Journal of Convex Analysis
Volume 23 (2016), No. 1, 291–311

Normality of Generalized Euler-Lagrange Conditions for State Constrained Optimal Control Problems

Piernicola Bettiol
Laboratoire de Mathématiques, Université de Bretagne Occidentale,
6 Avenue Victor Le Gorgeu, 29200 Brest, France
piernicola.bettiol@univ-brest.fr

Nathalie Khalil
Laboratoire de Mathématiques, Université de Bretagne Occidentale,
6 Avenue Victor Le Gorgeu, 29200 Brest, France
nathalie.khalil@univ-brest.fr

Richard B. Vinter
Department of Electrical and Electronic Engineering, Imperial College London,
Exhibition Road, London, SW7 2BT, UK
r.vinter@imperial.ac.uk

Dedicated to the memory of Jean Jacques Moreau.

Received: April 17, 2015
Revised manuscript received: October 6, 2015
Accepted: October 30, 2015

We consider state constrained optimal control problems in which the cost to minimize comprises both integral and end-point terms, establishing normality of the generalized Euler-Lagrange condition. Simple examples illustrate that the validity of the Euler-Lagrange condition (and related necessary conditions), in normal form, depends crucially on the interplay between velocity sets, the left end-point constraint set and the state constraint set. We show that this is actually a common feature for general state constrained optimal control problems, in which the state constraint is represented by closed convex sets and the left end-point constraint is a closed set. In these circumstances classical constraint qualifications involving the state constraints and the velocity sets cannot be used alone to guarantee normality of the necessary conditions. A key feature of this paper is to prove that the additional information involving tangent vectors to the left end-point and the state constraint sets can be used to establish normality.

Keywords: Optimal control, necessary conditions, differential inclusions, state constraints

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

1. Introduction

Consider the state constrained optimal control problem:

$$(P)\begin{cases}\text{Minimize } J(x(.)) := g(x(S), x(T)) + \int_S^T L(t, x(t), \dot{x}(t))\, dt \\ \quad \text{over } x(.) \in W^{1,1}([S,T];\mathbb{R}^n) \text{ satisfying} \\ \dot{x}(t) \in F(x(t)) \text{ a.e. } t \in [S,T]\,, \\ x(t) \in A \text{ for all } t \in [S,T]\,, \\ x(S) \in E_0\,, \end{cases}$$

in which $[S,T]$ is a given interval ($S < T$), $g(.,.) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and $L(.,.,.) : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are given functions, $F(.) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is a given multifunction with closed non-empty values, A and E_0 are given closed non-empty subsets of $\mathbb{R}^n$. An absolutely continuous arc $x(.) : [a,b] \to \mathbb{R}^n$ ($[a,b] \subset [S,T]$) which satisfies the differential inclusion $\dot{x}(t) \in F(x(t))$, a.e. $t \in [a,b]$, is called F-trajectory (on $[a,b]$). An F-trajectory $x(.)$ for which $x(S) \in E_0$ and $x(t) \in A$ for all $t \in [a,b]$ is called *feasible.*
The infimum of the functional $J(x(.))$ over all feasible state trajectories $x(.)$ is the infimum cost for (P). If no feasible state trajectories exist, a common interpretation of the infimum cost is to take the value $+\infty$. We say that an F-trajectory $\bar{x}(.)$ is a $W^{1,1}$ local minimizer for (P), if there exists $\delta > 0$ such that

$$J(x(.)) \geq J(\bar{x}(.))$$

for all feasible F-trajectories $x(.)$ satisfying $\|x(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \delta$. We shall also refer to $\bar{x}(.)$ as a $W^{1,1}$ *δ-local minimizer* for (P) when it is desirable to specify the positive number δ in the above definition.

Necessary conditions for optimal control problems in various forms have been known for many years (for example see the books [9], [22], and the references therein). But, for a significant class of problems these conditions can provide no useful information in finding minimizers (known as *degenerate* case), or apply in a form in which the cost multiplier vanishes (sometimes referred to as *abnormal* case). In the latter case the cost function would not intervene in detecting candidates to be minimizers. The importance of applying the necessary conditions for optimality with a non-zero cost multiplier, called the *normal* form, is confirmed also by the possibility to investigate regularity properties of minimizers, the non-occurrence of gaps between local minimizers and minimizers for the convexified dynamics, the non-occurrence of the Lavrentieff phenomenon, and other consequences of interest (cf. [22], [18], [15], [16]). A number of different techniques have been employed to obtain non-degeneracy (see for instance [1], [11], [12], [13], and [20]) or normality of necessary conditions (cf. [19], [5], [15], [16], [14], [7], and [4]). In this paper we use an approach suggested in [19] and successively developed in other papers: this is based on existence results of neighboring feasible trajectories (approximating a reference trajectory that possibly violates the state constraint) with associated linear estimates. The class of problems, which we consider (comprising in the cost an integral term which depends on $\dot{x}$), would require exhibiting linear estimates which are linear w.r.t. the $W^{1,1}$-norm for convex (in general non-smooth) state constraints. In addition, our constraint qualification involves a 'classical' (convexified) inward pointing condition: we do not invoke the existence of regular feedback controls (cf. [19]), or conditions requiring inward pointing vectors which are tangent to the velocity set (cf. [16]). But, known counter-examples (see [3] and [6]) clearly show that, in general, 'classical' inward pointing conditions are not enough to derive linear $W^{1,1}$-estimates for non-smooth state constraints. However, a recent paper [4] shows that linear $W^{1,1}$-estimates can be still considered for convex domains and Lipschitz continuous velocity sets, in either of the following two cases.

(a) the left end-point of the reference trajectory belongs to a region where the state constraint A is regular (i.e. we 'start away from corners');
(b) we are free to choose the left end-point for the approximating trajectory.

An application (shown in [4]) of this information is precisely the normality conditions for optimal control problems with convex state constraints when case (a) (whatever the left end-point constraint E_0 is), or alternatively (b) (that is $E_0 = \mathbb{R}^n$), occurs.

This paper contributes to the normality literature. It provides, in particular, new conditions for normality, covering cases when the initial state is located in a 'corner' of intersection of the left end-point and state constraint sets. Indeed, the results of this paper provide improvements on earlier demonstrated conditions for optimality, when the state constraint set A is non-smooth, in this respect: we allow the left end-point constraint set E_0 to be a possibly strict subset of $\mathbb{R}^n$, while in [4] E_0 was required to be the whole space $\mathbb{R}^n$. The conditions given here are also less restrictive than earlier conditions in [15], [16] for related cases, as is illustrated by an example. Furthermore, normality is established in relation to the extended Euler-Lagrange condition for optimal control problems for which the dynamic constraint is formulated as a differential inclusion, and not as a control-parameterized differential equation as in [15], [16].

Notation. We write $\mathbb{B}$ the closed unit ball in Euclidean space; and $|\ .\ |$ is the Euclidean norm. Given a non-empty set $Y \subset \mathbb{R}^n$, we write the interior, the boundary and the convex hull of Y respectively $\operatorname{int} Y$, ∂Y and $\operatorname{co} Y$. The Euclidean distance of the point x from the set Y (that is $\min_{y\in Y} |x-y|$) is written $d_Y(x)$. Y^* denotes the polar cone of Y, namely

$$Y^* := \{\xi \in \mathbb{R}^n \ :\ \xi \cdot y \leq 0 \text{ for all } y \in Y\}\,.$$

Given a closed non-empty set $E \subset \mathbb{R}^n$, the *limiting normal cone* $N_E(\bar{x})$ of E at $\bar{x} \in E$ and the *Clarke tangent cone* $T^C_E(\bar{x})$ of E at $\bar{x}$ are defined as follows

$$N_E(\bar{x}) := \left\{ \xi \mid \exists\, x_i \xrightarrow{E} \bar{x},\ \xi_i \longrightarrow \xi \text{ s.t. } \limsup_{x \xrightarrow{E} x_i} \frac{\xi_i \cdot (x - x_i)}{|x - x_i|} \leq 0 \text{ for all } i \in \mathbb{N} \right\}$$

and

$$T_E^C(\bar{x}) := \left\{ \xi \mid \forall\, x_i \xrightarrow{E} \bar{x} \text{ and } t_i \downarrow 0,\ \exists\, \{e_i\} \subset E \text{ s.t. } t_i^{-1}(e_i - x_i) \longrightarrow \xi \right\}.$$

The notation $x_i \xrightarrow{E} \bar{x}$ means that $x_i \to \bar{x}$ and $x_i \in E$ for all $i \in \mathbb{N}$. We recall that the Clarke tangent cone $T_E^C(\bar{x})$ and the limiting normal cone $N_E(\bar{x})$ are related according to the following polarity relation

$$(N_E(\bar{x}))^* = T_E^C(\bar{x}).$$

Whenever the closed set E is also convex, the (limiting) normal and the tangent cones above coincide with the normal and the tangent cones of Convex Analysis (and in addition we have $N_E(\bar{x}) = (T_E^C(\bar{x}))^*$). In this case, we denote by $\pi_E(x)$ the unique projection of the point $x \in \mathbb{R}^n$ onto E. The graph of a multifunction $F(.) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ is the set

$$\mathrm{Gr}\{F(.)\} := \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in \mathbb{R}^n \text{ and } v \in F(x)\}.$$

We recall that the subdifferential of a lower semi-continuous extended valued function φ at a point $x \in \operatorname{dom} \varphi := \{y \in \mathbb{R}^n \ :\ \varphi(y) < +\infty\}$ can be expressed in terms of the limiting normal cone of the epigraph of φ:

$$\partial\varphi(x) = \{\eta \,|\, (\eta, -1) \in N_{\operatorname{epi}\varphi}(x, \varphi(x))\}.$$

(For further details and discussion on these analytical tools we refer the reader to the books [2], [9] and [22] and the references therein.)

