Optimal Control Problems with Mixed and Pure State Constraints*

A. Boccia†, M.d.R. de Pinho‡ and R.B. Vinter§

Abstract

This paper provides necessary conditions of optimality for optimal control problems, in which the pathwise constraints comprise both 'pure' constraints on the state variable and also 'mixed' constraints on control and state variables. The proofs are along the lines of earlier analysis for mixed constraint problems, according to which Clarke's theory of 'stratified' necessary conditions is applied to a modified optimal control problem resulting from absorbing the mixed constraint into the dynamics; the difference here is that necessary conditions which now take account of the presence of pure state constraints are applied to the modified problem. Necessary conditions are given for a rather general formulation of the problem containing both forms of the constraints, and then these are specialized to apply to problems having special structure. While combined pure state and mixed control/state problems have been previously treated in the literature, the necessary conditions in this paper are proved under less restrictive hypotheses and for novel formulations of the constraints.

Key Words: Optimal Control, Maximum Principle, State Constraints, Nonsmooth Analysis.

AMS Subject Classifications: 49K15, 49K21

1 Introduction

The Pontryagin Maximum Principle (PMP), dating from the 1950's, is a cornerstone of Optimal Control Theory [15]. In its original form, the PMP provides necessary conditions that a solution $x(.) : [a, b] \to \mathbb{R}^n$ to a controlled differential equation (named a state trajectory) and the associated control function $u(.) : [a, b] \to \mathbb{R}^m$, minimize a cost function, subject to a pathwise constraint on the control function $u(.)$:

$$u(t) \in U(t) \quad \text{for all } t \in [a, b]$$

and to a constraint on the endpoints of $x(.)$:

$$(x(a), x(b)) \in E \,,$$

in which $U(t) \subset \mathbb{R}^m$, $a \le t \le T$, and $E \subset \mathbb{R}^n \times R^n$ are given sets. The PMP features an absolutely continuous function $p(.)$ (the costate trajectory) that satisfies a differential equation, named the costate equation (or a related differential inclusion, if the data is non-smooth) and boundary conditions named the transversality conditions. Information is then provided about the optimal control: 'the Hamiltonian

---

†Dept. of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, USA.

‡Faculadade de Engenharia, Universidade do Porto, DEEC, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal.

§EEE Dept., Imperial College London, Exhibition Road, London SW7 2BT, UK.

evaluated along the optimal state trajectory and costate trajectory is maximized at the minimizing control function, pointwise in time'.

In subsequent research into extensions to take account of pathwise constraints also for state trajectories, it was early recognised that different techniques were required to deal with the two types of constraints:

1. *mixed control/state constraints:*

$$\phi^1(t, x(t), u(t)) \leq 0 \quad \text{and} \quad \phi^2(t, x(t), u(t)) = 0, \tag{1.1}$$

2. *pure state constraints:* $h(t, x(t)) \leq 0$,

involving the given functions $\phi^1(.,.,.)$, $\phi^2(.,.,.)$ and $h(.,.)$.

In the mixed control/state constraint literature, typified by [13], [8] and [10], the aim is usually to show that the constraint can be accommodated, under a suitable constraint qualification, by means of Lagrange multipliers, namely suitably bounded measurable functions $\zeta_1(.)$ and $\zeta_2(.)$ on $[a, b]$. The modified PMP is, once again, expressed in terms of an absolutely continuous costate trajectory $p(.)$. It takes the form of the original PMP, in a restricted sense, when the cost is modified by the addition of the term:

$$\int_a^b \phi^1(t, x(t), u(t)) \cdot \zeta_1(t)dt + \int_a^b \phi^2(t, x(t), u(t)) \cdot \zeta_2(t)dt,$$

together with 'complementary slackness' conditions. (The conditions are 'restricted' in the sense that the Hamiltonian is maximized along the optimal state trajectory $\bar{x}(.)$, not over $U(t)$ but over the smaller set $U(t) \cap \{u' : \phi^1(t, \bar{x}(t), u') \leq 0 \text{ and } \phi^2(t, \bar{x}(t), u') = 0\}$.)

Implicit in the derivation of this modified PMP is the notion that, under a suitable 'constraint qualification', the mixed constraints can be absorbed into the dynamic constraint in such a manner that the hypotheses on the dynamic constraint invoked in the original PMP continue to be satisfied; application of the original PMP to the reformulated optimal control problem, with an 'absorbed' dynamic constraint, yields the desired, modified, necessary conditions.

It might be thought that the theory of necessary conditions for pure state constraints '$h(t, x(t)) \leq 0$' could simply be subsumed into that for mixed state constraints, by setting $\phi_1(t, x, u) = h(t, x)$. But this is not possible (if we are to use the proof techniques referred to above) because, when $\phi_1(t, x, u)$ does not depend on $u(.)$, the constraint qualification is violated, in consequence of which the mixed constraint cannot be used to eliminate control variable components and thereby generate an equivalent optimal control problem to which standard versions of the PMP can be applied. This is the reason why 'pure' state constraints have been treated separately from 'mixed control/state constraints', and accounts for the fundamentally different nature of modifications to the PMP that have been derived for optimal control problems with pure state constraints: the modified PMP, in the latter case, is formulated in terms of a 'measure multiplier' and a, possibly discontinuous, costate trajectory $q(.)$ that is of bounded variation. Papers rigorously treating pure state constraints and involving a discountinuous costate trajectory go back to the independent work of Dubovitskii and Milyutin [9] and of Warga [20] in the 1960's; references to the early literature appear, for example, in [19]. Necessary conditions to cover the combined occurence of the two types of constraints has been derived by Dmitruk [10] and other members of the Dubovitski-Milyutin school (see [11] for an overview of this work), and also by Makowski and Neustadt [13].

A breakthrough in the development of new tools to tackle a variety of differently structured optimal control problems, with nonsmooth data, was the publication of Clarke's memoir [4], the centrepiece of which was 'stratified' necessary conditions for optimal control problems, whose dynamic constraint took the form of a differential inclusion. The memoir introduced a new, and very useful, way of capturing the requisite

Lipschitz continuity-like properties of the differential inclusion for the validation of the Euler-Lagrange inclusion, and related necessary conditions, namely the 'bounded slope' hypothesis. Subsequently, Clarke and Pinho [5] examined the implications of these tools for 'mixed constraint problems'. Necessary conditions were provided for very general formulations of mixed constraint problems, by showing that they could be reduced to the optimal control problems treated in [4], by absorbing the mixed constraints into the dynamic constraint in such a manner that the bounded slope hypothesis continued to be satisfied. The results in [5] improve on many earlier-derived mixed constraint conditions, as described in detail in [5, Section 8], in some respects even when attention is restricted to problems with smooth data. However the presence of pure state constraints is excluded from the necessary conditions in [5]:

'. . the bounded slope condition excludes unilateral state constraints . . It is well-known that the necessary conditions of the type given . . fail, and their appropriate extensions involve measures and costate trajectories that are discontinuous'.

The aim of this paper is to provide extensions of the necessary conditions for mixed control/state constraints problems of [5], to allow also for pure state constraints. A key tool is the set of stratified necessary conditions of [1], that generalize the main necessary conditions in Clarke's memoir, to allow for unilateral state constraints. The proof technique for deriving necessary conditions, applying to optimal control problems with both mixed control/state and pure state constraints, is to reduce the problems to ones involving pure state constraints alone, by absorbing the mixed constraints into the dynamic constraint, and applying the stratified pure-state-constraint necessary conditions of [1].

The necessary conditions for combined mixed control/state and pure state constraints of this paper reduce (in important respects) to the main necessary conditions in [5] and [6] for mixed constraints alone (both in a general setting and when the mixed constraint has explicit representations including those of [6]), following removal of the pure state constraint. [5, Section 8] ('relation to the literature') gives details of the respects in which the necessary conditions in [5] improve on earlier necessary conditions, regarding hypotheses on the mixed state constraint data. These improvements are all the more in evidence in this paper, since the framework is broadened to include pure state constraints. On the other hand, this paper extends the earlier necessary conditions involving both mixed constraints and pure state constraints in [10] and [13], by allowing non-smooth data, by adopting a very general formulation of the mixed constraints of the form '$(x(t), v(t)) \in S(t, w(t))$' involving the controls $(u = (v, w))$ and states x in place of a collection of functional equality and inequality constraints, and by permitting general endpoint constraints.

This paper treats only optimal control problems with mixed constraints that are 'regular', in the sense that they separate into a constraint satisfying the bounded slope condition and pure control and state constraint. It fails to provide conditions for problems with mixed constraints that do not decompose in this way, for example '$|x|^2 + |u|^2 \leq 1$'. Necessary conditions for such problems have been studied by Milyutin et al. For an overview, references to this work and some open questions, see [11].

We mention that the 'stratified' necessary conditions in this paper are expressed in terms of an arbitrary radius multifunction $R(t)$, in place of the balls $\dot{\bar{x}}(t) + r(t)\mathbb{B}$ involving the radius function $r(t)$, as in [5]. This extra degree of generality in these necessary conditions for the general formulation of the optimal control problem of Section 2, simplifies the application of these conditions to take account of more structured formulations of the problem in Section 3, as compared with the analysis in [6].

Finally, at the suggestion of a reviewer, we comment on key differences between our proof techniques and the Dubovitskii/Milyutin scheme applied to control/state constraint problems as summarized, for example, in [11]. In the latter literature, the necessary conditions are deduced from an abstract Lagrange multiplier rule for an abstract optimization problem on Banach spaces, involving constraints associated with the original dynamic constraint, with the endpoint constraints, and with the mixed constraint (each parameterized by a control function). The constraint associated with the original dynamic constraint is

expressed in terms of a mapping from $W^{1,1}$ to L^1. It is straightforward to show that this mapping is 'regular' (i.e. Frechet differentiable and surjective) as required for the application of the abstract multiplier rule; it gives rise to a Lagrange multiplier in $L^\infty = (L^1)^*$, which, in consequence of the multiplier rule conditions, can be represented by an absolutely continuous function (in the case of no pure state constraints), interpreted as the costate trajectory. The endpoint constraints translate into transversality conditions via the abstract multiplier rule. The troublesome constraint is the mixed constraint which, in the abstract framework, gives rise to a constraint function with range space in L^∞. The Lagrange multiplier associated with this constraint is an element in the topological dual space $(L^\infty)^*$. An important step in the application of the Dubovitskii/Milyutin scheme in this context is to show that, under a constraint qualification (positive linear independence of the gradients of the 'mixed' constraint functional w.r.t. the control variable), this element can be represented by a point in the pre-dual space L^∞; this point appears as a Lagrange multiplier for the mixed constraint in the final statement of the necessary conditions. The proof techniques used in this paper (and in the related paper [6]), by contrast, circumvent altogether the difficulties associated with the fact that the natural space for the mixed-constraint Lagrange multiplier is in the difficult-to-deal-with space $(L^\infty)^*$; this is achieved by using the constraint qualification (now manifesting itself as the 'bounded slope' condition) *at the beginning* of the proof, to justify eliminating the mixed constraint by absorbing it into the dynamic constraint, instead of *at the end* of the proof to refine the consequences of the application of the abstract Lagrange multiplier. The dynamic constraint (in our framework) now becomes a differential inclusion but, with the help of nonsmooth analysis, this can be simply accommodated in the necessary conditions via a costate function, corresponding to the straightforward manner in which the differential equation constraint is dealt with in the Dubovistskii/Milyutin scheme. Advantages of our approach are simplicity (at least in avoiding consideration of the $(L^\infty)^*$ as a multiplier space) and that it permits consideration of nonsmooth data. An advantage of the Dubovistskii/Milyutin scheme is that it can be also used to derive necessary conditions for 'irregular' problems, namely problems for which the pathwise constraints do not separate into a pure state constraint and a mixed control/state constraint satisfying the bounded slope condition (see, e.g., [11]).

