c}$ lies in the interior of C, then the classical results guarantee the first-order necessary optimality condition $\phi'(\bar{c}) = 0$ (Fermat rule) and the second-order condition $\phi''(\bar{c}) \geq 0$. Let us assume now that $\bar{c}$ is a boundary point of C and denote by $T^\flat_C(\bar{c})$ the adjacent cone to C at $\bar{c}$. Recall that $u \in T^\flat_C(\bar{c})$ whenever there exist $\delta > 0$ and a "path" $x : [0, \delta] \to C$ satisfying $x(0) = \bar{c}$ and such that the difference quotients $\frac{1}{h}(x(h) - x(0))$ converge to u when $h \to 0+$. The Fermat rule becomes then:

$$\phi'(\bar{c})u \geq 0 \quad \forall\, u \in T^\flat_C(\bar{c}). \tag{2}$$

This is a first-order necessary optimality condition for problem (1). Furthermore, if the set C is convex and for some $\bar{c} \in C$, we have

$$\inf_{u \in T^\flat_C(\bar{c}),\ \|u\|=1} \phi'(\bar{c})u > 0,$$

then it is not difficult to justify that $\bar{c}$ is a local minimum for the problem (1). Observe that the second-order derivative of ϕ has no influence on this conclusion. In particular it may happen that $\phi''(\bar{c})(u, u) < 0$ for some $u \in T^\flat_C(\bar{c})$.

The second derivative starts to play a role in the expression of optimality conditions when $u \in T^\flat_C(\bar{c})$ is such that $\phi'(\bar{c})u = 0$. In this case it is natural to consider not only the second-order derivative of the function ϕ but also a second-order "linearization" of the constraint. To this end, one uses second-order tangents, whose definition we recall next.

With every $u \in T^\flat_C(\bar{c})$ we associate the set of second-order tangents to C at $(\bar{c}, u)$. Namely $v \in T^{\flat(2)}_C(\bar{c}, u)$ if we can find $\delta > 0$ and a "path" $x : [0, \delta] \to C$ satisfying $x(0) = \bar{c}$ such that the difference quotients $\frac{1}{h^2}(x(h) - x(0) - hu)$ converge to v when $h \to 0+$. The Taylor formula implies then that the following second-order optimality condition holds true:

$$(3) \quad \phi'(\bar{c})v + \frac{1}{2}\phi''(\bar{c})(u,u) \geq 0 \quad \forall\, u \in T_C^{\flat}(\bar{c}),\ v \in T_C^{\flat(2)}(\bar{c},u) \quad \text{such that} \quad \phi'(\bar{c})u = 0.$$

The optimality conditions (2) and (3) are classical in mathematical programming, see for instance [5] for an overview of this subject, and were largely explored in the optimal control theory, where they have led to various second-order necessary optimality conditions, mostly in the form of an integral inequality. See for instance[4, 16, 22, 27] and the bibliographies contained therein. The space X is then usually the space of essentially bounded controls and it is assumed that every control determines a unique state trajectory, see for instance [4, 16]. The map associating to each control the corresponding state trajectory is then, under suitable assumptions on the system dynamics, Fréchet differentiable. The set C incorporates restrictions imposed on trajectories and controls.

To derive necessary optimality conditions, the optimal control problem at hand is usually reformulated as an infinite dimensional abstract optimization problem involving inequality and equality constraints. Then first and second-order optimality conditions that are dual to (2) and (3) are obtained, see for instance [5, 7, 19]. This approach requires Robinson's like constraint qualifications. The obtained necessary conditions have to be translated then in terms of the original optimal control problem. Verification of Robinson's constraint qualifications and translation of the abstract necessary conditions in terms of control problems is not an easy task and it often requires strong assumptions on control systems and on optimal controls. This is the reason why in many published papers it is assumed that optimal controls are piecewise continuous, that the control system depends on time in a continuous way and the equality and inequality constraints satisfy qualification hypotheses like linear independence of gradients of active constraints. Moreover, in some works on second-order necessary optimality conditions in optimal control, see for instance [16, 22], the proposed first and second-order conditions are strictly linked, that is each u appearing in (3) gives rise to a pair of first/second-order conditions. This is usually inconvenient for the analysis of possible candidates for optimality. Let us also underline that, especially in the context of state constrained problems, it is more natural to expect optimal controls to be merely measurable and very general first-order necessary optimality conditions are already known in this context, see for instance [26]. This creates an important gap between generality of the available results on first and second-order necessary optimality conditions.

Avoiding the reformulation of optimal control problems under state constraints in an abstract way by using a direct variational approach allows to work under less restrictive hypotheses. In particular, in [26], several very general first-order necessary optimality conditions were derived from the Ekeland variational principle stated on metric spaces. The main tools used there are penalization, limiting gradients, generalized normals and their calculus.

An alternative approach was proposed in [6, 25] by considering the space $X = \mathcal{C}([0,1];\mathbb{R}^n)$ of continuous mappings defined on the time interval $[0,1]$ and taking values in $\mathbb{R}^n$ and the subset C equal to the set of trajectories of the control system satisfying state and endpoint constraints.

It was shown that solutions of a linearized control system under linearized constraints (along a given trajectory) are tangent to the set of trajectories of the original control system under state constraints. This, inequality (2) and the duality theory of convex analysis allowed then to obtain a direct proof of the maximum principle in the presence of state constraints.

The aim of our paper is to go beyond these first-order results and to investigate second-order necessary optimality conditions using second-order tangents.

In the context of optimal control theory this approach was applied in [17, 18] by considering second-order tangents to the sets of admissible controls. An important difference with the existing literature is the fact that the analysis takes place in the framework of measurable controls and therefore larger sets of second-order tangents are considered, since weaker convergence properties are imposed (L^1 versus L^∞). In particular, an optimal control may be merely measurable. The derived integral type second-order necessary conditions for weak optimality were obtained in primal form for general control and state constraints.

In the present work we address second-order necessary optimality conditions for strong local minima (that is we work with trajectories instead of controls). To investigate the second-order tangents to the set of trajectories of a control system under constraints, we derive a second-order variational differential inclusion. Applying it to the Mayer optimal control problem under state and end point constraints, we obtain a *pointwise second-order maximum principle* and a second-order necessary optimality condition in the form of an integral inequality which extends some earlier known results to the case of strong local minimizers. The second-order maximum principles derived in Theorems 5.10, 6.9 and Corollaries 5.11, 6.10 seem to be new even in the case without state constraints.

Furthermore, for the Mayer optimization problem involving a differential inclusion under state constraints, we also obtain a second-order maximum principle. Since the presence of end point constraints requires some additional assumptions, to simplify the discussion, we first state results not involving the end point constraints, postponing to the last section their analogues in a more general setting.

We would like to underline that, in the difference with the traditional approach to the first-order necessary optimality conditions based on normal cones, see for instance [26], in the present work we explore first- and second-order tangents and the associated *graphical derivatives/variations* of set-valued maps. In particular, we prove here several new properties of second order graphical variations, see Section 2. The first-order graphical derivatives having a wide range of applications, see for instance [1], [9] and the bibliographies contained therein, we believe that further analysis of second-order graphical variations may bring new results also to other areas of variational analysis.

The outline of the paper is as follows. In Section 2 we introduce some notations and provide a few preliminary results. In Section 3 a second-order variational differential inclusion is derived. Section 4 is devoted to a second-order approximation of differential inclusions under state constraints and to second-order necessary optimality conditions for a Mayer problem. Section 5 deals with the Mayer optimal control problem under state constraints. Finally Section 6 extends results of Sections 4 and 5 to the case when endpoint constraints are present.

2. Preliminaries. The spaces of continuous and essentially bounded maps from $[0,1]$ into $\mathbb{R}^n$ are denoted respectively by $\mathbb{C}([0,1];\mathbb{R}^n)$ and $L^\infty([0,1];\mathbb{R}^n)$ and their norms by $\|\cdot\|_\infty$, while $W^{1,1}([0,1];\mathbb{R}^n)$ and $L^1([0,1];\mathbb{R}^n)$ stand respectively for the spaces of absolutely continuous and of integrable maps from $[0,1]$ to $\mathbb{R}^n$ with the usual norms $\|\cdot\|_{W^{1,1}}$ and $\|\cdot\|_1$. The space of mappings from $[0,1]$ to $\mathbb{R}^n$ having bounded total variation that are right continuous on $]0,1[$ and vanish at zero is denoted by $NBV([0,1];\mathbb{R}^n)$. The norm $\|f\|_{TV}$ of $f \in NBV([0,1];\mathbb{R}^n)$ is the total variation of f on $[0,1]$.

Partial derivatives are denoted with subscripts, for instance $f_u := \dfrac{\partial f}{\partial u}$. Similarly, partial second-order derivatives are denoted by a double subscript, i.e. $f_{xu} := \dfrac{\partial^2 f}{\partial x \partial u}$. Moreover, for y_1, $y_2 \in \mathbb{R}^n$ we will abbreviate $f_{xx}(t_0,x_0,u_0)(y_1,y_2)$ by $f_{xx}(t_0,x_0,u_0)y_1y_2$. If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at x, then we denote by $\nabla f(x)$ its gradient.

We denote the norm in $\mathbb{R}^n$ by $|\cdot|$ and by $\langle\cdot,\cdot\rangle$ the inner product. For a set $K \subset \mathbb{R}^n$, let $\overline{K}$ be its closure, ∂K its boundary, $\operatorname{int} K$ its interior, K^c its complement and $\operatorname{co} K$ its convex hull. K^- stands for the (negative) polar cone to K, i.e. $K^- = \{v \in \mathbb{R}^n \mid \langle v,k\rangle \le 0,\ \forall\, k \in K\}$. The open unit ball in $\mathbb{R}^n$ is denoted by $B := \{x \in \mathbb{R}^n \mid |x| < 1\}$. Write $B(x,r)$ for an open ball with radius $r > 0$ and center $x \in \mathbb{R}^n$ and S^{n-1} for the unit sphere.

Let X be a Banach space and $K \subset X$. The distance between a point $x \in X$ and K is defined by $\text{dist}_K(x) := \inf_{k \in K} \|x - k\|_X$. When K is a nonempty proper subset of $\mathbb{R}^n$, the *oriented distance* $b_K(\cdot)$ from K is a real valued function defined by $b_K(x) := \text{dist}_K(x) - \text{dist}_{K^c}(x)$ for all $x \in \mathbb{R}^n$. We set $b_K(\cdot) \equiv 0$ if $K = \mathbb{R}^n$. Note that any closed set $K \subset \mathbb{R}^n$ can be represented via an inequality constraint involving the oriented distance function:

$$K = \{x \in \mathbb{R}^n \mid b_K(x) \leq 0\}.$$

Let

$$B_0 := \{v = (v_1, \ldots, v_n) \in B \mid v_n = 0\} \quad \& \quad B_- := \{v = (v_1, \ldots, v_n) \in B \mid v_n < 0\}.$$

A proper closed subset K of $\mathbb{R}^n$ is said to be of class $\mathbb{C}^2$ at $x_0 \in \partial K$ if there exist an open neighborhood $\mathcal{N}$ of x_0 and a bijective map $\vartheta : \mathcal{N} \to B$ such that $\vartheta(\cdot) \in \mathbb{C}^2(\mathcal{N}; B)$, its inverse $\vartheta^{-1}(\cdot) \in \mathbb{C}^2(B; \mathcal{N})$ and

$$\vartheta^{-1}(B_-) = (\text{int } K) \cap \mathcal{N}, \ \ \vartheta^{-1}(B_0) = \partial K \cap \mathcal{N} =: \Gamma, \ \ \vartheta(\Gamma) = B_0.$$

We say that K is of class $\mathbb{C}^2$ if it is so for every $x_0 \in \partial K$.

Remark 2.1. If K is of class $\mathbb{C}^2$ at $x_0 \in \partial K$, then there exists a neighborhood $\mathcal{W}$ of x_0 such that the oriented distance $b_K(\cdot) \in \mathbb{C}^2(\mathcal{W}; \mathbb{R})$, see for instance [8, Theorem 4.3].

We recall next some definitions concerning tangent sets to a nonempty subset K of a Banach space X.

Let $\mathcal{T}$ be a metric space and $\{K_\tau\}_{\tau \in \mathcal{T}}$ be a family of subsets of X. The lower limit of K_τ at $\tau_0 \in \mathcal{T}$ is defined by:

$$\operatorname*{Liminf}_{\tau \to \tau_0} K_\tau := \left\{ v \in X \;\middle|\; \lim_{\tau \to \tau_0} \text{dist}_{K_\tau}(v) = 0 \right\}.$$

Moreover, we will write $x_i \xrightarrow{K} x$, when a sequence x_i converges to x and $x_i \in K$ for all i. First and second-order adjacent subsets are defined next.

Definition. *Let K be a closed subset of X and $x \in K$. The* adjacent cone *to K at x is the set,*

$$T_K^\flat(x) := \operatorname*{Liminf}_{h \to 0+} \frac{K - x}{h}.$$

The second-order adjacent subset *to* K *at* $(x,u) \in K \times X$ *is the set defined by,*

$$T_K^{\flat(2)}(x,u) := \operatorname*{Liminf}_{h\to 0+} \frac{K - x - hu}{h^2}.$$

Note that $T_K^{\flat(2)}(x,u) \neq \emptyset$ implies $u \in T_K^{\flat}(x)$. If K is convex, then $T_K^{\flat}(x)$ coincides with the tangent cone of convex analysis to K at x and is denoted by $T_K(x)$.

Below $C_K(x)$ stands for the Clarke tangent cone to K at x and $N_K(x) = C_K(x)^-$. If $X = \mathbb{R}^n$ then it is convenient to abbreviate,

$$N_K^1(x) := N_K(x) \cap S^{n-1}.$$

A set K is called sleek if for all $x \in K$ the contingent cone and the Clarke tangent cone to K at x do coincide (see [1] for the definition of contingent cone). In this case also $T_K^{\flat}(x) = C_K(x)$ for every $x \in K$.

Remark 2.2. If $K \subset \mathbb{R}^n$ is of class $\mathbb{C}^2$ at some $x_0 \in \partial K$, then for all $x \in \partial K$ sufficiently close to x_0, $T_K^{\flat}(x) = \{u \in \mathbb{R}^n \mid \langle \nabla b_K(x), u\rangle \leq 0\}$ and $\nabla b_K(x)$ is the unit outward normal to K at x. Furthermore $v \in T_K^{\flat(2)}(x,u)$ if and only if either $\langle \nabla b_K(x), u\rangle < 0$ or $\langle \nabla b_K(x), u\rangle = 0$ and $\langle \nabla b_K(x), v\rangle + \frac{1}{2} b_K''(x)uu \leq 0$. Consequently, in this case, $T_K^{\flat(2)}(x,u)$ is a closed affine halfspace.