$C([a, b]; \mathbb{R}^n)$ is the set of continuous $\mathbb{R}^n$ valued functions on the time interval $[a, b]$. We denote $W^{1,1}([a, b]; \mathbb{R}^n)$ for the set of absolutely continuous $\mathbb{R}^n$ valued functions on $[a, b]$; we shall employ the following norm on $W^{1,1}([a, b]; \mathbb{R}^n)$:

$$\|x(.)\|_{W^{1,1}(a,b)} := \|x(.)\|_{L^\infty(a,b)} + \|\dot{x}(.)\|_{L^1(a,b)}\,, \quad \text{for all } x(.) \in W^{1,1}([a, b]; \mathbb{R}^n),$$

which is equivalent to the classical $W^{1,1}$-norm: $\|x(.)\| = \|x(.)\|_{L^1(a,b)} + \|\dot{x}(.)\|_{L^1(a,b)}$.

2. Main results

We write A_0 the set of all points $y \in A$ where the state constraint A is locally regular: more precisely, $y \in A_0$ if and only if there exists a radius $r > 0$ and a function $\varphi : \mathbb{R}^n \to \mathbb{R}$ of class C^{1+} (i.e. everywhere differentiable with locally Lipschitz continuous derivatives) such that

$$\nabla\varphi(y) \neq 0\,, \qquad A \cap (y + r \operatorname{int} \mathbb{B}) = \{x \ :\ \varphi(x) \leq 0\} \cap (y + r \operatorname{int} \mathbb{B})\,.$$

Observe that $A_0 \supset \operatorname{int} A$.

Theorem 2.1. *Let $\bar{x}(.)$ be a $W^{1,1}$ δ-local minimizer for problem (P), in which we assume that for some constants $c > 0$ and $k_F > 0$ the following hypotheses are satisfied:*

$(H1)$ *The set $A \subset \mathbb{R}^n$ is closed and convex.*

$(H2)$ (a) *The multifunction F has non-empty, closed values;*

(b)

$$F(x) \subset c\,\mathbb{B} \quad \textit{for all } x \in \bar{x}(t) + \delta\mathbb{B},\ t \in [S,T];$$

(c)

$$F(x) \subset F(y) + k_F|x-y|\mathbb{B}, \quad \textit{for all } x,y \in \bar{x}(t) + \delta\mathbb{B},\ t \in [S,T].$$

$(H3)$

$$\operatorname{co} F(x) \cap \operatorname{int} T_A^C(x) \neq \emptyset, \quad \textit{for all } x \in A \cap (\bar{x}(t) + \delta\mathbb{B}),\ t \in [S,T].$$

$(H4)$ (a) *g is Lipschitz continuous on $(\bar{x}(S) + \delta\mathbb{B}) \times (\bar{x}(T) + \delta\mathbb{B})$;*

(b) *$L(t,x,u)$ is measurable in t for fixed (x,u), and for some constants $k_L > 0$ we have*

$$|L(t,x,u) - L(t,x',u')| \leq k_L|(x,u)-(x',u')| \\ \textit{for all } (x,u),(x',u') \in (\bar{x}(t) + \delta\mathbb{B}) \times \mathbb{R}^n,$$

for all $t \in [S,T]$.

$(H5)$ *One of the following two conditions is satisfied:*

(a) *$\bar{x}(S) \in A_0$;*

(b) $\operatorname{int} T^C_{E_0\cap A}(\bar{x}(S)) \neq \emptyset$.

Then, there exist $p(.) \in W^{1,1}([S,T];\mathbb{R}^n)$ and a function of bounded variation $\nu(.) : [S,T] \to \mathbb{R}^n$, continuous from the right on (S,T), such that

(i) $\int_{[S,T]} \xi(t) \cdot d\nu(t) \leq 0$ *for all $\xi(.) \in C([S,T];\mathbb{R}^n)$ satisfying $\xi(t) \in T_A^C(\bar{x}(t))$,*

(ii) $\dot{p}(t) \in \operatorname{co}\{\eta \,|\, (\eta, q(t)) \in \partial L(t,\bar{x}(t),\dot{\bar{x}}(t)) + N_{\operatorname{Gr}\{F(.)\}}(\bar{x}(t),\dot{\bar{x}}(t))\}$ *a.e. $t \in [S,T]$,*

(iii) $(p(S), -q(T)) \in \partial g(\bar{x}(S),\bar{x}(T)) + N_{E_0\cap A}(\bar{x}(S)) \times \{0\}$,

(iv) $q(t)\cdot\dot{\bar{x}}(t) - L(t,\bar{x}(t),\dot{\bar{x}}(t)) = \max_{v\in F(\bar{x}(t))}\{q(t)\cdot v - L(t,\bar{x}(t),v)\}$ *a.e. $t \in [S,T]$,*

in which $q(.) : [S,T] \to \mathbb{R}^n$ is the function

$$q(t) := p(t) + \int_{[S,t]} d\nu(s) \quad \textit{for } t \in (S,T]\ .$$

Theorem 2.1 will be proved in Section 4.

Remark 2.2. (a) Very similar proof techniques to those employed here yield an alternative version of Theorem 2.1, which provides necessary conditions in normal form in cases when the state constraint set A is expressed in terms of a functional inequality, that is

$$A = \{x \in \mathbb{R}^n \ :\ h(x) \leq 0\},$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function, and when the necessary conditions are expressed directly in terms of the inequality state constraint function $h(.)$. More precisely, if assumptions $(H1)$–$(H5)$ of Theorem 2.1 are satisfied, then the (normal) extended Euler-Lagrange condition can be replaced by: there exist a costate arc $p(.) \in W^{1,1}([S,T];\mathbb{R}^n)$ associated with the state $\bar{x}(.)$, a Borel measure $\mu(.)$ on $[S,T]$ and a μ-integrable function $\gamma : [S,T] \to \mathbb{R}^n$, such that

(ia) $\gamma(t) \in \partial_x^{>} h(\bar{x}(t))$ μ-a.e., and $\mathrm{supp}\{\mu\} \subset \{t \ : \ h(\bar{x}(t)) = 0\}$,

(iia) $\dot{p}(t) \in \mathrm{co}\{\eta \,|\, (\eta, q(t)) \in \partial L(t, \bar{x}(t), \dot{\bar{x}}(t)) + N_{\mathrm{Gr}\{F(.)\}}(\bar{x}(t), \dot{\bar{x}}(t))\}$ a.e. $t \in [S,T]$,

(iiia) $(p(S), -q(T)) \in \partial g(\bar{x}(S), \bar{x}(T)) + N_{E_0 \cap A}(\bar{x}(S)) \times \{0\}$,

(iva) $q(t) \cdot \dot{\bar{x}}(t) - L(t, \bar{x}(t), \dot{\bar{x}}(t)) = \max_{v \in F(\bar{x}(t))} \{q(t) \cdot v - L(t, \bar{x}(t), v)\}$ a.e. $t \in [S,T]$,

in which $q(.) : [S,T] \to \mathbb{R}^n$ is the function

$$q(t) := p(t) + \int_{[S,t]} \gamma(s) d\mu(s) \quad \text{for } t \in (S,T].$$

Here, $\mathrm{supp}\,\mu$ denotes the support of the measure μ and $\partial_x^{>} h(x)$ is the hybrid subdifferential of h at x, which is defined as:

$$\begin{aligned} \partial_x^{>} h(x) := \ & \mathrm{co}\{\eta \,|\, \text{there exist } x_i \to x \text{ and } \eta_i \to \eta \text{ s.t.} \\ & \qquad \nabla_x h(x_i) \text{ exists}, \ \eta_i = \nabla_x h(x_i) \text{ and } h(x_i) > 0 \text{ for each } i \in \mathbb{N}\}. \end{aligned}$$

(b) The special case of Thm. 2.1, in which $(H5)(b)$ is replaced by the requirement '$E_0 = \mathbb{R}^n$', was recently proved in [4]. (Notice that, when $E_0 = \mathbb{R}^n$, $(H5)(b)$ is automatically satisfied, in consequence of $(H3)$.) The earlier literature provides conditions for normality also for the particular case in which '$\bar{x}(S) \in \mathrm{int}\, A$' (this implies that hypothesis $(H5)(a)$ of Thm. 2.1 is verified): see [4] for normality of the extended Euler-Lagrange condition in the context of state constrained differential inclusion problems; and cf. [5] and [15] for minimizers satisfying the maximum principle (for optimal control problems in which the dynamic constraint takes the form of a controlled differential equation).