Notation: The Euclidean norm of a vector $x \in \mathbb{R}^n$ is $|x|$. $\mathbb{B}$ indicates the open unit ball in $\mathbb{R}^n$ and the distance function of a point $x \in \mathbb{R}^n$ from a set $A \subset \mathbb{R}^n$ is defined as

$$d_A(x) := \inf\{|x - y| : y \in A\}.$$

The convex hull of the set A is written $\mathrm{co}\, A$. Given a multifunction $F(.) : \mathbb{R}^n \rightsquigarrow \mathbb{R}^k$, we denote by $\mathrm{Gr}\, F(.)$ the graph of $F(.)$.

$L^1([a,b]; \mathbb{R}^n)$ and $L^1([a,b]; \mathbb{R}^n)$ have their usual meanings as spaces of integrable and essentially bounded, measurable n-vector valued functions on $[a,b]$. $W^{1,1}([a,b]; \mathbb{R}^n)$ denotes the space of absolutely continuous functions $x : [a,b] \to \mathbb{R}^n$, equipped with the norm

$$\|x\|_{W^{1,1}} := |x(a)| + \int_a^b |\dot{x}(t)|\, dt.$$

We write L^1 in place of $L^1([a,b]; \mathbb{R}^n)$, etc., when no ambiguity arises. $NBV^+[a,b]$ denotes the space of nondecreasing, real-valued functions $\mu(.)$ on $[a,b]$ of bounded variation, vanishing at the point a and right continuous on (a,b). The total variation of a function $\mu(.) \in NBV^+[a,b]$ is written $\|\mu\|_{T.V.}$. As is well known, each point $\mu(.) \in NBV^+[a,b]$ defines a Borel measure on $[a,b]$. This associated measure is also denoted μ. We denote by *supp* the support of the measure μ.

We make use of some constructs from nonsmooth analysis, described in detail, for example, in [7, 19]: given a closed set $E \subset \mathbb{R}^n$ and $x \in E$, the *proximal normal cone of E at x* is

$$N_E^P(x) := \{\zeta \in \mathbb{R}^n : \exists\, \epsilon > 0 \text{ and } M > 0 \text{ s.t. } \zeta \cdot (y - x) \leq M|x - y|^2 \text{ for all } y \in E \cap x + \epsilon\mathbb{B}\}.$$

The *limiting normal cone of E at x* is

$$N_E^L(x) := \{\lim_{i \to \infty} \zeta_i : \zeta_i \in N_E^P(x_i) \text{ and } x_i \in E \text{ for all } i, \text{ and } x_i \to x\}.$$

The *Clarke normal cone of E at x* is $N_E^C(x) := co\, N_E^L(x)$. If E is convex, the two normal cones coincide with the normal of cone of convex analysis, and are written $N_E^C(x)$. More generally, a set E is said to be *regular* at $x \in E$ if $N_E^L(x) = N_E^C(x)$.

The *Clarke tangent cone of E at x* is the polar set of the limiting normal cone, i.e.

$$T_E^C(x) := \{\eta \in \mathbb{R}^n \mid \langle \eta, \zeta \rangle \leq 0 \text{ for all } \zeta \in N_E^L(x)\}.$$

Given a lower semicontinuous function $f(.) : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \text{dom } f(.) := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$, the *proximal subdifferential of $f(.)$ at x* is the set

$$\partial_P f(x) := \left\{ \begin{array}{c} \zeta \in \mathbb{R}^n : \quad \exists\, \sigma > 0 \text{ and } \epsilon > 0 \text{ such that } \forall\, y \in x + \epsilon \mathbb{B} \\ f(y) - f(x) \geq \langle \zeta, y - x \rangle - \sigma |y - x|^2 \end{array} \right\}.$$

The *limiting subdifferential of $f(.)$ at x* is

$$\partial_L f(x) := \{\lim_{i \to \infty} \zeta_i : \zeta_i \in \partial_P f(x_i), x_i \to x, f(x_i) \to f(x)\}.$$

In the case $f(.)$ is Lipschitz continuous on a neighborhood of x, we define Clarke generalized gradients $\partial_C f(x) := co\, \partial_L f(x)$. We say that $f(.)$ *is strictly differentiable at x* if $f(.)$ is Fréchet differentiable at x, Lipschitz continuous on a neighborhood of x and $\{\nabla f(x)\} = \partial_C f(x)$. When the function $f(.)$ depends on two variables (x, y) we write $\partial_C f(., \bar{y})(\bar{x})$ to denote the Clarke generalized gradient of the function $x \to f(x, \bar{y})$ at $\bar{x}$.

2 Necessary Conditions for a General Problem

Consider the following optimal control problem:

$$(P) \begin{cases} \text{Minimize} \quad J(x(.), u(.)) := \ell(x(a), x(b)) + \int_a^b \Lambda(t, x(t), u(t))\, dt \\ \text{over processes } (x(.), u(.)) \text{ such that} \\ \quad \dot{x}(t) = f(t, x(t), u(t)), \qquad\qquad\quad \text{a.e. } t \in [a, b], \\ \quad u(t) = (v(t), w(t)), \quad w(t) \in W(t), \quad \text{a.e. } t \in [a, b], \\ \quad (x(t), v(t)) \in S(t, w(t)), \qquad\qquad \text{a.e. } t \in [a, b], \\ \quad h(t, x(t)) \leq 0, \qquad\qquad\qquad\quad \text{for all } t \in [a, b], \\ \quad (x(a), x(b)) \in E\,. \end{cases}$$

the data for which comprises integers $n > 0$, $m > 0$, $m_v \geq 0$ and $m_w \geq 0$ such that $m = m_v + m_w$, functions $\ell(.,.) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\Lambda(.,.,.) : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $f(.,.,.) : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $h(.,.) : [a, b] \times \mathbb{R}^n \to \mathbb{R}$, and sets $S \subset \mathbb{R}^{1+n+m}$, $E \subset \mathbb{R}^n \times \mathbb{R}^n$ and $W \subset \mathbb{R}^{1+m_w}$. (Interpret $W(t) := \{w : (t, w) \in W\}$, $S(t, w) := \{(x, v) : (t, x, (v, w)) \in S\}$, $S(t, x) = \{u : (t, x, u) \in S\}$, etc.)

A *process* is a pair of functions $(x(.), u(.))$, in which $x(.)$ is a $W^{1,1}([a, b]; \mathbb{R}^n)$ function and $u(.) = (v(.), w(.)) : [a, b] \to \mathbb{R}^m$ is a measurable function, satisfying $\dot{x}(t) = f(t, x(t), u(t))$ and $w(t) \in W(t)$ a.e.. A process $(x(.), u(.))$ in which $x(.)$ and $u(.)$ satisfy the constraints of (P), and for which $t \to \Lambda(t, x(t), u(t))$ is integrable, is called a *feasible process*.

As in [4], we consider local minimizers for problem (P), with respect to a given 'radius multifunction' $R(.)$ in the following sense:

5

Definition 2.1

(a): a multifunction $R(.) : [a, b] \rightsquigarrow \mathbb{R}^m$ is called a *radius multifunction* if $R(t)$ is a non-empty, open, convex set for each t and there exists $r_0 > 0$ such that $r_0 \mathbb{B} \subset R(t)$ a.e..

(b): Given a radius multifunction $R(.) : [a, b] \rightsquigarrow \mathbb{R}^m$, a feasible process $(\bar{x}(.), \bar{u}(.))$ for (P) is said to be a $W^{1,1}$ local minimizer w.r.t. $R(.)$ if there exists $\epsilon > 0$ such that

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

over all the feasible processes $(x(.), u(.))$ satisfying

$$\|x - \bar{x}\|_{W^{1,1}} \leq \epsilon \quad \text{and} \quad u(t) \in R(t) \quad \text{a.e..}$$

If $R(t) \equiv \mathbb{R}^m$, we simply say that '$(\bar{x}(.), \bar{u}(.)))$ is a $W^{1,1}$ local minimizer'.

We shall invoke the following hypotheses, in which $(\bar{x}(.), (\bar{u}(.) = (\bar{v}(.), \bar{w}(.)))$ is a given feasible process. For some $\epsilon > 0$:

(H1): $\ell(.,.)$ is Lipschitz continuous on a neighborhood of $(\bar{x}(a), \bar{x}(b))$. E is closed. W and S are $\mathcal{L} \times \mathcal{B}^{m'}$ measurable subsets of $[a, b] \times \mathbb{R}^{m_w}$ and $[a, b] \times \mathbb{R}^{m+n}$ respectively, where $\mathcal{L} \times \mathcal{B}^{m'}$ denotes the product σ-algebra of $\mathcal{L}$ and $\mathcal{B}^{m'}$, in which $\mathcal{L}$ denotes the Lebesgue subsets of $[a, b]$ and $\mathcal{B}^{m'}$ the Borel subsets of $\mathbb{R}^{m'}$ for $m' = m_w$ or $m + n$.

(H2): $h(.,.)$ is upper semicontinuous and there exists $k_h > 0$ such that

$$|h(t, x') - h(t, x)| \leq k_h |x' - x|,$$

for all t in $[a, b]$ and all $x', x \in \bar{x}(t) + \epsilon \mathbb{B}$.

(H3): for each $(x, v) \in \mathbb{R}^n \times \mathbb{R}^{m_v}$, $f(., x, (v, .))$ and $\Lambda(., x, (v, .))$ are $\mathcal{L} \times \mathcal{B}^{m_w}$ measurable. For a.e. $t \in [a, b]$ and every $w \in W(t)$ the functions $(x, v) \mapsto f(t, x, v, w)$ and $(x, v) \mapsto \Lambda(t, x, v, w)$ are Lipschitz continuous on a neighborhood of

$$((\bar{x}(t) + \epsilon \mathbb{B}) \times R(t, w)) \cap S(t, w),$$

where $R(t, w) := \{v : (v, w) \in R(t)\}$. The Lipschitz constants, which may depend on t and w, can be chosen to be $\mathcal{L} \times \mathcal{B}^{m_w}$ measurable functions. We denote these Lipschitz constants by the symbols $k_x^f(t, w)$, $k_v^f(t, w)$, $k_x^\Lambda(t, w)$, and $k_v^\Lambda(t, w)$, thus:

$$|\Lambda(t, x_1, v_1, w) - \Lambda(t, x_2, v_2, w)| \leq k_x^\Lambda(t, w)|x_1 - x_2| + k_v^\Lambda(t, w)|v_1 - v_2|,$$
$$|f(t, x_1, v_1, w) - f(t, x_2, v_2, w)| \leq k_x^f(t, w)|x_1 - x_2| + k_v^f(t, w)|v_1 - v_2|$$

for any (x_1, v_1) and (x_2, v_2) in a neighbourhood of $((\bar{x}(t) + \epsilon \mathbb{B}) \times R(t, w)) \cap S(t, w)$ and $w \in W(t)$, a.e. $t \in [a, b]$.

(BS): (*'Bounded Slope Condition'*) The set $S(t, w)$ is closed for each $w \in W(t)$, a.e. $t \in [a, b]$. There exists a measurable function $k_S(.)$ such that, given any $w \in W(t)$ and $(x, v) \in ((\bar{x}(t) + \epsilon \mathbb{B}) \times R(t, w)) \cap S(t, w)$ we have

$$(\alpha, \beta) \in N_{S(t,w)}^P(x, v) \subset R^n \times \mathbb{R}^{m_v} \implies |\alpha| \leq k_S(t)|\beta|, \quad \text{a.e. } t \in [a, b],$$

and $k_S(t) \geq k_0 > 0$, a.e., for some $k_0 > 0$.