Example 2.3. Let $K = \cap_{i=1}^k K_i$, where $K_i \subset \mathbb{R}^n$ are of class $\mathbb{C}^2$ for every $i = 1, \dots, k$. Denote by b_i the oriented distance associated to K_i and by $I(x_0)$ the set of active indices at x_0, that is, $i \in I(x_0)$ if and only if $x_0 \in \partial K_i$. If $0 \notin \operatorname{co}\{\nabla b_i(x_0) \mid i \in I(x_0)\}$ for every $x_0 \in \partial K$, then it is well known that K is sleek and

$$T_K^{\flat}(y_0) = \bigcap_{i\in I(y_0)} T_{K_i}^{\flat}(y_0) \;\; \forall\, y_0 \in \partial K.$$

It is not difficult to check that if $x_0 \in \partial K$, then for every $u \in T_K^{\flat}(x_0)$, a vector $v \in T_K^{\flat(2)}(x_0,u)$ if and only if

$$\langle \nabla b_i(x_0), v\rangle + \frac{1}{2} b_i''(x_0)uu \leq 0 \qquad \forall\, i \in I^{(1)}(x_0,u),$$

where $I^{(1)}(x_0,u) = \{i \in I(x_0) \mid \langle \nabla b_i(x_0), u\rangle = 0\}$. That is if $I^{(1)}(x_0,u) \neq \emptyset$, then $T_K^{\flat(2)}(x_0,u)$ is a closed convex polytope in $\mathbb{R}^n$ (an intersection of a finite family of affine halfspaces in $\mathbb{R}^n$).

For any $K \subset X$ we adopt the following convention: $K + \emptyset = \emptyset + K = \emptyset$. A useful property of the second-order adjacent set is given in the following lemma.

Lemma 2.4. *Let $K \subset X$, $x \in K$ and $u \in T^{\flat}_K(x)$. Then,*

$$T^{\flat(2)}_K(x,u) = T^{\flat(2)}_K(x,u) + C_K(x).$$

P r o o f. Since $0 \in C_K(x)$, the inclusion $\subset$ is obvious. To prove the opposite, let $v \in T^{\flat(2)}_K(x,u)$ and $\tilde{u} \in C_K(x)$. We have to show that for all $h_i \to 0+$, there exists a sequence $w_i \to v + \tilde{u}$ such that,

$$x + h_i u + h_i^2 w_i \in K, \quad \forall\, i.$$

Fix a sequence $h_i \to 0+$. Since $v \in T^{\flat(2)}_K(x,u)$, there exist $v_i \to v$ such that,

$$x + h_i u + h_i^2 v_i \in K, \quad \forall\, i.$$

Further, by the very definition of $C_K(x)$, for every sequence $x_i \xrightarrow{K} x$, there exists $\tilde{u}_i' \to \tilde{u}$ such that $x_i + h_i^2 \tilde{u}_i' \in K$ for all i. Thus, in particular, there exists $\tilde{u}_i \to \tilde{u}$ such that,

$$x + h_i u + h_i^2 (v_i + \tilde{u}_i) \in K, \quad \forall\, i.$$

Setting $w_i = v_i + \tilde{u}_i$, we end the proof. $\square$

Let $K \subset \mathbb{R}^n$ be closed and define

$$\mathcal{K} := \{x \in \mathbb{C}([0,1];\mathbb{R}^n) \mid x(t) \in K \quad \forall\, t \in [0,1]\}.$$

Lemma 2.5. *Let $x, y, w, \bar{w} \in \mathbb{C}([0,1];\mathbb{R}^n)$ be such that $\bar{w} \in T^{\flat(2)}_{\mathcal{K}}(x,y)$ and for all $t \in [0,1]$, $w(t) \in C_K(x(t))$ and $\operatorname{int} C_K(x(t)) \neq \emptyset$. Then $\bar{w} + w \in T^{\flat(2)}_{\mathcal{K}}(x,y)$.*

P r o o f. Let x, y, w, $\bar{w}$ be as above. Since $\bar{w} \in T^{\flat(2)}_{\mathcal{K}}(x,y)$, for every $h > 0$ there exists $\bar{w}_h$ such that $\bar{w}_h \to \bar{w}$ uniformly when $h \to 0+$ and for all $h > 0$,

$$x + hy + h^2 \bar{w}_h \in \mathcal{K}. \tag{4}$$

It follows from [6, Lemma 4.1] that there exists $\hat{w} \in \mathbb{C}([0,1];\mathbb{R}^n)$ satisfying $\hat{w}(t) \in \operatorname{int} C_K(x(t))$ for all $t \in [0,1]$. Thus, by convexity of Clarke's tangent cone, for every $i \in \mathbb{N}$, $w_i(t) := \frac{1}{i}\hat{w}(t) + \left(1 - \frac{1}{i}\right) w(t) \in \operatorname{int} C_K(x(t))$ for all $t \in [0,1]$. We claim that for every $i \in \mathbb{N}$, there exists $\varepsilon_i > 0$ such that for all $t \in [0,1]$,

$$z + [0, \varepsilon_i]B(w_i(t); \varepsilon_i) \subset K, \quad \forall\, z \in B(x(t), \varepsilon_i) \cap K. \tag{5}$$

Indeed, from [24, Thm. 2] it follows that for all $i \in \mathbb{N}$ and all $t \in [0,1]$ there exists $\varepsilon_i^t > 0$ for which (5) is satisfied with ε_i replaced by ε_i^t. Then one can use the compactness of $x([0,1])$ to deduce (5) for some ε_i and all t. Applying (4) and (5) and using that $x + hy + h^2\bar{w}_h \to x$ uniformly when $h \to 0+$, we find that for all h small enough,

$$x(t) + hy(t) + h^2\left(\bar{w}_h(t) + w_i(t)\right) \in K \qquad \forall\, t \in [0,1].$$

It follows that $\bar{w} + w_i \in T_{\mathcal{K}}^{\flat(2)}(x,y)$ for all $i \in \mathbb{N}$. Finally, since $\bar{w} + w_i \to \bar{w} + w$ uniformly and the second-order adjacent set $T_{\mathcal{K}}^{\flat(2)}(x,y)$ is closed, we deduce that $\bar{w} + w \in T_{\mathcal{K}}^{\flat(2)}(x,y)$. □

Lemma 2.6. *Let $x, w \in \mathbb{C}([0,1]; \mathbb{R}^n)$ be such that for all $t \in [0,1]$, $w(t) \in C_K(x(t))$ and $\operatorname{int} C_K(x(t)) \neq \emptyset$. Then $w \in T_{\mathcal{K}}^{\flat}(x)$.*

Proof. Consider a sequence $h_j > 0$ converging to $0+$ and let $\hat{w}$, w_i be as in the proof of Lemma 2.5. Then by (5) for every i we have $x + h_j w_i \in \mathcal{K}$ for all j large enough and therefore $w_i \in T_{\mathcal{K}}^{\flat}(x)$. Since $T_{\mathcal{K}}^{\flat}(x)$ is closed in $\mathbb{C}([0,1]; \mathbb{R}^n)$ and w_i converge uniformly to w the proof follows. □

We shall also need the following tangent sets.

Definition. *Let K be a closed subset of $\mathbb{R}^n$ and $x \in K$. The Dubovitskii-Milyutin cone to K at x is the set,*

$$D_K(x) := \{v \in \mathbb{R}^n \mid \exists\, \varepsilon > 0, \forall\, h \in [0,\varepsilon],\ x + hB(v,\varepsilon) \subset K\}.$$

For any $(x,u) \in K \times \mathbb{R}^n$, define

$$D_K^2(x,u) := \{v \in \mathbb{R}^n \mid \exists\, \varepsilon > 0, \forall\, h \in [0,\varepsilon],\ x + hu + h^2B(v,\varepsilon) \subset K\}.$$

Remark 2.7. The set $D_K^2(x,u)$ was introduced in [20], see also [21]. Observe that $D_K(x)$ and $D_K^2(x,u)$ are open and

$$\operatorname{int} C_K(x) \subset D_K(x) \subset T_K^{\flat}(x), \quad D_K^2(x,u) \subset T_K^{\flat(2)}(x,u).$$

Lemma 2.8. *Let K be a closed subset of $\mathbb{R}^n$, $x \in K$ and $u \in \mathbb{R}^n$. Then,*

$$D_K^2(x,u) + \operatorname{int} C_K(x) \subset D_K^2(x,u).$$

Proof. Let $v \in D_K^2(x,u)$ and $\tilde{u} \in \operatorname{int} C_K(x)$. It is enough to show that there exists $\varepsilon > 0$ such that for all small $h > 0$

$$x + hu + h^2 B(v + \tilde{u}, \varepsilon) \subset K. \tag{6}$$

Since $v \in D_K^2(x,u)$, for some $\varepsilon_v > 0$ and for all $h \in [0, \varepsilon_v]$, $x + hu + h^2 B(v, \varepsilon_v) \subset K$. Moreover, since $\tilde{u} \in \operatorname{int} C_K(x)$, there exists $\varepsilon_{\tilde{u}} > 0$ such that for all $x' \in K \cap B(x, \varepsilon_{\tilde{u}})$ we have $x' + [0, \varepsilon_{\tilde{u}}]B(\tilde{u}, \varepsilon_{\tilde{u}}) \subset K$. Thus, in particular, for $\varepsilon = \min\{\varepsilon_v, \varepsilon_{\tilde{u}}\}$ and all $h \geq 0$ sufficiently small, (6) holds true. □

We define next the first and second-order directional derivatives/variations of set-valued maps.

Definition. *Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a set-valued map, locally Lipschitz around some $x \in \mathbb{R}^n$ and let $y \in F(x)$. The* adjacent derivative *$dF(x,y)$ is the set-valued map defined by,*

$$dF(x,y)(u) \coloneqq \operatorname*{Liminf}_{h \to 0+} \frac{F(x+hu) - y}{h} \quad \forall\, u \in \mathbb{R}^n.$$

For $v_1 \in dF(x,y)(u_1)$, the second-order adjacent variation *$d^2F(x,y,u_1,v_1)$ is the set-valued map defined by,*

$$d^2F(x,y,u_1,v_1)(u_2) \coloneqq \operatorname*{Liminf}_{h \to 0+} \frac{F(x + hu_1 + h^2 u_2) - y - hv_1}{h^2} \quad \forall\, u_2 \in \mathbb{R}^n.$$

Remark 2.9. It is well-known that if F has convex images and is Lipschitz around x, then for any $y \in F(x)$, the images of $dF(x,y)$ are convex and,

$$dF(x,y)(0) = T_{F(x)}(y) \quad \text{and} \quad dF(x,y)(u) + T_{F(x)}(y) = dF(x,y)(u), \tag{7}$$

see for instance [1, Prop. 5.2.6].

The adjacent variations can also be expressed using the distance function:

Proposition 2.10. *Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be Lipschitz around x, $y \in F(x)$ and $v_1 \in dF(x,y)(u_1)$. Then for all $u_2 \in \mathbb{R}^n$, the following holds true: $v_2 \in d^2F(x,y,u_1,v_1)(u_2)$ if and only if,*

$$\lim_{h \to 0+} \frac{\operatorname{dist}_{F(x+hu_1+h^2u_2)}(y + hv_1 + h^2 v_2)}{h^2} = 0.$$

The next result provides a useful property of the second-order adjacent variation.

Proposition 2.11. *Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a set-valued map with convex images which is Lipschitz around x, $y \in F(x)$ and $v_1 \in dF(x,y)(u_1)$. Then for all $u_2 \in \mathbb{R}^n$,*

$$(8)\quad T_{F(x)}(y) + T_{dF(x,y)(u_1)}(v_1) + d^2F(x,y,u_1,v_1)(u_2) = d^2F(x,y,u_1,v_1)(u_2).$$

In particular,

$$(9)\quad dF(x,y)(0)+dF(x,y)(u_1)+d^2F(x,y,u_1,v_1)(u_2) = v_1+d^2F(x,y,u_1,v_1)(u_2).$$

Proof. *Step 1*: We start by showing that,

$$(10)\qquad T_{F(x)}(y) + d^2F(x,y,u_1,v_1)(u_2) = d^2F(x,y,u_1,v_1)(u_2).$$

Since $0 \in T_{F(x)}(y)$, the inclusion "$\supset$" is obvious. To show the opposite, let $w \in T_{F(x)}(y)$, $v_2 \in d^2F(x,y,u_1,v_1)(u_2)$ and $h_i \to 0+$ be an arbitrary sequence. By the very definition of the second-order adjacent derivative, there exists a sequence $v_2^i \to v_2$ such that,

$$(11)\qquad y + h_i v_1 + h_i^2 v_2^i \in F(x + h_i u_1 + h_i^2 u_2) \qquad \forall\, i \in \mathbb{N}.$$

Moreover, since $v_1 \in dF(x,y)(u_1)$, by (7), $w + v_1 \in dF(x,y)(u_1)$. Hence for some $\beta_i \to w + v_1$ we have

$$(12)\qquad y + h_i\beta_i \in F(x + h_i u_1) \subset F(x + h_i u_1 + h_i^2 u_2) + o(h_i)B \qquad \forall\, i \in \mathbb{N},$$

where $o(\alpha)/\alpha \to 0$, when $\alpha \to 0+$. Then, multiplying (11) by $1 - h_i$ and (12) by h_i, and taking their sum, we obtain, thanks to the convexity of the images of F,

$$y + h_i v_1 + h_i^2\left(v_2^i + \beta_i - v_1\right) \in F(x + h_i u_1 + h_i^2 u_2) + o(h_i^2)B.$$

Using that $\beta_i \to v_1 + w$ when $i \to \infty$, we get $w + v_2 \in d^2F(x,y,u_1,v_1)(u_2)$.