(c) Assumption $(H5)(b)$ concerns the case in which $\bar{x}(S)$ belongs to a region where the state constraint A is not regular, and it involves tangent cones of the state and the left end-point constraints. An immediate consequence of $(H5)(b)$ is that the generalized Euler-Lagrange conditions necessarily apply in a non-degenerate form (in the sense of [20]); however, here, we are interested in the stronger property of normality. Recent papers make use of similar conditions (comprising the intersection of tangent cones to the constraints) to deal with the case in which $\bar{x}(S)$ turns out to be a 'corner' of A. In particular, [16] provides normality for a generalized version of the maximum principle, applicable for a wide class of optimal control problems, in which the state constraint is a closed set with non-empty interior tangent cones. Theorem 2.1 differs in the nature of the necessary conditions (here including generalized Euler-Lagrange conditions), and also in the nature of the inward pointing condition. Indeed, the constraint qualification in [16] takes into account also inward vectors which are tangents to the velocity set, whereas here $(H3)$ comprises only inward pointing vectors belonging only to the velocity set. The example at the end of the paper illustrates a typical case in which $(H3)$ holds true, $(H5)$ might be verified but the inward pointing condition of [16] is not satisfied.

3. $W^{1,1}$-linear estimates

We shall make use of the following theorem concerning existence of feasible F-trajectories, which is a refinement of a result in [4].

Theorem 3.1. *Fix $r_0 > 0$. Consider a non-empty set $A \subset \mathbb{R}^n$ and a multifunction $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ such that, for some constants $c > 0$ and $k_F > 0$, and for $R_0 := e^{c(T-S)}(r_0 + 2)$, the following assumptions are satisfied:*

$(H1)$ *The set $A \subset \mathbb{R}^n$ is closed and convex.*

$(H2)'$ (a) *The multifunction F has non-empty, closed values;*

(b)

$$F(x) \subset c(1 + |x|)\, \mathbb{B} \quad \text{for all } x \in \mathbb{R}^n.$$

(c)

$$F(x) \subset F(y) + k_F|x - y|\mathbb{B}, \quad \text{for all } x, y \in R_0\mathbb{B};$$

$(H3)'$

$$\operatorname{co} F(x) \cap \operatorname{int} T_A^C(x) \neq \emptyset, \quad \text{for all } x \in A \cap R_0\mathbb{B}.$$

Let $x_0 \in A \cap r_0\mathbb{B}$ be a given point. Then, the following properties hold true.

(i) *If $x_0 \in A_0$, then there exist $\theta \in (0, \frac{1}{4})$ and $K > 0$ such that for any F-trajectory $\hat{x}(.)$ on $[S, T]$ and $\varepsilon \geq 0$ satisfying*

$$\begin{cases} \max\limits_{t \in [S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in A \cap (x_0 + \theta\mathbb{B}), \end{cases}$$

we can find a feasible F-trajectory $x(.)$ on $[S, T]$ satisfying

$$\begin{cases} \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon \\ x(S) = \hat{x}(S). \end{cases}$$

(ii) *If $x_0 \in A \setminus A_0$, then, for any $w_0 \in \operatorname{int} T_A^C(x_0)$ fixed, there exist $\theta \in (0, \frac{1}{4})$, $\alpha \geq 0$ and $K > 0$ such that for any F-trajectory $\hat{x}(.)$ on $[S, T]$ and $\varepsilon \geq 0$ with*

$$\begin{cases} \max\limits_{t \in [S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in A \cap (x_0 + \theta\mathbb{B}), \end{cases} \tag{1}$$

we can find a feasible F-trajectory $x(.)$ on $[S, T]$ satisfying

$$\begin{cases} \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon \\ x(S) = \hat{x}(S) + \alpha\varepsilon w_0. \end{cases} \tag{2}$$

The proof of Theorem 3.1 will be provided in Section 5.

A consequence of Theorem 3.1 is the following result, in which the left end-point $\hat{x}(S)$ of the possibly 'violating' F-trajectory $\hat{x}(.)$ does not necessarily belong to the state constraint.

Corollary 3.2. *Suppose that all the assumptions of Theorem* 3.1 *are satisfied (for some constants $c > 0$ and $k_F > 0$, and for $R_0 := e^{c(T-S)}(r_0 + 2)$). Let $x_0 \in A \cap r_0\mathbb{B}$ be a given point. Then, the following properties hold true.*

(*i*) *If $x_0 \in A_0$, then there exist $\theta' > 0$ and $K' > 0$ such that for any F-trajectory $\hat{x}(.)$ on $[S,T]$ and $\varepsilon \geq 0$ satisfying*

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in x_0 + \theta'\mathbb{B}, \end{cases}$$

and for any given $y_0 \in A \cap (\hat{x}(S) + \varepsilon\mathbb{B})$, we can find a feasible F-trajectory $x(.)$ on $[S,T]$ satisfying

$$\begin{cases} \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K'\varepsilon \\ x(S) = y_0\,. \end{cases}$$

(*ii*) *If $x_0 \in A \setminus A_0$, then, for any $w_0 \in \operatorname{int} T_A^C(x_0)$ fixed, there exist $\alpha' \geq 0$, $\theta' > 0$ and $K' > 0$ such that for any F-trajectory $\hat{x}(.)$ on $[S,T]$ and $\varepsilon \geq 0$ with*

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in x_0 + \theta'\mathbb{B}, \end{cases} \tag{3}$$

and for any given $y_0 \in A \cap (\hat{x}(S) + \varepsilon\mathbb{B})$, we can find a feasible F-trajectory $x(.)$ on $[S,T]$ satisfying

$$\begin{cases} \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K'\varepsilon \\ x(S) = y_0 + \alpha'\varepsilon w_0\,. \end{cases} \tag{4}$$

Proof. We prove here just the case $x_0 \in A \setminus A_0$. The case $x_0 \in A_0$ is easier and can be similarly treated.
Fix any $x_0 \in A \setminus A_0$ and $w_0 \in \operatorname{int} T_A^C(x_0)$. Consider the constants $\alpha \geq 0$, $\theta > 0$ and $K > 0$ provided by Theorem 3.1 (case (*ii*)). Define $\theta' := \frac{\theta}{4}$. Take any $F-$trajectory $\hat{x}(.)$ and any number $\varepsilon \geq 0$ such that:

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in x_0 + \theta'\mathbb{B}. \end{cases}$$

It is not restrictive to assume that $\varepsilon \leq \theta'$. Consider any $y_0 \in A \cap (\hat{x}(S) + \varepsilon\mathbb{B})$. Observe that $y_0 \in A \cap (x_0 + \theta\mathbb{B})$. It is obvious that $|y_0 - \hat{x}(S)| \leq \varepsilon$. Then invoking the Filippov Existence Theorem (cf. [2] or [22]) to the reference trajectory $\hat{x}(.)$ on $[S,T]$, we obtain an F-trajectory $y(.)$ on $[S,T]$ such that $y(S)=y_0$ and

$$\|y(.)-\hat{x}(.)\|_{W^{1,1}(S,T)} \leq e^{k_F(T-S)}|y(S)-\hat{x}(S)| \leq e^{k_F(T-S)}\varepsilon\,. \tag{5}$$

Notice that

$$\max_{t\in[S,T]} d_A(y(t)) \leq \|y(.)-\hat{x}(.)\|_{L^\infty(S,T)} + \max_{t\in[S,T]} d_A(\hat{x}(t)) \leq (e^{k_F(T-S)}+1)\varepsilon\,. \tag{6}$$

