HAL
open science

Some results on optimal control with unilateral state constraints

Bernard Brogliato

To cite this version:

Bernard Brogliato. Some results on optimal control with unilateral state constraints. Nonlinear Analysis: Theory, Methods and Applications, 2009, 70 (10), pp.3626-3657. 10.1016/j.na.2008.07.020 . hal-01633237

HAL Id: hal-01633237

https://hal.science/hal-01633237v1

Submitted on 12 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Some results on optimal control with unilateral state constraints

Bernard Brogliato *

INRIA, Bipop team-project, ZIRST Montbonnot, 655 avenue de l'Europe, 38334 Saint Ismier cedex, France

In this paper, we study the problem of quadratic optimal control with state variables unilateral constraints, for linear time-invariant systems. The necessary conditions are formulated as a linear invariant system with complementary slackness conditions. Some structural properties of this system are examined. Then it is shown that the problem can benefit from the higher order Moreau's sweeping process, that is, a specific distributional differential inclusion, and from ten Dam's geometric theory [A.A. ten Dam, K.F. Dwarshuis, J.C. Willems, The contact problem for linear continuous-time dynamical systems: A geometric approach, IEEE Trans. Automat. Control 42 (4) (1997) 458–472; A.A. ten Dam, Unilaterally Constrained Dynamical Systems, Ph.D. Thesis, Rijsuniversiteit Groningen, NL, available at http://irs.ub.rug.nl/ppn/159407869, 1997] for partitioning of the admissible domain boundary (in particular for the case of multivariable systems). In fact, the first step may be also seen as follows: does the higher order Moreau's sweeping process (developed in Acary et al. [V. Acary, B. Brogliato, D. Goeleven, Higher order Moreau's sweeping process: Mathematical formulation and numerical simulation, Math. Programm. A 113 (2008) 133–217]) correspond to the necessary conditions of some optimal control problem with an extended integral action? The knowledge of the qualitative behaviour of optimal trajectories at junction times is improved with the approach, which also paves the way towards efficient time-stepping numerical algorithms to solve the optimal control boundary value problem.

Keywords:
Quadratic optimal control
Unilateral state constraints
Multivariable systems
Moreau's sweeping process
Complementarity problems
Distributional differential inclusion

1. Introduction

Optimal control with state constraints is a topic of major importance, and which has attracted the attention of researchers for a long time, see [12,24,26,33,58] to cite a few. Most of this work studies the first order necessary conditions stemming from Boltyanskii–Pontryagin's maximum principle. Dynamic programming under inequality state constraints is examined in [5,59], and second order optimality conditions are studied in [30]. The qualitative properties of optimal trajectories are studied in [23,26,33], while their regularity properties are examined in [59]. Numerical studies have received attention in [8, 40,57] and applications are presented in [5,7,9,10,46,56,60]. The optimal control of systems with inputs which are measures has also received attention [34,43,51,54]. The theory of systems with distributional inputs is progressing significantly [16, 20,27,41], so that it becomes possible to consider control problems involving inputs which are higher order distributions based on solid ground. A further motivation for considering higher degree distributions, will be explained next. However as we shall see later, even problems with a continuous optimal controller, may involve higher degree distributions in the costate differential equation.

The following Bolza problem is of interest

$$\underset{u(\cdot)\in\mathbf{U}}{\text{minimize}}\, I(u) = \frac{1}{2}\int_0^{T_1} [x(t)^{\mathrm{T}}Qx(t) + u(t)^{\mathrm{T}}Ru(t)]\mathrm{d}t + \frac{1}{2}x(T_1)^{\mathrm{T}}Fx(T_1) \tag{1}$$

* Tel.: +33 (0) 4 76 61 53 93; fax: +33 (0) 4 76 61 54 52 52.
E-mail address: bernard.brogliato@inrialpes.fr.

subject to

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = \bar{x}_0, \qquad x(T_1) = \bar{x}_1 \\ w(t) = Cx(t) + D \geq 0 \end{cases} \tag{2}$$

where A, B, $C$and D are constant matrices,[1] (A, B, C) is a minimal state space representation (i.e. (A, B) is controllable and (C, A) is observable [55]), $\bar{x}_0, \bar{x}_1 \in \Phi = \{x \mid Cx + D \geq 0\}$, $\mathbf{U}$ is the set of admissible inputs, $w(t) \in \mathbb{R}^m$, $u(t) \in \mathbb{R}^{n_u}$, $x(t) \in \mathbb{R}^n$ and $Q = Q^{\mathrm{T}} \geq 0$, $R = R^{\mathrm{T}} > 0$ (the Legendre–Clebsch condition). In the sequel, the case $m = n_u = 1$ will be called the monovariable case, whereas $m, n_u \geq 2$ is the multivariable case. This is a particular example of optimal control of a dynamical system with state constraints. The least requirement for problem (1) and (2) to possess a solution in $\mathbf{U}$ is that $\bar{x}_1$ belongs to the reachable set from $\bar{x}_0$, $0 \leq T_1 < +\infty$. A fundamental ingredient is therefore $\mathbf{U}$, as illustrated by Example 7 in Appendix C. This example shows that distributions can easily appear in unilaterally constrained controlled systems. Clearly then the formulation of the optimal control problem in (1) and (2) will have to be modified so that the integral action makes sense. This will be the object of the first part of this paper, and detailed in the next sections.

It is worth noting the strong similarities between this optimal control problem, and the so-called Continuous Linear Programming Problem (CTLP) [3,44], introduced by Bellman [6] to model some economic processes

$$\begin{cases} \text{maximize} \displaystyle\int_{[0,T_1]} c^{\mathrm{T}}(t)x(t)\mathrm{d}t \\ \text{subject to: } H(t)x(t) + \displaystyle\int_0^t G(s,t)x(s)\mathrm{d}s \leq a(t) \\ x(t) \geq 0, \quad t \in [0, T_1]. \end{cases} \tag{3}$$

When the integrand of $I(u)$ in (1) is linear in $x(\cdot)$ and $u(\cdot)$, then the linear optimal control problem with both state and input constraints, is a particular case of the CTLP in (3) [44]. The optimal control problem (1) and (2) and the dual problem of the CTLP in (3) share important features concerning the presence of higher degree distributions in their solution [3, Section 2.2], [44], and the need for time-stepping numerical algorithms to numerically solve them. Some macroeconomics models directly yield optimal control problems with unilateral state constraints and optimal controllers with singular measures [20]. This is why we will embed the optimisation problem (1) and (2) into a suitable framework which allows us to rigorously take into account the possible presence of higher degree distributions.

The first order necessary conditions for the optimal control problem in (1) and (2) can be formulated as [24,45,52,59]

$$\begin{cases} \begin{pmatrix} \dot{x}(t) \\ \dot{\eta}(t) \end{pmatrix} = \begin{pmatrix} A & BR^{-1}B^{\mathrm{T}} \\ Q & -A^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} x(t) \\ \eta(t) \end{pmatrix} + \begin{pmatrix} 0 \\ -C^{\mathrm{T}} \end{pmatrix} \lambda(t) & \text{(a)} \\ x(0) = \bar{x}_0,\ \eta(T_1) = Fx(T_1) + C^{\mathrm{T}}\gamma + \beta = F\bar{x}_1 + C^{\mathrm{T}}\gamma + \beta = \eta_1 & \text{(b)} \\ 0 \leq w(t) = Cx(t) + D \perp \lambda(t) \geq 0 & \text{(c)} \end{cases} \tag{4}$$

where $\eta(t) \in \mathbb{R}^n$, $0 \leq \gamma \in \mathbb{R}^m$, $\gamma^{\mathrm{T}}(Cx(T_1) + D) = 0$, $\beta \in \mathbb{R}^n$, and the optimal control is given on intervals where $\eta(\cdot)$ is a function by

$$u(t) = \operatorname*{argmax}_{u \in \mathbf{U}} \left[-\frac{1}{2}u^{\mathrm{T}}Ru + \eta(t)^{\mathrm{T}}Bu \right] = R^{-1}B^{\mathrm{T}}\eta(t). \tag{5}$$

An implicit assumption which allows one to write (1), (4) and (5) is that the multiplier λ is a measure. If not, these expressions have to be generalized to be meaningful. This is the first objective of this work. In this paper, the optimal trajectory and optimal controller will not be given a special notation in order to lighten the presentation. The values of the multipliers γ and β can be calculated from some boundary conditions [12, Section 2.8]. If $x(T_1)$ is not specified and $F = 0$ then $\eta(T_1) = 0$. We assume that the Slater constraint qualification ($\exists x$ such that $Cx + D > 0$) is satisfied, and that the problem (1) and (2) is normal, a property that is related to the controllability of the constrained system (see e.g. [45, Propositon 4.1] in case λ is a measure). We will assume that reachability and normality hold by taking $u(\cdot)$ in an enlarged set of distributions that will be described later (see Section 3.5). From a general point of view however, the controllability problem for (2) remains open, and we will not tackle it here (see Appendix C for some preliminary hints). In the sequel we shall denote $\tilde{x}^{\mathrm{T}} = (x^{\mathrm{T}}, \eta^{\mathrm{T}})$, $(\tilde{A}, \tilde{B}, \tilde{C})$ the triple associated to the system in (4)(a) (c), which we shall often refer to in the sequel as the *necessary conditions system*. It is clear that if λ is a measure, then from (4)(c)

$$-C^{\mathrm{T}}\lambda(t) \in \partial\psi_{\Phi}(x(t)) \tag{6}$$

(using [22, Proposition 1.3.11] on differentiation of convex functions) and the complementarity implies

$$\operatorname{supp}(\lambda) \subset \{t \mid Cx(t) + D = 0\} \tag{7}$$

[1] In the paper D will also be used to denote the distributional derivative.

It is noteworthy that when λ is not a function (e.g. a Dirac measure), then the meaning of the inclusion in (6) has to be stated rigorously, see [1, Section 2] or [31, p. 76]. Then the dynamical system in (4)(a) (c) may be viewed as a measure differential inclusion

$$\dot{\tilde{x}}(t) - \tilde{A}\tilde{x}(t) \in \partial\psi_{\tilde{\Phi}}(\tilde{x}(t)) \tag{8}$$

with $\tilde{\Phi} = \{\tilde{x} | \tilde{C}\tilde{x} + D \geq 0\}$, and where a state reinitialisation mapping is missing in (8) (let us notice from (7) that the control input $u(\cdot)$ must be a continuous function of time (even analytic) everywhere outside $\text{supp}(\lambda)$). However as we shall see in Section 3, in general one has to resort to a more complex formalism to correctly handle (4). Except if additional constraints are added to (4) to assure that λ is a measure, then the formalism in (4) and especially the *complementary slackness* condition (4)(c) are meaningless. When $m = 1$, the relative degree r of the triple $(\tilde{A}, \tilde{B}, \tilde{C})$ in (4) is twice the relative degree r^{wu} of the triple (A, B, C) in (2) [53] (what we call the relative degree in this paper is often referred to as the constraints order in the optimal control literature [26]). In case $n_u = m \geq 2$ we will say that the triple (A, B, C) has a uniform vector relative degree $(r^{wu}, \ldots, r^{wu})^{\mathrm{T}} \in \mathbb{R}^m$ denoted simply as r^{wu}, if the matrix $CA^{r^{wu}-1}B \in \mathbb{R}^{m\times m}$ whose entries are the scalars $C_iA^{r^{wu}-1}B_j$, is full rank, and $C_iA^{k-1}B_j = 0$ for all $1 \leq i \leq m$, $1 \leq j \leq m$, and $k \leq r^{wu} - 2$. This definition is useful to derive a special state space representation [1,48].

Remark 1. It is known [52,59] that (4)(a) (c) is a Hamiltonian system of the form $\dot{\tilde{x}}(t) = J\begin{pmatrix} \frac{\partial H(x,\eta)}{\partial x} \\ \frac{\partial H(x,\eta)}{\partial \eta} \end{pmatrix} + \begin{pmatrix} 0_{n\times n} \\ -I_n \end{pmatrix} C^{\mathrm{T}}\lambda$, $J = \begin{pmatrix} 0_{n\times n} & I_n \\ -I_n & 0_{n\times n} \end{pmatrix}$, whose unconstrained part has the Hamiltonian function

$$H(x, \eta) = \frac{1}{2}\tilde{x}^{\mathrm{T}}\begin{pmatrix} -Q & A^{\mathrm{T}} \\ A & BR^{-1}B^{\mathrm{T}} \end{pmatrix}\tilde{x}. \tag{9}$$

Notice that in (5) $u(t) = \text{argmax}_{u\in\mathbf{U}} H(x(t), u, \eta(t))$ with

$$H(x, u, \eta) = -\frac{1}{2}x^{\mathrm{T}}Qx - \frac{1}{2}u^{\mathrm{T}}Ru + \eta^{\mathrm{T}}(Ax + Bu). \tag{10}$$

Let us assume that λ is a measure. Defining an augmented Lagrangian function as $\bar{L}(x, u, \lambda) = L(x, u) - \lambda^{\mathrm{T}}(Cx + D)$, the augmented Hamiltonian function is $\bar{H}(x, \eta) = H(x, \eta) + \lambda^{\mathrm{T}}(Cx + D)$. One sees that (8) can also be rewritten as $(-\dot{\eta}(t), \dot{x}(t)) \in \partial\bar{H}(x(t), \eta(t))$, where ∂ is the subdifferential of convex analysis.

Remark 2. The support condition (7) is fundamental. It will also hold when λ is a distribution of higher degree. It implies that in the problem we are examining, the control u and costate η are allowed to have a degree larger than 1, only at junction times.

In this paper, we provide a framework that allows us to give a meaning to the dynamical system in (4)(a) (c) when λ is not a measure; in particular new multipliers are introduced and it is shown that this framework may be useful even if the optimal controller $u(\cdot)$ is a function (but the costate η may be a distribution). This allows us to introduce a generalized action $I(u)$ that handles distributional cases. Then we provide a detailed study of what happens at the entry times, using the geometrical approach of ten Dam [18,19]. This allows us to generalize some results on the existence (or non-existence) of entry states for odd relative degree systems, and to tackle the multivariable case. Finally a study of boundary arcs is proposed using complementarity theory. All these tools are introduced for the first time in the context of optimal control with inequality state constraints and are proved to significantly improve the qualitative knowledge of the optimal trajectories. The approach also paves the way towards the design of time-stepping numerical algorithms to solve the BVP in (4).

The paper is organised as follows: in Section 2 some properties of the system in (4)(a) are proved. In Section 3 a distributional differential inclusion is presented in which the system in (4)(a) (c) is embedded; the properties of its solutions are recalled. In Section 4 the geometric approach of ten Dam [19,18] is used to study the behaviour of the trajectories of (2) at junction times. In Section 5 the behaviour of the system when trajectories of (4)(a) evolve on the boundary of Φ, is analysed. A numerical scheme is outlined in Section 6. Conclusions are given in Section 7, followed by the notation used in the paper. Some mathematical definitions are recalled in Appendix A. An example of a function that belongs to the considered space of solutions is given in Appendix B. An example concerning the controllability of (2) and the set of admissible inputs is in Appendix C.

2. Some properties of the necessary conditions system

The next lemma is important for the well-posedness of (4)(a) seen as an IVP $(x(0^-) = x_0, \eta(0^-) = \eta_0)$, and will also be useful for the qualitative analysis of the BVP solutions, see Sections 4 and 5. We consider $R = I_n$ and $m = 1$ in Lemma 1.