Step 2: By (10), to prove (8) it suffices to show that,

$$(13)\qquad T_{dF(x,y)(u_1)}(v_1) + d^2F(x,y,u_1,v_1)(u_2) = d^2F(x,y,u_1,v_1)(u_2).$$

First of all, since $dF(x,y)(u_1)$ is a convex set, we know that $T_{dF(x,y)(u_1)}(v_1)$ is equal to the closure of the set $\mathbb{R}_+(dF(x,y)(u_1) - v_1)$. Hence, it is enough to show that for arbitrary $\lambda > 0$, $v \in dF(x,y)(u_1)$ and $v_2 \in d^2F(x,y,u_1,v_1)(u_2)$, we have $\lambda(v - v_1) + v_2 \in d^2F(x,y,u_1,v_1)(u_2)$. Fix such λ, v, v_2 and a sequence $h_i \to 0+$. Let $(v_2^i)_{i\in\mathbb{N}}$ be as above and $v^i \to v$ when $i \to \infty$, be such that,

$$(14) \qquad y + h_i v^i \in F(x + h_i u_1) \subset F(x + h_i u_1 + h_i^2 u_2) + o(h_i)B \qquad \forall\, i \in \mathbb{N}.$$

Then, multiplying (11) by $1 - \lambda h_i$ and (14) by λh_i, taking their sum and using that for i large enough $\lambda h_i < 1$, we have,

$$y + h_i v_1 + h_i^2 \left(v_2^i + \lambda v^i - \lambda v_1\right) \in F(x + h_i u_1 + h_i^2 u_2) + o(h_i^2)B,$$

which allows to conclude that $\lambda(v - v_1) + v_2 \in d^2F(x,y,u_1,v_1)(u_2)$. Hence the proof of (13) is complete.

Step 3: The inclusion "$\subset$" in (9) follows directly from (8) and the convexity of $dF(x,y)(u_1)$. The converse inclusion "$\supset$" is due to the fact that $0 \in dF(x,y)(u_1) - v_1$. The proof is complete. $\square$

Lemma 2.12. *Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a set-valued map with convex images which is Lipschitz around $x \in \mathbb{R}^n$, $y \in F(x)$ and $v_1 \in dF(x,y)(u_1)$, $v_2 \in T_{F(x)}(y)$. Then,*

$$T^{\flat(2)}_{F(x)}(y, v_2) \subset T_{dF(x,y)(u_1)}(v_1 + v_2).$$

P r o o f. By (7) we have $v_1 + T_{F(x)}(y) \subset dF(x,y)(u_1)$. It follows (see for instance [1, Table 4.3]) that,

$$T_{dF(x,y)(u_1)}(v_1 + v_2) \supset T_{v_1 + T_{F(x)}(y)}(v_1 + v_2) = T_{T_{F(x)}(y)}(v_2).$$

Since by [7, Prop. 3.1] it is known that $T^{\flat(2)}_{F(x)}(y, v_2) \subset T_{T_{F(x)}(y)}(v_2)$, the statement follows. $\square$

The graph of a set-valued map $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ is the set $\mathrm{Gr}\,(F) := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in F(x)\}$. Let $F\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ be a set-valued map which is Lipschitz on a neighborhood of some $x \in \mathbb{R}^n$ and let $y \in F(x)$. The *circatangent derivative* $CF(x,y)$ is the set-valued map defined by,

$$CF(x,y)(u) := \operatorname*{Liminf}_{\substack{(x',y') \xrightarrow{\mathrm{Gr}(F)} (x,y) \\ h \to 0+}} \frac{F(x' + hu) - y'}{h} \quad \forall\, u \in \mathbb{R}^n.$$

It is not difficult to realize that $\mathrm{Gr}\,(CF(x,y))$ is equal to the Clarke tangent cone to $\mathrm{Gr}\,(F)$ at (x,y). Hence $CF(x,y)\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$ is a *closed convex* *process*, i.e. a set-valued map whose graph is a closed convex cone. For a closed convex process $A\colon \mathbb{R}^n \rightsquigarrow \mathbb{R}^m$, its *adjoint process* $A^*\colon \mathbb{R}^m \to \mathbb{R}^n$ is defined by,

$$A^*(p) := \{q \in \mathbb{R}^n \mid \langle q, u\rangle \le \langle p, v\rangle \quad \forall\, (u,v) \in \mathrm{Gr}\,(A)\}\,.$$

Below, for a set-valued map $[0,1] \times \mathbb{R}^n \ni (t,x) \rightsquigarrow F(t,x)$ and $t_0 \in [0,1]$ such that $F(t_0,\cdot)$ is Lipschitz on a neighborhood of some $x_0 \in \mathbb{R}^n$, we denote the partial derivatives with respect to the second variable by a subscript x. That is $d_xF(t_0,x_0,y_0)$ is equal to the adjacent derivative of $F(t_0,\cdot)$ at (x_0,y_0) for any $y_0 \in F(t_0,x_0)$. Similarly, $C_xF(t_0,x_0,y_0)$ denotes the circatangent derivative of $F(t_0,\cdot)$ at $(x_0,y_0) \in \mathrm{Gr}\,(F(t_0,\cdot))$.

3. Second-order variational inclusion. We consider the following differential inclusion with state constraints:

$$\text{(15)} \qquad \begin{cases} \dot{x}(t) \in \widetilde{F}(t,x(t)) & \text{a.e. in } [0,1] \\ x(0) \in K_0 \\ x(t) \in K \quad \forall\, t \in [0,1], \end{cases}$$

where the set-valued map $\widetilde{F}\colon [0,1] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and the closed nonempty sets $K_0, K \subset \mathbb{R}^n$ are given. Denote by $F\colon [0,1] \times \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ the set-valued map defined by $F(t,x) := \mathrm{co}\,\widetilde{F}(t,x)$ for all $(t,x) \in [0,1] \times \mathbb{R}^n$. Moreover we introduce the following sets:

$$\text{(16)} \qquad \begin{gathered} \mathcal{S}(x_0) := \left\{x \in W^{1,1}([0,1];\mathbb{R}^n) \;\middle|\; \dot{x}(t) \in \widetilde{F}(t,x(t)) \text{ a.e. and } x(0) = x_0\right\} \\ \mathcal{S}^{rel}(x_0) := \left\{x \in W^{1,1}([0,1];\mathbb{R}^n) \mid \dot{x}(t) \in F(t,x(t)) \text{ a.e. and } x(0) = x_0\right\} \\ \mathcal{K} := \{x \in \mathbb{C}([0,1];\mathbb{R}^n) \mid x(t) \in K \quad \forall\, t \in [0,1]\} \\ \mathcal{S}_K(x_0) := \mathcal{S}(x_0) \cap \mathcal{K} \qquad \mathcal{S}_K^{rel}(x_0) := \mathcal{S}^{rel}(x_0) \cap \mathcal{K}. \end{gathered}$$

The following regularity assumptions on the dynamics are imposed throughout this section:

$$\text{(A)} \qquad \begin{cases} \widetilde{F} \text{ has nonempty, compact images and is locally bounded at every} \\ \qquad\qquad (t,x) \in [0,1] \times \partial K; \\ \widetilde{F}(\cdot,x) \text{ is Lebesgue measurable for every } x \in \mathbb{R}^n; \\ \exists\, a_1 \in L^1([0,1];\mathbb{R}_+) \text{ such that } \sup\limits_{v \in \widetilde{F}(t,x)} |v| \le a_1(t)(1+|x|) \\ \qquad\qquad \forall\, (t,x) \in [0,1] \times \mathbb{R}^n; \\ \forall\, R > 0, \exists\, k_R \in L^1([0,1];\mathbb{R}_+) \text{ such that } \widetilde{F}(t,\cdot) \text{ is } k_R(t)\text{-Lipschitz on} \\ \qquad\qquad RB \text{ for a.e. } t \in [0,1]. \end{cases}$$

In addition, we need the following inward pointing condition:

$$\text{(IPC)}\quad \begin{cases} \forall\, t \in [0,1],\ \forall\, x \in \partial K,\ \exists\, \Omega_{t,x} \subset [0,1] \text{ of zero Lebesgue measure such that} \\ \forall\, v \in \operatorname*{Limsup}_{\substack{(s,y)\to(t,x)\\ s\notin\Omega_{t,x}}} \widetilde{F}(s,y) \text{ with } \max_{n\in N_K^1(x)} \langle n, v\rangle \geq 0, \\ \exists\, w \in \operatorname*{Liminf}_{\substack{(s,y)\to(t,x)\\ s\notin\Omega_{t,x}}} F(s,y) \text{ satisfying } \max_{n\in N_K^1(x)} \langle n, w-v\rangle < 0. \end{cases}$$

Remark 3.1. Assumption (IPC) implies that for every $x \in K$, the interior of $C_K(x)$ is nonempty. The above inward pointing condition is needed to handle the case of $\tilde{F}$ merely measurable with respect to time. If moreover $\tilde{F}$ is left absolutely continuous in time, then a simpler condition proposed in [3] can be used instead of (IPC).

Consider a reference trajectory $\bar{x} \in \mathcal{S}_K(\bar{x}^0)$ where $\bar{x}^0 \in K_0$. The corresponding set of admissible first-order variations $\mathcal{V}^{(1)}(\bar{x})$ is the set of absolutely continuous maps $y \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying,

(i) $\dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t)) \quad$ for a.e. $t \in [0,1]$;

(ii) $y(0) \in T^\flat_{K_0}(\bar{x}^0)$;

(iii) $y \in T^\flat_{\mathcal{K}}(\bar{x})$;

(iv) $\exists\, a_2 \in L^1([0,1];\mathbb{R}_+)$, $\exists\, h_0 > 0$ such that for all $h \in [0,h_0]$ and a.e. $t \in [0,1]$,

$$\operatorname{dist}_{F(t,\bar{x}(t)+hy(t))}(\dot{\bar{x}}(t) + h\dot{y}(t)) \leq a_2(t)h^2.$$

For a given $y \in \mathcal{V}^{(1)}(\bar{x})$, we abbreviate $(t, \bar{x}(t), \dot{\bar{x}}(t), y(t), \dot{y}(t))$ by $[t]$ and define the set of admissible second-order variations $\mathcal{V}^{(2)}(\bar{x}, y)$ as the set of absolutely continuous maps $w \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying,

(i) $\dot{w}(t) \in d_x^2 F[t](w(t)) \quad$ for a.e. $t \in [0,1]$;

(ii) $w(0) \in T^{\flat(2)}_{K_0}(\bar{x}(0), y(0))$;

(iii) $w \in T^{\flat(2)}_{\mathcal{K}}(\bar{x}, y)$.

Remark 3.2. Let K be as in Example 2.3. Then it is known that $y \in T^\flat_{\mathcal{K}}(\bar{x})$ if and only if for all $t \in [0,1]$, $\langle \nabla b_i(\bar{x}(t)), y(t)\rangle \leq 0$ for all $i \in I(\bar{x}(t))$

(the set of active indices at $\bar{x}(t)$), see for instance [13, Remark 4.2]. The situation is more complicated for $T_{\mathcal{K}}^{\flat(2)}(\bar{x}, y)$ where the condition

$$\langle \nabla b_i(\bar{x}(t)), w(t)\rangle + \frac{1}{2} b_i''(\bar{x}(t))y(t)y(t) \leq 0 \qquad \forall\, i \in I^{(1)}(\bar{x}(t), y(t)) \quad \forall\, t \in [0,1], \tag{17}$$

in general, does not imply that $w \in T_{\mathcal{K}}^{\flat(2)}(\bar{x}, y)$. A pointwise characterization of $T_{\mathcal{K}}^{\flat(2)}(\bar{x}, y)$ is given in [19] for the case when $\mathcal{K}$ is the cone of nonnegative continuous functions. Further investigations of this subject can be found in [7, 20]. It turns out that a convenient pointwise condition is an appropriate strengthening of inequality (17), see for instance [18, (TV)].

We are ready to state the main result of this section:

Theorem 3.3. *Assume* (A) *and* (IPC). *Let* $\bar{x} \in \mathcal{S}_K(\bar{x}^0)$ *for some* $\bar{x}^0 \in K_0$, $y \in \mathcal{V}^{(1)}(\bar{x})$ *and* $w \in \mathcal{V}^{(2)}(\bar{x}, y)$. *Consider any sequences* $h_i \to 0+$, $w_i^0 \to w(0)$ *such that* $\bar{x}(0) + h_i y(0) + h_i^2 w_i^0 \in K_0$. *Then there exist*

$$x_i \in \mathcal{S}_K(\bar{x}(0) + h_i y(0) + h_i^2 w_i^0)$$

such that $\frac{1}{h_i^2}(x_i - \bar{x} - h_i y)$ *converge uniformly to* w *when* $i \to \infty$.

Proof. Let $\bar{x}, y, w$ be as above and fix any sequences $h_i \to 0+$, $w_i^0 \to w(0)$ such that $\bar{x}(0) + h_i y(0) + h_i^2 w_i^0 \in K_0$. To simplify the notations we set $y^0 := y(0)$ and $w^0 := w(0)$. Define for all $i \in \mathbb{N}$,

$$x_i^1(t) := \bar{x}(t) + h_i y(t) + h_i^2 w(t) + h_i^2(w_i^0 - w^0).$$

Note that with this definition $x_i^1(0) = \bar{x}^0 + h_i y^0 + h_i^2 w_i^0 \in K_0$ and $\dot{x}_i^1(t) = \dot{\bar{x}}(t) + h_i \dot{y}(t) + h_i^2 \dot{w}(t)$.