We write $\varepsilon' := (e^{k_F(T-S)}+1)\varepsilon$.
Theorem 3.1 is also applicable for the reference trajectory $y(.)$ and number ε': we obtain a feasible F-trajectory $x(.)$ such that

$$x(S) = y(S) + \alpha\varepsilon' w_0 = y_0 + \alpha\varepsilon' w_0 = y_0 + \alpha(e^{k_F(T-S)}+1)\varepsilon w_0$$

and

$$\|x(.)-y(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon'. \tag{7}$$

As a consequence, from (5), (6) and (7), we have

$$\begin{aligned}
\|x(.)-\hat{x}(.)\|_{W^{1,1}(S,T)} &\leq \|x(.)-y(.)\|_{W^{1,1}(S,T)} + \|y(.)-\hat{x}(.)\|_{W^{1,1}(S,T)}\\
&\leq K\varepsilon' + e^{k_F(T-S)}\varepsilon\\
&\leq (K(e^{k_F(T-S)}+1) + e^{k_F(T-S)})\varepsilon
\end{aligned}$$

and the validity of (4) follows taking the constants $\theta' := \frac{\theta}{4}$, $\alpha' := \alpha(e^{k_F(T-S)}+1)$ and $K' := K(e^{k_F(T-S)}+1) + e^{k_F(T-S)}$. □

We consider now a version of Theorem 3.1 in which the assumptions are satisfied only in a neighborhood of a reference feasible F-trajectory $\bar{x}(.)$.

Theorem 3.3. *Consider a set $A\subset\mathbb{R}^n$ and a multifunction $F:\mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ such that, for some constants $\delta>0$, $c>0$ and $k_F>0$, and for a given feasible F-trajectory $\bar{x}(.)$, hypotheses (H1)–(H3) of Theorem 2.1 are satisfied. Then, the following properties hold true.*

(i) *If $\bar{x}(S)\in A_0$, then there exist $\theta\in(0,\delta)$ and $K>0$ such that for any F-trajectory $\hat{x}(.)$ on $[S,T]$ and $\varepsilon\geq 0$ with*

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \|\bar{x}(.)-\hat{x}(.)\|_{W^{1,1}(S,T)} \leq \theta, \end{cases}$$

and for any given $y_0\in A\cap(\hat{x}(S)+\varepsilon\mathbb{B})$, we can find a feasible F-trajectory $x(.)$ on $[S,T]$ satisfying

$$\begin{cases} \|x(.)-\hat{x}(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon \\ x(S) = y_0\,. \end{cases}$$

(*ii*) *If $\bar{x}(S) \in A \setminus A_0$, then, for any $w_0 \in \operatorname{int} T_A^C(\bar{x}(S))$ fixed, there exist $\theta \in (0,\delta)$, $\alpha \geq 0$ and $K > 0$ such that for any F-trajectory $\hat{x}(.)$ on $[S,T]$ and $\varepsilon \geq 0$ with*

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \|\bar{x}(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq \theta, \end{cases}$$

and for any given $y_0 \in A \cap (\hat{x}(S) + \varepsilon\mathbb{B})$, we can find a feasible F-trajectory $x(.)$ on $[S,T]$ satisfying

$$\begin{cases} \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon \\ x(S) = y_0 + \alpha\varepsilon w_0\,. \end{cases}$$

4. Proof of Theorem 2.1

We provide here a proof supposing that $(H1)$–$(H4)$ and $(H5)(b)$ are satisfied. The proof when we assume $(H5)(a)$ instead of $(H5)(b)$ is simpler and has been already treated in [4] to which we refer the reader for a proof.

Step 1. Let $\bar{x}(.)$ be a $W^{1,1}$ δ-local minimizer for the optimal problem (P), for some $\delta > 0$. It is not restrictive to assume that $\delta \in (0, \frac{1}{2})$. Assumption $(H5)(b)$ guarantees the existence of a vector $w_0 \in \mathbb{R}^n$ such that

$$w_0 \in \operatorname{int} T^C_{E_0\cap A}(\bar{x}(S)).$$

Since A is convex, it follows that:

$$T^C_{E_0\cap A}(\bar{x}(S)) \subset T^C_A(\bar{x}(S)),$$

And therefore $w_0 \in \operatorname{int} T^C_A(\bar{x}(S))$, which means that all the assumptions of case (ii) of Theorem 3.3 are satisfied. From the characterization of the interior of the Clarke tangent cone (cf. [21]), there exists $\bar{\delta} \in (0,1)$ such that

$$y + [0,\bar{\delta}](w_0 + \bar{\delta}\mathbb{B}) \subset E_0 \cap A \quad \text{for all } y \in (\bar{x}(S) + 2\bar{\delta}\mathbb{B}) \cap (E_0 \cap A). \tag{8}$$

We take

$$\delta' = \min\left\{\theta;\ \frac{\delta}{4(K+1)};\ \frac{\bar{\delta}}{\alpha+1}\right\},$$

where α, θ and K are the positive constants provided by Theorem 3.3.

Step 2. We shall define the following sets of functions

$$\begin{aligned} \mathcal{W} := \{x(.) \in W^{1,1}([S,T],\mathbb{R}^n) : \dot{x}(t) \in F(x(t)) \text{ a.e. } x(S) \in E_0 \cap A,\\ \text{and } \|x(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \delta\}, \\ \mathcal{W}' := \{x(.) \in \mathcal{W} : \|x(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \delta'\}. \end{aligned}$$

From assumption $(H4)$ it immediately follows that the functional $J(.)$ is uniformly Lipschitz on $\mathcal{W}$ with respect to the $W^{1,1}$-norm. Write $K_J > 0$ the Lipschitz constant of $J(.)$:

$$|J(y(.)) - J(x(.))| \leq K_J \, \|y(.) - x(.)\|_{W^{1,1}(S,T)} \quad \text{for all } y(.), x(.) \in \mathcal{W}. \tag{9}$$

Combining a penalization method with the linear estimate provided by Theorem 3.3, we obtain that $\bar{x}(.)$ is also a $W^{1,1}$ local minimizer for a new optimal control problem, in which the state constraint in (P) is replaced by an extra penalty term in the cost.

Lemma 4.1. *Assume that all hypotheses of Theorem 2.1 are satisfied. Then, $\bar{x}(.)$ is a $W^{1,1}$ local minimizer for the problem*

$$\begin{cases} \textit{minimize } \tilde{J}(y(.)) = J(y(.)) + K_J K \max_{t\in[S,T]} d_A(y(t)) \\ \textit{over } y(.) \in \mathcal{W}'. \end{cases}$$

Proof. Take any $\hat{x}(.) \in \mathcal{W}'$. Notice that $\hat{x}(.)$ is an F-trajectory such that $\hat{x}(S) \in A$ and $\|\hat{x}(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \theta$. Write $\varepsilon := \max_{t\in[S,T]} d_A(\hat{x}(t))$. Then invoking Theorem 3.3, we can find an F-trajectory $x(.)$ such that

$$\begin{cases} x(t) \in A \quad \text{for all } t \in [S,T], \\ \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K\varepsilon \\ \text{and } x(S) = \hat{x}(S) + \alpha\varepsilon w_0 \,. \end{cases} \tag{10}$$

Using the fact that

$$\varepsilon = \max_{t\in[S,T]} d_A(\hat{x}(t)) \leq \|\hat{x}(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \delta' \tag{11}$$

and the linear estimate (10) we obtain

$$\begin{aligned} &\|x(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \\ \leq\; &\|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} + \|\hat{x}(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \\ \leq\; &K\varepsilon + \|\hat{x}(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq (K+1)\, \|\hat{x}(.) - \bar{x}(.)\|_{W^{1,1}(S,T)} \leq \delta \,. \end{aligned} \tag{12}$$

We prove that $x(S) \in E_0 \cap A$. Indeed, since $\hat{x}(.) \in \mathcal{W}'$, from the choice of δ' we also have $\hat{x}(S) \in (\bar{x}(S) + \bar{\delta}\mathbb{B}) \cap (E_0 \cap A)$. On the other hand from (11) it follows that $\alpha\varepsilon \leq \alpha\delta' \leq \bar{\delta}$. Therefore, we can apply inclusion (8), obtaining that $x(S) = \hat{x}(S) + \alpha\varepsilon w_0 \in E_0 \cap A$. As a consequence $x(.) \in \mathcal{W}$.