Lemma 1. *If $r^{wu} = 1$, then the leading Markov parameter of the triple $(\tilde{A}, \tilde{B}, \tilde{C})$ is $M^{(2)} = -CBB^{\mathrm{T}}C^{\mathrm{T}} < 0$. More generally if the transfer function of the triple (A, B, C) in (2) has a relative degree $r^{wu} \geq 2$, the leading Markov parameter of the triple $(\tilde{A}, \tilde{B}, \tilde{C})$ is $M^{(2r^{wu})} = (-1)^{r^{wu}} CA^{r^{wu}-1}B(CA^{r^{wu}-1}B)^{\mathrm{T}}$ $(= \tilde{C}\tilde{A}^{r-1}\tilde{B})$.*

Proof. Let us denote $\tilde{A} = \begin{pmatrix} A & BB^{\mathrm{T}} \\ Q & -A^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \tilde{A}_1 & \tilde{A}_2 \\ \tilde{A}_3 & \tilde{A}_4 \end{pmatrix}$ and $\tilde{A}^r = \begin{pmatrix} \tilde{A}_1^{(r)} & \tilde{A}_2^{(r)} \\ \tilde{A}_3^{(r)} & \tilde{A}_4^{(r)} \end{pmatrix}$ for some $r \geq 1$. The first assertion of the Lemma is a simple calculation with $\tilde{A}_2^{(1)} = BB^{\mathrm{T}}$ and $M^{(1)} = -C\tilde{A}_2^{(1)}C^{\mathrm{T}}$. The leading Markov parameter we want to compute is $M^{(2r+2)} = (C\ 0)\tilde{A}^{2r+1}(0\ -C)^{\mathrm{T}} = -C\tilde{A}_2^{(2r+1)}C^{\mathrm{T}}$. If $r^{wu} = \alpha + 1$, then $CA^iB = 0$ for all $0 \leq i \leq \alpha - 1$. Our objective is to show that $M^{(2\alpha+2)} = (-1)^{\alpha+1}CA^{\alpha}BB^{\mathrm{T}}(A^{\mathrm{T}})^{\alpha}C^{\mathrm{T}}$, $\alpha \geq 1$. By direct calculation we have for any integer $i \geq 1$

$$\begin{cases} \tilde{A}_1^{(i)} = \tilde{A}_1^{(i-1)}A + \tilde{A}_2^{(i-1)}Q \\ \tilde{A}_2^{(i+1)} = \tilde{A}_1^{(i)}BB^{\mathrm{T}} - \tilde{A}_2^{(i)}A^{\mathrm{T}}. \end{cases} \tag{11}$$

It follows that

$$\begin{cases} \tilde{A}_2^{(2\alpha)} = \displaystyle\sum_{i=1}^{2\alpha-1}(-1)^{i+1}\tilde{A}_1^{(i)}BB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i-1} - \tilde{A}_2^{(1)}(A^{\mathrm{T}})^{2\alpha-1} & \text{(a)} \\ \tilde{A}_2^{(2\alpha+1)} = \displaystyle\sum_{i=1}^{2\alpha}(-1)^{i}\tilde{A}_1^{(i)}BB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i} + \tilde{A}_2^{(1)}(A^{\mathrm{T}})^{2\alpha}. & \text{(b)} \end{cases} \tag{12}$$

This can be shown using (11) and then by induction. From (11), it follows by induction that $\tilde{A}_1^{(\alpha)} = A^{\alpha} + L_{\alpha}(Q)$, where $L_{\alpha}(Q)$ is some matrix depending on the matrix Q. Indeed, let us assume that $\tilde{A}_1^{(i-1)} = A^{i-1} + L_{i-1}(Q)$. Then $\tilde{A}_1^{(i)} = (A^{i-1} + L_{i-1}(Q))A + \tilde{A}_2^{(i-1)}Q = A^i + L_i(Q)$ with $L_i(Q) = L_{i-1}(Q)A + \tilde{A}_2^{(i-1)}Q$. Using (12) and since $\tilde{A}_1 = A$ and $\tilde{A}_1^{(2)} = A^2 + BB^{\mathrm{T}}Q$, the proof is complete. From (12)(b) one deduces that $M^{(2\alpha+2)} = \sum_{i=1}^{2\alpha}(-1)^{i+1}C\tilde{A}_1^{(i)}BB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i}C^{\mathrm{T}} + C\tilde{A}_2^{(1)}(A^{\mathrm{T}})^{2\alpha}C^{\mathrm{T}}$. Assume for the time being that $CL_i(Q)BB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i}C^{\mathrm{T}} = 0$ for all $1 \leq i \leq \alpha$ (the terms with $\alpha + 1 \leq i \leq 2\alpha$ need not be considered as $CA^iB = 0$ for all $0 \leq i \leq \alpha - 1$). Then we obtain $M^{(2\alpha+2)} = \sum_{i=1}^{2\alpha}(-1)^{i+1}CA^iBB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i}C^{\mathrm{T}}$, and the only nonzero term in this sum is for $i = \alpha$. Therefore $M^{(2\alpha+2)} = (-1)^{\alpha+1}CA^{\alpha}BB^{\mathrm{T}}(A^{\mathrm{T}})^{\alpha}C^{\mathrm{T}}$.

To end the proof, it can be shown again by induction using (12) that $L_i(Q)$ is composed of terms of the general form $A^jB\bigstar$ or $\bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^j$, with $0 \leq j \leq i-2$ and $\bigstar$ is some matrix. Indeed one has $L_i(Q) = L_{i-1}(Q)A + \tilde{A}_2^{(i-1)}Q$, as shown above. Assume that $L_{i-1}(Q) = \sum_{j=0}^{i-3}A^jB\bigstar + \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^j$. Then $L_i(Q) = \sum_{j=0}^{i-3}A^jB\bigstar A + \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^{j+1} + \tilde{A}_2^{(i-1)}Q$. One concludes from (12). Thus $CL_i(Q)BB^{\mathrm{T}}(A^{\mathrm{T}})^{2\alpha-i}C^{\mathrm{T}} = 0$ for all $1 \leq i \leq \alpha$ as required. ■

Lemma 1 continues to hold for systems with a uniform relative degree and $m \geq 1$ (recall that $C \in \mathbb{R}^{m\times n}$) since $CA^iB = 0_{m\times m}$ for all $0 \leq i \leq r^{wu} - 2$ in this multivariable case also. In such a case, having $M^{(2r^{wu})} \in \mathbb{R}^{m\times m}$ full-rank implies that the vectors C_i^{T} are independent, where C_i is the i-th row of C, and B has rank m.

Example 1. Consider the system

$$\begin{cases} \dot{x}_1(t) = x_1(t) + x_2(t) \\ \dot{x}_2(t) = x_3(t) \\ \dot{x}_3(t) = x_4(t) + x_3(t) \\ \dot{x}_4(t) = x_4(t) + x_6(t) \\ \dot{x}_5(t) = x_5(t) + x_1(t) \\ \dot{x}_6(t) = u(t) \\ w(t) = x_1(t) \geq 0. \end{cases} \tag{13}$$

Then $r^{wu} = 5$, $r = 10$, $CA^4B = 1$ and $M^{(10)} = -1$.

Let us introduce a canonical state space representation for the system $(\tilde{A}, \tilde{B}, \tilde{C}, D)$. This is a canonical representation which makes the so-called zero-dynamics explicitly appear. It will be useful for some subsequent developments. For a triple $(\tilde{A}, \tilde{B}, \tilde{C})$, with input $\lambda \in \mathbb{R}$, output $w \in \mathbb{R}$, and relative degree $r \geq 1$, it reads [48]:

$$
\begin{cases}
\dot{z}_1(t) = z_2(t) \\
\dot{z}_2(t) = z_3(t) \\
\dot{z}_3(t) = z_4(t) \\
\vdots \\
\dot{z}_{r-1}(t) = z_r(t) \\
\dot{z}_r(t) = \tilde{C}\tilde{A}^r\tilde{W}^{-1}\tilde{z}(t) + \tilde{C}\tilde{A}^{r-1}\tilde{B}\lambda(t) \\
\dot{\tilde{\xi}}(t) = \tilde{A}_\xi\tilde{\xi}(t) + \tilde{B}_\xi z_1(t) \\
w(t) = z_1(t) + D, z_1(t) = Cx(t) \\
z(0^-) = z_0.
\end{cases} \tag{14}
$$

We call it the $\tilde{z}$-dynamics, or $\tilde{z}$-canonical form, and the full-rank state transformation matrix from (4) to (14) is $\tilde{W}$: $\tilde{z} = \tilde{W}\tilde{x}$. A similar transformation can be applied to (A, B, C), and we then call the obtained representation, the z-dynamics. Then $z = Wx$ for a full-rank matrix W. We denote $\bar{\tilde{z}}^{\mathrm{T}} = (z_1, \ldots, z_r)$ and $\bar{z}^{\mathrm{T}} = (z_1, \ldots, z_{r^{wu}})$. Thus the $\tilde{z}$-dynamics is characterized by $(r, \tilde{C}\tilde{A}^r, \tilde{C}\tilde{A}^{r-1}\tilde{B}, \tilde{A}_\xi, \tilde{B}_\xi)$, and the z-dynamics by $(r^{wu}, CA^{r^{wu}}, CA^{r^{wu}-1}B, A_\xi, B_\xi)$. The $\tilde{z}$-dynamics may be seen as a dynamical extension of z-dynamics, taking into account $u = B^{\mathrm{T}}\eta$ and the costate equation. In the sequel, the components of $\bar{\tilde{z}}$ and of $\bar{z}$ will both be denoted as z_i. Since the zero dynamics play an important role in the systems we are dealing with, the following is of interest.

Lemma 2. *Let $m = 1$. Consider the z- and $\tilde{z}$-canonical forms that correspond to the systems (A, B, C) in (2) and $(\tilde{A}, \tilde{B}, \tilde{C})$ in (4), and let us denote the transitions matrices of their zero dynamics as A_ξ and $\tilde{A}_\xi$ respectively. Then if σ is an eigenvalue of A_ξ with multiplicity m_σ, $-\sigma$ and σ are both eigenvalues of $\tilde{A}_\xi$, both with multiplicities m_σ.*

Proof. Let $Q \geq 0$ and let $L \in \mathbb{R}^{n\times n}$ be such that $L^{\mathrm{T}}L = Q$. The transfer function that corresponds to the operator $\lambda \mapsto w$ in (4) is equal to $H_{w\lambda}(s) = \tilde{C}(sI_{2n} - \tilde{A})^{-1}\tilde{B} = G(s)G(-s)(1 - H^*(s)H(s))^{-1}$ [53], where $G(s)$ is the transfer of the triple (A, B, C) in (2). $H(s)$ is the transfer matrix of the triple (A, B, L), $H^*(s)$ is the transfer matrix of the triple $(-A^{\mathrm{T}}, L^{\mathrm{T}}, B^{\mathrm{T}})$, i.e. the adjoint system to (A, B, L) [55, p. 280]. Consider now the two canonical forms associated to the systems in (2) and (4), respectively. From Lemma 1 one has $\xi \in \mathbb{R}^{n-r^{wu}}$ and $\tilde{\xi} \in \mathbb{R}^{2(n-r^{wu})}$. From the fact that $H_{w\lambda}(s) = G(s)G(-s)(1 - H^*(s)H(s))^{-1}$ and since the ξ-dynamics corresponds to the numerator $B(s)$ of the transfer function $G(s)$, we deduce that the roots of the polynomial $B(\alpha)B(-\alpha) = 0$ are modes of the matrix $\tilde{A}_\xi$. Since the order of $B(\alpha)B(-\alpha)$ is precisely $2(n - r^{wu})$, the result follows. ■

We therefore have a complete description of the triple $(\tilde{A}, \tilde{B}, \tilde{C})$ in (4) in terms of its relative degree and zero dynamics. The role played by the zero dynamics will prove to be important, see Section 5.

3. The higher order Moreau's sweeping process

3.1. Presentation of the differential inclusion

Let us consider the system in (4)(a) (c), and let us forget for the time being, that it may represent the necessary conditions of the maximum principle. An important point is first to understand the dynamics of such a dynamical system involving complementary-slackness conditions. For instance, what is the meaning of $\lambda \geq 0$ if λ is not a measure (distributional multipliers may easily be needed to integrate (4) (a) (c))? The answer to this question is given later in Proposition 2. Before this, let us recall some facts about the higher order sweeping process (HOSP) as it is introduced in [1], which provides a rigorous framework to study the dynamics in (4)(a) (c), and extends the well-known first and second order sweeping process (see references in [1] and [11, Section 5.3] for a non-mathematical introduction). Roughly speaking, the HOSP is a specific differential inclusion, whose solutions are distributions, and which permits to give a meaning to the system (4) (especially the positiveness of λ when it is not a measure). Let I be a real nondegenerate interval. We denote as $\mathcal{T}_n(I)$ the set of distributions of degree $n + 1$ [21] which are generated by RCSLBV functions on I (see (62) and (63) in Appendix A, whose successive derivatives possess an absolutely continuous part (denoted as $[\cdot]$) that is also RCSLBV on I. The right derivative of $[h]$ is denoted as $\hat{h}^{(1)} = \frac{d^+[h]}{dt}(t) = \lim_{\sigma\to 0^+} \frac{[h](t+\sigma)-[h](t)}{\sigma}$. The set of such functions is denoted as $\mathcal{F}_\infty(I;\mathbb{R}) = \cap_{k\in\mathbb{N}} \mathcal{F}_k(I;\mathbb{R})$, with $\mathcal{F}_k(I;\mathbb{R}) = \{h \in \mathcal{F}_{k-1}(I;\mathbb{R}) : \hat{h}^{(k)} := \frac{d^+}{dt}[\hat{h}^{(k-1)}] \in \mathrm{RCSLBV}(I;\mathbb{R})\}$. In particular $\mathcal{F}_0(I;\mathbb{R}) = \mathrm{RCSLBV}(I;\mathbb{R})$. If $T \in \mathcal{T}_n(I)$ and is generated by a function $F \in \mathcal{F}_\infty(I;\mathbb{R})$, it has a "function" part denoted as $\{T\}(\cdot) = [\hat{F}^{(n)}](\cdot)$, and a "measure" part denoted as $\langle\langle T\rangle\rangle$ such that $\langle\langle T\rangle\rangle, \varphi\rangle = \int_{-\infty}^{+\infty} \varphi \mathrm{d}[\hat{F}^{(n-1)}]$, $\forall\ \varphi \in C_0^\infty(I)$. D denotes the distributional derivative, and dz denotes the Stieltjes or differential measure generated by a function z of local bounded variation (see Appendix A). Thus $\mathcal{T}_n(I)$ denotes the set of all Schwartz' distributions, such that there exists a function $F \in \mathcal{F}_\infty(I;\mathbb{R})$ such that $T = D^nF$. Let n be the smallest integer such that $T \in \mathcal{T}_n(I)$, we set

$$
\deg(T) = \begin{cases}
n+1 \quad \text{if } n \geq 1 \\
1 \quad \text{if } n = 0 \quad \text{and} \quad E_0(\{T\}) \neq \emptyset \\
0 \quad \text{if } n = 0 \quad \text{and} \quad E_0(\{T\}) = \emptyset.
\end{cases} \tag{15}
$$

The Dirac measure δ_0 has degree 2, its derivative $\dot{\delta}_0$ has degree 3, etc. Continuous functions have degree 0, and discontinuous functions have degree 1. The distributions we work with originate from LBV functions. This is crucial for the characterization of their support. Since the notion of a solution for the formalism in which we shall embed (4) is crucial, an example is provided in detail in Appendix B. We insist here on the fact that that the solutions that will be considered next, are of special bounded variation. In another words, their derivatives do not contain any singular Lebesgue integrable part. See Definition 1 and Eq. (32).

The core of the HOSP lies in the definition of the following set of tangent cones. Here, and in the following, we suppose that $m = 1$ and $D = 0$ in (2), in order to fit within the framework of [1]. A non zero D can be considered with slight modifications of the tangent cones below, where $\Phi = \mathbb{R}^+$ should be replaced by $\Phi = [-D, +\infty)$. Let Φ be a nonempty closed convex subset of $\mathbb{R}$. We denote by $T_\Phi(x)$ the tangent cone of Φ at $x \in \mathbb{R}$ defined by

$$T_\Phi(x) = \overline{\text{cone}}(\Phi - \{x\}) \tag{16}$$

where $\text{cone}(\Phi - \{x\})$ denotes the cone generated by $\Phi - \{x\}$, defined as in [36] to take into account constraint violations, and $\overline{\text{cone}}(\Phi - \{x\})$ is its closure. Given a closed nonempty convex set $\Phi \subset \mathbb{R}$, we set

$$T^0_\Phi(0) = \Phi, \qquad T^1_\Phi(z_1) = T_\Phi(z_1), \qquad T^2_\Phi(z_1, z_2) = T_{T^1_\Phi(z_1)}(z_2),$$
$$T^i_\Phi(z_1, \ldots, z_i) = T_{T^{i-1}_\Phi(z_1,\ldots,z_{i-1})}(z_i), \quad i \geq 0.$$

In the sequel we will fix $\Phi = \mathbb{R}^+$ when $m = 1$.

Remark 3. If $m \geq 2$ and $\Phi = (\mathbb{R}^+)^m$ then

$$T^i_\Phi(z_1, \ldots, z_i) = \times_{l=1}^m T^i_\Phi(z^l_1, \ldots, z^l_i).$$

Note that

$$T_{\mathbb{R}^+}(x) = \begin{cases} \mathbb{R} & \text{if } x > 0 \\ \mathbb{R}^+ & \text{if } x \leq 0 \end{cases}$$

and

$$T_{\mathbb{R}}(x) = \mathbb{R}.$$

Remark 4. (i) The subdifferential of the indicator function of the cone $T^{i-1}_\Phi(z_1, \ldots, z_{i-1})$ is given by [22, Section 1.3.1]

$$\partial\psi_{T^{i-1}_\Phi(z_1,\ldots,z_{i-1})}(z_i) = \{w \in \mathbb{R} : \langle w, v - z_i\rangle \leq 0, \forall v \in T^{i-1}_\Phi(z_1, \ldots, z_{i-1})\}$$

and is the outward normal cone to $T^{i-1}_\Phi(z_1, \ldots, z_{i-1})$ at z_i.
(ii) It can be shown that

$$T^{i-1}_\Phi(z_1, \ldots, z_{i-1}) = \mathbb{R} \Rightarrow \partial\psi_{T^{i-1}_\Phi(z_1,\ldots,z_{i-1})}(z_i) = \{0\},$$
$$T^{i-1}_\Phi(z_1, \ldots, z_{i-1}) = \mathbb{R}^+ \quad \text{and} \quad z_i > 0 \Rightarrow \partial\psi_{T^{i-1}_\Phi(z_1,\ldots,z_{i-1})}(z_i) = \{0\},$$
$$T^{i-1}_\Phi(z_1, \ldots, z_{i-1}) = \mathbb{R}^+ \quad \text{and} \quad z_i \leq 0 \Rightarrow \partial\psi_{T^{i-1}_\Phi(z_1,\ldots,z_{i-1})}(z_i) = \mathbb{R}^-.$$

Let us introduce now two mathematical formalisms of the HOSP. For this we rely upon the special state space representation in (14), the reason for this particular choice being explained in [1].