First, we show that there exist $x_i^2 \in \mathcal{S}^{rel}(\bar{x}^0 + h_i y^0 + h_i^2 w_i^0)$ such that,

$$\frac{1}{h_i^2} \left\| x_i^1 - x_i^2 \right\|_{W^{1,1}} \to 0, \qquad \text{when } i \to \infty. \tag{18}$$

For this aim define,

$$\gamma_i(t) := \mathrm{dist}_{F(t, x_i^1(t))}(\dot{x}_i^1(t)).$$

Let $R > 0$ be such that for all $t \in [0,1]$ and $i \geq 1$,

$$|\bar{x}(t) + h_i y(t)| + |\bar{x}(t) + h_i y(t) + h_i^2 w(t)| + |x_i^1(t)| \leq R.$$

Then it follows from the $k_R(t)$-Lipschitz continuity of $F(t,\cdot)$ that for all i large enough,

(19)

$$\begin{aligned}\gamma_i(t) &\leq \text{dist}_{F(t,\bar{x}(t)+h_iy(t)+h_i^2w(t))}(\dot{\bar{x}}(t)+h_i\dot{y}(t)+h_i^2\dot{w}(t))+k_R(t)h_i^2\left|w_i^0-w^0\right| \\ &\leq \text{dist}_{F(t,\bar{x}(t)+h_iy(t))}(\dot{\bar{x}}(t)+h_i\dot{y}(t))+h_i^2\left(k_R(t)\left(\|w\|_\infty+\left|w_i^0-w^0\right|\right)+|\dot{w}(t)|\right) \\ &\leq h_i^2\left(k_R(t)\left(\|w\|_\infty+\left|w_i^0-w^0\right|\right)+|\dot{w}(t)|+a_2(t)\right).\end{aligned}$$

By (19), the sequence $\left(\dfrac{\gamma_i(t)}{h_i^2}\right)_{i\in\mathbb{N}}$ is integrably bounded. Since $w(\cdot)\in\mathcal{V}^{(2)}(\bar{x},y)$,

$$\frac{1}{h_i^2}\text{dist}_{F(t,\bar{x}(t)+h_iy(t)+h_i^2w(t))}(\dot{\bar{x}}(t)+h_i\dot{y}(t)+h_i^2\dot{w}(t))\to 0 \qquad \text{for a.e. } t\in[0,1].$$

Hence

$$\lim_{i\to\infty}\frac{1}{h_i^2}\int_0^1\gamma_i(t)dt=\int_0^1\left(\lim_{i\to\infty}\frac{\gamma_i(t)}{h_i^2}\right)dt=0.$$

This and Filippov's theorem, see for instance [1, Thm. 10.4.1], imply the existence of $x_i^2\in\mathcal{S}^{rel}(\bar{x}^0+h_iy^0+h_i^2w_i^0)$ satisfying (18).

Next, [13, Thm. 3.3] implies that for a constant $L>0$ and for every $i\in\mathbb{N}$, there exists $x_i^3\in\mathcal{S}_K^{rel}(\bar{x}^0+h_iy^0+h_i^2w_i^0)$ such that,

$$\left\|x_i^2-x_i^3\right\|_{W^{1,1}}\leq L\max_{t\in[0,1]}\text{dist}_K(x_i^2(t)).$$

Note that for all i,

(20) $\text{dist}_K(x_i^2(t))\leq\text{dist}_K(\bar{x}(t)+h_iy(t)+h_i^2w(t))+h_i^2\left|w_i^0-w^0\right|+\left|x_i^2(t)-x_i^1(t)\right|,$

and for every $x\in\mathbb{C}([0,1];\mathbb{R}^n)$ and all $t\in[0,1]$,

$$\text{dist}_K(x(t))=\inf_{k\in K}|x(t)-k|\leq\inf_{\kappa\in\mathcal{K}}|x(t)-\kappa(t)|\leq\text{dist}_{\mathcal{K}}(x).$$

Thus, $\max_{t\in[0,1]}\text{dist}_K(\bar{x}(t)+h_iy(t)+h_i^2w(t))\leq\text{dist}_{\mathcal{K}}(\bar{x}+h_iy+h_i^2w)$. Since $w\in\mathcal{V}^{(2)}(\bar{x},y)$, this implies by (18) and (20) that,

(21) $$\frac{1}{h_i^2}\left\|x_i^2-x_i^3\right\|_{W^{1,1}}\to 0, \qquad \text{when } i\to+\infty.$$

Furthermore, by [13, Cor. 3.4], for every $i\in\mathbb{N}$, there exists $x_i\in\mathcal{S}_K(\bar{x}^0+h_iy^0+h_i^2w_i^0)$ such that,

(22) $$\left\|x_i^3-x_i\right\|_\infty\leq h_i^3.$$

Finally, by the very definition of x_i^1, we find that,

$$\left\|\bar{x} + h_i y + h_i^2 w - x_i\right\|_\infty \le h_i^2 \left|w_i^0 - w^0\right| + \left\|x_i^1 - x_i^2\right\|_{W^{1,1}} + \left\|x_i^2 - x_i^3\right\|_{W^{1,1}} + \left\|x_i^3 - x_i\right\|_\infty .$$

Thus, from (18), (21) and (22) it follows that,

$$\lim_{i\to\infty} \frac{\left\|\bar{x} + h_i y + h_i^2 w - x_i\right\|_\infty}{h_i^2} = 0. \qquad \square$$

4. Second-order optimality conditions: differential inclusions. We study here necessary optimality conditions for a Mayer optimization problem involving the differential inclusion (15). Throughout this section we associate with any $\bar{x} \in \mathcal{S}_K(x_0)$ the linearized differential inclusion:

$$\begin{cases} \dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t)) \quad a.e. \\ y(0) \in T^\flat_{K_0}(\bar{x}(0)). \end{cases} \tag{23}$$

Our first goal is to define subsets of the sets of admissible variations that are convenient for the expression of necessary optimality conditions. For every $t \in [0,1]$ such that $\dot{\bar{x}}(t) \in F(t, \bar{x}(t))$ define the closed convex process

$$\mathcal{A}(t) := C_x F(t, \bar{x}(t), \dot{\bar{x}}(t))$$

and the closed convex cone

$$\mathcal{L}(t) := T_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)).$$

Moreover, for every trajectory $y(\cdot)$ of (23) and $t \in [0,1]$ such that $\dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t))$ we write

$$\mathcal{E}(y;t) := T_{d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))}(\dot{y}(t)).$$

Since $d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t)) = d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t)) + \mathcal{L}(t)$ we have

$$\mathcal{E}(y;t) + \mathcal{L}(t) = \mathcal{E}(y;t). \tag{24}$$

For a fixed solution y of (23), we introduce the set-valued map $\mathcal{F}_I(y;\cdot)\colon [0,1] \rightsquigarrow \mathbb{R}^n$ given by

$$\mathcal{F}_I(y;t) := \operatorname*{Liminf}_{h\to 0+} \frac{F(t, \bar{x}(t) + hy(t)) - h\dot{y}(t) - \dot{\bar{x}}(t)}{h^2} \quad \forall\, t \in [0,1],$$

and define the second-order approximation of the differential inclusion from (15) as follows:

$$\begin{cases} \dot{w}(t) \in \mathcal{A}(t)w(t) + \mathcal{F}_I(y;t) + \mathcal{E}(y;t) \quad \text{a.e.} \\ w(0) \in T^{\flat(2)}_{K_0}(\bar{x}(0), y(0)). \end{cases} \tag{25}$$

This definition is motivated by (8) and the following fact:

Proposition 4.1. *Assume* (A) *and let* $\bar{x} \in \mathcal{S}(x_0)$, $y \in W^{1,1}([0,1];\mathbb{R}^n)$ *be a solution of* (23). *Then for a.e.* $t \in [0,1]$,

$$\mathcal{A}(t)w + \mathcal{F}_I(y;t) \subset d_x^2 F[t](w) \quad \forall\, w \in \mathbb{R}^n.$$

P r o o f. Let $R := \|x\|_\infty + 1$, $k_R \in L^1([0,1];\mathbb{R}_+)$ be as in (A) and $t \in [0,1]$ be such that $F(t,\cdot)$ is $k_R(t)$-Lipschitz on RB, $\dot{\bar{x}}(t) \in F(t,\bar{x}(t))$, $\dot{y}(t) \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))$. If $\mathcal{A}(t)w$ or $\mathcal{F}_I(y;t)$ is empty, then there is nothing to prove.

Fix $w \in \mathbb{R}^n$, $v \in \mathcal{F}_I(y;t)$, $\alpha \in \mathcal{A}(t)w$ and a sequence $h_i \to 0+$. By the very definition of $\mathcal{F}_I(y;t)$, there exists a sequence $v_i \to v$ such that,

$$\dot{\bar{x}}(t) + h_i\dot{y}(t) + h_i^2 v_i \in F(t, \bar{x}(t) + h_i y(t)) \quad \forall\, i \in \mathbb{N}.$$

On the other hand, since $\mathrm{Gr}\,(\mathcal{A}(t))$ is equal to the Clarke tangent cone to $\mathrm{Gr}\,(F(t,\cdot))$ at $(\bar{x}(t),\dot{\bar{x}}(t))$, there exist sequences $(w_i,\alpha_i) \to (w,\alpha)$ satisfying,

$$\begin{aligned} \mathrm{Gr}\,(F(t,\cdot)) &\ni \left(\bar{x}(t) + h_i y(t), \dot{\bar{x}}(t) + h_i\dot{y}(t) + h_i^2 v_i\right) + h_i^2(w_i,\alpha_i) \\ &= \left(\bar{x} + h_i y(t) + h_i^2 w_i, \dot{\bar{x}}(t) + h_i\dot{y}(t) + h_i^2(v_i + \alpha_i)\right), \end{aligned}$$

which implies, by the Lipschitz continuity of $F(t,\cdot)$, that $v + \alpha \in d_x^2 F[t](w)$. The proof is complete. □

Thanks to Propositions 2.11 and 4.1, for any $y \in \mathcal{V}^{(1)}(\bar{x})$ we can define the following subset of second-order admissible variations:

$$\mathcal{V}_I^{(2)}(\bar{x},y) := \left\{ w \in \mathcal{V}^{(2)}(\bar{x},y) \;\middle|\; w \text{ solves } (25) \right\}.$$

Consider now the Mayer problem

$$\text{Minimize } \{\varphi(x(1)) \mid x(\cdot) \in \mathcal{S}_K(x_0),\ x_0 \in K_0\}, \tag{26}$$

where $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is a given twice differentiable function.

A trajectory $\bar{x}(\cdot) \in \mathcal{S}_K(K_0)$ is called a *strong local minimizer* of the above Mayer problem if there exists $\varepsilon > 0$ such that for every $x(\cdot) \in \mathcal{S}_K(K_0)$ satisfying $\|x - \bar{x}\|_\infty < \varepsilon$ we have $\varphi(\bar{x}(1)) \leq \varphi(x(1))$.

Let us start by recalling the first-order necessary optimality conditions of [6, Corollary 3.8] in the form of a maximum principle. We need the following additional assumption:

$$(\widetilde{A}) \qquad \exists\, \ell \geq 0 \text{ such that } \mathcal{A}(t) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n \text{ is } \ell - \text{Lipschitz for a.e. } t \in [0,1].$$

Remark 4.2. Assumption $(\widetilde{A})$ concerns a reference trajectory $\bar{x}$. It was imposed in [6] to apply the duality theory of convex analysis to closed convex processes. In the next section, when dealing with control systems, we do not need such assumption, because in this case a more direct approach not involving abstract duality theorems is used.

Recall that the domain of $\mathcal{A}(t)$ is equal to $\mathbb{R}^n$ if and ony if $\mathcal{A}(t)$ is Lipschitz with a constant $c(t) \geq 0$. Assumption $(\widetilde{A})$ requires $c(\cdot)$ to be essentially bounded.

If for almost every $t \in [0,1]$, the set $\mathrm{Gr}(F(t,\cdot))$ is sleek and $k_R(\cdot)$ in (A) is essentially bounded for every $R > 0$, then $(\widetilde{A})$ is satisfied for any reference trajectory $\bar{x}$.

Theorem 4.3 (Maximum Principle)**.** *Let $\bar{x}$ be a strong local minimizer of Problem* (26) *and for all $t \in [0,1]$,* $\mathrm{int}\, C_K(\bar{x}(t)) \neq \emptyset$. *Assume $(\widetilde{A})$ and that* (A) *holds true with a bounded $a_1(\cdot)$. If* $\mathrm{int}\, C_{K_0}(\bar{x}(0)) \cap \mathrm{int}\, C_K(\bar{x}(0)) \neq \emptyset$*, then there exist $\lambda \in \{0,1\}$, $\psi \in NBV([0,1];\mathbb{R}^n)$ and $p \in W^{1,1}([0,1];\mathbb{R}^n)$ such that $\lambda + \|\psi\|_{TV} \neq 0$ and*

(*i*) $\dot{p}(t) \in \mathcal{A}(t)^*(-p(t) - \psi(t))$ *a.e. in* $[0,1]$

(*ii*) $p(0) \in \lambda\left(N_{K_0}(\bar{x}(0)) + N_K(\bar{x}(0))\right)$

(*iii*) $p(1) = -\lambda \nabla\varphi(\bar{x}(1)) - \psi(1)$

(*iv*) $\langle p(t) + \psi(t), \dot{\bar{x}}(t)\rangle = \max_{z \in F(t,\bar{x}(t))} \langle p(t) + \psi(t), z\rangle$ *a.e. in* $[0,1]$.

Furthermore,

$$\psi(0+) \in N_K(\bar{x}(0)),\ \psi(t) - \psi(t-) \in N_K(\bar{x}(t)),\ \psi(t) = \int_{[0,t]} \nu(s)d\mu(s)\ \forall\, t \in\,]0,1],$$

for a non-negative (scalar) Borel measure μ on $[0,1]$ and a Borel measurable mapping $\nu\colon [0,1] \to \mathbb{R}^n$ satisfying

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

Moreover, the following non degeneracy conditions hold true

$$\lambda + \sup_{t\in]0,1[} |p(t)+\psi(t)| \neq 0 \quad \textit{and} \quad \lambda + var(\psi,]0,1]) \neq 0,$$

where $var(\psi,]0,1])$ *denotes the total variation of* ψ *on* $]0,1]$. *Finally, if there exists a solution to the constrained differential inclusion*

$$\begin{cases} \dot{y}(t) \in \overline{\mathcal{A}(t)y(t)+\mathcal{L}(t)} & a.e. \\ y(t) \in \operatorname{int} C_K(\bar{x}(t)) & \forall\, t \in [0,1] \\ y(0) \in \operatorname{int} C_{K_0}(\bar{x}(0)), \end{cases} \tag{27}$$

then the above holds true with $\lambda = 1$.

An *extremal* is a tuple $(\bar{x}, p, \psi, \mu, \nu, \lambda)$, where $\bar{x}$ is a feasible trajectory of (15) and p, ψ, μ, ν, λ are as in the maximum principle (Theorem 4.3). An extremal is *normal* if $\lambda = 1$.