Now, recalling that $K_J > 0$ is the Lipschitz constant of $J(.)$ and that $\bar{x}(.)$ is a $W^{1,1}$ minimizer for the reference problem (P) (and in particular $\max_{t\in[S,T]} d_A(\bar{x}(t))$ $= 0$ for it is feasible), for any $\hat{x}(.) \in \mathcal{W}'$ we obtain

$$\begin{aligned} \tilde{J}(\bar{x}(.)) = J(\bar{x}(.)) \leq J(x(.)) &\leq J(\hat{x}(.)) + K_J \, \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \\ &\leq J(\hat{x}(.)) + K_J K\varepsilon = J(\hat{x}(.)) + K_J K \max_{t\in[S,T]} d_A(\hat{x}(t)) = \tilde{J}(\hat{x}(.)). \end{aligned}$$

This confirms the lemma statement. $\square$

Step 3. Consider the auxiliary state constrained optimal control problem

$$(P)' \begin{cases} \text{Minimize } \tilde{g}(X(S), X(T)) \\ \text{over } X(.) = (x(.), y(.), z(.)) \in W^{1,1}([S,T]; \mathbb{R}^{n+2}) \text{ satisfying} \\ \dot{X}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t)) \in G(t, X(t)) \text{ a.e. } t \in [S,T], \\ \tilde{h}(X(t)) \leq 0 \text{ for all } t \in [S,T], \\ (x(S), y(S), z(S)) \in (E_0 \cap A) \times \mathbb{R}_+ \times \{0\}. \end{cases}$$

The cost function $\tilde{g}$ is defined by

$$\tilde{g}(X(S), X(T)) := g(x(S), x(T)) + K_J K \max\{0, y(T)\} + z(T).$$

The multivalued function is defined by

$$G(t, X) := \{(v, 0, r) : v \in F(x), r \geq L(t, x, v)\},$$

and the function $\tilde{h} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ (which provides the state constraint by means of a functional inequality) is defined by:

$$\tilde{h}(X) = \tilde{h}(x, y, z) := d_A(x) - y.$$

We added two extra state variables y and z, and applying known necessary optimality conditions to a minimizer for $(P)'$, we shall derive normality for the reference problem (P).
Define the arc

$$\bar{z}(t) := \int_S^t L(s, \bar{x}(s), \dot{\bar{x}}(s))ds, \quad \text{for all } t \in [S,T].$$

Observe that Lemma 4.1 implies that the G-trajectory

$$\bar{X}(.) := (\bar{x}(.), \bar{y} \equiv 0, \bar{z}(.)),$$

is a $W^{1,1}$ $\delta'-$local minimizer for problem $(P)'$.
Indeed, suppose by contradiction, that there exists a state trajectory $X(.) = (x(.), y(.), z(.))$ satisfying the state and dynamic constraints of $(P)'$, such that

$$\|(x(.), y(.), z(.)) - (\bar{x}(.), \bar{y}(.), \bar{z}(.))\|_{W^{1,1}(S,T)} \leq \delta' \ ,$$

and $\tilde{g}(X(S), X(T)) < \tilde{g}(\bar{X}(S), \bar{X}(T))$.
Observe that $y(.) \equiv y \geq 0$, and condition

$$\tilde{h}(x(t), y, z(t)) \leq 0 \ \text{ for all } t \in [S,T]$$

can be equivalently written as

$$\max_{t \in [S,T]} d_A(x(t)) \leq y.$$

Then we would obtain:

$$\begin{aligned}\tilde{J}(x(.)) &\leq g(x(S),x(T)) + K_J K \max\{0,y\} + \int_S^T L(t,x(t),\dot{x}(t))ds \\ &\leq \tilde{g}(X(S),X(T)) < \tilde{g}(\bar{X}(S),\bar{X}(T)) = \tilde{J}(\bar{x}(.)).\end{aligned}$$

This contradicts the fact that $\bar{x}(.)$ is a $W^{1,1}$ minimizer over $\mathcal{W}'$.

Step 4. Known necessary conditions for optimality are now applicable. In particular [22, Theorem 10.3.1] implies that there exist a Lagrange multiplier $\lambda \geq 0$, a costate arc $P(.) = (p(.), \psi(.), \phi(.)) \in W^{1,1}([S,T], \mathbb{R}^{n+2})$ (associated with the minimizer $\bar{X}(.) = (\bar{x}(.), \bar{y} \equiv 0, \bar{z}(.)))$, a Borel measurable function $\mu(.)$ on $[S,T]$ and a $\mu-$integrable function $\gamma(.) : [S,T] \to \mathbb{R}^{n+2}$ satisfying

(a) $\|P\|_{L^\infty} + \lambda + \int_S^T d\mu(s) = 1$,

(b) $\dot{P}(t) \in \text{co}\{w \in \mathbb{R}^{n+2} | (w, Q(t)) \in N_{Gr\{G(t,.)\}}((\bar{x}(t), \bar{y} \equiv 0, \bar{z}(t)), (\dot{\bar{x}}(t), 0, \dot{\bar{z}}(t)))\}$ a.e. $t \in [S,T]$,

(c) $(P(S), -Q(T)) \in \lambda \partial \tilde{g}((\bar{x}(S), \bar{y} \equiv 0, \bar{z}(S)), (\bar{x}(T), \bar{y} \equiv 0, \bar{z}(T))) + N_{E_0 \cap A}(\bar{x}(S)) \times \mathbb{R}_- \times \mathbb{R} \times \{(0,0,0)\}$,

(d) $Q(t) \cdot (\dot{\bar{x}}(t), 0, \dot{\bar{z}}(t)) = \max_{\tilde{v} \in G(t,(\bar{x}(t),\bar{y}\equiv 0,\bar{z}(t)))} Q(t) \cdot \tilde{v}$ a.e. $t \in [S,T]$,

(e) $\gamma(t) \in \partial^> \tilde{h}(\bar{x}(t), \bar{y} \equiv 0, \bar{z}(t))$ $\mu-$ a.e., and $\text{supp}\{\mu\} \subset \{t : \tilde{h}(\bar{x}(t), \bar{y} \equiv 0, \bar{z}(t)) = 0\}$,

in which $Q(.) : [S,T] \to \mathbb{R}^{n+2}$ is the function

$$Q(t) := P(t) + \int_{[S,t]} \gamma(s) d\mu(s) \quad \text{for } t \in (S,T].$$

Step 5. One can easily derive from relations (a)–(e) given above for the auxiliary problem $(P)'$, the desired necessary conditions for optimality for the reference problem (P):

$$(p(S), -q(T)) \in \lambda \partial g(\bar{x}(S), \bar{x}(T)) + N_{E_0 \cap A}(\bar{x}(S)) \times \{0\},$$

$$\dot{\psi} \equiv 0, \quad \psi := \psi(S) \leq 0, \quad \text{and} \quad -\psi(T) + \int_{[S,T]} d\mu(s) \in [0, \lambda K_J K],$$

$$\dot{\phi} \equiv 0 \quad \text{and} \quad \phi(T) = -\lambda,$$

$$\dot{p}(t) \in \text{co}\{\eta \mid (\eta, q(t)) \in \lambda \partial L(t, \bar{x}(t), \dot{\bar{x}}(t)) + N_{Gr\{F(.)\}}(\bar{x}(t), \dot{\bar{x}}(t))\} \ \text{ a.e. } t \in [S,T].$$

Moreover, the necessary conditions apply in the normal form, for if $\lambda = 0$, then we would obtain

$$-\psi + \int_{[S,T]} d\mu(s) = 0, \quad p \equiv 0, \quad \psi \equiv 0 \quad \text{and} \quad \phi \equiv 0,$$

But this contradicts (a) of Step 4. Therefore, relations (a)–(d) above yield conditions (ii)–(iv) of Theorem 2.1. Finally, since the function $\tilde{h}$ is defined using the distance function to A, $d_A(.)$, standard analysis (cf. [22] pages 332 and 370) allows to conclude the proof, confirming also property (i) of the theorem. □

5. Proofs of Theorems 3.1 and 3.3

Proof of Theorem 3.1. We build up the proof of Theorem 3.1 in three steps. In the first step we provide upper bounds for the magnitude of state trajectories and their velocities having initial state on a bounded region. This allows us to use a standard compactness argument, and consequently apply a local analysis. In the second step, we treat separately the cases in which $x_0 \in A_0$ and $x_0 \in A \setminus A_0$. We shall provide two lemmas giving local $W^{1,1}-$estimate results for an initial subinterval. Step 3 is devoted to a recursive argument ensuring the construction of the desired feasible $F-$trajectories.