Distributional formalism. Find $z_1, \ldots, z_r \in \mathcal{T}_{r-1}(\mathbb{R}^+)$ and $\xi_i \in \mathcal{T}_{r-1}(\mathbb{R}^+)$ $(1 \leq i \leq n - r)$ satisfying the distributional equations

$$\begin{cases} Dz_1 - z_2 = 0 \\ Dz_2 - z_3 = 0 \\ Dz_3 - z_4 = 0 \\ \vdots \\ Dz_{r-1} - z_r = 0 \\ Dz_r - \tilde{C}\tilde{A}^r\tilde{W}^{-1}\tilde{z} = \tilde{C}\tilde{A}^{r-1}\tilde{B}\lambda \\ D\xi = \tilde{A}_\xi\xi + \tilde{B}_\xi z_1 \end{cases} \tag{17}$$

$$\lambda = (\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}[D^{(r-1)}\langle\langle Dz_1 - \{z_2\}\rangle\rangle + \cdots + +D\langle\langle Dz_{r-1} - \{z_r\}\rangle\rangle] + \langle\langle Dz_r - \tilde{C}\tilde{A}^r\tilde{W}^{-1}\{z\}\rangle\rangle \tag{18}$$

and

$$\begin{cases} \mathrm{d}\{z_1\} - \{z_2\}(t)\mathrm{d}t \in -\partial\psi_{T_\Phi^0}(\{z_1\}(t^+)) \\ \vdots \\ \mathrm{d}\{z_i\} - \{z_{i+1}\}(t)\mathrm{d}t \in -\partial\psi_{T_\Phi^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)) \\ \vdots \\ \mathrm{d}\{z_{r-1}\} - \{z_r\}(t)\mathrm{d}t \in -\partial\psi_{T_\Phi^{r-2}(\{z_1\}(t^-),\ldots,\{z_{r-2}\}(t^-))}(\{z_{r-1}\}(t^+)) \\ (\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}[\mathrm{d}\{z_r\} - \tilde{C}\tilde{A}^r\tilde{W}^{-1}\{z\}(t)\mathrm{d}t] \in -\partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\ldots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+)). \end{cases} \tag{19}$$

More compactly (17) is rewritten as $D\tilde{z} - \tilde{W}\tilde{A}\tilde{W}^{-1}\tilde{z} - \tilde{W}\tilde{B}\lambda = 0$ or $D\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\lambda = 0$. The relations given in (19) have to be interpreted in the following sense: find nonnegative real-valued Radon measures $\mathrm{d}\mu_i$ $(1 \le i \le r)$ relative to which the Lebesgue measure $\mathrm{d}t$ and the Stieltjes measure $\mathrm{d}\{z_i\}$ possess densities $\frac{\mathrm{d}t}{\mathrm{d}\mu_i}$ and $\frac{\mathrm{d}\{z_i\}}{\mathrm{d}\mu_i}$ respectively such that:

$$\frac{\mathrm{d}\{z_i\}}{\mathrm{d}\mu_i}(t) - \{z_{i+1}\}(t)\frac{\mathrm{d}t}{\mathrm{d}\mu_i}(t) \in -\partial\psi_{T_\Phi^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)), \quad \mathrm{d}\mu_i\text{-a.e. } t \in \mathbb{R}\ (1 \le i \le r-1) \tag{20}$$

and

$$(\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}\left[\frac{\mathrm{d}\{z_r\}}{\mathrm{d}\mu_r}(t) - \tilde{C}\tilde{A}^r\tilde{W}^{-1}\{z\}(t)\frac{\mathrm{d}t}{\mathrm{d}\mu_r}(t)\right] \in -\partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\ldots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+)), \quad \mathrm{d}\mu_r\text{-a.e. } t \in \mathbb{R}. \quad \blacksquare \tag{21}$$

The solutions of the distributional formalism will be shown to be distributions of degree possibly larger than 1. Let us now introduce a second formalism whose solutions are functions, independently of the relative degree r^{wu}. This will be quite useful when the optimal control problem is embedded into the HOSP, see Sections 3.5 and 3.6.

Measure differential formalism. Find $z_i \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})$ $(1 \le i \le r)$ and $\xi_i \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})(1 \le i \le 2n-r)$ such that[2]

$$\mathrm{d}z_i - z_{i+1}(t)\mathrm{d}t \in -\partial\psi_{T_\Phi^{i-1}(z_1(t^-),\ldots,z_{i-1}(t^-))}(z_i(t^+)) \quad (1 \le i \le r-1) \tag{22}$$

$$(\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}[\mathrm{d}z_r - \tilde{C}\tilde{A}^r\tilde{W}^{-1}z(t)\mathrm{d}t] \in -\partial\psi_{T_\Phi^{r-1}(z_1(t^-),\ldots,z_{r-1}(t^-))}(z_r(t^+)) \tag{23}$$

and

$$\dot{\xi}(t) - \tilde{A}_\xi\xi(t) - \tilde{B}_\xi z_1(t) = 0, \quad \mathrm{d}t - \text{a.e.} t \in \mathbb{R}. \tag{24}$$

The system in (22) and (23) has to be interpreted in the following sense: find nonnegative real-valued Radon measure $\mathrm{d}\mu_i$ relative to which the Lebesgue measure $\mathrm{d}t$ and the Stieltjes measure $\mathrm{d}z_i$ possess densities $\frac{\mathrm{d}t}{\mathrm{d}\mu_i}$ and $\frac{\mathrm{d}z_i}{\mathrm{d}\mu_i}$ respectively such that

$$\frac{\mathrm{d}z_i}{\mathrm{d}\mu_i}(t) - z_{i+1}(t)\frac{\mathrm{d}t}{\mathrm{d}\mu_i}(t) \in -\partial\psi_{T_\Phi^{i-1}(z_1(t^-),\ldots,z_{i-1}(t^-))}(z_i(t^+)), \quad \mathrm{d}\mu_i\text{-a.e. } t \in \mathbb{R}\ (1 \le i \le r-1) \tag{25}$$

and

$$(\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}\left[\frac{\mathrm{d}z_r}{\mathrm{d}\mu_r}(t) - \tilde{C}\tilde{A}^r\tilde{W}^{-1}z(t)\frac{\mathrm{d}t}{\mathrm{d}\mu_r}(t)\right] \in -\partial\psi_{T_\Phi^{r-1}(z_1(t^-),\ldots,z_{r-1}(t^-))}(z_r(t^+)), \quad \mathrm{d}\mu_r\text{-a.e. } t \in \mathbb{R}. \tag{26}$$

Another, more intuitive way to see the measure differential formalism, is to consider the following:
Find $z_1, \ldots, z_r, \xi_1, \ldots, \xi_{n-r} \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})$ and measures $\mathrm{d}\nu_1, \ldots, \mathrm{d}\nu_r$ such that

$$\begin{cases} \mathrm{d}z_1 = z_2(t)\mathrm{d}t + \mathrm{d}\nu_1 \\ \mathrm{d}z_2 = z_3(t)\mathrm{d}t + \mathrm{d}\nu_2 \\ \mathrm{d}z_3 = z_4(t)\mathrm{d}t + \mathrm{d}\nu_3 \\ \vdots \\ \mathrm{d}z_i = z_{i+1}(t)\mathrm{d}t + \mathrm{d}\nu_i \\ \vdots \\ \mathrm{d}z_{r-1} = z_r(t)\mathrm{d}t + \mathrm{d}\nu_{r-1} \\ \mathrm{d}z_r = \tilde{C}\tilde{A}^r\tilde{W}^{-1}z(t)\mathrm{d}t + \tilde{C}\tilde{A}^{r-1}\tilde{B}\mathrm{d}\nu_r \\ \dot{\xi}(t) = \tilde{A}_\xi\xi(t) + \tilde{B}_\xi z_1(t) \end{cases} \tag{27}$$

[2] In this formalism, $\{z_i\} = z_i$ because z_i is a function and $\mathrm{d}z_i$ is a Stieltjes measure.

with the inclusions

$$\mathrm{d}\nu_i \in -\partial\psi_{T_\Phi^{i-1}(\{z_1\}(t^-),\dots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)) \quad \text{for all } 1 \le i \le r \tag{28}$$

and λ is in (18):

$$\lambda = (\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}[D^{(r-1)}\nu_1 + \cdots + D\nu_{r-1}] + \nu_r. \tag{29}$$

In other words λ is a distribution whose measure part is ν_r and its positivity is understood as the positivity of ν_r. If one imposes that λ is a measure, then $\nu_1 = \nu_2 = \cdots = \nu_{r-1} = 0$. The measures $\nu_i = \int \mathrm{d}\nu_i$ are multipliers whose meaning in the context of optimality will be made clear later (see Corollaries 8 and 9, and Remark 9). From (27) and (28) it follows that

$$\mathrm{d}\nu_i(\{t\}) = \{z_i\}(t^+) - \{z_i\}(t^-) \in -\partial\psi_{T_\Phi^{i-1}(\{z_1\}(t^-),\dots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)),\ (1 \le i \le r-1) \tag{30}$$

and

$$\{z_r\}(t^+) - \{z_r\}(t^-) = M^{(r)}\mathrm{d}\nu_r\{t\} \in -\tilde{C}\tilde{A}^{r-1}\tilde{B}\ \partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\dots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+)). \tag{31}$$

If $\bar{\tilde{z}}(\tau^-) \prec 0$, then from (30) and (31) a state jump occurs and $\deg(\lambda) \ge 2$. We recall that given a closed convex nonempty set K and $M = M^{\mathrm{T}} > 0$, $\mathrm{prox}_M[K;\tilde{z}]$ is the closest point to $\tilde{z}$ in K, in the metric defined by M. To better understand (30) and (31) it is useful to recall here the equivalences for two vectors of appropriate dimension, $M = M^{\mathrm{T}} > 0$ and K a closed convex set: $x - y \in -M^{-1}\partial\psi_K(x) \iff x = \mathrm{prox}_M[K, y] \iff x = \mathrm{argmin}_{z\in K}\frac{1}{2}(z-y)^{\mathrm{T}}M(z-y) \iff \langle x-y, v-x\rangle \ge 0$ for all $v \in K$. The measure differential formalism is useful in the context of optimal control, see Section 3.5. Indeed it is a formalism which retains only the "measure part" of the distributional formalism, i.e. the states discontinuities. Most importantly its solutions are functions of time independently of the degree of λ. Both formalisms are related through the following.

Proposition 1 *([1]).* (i) *Let $(z_1,\dots,z_r,\xi) \in (\mathcal{T}_{r-1}(\mathbb{R}^+))^n$ be a solution of Problem* (17)–(19)*, with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$. Then $z_1 = \{z_1\} \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})$, $z_i \in \mathcal{T}_{i-1}(\mathbb{R}^+)$ $(2 \le i \le r)$, $\xi = \{\xi\} \in (\mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R}))^{n-r}$ and $(\{z_1\},\dots,\{z_r\},\xi)$ is a solution of Problem* (22)–(24)*, with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$.*

(ii) *Let $(w_1,\dots,w_r,\xi) \in (\mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R}))^n$ be a solution of Problem* (22)–(24)*, with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$. Then $(z_1,\dots,z_r,\xi) \in (\mathcal{T}_{r-1}(\mathbb{R}^+))^n$, where $z_1 := w_1$ and*

$$z_i = w_i + \sum_{j=1}^{i-1}\left(\sum_{t_k\in E_0(w_j)}(w_j(t_k^+) - w_i(t_k^-))\delta_{t_k}^{(i-j-1)}\right) \quad (2 \le i \le r),$$

is a solution of Problem (17)–(19)*, with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$.* ■

An example is treated in detail in Section 4.4.3 which helps understanding how (17)–(19) is integrated in time and how the state jump mappings (30) and (31) work. Notice that Proposition 1 applies to the solutions of the BVP as well. It is noteworthy that if the distribution $\tilde{z}$ is a solution of the distributional problem in (17)–(19) and if the function $\tilde{\zeta}(\cdot)$ is a solution of the measure problem (22)–(24), then $\{\tilde{z}\}(\cdot) = \tilde{\zeta}(\cdot)$ almost everywhere on $[0, T_1]$. It is also clear that if λ is a measure, then both formalisms are equivalent one to each other. From a Control engineer point of view, the "real" solution is the solution of the distributional formalism. However the measure differential formalism will prove to be quite useful to formulate an extended optimal control problem.

Definition 1. Let $0 \le a < b \in \mathbb{R} \cup \{+\infty\}$ be given. We say that a solution $z \in (\mathcal{T}_{r-1}(\mathbb{R}^+))^n$ of (17)–(19), with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$, is regular on $[a, b)$ if for each $t \in [a, b)$, there exists a right neighborhood $[t, \sigma)$ $(\sigma > 0)$ such that the restriction of $\{z\}$ to $[t, \sigma)$ is analytic.

Regular solutions may possess accumulations of jumps, but only on the left of any $t \in [a, b)$, as right accumulations cannot exist by definition. They encompass right-analytic solutions [44]. The following is proved in [1] and concerns the IVP in (17)–(19).

Theorem 1. *Suppose that $\tilde{C}\tilde{A}^{r-1}\tilde{B} > 0$ and $m = 1$. For each $z_0 \in \mathbb{R}^n$, the system in* (17)–(19)*, with $\{z\}(0^-) = z_0 \in \mathbb{R}^n$ has at least one regular solution.*

Moreover:

(i) *$z_1(\cdot) \equiv \{z_1\}(\cdot) \ge 0$ on $\mathbb{R}^+$*
(ii) *$\|\{\tilde{z}\}(t)\| \le \sqrt{\mathrm{e}^{ct}}\|\tilde{z}_0\|,\ \forall\, t \in \mathbb{R}^+$, for some $c > 0$,*
(iii) *If $\tilde{z}^1$ and $\tilde{z}^2$ are two regular solutions of* (17)–(19)*, with $\{\tilde{z}\}(0^-) = \tilde{z}_0 \in \mathbb{R}^n$ then $\langle\tilde{z}^1,\varphi\rangle = \langle\tilde{z}^2,\varphi\rangle, \forall\varphi \in C_0^\infty(\mathbb{R}^+;\mathbb{R}^n)$;*

where $\tilde{z}_0'$ is uniquely defined by

$$z_{0,i}' = \text{prox}\left[T_\Phi^{i-1}(z_{0,1},\ldots,z_{0,i-1});z_{0,i}\right], \quad \forall\ 1 \le i \le r-1,$$
$$z_{0,r}' = \text{prox}_{(\tilde{C}\tilde{A}^{r-1}\tilde{B})^{-1}}\left[T_\Phi^{r-1}(z_{0,1},\ldots,z_{0,r-1});z_{0,r}\right]$$

and

$$z_{0,l}' = z_{0,l}, \quad (r+1 \le l \le 2n).$$

It is proved that uniqueness holds in the class of regular solutions only. A crucial step is to understand the relationship between **U** and the class of solutions of the HOSP. Indeed it is possible that reachability of $\bar{x}_1$ from $\bar{x}_0$ in (2), holds for a u which belongs to a set **U** of distributions. However is there a correspondence between this **U** and $\mathcal{T}_{r-1}([0,T_1])$? An element of answer will be given in Section 3.5. An important property which helps understanding how the HOSP works is the following. Notice first, that we may write $\mathrm{d}\nu_i$ in (27) and (28) as

$$\mathrm{d}\nu_i = g_i(t)\mathrm{d}t + \mathrm{d}\mathcal{J}_i, \tag{32}$$

where $g_i \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})$ and $\mathrm{d}\mathcal{J}_i$ is an atomic measure with countable and orderable set of atoms generated by the right continuous jump function $\mathcal{J}_i$. The following holds.

Proposition 2 *([1]). Let $z(\cdot)$ be a solution of* (20) *and* (21)*. We have*

$$g_i(t) = 0, \quad \mathrm{d}t\text{-a.e. } t \in \mathbb{R}^+, \quad (1 \le i \le r-1), \tag{33}$$
$$g_r(t) \in -\partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\ldots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+)), \quad \mathrm{d}t\text{-a.e. } t \in \mathbb{R}^+, \tag{34}$$

and

$$0 \le z_1(t^+) \perp \mathrm{d}\nu_r(t^+) \ge 0, \quad \text{for each } t \in \mathbb{R}^+ \tag{35}$$
$$0 \le z_1(t^+) \perp g_r(t^+) \ge 0, \quad \mathrm{d}t\text{-a.e. } t \in \mathbb{R}^+. \tag{36}$$
$$T_\Phi^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-)) \ni \{z_i\}(t^+) \perp -\mathrm{d}\nu_i(\{t\}) \in \partial\psi_{T_\Phi^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)) \tag{37}$$

for all $1 \le i \le r$ and all $t \in \mathbb{R}^+$.

Thus, the complementarity in (4)(c) is given a meaning with (36) and (37). Notice that (36) may equivalently be rewritten as $\int z_1(t)g_r(t)\mathrm{d}t = 0$, $0 \le z_1(t^+), g_r(t^+) \ge 0$. The cone CP in (37) shows what CP the measures $\mathrm{d}\nu_i$ satisfy (the rigorous meaning of the inclusion in the right-hand-side of (37) is explained in [1, Proposition 1] or [31, p. 76]). It is important to notice that given $\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-)$, the measure value depends on $\{z_i\}(t^+)$. We shall come back on (37) in Section 4.2. From (33) it follows that the non-atomic parts of the Lagrange multipliers are zero almost everywhere, except for $g_r(\cdot)$. Consequently the Lagrange multipliers associated to additional constraints to be imposed on the problem, are purely atomic. This may explain why they have been ignored in the literature, since the condition that λ be a measure has always been imposed. It is crucial to keep in mind that the unilateral constraint $Cx \ge 0$ of the original formulation, is replaced now by $\{z_i\}(t^+) \in T_\Phi^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-))$. The definition of the cones and the inclusions (28) guarantee in particular that $z_1(t) = w(t) \ge 0$, see (36).