Note that (iv) of Theorem 4.3 is equivalent to $\langle p(t)+\psi(t), z - \dot{\bar{x}}(t)\rangle \leq 0$ for all $z \in F(t, \bar{x}(t))$ and a.e. $t \in [0,1]$. In other words,

$$p(t)+\psi(t) \in N_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t)) \qquad \text{for a.e. } t \in [0,1]. \tag{28}$$

Corollary 4.4. *Under all the assumptions of Theorem* 4.3 *consider* λ, p, ψ *as in its conclusions. Then for almost every* $t \in [0,1]$ *and all* $v \in \mathcal{L}(t)$ *satisfying* $\langle p(t)+\psi(t), v\rangle = 0$ *we have*

$$\langle p(t)+\psi(t), w\rangle \leq 0 \qquad \forall\, w \in T^{\flat(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t), v). \tag{29}$$

Proof. Indeed, let $t \in [0,1]$ be such that (iv) of Theorem 4.3 holds true and $w \in T^{\flat(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t), v)$ for some $v \in \mathcal{L}(t)$ satisfying $\langle p(t)+\psi(t), v\rangle = 0$. By the very definition of the second-order adjacent set, for every $h > 0$ there exists w_h such that $w_h \to w$ when $h \to 0+$ and $\dot{\bar{x}}(t) + hv + h^2 w_h \in F(t, \bar{x}(t))$ for all h. Since $\langle p(t)+\psi(t), v\rangle = 0$ it follows from (28) that $h^2 \langle p(t)+\psi(t), w_h\rangle \leq 0$. Dividing by h^2 and passing to the limit, we get (29). □

Consider the following linearization of the differential inclusion in (15):

$$\begin{cases} \dot{y}(t) \in \mathcal{A}(t)y(t) + F(t,\bar{x}(t)) - \dot{\bar{x}}(t) & a.e. \\ y(0) \in T^{\flat}_{K_0}(\bar{x}(0)). \end{cases} \tag{30}$$

Remark 4.5. It is clear that $\mathrm{Gr}\,(\mathcal{A}(t)) \subset \mathrm{Gr}\,(d_x F(t, \bar{x}(t), \dot{\bar{x}}(t)))$. Hence, by (7), for any $y(\cdot)$ solving (30) we have $\dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t))$ a.e.

Remark 4.6. Let $(\bar{x}, p, \psi, \mu, \nu)$ be a normal extremal. Consider $y \in \mathcal{V}^{(1)}(\bar{x})$ satisfying (30) with $y(0) \in C_{K_0}(\bar{x}(0))$, $y(t) \in C_K(\bar{x}(t))$ for all $t \in [0,1]$ and integrable selections $\gamma(t) \in \mathcal{A}(t)y(t)$, $v(t) \in F(t, \bar{x}(t))$ such that $\dot{y}(t) = \gamma(t) + v(t) - \dot{\bar{x}}(t)$ a.e. Then we have,

$$\begin{aligned}
\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle &= \langle -p(1) - \psi(1), y(1)\rangle \\
&= \langle -\psi(1), y(1)\rangle + \langle -p(0), y(0)\rangle \\
&\qquad + \int_0^1 \big(\langle -p(t), \dot{y}(t)\rangle + \langle -\dot{p}(t), y(t)\rangle \big) dt \\
&\geq \langle -\psi(1), y(1)\rangle + \int_{[0,1]} \langle \nu(t), y(t)\rangle \, d\mu(t) \\
&\qquad + \int_0^1 \big(\langle -p(t), \dot{y}(t)\rangle + \langle -\dot{p}(t), y(t)\rangle \big) dt \\
&= \int_0^1 \big(\langle -p(t) - \psi(t), \dot{y}(t)\rangle + \langle -\dot{p}(t), y(t)\rangle \big) dt \\
&\geq \int_0^1 \langle -p(t) - \psi(t), \dot{y}(t)\rangle \, dt + \int_0^1 \langle p(t) + \psi(t), \gamma(t)\rangle \, dt \\
&= \int_0^1 \langle -p(t) - \psi(t), v(t) - \dot{\bar{x}}(t)\rangle \, dt \geq 0.
\end{aligned}$$

Therefore if $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, then (iv) of Theorem 4.3 yields $\langle p(t) + \psi(t), v(t) - \dot{\bar{x}}(t)\rangle = 0$ a.e. This and Corollary 4.4 imply that for almost every $t \in [0,1]$,

$$\langle p(t) + \psi(t), w\rangle \leq 0 \qquad \forall\, w \in T^{\flat(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t), v(t) - \dot{\bar{x}}(t)). \tag{31}$$

Consider the sets

$$\Lambda(t) := \{ v \in F(t, \bar{x}(t)) \mid \langle p(t) + \psi(t), \dot{\bar{x}}(t)\rangle = \langle p(t) + \psi(t), v\rangle \},$$

and the constrained differential inclusion

$$\begin{cases} \dot{y}(t) \in \mathcal{A}(t)y(t) + v(t) - \dot{\bar{x}}(t), \;\; v(t) \in F(t, \bar{x}(t)) \quad \text{a.e.} \\ y(0) \in C_{K_0}(\bar{x}(0)) \\ y(t) \in C_K(\bar{x}(t)) \quad \forall\, t \in [0,1]. \end{cases} \tag{32}$$

Therefore we have derived the following alternative: If $F(t,\bar{x}(t)) \neq \{\dot{\bar{x}}(t)\}$ on a set of positive measure, then either for every $y(\cdot) \in \mathcal{V}^{(1)}(\bar{x})$ satisfying (32) with $v(\cdot) \neq \dot{\bar{x}}(\cdot)$ on a set of positive measure we have $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle > 0$, or $\Lambda(t) \neq \{\dot{\bar{x}}(t)\}$ on a set of positive measure. In this later case, such normal extremals can then be considered as singular.

Using the second-order variational equation from Section 3, we find the following necessary optimality conditions for problem (26):

Theorem 4.7. *Assume* (A), (IPC) *and let $\bar{x}$ be a strong local minimizer of problem* (26). *Then,*

$$\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle \geq 0 \qquad \forall\, y \in \mathcal{V}^{(1)}(\bar{x}). \tag{33}$$

Moreover, for all $y \in \mathcal{V}^{(1)}(\bar{x})$ such that $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, we have

$$\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle + \frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) \geq 0 \qquad \forall\, w \in \mathcal{V}^{(2)}(\bar{x}, y). \tag{34}$$

P r o o f. To show (33), let $y \in \mathcal{V}^{(1)}(\bar{x})$ and fix a sequence $h_i \to 0+$. Since $y(0) \in T^\flat_{K_0}(\bar{x}(0))$, there exists a sequence $y_i^0 \to y(0)$ such that $\bar{x}(0) + h_i y_i^0 \in K_0$ for all i. Thus, by [13, Thm. 4.3], we can find a sequence $(y_i)_{i\in\mathbb{N}}$ converging uniformly to y, such that $\bar{x} + h_i y_i \in \mathcal{S}_K(\bar{x}(0) + h_i y_i^0)$ for all i. Since $\bar{x}$ is a strong local minimizer, for all large i,

$$0 \leq \varphi(\bar{x}(1) + h_i y_i(1)) - \varphi(\bar{x}(1)) = h_i \langle \nabla\varphi(\bar{x}(1)), y(1)\rangle + o(h_i),$$

where $o(h_i)/h_i \to 0+$ when $i \to \infty$. Dividing by h_i and passing to the limit, we find (33).

Next, let $y \in \mathcal{V}^{(1)}(\bar{x})$ be such that $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, $w \in \mathcal{V}^{(2)}(\bar{x}, y)$ and $h_i \to 0+$. By Theorem 3.3 there exists a sequence $w_i \to w$ uniformly, when $i \to \infty$ such that $\bar{x} + h_i y + h_i^2 w_i \in \mathcal{S}_K(K_0)$ for all i. Since $\bar{x}$ is a strong local minimizer, it follows that for all $i \in \mathbb{N}$ large enough,

$$\begin{aligned} 0 \leq \varphi(\bar{x}(1) + h_i y(1) + h_i^2 w_i(1)) &- \varphi(\bar{x}(1)) \\ &= h_i^2 \langle \nabla\varphi(\bar{x}(1)), w(1)\rangle + \frac{h_i^2}{2}\varphi''(\bar{x}(1))y(1)y(1) + o(h_i^2), \end{aligned}$$

where $o(h_i^2)/h_i^2 \to 0$, when $i \to \infty$. Dividing by h_i^2 and passing to the limit, we get (34). $\square$

Consider the following "second-order linearization" of (15):

$$(35)\qquad \begin{cases} \dot{w}(t) \in \mathcal{A}(t)w(t) + \mathcal{E}(y;t) & \text{for a.e. } t \in [0,1] \\ w(0) \in C_{K_0}(\bar{x}(0)) \\ w(t) \in C_K(\bar{x}(t)) & \text{for all } t \in [0,1]. \end{cases}$$

Theorem 4.8 (Second Order Maximum Principle)**.** *Let $\bar{x}$ be a strong local minimizer of Problem (26) and* ($\widetilde{A}$), (IPC), (A) *hold true with a bounded $a_1(\cdot)$. Assume that $y \in \mathcal{V}^{(1)}(\bar{x})$ is such that $\mathcal{V}_I^{(2)}(\bar{x}, y) \neq \emptyset$ and $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$. Then for every solution w of* (35) *we have $\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle \geq 0$.*

Furthermore, if $\operatorname{int} C_{K_0}(\bar{x}(0)) \cap \operatorname{int} C_K(\bar{x}(0)) \neq \emptyset$*, then there exist $\lambda \in \{0,1\}$, $\psi \in NBV([0,1];\mathbb{R}^n)$ and $p \in W^{1,1}([0,1];\mathbb{R}^n)$ satisfying all the conclusions of Theorem* 4.3 *such that in addition*

$$\max_{v \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))} \langle p(t) + \psi(t), v\rangle = \langle p(t) + \psi(t), \dot{y}(t)\rangle \ \text{ a.e. in } [0,1].$$

Finally, if the constrained differential inclusion

$$(36)\qquad \begin{cases} \dot{z}(t) \in \overline{\mathcal{A}(t)z(t) + \mathcal{E}(y;t)} & a.e. \\ z(t) \in \operatorname{int} C_K(\bar{x}(t)) & \forall\, t \in [0,1] \\ z(0) \in \operatorname{int} C_{K_0}(\bar{x}(0)) \end{cases}$$

has a solution, then the above holds true with $\lambda = 1$.

Proof. Let $\bar{w} \in \mathcal{V}_I^{(2)}(\bar{x}, y)$ and w be a solution of (35). Since $\mathcal{A}(t)$ is a closed convex process and $\mathcal{E}(y;t)$ is a closed convex cone for a.e. $t \in [0,1]$, it follows from Lemmas 2.4, 2.5 and Propositions 2.11, 4.1 that $\bar{w} + w \in \mathcal{V}_I^{(2)}(\bar{x}, y)$. Therefore, by Theorem 4.7,

$$(37)\qquad \langle \nabla\varphi(\bar{x}(1)), w(1)\rangle + \langle \nabla\varphi(\bar{x}(1)), \bar{w}(1)\rangle + \frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) \geq 0.$$

The first conclusion follows from the fact that the set of solutions to (35) is a cone and the last two terms of the above inequality do not depend on w. Therefore $w \equiv 0$ is an optimal solution of the problem

$$\text{Minimize } \ \langle \nabla\varphi(\bar{x}(1)), w(1)\rangle\,,$$

over solutions w of (35). For all $t \in [0,1]$ consider the closed convex process $\mathcal{A}_1(t) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ defined by

$$\mathcal{A}_1(t)x = \overline{\mathcal{A}(t)x + \mathcal{E}(y;t)} \quad \forall\, x \in \mathbb{R}^n.$$

The same arguments as in [6, pp. 687, 693-697] applied to the closed convex process $\mathcal{A}_1(t)$ instead of $B(t,\cdot)$ introduced in [6, p. 693] and (24) lead to the desired result. □

5. Second-order optimality conditions: control systems. We investigate here the following constrained control system:

$$(38)\qquad \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U(t) \qquad \text{for a.e. } t \in [0,1], \\ x(0) \in K_0 \\ x(t) \in K \qquad \forall\, t \in [0,1], \end{cases}$$

where $f\colon [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $K_0, K \subset \mathbb{R}^n$ are nonempty closed sets and $U : [0,1] \rightsquigarrow \mathbb{R}^m$ is a set-valued map. Throughout this section $(\bar{x}, \bar{u})$ denotes a trajectory control pair of the control system (38). In order to simplify the notations, we will abbreviate $(t, \bar{x}(t), \bar{u}(t))$ by $[t]$, so for instance $f[t] = f(t, \bar{x}(t), \bar{u}(t))$. We assume that

(H1) (a) $\forall\, x \in \mathbb{R}^n$, $\forall\, u \in \mathbb{R}^m$, $f(\cdot, x, u)$ is measurable and for almost all $t \in [0,1]$, $f(t,\cdot,\cdot)$ is continuous, and $f(t, x, U(t))$ is closed for all t, x. Moreover, $(t, x) \rightsquigarrow f(t, x, U(t))$ is locally bounded on $[0,1] \times \partial K$;

(b) $\forall\, R > 0$, there exists $k_R \in L^1([0,1];\mathbb{R}_+)$ such that for a.e. $t \in [0,1]$, $f(t,\cdot,u)$ is $k_R(t)$-Lipschitz on RB for all $u \in U(t)$;

(c) There exists $a_1 \in L^1([0,1];\mathbb{R}_+)$ such that for a.e. $t \in [0,1]$ and all $x \in \mathbb{R}^n$,

$$\sup_{u \in U(t)} |f(t, x, u)| \le a_1(t)(1 + |x|);$$

(d) The set-valued map $U : [0,1] \rightsquigarrow \mathbb{R}^m$ is measurable with closed, nonempty images.