Step 1. Fix any $r_0 > 0$ and $x_0 \in A \cap r_0\mathbb{B}$. Write $R_0 := e^{c(T-S)}(r_0+2)$ and $M_0 := c(1+R_0)$. Observe that assumption $(H2)'$ ensures the following a-priori bounds for the magnitude of all F-trajectories $x(.)$ starting from $(r_0+1)\mathbb{B}$ and their velocities: if $x(.)$ is an F-trajectory on $[S,T]$ with $x(S) \in (r_0+1)\mathbb{B}$, then

$$x(t) \in R_0\mathbb{B} \quad \text{for all } t \in [S,T]$$

and

$$\dot{x}(t) \in M_0\mathbb{B} \quad \text{a.e. } t \in [S,T].$$

Using the characterization of the interior of the Clarke tangent cone (cf. [9] and [21]) and applying a standard compactness argument, we can find a finite number of points $y_i \in \partial A \cap R_0\mathbb{B}$ and numbers $\bar{\varepsilon} > 0,\ \delta_i \in (0,1),\ \rho_i \in (0, \frac{1}{4})$, for $i = 0, \cdots, N$, such that

$$(\partial A + \bar{\varepsilon}\mathbb{B}) \cap R_0\mathbb{B} \subset \bigcup_{i=0}^{N} \left(y_i + \frac{\delta_i}{4}\operatorname{int}\mathbb{B} \right),$$

and for all $i = 0, \cdots, N$, there exists a vector $v_i \in \operatorname{co} F(y_i) \cap \operatorname{int} T_A^C(y_i)$ such that

$$y + [0, 2\delta_i](v_i + \rho_i\mathbb{B}) \subset A, \quad \text{for all } y \in (y_i + \delta_i\mathbb{B}) \cap A.$$

Step 2. We treat the two cases (i) $(x_0 \in A_0)$ and (ii) $(x_0 \in A \setminus A_0)$ separately to obtain a local $W^{1,1}$-estimate.

Lemma 5.1 (Case $x_0 \in A_0$). *Fix any $r_0 > 0$. Consider a multifunction $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and a closed non-empty set A. Assume that assumptions $(H2)'$ and $(H3)'$ of Theorem 3.1 are satisfied (for some constants $c > 0$, $k_F > 0$ and $R_0 = e^{c(T-S)}(r_0+2)$). Take any $x_0 \in A_0 \cap r_0\mathbb{B}$. Then, we can find $\theta_0 \in (0, \frac{1}{4})$, $\tau_0 \in (0, T-S]$ and $K_0 > 0$ such that the following property is satisfied: given any F-trajectory $\hat{x}(.) : [S,T] \to \mathbb{R}^n$ and $\varepsilon \geq 0$ such that*

$$\begin{cases} \max\limits_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in A \cap (x_0 + \theta_0\mathbb{B}), \end{cases}$$

we can find an F-trajectory $x(.)$ on $[S,T]$ such that:

$$\begin{cases} x(S) = \hat{x}(S), \\ x(t) \in A \text{ for all } t \in [S, S+\tau_0], \\ \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K_0\varepsilon. \end{cases}$$

Proof. Since $x_0 \in A_0$, there exist a C^{1+} function $h_0 : \mathbb{R}^n \to \mathbb{R}$ and a constant $\epsilon_0 \in (0,1)$ such that we have:

$$\nabla h_0(x_0) \neq 0, \qquad A \cap (x_0 + \epsilon_0 \operatorname{int} \mathbb{B}) = \{x \ : \ h_0(x) \leq 0\} \cap (x_0 + \epsilon_0 \operatorname{int} \mathbb{B}) \subset A_0 \, .$$

Take $\theta_0 := \frac{\epsilon_0}{4}$ and $\tau_0 := \frac{\epsilon_0}{2M_0}$, where $M_0 > 0$ is the constant provided in Step 1 of the proof of Theorem 3.1. This ensures that any F-trajectory on $[S,T]$ starting from $x_0 + \theta_0\mathbb{B}$ remains in $x_0 + \frac{3}{4}\epsilon_0\mathbb{B}$ for all $t \in [S, S+\tau_0]$. Then, Lemma 5.1 represents a local version of known linear $W^{1,1}$-estimates for smooth state constraints (cf. [3]). □

We can now restrict attention to the case in which x_0 belongs to a non-smooth region of A, i.e. $x_0 \in A \setminus A_0$.

Lemma 5.2. *Fix any $r_0 > 0$. Consider a multifunction $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and a non-empty closed set $A \subset \mathbb{R}^n$. Suppose that hypothesis $(H2)'$ of Theorem 3.1 is satisfied (for some constants $c > 0$, $k_F > 0$ and $R_0 = e^{c(T-S)}(r_0+2)$). Take any $x_0 \in A \cap r_0\mathbb{B}$, and any $w_0 \in \operatorname{int} T_A^C(x_0)$. Then we can find $\alpha \geq 0$, $\theta_1 \in (0, \frac{1}{16})$, $\tau_1 \in (0, T-S]$ and $K_1 > 0$ such that the following property is satisfied: given any F-trajectory $\hat{x}(.)$ on $[S,T]$ and $\varepsilon \geq 0$ such that*

$$\begin{cases} \max\limits_{t \in [S,T]} d_A(\hat{x}(t)) \leq \varepsilon \\ \hat{x}(S) \in A \cap (x_0 + \theta_1\mathbb{B}), \end{cases} \tag{13}$$

we can find an F-trajectory $x(.)$ on $[S,T]$ such that

$$\begin{cases} x(S) = \hat{x}(S) + \alpha\varepsilon w_0, \\ x(t) \in A \text{ for all } t \in [S, S+\tau_1], \\ \|x(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K_1\varepsilon \, . \end{cases}$$

Proof. From the characterization of the interior of the Clarke tangent cone of Rockafellar [21], since $w_0 \in \operatorname{int} T_A^C(x_0)$, there exists $\epsilon_1 \in (0, \frac{1}{4})$ such that

$$y + [0, \epsilon_1](w_0 + 2\epsilon_1\mathbb{B}) \subset A \ \text{ for all } y \in (x_0 + \epsilon_1\mathbb{B}) \cap A \, . \tag{14}$$

Recall that the positive constant $M_0 = c(1+R_0)$ (provided in Step 1 of the proof of Theorem 3.1) represents an upper bound for the velocities of all F-trajectories with initial state in $(r_0+1)\mathbb{B}$.

It is not restrictive to assume that $\varepsilon \leq \epsilon_1^2$. For, otherwise, the statement of the

lemma immediately follows taking $\alpha = 0$, $K_1 = \frac{2c(1+R_0)(T-S)}{\epsilon_1^2}$, and any feasible F-trajectory $x(.)$ with $x(S) = \hat{x}(S)$.
Define $\theta_1 := \frac{\epsilon_1}{4}$, $\alpha := \frac{1}{\epsilon_1}$ and $\tau_1 \in (0, \frac{\epsilon_1}{2M_0}]$ such that $\alpha|w_0|(e^{k_F\tau_1} - 1) \leq 1$. Take any F-trajectory $\hat{x}(.)$ on $[S,T]$ and any $\varepsilon \geq 0$ satisfying (13). Consider the following arc $y(.)$:

$$y(t) := \alpha\varepsilon w_0 + \hat{x}(t), \quad \text{for all } t \in [S,T].$$

Observe that $y(t) = \hat{y}(t) + \alpha\varepsilon w_0 + (\hat{x}(t) - \hat{y}(t))$, for all $t \in [S,T]$, where $\hat{y}(t) \in \pi_A(\hat{x}(t))$ ($\pi_A(x)$ denotes the projection of x, possibly non-unique, on the convex set A). Bearing in mind that $\hat{x}(.)$ satisfies (13), we have

$$|\hat{x}(t) - \hat{y}(t)| \leq \varepsilon, \quad \text{for all } t \in [S,T],$$

and therefore, for all $t \in [S,T]$,

$$\begin{aligned} y(t) + \varepsilon\mathbb{B} &\subset \hat{y}(t) + \alpha\varepsilon w_0 + 2\varepsilon\mathbb{B} = \hat{y}(t) + \alpha\varepsilon(w_0 + 2\epsilon_1\mathbb{B}) \\ &\subset \hat{y}(t) + \epsilon_1(w_0 + 2\epsilon_1\mathbb{B}), \end{aligned}$$

where in the last inclusion we have used the fact that $\varepsilon \leq \epsilon_1^2$. As a consequence, as long as $\hat{y}(t) \in (x_0 + \epsilon_1\mathbb{B}) \cap A$, from (14) we deduce that $y(t) + \varepsilon\mathbb{B} \subset A$ for all $t \in [S, S+\tau_1]$. But this is guaranteed for all $t \in [S, S+\tau_1]$ owing to the choice of θ_1 and τ_1; indeed, we have:

$$\begin{aligned} |x_0 - \hat{y}(t)| &\leq |x_0 - \hat{x}(S)| + |\hat{x}(S) - \hat{x}(t)| + |\hat{x}(t) - \hat{y}(t)| \\ &\leq \theta_1 + \tau_1 M_0 + \varepsilon \leq \epsilon_1/4 + \epsilon_1/2 + \epsilon_1^2 \\ &< \epsilon_1 . \end{aligned} \tag{15}$$