Remark 5. (i) Theorem 1 remains true if $m \ge 2$ and $\tilde{C}\tilde{A}^{r-1}\tilde{B}$ is a nonsingular symmetric **M**-matrix (see [17] for a definition) [1].
(ii) The sets $\text{supp}(\mathrm{d}\mathcal{J}_i)$ are countable and orderable sets. In other words the set of times at which λ possesses atoms is countable and orderable.
(iii) If $C\bar{x}_0 \ge 0$, then along the solutions of the HOSP IVP one has $\deg(w(\cdot)) = 0$ on $[0,T_1]$.

Another important tool is the following LCP which holds on time intervals $[\tau-\epsilon,\tau)$, $\epsilon > 0$, on which $z_1(t)$ is identically zero:

$$0 \le \lambda(t) \perp \tilde{C}\tilde{A}^r\tilde{W}^{-1}\tilde{z}(t^+) + \tilde{C}\tilde{A}^{r-1}\tilde{B}\lambda(t) \ge 0. \tag{38}$$

In (38) one could write equivalently $g_r(t)$ instead of $\lambda(t)$. This LCP monitors the evolution of the multiplier $\lambda(\cdot)$ for $t \in [\tau-\epsilon,\tau)$. The matrix of this LCP is the leading Markov parameter $M^{(2r^{wu})}$. We therefore have at our disposal a complete dynamical system, which allows us to give a meaning to the system in (4), for any initial data and for all $t \ge 0$. This is a first crucial step for the understanding of the necessary conditions of optimality.

Corollary 1. *Along trajectories of* (17)–(19) *one has $\frac{\mathrm{d}}{\mathrm{d}t}H(\{x\}(t),\{\eta\}(t)) = 0$ almost everywhere.*

Proof. Follows from easy calculations and the fact that the support of the measures $\mathrm{d}\mathcal{J}_i$ is of zero Lebesgue measure on $[0,T_1]$. ■

However, along the optimal trajectory the Hamiltonian may undergo jumps at atoms of λ (see Corollary 3).

3.2. Solutions of the necessary conditions IVP

From now on we will choose to embed the system in (4)(a) (c) in the HOSP. In particular, the degree of distributions refers to the degree of distributions in spaces $\mathcal{T}_n([0, T_1])$ for some n. The following is going to be useful subsequently and follows from the state variable change $\tilde{z} = \tilde{W}\tilde{x}$ and (17)–(19) and (22)–(24).

Costate equation distributional formalism

$$D\eta - Qx + A^{\mathrm{T}}\eta = [0_{n\times n} I_n]\tilde{W}^{-1}\begin{pmatrix} 0^{r-1} \\ M^{(r)}\lambda \\ 0^{2n-r} \end{pmatrix}. \tag{39}$$

Costate equation measure formalism

$$\mathrm{d}\eta - Qx(t)\mathrm{d}t + A^{\mathrm{T}}\eta(t)\mathrm{d}t = [0_{n\times n} I_n]\tilde{W}^{-1}\begin{pmatrix} \mathrm{d}\nu_1 \\ \vdots \\ \mathrm{d}\nu_{r-1} \\ M^{(r)}\mathrm{d}\nu_r \\ 0^{2n-r} \end{pmatrix}. \tag{40}$$

Both $x(\cdot)$ and $\eta(\cdot)$ are $\mathcal{F}_\infty([0, T_1], \mathbb{R})$-functions in (40), while in (39) x and η are distributions in $\mathcal{T}_{r-1}([0, T_1])$. We recall that Proposition 1 guarantees that there is a bijective relation between the solutions of the formalisms in (39) and (40). We also denote $\mathrm{d}\nu_\eta = [0_{n\times n} I_n]\tilde{W}^{-1}(\mathrm{d}\nu_1 \cdots \mathrm{d}\nu_{r-1}; M^{(r)}\mathrm{d}\nu_r O_{2n-r})^{\mathrm{T}}$. In a first step, it is necessary to give a meaning to (4)(a) (c), as a well-posed dynamical system.

Lemma 3. *Let us embed the system in* (4) *(a) (c) in the framework of the HOSP. Assume that* $r^{wu} = 2k$, $k \geq 1$ *and* $m = 1$. *Then there exists a global solution starting from any initial data* $(x(0^-), \eta(0^-))$ *which is regular in the sense of Definition* 1, *and uniqueness holds in the set of regular solutions.*

Proof. From Lemma 1, the leading Markov parameter $M^{(2r^{wu})} = \tilde{C}\tilde{A}^{2r^{wu}-1}\tilde{B} > 0$. Apply Theorem 1 and the fact that the solution $\tilde{x}$ of (4)(a) (c) is regular on $\mathbb{R}^+$ for any initial data. ■

It is crucial to realise that Lemma 3 neither states uniqueness of solutions of the BVP (4)(a) (b) (c) when r^{wu} is even, nor says that when r^{wu} is odd, the BVP (4)(a) (b) (c) has no solution. The well-posedness of the IVP (4)(a) (c) and of the BVP (4)(a) (b) (c), are two very distinct notions. However studying a good formalism for the IVP, will allow us to better understand the BVP. In particular, the study of local properties of the optimal solution can take advantage of the IVP study, as well as for numerical integration. As pointed out in the introduction, the existence of solutions to the minimisation problem (1) and (2), is a reachability problem. Concerning the BVP in (4)(a)–(c), we make the choice to embed it into the HOSP, i.e. we allow for its solutions to be regular in the sense of Definition 1. The consequences of such a choice are important, as it may imply to extend the integral action $I(u)$ in (1) since x, η and u may no longer be functions at contact times. Actually the central question is: let us allow the solutions of the BVP in (4) to be regular. Then how should one extend the minimisation problem in (1) and (2) so that the set of equations/conditions in (17)–(19) represents the necessary optimality conditions?

3.3. Motivation example

Let us examine the simple case where the system (2) is given in the z-dynamics representation, $Q = I_n$ and $R = I_{n_u}$. Let us rewrite $I(u)$ as the sum of a regular part $\frac{1}{2}\int_0^{T_1}[\{z\}(t)^{\mathrm{T}}\{z\}(t) + \frac{1}{2}\{u\}(t)^{\mathrm{T}}R\{u\}(t)]\mathrm{d}t$ and an algebraic part $\frac{1}{2}\sum_k \sigma_{\{z\}}^{\mathrm{T}}(t_k)\sigma_{\{z\}}(t_k)$ which takes into account the norm of the state vector jump. We notice that in addition, the minimisation of the algebraic part has to take into account some constraints on $\{z\}(t^+)$, due to the unilateral constraints. In other words, we disregard what happens at instants t_k, and retain only the right and left-limits of the function part of the state: we choose to work with the measure differential formalism (22)–(24) or (25) and (26). Assume now that $\{z_1\}(t) = 0$ and $\{z_i\}(t) < 0$, for some $t \in [0, T_1]$ and all $2 \leq i \leq r$. Thus from (30) and (31) a jump in $\{z_i\}(\cdot)$, $2 \leq i \leq r$, occurs at t while the other variables are continuous. Then we have $\{z_i\}(t^+) - \{z_i\}(t^-) = \mathrm{argmin}_{\sigma_i + \{z_i\}(t^-)\geq 0}\frac{1}{2}\sigma_i^2$. The HOSP jump rule therefore minimizes $\frac{1}{2}\sum_k\sum_i(\sigma_{\{z_i\}}(t_k))^2$, over the admissible values of the post-impact state.

3.4. Meaning of the costate $\eta(\cdot)$ jump condition

In this subsection, we rely on the embedding of the system (4)(a)–(c) in the HOSP formalism to characterise both x and η as distributions. The condition $\eta(\tau^+) = \eta(\tau^-) - C^{\mathrm{T}}\lambda_1$ with $\lambda_1 \geq 0$ is usually given in the set of the necessary conditions. It is also sometimes indicated that if the optimal controller u is discontinuous, then a jump in $\eta(\cdot)$ occurs. In the light of the HOSP distributional and measure differential formalisms in (17)–(21) and (22)–(26) respectively, one may wonder what

this jump condition really means[3]: is λ a measure at τ (i.e., λ_1 is the magnitude of the atom of $\mathrm{d}\mathcal{J}_r$ at τ, see (32)), or is it $\{\eta\}(\tau^+) = \{\eta\}(\tau^-) - C^{\mathrm{T}}\lambda_1$? This is not at all equivalent.

Proposition 3. *The degree of λ at a time t is ≤ 2 if and only if all the $z_i(\cdot)$, $1 \leq i \leq r-1$, are continuous at t, while $z_r(\cdot)$ is a possibly discontinuous function at t. Moreover in such a case $x(\cdot) = \{x\}(\cdot)$ is a continuous function at t while $\eta(\cdot) = \{\eta\}(\cdot)$ jumps at t.*

Proof. The first assertion comes from the $\tilde{z}$-dynamics. Now if $x(\cdot)$ is a function and has a jump, necessarily from (4) $\deg(\eta) \geq 2$. Thus necessarily $\deg(\lambda) \geq 3$ which is a contradiction. Finally $\tilde{z} = \tilde{W}\tilde{x}$ and the transformation matrix $\tilde{W}$ is square full rank. Thus if $\eta(\cdot)$ is continuous at t, so is $\tilde{z}(\cdot)$. So necessarily $\eta(\cdot)$ has to be discontinuous if $z_r(\cdot)$ is. The second part of the Proposition is proved. ■

Proposition 3 says nothing on the definition of the jump: clearly if $M^{(r)} \leq 0$ and if we admit that the system in (4) is embedded in the HOSP, then the post-jump state is not well defined. Finally the jump in $\eta(\cdot) = \{\eta\}(\cdot)$ has to satisfy stringent conditions to assure that only $z_r(\cdot)$ jumps while the lower order variables $z_i(\cdot)$ remain continuous.

Proposition 4. *Let $m = n_u = 1$. All the $z_i(\cdot)$, $1 \leq i \leq r-1$, are continuous at t, while $z_r(\cdot)$ is a discontinuous function at t, if and only if $(A^{\mathrm{T}})^j\sigma_\eta(t) \in \mathrm{Ker}(B^{\mathrm{T}})$ for $0 \leq j \leq r^{wu}-1$.*

Proof. Let us consider (4) and the state transformation in (14). It is easy to see that $\dot{z}_i(t) = z_{i+1}(t) = CA^i x(t)$ for all $1 \leq i \leq r^{wu}-1$ when $r^{wu} \geq 2$. Let us assume that $\dot{z}_i(t) = M_i x(t) + \sum_{j=0}^{i-r^{wu}} \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^j\eta(t)$, for some $r^{wu} \leq i \leq r-1$, where $\bigstar$ generically denotes some constant scalar and $M_i \in \mathbb{R}^{1\times n}$ is a row vector. Then $\dot{z}_{i+1}(t) = M_i(Ax(t) + BB^{\mathrm{T}}\eta(t)) + \sum_{j=0}^{i-r^{wu}} \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^j(Qx(t) - A^{\mathrm{T}}\eta(t)) = M_{i+1}x(t) + \sum_{j=0}^{i-r^{wu}+1} \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^j\eta(t)$, where $M_{i+1} = M_iA + \sum_{j=0}^{i-r^{wu}} \bigstar B^{\mathrm{T}}(A^{\mathrm{T}})^jQ$ (notice that $M_iB \in \mathbb{R}$). Since $\dot{z}_{r^{wu}}(t) = CA^{r^{wu}}x(t) + CA^{r^{wu}-1}BB^{\mathrm{T}}\eta(t)$ the proof is complete by induction with $M^{r^{wu}} = CA^{r^{wu}}$ and $\bigstar = CA^{r^{wu}-1}B$. From Proposition 3 $x(\cdot)$ is continuous. The "if" part of the proof thus follows by letting j vary from 0 to $i - r^{wu}$ and $r^{wu} \leq i \leq r-1$. The "only if" part follows from the fact that the pair (A, B) is controllable. Therefore the rows $B^{\mathrm{T}}(A^{\mathrm{T}})^j$ are independent for all $0 \leq j \leq n-1$. Since M_i is an n-row and $\bigstar$ is a scalar, and since $r^{wu} \leq n$, the proof follows. ■

Corollary 2. *Let $\eta(\cdot)$ have a jump $\sigma_\eta(t)$ at t and $(A^{\mathrm{T}})^j\sigma_\eta(t) \in \mathrm{Ker}(B^{\mathrm{T}})$ for $0 \leq j \leq r^{wu}-1$. Then $u^{(r^{wu}-1)}(\cdot)$ is the lowest order derivative of $u(\cdot)$ which is discontinuous at t.*

Proof. From Proposition 4 it follows that $\deg(\lambda) \leq 2$ so that (4) does represent the optimality conditions for (1) (2) and (5) holds [45]. Since $u(\cdot) = B^{\mathrm{T}}\eta(\cdot)$, we deduce that $\sigma_u(t) = B^{\mathrm{T}}\sigma_\eta(t) = 0$. Since $\dot{u}(\cdot) = B^{\mathrm{T}}\dot{\eta}(\cdot) = B^{\mathrm{T}}(Qx(t) - A^{\mathrm{T}}\eta(t))$,[4] we deduce that $\sigma_{\dot{u}}(t) = -B^{\mathrm{T}}A^{\mathrm{T}}\sigma_\eta(t) = 0$. The reasoning can be continued until one attains $u^{(r^{wu}-1)}$, recalling that $z_1^{(r-1)}(\cdot) = z_r(\cdot)$ depends on $u^{(r^{wu}-1)}(\cdot)$. ■

One notices from Corollary 2 that the sum of the lowest derivative of $u(\cdot)$ that is discontinuous and of the relative degree are always equal to $2r^{wu}-1$ whenever there is a jump in $\eta(\cdot)$ at t (i.e. $\lambda_1 > 0$) and λ is a measure at t. This is in agreement with [26, Theorem 6] which states that $\lambda_1 \neq 0$ if and only if $u^{(r^{wu}-1)}(\cdot)$ is discontinuous at t, at an entry time. However our proof completely differs from that in [26].

Corollary 3. *Let λ be a measure, and consider the Hamiltonian function in (9) evaluated along the solutions of the HOSP in (17)–(19) (equivalently along the solutions of (22)–(24)). Then $H(t^+) = H(t^-)$ at $t \in [0, T_1]$ if and only if $(\eta(t^+) - \eta(t^-))^{\mathrm{T}}Ax(t) = 0$.*

Proof. One finds from Proposition 3 that $H(t^+) - H(t^-) = -\frac{1}{2}(\eta(t^+) - \eta(t^-))^{\mathrm{T}}BB^{\mathrm{T}}(\eta(t^+) - \eta(t^-)) + (\eta(t^+) - \eta(t^-))^{\mathrm{T}}Ax(t)$. Thus from Proposition 4 one has $H(t^+) - H(t^-) = (\eta(t^+) - \eta(t^-))^{\mathrm{T}}Ax(t)$. ■

Therefore requiring the continuity of the Hamiltonian at contact states usually implies that $\deg(\lambda) \leq 1$. It is of interest to investigate whether or not conditions exist such that the Hamiltonian function, the optimal trajectory and the optimal control are functions of time (so that $I(u)$ in (1) and (5) have a meaning), while λ is a distribution of degree ≥ 3. In a sense, the costate η should "incorporate" all the higher degree distributions. Let us denote the transformation matrix $\tilde{z} = \tilde{W}\tilde{x}$ as $\tilde{W} = \begin{pmatrix} \tilde{W}_1 & 0_{n\times n} \\ \tilde{W}_2 & \tilde{W}_3 \end{pmatrix}$, where $\tilde{W}_j \in \mathbb{R}^{n\times n}$, $1 \leq j \leq 3$, $\tilde{x}^{\mathrm{T}} = (x^{\mathrm{T}}, \eta^{\mathrm{T}})$.

Proposition 5. *Let $m = n_u = 1$ and $r^{wu} = n$. Then* **(a)** $\deg(\lambda) \leq 2 \Leftrightarrow$ **(b)** *$\tilde{z}$ and $\tilde{x}$ are functions* $\Leftrightarrow$ **(c)** *$B^{\mathrm{T}}\eta$ and $A^{\mathrm{T}}\eta$ are functions. Also* $\deg(\lambda) \geq 3 \Leftrightarrow \deg(B^{\mathrm{T}}\eta) \geq 2$ *or* $\deg(A^{\mathrm{T}}\eta) \geq 2$.

[3] We say that a vector jumps if at least one of its components jumps.

[4] We can safely differentiate $B^{\mathrm{T}}\eta(t)$ as $B^{\mathrm{T}}\dot{\eta}(t)$, as we know that the function $B^{\mathrm{T}}\dot{\eta}(\cdot)$ is continuous at t.