Given a trajectory control pair $(\bar{x}, \bar{u})$, in some results below we assume that

$$(39)\quad f_x(t,\cdot,\cdot) \text{ exists and is continuous on a neighborhood of } (\bar{x}(t), \bar{u}(t)) \text{ for almost all } t \in [0,1].$$

One readily checks that (H1) implies that the set-valued map $(t, x) \rightsquigarrow f(t, x, U(t))$ satisfies (A). Thus (38) is a special case of (15). The set-valued map $x_0 \rightsquigarrow \mathcal{S}_K(x_0)$, defined in (16), is in this section understood with respect to $\tilde{F}(t, x) := f(t, x, U(t))$. As before, consider the set-valued map

$$(t, x) \rightsquigarrow F(t, x) := \operatorname{co}\{f(t, x, u) \mid u \in U(t)\}.$$

Let us now fix a solution y of (23) and introduce the sets $\mathcal{F}_C(y;t) \subset \mathbb{R}^n$ defined for all $t \in [0,1]$ such that $\dot{y}(t)$ does exist by,

$$\begin{aligned}\mathcal{F}_C(y;t) :=& \{v \in \mathbb{R}^n \mid \forall\, h > 0,\ \exists\, u_h \in U(t),\ v_h \in \mathbb{R}^n,\ \ (u_h, v_h) \to (\bar{u}(t), v)\\ &\text{when } h \to 0+ \ \text{ such that } f(t, \bar{x}(t) + hy(t), u_h) = f[t] + h\dot{y}(t) + h^2 v_h\},\end{aligned}$$

and consider the following second-order approximation of the control system from (38):

$$\text{(40)} \qquad \begin{cases} \dot{w}(t) \in f_x[t]w(t) + \mathcal{F}_C(y;t) + \mathcal{E}(y;t) & \text{for a.e. } t \in [0,1] \\ w(0) \in T^{\flat(2)}_{K_0}(\bar{x}(0), y(0)). \end{cases}$$

This definition is motivated by the following observation:

Proposition 5.1. *Let $(\bar{x}, \bar{u})$ be a trajectory control pair of the control system* (38) *and y be a solution of* (23)*. If* (39) *holds true, then for almost all $t \in [0,1]$ and all $w \in \mathbb{R}^n$,*

$$f_x[t]w + \mathcal{F}_C(y;t) + \mathcal{E}(y;t) \subset d_x^2 F[t](w).$$

Proof. By Proposition 2.11 it is enough to show that for a.e. $t \in [0,1]$ and for all $w \in \mathbb{R}^n$ we have

$$f_x[t]w + \mathcal{F}_C(y;t) \subset d_x^2 F[t](w).$$

Fix $t \in [0,1]$ such that $\dot{\bar{x}}(t) = f(t, \bar{x}(t), \bar{u}(t))$ and $\dot{y}(t) \in d_x F(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t))$. Let $w \in \mathbb{R}^n$, $v \in \mathcal{F}_C(y;t)$ and $h_i \to 0+$. By the very definition of $\mathcal{F}_C(y;t)$, there exist $u_i \xrightarrow{U(t)} \bar{u}(t)$ and $v_i \to v$ such that,

$$f(t, \bar{x}(t) + h_i y(t), u_i) = f[t] + h_i \dot{y}(t) + h_i^2 v_i \quad \forall\, i \in \mathbb{N}.$$

Moreover, using the continuity of $f_x(t,\cdot,\cdot)$ on a neighborhood of $(\bar{x}(t), \bar{u}(t))$, for all sufficiently large i,

$$\begin{aligned} F(t, \bar{x}(t) + h_i y(t) + h_i^2 w) &\ni f(t, \bar{x}(t) + h_i y(t) + h_i^2 w, u_i)\\ &= f(t, \bar{x}(t) + h_i y(t), u_i)\\ &\qquad + \int_0^1 f_x(t, \bar{x}(t) + h_i y(t) + \theta h_i^2 w, u_i) h_i^2 w d\theta\\ &= f(t, \bar{x}(t) + h_i y(t), u_i) + f_x[t] h_i^2 w\\ &\qquad + \int_0^1 \left(f_x(t, \bar{x}(t) + h_i y(t) + \theta h_i^2 w, u_i) - f_x[t]\right) h_i^2 w d\theta\\ &= f[t] + h_i \dot{y}(t) + h_i^2 \left(v_i + f_x[t]w\right) + o(h_i^2), \end{aligned}$$

where $o(h_i^2)/h_i^2 \to 0$ when $i \to \infty$, completing the proof. □

Consider the following classical linearization of control system (38):

$$\begin{cases} \dot{y}(t) = f_x[t]y(t) + v(t) - \dot{\bar{x}}(t), \quad v(t) \in F(t, \bar{x}(t)) \qquad \text{for a.e. } t \in [0,1] \\ y(0) \in T^{\flat}_{K_0}(\bar{x}(0)). \end{cases} \tag{41}$$

Remark 5.2. It is not difficult to check that for a.e. $t \in [0,1]$, $f_x[t]y(t) \in d_xF(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t))$. Thus by (7), any solution $y \in W^{1,1}([0,1];\mathbb{R}^n)$ of (41) satisfies $\dot{y}(t) \in d_xF(t, \bar{x}(t), \dot{\bar{x}}(t))(y(t))$ a.e.

Proposition 5.3. *Assume (H1) and that for some $\varepsilon > 0$, $a_0 \in L^1([0,1];\mathbb{R}_+)$ and for a.e. $t \in [0,1]$, $f_x(t, \cdot, \bar{u}(t))$ is Lipschitz on $B(\bar{x}(t), \varepsilon)$ with Lipschitz constant $a_0(t)$. Then for any solution y of (41), there exists $h_0 > 0$ such that for $R := \|\bar{x}\|_\infty + \|y\|_\infty$ and for every $h \in [0, h_0]$ the following inequality holds true:*

$$\mathrm{dist}_{F(t,\bar{x}(t)+hy(t))}(\dot{\bar{x}}(t) + h\dot{y}(t)) \le (2k_R(t)R + a_0(t)R^2)h^2 \quad a.e.$$

Moreover, if for all $t \in [0,1]$, $y(t) \in C_K(\bar{x}(t))$ and $\mathrm{int}\, C_K(\bar{x}(t)) \neq \emptyset$, *then $y \in \mathcal{V}^{(1)}(\bar{x})$.*

Proof. There exists $0 < h_0 < 1$ such that for a.e. $t \in [0,1]$ and for all $h \in [0, h_0]$,

$$\begin{aligned} \dot{\bar{x}}(t) + hf_x[t]y(t) &\in f(t, \bar{x}(t) + hy(t), \bar{u}(t)) + a_0(t)h^2|y(t)|^2\overline{B} \\ &\subset F(t, \bar{x}(t) + hy(t)) + a_0(t)h^2R^2\overline{B} \end{aligned} \tag{42}$$

and

$$\begin{aligned} \dot{\bar{x}}(t) + v(t) - \dot{\bar{x}}(t) \in F(t, \bar{x}(t)) &\subset F(t, \bar{x}(t) + hy(t)) + hk_R(t)|y(t)|\overline{B} \\ &\subset F(t, \bar{x}(t) + hy(t)) + hk_R(t)R\overline{B}. \end{aligned} \tag{43}$$

Multiplying (42) by $1-h$ and (43) by h, adding them and using the convexity of $F(t, \bar{x}(t) + hy(t))$, we obtain

$$\begin{aligned} \dot{\bar{x}}(t) + hf_x[t]y(t) + h(v(t) - \dot{\bar{x}}(t)) \in h^2 f_x[t]y(t) + F(t, \bar{x}(t) + hy(t)) \\ + a_0(t)h^2R^2\overline{B} + k_R(t)h^2R\overline{B}, \end{aligned}$$

and the first statement of our proposition follows. Lemma 2.6 completes the proof. □

Corollary 5.4. *Under all the assumptions of Proposition* 5.3*, suppose that K is sleek and that* $\operatorname{int} T_K^\flat(\bar{x}(t)) \neq \emptyset$ *for every $t \in [0,1]$. Then every solution y of*

$$(44)\quad \begin{cases} \dot{y}(t) = f_x[t]y(t) + v(t) - \dot{\bar{x}}(t), \quad v(t) \in F(t, \bar{x}(t)) \qquad \text{for a.e. } t \in [0,1] \\ y(t) \in T_K^\flat(\bar{x}(t)) \ \ \forall\, t \in [0,1] \\ y(0) \in T_{K_0}^\flat(\bar{x}(0)) \end{cases}$$

satisfies $y \in \mathcal{V}^{(1)}(\bar{x})$.

For every $y \in \mathcal{V}^{(1)}(\bar{x})$ define the following subset of the set of admissible second-order variations:

$$\mathcal{V}_C^{(2)}(\bar{x}, y) := \left\{ w \in \mathcal{V}^{(2)}(\bar{x}, y) \;\middle|\; w \text{ is solution of } (40) \right\}.$$

Let us consider the following Mayer optimal control problem:

$$(45)\qquad \text{Minimize } \{\varphi(x(1)) \mid x(\cdot) \in \mathcal{S}_K(x_0),\ \ x_0 \in K_0\},$$

where $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is a given twice differentiable function. Next, we recall the celebrated maximum principle, see for instance [26, Thm. 9.5.1]:

Theorem 5.5 (Maximum Principle)**.** *Assume (H1) and let $(\bar{x}, \bar{u})$ be a strong local minimizer of problem* (45) *such that* $\operatorname{int} C_K(\bar{x}(t)) \neq \emptyset$ *for every $t \in [0,1]$. If* (39) *holds true, then there exist $p \in W^{1,1}([0,1];\mathbb{R}^n)$, $\lambda \in \{0,1\}$, a non-negative Borel measure μ on $[0,1]$ and a Borel measurable $\nu\colon [0,1] \to \mathbb{R}^n$ satisfying,*

$$(46)\qquad \nu(t) \in N_K(\bar{x}(t)) \cap \overline{B} \quad \mu\text{-a.e.},$$

such that for $\psi : [0,1] \to \mathbb{R}^n$ defined by $\psi(t) := \int_{[0,t]} \nu(s)d\mu(s)$ if $t \in\,]0,1]$ and $\psi(0) = 0$ we have $(p, \psi, \lambda) \neq 0$,

(*i*) $-\dot{p}(t) = f_x(t, \bar{x}(t), \bar{u}(t))^*(p(t) + \psi(t)) \quad a.e.$

(*ii*) $p(0) \in N_{K_0}(\bar{x}(0))$

(*iii*) $-p(1) = \lambda \nabla\varphi(\bar{x}(1)) + \psi(1)$

(*iv*) $\langle p(t) + \psi(t), f(t, \bar{x}(t), \bar{u}(t))\rangle = \max_{u \in U(t)} \langle p(t) + \psi(t), f(t, \bar{x}(t), u)\rangle \quad a.e.$

Remark 5.6. Note that in [26, Thm. 9.5.1] integrals are taken over $[0, t[$ in the definition of ψ. However, since ψ is of bounded total variation, it has only a countable number of jumps, thus (i) and (iv) remain both valid a.e. Moreover, (46) follows from [26, Thm. 9.5.1] since $K = \{\xi \in \mathbb{R}^n \mid \mathrm{dist}_K(\xi) \leq 0\}$ and by [26, Prop. 4.7.6, Thm. 4.8.5] we always have $\partial \mathrm{dist}_K(\xi) \subset N_K(\xi)$ for all $\xi \in K$. Here ∂ denotes the Clarke subdifferential.

A tuple $(\bar{x}, \bar{u}, p, \psi, \mu, \nu, \lambda)$ such that $(\bar{x}, \bar{u})$ is a trajectory control pair of (38) and $(p, \psi, \mu, \nu, \lambda)$ is as in the maximum principle, is called an *extremal.* An extremal is normal if $\lambda = 1$. There are several results on normality of the maximum principle for optimal control problems, see for instance [2, 10, 11, 12, 13, 14, 15, 23] and the references therein. We provide next second-order necessary optimality conditions for problem (45):

Theorem 5.7. *Assume* (H1) *and* (IPC) *and let* $(\bar{x}, \bar{u})$ *be a strong local minimizer. Then*

$$\langle \nabla \varphi(\bar{x}(1)), y(1) \rangle \geq 0 \qquad \forall\, y \in \mathcal{V}^{(1)}(\bar{x}).$$

If the maximum principle of Theorem 5.5 *holds true with* $\lambda = 1$ *and some* (p, ψ, μ, ν) *and* $y \in \mathcal{V}^{(1)}(\bar{x})$ *is such that* $\langle \nabla \varphi(\bar{x}(1)), y(1) \rangle = 0$, *then for every* $w \in \mathcal{V}_C^{(2)}(\bar{x}, y)$,

$$\frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) - \langle p(0), w(0) \rangle - \int_{[0,1]} \langle \nu(t), w(t) \rangle \, d\mu(t) - \int_0^1 \langle p(t) + \psi(t), \hat{v}(t) \rangle \, dt \geq 0,$$

where $\hat{v}(t) := \dot{w}(t) - f_x[t]w(t) \in \mathcal{F}_C(y; t) + \mathcal{E}(y; t)$ *a.e.*

P r o o f. Theorem 4.7 implies the first statement. Let $y \in \mathcal{V}^{(1)}(\bar{x})$ be such that $\langle \nabla \varphi(\bar{x}(1)), y(1) \rangle = 0$. Then, by Theorem 4.7, for all $w \in \mathcal{V}^{(2)}(\bar{x}, y)$ we have

$$\langle \nabla \varphi(\bar{x}(1)), w(1) \rangle + \frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) \geq 0. \tag{47}$$

If $w \in \mathcal{V}_C^{(2)}(\bar{x}, y)$, then $\dot{w}(t) = f_x[t]w(t) + \hat{v}(t)$ for some $\hat{v}(t) \in \mathcal{F}_C(y; t) +$

$\mathcal{E}(y;t)$ and

$$\begin{aligned}\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle &= \langle -p(1)-\psi(1), w(1)\rangle \\ &= \int_0^1 \big(\langle -p(t)-\psi(t), \dot{w}(t)\rangle + \langle -\dot{p}(t), w(t)\rangle \big) dt \\ &\qquad - \int_{[0,1]} \langle \nu(t), w(t)\rangle \, d\mu(t) - \langle p(0), w(0)\rangle \\ &= -\int_0^1 \langle p(t)+\psi(t), \hat{v}(t)\rangle \, dt \\ &\qquad - \int_{[0,1]} \langle \nu(t), w(t)\rangle \, d\mu(t) - \langle p(0), w(0)\rangle \, .\end{aligned}$$

This and (47) complete the proof. □

Remark 5.8. Let $(\bar{x}, \bar{u}, p, \psi, \mu, \nu)$ be a normal extremal, $y \in \mathcal{V}^{(1)}(\bar{x})$ satisfy (41), $y(t) \in C_K(\bar{x}(t))$ for all $t \in [0,1]$ and $y(0) \in C_{K_0}(\bar{x}(0))$. Then $\dot{y}(t) = f_x[t]y(t) + v(t) - \dot{\bar{x}}(t)$ for some integrable selection $v(t) \in F(t, \bar{x}(t))$ a.e. By (iv) of Theorem 5.5 we have that for a.e. $t \in [0,1]$,

$$\langle p(t)+\psi(t), f(t,\bar{x}(t),u) - f[t]\rangle \le 0 \qquad \forall\, u \in U(t), \tag{48}$$

which implies that $p(t)+\psi(t) \in \mathcal{L}(t)^-$ for almost every t. This in turn yields,