Consequently,

$$y(t) + \varepsilon\mathbb{B} \subset A, \quad \text{for all } t \in [S, S+\tau_1]. \tag{16}$$

Invoking the Filippov Existence Theorem (cf. [2] or [22]) to the reference trajectory $\hat{x}(.)$, we obtain an F-trajectory $x(.)$ on $[S,T]$ such that:

$$x(S) := y(S) = \hat{x}(S) + \alpha\varepsilon w_0$$

and for all $t \in (S,T]$:

$$\begin{aligned} \|x(.) - \hat{x}(.)\|_{L^\infty(S,t)} &\leq |x(S) - \hat{x}(S)| + \int_S^t |\dot{x}(s) - \dot{\hat{x}}(s)| ds \\ &\leq e^{k_F(t-S)}|x(S) - \hat{x}(S)|. \end{aligned} \tag{17}$$

Thus, for all $t \in (S,T]$, we have:

$$\|x(.) - \hat{x}(.)\|_{L^\infty(S,t)} \leq \alpha|w_0|e^{k_F(t-S)}\varepsilon$$

and

$$\|x(.) - \hat{x}(.)\|_{W^{1,1}(S,t)} \leq \alpha|w_0|(2e^{k_F(t-S)} - 1)\varepsilon . \tag{18}$$

Since the arcs $x(.)$ and $y(.)$ have the same left end-point ($x(S) = y(S) = \hat{x}(S) + \alpha\varepsilon w_0$), and $\dot{y}(t) = \dot{\hat{x}}(t)$ for a.e. $t \in [S,T]$, from (17) we also obtain that

$$|x(t) - y(t)| \leq \int_S^t |\dot{x}(s) - \dot{\hat{x}}(s)|ds \leq \alpha|w_0|(e^{k_F(t-S)} - 1)\varepsilon, \quad \text{for all } t \in (S,T]$$

and then, from the choice of τ_1,

$$|x(t) - y(t)| \leq \varepsilon, \quad \text{for all } t \in [S, S + \tau_1]. \tag{19}$$

As a consequence, the statement of the lemma is confirmed by relations (16) and (19), which yield the feasibility of the F-trajectory $x(.)$ on $[S, S + \tau_1]$, and (18), which provides the desired linear $W^{1,1}$-estimate with constant $K_1 := \alpha|w_0|(2e^{k_F(T-S)} - 1)$. □

Step 3. We shall apply an iterative construction based on the following technical result.

Lemma 5.3. *Fix $r_0 > 0$. Consider a multifunction $F : \mathbb{R}^n \rightsquigarrow \mathbb{R}^n$ and a non-empty closed set $A \subset \mathbb{R}^n$ satisfying the assumptions of Theorem* 3.1 *(for some constants $c > 0$, $k_F > 0$, and $R_0 = e^{c(T-S)}(r_0 + 2)$). Then there exists $\tau_2 \in (0, T - S]$ such that for each $0 < \tau \leq \tau_2$, we can find $\varepsilon_2 > 0$ and $K_2 > 0$ satisfying the following property.*
Take any $t_0 \in [S,T)$, any F-trajectory $\hat{x}(.)$ on $[t_0, T]$, and $\varepsilon \in [0, \varepsilon_2]$ satisfying:

$$\begin{cases} \hat{x}(t_0) \in A \cap (e^{c(t_0-S)}(r_0 + 1) - 1)\mathbb{B}, \\ \hat{x}(t) \in A \textit{ for all } t \in [t_0, t_0 + \tau], \\ d_A(\hat{x}(t)) \leq \varepsilon \textit{ for all } t \in [t_0 + \tau, t_0 + 2\tau]. \end{cases}$$

Then, there exists an F-trajectory $x(.)$ on $[t_0, T]$ such that

$$\begin{cases} x(t_0) = \hat{x}(t_0), \\ x(t) \in A \textit{ for all } t \in [t_0, t_0 + 2\tau], \\ \|x(.) - \hat{x}(.)\|_{W^{1,1}(t_0,T)} \leq K_2\varepsilon \ . \end{cases}$$

Proof. The proof is exactly the same as the proof of [4, Lemma 1]. We highlight only the fact that the properties listed in step 1 of the proof of Theorem 3.1 (which are an immediate consequence of assumptions $(H2)'$ and $(H3)'$) represent a localized version of the hypotheses invoked in [4, Lemma 1]. □

Assume first that $x_0 \in A_0$.
Define

$$\theta := \theta_0, \qquad \tau^* := \min\{\tau_0, \tau_2\}, \qquad K^* := \max\{K_0, K_2\},$$

where θ_0, τ_0 and K_0 are provided by Lemma 5.1, whereas τ_2 and K_2 are obtained using Lemma 5.3 (K_2 is related to time $\tau \leq \tau_2$). Write m the smaller integer such that $m\tau^* \geq T - S$, and consider the following standard iterative construction. We apply Lemma 5.1 to the reference F-trajectory $\hat{x}(.)$ and $\varepsilon \geq 0$ satisfying $\max_{t\in[S,T]} d_A(\hat{x}(t)) \leq \varepsilon$ and $\hat{x}(S) \in (x_0 + \theta\mathbb{B}) \cap A$. Therefore, we obtain an F-trajectory $z_0(.)$ on $[S,T]$ such that

$$\begin{cases} z_0(t) \in A, \text{ for all } t \in [S, S+\tau^*] \\ \|z_0(.) - \hat{x}(.)\|_{W^{1,1}(S,T)} \leq K^*\varepsilon \\ \max\limits_{t\in[S,T]} \; d_A(z_0(t)) \leq (K^*+1)\varepsilon \; . \end{cases}$$

Observe that the F-trajectory $z_0(.)$ verifies all the assumptions of Lemma 5.3 (for $t_0 = S$). As a result, we deduce the existence of an F-trajectory $z_1(.)$ on $[S,T]$ satisfying the following relations:

$$\begin{cases} z_1(t) \in A, \text{ for all } t \in [S, S+2\tau^*] \\ \|z_1(.) - z_0(.)\|_{W^{1,1}(S,T)} \leq K^*(K^*+1)\varepsilon \\ \max\limits_{t\in[S,T]} \; d_A(z_1(t)) \leq (K^*+1)^2\varepsilon \; . \end{cases}$$

The above construction can be iteratively applied m-times, obtaining a finite sequence of F-trajectories $z_k(.)$, for $k = 1, \dots, m$, such that:

$$\begin{cases} z_k(t) = z_{k-1}(t), \text{ for all } t \in [S, S+(k-1)\tau^*] \\ z_k(t) \in A, \text{ for all } t \in [S+(k-1)\tau^*, S+k\tau^*] \\ \|z_k(.) - z_{k-1}(.)\|_{W^{1,1}(S,T)} \leq K^*(K^*+1)^k\varepsilon \\ \max\limits_{t\in[S,T]} \; d_A(z_k(t)) \leq (K^*+1)^{k+1}\varepsilon \, . \end{cases}$$

The F-trajectory $x(.) := z_m(.)$ satisfies the theorem assertions in case (i), where for the $W^{1,1}$-estimate we take the constant $K := (K^*+1)^{m+1}$.