There follows a set of necessary conditions for a feasible process to be a $W^{1,1}$ local minimizer w.r.t. $R(.)$, a given radius multifunction. For $\lambda^0 \geq 0$ define the Hamiltonian:

$$H^{\lambda^0}(t, x, v, w, p) := p \cdot f(t, x, v, w) - \lambda^0 \Lambda(t, x, v, w) .$$

Theorem 2.2 (General Necessary Conditions) *Let $R(.)$ be a radius multifunction and let $(\bar{x}(.), \bar{u}(.) = (\bar{v}(.), \bar{w}(.)))$ be a $W^{1,1}$ local minimizer w.r.t. $R(.)$. Assume hypotheses (H1)-(H3) and (BS) are satisfied. Assume also that*

$$k_x^f(t, \bar{w}(t)), \quad k_x^\Lambda(t, \bar{w}(t)), \quad k_S(t)[k_v^f(t, \bar{w}(t)) + k_v^\Lambda(t, \bar{w}(t))] \text{ are integrable ,} \tag{2.1}$$

and

$$\exists \ \eta > 0 \text{ such that } \bar{u}(t) + \eta k_S(t)\mathbb{B} \subset R(t), \text{ for a.e. } t \in [a, b]. \tag{2.2}$$

Then there exists a multiplier set $(p(.), \mu(.), \lambda^0) \in W^{1,1} \times NBV^+[a, b] \times \mathbb{R}^+$ and a μ-integrable function $m(.) : [a, b] \to \mathbb{R}^n$ such that

(a): $\lambda^0 + \|p(.)\|_{L^\infty} + \|\mu\|_{T.V.} = 1$,

(b): $(-\dot{p}(t), 0) \in \partial_C H^{\lambda^0}(t, ., ., \bar{w}(t), q(t))(\bar{x}(t), \bar{v}(t)) - N_{S(t, \bar{w}(t))}^C(\bar{x}(t), \bar{v}(t))$ a.e.,

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

(d): $(q(a), -q(b)) \in \lambda^0 \partial_L \ell(\bar{x}(a), \bar{x}(b)) + N_E^L(\bar{x}(a), \bar{x}(b))$.

(e): For any $u = (v, w) \in R(t) \cap S(t, \bar{x}(t))$ such that $w \in W(t)$,

$$H^{\lambda^0}(t, \bar{x}(t), \bar{u}(t), q(t)) \geq H^{\lambda^0}(t, \bar{x}(t), u, q(t)) \quad a.e. \ t \in [a, b].$$

Here,
$$q(t) := \begin{cases} p(t) + \int_{[a,t]} m(s) \, \mu(ds) & if \ t \in (a, b] \\ p(a) & if \ t = a \end{cases}, \tag{2.3}$$

$S(t, x) = \{(v, w) : (t, x, (v, w)) \in S\}$ *(consistent with earlier notation), and $\partial_x^> h(t, x)$ is the set*

$$\partial_x^> h(t, x) := \{\lim_{(t_i, x_i) \to (t, x)} \nabla h(t_i, x_i) : \nabla h(t_i, x_i) \text{ exists and } h(t_i, x_i) > 0 \text{ for each } i\}.$$

A proof of Theorem 2.2 is given in Section 4. These conditions are expressed in terms of the function $q(.)$ that is right continuous on (a, b). The assertions of the theorem remain the same if $q(.)$ is replaced by the function $q'(.)$, which coincides with $q(.)$ at the end-times, and which, at interior points $t \in (a, b)$ is left continous and given by $q'(t) = p(t) + \int_{[a,t)} m(s)\mu(ds)$, as in [19].

Comments.

(i): The requirement that $k_S(.)$ in (BS) satisfies $k_S(t) \geq k_0$, for some $k_0 > 0$, is essential. The minorization condition might seem superfluous because, if it is violated, we can always arrange that (BS) is satisfied by addition of a positive constant. But this ignores the fact that $k_S(.)$ is also required to satisfy the condition $\bar{u}(t) + \eta k_S(t) \subset R(t)$, and addition of a positive constant to $k_S(.)$ might result in violation of this latter condition. A simple counter example can be constructed along the lines of [1, Ex. 1] illustrating that the assertions of Th. 2.2 are in general false if the minorization condition is omitted.

(ii): Because the necessary conditions (a)-(e) are invariant under positive scaling of the Lagrange multipliers $(\lambda^0, p(.), \mu)$, they can be replaced by equivalent conditions, in which (a) now takes the form $\lambda^0 + \|p(.)\|_{L^\infty} + \|\mu\|_{T.V.} > 0$. This alternative form is convenient in some applications.

(iii): In the formulation (P) of the optimal control problem, we interpret '$(x(t), v(t)) \in S(t, w(t))$ a.e.' as a mixed control/state constraint, as in the previous work of Clarke and Pinho [5]. This description of the constraint reduces to the standard mixed constraint description (1.1), in the form of a set of functional inequality and equality constraints, when the data does not depend on w, so that we can identify u with the control component v alone, and $S(t)$ (no longer dependent on w) is chosen to be

$$S(t) := \{(x, u) : \phi^1(t, x, u) \leq 0 \text{ and } \phi^2(t, x, u) = 0\} . \tag{2.4}$$

While the special case (2.4) covers most applications involving mixed control/state constraints, the more general description '$(x(t), v(t)) \in S(t, w(t)))$' is a useful starting point for formulating hypotheses under which necessary conditions of optimality may be derived for a wide variety of problems.

(iv): The pure state constraint description '$h(t, x(t)) \leq 0$' of problem (P), involving the function $h(t, x)$ which is upper semicontinuous and uniformly Lipschitz w.r.t. the x variable, has been widely employed since its introduction in [3], because of its versality. We recall that it covers multiple functional state inequality constraints of the type '$h_i(t, x(t)) \leq 0$', $i = 1, \ldots, k_s$, which can be accommodated by setting $h(t, x) := \max_i h_i(t, x)$, and it subsumes constraints imposed on a given closed subset $I \subset [a, b]$ of the time interval, of the form

$$x(t) \in A(t) \quad \text{for all } t \in I,$$

in which $A(.) : [a, b] \to \mathbb{R}^n$ is a given upper semicontinuous multifunction, taking values closed subsets. In this case we choose

$$h(t, x) := \begin{cases} d_{A(t)}(x) & \text{if } t \in I \\ -1 & \text{otherwise} . \end{cases}$$

(v): Following [10] and [5], we have partitioned the control variable as $u = (v, w)$; the components v and w will each have a different status, regarding the hypotheses that are imposed on them for the derivation of necessary conditions. The purpose here is to capture within a single framework, and thereby achieve greater generality, problems involving a simple set inclusion constraint on the control '$w(t) \in W(t)$', problems in which the contol/state variables (x, v) are required to satisfy, say, pathwise functional equality and inequality constraints, and a combination of such problems. A version of the theorem, valid for the case $m_w = 0$ is obtained by eliminating, in an obvious way, all reference to w variable in the theorem statement and accompaning hypotheses. Likewise, a version valid for the case $m_v = 0$ is obtained by removing all reference to the v variable.

(vi): Theorem 2.2 extends [5, Theorem 3.2], to allow for the presence of additional, pure state constraints. Notice that, as in [5], Theorem 2.2 incorporates information about the dependence of the Hamiltonian on the control, along a minimizing state trajectory and costate trajectory both in the form of the extended costate inclusion (c.f. [16]) and the Weierstrass condition (conditions (b) and (e) of the theorem statement). Earlier necessary conditions for problem (P), involving both pure and mixed constraints, for problems with nonsmooth data and for which 'bounded slope' type hypotheses on the mixed constraint are invoked, appear in [2]. Our necessary conditions, embodied in Theorem 2.2, broadly confirm the assertions of [2, Theorem 2] (the convergence analysis in [2], justifying the costate inclusion, is incomplete) and extend them in numerous respects. The conditions in this paper allow for a more general description of the mixed constraint (not just one involving mixed equality and inequality constraints as in [2]), for partitioned control variables $u = (v, w)$ in which very much weaker hypotheses are imposed on the data regarding w dependence and for 'stratification' (i.e. minimizers w.r.t. a radius multifunction). In our applications of the necessary conditions in the next section, we also allow non-compact control constraint sets.

3 Necessary Condition for Mixed Constraints in Explicit Form

Theorem 2.2 of the previous section provides necessary conditions for optimal control problems involving mixed state and control constraints and pure state constraints, when the mixed constraint is captured by the condition

$$(x(t), u(t)) \in S(t) ,$$

in which $S(.)$ is a given multifunction. Theorem 2.2 can be used as a starting point for the derivation of necessary conditions, in which the set $S(.)$ is of the form

$$S(t) := \{(x, u) \in \mathbb{R}^n \times \mathbb{R}^m : \phi(t, x, u) \in \Phi(t) \text{ and } u \in U(t)\}, \tag{3.1}$$

in which $\phi : [a, b] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^\kappa$ and $\Phi(.) : [a, b] \rightsquigarrow \mathbb{R}^\kappa$ is a Lebesgue measurable multifunction that takes values closed sets. The most widely considered special case of (3.1) (functional inequality and equality constraints) is that when, for some integers $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$ such that $\kappa_1 + \kappa_2 = \kappa$, $\phi(.,.,.)$ is partitioned as $\phi(.,.,.) = (\phi^1(.,.,.), \phi^2(.,.,.))$ into $\mathbb{R}^{\kappa_1}$ and $\mathbb{R}^{\kappa_2}$ valued functions respectively and

$$\Phi(t) \equiv \overbrace{(-\infty, 0] \times \ldots \times (-\infty, 0]}^{\kappa_1} \times \overbrace{\{0\} \times \ldots \times \{0\}}^{\kappa_2}. \tag{3.2}$$

The idea is to express the conditions directly in terms of $\phi(.,.)$ and $\Phi(.)$, under hypotheses that generalize earlier conditions governing the u-dependence of $\phi(x, u)$ (classical rank conditions or, more generally, Mangasarian-Fromowitz type conditions on the gradients of this function, and 'surjectivity conditions' originating in the work of Schwarzkopf [17].) We consider henceforth the variant of problem (P), labelled (P_S), in which we no longer distinguish block components v and w of the control variable u, and when we impose both a mixed constraint of the type (3.1), an implicit control constraint $u \in U(t)$ (in which $U(.) : [a, b] \rightsquigarrow \mathbb{R}^m$), and a pure state constraint:

$$(P_S) \begin{cases} \text{Minimize } \ell(x(a), x(b)) + \int_a^b \Lambda(t, x(t), u(t))dt \\ \text{over processes } (x(.), u(.)) \text{ satisfying} \\ \dot{x}(t) = f(t, x(t), u(t)) \text{ a.e.,} \\ \phi(t, x(t), u(t)) \in \Phi(t) \text{ and } u(t) \in U(t) \text{ a.e.,} \\ h(t, x(t)) \leq 0 \text{ for all } t \in [a, b], \\ (x(a), x(b)) \in E. \end{cases}$$

The following theorem provides necessary conditions for optimal control problems involving mixed constraints formulated as (3.1), under hypotheses on the mixed constraint data, expressed directly in terms of $\phi(.,.,.)$ and $\Phi(.)$, namely:

(H3)$'$: $\Phi(.)$ is a Lebesgue measurable multifunction taking values non-empty, closed subsets of $\mathbb{R}^\kappa$. $U(.)$ has $\mathcal{L} \times \mathcal{B}^m$ measurable graph and takes values closed sets. $f(t,.,.)$, $\Lambda(t,.,.)$ and $\phi(t,.,.)$ are Lipschitz continuous on a neighbourhood of $(\bar{x}(t) + \epsilon\mathbb{B} \times R(t)) \cap S(t)$, for a.e. $t \in [a, b]$,

in which $(\bar{x}(.)), \bar{u}(.))$ is the process of interest, and (for a given radius multifunction $R(.)$), the following constraint qualification is imposed, in place of the 'bounded slope' condition (BS):

(CQ): There exists a measurable function $M(.) : [a, b] \to \mathbb{R}$ such that, for every $x \in \bar{x}(t) + \epsilon\mathbb{B}$ and $u \in R(t) \cap U(t)$ satisfying $\phi(t, x, u) \in \Phi(t)$,

$$\left. \begin{array}{l} \lambda \in N^L_{\Phi(t)}(\phi(t, x, u)), \ (\alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}^m \\ (\alpha, \beta) \in \partial_L (\lambda \cdot \phi)(t,.,.)(x, u) + \{0\} \times N^L_{U(t)}(u) \end{array} \right\} \implies |\lambda| \leq M(t)|\beta| \quad a.e..$$

Theorem 3.1 (Explicit Mixed Constraints Conditions I) *Let $(\bar{x}(.), \bar{u}(.))$ be a $W^{1,1}$ local minimizer for (P_S), w.r.t. some radius multifunction $R(.)$. Assume that, for some $\epsilon > 0$, hypotheses (H1), (H2), (H3$'$) and (CQ) are satisfied. Assume also*

$$k^f_x(t), \ k^\Lambda_x(t), \ M(t)k^\phi_x(t)[k^f_u(t) + k^\Lambda_u(t)] \ \text{are integrable}$$

and there exists $\eta > 0$ such that $\bar{u}(t) + \eta M(t)k^\phi_x(t)\mathbb{B} \subset R(t)$ a.e..

Then there exists a multiplier set $(p(.), \mu, \lambda^0) \in W^{1,1} \times NBV^+ \times \mathbb{R}^+$ and a bounded μ-measurable function $m(.) : [a, b] \to \mathbb{R}^n$ for which all the assertions of Theorem 2.2 are valid, when (P_S) is interpreted as a

9

special case of (P), in which the w component of the control is absent (write u for v) and S(t) is as defined in (3.1).

If in addition $\phi(t,.,.)$ is strictly differentiable at $(\bar{x}(t),\bar{u}(t))$ and $U(t)$ and $\Phi(t)$ are regular at $\bar{u}(t)$ and at $\phi(t,\bar{x}(t),\bar{u}(t))$ respectively, a.e. $t \in [a,b]$, then the costate inclusion condition (b) in Theorem 2.2 can be replaced by

$$(-\dot{p}(t),0) \in \partial_C\{q(t)\cdot f(t,.,.) - \lambda^0\Lambda(t,.,.) - \langle\lambda(t),\phi(t,.,.)\rangle\}(\bar{x}(t),\bar{u}(t)) - \{0\}\times N^C_{U(t)}(\bar{u}(t)) \ a.e., \quad (3.3)$$

in which $\lambda(.) : [a,b] \to \mathbb{R}^\kappa$ is some measurable function that satisfies

$$\lambda(t) \in N^C_{\Phi(t)}(\phi(t,\bar{x}(t),\bar{u}(t))) \ a.e.. \quad (3.4)$$

Furthermore, for $M(.)$ as in (CQ),

$$|\lambda(t)| \le M(t)\{|q(t)|k^f_u(t) + \lambda^0 k^\Lambda_u(t)\} \ a.e..$$

A proof of Theorem 3.1 is given in Section 5.

Comments.

(i): As we have earlier observed, the standard formulation of mixed constraints, in the form of a combination of functional equality and inequality constraints, is captured by choosing $\phi(.,.,.)$ and $\Phi(.)$ according to (3.1) and (3.2). In this case, the necessary conditions of Theorem 3.1 take a familiar form, involving a multiplier function $\lambda(.) = (\lambda^1(.),\lambda^2(.))$, partitioned into multiplier functions associated with the inequality constraint, $\lambda^1(.)$, and with the equality constraint, $\lambda^2(.)$. If the data is smooth, condition (3.3) becomes the costate equation with added multiplier terms

$$-\dot{p}(t) = q(t)\cdot f_x(t,\bar{x}(t),\bar{u}(t)) - \lambda^1(t)\cdot\phi^1_x(t,\bar{x}(t),\bar{u}(t)) - \lambda^2(t)\cdot\phi^2_x(t,\bar{x}(t),\bar{u}(t))$$

(as, for example, in [8]) and (3.4) is simply the complementary slackness condition, relating to the inequality constraint multiplier $\lambda^1(.)$:

$$\lambda^1_i(t) \ge 0 \text{ and } `\phi^1_i(t,\bar{x}(t),\bar{u}(t)) < 0 \implies \lambda^1_i(t) = 0' \text{ for } i = 1,\ldots,\kappa_1, \ a.e..$$

(ii): Necessary conditions for nonsmooth optimal control problems with mixed constraints formulated as '$\phi(t,x(t),u(t)) \in \Phi(t)$' and in which constraint qualification (CQ) is invoked, are given in [5]. Theorem 2.2 extends [5, Theorem 4.3] to allow also for the presence of a pure state constraint $(h(t,x(t)) \le 0)$.

Typically in the earlier mixed constraints literature, when $\Phi(.)$ is given by (3.2), necessary conditions are derived under hypotheses on gradients of the function $\phi(t,x,u)$ w.r.t. the u variable, which ensure satisfaction of hypothesis (CQ). An exception is the set of necessary conditions due to Schwarzkopf [17], [18] in which hypotheses imposed on the gradients of ϕ are replaced by a convexity hypothesis on the generalized velocity sets involved, together with a surjectivity hypothesis on the mixed constraint functional $\phi(.,.,.)$.

The following theorem provides necessary conditions for optimal control problem (P_S), involving both a mixed constraint and an implicit control constraint under regularity hypotheses on the data (labelled (H3″) that replace (H3′)), and under hypotheses imposed on the mixed constraint data (labelled (SC) 'Schwarzkopf Condition'), which retain the character of the hypotheses in [17], namely

(H3″) (a) the function $(t,u) \mapsto (f(t,x,u),\phi(t,x,u))$ is $\mathcal{L}\times\mathcal{B}^m$ measurable for each x, $U(.)$ has $\mathcal{L}\times\mathcal{B}^m$ measurable graph. $\Phi(.)$ is a Lebesgue measurable multifunction taking values non-empty, closed subsets of $\mathbb{R}^k$,

(b) there exists a function $k_f(.) \in L^1$ such that, for every $u \in U(t)$ the function $f(t, ., u)$ is Lipschitz continuous with Lipschitz constant $k_f(t)$ on $\bar{x}(t) + \epsilon \mathbb{B}$, a.e.,

(c) there exists a constant k_ϕ such that for a.e. t, every $u \in U(t)$, the function $\phi(t, ., u)$ is Lipschitz continuous with Lipschitz constant k_ϕ on $\bar{x}(t) + \epsilon \mathbb{B}$, and this function is strictly differentiable at $\bar{x}(t)$,

and

(SC) There exists an integrable function $r(.)$, and a constant $\delta > 0$ such that the multifunction

$$U_r(t) := \{u \in U(t) : |f(t, \bar{x}(t), u)| \le r(t) \text{ a.e. }\}$$

satisfies

$$\phi(t, \bar{x}(t), \bar{u}(t)) + \delta \mathbb{B} \subset \phi(t, \bar{x}(t), U_r(t)) - \mathcal{T}_t^\eta \quad \text{a.e.} \tag{3.5}$$

where, for each t, $\mathcal{T}_t^\eta$ is a subset of $\mathbb{R}^\kappa$ satisfying

$$\mathcal{T}_t^\eta \subset \bigcap_{z \in S_t^\eta} \left(T^C_{\Phi(t)}(z) \right)$$

in which

$$S_t^\eta := \{z \in \Phi(t) \; : \; |z - \phi(t, \bar{x}(t), \bar{u}(t))| \le \eta \times (1 + d_r(t))\}$$

for some $\eta > 0$. Here $d_r(.)$ is the (possibly infinite valued) function

$$d_r(t) := 2 \cdot \sup\{|\phi(t, \bar{x}(t), u)| \; : \; u \in U_r(t)\}. \tag{3.6}$$

(Notice that a possible choice of $\mathcal{T}_t^\eta$ in (SC) is $\mathcal{T}_t^\eta = \{0\}$. But other choices may be possible, such as that discussed in Comment (i) below, that make condition (3.5) less restrictive.)

Define the extended Hamiltonian by

$$H^{\lambda^0}_{(E)}(t, x, u, p, \lambda) = H^{\lambda^0}(t, x, u, p) - \langle \lambda, \phi(t, x, u) \rangle.$$

Theorem 3.2 (Explicit Mixed Constraints Conditions II) *Let $(\bar{x}(t), \bar{u}(t))$ be a $W^{1,1}$ local minimizer for (P_S). Assume, for some $\epsilon > 0$, (H1), (H2), (H3″) and (SC) are satisfied. Assume also*

(C): $\{(f(t, x, u), \Lambda(t, x, u), \phi(t, x, u)) : u \in U(t)\}$ is convex, for all $x \in \bar{x}(t) + \epsilon \mathbb{B}$, a.e. $t \in [a, b]$.

Then there exists a multiplier set $(p(.), \mu, \lambda(.), \lambda^0)$, in which $(p(.), \mu, \lambda^0) \in W^{1,1} \times NBV^+ \times \mathbb{R}^+$ and $\lambda(.) : [a, b] \to \mathbb{R}^\kappa$ is an integrable function, such that

(a) $\lambda^0 + \|p\|_{L^\infty} + \|\mu\|_{T.V.} = 1$.

(b) $-\dot{p}(t) \in \partial_C H^{\lambda^0}_{(E)}(t, ., \bar{u}(t), q(t), \lambda(t))(\bar{x}(t)) \quad$ a.e.

(c) $m(t) \in \text{co } \partial_x^> h(t, \bar{x}(t)) \; \mu\text{-a.e. and } \text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$.

(d) $(q(a), -q(b)) \in \lambda^0 \partial_L \ell(\bar{x}(a), \bar{x}(b)) + N^L_E(\bar{x}(a), \bar{x}(b))$.

(e) $H^{\lambda^0}_{(E)}(t, \bar{x}(t), \bar{u}(t), q(t), \lambda(t)) \ge H^{\lambda^0}_{(E)}(t, \bar{x}(t), u, q(t), \lambda(t)) \text{ for all } u \in U(t) \text{ a.e.,}$

in which $q(.)$ is defined by (2.3) and $\lambda(t) \in N^C_{\Phi(t)}(\phi(t, \bar{x}(t), \bar{u}(t))) \text{ a.e.}$.

If the 'convexity' hypothesis (C) is replaced by

(C+L): There exists $\sigma_0 > 0$ such that, for a.e. $t \in [a, b]$, $\Phi(t)$ and $U(t)$ are closed, convex sets, $U(t) \subset \sigma_0 \mathbb{B}$ and $\phi(t, x, .)$ is an affine function, i.e.

$$\phi(t, x, u) := \theta^0(t, x) + \sum_{k=1}^m u^i \theta^i(t, x), \text{ for some functions } \theta^i(t, x) : [a, b] \times \mathbb{R}^n \to \mathbb{R}^\kappa, i = 0, \ldots, m.$$

Then the above assertions remain valid, except that (b) is replaced by (3.3) *and the 'extended' Weierstrass condition (e) is replaced by the weaker Weierstrass condition of Theorem 2.2, namely*

$$H^{\lambda^0}(t, \bar{x}(t), \bar{u}(t), q(t)) \geq H^{\lambda^0}(t, \bar{x}(t), u, q(t)) \text{ for all } u \in U(t) \cap \{u' \, : \, \phi(t, \bar{x}(t), u') \in \Phi(t)\}, \text{ a.e. } t \in [a, b].$$

A proof of Theorem 3.1 is given in Section 5.