(49)

$$\begin{aligned} 0 &\le \int_0^1 - \langle p(t)+\psi(t), v(t) - f[t]\rangle \, dt \\ &= \int_0^1 \big(- \langle \dot{p}(t), y(t)\rangle - \langle p(t)+\psi(t), \dot{y}(t)\rangle \big) dt \\ &= \int_0^1 \big(- \langle \dot{p}(t), y(t)\rangle - \langle p(t), \dot{y}(t)\rangle \big) dt + \int_{[0,1]} \langle \nu(t), y(t)\rangle \, d\mu(t) - \langle \psi(1), y(1)\rangle \\ &\le \langle -p(1)-\psi(1), y(1)\rangle + \langle p(0), y(0)\rangle \le \langle \nabla\varphi(\bar{x}(1)), y(1)\rangle \, . \end{aligned}$$

If $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, then it follows from (49) and (iv) of Theorem 5.5 that $\langle p(t)+\psi(t), v(t)-f[t]\rangle = 0$ a.e. In the same way as in Remark 4.6 we deduce that in this case for almost every $t \in [0,1]$,

$$\langle p(t)+\psi(t), w\rangle \le 0 \qquad \forall\, w \in T^{\flat(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t), v(t) - f[t]).$$

Consider the sets

$$\overline{U}(t) := \{u \in U(t) \mid \langle p(t)+\psi(t), f(t,\bar{x}(t),u)\rangle = \langle p(t)+\psi(t), f(t,\bar{x}(t),\bar{u}(t))\rangle\} \, ,$$

and the constrained control system

$$(50)\qquad \begin{cases} \dot{y}(t) = f_x[t]y(t) + f(t,\bar{x}(t),u(t)) - f[t], \quad u(t) \in U(t) \quad \text{a.e.} \\ y(0) \in C_{K_0}(\bar{x}(0)) \\ y(t) \in C_K(\bar{x}(t)) \quad \forall\, t \in [0,1]. \end{cases}$$

We deduce that if $f(t,\bar{x}(t),U(t)) \neq f(t,\bar{x}(t),\bar{u}(t))$ on a set of positive measure, then the following alternative holds true: either for every $y(\cdot) \in \mathcal{V}^{(1)}(\bar{x})$ satisfying (50) with $f(\cdot,\bar{x}(\cdot),u(\cdot)) \neq f[\cdot]$ on a set of positive measure we have $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle > 0$, or $f(t,\bar{x}(t),\overline{U}(t)) \neq \{f(t,\bar{x}(t),\bar{u}(t))\}$ on a set of positive measure. Consequently, also $\overline{U}(t) \neq \{\bar{u}(t)\}$ on a set of positive measure. We can consider then such normal extremals as singular.

Example 5.9. Let $f(t,\cdot,\cdot)$ be twice differentiable for all $t \in [0,1]$ and $f'(t,\cdot,\cdot)$, $f''(t,\cdot,\cdot)$ denote the derivative, respectively the Hessian of the map $(x,u) \mapsto f(t,x,u)$. Assume that for some $\varepsilon > 0$, $a_3 \in L^1([0,1];\mathbb{R}^n)$ and for a.e. $t \in [0,1]$, the mappings $f'(t,\cdot,\cdot)$, $f''(t,\cdot,\cdot)$ are Lipschitz on $B(\bar{x}(t),\varepsilon) \times B(\bar{u}(t),\varepsilon)$ with the Lipschitz constant $a_3(t)$. Let $u \in L^\infty([0,1];\mathbb{R}^m)$ be such that $u(t) \in T^\flat_{U(t)}(\bar{u}(t))$ a.e. and for some $c, h_0 > 0$, $\mathrm{dist}_{U(t)}(\bar{x}(t) + hu(t)) \leq ch^2$ for every $h \in [0,h_0]$ and a.e. $t \in [0,1]$. Then, by [18, Prop. 4.1], for all $h > 0$, there exists $u_h \in L^\infty([0,1];\mathbb{R}^m)$ such that $u_h(\cdot) \to u(\cdot)$ a.e. when $h \to 0+$, $\|u_h\|_\infty \leq 2\,\|u\|_\infty$ and $\bar{u}(t) + hu_h(t) \in U(t)$ a.e. Using this fact, it is not difficult to check that $f_u[t]u(t) \in \mathcal{L}(t)$ a.e. Consider a solution $y \in W^{1,1}([0,1];\mathbb{R}^n)$ of

$$(51)\qquad \begin{cases} \dot{y}(t) = f_x[t]y(t) + f_u[t]u(t) \quad \text{for a.e. } t \in [0,1] \\ y(0) \in T^\flat_{K_0}(\bar{x}(0)) \\ y \in T^\flat_{\mathcal{K}}(\bar{x}), \end{cases}$$

and assume that there exists $\bar{v} \in L^\infty([0,1];\mathbb{R}^m)$ such that $\bar{v}(t) \in T^{\flat(2)}_{U(t)}(\bar{x}(t),u(t))$ a.e. Then [18, Prop. 4.2] states that for every $h > 0$ there exists $\bar{v}_h \in L^\infty([0,1];\mathbb{R}^m)$ satisfying $\bar{v}_h(\cdot) \to \bar{v}(\cdot)$ a.e. when $h \to 0+$, $\|\bar{v}_h\|_\infty \leq 2\,\|\bar{v}\|_\infty + c$ and $\bar{u}(t) + hu(t) + h^2\bar{v}_h(t) \in U(t)$ a.e. It is not difficult to verify then that $y \in \mathcal{V}^{(1)}(\bar{x})$.

By the Taylor formula we find that

$$f_u[t]\bar{v}(t) + \frac{1}{2}f_{xx}[t]y(t)y(t) + f_{xu}[t]y(t)u(t) + \frac{1}{2}f_{uu}[t]u(t)u(t) \in \mathcal{F}_C(y;t) \quad \textit{a.e.}$$

Hence the second-order approximation obtained in [18] is a special case of the second-order approximation (40) introduced in this section. Consequently, in the case of strong local minimizers, the second-order necessary optimality conditions of Theorem 5.7 generalize those of [18, Thm. 3.5].

Next we deduce from the above results a pointwise second-order necessary condition for optimality. Consider the following "second-order linearization" of (38):

$$\text{(52)}\qquad \begin{cases} \dot{w}(t) \in f_x[t]w(t) + \mathcal{E}(y;t) & \text{for a.e. } t \in [0,1] \\ w(0) \in C_{K_0}(\bar{x}(0)) \\ w(t) \in C_K(\bar{x}(t)) \quad \text{for all} \quad t \in [0,1]. \end{cases}$$

Theorem 5.10 (Second-Order Maximum Principle). *Let* $(\bar{x}, \bar{u})$ *be a strong local minimizer of problem* (45) *and* (H1), (IPC), (39) *hold true. If* $y \in \mathcal{V}^{(1)}(\bar{x})$ *is such that* $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$ *and* $\mathcal{V}^{(2)}_C(\bar{x}, y) \neq \emptyset$*, then for every solution* w *of* (52) *we have* $\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle \geq 0$.

Furthermore, there exist $\lambda \in \{0,1\}$, $\psi \in NBV([0,1];\mathbb{R}^n)$ *and* $p \in W^{1,1}([0,1];\mathbb{R}^n)$ *satisfying all the conclusions of Theorem* 5.5 *such that in addition*

$$\text{(53)}\qquad \max_{v \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))} \langle p(t) + \psi(t), v\rangle = \langle p(t) + \psi(t), \dot{y}(t)\rangle \quad \text{a.e. in } [0,1].$$

Finally, if the following linear system under state constraint has a solution:

$$\text{(54)}\qquad \begin{cases} \dot{z}(t) \in f_x[t]z(t) + \mathcal{E}(y;t) & a.e. \\ z(0) \in C_{K_0}(\bar{x}(0))) \\ z(t) \in \operatorname{int} C_K(\bar{x}(t)) & \textit{for all} \quad t \in [0,1], \end{cases}$$

then the above holds true with $\lambda = 1$.

Proof. Let $\bar{w} \in \mathcal{V}^{(2)}_C(\bar{x}, y)$ and w be a solution of (52). Then it follows from Lemmas 2.4, 2.5 and Proposition 5.1 that $\bar{w} + w \in \mathcal{V}^{(2)}_C(\bar{x}, y)$. Therefore, by Theorem 4.7 we have (37). In the same way as in the proof of Theorem 4.8 we deduce that $\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle \geq 0$ for every solution w of (52). To complete the proof it is enough to apply the same arguments as in [25, pp. 358-361] with $V(t)$ defined in [25, pp. 356] replaced by $\mathcal{E}(y;t)$. □

Corollary 5.11. *Let* $(\bar{x}, \bar{u})$ *be a strong local minimizer of Problem* (45) *and* (H1), (IPC), (39) hold true. Let a solution y of (41) satisfy $y \in \mathcal{V}^{(1)}(\bar{x})$, $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, $\mathcal{V}^{(2)}_C(\bar{x}, y) \neq \emptyset$ and consider p, ψ as in the conclusions of Theorem 5.10. Then for a.e. $t \in [0,1]$,

$$\langle p(t)+\psi(t), f_x[t]y(t)\rangle = \max\Big\{ \langle p(t)+\psi(t), \sum_{i=1}^{k} c_i f_x(t,\bar{x}(t),u_i)y(t)\rangle \mid c_i \geq 0, \sum_{i=1}^{k} c_i = 1, u_i \in U(t), f[t] = \sum_{i=1}^{k} c_i f(t,\bar{x}(t),u_i)\Big\}.$$

Proof. Let $v(t) \in F(t,\bar{x}(t))$ be such that $\dot{y}(t) = f_x[t]y(t) + v(t) - \dot{\bar{x}}(t)$ a.e. Fix $t \in [0,1]$ such that (53) is satisfied. Consider an integer $k > 0$, $u_i \in U(t)$ and $c_i \geq 0$, $i = 1, \ldots, k$ such that $\sum_{i=1}^{k} c_i = 1$ and $f(t,\bar{x}(t),\bar{u}(t)) = \sum_{i=1}^{k} c_i f(t,\bar{x}(t),u_i)$. Then

$$\sum_{i=1}^{k} c_i f_x(t,\bar{x}(t),u_i)y(t) \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))$$

and therefore

$$\sum_{i=1}^{k} c_i f_x(t,\bar{x}(t),u_i)y(t) + v(t) - \dot{\bar{x}}(t) \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t)).$$

Consequently, for a.e. $t \in [0,1]$,

$$\left\langle p(t)+\psi(t), \sum_{i=1}^{k} c_i f_x(t,\bar{x}(t),u_i)y(t) \right\rangle \leq \langle p(t)+\psi(t), f_x[t]y(t)\rangle. \qquad \square$$

Remark 5.12. The above corollary can be stated in a much more general way. Namely let p, ψ be as in the conclusions of Theorem 5.10 and t be such that (53) is satisfied. Consider any mappings $c_i \colon \mathbb{R}^n \to \mathbb{R}_+$, $i = 1, \ldots, k$ that are continuously differentiable on a neighborhood of $\bar{x}(t)$ and such that $\sum_{i=1}^{k} c_i(x) = 1$ for every x sufficiently close to $\bar{x}(t)$. Then for all $u_i \in U(t)$ satisfying $f[t] = \sum_{i=1}^{k} c_i(\bar{x}(t)) f(t,\bar{x}(t),u_i)$ the following inequality holds true:

$$\langle p(t)+\psi(t), f_x[t]y(t)\rangle \geq \left\langle p(t)+\psi(t), \sum_{i=1}^{k} c_i(\bar{x}(t)) f_x(t,\bar{x}(t),u_i)y(t) + \sum_{i=1}^{k} (c_i'(\bar{x}(t))y(t)) f(t,\bar{x}(t),u_i) \right\rangle.$$

6. Optimality conditions in the presence of endpoint constraints. In this section we impose an additional endpoint constraint:

$$x(1) \in K_1, \tag{55}$$

where $K_1 \subset \mathbb{R}^n$ is nonempty and closed. Consider again the differential inclusion (15). As before, $F(t,x) := \operatorname{co} \widetilde{F}(t,x)$ for all $(t,x) \in [0,1] \times \mathbb{R}^n$.

Theorem 6.1. *Assume* (A), (IPC) *and let* $\bar{x} \in \mathcal{S}_K(\bar{x}^0)$ *for some* $\bar{x}^0 \in K_0$ *satisfy* $\bar{x}(1) \in K_1$. *Suppose that* $y \in \mathcal{V}^{(1)}(\bar{x})$ *and* $w \in \mathcal{V}^{(2)}(\bar{x}, y)$ *are such that* $w(1) \in D^2_{K_1}(\bar{x}(1), y(1))$. *Consider any sequences* $h_i \to 0+$, $w_i^0 \to w(0)$ *with* $\bar{x}(0) + h_i y(0) + h_i^2 w_i^0 \in K_0$. *Then there exist* $x_i \in \mathcal{S}_K(\bar{x}(0) + h_i y(0) + h_i^2 w_i^0)$ *satisfying* $x_i(1) \in K_1$ *and such that* $\dfrac{1}{h_i^2}(x_i - \bar{x} - h_i y)$ *converge uniformly to* w *when* $i \to \infty$.

P r o o f. Let x_i be as in Theorem 3.3. It is enough to observe that since $w(1) \in D^2_{K_1}(\bar{x}(1), y(1))$ and $\dfrac{1}{h_i^2}(x_i(1) - \bar{x}(1) - h_i y(1))$ converge to $w(1)$ for $i \to \infty$, we have $x_i(1) \in K_1$, for i sufficiently large. □

For any $y \in \mathcal{V}^{(1)}(\bar{x})$ define the following subset of second-order admissible variations:

$$\mathcal{V}_I^{(2)}(\bar{x}, y, K_1) := \left\{ w \in \mathcal{V}_I^{(2)}(\bar{x}, y) \;\middle|\; w(1) \in D^2_{K_1}(\bar{x}(1), y(1)) \right\}.$$

Consider now the Mayer problem

$$\text{Minimize } \{\varphi(x(1)) \mid x(\cdot) \in \mathcal{S}_K(x_0),\ x_0 \in K_0,\ x(1) \in K_1\}, \tag{56}$$

where $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ is a given twice differentiable function. Let us start by recalling the first-order necessary optimality conditions from [6, Corollary 3.8] in this more general case.