The case in which $x_0 \in A \setminus A_0$ can be treated applying exactly the same argument, except the fact that we use Lemma 5.2 instead of Lemma 5.1. □

Proof of Theorem 3.3. The proof is analogous to that one of Theorem 3.1 (and Corollary 3.2 whenever the initial state $\hat{x}(S)$ does not necessarily belong to the state constraint A). Indeed, we can use the same argument of step 1 of the proof of Theorem 3.1, which now is applied to the compact set

$$A \cap \{x \; : \; |x - \bar{x}(t)| \leq \delta, \; t \in [S,T]\}.$$

Estimates on an initial time interval (provided by Lemmas 5.1 and 5.2 in the proof of Thm 3.1) are still valid if we replace the condition $\hat{x}(S) \in A \cap (x_0 + \theta\mathbb{B})$ with the condition

$$||\hat{x}(.) - \bar{x}(.)||_{W^{1,1}(S,T)} \leq \theta,$$

for some $0 < \theta < \delta$. Observe that the latter is much 'stronger' in the sense that we require that the F-trajectory $\hat{x}(.)$ is close w.r.t. the $W^{1,1}$-norm to the reference arc $\bar{x}(.)$, rather than the mere left end-point $\hat{x}(S)$ is close to a given point x_0: we can take $0 < \theta < \delta$ small enough in such a manner that all the involved arcs and F-trajectories remain in the set

$$\{x \ : \ |x - \bar{x}(t)| \le \delta/2, \ t \in [S, T]\},$$

in which assumptions $(H2)'$ and $(H3)'$ are always applicable. This guarantees the validity of all estimates of the iterative argument (cf. step 3 in the proof of Theorem 3.1) confirming all theorem assertions. $\square$

6. An example

We consider a similar example to that considered in [4], but modified to include a more general left end-point constraint E_0 to illustrate the crucial role played by hypothesis $(H5)$ in providing normality of the necessary conditions of optimality. More precisely we look at the optimal control problem with state constraints:

$$(P2) \begin{cases} \text{Minimize } \int_0^1 L(t, x(t), \dot{x}(t))dt \\ \text{over arcs } x(.) \in W^{1,1}([0,1]; \mathbb{R}^2) \text{ s.t.} \\ \dot{x}(t) \in F \text{ a.e. } t \in [0,1], \\ x(t) \in A \text{ for all } t \in [0,1], \\ x(0) \in E_0, \end{cases}$$

where

$$A := \{x = (x_1, x_2) \in \mathbb{R}^2 : h(x) \le 0\} = \{(x_1, x_2) \in \mathbb{R}^2 : |x_2| \le x_1\}.$$

$$F := \{(1, u) \ : \ u \in \{-2, 2\}\}, \qquad L(t, u) := \begin{cases} 2 - u & \text{if } \dot{z}(t) = 2 \\ u + 2 & \text{if } \dot{z}(t) = -2. \end{cases}$$

The arc $z(.) : [0,1] \to \mathbb{R}$ is defined by: $z(0) = 0$, $z(t_k) := (-1)^k t_k$, and $\dot{z}(t) = (-1)^k 2$, for all $t \in (t_{k+1}, t_k]$, where $t_k := \frac{1}{3^k}$, for $k = 0, 1, 2, 3, \ldots$. Here, E_0 is any closed subset of $\mathbb{R}^2$ containing $\{(0,0)\}$ and such that assumption $(H5)(b)$ is verified for $\bar{x}(S) = (0,0)$ (not just $E_0 = \mathbb{R}^2$). The F-trajectory

$$\bar{x}(t) := (t, z(t)), \quad \forall t \in [0,1],$$

is a $W^{1,1}$-minimizer for problem $(P2)$, no matter the left end-point constraint is given, for we are assuming that $\bar{x}(0) = (0,0) \in E_0$.

Applying the necessary conditions of optimality for the reference arc $\bar{x}(.)$, if we suppose that the Lagrange multiplier (associated with the cost) vanishes, i.e. $\lambda = 0$, we obtain the existence of a co-state arc $p(.) \equiv p(0) \neq 0$ such that (cf. [4] for the details)

$$p(0) \in N_{E_0 \cap A}(\bar{x}(0)) \tag{20}$$

and

$$p(0) \in -N_A(\bar{x}(0)). \tag{21}$$

(We recall that (20) is just the left end-point transversality condition, whereas (21) follows from the Weierstrass and the right-end point transversality conditions.)

Now, assumption $(H5)(b)$ ensures the existence of a vector $w_0 \in \operatorname{int} T^C_{E_0 \cap A}(\bar{x}(0))$. Notice that the convexity of A yields the inclusion $\operatorname{int} T^C_{E_0\cap A}(\bar{x}(0)) \subset \operatorname{int} T^C_A(\bar{x}(0))$. Using the polar relation between the Clarke tangent cone and the (limiting) normal cone, since $w_0 \in \operatorname{int} T^C_{E_0\cap A}(\bar{x}(0))$, we have

$$w_0 \cdot \xi < 0, \quad \text{for all } \xi \in N_{E_0\cap A}(\bar{x}(0)) \setminus \{0\}.$$

Taking $\xi = p(0)$ we obtain the following inequality:

$$w_0 \cdot p(0) < 0. \tag{22}$$

And recalling also that $w_0 \in \operatorname{int} T^C_A(\bar{x}(0))$:

$$w_0 \cdot \zeta < 0, \quad \text{for all } \zeta \in N_A(\bar{x}(0)) \setminus \{0\}.$$

Then, for $\zeta = -p(0)$, we have the inequality $w_0 \cdot (-p(0)) < 0$, which contradicts the previous relation (22). This confirms normality of the necessary conditions once assumption $(H5)(b)$ is satisfied. The earlier work in [4] does not yield the information that there exists a normal minimizer in this case, since the left end-point constraint E_0 is not the whole space $\mathbb{R}^2$.

Remark 6.1. Observe that if we had $E_0 = \{(0,0)\}$, then hypothesis $(H5)(b)$ would not be satisfied, and the necessary conditions for optimality would apply in the degenerate form for the minimizer $\bar{x}(.)$ (cf. [4]), confirming the relevance of assumption $(H5)(b)$.

References

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

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

[3] P. Bettiol, A. Bressan, R. B. Vinter: On trajectories satisfying a state constraint: $W^{1,1}$ estimates and counter-examples, SIAM J. Control Optim. 48(7) (2010) 4664–4679.

[4] P. Bettiol, G. Facchi: Linear estimates for trajectories of state-constrained differential inclusions and normality conditions in optimal control, J. Math. Anal. Appl. 414(2) (2014) 914–933.

[5] P. Bettiol, H. Frankowska: Normality of the maximum principle for non convex constrained Bolza problems, J. Differ. Equations 243(2) (2007) 256–269.

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

[7] P. Bettiol, R. B. Vinter: Trajectories satisfying a smooth state constraint: improved estimates, IEEE Trans. Autom. Control. 56(5) (2011) 1090–1096.

[8] A. Bressan, G. Facchi: Trajectories of differential inclusions with state constraints, J. Diff. Equations 250 (2011) 2267–2281.

[9] F. H. Clarke: Optimization and Nonsmooth Analysis, Wiley, New York (1983).

[10] A. Yu. Dubovitskii, A. A. Milyutin: Extremum problems in the presence of restrictions, U.S.S.R. Comput. Math. Math. Phys. 5(3) (1965) 1–80.

[11] M. M. A. Ferreira, F. A. C. C. Fontes, R. Vinter: Nondegenerate necessary conditions for nonconvex optimal control problems with state constraints, J. Math. Anal. Appl. 233(1) (1999) 116–129.

[12] M. M. A. Ferreira, R. Vinter: When is the maximum principle for state constrained problems nondegenerate?, J. Math. Anal. Appl. 187(2) (1994) 438–467.

[13] F. A. C. C. Fontes, S. O. Lopes, R. de Pinho: On constraint qualifications for nondegenerate neccesary conditions of optimality applied to optimal control problems, Discrete Contin. Dyn. Syst. 29(2) (2011) 559–575.

[14] F. A. C. C. Fontes, S. O. Lopes: Normal forms of necessary conditions for dynamic optimization problems with pathwise inequality constraints, J. Math. Anal. Appl. 399(1) (2013) 27–37.

[15] H. Frankowska: Normality of the maximum principle for absolutely continuous solutions to Bolza problems under state constraints, Control Cybern. 38(4B) (2009) 1327–1340.

[16] H. Frankowska, D. Tonon: Inward pointing trajectories, normality of the maximum principle and the non occurrence of the Lavrentieff phenomenon in optimal control under state constraints, J. Convex Analysis 20(4) (2013) 1147–1180.

[17] R. A. Poliquin, R. T. Rockafellar, L. Thibault: Local differentiability of distance functions, Trans. Amer. Math. Soc. 352 (2000) 5231–5249.

[18] M. Palladino, R. B. Vinter: Minimizers that are not also relaxed minimizers, SIAM J. Control Optim. 52(4) (2014) 2164–2179.

[19] 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 Inf. 16 (1999) 335–351.

[20] F. Rampazzo, R. B. Vinter: Degenerate optimal control problems with state constraints, SIAM J. Control Optim. 39 (2000) 989–1007.

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

[22] R. B. Vinter: Optimal Control, Birkhäuser, Boston (2000).