Comments

(i): Necessary conditions for non‑smooth optimal control problems with mixed constraints, under Schwartzkopf type surjectivity hypotheses, are given in [6]. Theorem 3.2 extends [6, Theorem 3.2] to allow also for pure state constraints. Even when the pure state constraint is absent, Theorem 3.2 refines [6, Theorem 3.2], because the condition (3.5) in hypothesis (SC) now involves the set $\mathcal{T}_t^\eta$. In the most common case when $\Phi(.)$ is given by (3.2) (mixed equality and inequality constraints), we can choose

$$\mathcal{T}_t^\eta = \overbrace{(-\infty, 0] \times \ldots \times (-\infty, 0]}^{\kappa_1} \times \overbrace{\{0\} \times \ldots \times \{0\}}^{\kappa_2}.$$

With this choice, condition (3.5) becomes:

$$\begin{aligned}(\phi^1, \phi^2)(t, \bar{x}(t), \bar{u}(t)) + \delta \cdot (\mathbb{B} \times \mathbb{B}) \subset \\ (\phi^1, \phi^2)(t, \bar{x}(t), U_r(t)) + ([0, +\infty) \times \ldots \times [0, +\infty)) \times (\{0\} \times \ldots \times \{0\}).\end{aligned} \qquad (3.7)$$

In the special case when the equality constraint ('$\phi^2(t, x(t), u(t)) = 0$') is absent, condition (3.7) is simply the Slater type condition: there exists $\delta > 0$ and a measurable function $u(.)$ such that $u(t) \in U_r(t)$ and $\phi_i^1(t, \bar{x}(t), u(t)) \leq -\delta$ a.e..

Notice that the analogous hypothesis in [6, Theorem 3.1] can be interpreted as (SC) when we take $T_t^\eta = \{0\}$, that is, it replaces (3.7) by the more restrictive hypothesis

$$(\phi^1, \phi^2)(t, \bar{x}(t), \bar{u}(t)) + \delta \cdot (\mathbb{B} \times \mathbb{B}) \subset (\phi^1, \phi^2)(t, \bar{x}(t), U_r(t)).$$

The possibility of invoking less restrictive hypotheses, expressed in terms of the tangent cone of $\Phi(t)$ is discussed in [6, Section 3] under the heading 'tangential covering conditions'.

(ii): Theorems 2.2, 3.1 and 3.2 provide three sets of necessary conditions that result from bringing together the 'pure state constraints' necessary conditions of [1] and the techniques of [6] for treating mixed state-control constraints by absorbing them into the dynamic constraint. Similar extensions can be achieved for all the sets of necessary conditions in [6], and elsewhere, for different formulations of optimal control problems involving mixed state-control constraints, to allow for the presence of state constraints, including differential algebraic formulations and calculus of variations formulations.

(iii) Theorem 3.2 under hypothesis (C+L) partially answers an open question raised in [6, pag. 10] regarding the validity of Theorem 3.2 without convexity hypotheses on the extended velocity set.

4   Proof of Theorem 2.2

The proof is an adaptation of the proof of [5, Theorem 2.1], to allow for the presence of state constraints and for the more general radius multifunction here considered. The idea is to reformulate Problem (P) as a differential inclusion optimal control problem with state constraints by absorbing the mixed constraint into the dynamics. We apply known necessary conditions to the reformulated problem, and express these necessary conditions directly in terms of the data for the original problem (P). The main difference is

that we make use of the necessary conditions in [1, Corollary 2.2] for 'pure' state constrained problems, in place of the necessary conditions for pure state constraint free problems employed in [5].

Observe at the outset that we can, without loss of generality, restrict attention to the case when the integral cost is absent, since the full statement can be derived from the special case when $L(.,.,.) \equiv 0$ by means of standard state augmentation techniques. We can modify the Lipschitz constants to ensure that $k_x^f(.,w)$ is positive valued and $k_S(t) \geq 1$ a.e., without violating the hypotheses. We can assume, finally, that

(A1): For every $t \in [a,b]$, $W(t)$ is a finite set.

(A2): There exist an integrable function $C(.) : [a,b] \to \mathbb{R}$ such that, for a.e. $t \in [a,b]$ and $w \in W(t)$:

$$(x,v) \in (\bar{x}(t) + \epsilon \mathbb{B}) \times R(t,w) \cap S(t,w) \Longrightarrow$$
$$|k_x^f(t,w) - k_x^f(t,\bar{w}(t))| + k_S(t)|k_v^f(t,w) - k_v^f(t,\bar{w}(t))| + |f(t,x,(v,w)) - \dot{\bar{x}}(t)| \leq C(t).$$

Reductions to special cases, involving hypotheses $(A1)$ and $(A2)$, introduced in [3], are now standard in the derivation of necessary conditions (See [3, p. 201 et seq.] or [19, Lemma 6.3.1]). Justification for using them depends on introducing finite, inner approximations of the sets $W(t)$ and robustness properties of the necessary conditions under limit taking: the 'state constraint' necessary conditions here considered have the required robustness properties for this procedure to apply.

Let $(\bar{x}(.), \bar{u}(.))$ be a $W^{1,1}$ local minimizer for (P) w.r.t. the radius multifunction $R(.)$. Define

$$k(t) := \max_{w \in W(t)} k_x^f(t,w) + k_S(t) \times \max_{w \in W(t)} k_v^f(t,w) \text{ and } c(t) := k(t)/k_S(t), \tag{4.1}$$

which are integrable by (2.1) and (A2) and since $k_S(t) \geq 1$, and let $F(.,.)$ be the multifunction

$$F(t,x) := \{(f(t,x,u), r(t,u)) : u = (v,w), (x,v) \in S(t,w), w \in W(t)\}. \tag{4.2}$$

in which $r(t,u) := c(t)(u - \bar{u}(t))$.

We remark that $F(t,x)$ is non-empty for all $x \in \bar{x}(t) + \epsilon \mathbb{B}$, a.e. $t \in [a,b]$ following, if required, a reduction in the size of $\epsilon > 0$. This is because we can deduce from the bounded slope condition (BS) and (2.2), with the help of [4, Theorem 3.5.2.], that $S(t,x,w)$ is non-empty for all $w \in W(t)$, $x \in \bar{x}(t) + \epsilon' \mathbb{B}$. a.e. $t \in [a,b]$ (for some suitably small $\epsilon' > 0$).

Consider the following optimal control problem in which the dynamics are modelled as a differential inclusion:

$$(P') \quad \begin{cases} \text{Minimize } \ell(x(a), x(b)) \\ \text{over } (x(.), z(.)) \in W^{1,1} \text{ satisfying} \\ (\dot{x}(t), \dot{z}(t)) \in F(t,x(t)) \text{ a.e. } t \in [a,b], \\ h(t,x(t)) \leq 0 \text{ for all } t \in [a,b], \\ (x(a), x(b)) \in E, \quad z(a) = 0. \end{cases}$$

With the help of measurable select theory, we can show that the feasible F-trajectory $(\bar{x}(t), \bar{z}(t) \equiv 0)$ is a $W^{1,1}$ local minimizer w.r.t. the radius multifunction:

$$R_F(t) := \mathbb{R}^n \times c(t)\left(R(t) - \bar{u}(t)\right).$$

(Note that the 'extra' state variable z is introduced into the dynamic constraint $(\dot{x}, \dot{z}) \in F$ of (P'), in order to derive stratified necessary conditions involving the radius multifunction $R(.)$.) The hypotheses are satisfied under which the necessary conditions of [1, Corollary 2.2] are valid, w.r.t. the $W^{1,1}$ local

13

minimizer $(\bar{x}(.), \bar{z}(.) \equiv 0)$. The key conditions to check, regarding the required properties of $F(.,.)$, are the *bounded slope* condition (with reference to the radius multifunction $R_F(.)$), which is here conveniently expressed in the form: for some possibly readjusted $\epsilon > 0$ and some $k(.) \in L^1$ such that for any $x \in \bar{x}(t) + \epsilon \mathbb{B}$ and and $u \in \mathbb{R}^m$ s.t. $(f(t,x,u), r(t,u)) \in F(t,x) \cap R_F(t)$,

$$(\alpha, (\beta, \tau)) \in N^P_{Gr\,F(t,.)}(x, f(t,x,u), r(t,u)) \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \implies |\alpha| \le k(t)|(|\beta| + |\tau|), \text{ a.e.} \quad (4.3)$$

and the compatibility condition: there exists $\eta > 0$ such that

$$(\dot{\bar{x}}(t), 0) + \eta k(t)\,\mathbb{B} \subset R_F(t) \quad \text{a.e..} \tag{4.4}$$

Condition (4.4) is satisfied when $k(.)$ is chosen according to (4.1), and since $(\bar{u}(t) + \eta k_S(t)\mathbb{B}) \subset R(t)$, which permit us to conclude

$$(\dot{\bar{x}}(t), 0) + \eta k(t)\mathbb{B} \subset (\dot{\bar{x}}(t), 0) + \eta \, (\mathbb{R}^n \times c(t)k_S(t)\mathbb{B}) \subset \mathbb{R}^n \times c(t)(R(t) - \bar{u}(t)) = R_F(t) \text{ a.e.}$$

as required. Condition (4.3) is also satisfied, in view of the following lemma:

Lemma 4.1 *Fix* $w \in W(t)$ *and* $(x,v) \in ((\bar{x}(t) + \epsilon\mathbb{B}) \times R(t,w)) \cap S(t,w)$, *where, consistent with our notation,* $R(t,w) := \{v : (v,w) \in R(t)\}$. *Take any* $(\alpha, \beta, \tau) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m$ *such that* $(\alpha, \beta, \tau) \in N^L_{GrF(t,.)}(x, f(t,x,v,w), r(t,v,w))$. *Then,*

$$|\alpha| \le k(t)(|\beta| + |\tau|) \quad a.e.$$

Furthermore

$$(\alpha, 0) \in \partial_L \left\{ -\langle \beta, f(t,.,.,w)\rangle - \langle \tau, r(t,.,w)\rangle + 2k(t)[|\beta| + |\tau|]d_{S(t,w)}(.,.)\right\}(x,v).$$

This lemma in proved in [5]. (The proof of [5, Prop. 9.1] covers the case when the w-control is absent and the analysis on [5, pp. 23-24] extends the estimates to allow for w-dependent data). Notice we have taken advantage of the supplementary hypothesis '$k_S(t) \ge 1$' to simplify the estimates in [5, Prop. 9.1]. We deduce from [1, Corollary 2.2] the following information: there exist an absolutely continuous function $p(.) \in W^{1,1}([a,b]; \mathbb{R}^n)$, $\lambda^0 \ge 0$, $\mu(.) \in NBV^+[a,b]$ and a μ-integrable function $m(.)$, such that

(i): $\lambda^0 + \|p\|_{L^\infty} + \|\mu\|_{T.V.} = 1$,

(ii): $\dot{p}(t) \in co\left\{\eta : (\eta, q(t), 0) \in N^L_{GrF(t,.)}(\bar{x}(t), \dot{\bar{x}}(t), 0)\right\}$ a.e.,