Proof. We first notice that under the stated conditions and from (4)(a), $\tilde{W}$ possesses the above structure, and since it is full rank then necessarily $\tilde{W}_3$ and $\tilde{W}_1$ are full rank [28, p. 38]. Let us denote $\zeta_i^{\mathrm{T}} = (z_i, \ldots, z_r)$, $i \geq r^{wu} + 1$. We first prove that **(c)** $\Rightarrow$ **(a)**. If $z_j(\cdot)$, $n+1 \leq j \leq i-1$ are functions and $z_{i-1}(\cdot)$ is discontinuous, then $\deg(\lambda) = r + 2 - i$, and reciprocally. We notice that $\deg(\lambda) = r + 2 - i$, $i \leq r - 1$, implies that $\deg(\lambda) \geq 3$. The vector ζ_i may be considered as a $(n-1)$-vector of distributions of degrees ≥ 2. Since $\eta = -(\tilde{W}_3^{-1}\tilde{W}_2\tilde{W}_1^{-1})\bar{z} + \tilde{W}_3^{-1}\begin{pmatrix} z_{n+1} \\ \vdots \\ z_{i-1} \\ \zeta_i \end{pmatrix}$, then $B^{\mathrm{T}}\eta$ and $A^{\mathrm{T}}\eta$ are time functions if and only if $\tilde{W}_3^{-1}\begin{pmatrix} 0^{i-n+1} \\ \zeta_i \end{pmatrix}$ belongs to $\mathrm{Ker}(B^{\mathrm{T}})$ and to $\mathrm{Ker}(A^{\mathrm{T}})$. However since the pair (A, B) is controllable, the Kalman controllability matrix $\mathcal{K}$ and its transpose are square full rank $(n \times n)$ matrices. Since $\mathcal{K}^{\mathrm{T}} = \begin{pmatrix} B^{\mathrm{T}} \\ B^{\mathrm{T}}A^{\mathrm{T}} \\ B^{\mathrm{T}}(A^{\mathrm{T}})^2 \\ \vdots \\ B^{\mathrm{T}}(A^{\mathrm{T}})^{n-1} \end{pmatrix}$, one sees that $\mathbb{R}^n \ni v \in \mathrm{Ker}(B^{\mathrm{T}}) \cap \mathrm{Ker}(A^{\mathrm{T}})$ implies that $\mathcal{K}^{\mathrm{T}}v = 0$. Hence $v = 0$ and we deduce that $\zeta_i = 0$ which is a contradiction. Therefore $\deg(B^{\mathrm{T}}\eta) \leq 1$ and $\deg(A^{\mathrm{T}}\eta) \leq 1$ implies that $\deg(\lambda) \leq 2$. The other implications/equivalences follow from the $\tilde{z}$-dynamics. The last equivalence is just a rewriting of **(a)** $\Leftrightarrow$ **(c)**. ■

Consequently one sees that the only case where $B^{\mathrm{T}}\eta$ and $A^{\mathrm{T}}\eta$ are functions, is when all $z_i(\cdot)$, $r^{wu}+1 \leq i \leq r$, are functions. In this latter case $z_r(\cdot)$ may be time discontinuous, so that λ is a measure. However there may exist cases when $\deg(\lambda) \geq 3$, $\deg(B^{\mathrm{T}}\eta) \leq 1$ and $\deg(A^{\mathrm{T}}\eta) \geq 2$.

Proposition 6. *Let $m = n_u = 1$ and $r^{wu} = n$ ($\Rightarrow r = 2n$). Let us embed the necessary condition system in* (4) *in the HOSP in* (17)–(21). *Let a trajectory of* (4) *make contact with* $\mathrm{bd}\Phi$ *at $t = \tau$. Assume that $\bar{z}(\tau^-) = 0$. Then* $\deg(\lambda) \leq n$ *and $x(\cdot)$ is a continuous function at τ while* $\deg(B^{\mathrm{T}}\eta) \leq 1$. *Finally* $\deg(A^{\mathrm{T}}\eta) \leq n - 1$ *at t.*

Proof. Recall that $\bar{z} = (z_1, \ldots, z_{r^{wu}})^{\mathrm{T}}$. We notice first that $\bar{z}(\tau^-) = 0$ ($=\bar{z}(\tau^+)$ from (28)–(31)) is equivalent to having $x(\cdot)$ continuous at τ, because of the minimality of (A, B, C) and consequently the structure of the transformation matrix $\tilde{W}$ (since $r^{wu} = n$, the matrix $\tilde{W}_1 = [C^{\mathrm{T}}(CA)^{\mathrm{T}} \cdots (CA^{n-1})^{\mathrm{T}}]^{\mathrm{T}}$ is the Kalman observability matrix). Using the $\tilde{z}$-dynamics (14) and (30) it is deduced from $\bar{z}(\tau^-) = 0$ that $\deg(\lambda) \leq r^{wu}$ at τ and that $\deg(B^{\mathrm{T}}\eta) \leq 1$. The last statement is a consequence of the structure of $\tilde{A}$ in (4). ■

It is worth recalling that all the distributions which appear in the problem, have an atomic part whose support is of zero Lebesgue measure, see Remark 5 (ii). As a consequence of Proposition 6, when $\bar{z}(\tau^-) = 0$ at all entry times τ, the state jumps are jumps of $\{\eta\}(\cdot)$, i.e. $\{\eta\}(t_k^+) \neq \{\eta\}(t_k^-)$. But imposing the constraint $\bar{z}(\tau^-) = 0$ at all entry times τ does not imply that λ is a measure.

Example 2. Let us consider a triple integrator $x^{(3)}(t) = u(t)$ with $x_1(t) \geq 0$. We have $\tilde{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{pmatrix}$, $\tilde{W} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 \end{pmatrix}$, $\tilde{B}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}$, $B^{\mathrm{T}}\eta = \eta_3$, $A^{\mathrm{T}}\eta = \begin{pmatrix} 0 & \eta_1 & \eta_2 \end{pmatrix}^{\mathrm{T}}$. Assume that $\lambda = \dot{\delta}_t$, which happens if $z_5(\cdot)$ jumps at t while $z_6(\cdot)$ is continuous (for instance if $z_1(t^-) = z_2(t^-) = z_3(t^-) = z_4(t^-) = 0$, $z_5(t^-) < 0$, $z_6(t^-) > 0$). From the HOSP dynamics, this implies that $z_1(\cdot)$, $z_2(\cdot)$, $z_3(\cdot)$, $z_4(\cdot)$ and $z_6(\cdot)$ are continuous at t, while $z_5(t^+) = 0$. Thus $\deg(\eta_1) = 0$, $\deg(\eta_2) = 2$ and $\deg(\eta_3) = 0$. This is a case where $\deg(B^{\mathrm{T}}\eta) = 0$ while $\deg(A^{\mathrm{T}}\eta) = 2$, and $\deg(\lambda) = 3$. This is in agreement with Proposition 5: if $\deg(\lambda) = 3$, necessarily either $\deg(B^{\mathrm{T}}\eta) \geq 2$ or $\deg(A^{\mathrm{T}}\eta) \geq 2$. Since $\deg(B^{\mathrm{T}}\eta) \leq 1$ necessarily $\deg(A^{\mathrm{T}}\eta) \geq 2$.

The preceding analysis suggests that there are two basic situations concerning the system in (4)(a) (c):

(1) λ is a measure, i.e. $\lambda = g_r(t)\mathrm{d}t + \mathrm{d}\mathcal{J}_r$,
(2) $\deg(\lambda) \geq 3$ and
  (2.1) $B^{\mathrm{T}}\eta(\cdot)$ and $x(\cdot)$ are time functions on $[0, T_1]$, $\deg(A^{\mathrm{T}}\eta) \geq 2$,
  (2.2) $\deg(\lambda) \geq 3$, $\deg(B^{\mathrm{T}}\eta) \geq 2$.

Quantum electronics and laser control, portfolio optimisation, optimisation of the loan policy of a company, are problems which involve optimal controls with $\deg(u(\cdot)) = 2$ [20]. Considering optimal inputs which are not functions, therefore, makes sense from the application point of view. There are several basic ingredients in the optimality conditions, among

which: (i) the value of the optimal control in (5) which comes from $\frac{\partial H}{\partial u}(x, u, \eta) = 0$, (ii) the costate equation $\dot{\eta}(t) = Qx(t) - A^{\mathrm{T}}\eta(t) - C^{\mathrm{T}}\lambda$ which comes from calculus of variations and Kuhn–Tucker conditions stating that $\dot{\tilde{x}}(t) = J\frac{\partial H(\tilde{x})}{\partial \tilde{x}}(t)$ (see (9) and paragraph after), (iii) the expression of the integral action (or cost) in (1).

In case **(1)** the three ingredients have a meaning and the BVP in (4) represents the optimality necessary conditions with $I(u)$ in (1) [4, Theorem 1] [9] [45, (4.1)–(4.5)]. From [45, Proposition 4.1] normal extremals are minimisers of (1) and (2) in such a case. Also we note that $x = \{x\}$, $u = \{u\}$, and $\eta = \{\eta\}$ on $[0, T_1]$.

Corollary 4. *Assume the conditions of Proposition* 6 *are satisfied and that* $\mathbf{U}$ *is a set of functions of time. Let* $\tau \in [0, T_1]$ *be an entry time. Let us consider the solution* $\tilde{x} = \tilde{W}^{-1}\tilde{z}$ *of the distributional formalism in* (17)–(19) *and the function* $H(x, u, \eta)$ *in* (10). *Then the control which satisfies* $\frac{\partial H}{\partial u}(x, u, \eta) = 0$ *at* τ *is given in* (5) *if and only if* $(\sigma_{z_{n+1}}(\tau), \ldots, \sigma_{z_r}(\tau))\tilde{W}_3^{-T} \in \mathrm{Ker}(A)$. *In this case* $\deg(\lambda) \leq 2$.

Proof. From Proposition 6, one sees that $\deg(A^{\mathrm{T}}\eta) \leq r^{wu} - 1$ at τ. One may rewrite the Hamiltonian as $H(x, u, \eta) = -\frac{1}{2}x^{\mathrm{T}}Qx - \frac{1}{2}u^{\mathrm{T}}Ru + \eta^{\mathrm{T}}Bu + \langle A^{\mathrm{T}}\eta, x\rangle$ since $A^{\mathrm{T}}\eta \in \mathcal{T}_{n-2}$ is a distribution. At τ we may write η as $\eta = \sum_{i=0}^{n-3} \eta_i\delta_\tau^{(i)}$ for some η_i. We know that $\eta = -\tilde{W}_3^{-1}\tilde{W}_2\tilde{W}_1^{-1}\bar{z} + \tilde{W}_3^{-1}\begin{pmatrix} z_{n+1} \\ \vdots \\ z_r \end{pmatrix}$. Since $\bar{z}(\cdot)$ is a continuous function at τ, we deduce that $\eta_i = \tilde{W}_3^{-1}\beta_i$, with $\beta_i = \begin{pmatrix} 0^{r-i-1} \\ \sigma_{z_{r-i}}(\tau) \\ 0^{2n-r+i} \end{pmatrix}$ (from (30) and (31) either $\beta_i = 0$ or $\beta_i = -z_{r-i}(\tau^-) > 0$). The result follows because $\deg(A^{\mathrm{T}}\eta) \leq 1$ if and only if $(\beta_{n-1}, \ldots, \beta_0)^{\mathrm{T}} \in \mathrm{Ker}(A^{\mathrm{T}})$ (otherwise the product $\langle A^{\mathrm{T}}\eta, x\rangle$ necessarily involves $u(\cdot)$ and its derivatives). The last result follows from Propositions 5 and 6. ■

In other words, the optimal input computed along the solution of the distributional problem (17)–(19) is given by (5) at a tangential contact if and only if $\deg(\lambda) \leq 2$. When $\deg(\lambda) \geq 3$ one will have to resort to the measure differential formalism (22)–(24) whose solutions are time functions in $\mathcal{F}_\infty$ to characterize the optimality, see Sections 3.5 and 3.6.

3.5. An extended integral action $I(u)$

Let us deal now with case **(2)**, i.e. $\deg(\lambda) \geq 3$. The case **(2.1)** is of particular interest, because it shows that the optimal controller may be a function of time, while the costate η is a distribution of degree ≥ 3. Thus decreasing the "regularity" of $u(\cdot)$ while keeping $\deg(u) \leq 1$ may create serious difficulties in the analysis of this optimisation problem, the costate equation being a distributional differential equation as (39). On the sets $\mathrm{supp}(\mathrm{d}\mathcal{J}_i)$ with $\mathrm{d}\mathcal{J}_i$ in (32), the dynamics of the system in (4)(a) (c) becomes algebraic, see (30) and (31). Consequently when $\deg(\lambda) \geq 3$ it is necessary that the optimal control problem in (1) and (2), involves an algebraic action in addition to an integral action. One path is the following: u^2 has no meaning if $u = \delta_0$. However $\langle\delta_0, \varphi(t)\rangle^2 = \varphi(0)^2$ has a meaning for any test function $\varphi(\cdot) \in C_0^\infty$. Also $\langle a_1\delta_0 + a_2\dot{\delta}_0, \varphi\rangle^2 = a_1^2\langle\delta_0, \varphi\rangle^2 - 2a_1a_2\langle\delta_0, \varphi + \dot{\varphi}\rangle + a_2^2\langle\delta_0, \dot{\varphi}\rangle^2 = a_1^2\varphi^2(0) - 2a_1a_2(\varphi(0) + \dot{\varphi}(0)) + a_2^2\dot{\varphi}^2(0)$. This way of introducing the singular distributions contributions in the quadratic cost, is different from that considered in [15]. Let $\{\delta_0^n(\cdot)\}$ be a fundamental sequence for the Dirac measure δ_0 [11, Appendix A]. In [15] it is argued that since $\int_0^{+\infty}(\delta_0^n)^2(t)\mathrm{d}t \to +\infty$ as $n \to +\infty$, then the cost to be associated to the distributional parts is $+\infty$. It is concluded in that only those optimal control problems involving no distributions of degree ≥ 2 make sense. Our approach of considering the cost function (or action) when higher degree distributions are present in the optimal control problem, rather follows Moreau's results in [35]. An important tool in the developments which follow is the measure differential formalism of the HOSP in (22)–(26).

Let us denote $\mathrm{d}\nu_{\bar{z}} = (\mathrm{d}\nu_1 \cdots \mathrm{d}\nu_r)^{\mathrm{T}} \in \mathbb{R}^r$. Similarly for $\mathrm{d}\mathcal{J}_{\bar{z}}$ which denotes the atomic part of the vector measure $\mathrm{d}\nu_{\bar{z}} = g_{\bar{z}}(t)\mathrm{d}t + \mathrm{d}\mathcal{J}_{\bar{z}}$ (see (32)). Let us define $\mathcal{M} = \begin{pmatrix} I_{r-1} & 0^{r-1} \\ 0_{r-1} & M^{(r)} \end{pmatrix} \in \mathbb{R}^{r\times r}$, and $\tilde{\mathcal{M}} = \begin{pmatrix} \mathcal{M} & 0_{r\times(2n-r)} \\ 0_{(2n-r)\times r} & 0_{(2n-r)\times(2n-r)} \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$. From (27) we have $\mathrm{d}\tilde{z} = \tilde{A}\tilde{z}(t)\mathrm{d}t + \tilde{\mathcal{M}}\begin{pmatrix} \mathrm{d}\nu_{\bar{z}} \\ 0_{2n-r} \end{pmatrix}$. Thus $\mathrm{d}\nu_{\tilde{x}} = \tilde{W}^{-1}\tilde{\mathcal{M}}\begin{pmatrix} \mathrm{d}\nu_{\bar{z}} \\ 0_{2n-r} \end{pmatrix} = g_{\tilde{x}}(t)\mathrm{d}t + \mathrm{d}\mathcal{J}_{\tilde{x}}$.

Lemma 4. *Let* $r^{wu} = 2k$, $k \geq 0$, *and* $m = 1$. *The HOSP jump rule in* (30) *and* (31) *minimises the quadratic term* $\frac{1}{2}\langle\mathrm{d}\mathcal{J}_{\bar{z}}, \varphi\rangle^{\mathrm{T}}\mathcal{M}\langle\mathrm{d}\mathcal{J}_{\bar{z}}, \varphi\rangle$ *for any test function* $\varphi(\cdot) \in C^0[\mathbb{R}; \mathbb{R}]$ *with support containing* $[0, T_1]$, *under the constraints in* (36) *and* (37).