Theorem 6.2. *Let* $\bar{x}$ *be a strong local minimizer of Problem* (56) *and for all* $t \in [0,1]$, $\operatorname{int} C_K(\bar{x}(t)) \neq \emptyset$. *Assume* $(\widetilde{A})$ *and that* (A) *holds true with a bounded* $a_1(\cdot)$. *If* $\operatorname{int} C_{K_0}(\bar{x}(0)) \cap \operatorname{int} C_K(\bar{x}(0)) \neq \emptyset$ *and* $\operatorname{int} C_{K_1}(\bar{x}(1)) \neq \emptyset$, *then there exist* λ, ψ, p *as in Theorem* 4.3 *with* (*iii*) *replaced by*

$$p(1) \in -\lambda \nabla\varphi(\bar{x}(1)) - \psi(1) - N_{K_1}(\bar{x}(1)).$$

Moreover, if $C_{K_1}(\bar{x}(1)) \cap \operatorname{int} C_K(\bar{x}(1)) \neq \emptyset$, *then* $\lambda + \sup_{t \in]0,1[} |p(t) + \psi(t)| \neq 0$. *Finally, if there exists a solution* y *of* (27) *satisfying* $y(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$, *then the above holds true with* $\lambda = 1$.

The very same proof as the one of Corollary 4.4 implies the following result.

Corollary 6.3. *Under all the assumptions of Theorem* 6.2 *consider any* λ, p, ψ *as in its conclusions. Then for almost every* $t \in [0,1]$ *and all* $v \in \mathbb{R}^m$ *satisfying* $\langle p(t)+\psi(t), v\rangle = 0$ *we have*

$$\langle p(t)+\psi(t), w\rangle \leq 0 \qquad \forall\, w \in T^{\flat(2)}_{F(t,\bar{x}(t))}(\dot{\bar{x}}(t), v).$$

Remark 6.4. Let $(\bar{x}, p, \psi, \mu, \nu)$ be a normal extremal. Consider $y \in \mathcal{V}^{(1)}(\bar{x})$ satisfying (30) with $y(0) \in C_{K_0}(\bar{x}(0))$, $y(1) \in C_{K_1}(\bar{x}(1))$, $y(t) \in C_K(\bar{x}(t))$ for all $t \in [0,1]$. Then conclusions similar to those of Remark 4.6 hold true.

Using Theorem 6.1 and the same proof strategy as for Theorem 4.7 we find the following necessary optimality conditions for problem (56):

Theorem 6.5. *Assume* (A), (IPC) *and let* $\bar{x}$ *be a strong local minimizer of problem* (56). *Then,*

$$\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle \geq 0 \qquad \forall\, y \in \mathcal{V}^{(1)}(\bar{x}) \ \text{ satisfying } \ y(1) \in D_{K_1}(\bar{x}(1)). \tag{57}$$

Moreover, for all $y \in \mathcal{V}^{(1)}(\bar{x})$ *such that* $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, *we have*

$$\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle + \frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) \geq 0 \ \ \forall\, w \in \mathcal{V}^{(2)}(\bar{x}, y) \quad \text{satisfying } \ w(1) \in D^2_{K_1}(\bar{x}(1), y(1)). \tag{58}$$

As in Section 4, we can prove the following second-order maximum principle.

Theorem 6.6. *Let* $\bar{x}$ *be a strong local minimizer of problem* (56), $(\widetilde{A})$, (IPC), (A) *hold with a bounded* $a_1(\cdot)$ *and* $\operatorname{int} C_{K_1}(\bar{x}(1)) \neq \emptyset$. *Let* $y \in \mathcal{V}^{(1)}(\bar{x})$ *be such that* $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$ *and* $\mathcal{V}_I^{(2)}(\bar{x}, y, K_1) \neq \emptyset$. *Then for every solution* w *of* (35) *such that* $w(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$, *we have* $\langle \nabla\varphi(\bar{x}(1)), w(1)\rangle \geq 0$.

Furthermore, if $\operatorname{int} C_{K_0}(\bar{x}(0)) \cap \operatorname{int} C_K(\bar{x}(0)) \neq \emptyset$, *then there exist* λ, ψ, p *as in the conclusions of Theorem 6.2 such that in addition*

$$\max_{v \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))} \langle p(t)+\psi(t), v\rangle = \langle p(t)+\psi(t), \dot{y}(t)\rangle \ \text{ a.e. in } [0,1].$$

Finally, if the constrained differential inclusion (36) *has a solution* z *satisfying* $z(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$, *then the above holds true with* $\lambda = 1$.

Proof. Let $\bar{w} \in \mathcal{V}_I^{(2)}(\bar{x}, y, K_1)$ and w be a solution of (35) such that $w(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$. By Lemma 2.8, $\bar{w}(1)+w(1) \in D^2_{K_1}(\bar{x}(1), y(1))$. Arguments similar to those of the proof of Theorem 4.8 imply the result. □

Consider the constrained control system (38) under an additional endpoint constraint (55) and define the set-valued maps $\tilde{F}$, F as in Section 5. For a solution y of (23) let the sets $\mathcal{F}_C(y;t) \subset \mathbb{R}^n$ be as in Section 5. For every $y \in \mathcal{V}^{(1)}(\bar{x})$ we introduce the following subset of the set of admissible second-order variations:

$$\mathcal{V}_C^{(2)}(\bar{x}, y, K_1) := \left\{ w \in \mathcal{V}_C^{(2)}(\bar{x}, y) \;\middle|\; w(1) \in D^2_{K_1}(\bar{x}(1), y(1)) \right\}.$$

We investigate next the Mayer optimal control problem (56) where $x_0 \rightsquigarrow \mathcal{S}_K(x_0)$ is understood with respect to the control system (38).

Theorem 6.7 ([26]). *Let $(\bar{x}, \bar{u})$ be a strong local minimizer of problem* (56) *such that* $\operatorname{int} C_K(\bar{x}(t)) \neq \emptyset$ *for every $t \in [0,1]$. If (H1) and* (39) *hold true, then there exist λ, ψ, p, μ, ν as in Theorem 5.5, with (iii) replaced by* $-p(1) \in \lambda \nabla\varphi(\bar{x}(1)) + \psi(1) + N_{K_1}(\bar{x}(1))$.

The second-order necessary optimality conditions for problem (56) are similar to those derived in Section 5.

Theorem 6.8. *Let $(\bar{x}, \bar{u})$ be a strong local minimizer of problem* (56) *and* (H1), (IPC) *be satisfied. Then,*

$$\langle \nabla\varphi(\bar{x}(1)), y(1) \rangle \geq 0 \qquad \forall\, y \in \mathcal{V}^{(1)}(\bar{x}) \;\text{ satisfying }\; y(1) \in D_{K_1}(\bar{x}(1)).$$

Moreover, if the maximum principle of Theorem 6.7 holds true with $\lambda = 1$ and some (p, ψ, μ, ν) and $y \in \mathcal{V}^{(1)}(\bar{x})$ is such that $\langle \nabla\varphi(\bar{x}(1)), y(1) \rangle = 0$, then for every $w \in \mathcal{V}_C^{(2)}(\bar{x}, y, K_1)$,

$$\frac{1}{2}\varphi''(\bar{x}(1))y(1)y(1) - \langle p(0), w(0) \rangle - \langle \bar{n}, w(1) \rangle \\ - \int_{[0,1]} \langle \nu(t), w(t) \rangle \, d\mu(t) - \int_0^1 \langle p(t) + \psi(t), \hat{v}(t) \rangle \, dt \geq 0,$$

where $\bar{n} = -p(1) - \nabla\varphi(\bar{x}(1)) - \psi(1) \in N_{K_1}(\bar{x}(1))$ and $\hat{v}(t) := \dot{w}(t) - f_x[t]w(t) \in \mathcal{F}_C(y;t) + \mathcal{E}(y;t)$ a.e.

Theorem 6.9. *Let $(\bar{x}, \bar{u})$ be a strong local minimizer of problem* (56) *and* (H1), (IPC), (39) *hold true. If $y \in \mathcal{V}^{(1)}(\bar{x})$ is such that $\langle \nabla\varphi(\bar{x}(1), y(1) \rangle = 0$ and $\mathcal{V}_C^{(2)}(\bar{x}, y, K_1) \neq \emptyset$, then for every solution w of* (52) *satisfying* $w(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$, *we have $\langle \nabla\varphi(\bar{x}(1)), w(1) \rangle \geq 0$.*

Furthermore, there exist λ, ψ, p *as in Theorem* 6.7 *such that in addition*

$$\max_{v \in d_x F(t,\bar{x}(t),\dot{\bar{x}}(t))(y(t))} \langle p(t)+\psi(t), v\rangle = \langle p(t)+\psi(t), \dot{y}(t)\rangle \quad \text{a.e. in } [0,1]. \tag{59}$$

Finally, if (54) *has a solution* z *with* $z(1) \in \operatorname{int} C_{K_1}(\bar{x}(1))$, *then the above holds true with* $\lambda = 1$.

Corollary 6.10. *Let* $(\bar{x},\bar{u})$ *be a strong local minimizer of problem* (56) *and* (H1), (IPC), (39) *hold true. If* y *is a solution of* (41) *such that* $y \in \mathcal{V}^{(1)}(\bar{x})$, $\langle \nabla\varphi(\bar{x}(1)), y(1)\rangle = 0$, $\mathcal{V}_C^{(2)}(\bar{x}, y, K_1) \neq \emptyset$ *and* λ, p, ψ *are as in the conclusions of Theorem* 6.9, *then for a.e.* $t \in [0,1]$,

$$\begin{aligned}\langle p(t)+\psi(t), f_x[t]y(t)\rangle &= \\ \max\Bigg\{ \langle p(t)+\psi(t), &\sum_{i=1}^{k} c_i f_x(t,\bar{x}(t),u_i)y(t)\rangle \,|\, c_i \geq 0, \\ &\sum_{i=1}^{k} c_i = 1, u_i \in U(t), f[t] = \sum_{i=1}^{k} c_i f(t,\bar{x}(t),u_i) \Bigg\}.\end{aligned}$$

REFERENCES

[1] J.-P. Aubin, H. Frankowska. Set-Valued Analysis, Systems Control Found. Appl. Boston, MA, Birkhäuser Boston, Inc., 1990.

[2] P. Bettiol, H. Frankowska. Normality of the maximum principle for nonconvex constrained Bolza problems. *J. Differential Equations* **243** (2007), 256–269.

[3] P. Bettiol, H. Frankowska, R. B. Vinter. L^∞ estimates on trajectories confined to a closed subset. *J. Differential Equations* **252** (2012), 1912–1933.

[4] J. F. Bonnans, A. Hermant. Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** (2009), 561–598.

[5] J. F. Bonnans, A. Shapiro. Perturbation Analysis of Optimization Problems. Springer Series in Operations Research. New York, Springer-Verlag, 2000.

[6] A. Cernea, H. Frankowska. A connection between the maximum principle and dynamic programming for constrained optimal control problems. *SIAM J. Control Optim.* **44** (2005), 673–703.

[7] R. Cominetti. Metric regularity, tangent sets and second order optimality conditions. *Appl. Math. Optim.* **21** (1990), 265–287.

[8] M. C. Delfour, J.-P. Zolésio. Shapes and Geometries: Analysis, Differential Calculus and Optimization. Advances in Design and Control, vol. **4**. Philadelphia, PA, SIAM, 2001.

[9] A. L. Dontchev, R. T. Rockafellar. Implicit Functions and Solution Mappings. A view from variational analysis. Springer Monographs in Mathematics. Dordrecht, Springer, 2009.

[10] F. Fontes, S. Lopez. Normal forms of necessary conditions for dynamic optimization problems with pathwise inequality constraints. *J. Math. Anal. Appl.* **399** (2013), 27–37.

[11] H. Frankowska. Regularity of minimizers and of adjoint states in optimal control under state constraints. *J. Convex Anal.* **13** (2006), 299–328.

[12] H. Frankowska. Normality of the maximum principle for absolutely continuous solutions to Bolza problems under state constraints. *Control Cybernet.* **38** (2009), 1327–1340.

[13] H. Frankowska, M. Mazzola. On relations of the adjoint state to the value function for optimal control problems with state constraints. *NoDEA Nonlinear Differential Equations Appl.* **20**, *2* (2013), 361–383.

[14] H. Frankowska, D. Tonon. Inward pointing trajectories, normality of the maximum principle and the non occurence of the Lavrentieff phenomenon in optimal control under state constraints. *J. Convex Anal.* **20**, 2013.

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

[16] E. G. Gilbert, D. S. Bernstein. Second-order necessary conditions in optimal control: Accessory-problem results without normality conditions. *J. Optim. Theory Appl.* **41** (1983), 75–106.

[17] D. Hoehener. Second-order optimality conditions for a Bolza problem with mixed constraints. In: Proceedings of the 18th IFAC World Congress (Eds S. Bittanti, A. Cenedese, S. Zampieri), 2011, 2594–2599.

[18] D. Hoehener. Variational approach to second-order optimality conditions for control problems with pure state constraints. *SIAM J. Control Optim.* **50** (2012), 1139–1173.

[19] H. Kawasaki. Second order necessary optimality conditions for minimizing a sup-type function. *Math. Program.* **49** (1991), 213–229.

[20] Z. Páles, V. Zeidan. Nonsmooth optimum problems with constraints. *SIAM J. Control Optim.* **32** (1994), 1476–1502.

[21] Z. Páles, V. Zeidan. Optimum problems with certain lower semicontinuous set-valued constraints. *SIAM J. Optim.* **8** (1998), 707–727.

[22] Z. Páles, V. Zeidan. Optimal control problems with set-valued control and state constraints. *SIAM J. Optim.* **14** (2003), 334–358.

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

[24] R. T. Rockafellar. Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$. *Nonlinear Anal.* **3**, *1* (1979), 145–154.

[25] M. Tamzali-Lafond. Variational inclusions under state constraints. *SIAM J. Control Optim.* **42** (2003), 342–362.

[26] R. B. Vinter. Optimal Control. Boston, MA, Birkhäuser, 2000.

[27] J. Warga. A second-order condition that strengthens Pontryagin's maximum principle. *J. Differential Equations* **28** (1979), 284–307.

Combinatoire & Optimisation
Institut de Mathématiques de Jussieu
Université Pierre et Marie Curie
4, place Jussieu
75252 Paris cedex 05, France
e-mail: `frankowska@math.jussieu.fr`
`daniel-hoehener@bluewin.ch`
`tonon@ceremade.dauphine.fr`

Received April 23, 2013