(iii): $(q(a), -q(b)) \in \lambda^0 \partial_L \ell(\bar{x}(a), \bar{x}(b)) + N^L_E(\bar{x}(a), \bar{x}(b))$,

(iv): For any $(e_1, e_2) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $(e_1, e_2) \in F(t, \bar{x}(t)) \cap R_F(t)$,

$$\langle q(t), \dot{\bar{x}}(t)\rangle \ge \langle q(t), e_1\rangle + \langle p_0, e_2\rangle \quad \text{a.e. } t \in [a,b],$$

(v): $m(t) \in co\,\partial_x^> h(t, \bar{x}(t))$ μ-a.e. and $\text{supp}\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}$,

in which $q(.) : [a,b] \to \mathbb{R}^n$ is as defined in (2.3). By condition (ii) and Lemma 4.1,

$$(\eta, 0) \in \partial_L\{-\langle q(t), f(t,.,.,\bar{w}(t))\rangle + 2k(t)|q(t)|d_{S(t,\bar{w}(t))}(.,.)\}(\bar{x}(t), \bar{v}(t)).$$

(Note that the costate associated with the z variable is identically zero and so drops out of the above conditions.) Reviewing these conditions, we see that the proof of the theorem will be complete if we can show that condition (iv) implies the Weierstrass condition (e) and condition (ii) implies the costate inclusion (b). But (iv) tells

$$\langle q(t), \dot{\bar{x}}(t)\rangle \ge \langle q(t), f(t, \bar{x}(t), u)\rangle$$

for a.e. t and all $u = (w, v)$ such that $(\bar{x}(t), v) \in S(t, w)$, $w \in W(t)$ and

$$(f(t, \bar{x}(t), u), r(t, u)) \in R_F(t) = \mathbb{R}^n \times c(t)(R(t) - \bar{u}(t)).$$

This is precisely condition (e) expressed in terms of the radius multifunction $R(.)$, since this last condition can be equivalently stated '$c(t)(u - \bar{u}(t)) \in c(t)(R(t) - \bar{u}(t))$', i.e '$u \in R(t)$' since $c(t) > 0$, a.e..

Finally, we look at the implications of (ii). Using the facts that $\partial_C = co\, \partial_L$ and, given two Lipschitz functions f and g, we have $\partial_C(-f) = -\partial_C f$ and $\partial_C(f + g) \subset \partial_C f + \partial_C g$ we conclude

$$(-\dot{p}(t), 0) \in \partial_C \{\langle q(t), f(t, ., ., \bar{w}(t)) \rangle\}(\bar{x}(t), \bar{v}(t)) - \partial_C \{2k(t)|q(t)|d_{S(t, \bar{w}(t))}(., .)\}(\bar{x}(t), \bar{v}(t)), \qquad (4.5)$$

which implies condition (b) in the theorem statement, since the Clarke generalized gradient is positively homogeneous and, for fixed t, $\partial_C d_{S(t, \bar{w}(t))}(\bar{x}(t), \bar{v}(t)) \subset N^C_{S(t, \bar{w}(t))}(\bar{x}(t), \bar{v}(t))$.

5 Proof of Theorems 3.1 and 3.2

Proof of Theorem 3.1.

Problem (P_S) is a special case of (P), in which the w-component of the control variable is absent (we indentify v with u) and $S(t) = \{(x, u) : \tilde{\phi}(t, x, u) \in \tilde{\Phi}(t)\}$, in which

$$\tilde{\phi}(t, x, u) = (\phi(t, x, u), u) \text{ and } \tilde{\Phi}(t) = \Phi(t) \times U(t).$$

Proving the first part of the theorem requires us to confirm that the data for the special case satisfies the hypotheses of Theorem 2.2, with reference to the W^{11} local minimizer $(\bar{x}(.), \bar{u}(.))$ and radius multifunction $R(.)$. We attend only to the verification of the key 'bounded slope' hypothesis (BS) of Theorem 2.2, with the above identification of $S(t)$. A straightforward refinement of the analysis in [5, Proof of Prop. 4.2] (to taken account of the special structure of $\tilde{\phi}(., ., .)$ and $\tilde{\Phi}(.)$) permits us to deduce the following: under the hypotheses of Theorem 3.1, for a.e. t, and $(x, u) \in S(t)$ such that $x \in \bar{x}(t) + \epsilon \mathbb{B}$, $u \in R(t)$, the relation $(\alpha, \beta) \in N^P_{S(t)}(x, u)$ implies

$$|\alpha| \leq M(t)k^\phi_x(t)|\beta|, \qquad (5.1)$$

where $M(t)$ is as in hypothesis (CQ); furthermore, for a.e. $t \in [a, b]$ there exists $\lambda \in N^L_{\Phi(t)}(\phi(t, x, u)$ such that

$$(\alpha, \beta) \in \partial_L \langle \lambda, \phi(t, ., .) \rangle(x, u) + \{0\} \times N^L_{U(t)}(u). \qquad (5.2)$$

We deduce from property (5.1) that (BS) is indeed satisfied, when we identify $k_S(t) = M(t)k^\phi_x(t)$.

Now suppose that $\phi(t, ., .)$ is strictly differentiable at $(\bar{x}(t), \bar{u}(t))$ and that $U(t)$ and $\Phi(t)$ are regular at $\bar{u}(t)$ and at $\phi(t, \bar{x}(t), \bar{u}(t))$. Making use of relation (5.2) which, at the point $(x, u) = (\bar{x}(t), \bar{u}(t))$, becomes (by strict differentiability and regularity)

$$(\alpha, \beta) \in \nabla \langle \lambda, \phi(t, ., .) \rangle(x, u) + \{0\} \times N^C_{U(t)}(u),$$

and also the sum rule for Clarke generalized gradients, we see that the costate inclusion (b) of Theorem 2.2 can be replaced by

$$(-\dot{p}(t), 0) \in \partial_C \{q(t) \cdot f(t, ., .) - \lambda^0 \Lambda(t, ., .) - \langle \lambda(t), \phi(t, ., .) \rangle\}(\bar{x}(t), \bar{u}(t)) - \{0\} \times N^C_{U(t)}(\bar{u}(t)),$$

for some function $\lambda(.)$, a.e. t. (The function $\lambda(.)$ can be chosen to be measurable.) We have arrived at the costate inclusion of Theorem 3.1. Finally note that, any $(\zeta_1, \zeta_2) \in \partial_C \{q(t) \cdot f(t, ., .) - \lambda^0 \Lambda(t, ., .)\}\}(\bar{x}(t), \bar{u}(t))$ satisfies $|\zeta_2| \leq |q(t)|k^f_u(t) + \lambda^0 k^\Lambda_u(t)$. Examination of the adjoint inclusion in combination with hypothesis (CQ) yields the estimate

$$|\lambda(t)| \leq M(t) \left(|q(t)|k^f_u(t) + \lambda^0 k^\Lambda_u(t)\right), \quad a.e..$$

15

Note that the regularity hypothesis of the set $U(t)$ and $\Phi(t)$ at $\bar{u}(t)$ and at $\phi(t, \bar{x}(t), \bar{u}(t))$, a.e., is required in order to make use of hypothesis (CQ).

Proof of Theorem 3.2.

We first confirm the assertions of the theorem excluding the final assertion. Accordingly assume (C), in addition to the preceding hypotheses in the theorem statment. Following [6], we obtain, as a corollary to Theorem 2.2, necessary conditions for a special case of (P), in which the data is affine w.r.t. the control variable. These necessary conditions are used, subsequently, to derive necessary conditions for problem (P_S) under the Schwartzkopf type surjectivity hypothesis in place of a constraint qualification such as (CQ), governing the mixed constraint data. The difference is that our analysis takes as starting point the state constrained necessary conditions of Theorem 2.2 of this paper in place of the state constraint-free conditions of [6, Theorem 3.2]. The special case of (P) that we need to consider is:

$$(P'') \quad \begin{cases} \text{Minimize } \ell(x(a), x(b)), \\[2mm] \dot{x}(t) = \sum_{i=0}^{N} c_i(t) g^i(t, x(t)), \quad a.e. \ \ t \in [a, b], \\[2mm] \sum_{i=0}^{N} c_i(t) \theta^i(t, x(t)) \in \Phi(t), \quad c \in \Sigma, \quad a.e. \ \ t \in [a, b], \\[2mm] h(t, x(t)) \leq 0, \quad \text{for all } t \in [a, b], \\[2mm] (x(a), x(b)) \in E. \end{cases}$$

Here $g^i(.,.) : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\theta^i(.,.) : [a, b] \times \mathbb{R}^n \to \mathbb{R}^\kappa$ (for $i = 0, \ldots, N$), $\ell(.,.) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $h(.,.) : [a, b] \times \mathbb{R}^n \to \mathbb{R}$ are given functions, and $\Phi \subset \mathbb{R} \times \mathbb{R}^\kappa$, $\Sigma \subset \mathbb{R}^{N+1}$, $E \subset \mathbb{R}^n \times \mathbb{R}^n$ are given sets.

To fit (P'') to the framework of problem (P) (when the control variable u is not partitioned, i.e. $u = (v, w)$ comprises only the variable v, which we write here as c), we take the dynamic constraint to be $\dot{x}(t) = f(t, x(t), c(t))$ with

$$f(t, x, c) = \sum_{i=0}^{N} c_i g^i(t, x)$$

and the sets $S(t)$ in the 'mixed constraint' $(x(t), c(t)) \in S(t)$ to be

$$S(t) = \{(x, c) : (\phi(t, x, c), c) \in \Phi(t) \times \Sigma\},$$

in which

$$\phi(t, x, c) = \sum_{i=0}^{N} c_i \theta^i(t, x) \ .$$

Lemma 5.1 *Let $(\bar{x}(.), \bar{c}(.))$ be a $W^{1,1}$ local minimizer for (P'') w.r.t. the radius multifunction $R(t) = \bar{c}(t) + R\mathbb{B}$ a.e., for some $R > 0$. Assume that, for some $\epsilon > 0$,*

(L1): $\ell(.,.)$ is Lipschitz continuous on a neighborhood of $(\bar{x}(a), \bar{x}(b))$. The function $h(.,.)$ satisfies hypothesis (H2). E is closed. Σ is a compact, convex set.

(L2): The functions $g^i(.,.)$ and $\theta^i(.,.)$ are $\mathcal{L} \times \mathcal{B}^n$ measurable. There exist an integrable function $k^g(.)$ and a constant k^θ such that $g^i(t, .)$ and $\theta^i(t, .)$ are Lipschitz continuous on $\bar{x}(t) + \epsilon\mathbb{B}$, a.e. $t \in (a, b)$, $i = 0, \ldots, N$. $t \to g^i(t, \bar{x}(t))$ is integrable and $\theta^i(t, .)$ is strictly differentiable at $\bar{x}(t)$, for a.e. $t \in [a, b]$. $\Phi(.)$ is a Lebesgue measurable multifunction taking values non-empty, closed subsets of $\mathbb{R}^\kappa$.