Proof. One has $\mathrm{d}\nu_i(\{t\}) = \{z_i\}(t^+) - \{z_i\}(t^-)$ for all $t \in [0, T_1]$. If t is not an atom of $\mathrm{d}\nu_i$ then $\mathrm{d}\nu_i(\{t\}) = \mathrm{d}\mathcal{J}_i(\{t\}) = 0$. If t is an atom of $\mathrm{d}\nu_i$ this implies that all the tangent cones $T_\Phi^k(\{z_1\}(t^-), \ldots, \{z_k\}(t^-))$ satisfy $T_\Phi^k(\{z_1\}(t^-), \ldots, \{z_k\}(t^-)) = \mathbb{R}^+$ for $0 \leq k \leq i - 1$. Then $\mathrm{d}\nu_i(\{t\}) = \mathrm{d}\mathcal{J}_i(\{t\}) = \mathrm{argmin}_{\sigma_i + \{z_i\}(t^-)\geq 0}\frac{1}{2}\sigma_i^2$. Indeed from (30) it follows that $\{z_i\}(t^+) = \mathrm{prox}[T_\Phi^{i-1}(z_1(t^-), \ldots, \{z_{i-1}\}(t^-)); \{z_i\}(t^-)]$. We deduce that $\{z_i\}(t^+) = \mathrm{argmin}_{\sigma\in T_\Phi^{i-1}(z_1(t^-),\ldots,\{z_{i-1}\}(t^-))}\frac{1}{2}(\sigma - \{z_i\}(t^-))^2$. In case $\{z_i\}(t^-) < 0$ and $T_\Phi^{i-1}(z_1(t^-), \ldots, \{z_{i-1}\}(t^-)) = \mathbb{R}^+$ (which are the conditions required so that $\mathrm{d}\mathcal{J}_i$ possesses an atom at t) the result follows. We have $\langle\mathrm{d}\mathcal{J}_{\bar{z}}, \varphi\rangle^{\mathrm{T}}\mathcal{M}\langle\mathrm{d}\mathcal{J}_{\bar{z}}, \varphi\rangle = \sum_{i=1}^{r-1}\langle\mathrm{d}\mathcal{J}_i, \varphi\rangle^2 + M^{(r)}\langle\mathrm{d}\mathcal{J}_r, \varphi\rangle^2$, and from Lemma 1 it follows that $M^{(r)} > 0$. The result is proved. ■

Let us illustrate the second argument of the proof. If $\{z_i\}(t^-) < 0$ then $\operatorname{argmin}_{\sigma_i+\{z_i\}(t^-)\geq 0}\frac{1}{2}\sigma_i^2 = -\{z_i\}(t^-)$. Since $\mathrm{d}\nu_i(\{t\}) = \{z_i\}(t^+) - \{z_i\}(t^-)$ we deduce that $\{z_i\}(t^+) = 0$ (a plastic jump). The parallel with unilateral Mechanics and Moreau's second order sweeping process can be done, where the post-impact velocity satisfies $\dot{q}(t^+) = \operatorname{argmin}_{w\in T_\Phi(q(t))}\frac{1}{2}(w-\dot{q}(t^-)^{\mathrm{T}})M(q(t))(w-\dot{q}(t^-))$ when the impacts are plastic [11, Remark 5.11].

Lemma 4 extends to odd-r^{wu} systems, provided $\mathrm{d}\mathcal{J}_r = 0$. We shall find again the specificity of odd-r^{wu} systems in Sections 4.2 and 5. Let us propose the following extended action to be minimised:

$$\underset{u(\cdot)\in\mathbf{U}}{\text{minimize}}\ \frac{1}{2}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle^{\mathrm{T}}\tilde{W}^{\mathrm{T}}\tilde{\mathcal{M}}\tilde{W}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle + \frac{1}{2}\int_0^{T_1}[\{x\}(t)^{\mathrm{T}}Q\{x\}(t)+\{u\}(t)^{\mathrm{T}}\{u\}(t)]\mathrm{d}t + \frac{1}{2}\{x\}^{\mathrm{T}}(T_1)F\{x\}(T_1) \tag{41}$$

for any test function $\varphi(\cdot)\in C^0[\mathbb{R},\mathbb{R}]$, with support containing $[0,T_1]$,

Definition 2. Let $\mathbf{U}$ be a set of distributions in $\mathcal{T}_{r^{wu}}([0,T_1])$ such that the atoms and points of non-analyticity of any $u\in\mathbf{U}$ are in the set $\bigcup_{i=1}^{r}\operatorname{supp}(\mathrm{d}\mathcal{J}_i)$. We denote it as $\mathbf{U}_{\mathcal{J}}^{r^{wu}}$.

The definition of $\mathbf{U}_{\mathcal{J}}^{r^{wu}}$ is motivated by the fact that the HOSP solutions are constrained to exist in the set $\mathcal{T}_{r-1}([0,T_1])$, see Proposition 1. In particular $\deg(z_1)\leq 1$. Having $u\in\mathbf{U}_{\mathcal{J}}^{r^{wu}}$ means that $\deg(u)\leq r^{wu}+1$, while along solutions of the HOSP $\deg(\lambda)\leq r+1=2r^{wu}+1$. The inspection of the z-dynamics shows that $\deg(z_1)\leq 1\Rightarrow\deg(u)\leq r^{wu}+1$. It is consequently useless to look for a larger set (in the set of distributions as defined in Section 3.1) of inputs for this optimal control problem. However as Example 7 shows it is easy to construct cases where $\bar{x}_1$ is not reachable from $\bar{x}_0$ in the set $\mathbf{U}_{\mathcal{J}}^{r^{wu}}$. This is therefore an important issue that is not tackled in this paper. By taking u in $\mathbf{U}_{\mathcal{J}}^{r^{wu}}$ we assume that there is no need for the zero dynamics state ξ to jump in order for $\bar{x}_1$ to be reachable from $\bar{x}_0$. Finally notice that $\{u\}(\cdot)\in L_2([0,T_1];\mathbb{R}^{n_u})$ since quadratic forms of LBV functions are LBV [37].

Let us notice that $\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle^{\mathrm{T}}\tilde{W}^{\mathrm{T}}\tilde{\mathcal{M}}\tilde{W}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle$ contains terms of the form $\varphi^2(t_k)$ for all $t_k\in\operatorname{supp}(\mathrm{d}\mathcal{J}_{\tilde{x}})$, since $\mathrm{d}\mathcal{J}_{\tilde{x}}$ is an atomic measure generated by a right continuous jump function $\mathcal{J}_{\tilde{x}}(\cdot)$. Let us recall that in (41) the state jump times are totally free and are not a priori fixed. The only thing that we impose through the HOSP formalism is that the supports of the atomic measures $\mathrm{d}\mathcal{J}_i$ are orderable and countable sets, and are in $\{t\in[0,T_1]: Cx(t)=0\}$ (even more: the atoms are in the set of junction times). It is noteworthy that even in case **(2.1)**, the addition of an algebraic action is necessary. At a time $t\in\operatorname{supp}(\mathrm{d}\nu_i)$ where **(2.1)** holds, then the value of $\{u\}(\cdot)=B^{\mathrm{T}}\{\eta\}(\cdot)=B^{\mathrm{T}}\eta$ is irrelevant (and this is the case on the set of all atoms of λ as this set is of zero Lebesgue measure). The optimisation is done at such a t on the costate η.

Remark 6. The idea of separating functions and distributions in the action is not new, see e.g. [38] and [3, equ. (2.5)]. Also in [12, Section 3.7] $I(u)$ is split into an integral part and an algebraic part taking into account state jumps. However the works [12,38] do not apply to the problem considered in this paper as they do not involve state unilateral constraints.

Proposition 7. *Suppose that $\bar{x}_1$ is reachable from $\bar{x}_0$ in a finite time $T_1\geq 0$, with $u\in\mathbf{U}_{\mathcal{J}}^{r^{wu}}$. Let r^{wu} be even. Then the HOSP measure differential inclusion in* (22)–(24) *represents the optimality necessary conditions for the extended integral action $I(u)$ in* (41)*, with $u=B^{\mathrm{T}}\eta$ for all $t\in[0,T_1]$ and subject to the constraints in* (36) *and* (37)*. If r^{wu} is odd, then the result holds provided the additional constraint $\mathrm{d}\mathcal{J}_r=0$ holds along the optimal trajectory.*

Proof. In view of the material of Sections 3.1, 3.2 and 3.4, one sees that if $\deg(\lambda)\geq 2$ an algebraic action has to be added to $I(\{u\})$ in (1), taking into account the state $\{x\}(\cdot)$ and/or costate $\{\eta\}(\cdot)$ jumps (where $\{x\}$ and $\{\eta\}$ mean here the function part of the solution of the distributional formalism (17)–(19)). From the measure differential formalism (40), $u=B^{\mathrm{T}}\eta$ for all $t\in[0,T_1]$ means that

$$u(t)=B^{\mathrm{T}}\eta(t)+B^{\mathrm{T}}\int_{[0,t)}\mathrm{d}\nu_\eta, \tag{42}$$

where $\mathrm{d}\nu_\eta$ is defined after (40) and $\eta(\cdot)$ in (42) is to be understood as the solution of the ODE: $\dot{\eta}(t)-Qx(t)+A^{\mathrm{T}}\eta(t)=0$ with $\dot{x}(t)-Ax(t)-BR^{-1}B^{\mathrm{T}}\eta(t)=0$. In other words (42) stems from

$$u(t)=\underset{u\in\mathbf{U}_{\mathcal{J}}^{r^{wu}}}{\operatorname{argmax}}\,H\left(x(t)+\int_{[0,t)}\mathrm{d}\nu_x,\,\eta(t)+\int_{[0,t)}\mathrm{d}\nu_\eta,\,u(t)\right) \tag{43}$$

and it is worthwhile noting that the arguments in the Hamiltonian function in (43), are functions of time (the fundamental role played by the measure differential formalism (22)–(24) and Proposition 1 is clear in this context). Let us recall that solutions of the measure differential formalism are $\mathcal{F}_\infty$ functions, and so is $u(\cdot)$ in (42). From the distributional formalism (39) $u=B^{\mathrm{T}}\eta$ means that $\langle u,\varphi\rangle=\langle B^{\mathrm{T}}\eta,\varphi\rangle$ for all $\varphi\in C_0^\infty([0,T_1])$. Therefore, if the necessary condition system (4) embedded in the distributional HOSP (17)–(19), possesses a distributional solution in $\mathcal{T}_{r-1}([0,T_1])$, the corresponding solution of the measure HOSP in (22)–(24) decomposed as in (42) into an unconstrained and a constrained parts, satisfies (42) and (43) (by Proposition 1 there is a bijective correspondence between solutions of the distributional and the measure

formalisms of the HOSP). The Hamiltonian function $H(\{x\},\{u\},\{\eta\})$ evaluated along the optimal trajectory (and optimal control) satisfies

$$H(\{x\},\{u\},\{\eta\}) = \sup_{\{v\}\in\mathbf{U}^r_{\mathcal{J}}} -\frac{1}{2}\{x\}^{\mathrm{T}}Q\{x\} - \frac{1}{2}\{v\}^{\mathrm{T}}R\{v\} + \{\eta\}^{\mathrm{T}}(A\{x\}+B\{v\}). \tag{44}$$

Lebesgue almost everywhere (equivalently $\{u\}(\cdot) = B^{\mathrm{T}}\{\eta\}(\cdot)$ almost everywhere on $[0,T_1]$). Also the costate equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}(\{\eta\})(t^+) = Q\{x\}(t^+) - A^{\mathrm{T}}\{\eta\}(t^+) - C^{\mathrm{T}}(0_{2n-1}g_r(t^+))^{\mathrm{T}} \tag{45}$$

on intervals $(t_k, t_{k+1}]$ since the solutions are right-continuous. We recall that the values of the functions $\{x\}(\cdot)$, $\{u\}(\cdot)$, $\{\eta\}(\cdot)$ are irrelevant at the atoms t_k of the distribution λ. Let us now deal with the algebraic action in (41). One has $\tilde{z} = \tilde{W}\tilde{x}$, and consequently $\langle \mathrm{d}\mathcal{J}_{\tilde{z}},\varphi\rangle^{\mathrm{T}}\mathcal{M}\langle \mathrm{d}\mathcal{J}_{\tilde{z}},\varphi\rangle = \langle \mathrm{d}\mathcal{J}^{\mathrm{T}}_{\tilde{z}},\varphi\rangle\tilde{\mathcal{M}}\langle \mathrm{d}\mathcal{J}_{\tilde{z}},\varphi\rangle = \langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle^{\mathrm{T}}\tilde{W}^{\mathrm{T}}\tilde{\mathcal{M}}\tilde{W}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle$. The solutions are distributions in $\mathcal{T}_{r-1}([0,T_1])$, solutions of the HOSP in (17)–(19). From Proposition 1 we can consider equivalently the solutions of the measure formalism in (22)–(24). From Lemma 4, the solution of (22)–(24) minimizes the quadratic term $\frac{1}{2}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle^{\mathrm{T}}\tilde{W}^{\mathrm{T}}\tilde{\mathcal{M}}\tilde{W}\langle \mathrm{d}\mathcal{J}_{\tilde{x}},\varphi\rangle$. All the above applies to even r^{wu}, and to odd r^{wu}, provided the constraint $\mathrm{d}\mathcal{J}_r = 0$ is added. ■

In case $u(\cdot)$ is a function, its derivatives may be distributions. This is taken into account by the algebraic action which is a function of η, hence implicitly of $u(\cdot)$ and its derivatives. Consider Example 2. One has at the considered time $u(t) = \eta_3(t)$. From (17) the distributional equality $Du = D\eta_3 = x_3 - \eta_2 = z_3 - \eta_2$ shows that $\deg(Du) = 2$ since $\deg(\eta_2) = 2$. This is the consequence of a jump in $z_5(\cdot)$ at t which is the optimal jump according to the action in (41).

Most importantly, when λ is a measure, (42) and (43) coincide with [59, equ. (11.42)] and [59, equ. (11.41)], respectively. We note however that the solution and optimal controller regularity conditions as studied in [59, Sections 9–11] are not relevant in our framework since the data in (1) and (2) are analytic and the solution we are looking for is regular (in the sense of Definition 1). Similarly (39) and (40) reduce to an equality of measures as in [8, equ. (2.21)] when λ is a measure.

It is worthwhile pointing out that in case **(2.1)**, the integral term of the action $I(u)$ is equal to $\int_0^{T_1}[x(t)^{\mathrm{T}}Qx(t) + u(t)^{\mathrm{T}}u(t)]\mathrm{d}t + \frac{1}{2}x^{\mathrm{T}}(T_1)Fx(T_1)$. However from Corollary 4 one does not have $\frac{\partial H}{\partial u} = 0$ at atoms of $\mathrm{d}\nu_i$, $1 \le i \le r-1$, but only almost everywhere on $[0,T_1]$ (see also (42)). The algebraic action is then equal to $\frac{1}{2}\langle (0_n\,\mathrm{d}\mathcal{J}^{\mathrm{T}}_\eta),\varphi\rangle\tilde{W}^{\mathrm{T}}\tilde{\mathcal{M}}\tilde{W}\langle (0_n\,\mathrm{d}\mathcal{J}^{\mathrm{T}}_\eta)^{\mathrm{T}},\varphi\rangle$, and takes into account possible jumps in the costate $\eta(\cdot)$.

Example 3. Following [12, Section 3.11] let us consider the system with relative degree $r^{wu} = 2$, $l > 0$,

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = u(t) \\ w(t) = l - x_1(t) \ge 0 \end{cases} \tag{46}$$

with $Q = 0$ and $R = 1$. Then (4) becomes in the $\tilde{z}$-canonical representation

$$\begin{cases} \dot{z}_1(t) = z_2(t) \quad (= -x_2(t)) \\ \dot{z}_2(t) = z_3(t) \quad (= -\eta_2(t)) \\ \dot{z}_3(t) = z_4(t) \quad (= \eta_1(t)) \\ \dot{z}_4(t) = \lambda(t) \\ 0 \le w(t) = l + z_1(t), \quad z_1(t) = -x_1(t). \end{cases} \tag{47}$$

Let us consider the end-point conditions $z_1(0^-) = z_1(T_1^+) = -l$ and $z_2(0^-) = z_2(T_1^+) < 0$. Then the optimal solution is $\tilde{z}(t) = 0$ for all $t \in (0,T_1)$ (so that $u(t) = 0$ for all $t \in (0,T_1)$ and the integral action is zero along the optimal path), and $z_2(\cdot)$ jumps at $t = 0$ and $t = T_1$ (here $z_2(\cdot)$ is to be considered as a solution of the measure differential formalism (22)–(24)). Thus $\mathrm{d}\nu_2 > 0$ and $\mathrm{d}\nu_1 = \mathrm{d}\nu_3 = \mathrm{d}\nu_4 = 0$. The algebraic action is equal to $\langle \mathrm{d}\nu_2,\varphi\rangle^2$. Along the solutions of the distributional formalism (17)–(19) one has $u = \eta_2 = -z_3$ as an equality of distributions, and $\deg(u) = 2$. The optimal solution that is presented in [12, Section 3.11] is such that $z_4(t_1^-) < 0$ at an entry time t_1, so that $\deg(\lambda) = 2$. An open issue is to find whether or not admitting jumps in $z_3(\cdot) = -\eta_2(\cdot)$ (solutions of (22)–(24)), hence a discontinuous optimal controller, would allow one to decrease the action value compared to the case when λ is constrained to be a measure. In this case, the distributional formalism solution yields $\deg(\lambda) = 3$. In this simple example, the test merely amounts to varying the entry and exit time, allowing for $\deg(\lambda) \ge 3$.