(BS'): There exists a constant $M > 0$ such that, for a.e. $t \in [a, b]$, the following condition is satisfied: given any $x \in \bar{x}(t) + \epsilon \mathbb{B}$, $c \in \Sigma \cap (\bar{c}(t) + R\mathbb{B})$ such that $\phi(t, x, c) \in \Phi(t)$, $\lambda \in N^C_{\Phi(t)}(\phi(t, x, c))$ and $\gamma \in N^C_{\Sigma}(c)$, we have

$$\beta = (\lambda \cdot \theta^0(t, x), \ldots, \lambda \cdot \theta^N(t, x)) + \gamma \implies |\lambda| \leq M|\beta|. \tag{5.3}$$

Then there exist $p(.) \in W^{1,1}$, $\lambda^0 \geq 0$, $\mu(.) \in NBV^+[a, b]$, a μ-integrable function $m(.) \in \partial^>_x h(t, \bar{x}(t))$ μ-a.e. and a integrable function $\lambda : [a, b] \to \mathbb{R}^\kappa$ satisfying $\lambda(t) \in N^C_{\Phi(t)}(\phi(t, \bar{x}(t), \bar{c}(t)))$ such that

(i): $(p(.), \lambda^0, \mu) \neq (0, 0, 0)$

(ii): $-\dot{p}(t) \in \partial_C \{\langle q(t), f(t, ., \bar{c}(t)) \rangle - \langle \lambda(t), \phi(t, ., \bar{c}(t)) \rangle\}(\bar{x}(t))$ a.e.,

(iii): $\langle q(t), f(t, \bar{x}(t), \bar{c}(t)) \rangle - \langle \lambda(t), \phi(t, \bar{x}(t), \bar{c}(t)) \rangle$

$$\geq \langle q(t), f(t, \bar{x}(t), c) \rangle - \langle \lambda(t), \phi(t, \bar{x}(t), c) \rangle \text{ for all } c \in \Sigma, \text{ a.e.,}$$

(iv): $(q(a), -q(b)) \in \lambda^0 \partial_L \ell(\bar{x}(a), \bar{x}(b)) + N^L_E(\bar{x}(a), \bar{x}(b))$,

in which $\quad q(t) := \begin{cases} p(t) + \int_{[a,t]} m(s)\, \mu(ds) & \text{if } t \in (a, b] \\ p(a) & \text{for } t = a. \end{cases}$

The lemma is proved as [6, Cor. 2.2], in the special case that there are no state constraints, by showing that the data for problem (P''), regarded as a special case of P, satisfies the hypotheses of the state constraint free version of Theorem 2.2 and by applying the necessary conditions of this theorem to (P''). The analysis in [6] transcribes directly and without alteration, when the full version of Theorem 2.2 is substituted for the state constraint-free version, to furnish a proof of the lemma.

We are now ready to prove Theorem 3.2. Consider then a $W^{1,1}$ local minimizer $(\bar{x}(t), \bar{u}(t))$ for (P_S) under the hypotheses of Theorem 3.2. We notice immediately that we can assume, without loss of generality, that $\phi(t, \bar{x}(t), \bar{u}(t)) \equiv 0$ since replacing $\phi(t, x, u)$ by $\phi(t, x, u) - \phi(t, \bar{x}(t), \bar{u}(t))$ and $\Phi(t)$ by $\Phi(t) - \phi(t, \bar{x}(t), \bar{u}(t))$ ensures that this condition is satisfied, yet has no effect on the assertions of the theorem or validity of the hypotheses under which it applies. The use of state augmentation techniques permit us to restrict attention to the case when $\Lambda(., ., .) \equiv 0$.

Write e_i, $i = 1, \ldots, \kappa$, for the canonical basis vectors of R^κ. In consequence of (SC), there exist controls $u^+_i(.)$ and $u^-_i(.)$ (we can choose them to be measurable) such that

$$\phi(t, \bar{x}(t), u^+_i(t)) - \phi(t, \bar{x}(t), \bar{u}(t)) \in \delta e_i + \mathcal{T}^\eta_t \tag{5.4}$$
$$\phi(t, \bar{x}(t), u^-_i(t)) - \phi(t, \bar{x}(t), \bar{u}(t)) \in -\delta e_i + \mathcal{T}^\eta_t \tag{5.5}$$
$$|f(t, \bar{x}(t), u^+(t)) - f(t, \bar{x}(t), \bar{u}(t))| \leq r(t) + |\dot{\bar{x}}(t)|,$$
$$|f(t, \bar{x}(t), u^-(t)) - f(t, \bar{x}(t), \bar{u}(t))| \leq r(t) + |\dot{\bar{x}}(t)|,$$

for $i = 1, \ldots, \kappa$, where $r(.)$, δ, and $\mathcal{T}^\eta$ are the same as in hypothesis (SC). In view of hypothesis $(H3)''$, we can further arrange, by reducing the size of $\epsilon > 0$ if necessary, that

$$|\phi(t, x, u^+_i(t)) - \phi(t, \bar{x}(t), u^+_i(t))| \leq \delta/(12\kappa), |\phi(t, x, u^-_i(t)) - \phi(t, \bar{x}(t), u^-_i(t))| \leq \delta/(12\kappa), \tag{5.6}$$

and

$$|\phi(t, x, \bar{u}(t)) - \phi(t, \bar{x}(t), \bar{u}(t))| \leq \delta/(12\kappa), \tag{5.7}$$

for all $x \in \bar{x}(t) + \epsilon \mathbb{B}$, $j = 1, \ldots, \kappa$, a.e.

Fix any $d > 0$ and define

$$U^d(t) := \{u \in U(t) : |(f(t, \bar{x}(t), u) - f(t, \bar{x}(t), \bar{u}(t)), \phi(t, \bar{x}(t), u))| \leq 1/d\}$$
$$h'(t, p, \lambda, u) := \langle p, f(t, \bar{x}(t), u) - f(t, \bar{x}(t), \bar{u}(t)) \rangle - \langle \lambda, \phi(t, \bar{x}(t), u) \rangle$$
$$H^d(t, p, \lambda) := \sup\{h'(t, p, \lambda, u) : u \in U^d(t)\}.$$

Let $\{(p_j, \lambda_j) \,|\, j = 2\kappa + 1, \ldots, N_d\}$ be a collection of vectors in $\mathbb{B}$ such that

$$\mathbb{B} \subset \bigcup_j \left((p_j, \lambda_j) + d^2\mathbb{B}\right). \tag{5.8}$$

We can choose measurable controls $v_j(.)$, $j = 2\kappa + 1, \ldots, N_d$ with values in $U^d(.)$ and satisfying

$$H^d(t, p_j, \lambda_j) - d < h'(t, p_j, \lambda_j, v_j(t)) \quad \text{a.e..}$$

Consider now problem (P'') when $\Phi(.)$, $\ell(.,.)$, $h(.)$ and E are as in problem (P_S),

$$
\begin{aligned}
g^0(t, x) &:= f(t, x, \bar{u}(t)), & \theta^0(t, x) &= \phi(t, x, \bar{u}(t)) \\
g^i(t, x) &:= f(t, x, u^+_{(i+1)/2}(t)) - f(t, x, \bar{u}(t)), & \theta^i(t, x) &= \phi(t, x, u^+_{(1+i)/2}(t)) - \phi(t, x, \bar{u}(t)), \\
& & & \quad i = 1, 3, \ldots, 2\kappa - 1 \\
g^i(t, x) &:= f(t, x, u^-_{i/2}(t)) - f(t, x, \bar{u}(t)), & \theta^i(t, x) &= \phi(t, x, u^-_{i/2}(t)) - \phi(t, x, \bar{u}(t)), \\
& & & \quad i = 2, 4, \ldots, 2\kappa \\
g^i(t, x) &:= f(t, x, v_i(t)) - f(t, x, \bar{u}(t)), & \theta^i(t, x) &= \phi(t, x, v_i(t)) - \phi(t, x, \bar{u}(t)), \\
& & & \quad i = 2\kappa + 1, \ldots, N_d,
\end{aligned}
$$

and

$$\Sigma := \left\{c = (c_0, \ldots, c_{N_d}) : c_0 = 1, c_i \geq 0, \ i = 1, 2, \ldots, N_d \text{ and } \sum_1^{N_d} c_i \leq 1\right\}.$$

Making use of the 'convexity' hypothesis (C) in Theorem 3.2 and appealing to measurable selection theory, we can show that $(\bar{x}(.), \bar{c}(t) \equiv (1, 0, \ldots, 0))$ is a $W^{1,1}$ local minimizer w.r.t. to the constant radius multifunction $\bar{c}(t) + R\mathbb{B}$ for any positive constant $R > 0$. Following reduction in the size of ϵ, we can choose $R > 0$ to satisfy:

$$0 < R < 1 \quad \text{and} \quad \max\{\epsilon k_\phi(1 + 2R), R\} < \eta, \tag{5.9}$$

in which $\eta > 0$ is the constant of Hypothesis (SC).

Our aim now is to apply the necessary conditions of Lemma 5.1, with respect to the minimizer $(\bar{x}(.), \bar{c}(.))$, for radius multfunction $R(.) \equiv R\mathbb{B}$, checking first that the relevant hypotheses are satisfied. Consider hypothesis (BS$'$). Take $(x, c) \in (\bar{x}(t) + \epsilon\mathbb{B}) \times \Sigma$ such that $\sum_{i=1}^{N_d} c_i < R$, $\sum_{i=0}^{N_d} c_i\theta^i(t, x) \in \Phi(t)$, $\lambda \in N^C_{\Phi(t)}(\phi(t, x, c))$, $\gamma \in N^C_\Sigma(c)$ and

$$\beta = \left(\langle \lambda, \theta^0(t, x) \rangle, \ldots, \langle \lambda, \theta^{N_d}(t, x) \rangle\right) + \gamma. \tag{5.10}$$

We must show, for a.e. $t \in [a, b]$,

$$|\lambda| \leq M|\beta|, \tag{5.11}$$

for some M independent of (x, c) and γ. Notice that γ is a vector of non-positive numbers because $\gamma \in N_\Sigma(c)$ and $\sum_{i=1}^{N_d} c_i < 1$. Examination of the $i = 1$ and 2 components of the $N_d + 1$-vector relation (5.10), in the light of (5.4), (5.5), (5.6) and (5.7), yields:

$$\beta_1 = \delta\lambda_1 + \mu_1|\lambda| + \langle \lambda, \xi_1 \rangle + \gamma_1 \quad \text{and} \quad \beta_2 = -\delta\lambda_1 + \mu_2|\lambda| + \langle \lambda, \xi_2 \rangle + \gamma_2, \tag{5.12}$$

for some numbers μ_i, γ_i such that $|\mu_i| \le \delta/6\kappa$ and $\gamma_i \le 0$ for $i = 1, 2$, and vectors $\xi_i \in \mathcal{T}_t^\eta$ for $i = 1, 2$. Notice however that, in consequence of (5.9),

$$\left| \sum_{i=0}^{N_d} c_i \theta^i(t, x) - \theta_0(t, \bar{x}(t)) \right| \le \epsilon k_\phi + (\sum_{i=1}^{N_d} c_i)(2\epsilon k_\phi + d_r(t)) \le \epsilon k_\phi(1 + 2R) + R d_r(t)$$

$$\le \eta \cdot (1 + d_r(t))$$

in which $d_r(.)$ is the function in (3.6).