3.6. Hamilton's principle of Mechanics

Let us briefly expose here some facts about Hamilton's principle of Mechanics for Lagrangian systems subjected to complementarity relations and impact laws. As recalled in [11, Section 3.5], the addition of a unilateral constraint and of an impact law, implies that Hamilton's principle has to be modified in order to keep its meaning (otherwise it cannot represent

the nonsmooth feature of the dynamics). The position is supposed to be constrained in $\Phi = \{q \mid h(q(t)) \geq 0\}$ with $h : \mathbb{R}^n \to \mathbb{R}$, and we take $T_1 = 1$. The Lagrangian dynamics is embedded into Moreau's second order sweeping process [2]. Let us first consider the following minimization problem with $e \in [0, 1]$.

$$\underset{\substack{q(0)=q_0\in\Phi \\ q(1)=q_1\in\Phi \\ \dot{q}(t^+)+e\dot{q}(t^-)\in T_\Phi(q(t))}}{\text{minimize}} \quad \frac{1}{2}\sigma_{\dot{q}}(t)^{\mathrm{T}}M(q(t))\sigma_{\dot{q}}(t) + \int_0^1 L(q(t), \dot{q}(t))\mathrm{d}t. \tag{48}$$

The solution of the sweeping process on [0, 1] which satisfies $q(0) = q_0$ and $q(1) = q_1$, minimises the action (48) as it satisfies the smooth Euler–Lagrange dynamics outside impacts, and minimises the quadratic term at impacts. However the writing in (48) is not satisfactory because it explicitly resorts to impact times in the formulation of the algebraic term of the action (notice that if t is not an impact time then the corresponding algebraic action becomes trivial). Let us define the Stieltjes measure $\mathrm{d}v$ associated to the acceleration, whose atoms are the impact times so that $\mathrm{d}\mathcal{J}_v(\{t\}) = \dot{q}(t^+) - \dot{q}(t^-)$. Then one may rewrite (48) as

$$\underset{\substack{q(0)=q_0\in\Phi \\ q(1)=q_1\in\Phi \\ \dot{q}(t^+)+e\dot{q}(t^-)\in T_\Phi(q(t))}}{\text{minimize}} \quad \frac{1}{2}\langle \mathrm{d}\mathcal{J}_v, \varphi\rangle^{\mathrm{T}}M(q(t))\langle \mathrm{d}\mathcal{J}_v, \varphi\rangle + \int_0^1 L(q(t), \dot{q}(t))\mathrm{d}t \tag{49}$$

with $\varphi(\cdot) \in C^0$ a test function whose support contains [0, 1]. As shown in [13] for the case $e = 1$, the extremal of the functional

$$\int_0^1 (L(q(t), \dot{q}(t))\mathrm{d}t + h(q(t))\lambda) \tag{50}$$

under the constraints $h(q(t)) \geq 0$ for all $t \in [0, 1]$, some initial data and no kinetic energy loss, is a solution to the bounce problem, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(q,\dot{q})}{\partial \dot{q}}\right)(t) - \frac{\partial L(q,\dot{q})}{\partial q}(t) = \nabla h(q(t))\lambda$ on [0, 1], and vice-versa. The multiplier λ in (50) is a measure whose support is in the set $\{t \in [0, 1] \mid h(q(t)) = 0\}$.

One sees that in (50) the Lagrangian function is augmented to cope with the unilaterality, and the multiplier is a measure. The basic idea is to augment the Lagrangian function $L(x, u) = \frac{1}{2}x^{\mathrm{T}}Qx + \frac{1}{2}u^{\mathrm{T}}Ru$, taking into account the specific features of the problem. Let us recall that $\mathrm{d}\nu_{\tilde{z}} = (\mathrm{d}\nu_1 \cdots \mathrm{d}\nu_r)^{\mathrm{T}}$, $\mathrm{d}\nu_{\tilde{x}} = \tilde{W}^{-1}\tilde{\mathcal{M}}\begin{pmatrix}\mathrm{d}\nu_{\tilde{z}} \\ 0_{2n-r}\end{pmatrix}$, and $\mathrm{d}\tilde{x} - \tilde{A}\tilde{x}(t)\mathrm{d}t - \mathrm{d}\nu_{\tilde{x}} = 0$ from (22)–(24). We define the augmented Lagrangian as

$$\bar{L}(\{x\}, \{u\}, \mathrm{d}\nu_{\tilde{x}}, \mathrm{d}t) = L(\{x\}, \{u\})\mathrm{d}t + \{\tilde{x}\}(t^+)^{\mathrm{T}}J\mathrm{d}\nu_{\tilde{x}}$$

where $J \in \mathbb{R}^{2n\times 2n}$ is defined in Remark 1. The product $\{\tilde{x}\}(t^+)^{\mathrm{T}}J\mathrm{d}\nu_{\tilde{x}}$ is written here with a strong abuse of notation and requires some care. The function $\{\tilde{x}\}(.)$ is right-continuous and $\mathrm{d}\nu_{\tilde{x}}$ is a measure with possible atoms at times t_k, $k \geq 0$. The space of functions which are $\mathrm{d}\nu_{\tilde{x}}$-integrable contains functions continuous at t_k, and also the functions which are $\mathrm{d}\nu_{\tilde{x}}$-almost everywhere equal to an integrable and continuous function $g(.)$. Since the supports of the atoms are the singletons $\{t_k\}$, it is sufficient that $\{\tilde{x}\}(t_k) = g(t_k)$. Then, denoting the atoms of $\mathrm{d}\nu_{\tilde{x}}$ as δ_{t_k} one has

$$\int \{\tilde{x}\}\mathrm{d}\delta_{t_k} = \int g\mathrm{d}\delta_{t_k} = g(t_k) = \{\tilde{x}\}(t_k) = \{\tilde{x}\}(t_k^+). \tag{51}$$

Equality (51) shows that $\int \bar{L}(\{x\}, \{u\}, \mathrm{d}\nu_{\tilde{x}}, \mathrm{d}t)$ is meaningful. We therefore propose an alternative to Proposition 7.

Proposition 8. *Suppose that $\bar{x}_1$ is reachable from $\bar{x}_0$ in a finite time $T_1 \geq 0$, with $u \in \mathbf{U}_{\mathcal{J}}^{r^{wu}}$. Let r^{wu} be even. The HOSP measure differential formalism in* (22)–(24) *represents necessary conditions satisfied by the extremals of the functional*

$$I(u) = \int_0^{T_1} \bar{L}(\{x\}, \{u\}, \mathrm{d}\nu_{\tilde{x}}, \mathrm{d}t) + \frac{1}{2}\{x\}(T_1)^{\mathrm{T}}F\{x\}(T_1) \tag{52}$$

where the measures $\mathrm{d}\nu_i$ satisfy the inclusions (28) *and $u = B^{\mathrm{T}}\eta$.*

Proof. Let us consider the augmented Hamiltonian

$$\bar{H}(\{x\}, \{u\}, \nu_{\tilde{x}}) = -L(\{x\}, \{u\}) + \{\tilde{x}\}(t^+)^{\mathrm{T}}J\nu_{\tilde{x}} + \{\eta\}^{\mathrm{T}}(A\{x\} + B\{u\}). \tag{53}$$

It is noteworthy that in the framework of (22)–(24) one can drop all brackets as $\{\tilde{x}\}(\cdot) = \tilde{x}(\cdot)$, because the solution is considered as a function in $\mathcal{F}_\infty([0, T_1]; \mathbb{R})$. Since the set of controllers is $\mathbf{U}_{\mathcal{J}}^{r^{wu}}$ and $u = B^{\mathrm{T}}\eta$ (whose meaning is as in (42)), the

support of the measure $\mathrm{d}\nu_{\tilde{x}}$ is orderable and countable. Mimicking [13, Remark 1.1] and using the augmented Hamiltonian in (53), we deduce that the extremals of the functional $I(u)$ satisfy $\mathrm{d}\tilde{x} = J\mathrm{d}\bar{H}$, i.e.:

$$\begin{cases} \mathrm{d}x = \dfrac{\partial}{\partial \eta}\left(-\bar{L}(x,u)\right)(t)\mathrm{d}t + Ax(t)\mathrm{d}t + Bu(t)\mathrm{d}t = Ax(t)\mathrm{d}t + Bu(t)\mathrm{d}t + \mathrm{d}\nu_x \\ \mathrm{d}\eta = \dfrac{\partial}{\partial x}\left(\bar{L}(x,u)\right)(t)\mathrm{d}t - A^{\mathrm{T}}\eta(t)\mathrm{d}t = Qx(t)\mathrm{d}t - A^{\mathrm{T}}\eta(t)\mathrm{d}t + \mathrm{d}\nu_\eta \\ u = R^{-1}B^{\mathrm{T}}\eta \end{cases} \tag{54}$$

as an equality of differential measures for the first two lines of (54). The third equality has the meaning explained in the proof of Proposition 7 (see (42)) and allows one to calculate the optimal input using Proposition 1. One sees that (54) is the measure differential formalism (22)–(24). One therefore has along solutions of the measure differential inclusion (54) $u(t) = \mathrm{argmax}_{u\in\mathcal{F}_\infty} \bar{H}(x, u, \mathrm{d}\nu_{\tilde{x}}, \mathrm{d}t)$ Lebesgue almost everywhere (notice that if $u \in \mathbf{U}_{\mathcal{J}}^{r^{wu}}$ then $\{u\}(\cdot) \in \mathcal{F}_\infty$). ■

It is noteworthy that if a solution of the measure differential formalism is a solution of the minimisation problem in (52), then it satisfies (54), and equivalently from Proposition 1 the corresponding solution of the distributional formalism will satisfy (44) and (45) Lebesgue a.e. on $[0, T_1]$. We have therefore extended the measure differential inclusion necessary conditions in (4) and (6) to the following BVP (named a Measure BVP)

MBVP :

$$-\mathrm{d}\tilde{x} + \tilde{A}\tilde{x}(t)\mathrm{d}t + \mathrm{d}\nu_{\tilde{x}} = 0$$

$$\mathrm{d}\nu_{\tilde{x}} = \tilde{W}^{-1}\tilde{\mathcal{M}}\begin{pmatrix} \mathrm{d}\nu_{\tilde{z}} \\ 0_{2n-r} \end{pmatrix}$$

$$\mathrm{d}\nu_{\tilde{z}} = (\mathrm{d}\nu_1 \ \cdots \ \mathrm{d}\nu_r)^{\mathrm{T}}$$

$$\mathrm{d}\nu_i \in -\partial\psi_{T_\Phi^{i-1}(z_1(t^-),\ldots,z_{i-1}(t^-))}(z_i(t^+)) \ \text{ for all } \ 1 \le i \le r$$

$$\tilde{z} = \tilde{W}\tilde{x}$$

$$x(0) = \bar{x}_0,\ \eta(T_1) = Fx(T_1) + C^{\mathrm{T}}\gamma + \beta = F\bar{x}_1 + C^{\mathrm{T}}\gamma + \beta = \eta_1$$

which reduces to (4) and (6) when λ is a measure (see (36)). Let us recall that if $x(T_1)$ is not specified and $F = 0$ then $\eta(T_1) = 0$, and that both $\tilde{x}$ and $\tilde{z}$ are time functions since they are the solution of the measure differential formalism of the HOSP in (22)–(24). This will permit to settle a time discretisation of MBVP (see Section 6). It is also noteworthy to recall that the atoms of all the distributions of degree ≥ 2 belong to the set of junction times in the setting developed in this paper.

Remark 7. Comparing the "optimality" of the optimal trajectories obtained with $\deg(\lambda) \le 2$, and of those obtained with $\deg(\lambda) \ge 3$, is not easy because the two problems involve different cost functions. Depending on the application, one may anyway want to compare the cost of the problem with $\deg(\lambda) \le 2$, and the continuous part of the cost with $\deg(\lambda) \ge 3$.

4. Behaviour of the optimal trajectories at junctions with $bd\Phi$

There are two major points: what happens at an entry time, and what happens after an entry time. In this section we use the partitioning of $bd\Phi$ as proposed in [18,19]. More precisely, we study the *qualitative* behaviour of optimal solutions which attain $bd\Phi$, and possibly leave it or remain on it. The control input $u(\cdot)$ in (2) is assumed to be piecewise smooth (a piecewise $C^\infty([0, T_1]; \mathbb{R})$ time function). Roughly, the process consists of subdivising $bd\Phi$ into subsets and especially trying to select that portion of $bd\Phi$ such that no input $u(\cdot)$ exists that can keep an entry state $x(\tau)$ inside Φ. Then surely an optimal trajectory which possesses a boundary arc, cannot pass through this $x(\tau)$.

4.1. The admissible domain boundary partitioning

The following assumptions are in order in Sections 4.1–4.3:

(i) The pair (A, B) is controllable,
(ii) $\mathrm{Im}(B) \subseteq \mathrm{Ker}(C)$,
(iii) $D = 0$, $C \neq 0$ and $m = 1$,
(iv) $\Phi \neq \emptyset$.

One sees that (ii) implies that $CB = 0$, i.e. $r^{wu} \geq 2$. The case $r^{wu} = 1$ will be treated, but deserves special attention. We restrict ourselves to the case $D = 0$, because as shown in [18, Section VII] it is always possible to transform the dynamics to recover this case. The subsets of $bd\Phi$ are invariant under linear state feedback. Three sets are of interest here: χ_{con} the subset of $bd\Phi$ at which trajectories coming from $\text{Int}(\Phi)$ can make contact with $bd\Phi$, χ_{rel} the subset of $bd\Phi$ at which trajectories starting on $bd\Phi$ can leave $bd\Phi$ and stay on a positive time interval in Φ, and ν^* the subset of $bd\Phi$ such that there exists one $u(\cdot) = u^*(\cdot)$ that can keep (locally) the trajectory on $bd\Phi$ (ν^* is the largest controlled invariant subspace in $bd\Phi$). Trajectories that make contact with $bd\Phi$ in $\chi_{\text{con}} \setminus \nu^*$, leave Φ. Let

$$r(x, u) : bd\Phi \times \mathbf{U} \to \mathbb{N} \cup \{+\infty\}$$

and

$$r(x, u) = \min\{i \in \mathbb{N} \mid w^{(i)}(x, u) \neq 0\}$$

with $r(x, u) = +\infty$ if $w^{(i)}(x, u) = 0$ for all $i \in \mathbb{N}$, where $\mathbf{U}$ is the set of piecewise smooth inputs. According to the framework developed in Section 3, we are looking for a solution that is regular in the sense of Definition 1. Thus $x(\cdot)$ is analytic on $[0, T_1] \setminus \{t_k\}$, and so is the optimal control $u(\cdot)$. Moreover the times at which the solution is not analytic are the jump times $t_k \in \text{supp}(\lambda)$ and are junction times. But left accumulations of state $\tilde{z}$ jumps may occur. As a consequence, to derive most of the results on the qualitative behaviour of optimal trajectories at an entry time τ, one has to make the additional assumption that there is no left accumulation at τ (such accumulations of touch times on the left of an entry time may however exist ni some optimal trajectories, see [47]). In view of Proposition 7, this is equivalent to assuming that the set of admissible controls is restricted to piecewise analytic functions $\{u\}(\cdot)$ (i.e. the support of the atomic part of λ is finite) and that reachability holds within this set. Such an assumption is made for instance in [33, Theorem 1] (in a different context, though), where junction states are then called *analytic* junctions. Finally most of the results stated below hold for $m \geq 2$, but assuming that there is only one constraint that is active at the considered time.

Lemma 5 (*[18]*). *The following holds.*

- $\chi_{\text{con}} = \{x \in bd\Phi \mid \exists u \in \mathbf{U}$ *such that* $\{r(x, u) < +\infty$ *and even, and* $w^{(r(x,u))} > 0\}$, *or* $\{r(x, u) < +\infty$ *and odd, and* $w^{(r(x,u))} < 0\}\}$,
- $\chi_{\text{rel}} = \{x \in bd\Phi \mid \exists u \in \mathbf{U}$ *such that* $r(x, u) < +\infty$ *and* $w^{(r(x,u))} > 0\}$
- $\nu^* = \{x \in bd\Phi \mid \exists u \in \mathbf{U}$ *such that* $r(x, u) = +\infty\}$.

With some abuse of notation u in $r(x, u)$, means that the involved functions depend also on the derivatives of u. The following result is obvious from the definition of the three subsets of $\text{bd}\Phi$.

Proposition 9. *Boundary arcs of optimal trajectories exist only in the subset ν^* of $bd\Phi$. Contact states exist only in χ_{con}, and exit states exist only in χ_{rel}. Entry states satisfy $z(t^-) \in \chi_{\text{con}}$ and $z(t^+) \in \nu^*$. Touch states satisfy $z(t^-) \in \chi_{\text{con}}$ and $z(t^+) \in \chi_{\text{rel}}$.*

This shows that an entry time may exist in $\chi_{\text{con}} \setminus \nu^*$ only if $u(\cdot)$ is not smooth (i.e. not infinitely differentiable), since there must exist a higher derivative that jumps to zero at the entry time. These results say nothing on the relationships between the regularity of the control and the nature of the state (contact, touch, entry). More details are given below. As the following example shows it may even be necessary to apply a distributional input.

Example 4. Let us choose $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $C = (1\ 0)$. Then $\chi_{\text{con}} = \{x | x_1 = 0, x_2 \leq 0\}$, $\chi_{\text{rel}} = \{x | x_1 = 0, x_2 \geq 0\}$, and $\nu^* = \{x | x_1 = x_2 = 0\}$. If $u(\cdot)$ is restricted to be a function everywhere, all trajectories entering $bd\Phi$ in $\chi_{\text{con}} \setminus \nu^*$ leave Φ. All trajectories initialised in $\chi_{\text{rel}} \setminus \nu^*$ enter $\text{Int}(\Phi)$. Thus boundary arcs are restricted to ν^* and $u^* = 0$. If an entry time τ exists in $\chi_{\text{con}} \setminus \nu^*$, then it is necessary to apply a Dirac measure u_τ to keep the state in Φ and make it jump in χ_{rel}.

Let us notice that the case $x_2(\cdot) \geq 0$ (i.e. $C = (0\ 1)$) does not satisfy assumption (ii). In such a case $\nu^* = bd\Phi$ [19, Lemma 6.7.1] [18, lemma VII; 1].