It follows from the definition of $\mathcal{T}_t^\eta$ as an intersection of tangent cones to $\Phi(t)$ over base points in the set S_t^η, and the fact that $\sum_{i=0}^{N_d} c_i \theta^i(t, x)$ lies in this set, that

$$\xi_i \in \mathcal{T}_t^\eta \subset T_{\Phi(t)}^C (\sum_{i=0}^{N_d} c_i \theta^i(t, x)) \qquad \text{for } i = 1, 2.$$

Since $\lambda \in N_{\Phi(t)}^C(\sum_{i=0}^{N_d} c_i \theta^i(t, x))$, we have $\langle \lambda, \xi_i \rangle \le 0$, for $i = 1, 2$. From (5.12) then,

$$\beta_1 = \delta \lambda_1 + \mu_1 |\lambda| + \tilde{\gamma}_1 \quad \text{and} \quad \beta_2 = -\delta \lambda_1 + \mu_2 |\lambda| + \tilde{\gamma}_2, \tag{5.13}$$

for some $\tilde{\gamma}_i \le 0$, $i = 1, 2$. Adding the two preceding relations gives $\beta_1 + \beta_2 = (\mu_1 + \mu_2)|\lambda| + \tilde{\gamma}_1 + \tilde{\gamma}_2$. Noting that $\tilde{\gamma}_i$'s are non-positive, and using the bounds on the μ_i's, we conclude $|\tilde{\gamma}_i| \le 2|\beta| + (\delta/3\kappa)|\lambda|$, for $i = 1, 2$. But then, by (5.13),

$$|\lambda_i| \le |\beta|/\delta + (1/6\kappa)|\lambda| + 2|\beta|/\delta + (1/3\kappa)|\lambda| = 3|\beta|/\delta + (1/2\kappa)|\lambda| \text{ for } i = 1, 2.$$

Analogous bounds can be established for $\lambda_2, \ldots, \lambda_{2\kappa}$. Summing such bounds over i, yields $|\lambda| \le (3\kappa/\delta)|\beta| + (1/2)|\lambda|$, whence $|\lambda| \le (6\kappa/\delta)|\beta|$. We have confirmed (5.11) with $M = (6\kappa/\delta)$.

Verification of the remaining hypotheses is straighforward. Applying Lemma 5.1, we deduce immediately that, for some $(p(.), \lambda^0, \mu)$ and some integrable $\lambda(.)$:

$$\lambda^0 + \|p\|_{L^\infty} + \|\mu\|_{T.V.} = 1 \tag{5.14}$$

$$-\dot{p}(t) \in \partial_C \{ \langle q(t), f(t, ., \bar{u}(t)) \rangle - \langle \lambda(t), \phi(t, ., \bar{u}(t)) \rangle \}(\bar{x}(t)) \quad \text{a.e.} \tag{5.15}$$

$$m(t) \in \text{co} \, \partial_x^> h(t, \bar{x}(t)) \ \mu\text{-a.e. and } \ supp\{\mu\} \subset \{t : h(t, \bar{x}(t)) = 0\}, \tag{5.16}$$

$$(q(a), -q(b)) \in \lambda^0 \partial_L \ell(\bar{x}(a), \bar{x}(b)) + N_E^L(\bar{x}(a), \bar{x}(b)). \tag{5.17}$$

in which $q(.)$ is defined by (2.3). Following the analysis, in [6] we deduce from the Weierstrass condition in the Lemma that the 'multiplier' $\lambda(.)$ is bounded according to:

$$|\lambda(t)| \le 12 \cdot (\kappa^2/\delta) \cdot \|q(.)\|_{L^\infty} \cdot (r(t) + |\dot{\bar{x}}(t)|), \tag{5.18}$$

The Weierstrass condition and (5.8) can also be shown to imply (see [6]):

$$\langle q(t), f(t, \bar{x}(t), \bar{u}(t)) \rangle - \langle \lambda(t), f(t, \bar{x}(t), u) \rangle$$
$$\ge -3d \left(12\kappa^2/\delta \times (r(t) + |\dot{\bar{x}}(t)|)(k_h + 1) + 1 \right) \quad \text{for each } u \in U^d(t), \text{ a.e.} \tag{5.19}$$

(Here, $r(.)$ and δ are as in hypothesis (SC) and k_h is as in hypothesis (H2)). Reviewing (5.14) - (5.19), we see that the assertions of the Theorem have been confirmed except that the Weierstrass condition (5.19) appears in a restricted, approximate form, in which $U(t)$ is replace by the subset $U^d(t)$. The final step is to take a sequence $d_i \downarrow 0$ and, for each i, carry out the preceding analysis, expressing the multipliers (in relations (5.14) - (5.19)) as $(p_i(.), \lambda_i^0, \mu_i, \lambda_i)$. Define $q_i(.)$ according to (2.3) (when $p_i(.)$ replaces $p(.)$, etc). Using the fact that, for a.e. t,

$$u \in U(t) \implies u \in U^d(t) \text{ for all } i \text{ sufficiently large,}$$

19

we carry out a standard convergence analysis to confirm the validity of (5.14)-(5.19), now involving the 'full' Weierstrass condition, in terms of multipliers $(p(.), \mu, \lambda(.))$ and some selection of $\operatorname{co} \partial^> h(t, \bar{x}(t))$, where the multipliers are cluster points of $p_i(.)$, etc. (See, e.g. [19]). (Observe that (5.18), implying that the $\dot{p}_i(.)$'s are uniformly integrably bounded, has a crucial role here.)

Finally, we need to consider the case when hypothesis (C+L) replaces (C). Let us assume then that $\phi(t, x, u)$ has the special structure $\phi(t, x, u) := \theta^0(t, x) + \sum_{k=1}^m u^i \theta^i(t, x)$, and $\Phi(t)$ and $U(t)$ are convex sets. The assertions of Theorem 3.2 will follow, in this case, from an application of Theorem 3.1 to problem (P_S). provided, of course, the relevant hypotheses are satisfied. We attend only to the verification of (CQ), since satisfaction of the remaining hypotheses is automatic. This requires us to verify:

Claim: There exists a constant $M > 0$ such that, for every $x \in \bar{x}(t) + \epsilon \mathbb{B}$ and $u \in U(t)$ such that $\phi(t, x, u) \in \Phi(t)$, we have:

$$\lambda \in N_{\Phi(t)}^C(\phi(t, x, u)), \ \gamma \in N_{U(t)}^C(u), \ \beta = (\lambda \cdot \theta^1(t, x), \dots, \lambda \cdot \theta^m(t, x)) + \gamma \Rightarrow |\lambda| \leq M|\beta|.$$

for a.e. $t \in [a, b]$.

To accomplish this task, let us take any time t located in the set of full measure on which relevant conditions are satisfied, and (x, u), (λ, γ) and β with the properties listed in the above 'claim'. Since $\gamma \in N_{U(t)}^C(u)$ and $U(t)$ is convex,

$$\gamma \cdot (u' - u) \leq 0 \ \text{ for all } u' \in U(t). \tag{5.20}$$

Since $\phi(t, \bar{x}(t), \bar{u}(t)) \in \Phi(t)$, $\lambda \in N_{\Phi(t)}^C(\phi(t, x, u))$ and $\Phi(t)$ is convex,

$$\lambda \cdot (\phi(t, \bar{x}(t), \bar{u}(t)) - \phi(t, x, u)) \leq 0. \tag{5.21}$$

Since, finally, $\lambda \in N_{\Phi(t)}^C(\phi(t, x_0, u_0))$ and $\xi \in \mathcal{T}_t^\eta \subset T_{\Phi(t)}^C(\phi(t, x_0, u_0))$, we have

$$\lambda \cdot \xi \leq 0. \tag{5.22}$$

We may assume, without loss of generality, that $\lambda \neq 0$ (otherwise the claimed property is true with $M = 0$). By hypothesis (SC) (in which, we may choose $U_r(t) \equiv U(t)$ since $U(.)$ is bounded) there exists $u' \in U(t)$ such that

$$\phi(t, \bar{x}(t), \bar{u}(t)) - \frac{\lambda}{|\lambda|}\delta = \phi(t, \bar{x}(t), u') - \xi. \tag{5.23}$$

By (5.23), (5.20), (5.21) and (5.22), and since $\beta = (\lambda \cdot \theta^1(t, x), \dots, \lambda \cdot \theta^m(t, x)) + \gamma$, we have

$$\begin{aligned}
\beta \cdot (u' - u) &= (\lambda \cdot \theta^1(t, x), \dots, \lambda \cdot \theta^m(t, x)) \cdot (u' - u) + \gamma \cdot (u' - u) \\
&\leq \lambda \cdot \phi(t, \bar{x}(t), u') - \lambda \cdot \phi(t, \bar{x}(t), \bar{u}(t)) + k_\phi \epsilon |\lambda| \\
&\leq -|\lambda|(\delta - k_\phi \epsilon).
\end{aligned}$$

Reducing the size of ϵ, if necessary, we can arrange that $\delta > k_\phi \epsilon$. (δ is as in hypothesis (SC) and k_ϕ is as in hypothesis (H3$''$)(c)). Since $U(t) \subset \sigma_0 \mathbb{B}$, we have from the preceding relation

$$|\lambda| \leq M|\beta|,$$

with $M = 2\sigma_0/(\delta - k_\phi \epsilon)$. We have justified the claim, and the proof of the theorem is complete.

$\square$

References

[1] P. Bettiol, A. Boccia and R. B. Vinter, *Stratified Necessary Conditions for Differential Inclusions with State Constraints*, SIAM J. Control Optimization, 5, 2013, pp. 3903–3917

[2] H. A. Biswas, M. do R. de Pinho, *A Maximum Principle for Optimal Control Problems with State and Mixed Constraints*, ESAIM: Control, Optimisation and Calculus of Variations, to appear.

[3] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983; reprinted as vol.5 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1990.

[4] F. H. Clarke, *Necessary Conditions in Dynamic Optimization*, Mem. Amer. Math. Soc., 173, 2005.

[5] F. H. Clarke and M. do R. de Pinho, *Optimal Control Problems with Mixed Constraints*, SIAM J. Control Optim., 48, 7, 2010, pp. 4500–4524.

[6] F. H. Clarke, Y. Ledyaev and M. do R. de Pinho, *An Extension of the Schwarzkopf Multiplier Rule in Optimal Control*, SIAM J. Control Optim., 49, 2, 2011, pp. 599–610.

[7] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, Graduate Texts in Mathematics, New York, vol. 178, 1998.

[8] E. N. Devdariani and Yu. S. Ledyaev, *Maximum Principle for Implicit Control Systems*, Appl. Math. Optim., 40, 1999, pp. 79–103.

[9] A.Ya. Dubovitskii and A. A. Milyutin, *Extremum Problems with Constraints*, Zh. Vychisl. Mat. Fiz., 5, 1965, pp. 395-45.

[10] A. V. Dmitruk, *Maximum Principle for a General Optimal Control Problem with State and Regular Mixed Constraints*, Comput. Math. Model., 4, 1993, pp. 364–377.

[11] A. V. Dmitruk, *On the development of Pontryagins Maximum Principle in the works of A.Ya. Dubovitskii and A.A. Milyutin*, Control and Cybernetics, 38, 4A, 2009, pp. 923-957.

[12] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory*, John Wiley and Sons, New York, 1966.

[13] K. Makowski and L. W. Neustadt, *Optimal Control Problems with Mixed Control-Phase Variable Equality and Inequality Constraints*, SIAM J. Control, 12, 1974, pp. 184–228.

[14] A. A. Milyutin and N. P. Osmolovskii, *Calculus of Variations and Optimal Control*, AMS, Providence, RI, 1998.

[15] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, *The Mathematical Theory of Optimal Processes, English Translation*, John Wiley and Sons, New York, 1962.

[16] M. do R. de Pinho and R. B. Vinter, *An Euler-Lagrange Inclusion for Optimal Control Problems*, IEEE Trans. Automat. Control, 40, 1995, pp. 1191–1198.

[17] A. B. Schwarzkopf, *Relaxed control problems with state equality constraints*, SIAM J. Control, 13, 1975, pp. 677–694.

[18] A. B. Schwarzkopf, *Optimal Controls with Equality State Constraints*, J. Optim. Th. Appl., 19, 1976, pp. 455–468.

[19] R. B. Vinter, *Optimal Control*, Birkhäuser, Systems and Control: Foundations and Applications, Boston MA, 2000.

[20] J. Warga, *Control Functions with with Functional Restrictions*, SIAM J. Control 8, 1970, pp.360-371.