Algorithms exist [18, Algorithm A.8] [19, Algorithm A.1.6] which allow one to compute these three subsets in a finite number of steps. For instance the sequence $\nu^{k+1} = \text{Ker}(C) \cap A^{-1}(\nu^k + \text{Im}(B)) = \{x \in \nu^k \mid CA^k x = 0\}$ converges towards ν^* in at most r^{wu} steps.[5] This is extremely important in view of the development of a numerical analysis of the problem of interest here.

4.2. Entry and contact times and states

In this subsection, we essentially focus on the qualitative analysis of optimal trajectories and control at entry and contact times, when $m = n_u = 1$. We denote u^* the control inside ν^*.

[5] $A^{-1}\nu = \{x \in \mathbb{R}^n |; Ax \in \nu\}$. It is not required that A be full rank.

Proposition 10. *Let assumptions* (i)–(iv) *stand.* **(i)** *Let $r^{wu} = n$. Then $\nu^* = \{0\}$. If the optimal trajectory has a boundary arc on $(\tau, \tau+\epsilon)$, $\epsilon \geq 0$, then $u^*(t) = 0$ on $(\tau, \tau+\epsilon)$.* **(ii)** *If $r^{wu} < n$ then $\nu^* = \{x^*(t) = W^{-1}\begin{pmatrix} 0^{r^{wu}} \\ \xi(t) \end{pmatrix}\} = \{z^*(t) \mid \bar{z}^* = 0\}$, $\dot{\xi}(t) = A_\xi \xi(t)$, and $u^*(t) = -(CA^{r^{wu}-1}B)^{-1}CA^{r^{wu}}W^{-1}\begin{pmatrix} 0^{r^{wu}} \\ \xi(t) \end{pmatrix}$.*

Proof. **(i)** From the definition of ν^*, given a state trajectory $x(t)$ such that $Cx(t) = 0$ on a positive time interval, one has to look for a control $u^*(\cdot)$ that makes $w^{(i)}(\cdot)$ equal to 0 for all integers $i \geq 0$. On $(\tau, \tau+\epsilon)$ one has $z_1^{(i)} = w^{(i)} = 0$. Since $r^{wu} = n$, $\bar{z} = z$ and thus $\nu^* \ni x^* = 0$ since $z = Wx$ with W full rank. Since $w^{(i)}(t) = CA^i x(t) + \sum_{j=1}^{i} CA^{j-1}Bu^{*(i-j)}(t)$, for all $i \geq 1$, we deduce that on $(\tau, \tau+\epsilon)$ one has $0 = \sum_{j=1}^{i} CA^{j-1}Bu^{*(i-j)}(t) = \sum_{j=r^{wu}}^{i} CA^{j-1}Bu^{(i-j)}$ for all $i \geq r^{wu}$. Since $m = n_u = 1$ it follows that all the terms CA^iB are scalars and $u^* = 0$. **(ii)** is proved from the z-dynamics and the fact that $\nu^* = \lim_{k \to r^{wu}} \nu^k$, with $\nu^{k+1} = \{x \in \nu^k \mid CA^k x = 0\}$ [18]. ■

When assumption (ii) is not satisfied, $r^{wu} = 1$ and $\nu^* = \{z \mid z_1 = 0\} = bd\Phi$. Point **(ii)** of Proposition 10 holds for any $1 \leq r^{wu} \leq n$.

Corollary 5. *Let $r^{wu} = n$ and the above assumptions hold. If the unconstrained optimal trajectory $x_{un}(\cdot)$ (the solution of* (4)(a) (b) *with $C = 0$) satisfies $x_{un}(t) \in \mathrm{Int}(\Phi)$ for all $t \in [0, T_1] \setminus \{t_k\}_{1 \leq k \leq l}$, $l < +\infty$, and $x_{un}(t_k) \in \nu^*$, then the solution $x(\cdot)$ of* (4) *is equal to $x_{un}(\cdot)$ if $u(\cdot)$ is restricted to be a smooth time function on* $[0, T_1]$.

Proof. We know from Proposition 9 that boundary arcs and contact states belong to ν^*, and from Proposition 10 that on boundary arcs and contact states $x = 0$ and $u = 0$. Any other state trajectory linking $\bar{x}_0$ to $\bar{x}_1$ in Φ with a smooth input u', will yield an integral action $I(u') > I(u)$. ■

If $D \neq 0$, one may recover the previous case by a simple transformation, however the new pair $(\bar{A}, \bar{B})$ is never controllable. The subset ν^* can nevertheless be computed [19, Section 6.7] [18, Section VII], as the following example shows.

Example 5. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and $w(t) = -x_1(t) + 1$. Then $\nu^* = \{(1, 0)\}$, and $u^* = 0$.

Proposition 11. *If assumption* (ii) *does not hold, i.e. $r^{wu} = 1$, then a (possibly discontinuous) time function $u(\cdot)$ is sufficient to keep $x(\cdot)$ inside Φ at entry times. At such instants $2 \geq \deg(\lambda)$.*

Proof. Since $\nu^* = bd\Phi = \{z_1 = 0\}$, it is obvious from the z-dynamics of (A, B, C) that a discontinuous $u(\cdot)$ is sufficient to keep $x(\cdot)$ in Φ. From (14), which is the $\tilde{z}$-dynamics of $(\tilde{A}, \tilde{B}, \tilde{C})$, we deduce that λ is a Dirac measure. Indeed $r = 2$ and $z_1(\cdot) = \{z_1\}(\cdot)$ is a function, and its derivative may be discontinuous at an entry time τ. Thus Dz_2 is a Dirac measure at τ, and so is λ. We note also that $x(\cdot)$ is continuous at τ, so that $\eta(\cdot)$ (which is a function) "concentrates" all the state $\tilde{x}$ jumps. ■

Proposition 11 will be refined when we examine what happens along optimal trajectories for odd-r^{wu} systems, see Corollary 10. Proposition 11 suggests that if assumption ii) is true (i.e. $r^{wu} \geq 2$), then at entry times in $\chi_{\mathrm{con}} \setminus \nu^*$ a distributional input with degree ≥ 2 will be needed to keep $x(\cdot)$ inside Φ. This is a little more subtle. The following results show that trajectories attain bdΦ in a specific way.

The sets inside $\chi_{\mathrm{con}} \setminus \nu^*$ such that the first nonzero derivative of $w(\cdot) = z_1(\cdot)$ is of order $2k+1$ are denoted as

$$\chi_{\mathrm{con}}^{2k+1} = \{z \in \mathbb{R}^n, z \prec 0 \mid z_1^{(i)} = 0 \text{ for all } 0 \leq i \leq 2k,\ \ w^{(2k+1)} = z_1^{(2k+1)} < 0\}.$$

In the same way

$$\chi_{\mathrm{con}}^{2k} = \{z \in \mathbb{R}^n, z \succ 0 \mid z_1^{(i)} = 0 \text{ for all } 0 \leq i \leq 2k-1,\ \ w^{(2k)} = z_1^{(2k)} > 0\}$$

and

$$\chi_{\mathrm{rel}}^{2k+1} = \{z \in \mathbb{R}^n, z \succ 0 \mid z_1^{(i)} = 0 \text{ for all } 0 \leq i \leq 2k, z_1^{(2k+1)} > 0\}.$$

From $z_1^{(i)}(t) = CA^i x(t) + \sum_{k=r^{wu}-1}^{i-1} CA^k Bu^{(i-1-k)}(t)$ one sees that these sets may depend on $u(\cdot)$ and its derivatives. A trajectory which attains $bd\Phi$ in χ_{con}^{2k} detaches immediately from $bd\Phi$ if the control $u(\cdot)$ is smooth, since the first nonzero derivative will keep its positive sign in a right-neighborhood of the contact time (hence it is a touch state). It is noteworthy that a trajectory which comes from $\mathrm{Int}(\Phi)$ and enters $\chi_{\mathrm{con}}^{2k+1}$ at τ satisfies $z_1^{(2j)}(\tau^-) \geq 0$ and $z_1^{(2j-1)}(\tau^-) \leq 0$ for all $0 \leq j \leq k$. There is a sign inversion at each differentiation order due to the left analyticity in the neighborhood of τ and the fact that all derivatives are zero at τ up to the order $2k$. An illustrative example with $r^{wu} = 3$ can be found in [11, Example 1.8]. A trajectory that leaves $bd\Phi$ at τ after a boundary arc has to do it with a non-analytic $z_1(\cdot)$ (otherwise by a simple Taylor expansion it follows that the trajectory remains in ν^*) and such that the first non-zero derivative of $z_1(\cdot)$ is positive, i.e. in a set $\chi_{\mathrm{rel}}^{2k+1}$.

Corollary 6. *Let a contact time $\tau \in [0, T_1]$ be such that $x(\tau) \in \chi_{\text{con}}^{2k+1}$ and $x(\cdot) \in \Phi$ in a right neighborhood of τ. Then if $2k+2 \leq r^{wu}$, $k \geq 0$, u has to be a distribution of degree $2 + r^{wu} - 2k$ with an atom at τ. If $2k+2 > r^{wu}$ then $u(\cdot)$ is a function of time in a neighborhood of τ, with its first nonzero derivative that is of order $2k+1-r^{wu}$ and is discontinuous at τ.*

Proof. From the z-dynamics, one has that $z_1^{(r^{wu}-1)}(\cdot) = z_{r^{wu}}(\cdot)$ depends explicitly on $u(\cdot)$. In particular $z_1^{(2k+1)}(\cdot) = \dot{z}_{2k+1}(\cdot) = z_{2k+2}(\cdot)$ and is a function of $u^{(2k+1-r^{wu})}(\cdot)$, which has then to be discontinuous to keep the trajectory inside Φ. ■

The result of Corollary 6 will be refined next depending on r^{wu}. In most of the studies (see e.g. [23,26,9,12]) the constraint that the contact with $bd\Phi$ occurs "tangentially" is added. This means simply that at a time τ at which the optimal trajectory attains $bd\Phi$, then $\bar{z}(\tau^-) = 0$, where we take the left limit as it could be that jumps occur (see Proposition 6).

Definition 3. A contact state at time $\tau \in [0, T_1]$ is said *tangential* if $\bar{z}(\tau^-) = 0$, and *hypertangential* if $\bar{\tilde{z}}(\tau^-) = 0$.

The meaning of the overbar is given in the paragraph after (14). If the contact is hypertangential, it follows from (30) and (31) that $\bar{\tilde{z}}(\tau^+) = 0$ (which does not mean that $\tilde{z}(\tau^+) \in \nu^*$). An hypertangential contact necessarily occurs in a set χ_{con}^i with $i \geq r$. If the contact states are forced to be hypertangential, then $\mathrm{d}\mathcal{J}_{\tilde{x}} = 0$ and the action in (41) is purely integral.

Proposition 12. *Let $r^{wu} = 2k$, $k \geq 1$, $\deg(\lambda) \leq 2$, and $i \geq 0$. Then all optimal trajectories which make contact with $bd\Phi$ at $t = \tau$ in the set $\chi_{\text{con}}^{r^{wu}+2i}$ are touch states if $u^{(2i)}(\cdot)$ is a continuous function at $t = \tau$. If contact occurs in $\chi_{\text{con}}^{r^{wu}+2i+1}$, then necessarily $i \geq k-1$ and necessarily $u^{(2i+1)}(\cdot)$ is discontinuous at τ.*

Proof. If contact occurs in $\chi_{\text{con}}^{r^{wu}+2i}$, then $z_1^{(r^{wu}+2i)} > 0$. If $u^{(2i)}(\cdot)$ is a continuous function at $t = \tau$, so are all $u^{(j)}(\cdot)$ with $0 \leq j \leq 2i-1$. Therefore from $z_1^{(r^{wu}+2i)} = CA^{r^{wu}+2i}BW^{-1}z + \sum_{k=0}^{2i} CA^{r^{wu}-1-k+2i}Bu^{(k)}$, we deduce that $z_1^{(r^{wu}+2i)}(\cdot)$ is a continuous function at τ. From the definition of $\chi_{\text{con}}^{r^{wu}+2i}$, it results that in an open right neighbourhood of τ, $z_1(t) > 0$. If contact occurs in $\chi_{\text{con}}^{r^{wu}+2i+1}$, then $z_1^{(r^{wu}+2i+1)} < 0$. If $u^{(j)}(\cdot)$, $0 \leq j \leq 2i+1$, are continuous at τ, then in an open right neighbourhood of τ, $z_1(t) < 0$. Also $\deg(\lambda) \leq 2$ implies that $\deg(u^{(r^{wu})}) \leq 2$. If $i \leq k-2$, then the trajectory escapes from Φ in an open right neighbourhood of τ, because the lower order derivatives of $u(\cdot)$ must be continuous at τ, and so is $z_1^{(r^{wu}+2i+1)}(\cdot)$. If $i \geq k-1$, the degree condition on λ can be respected even if $u^{(2i+1)}(\cdot)$ is discontinuous at τ. Moreover if $u^{(2i+1)}(\cdot)$ is continuous at τ then the trajectory escapes from Φ in an open right neighbourhood of τ. ■

It is noteworthy that $u^{(2i+1)}(\cdot)$ be discontinuous at τ when $i \geq k-1$ is not only a necessary condition, but this is also sufficient. Indeed the conditions secure that (4) does represent the optimal conditions for (1) and (2), and the positivity of $M^{(2r^{wu})}$ (see Lemma 1) assures that Lemma 3 holds. The case $r^{wu} = 2k+1$ can be analysed similarly. However it requires more care as Proposition 13 and Corollary 7 show. The case $r^{wu} = 1$ is treated in Proposition 11 (since then $\nu^* = bd\Phi$ we get that $\chi_{\text{con}} \setminus \nu^* = \emptyset$ so that the condition of Proposition 12 degenerates).

The next step is to study optimal trajectories which make contact with $bd\Phi$ in $\chi_{\text{con}} \setminus \nu^*$, and then are sent into ν^* (where possibly a boundary arc exists).

Proposition 13. *Let $m = n_u = 1$. Let an optimal trajectory make contact with $\mathrm{bd}\Phi$ at $t = \tau$, in the subset $\chi_{\text{con}}^{2r^{wu}-1}$, and be such that $z(\tau^+) \in \nu^*$ (thus τ is an entry time). Then necessarily r^{wu} is even. The same holds if $z(\tau^+) \in \chi_{\text{rel}}^{2r^{wu}-1}$.*

Proof. From the $\tilde{z}$-dynamics and Lemma 1, it follows that if λ is a measure at a time t_k, then $\mathrm{d}\mathcal{J}_r$ has an atom $\lambda_\tau = (-1)^{r^{wu}}(CA^{r^{wu}-1}B)^{-2}(z_r(\tau^+) - z_r(\tau^-))\delta_\tau$. If $z(\tau^+) \in \nu^*$, then $z_r(\tau^+) = 0$ from the definition of ν^*. Now since contact is made in $\chi_{\text{con}}^{2r^{wu}-1}$ we have that $z_1^{(r-1)}(\tau^-)\ (= z_r(\tau^-)) < 0$, whereas $z(\tau^+) \in \nu^*$ implies that $z_1^{(r-1)}(\tau^+) = 0$. Therefore if $r^{wu} = 2k+1$ for some $k \geq 0$, λ cannot be a positive measure. The last point can be proved similarly since then $z_r(\tau^+) > 0$. The proof is complete. ■

Let us notice that under the conditions of Proposition 13, a jump in $z_r(\cdot)$ is equivalent to a jump in $u^{(r^{wu}-1)}(\cdot)$, since the derivatives of smaller order of $u(\cdot)$ are time-continuous functions. Proposition 13 says that under the stated conditions (which in particular imply that $\deg(\lambda) = 2$ since $z_r(\cdot)$ jumps at τ), only systems with an even relative degree possess entry states, which is in accordance with the result of [26]. The second point of Proposition 13 implies in particular that odd-r^{wu} systems do not even possess touch points with $z(\tau^-) \in \chi_{\text{con}}^{2r^{wu}-1}$ and $z(\tau^+) \in \chi_{\text{rel}}^{2r^{wu}-1}$. Consequently when $\deg(\lambda) = 2$, it is excluded to get an entry time that is the accumulation of touch points. Proposition 13 concerns conditions under which λ is a singular measure. If the contact occurs in a set χ_{con}^i with $i \geq r$, then $\bar{\tilde{z}}(\tau^-) = 0$ and the problem of existence of a boundary arc is equivalent to the well-posedness of an LCP (see Section 5). The next Corollary is a consequence of Proposition 13.

Corollary 7. *Let $m = 1$, λ be a measure and $\{\eta\}(\tau^+) - \{\eta\}(\tau^-) = -C^{\mathrm{T}}\lambda_1$. If $\lambda_1 > 0$, then r^{wu} is an even integer.*

Compiling Corollary 7 and Proposition 11, one deduces that if $r^{wu} = 1$ and if $u(\cdot)$ has a discontinuity at an entry time τ, then necessarily $\lambda_1 = 0$, i.e. the costate $\eta(\cdot)$ is continuous at τ. Hence we retrieve the result of [9, Lemma 2.4 (2)]. Also from Proposition 4 and Corollaries 2 and 7, it follows that if $(A^{\mathrm{T}})^j\sigma_\eta(\tau) \in \mathrm{Ker}(B^{\mathrm{T}})$ for $0 \le j \le r^{wu} - 2$ and $u^{r^{wu}-1}(\cdot)$ jumps at τ, then r^{wu} is even. Notice that if $\deg(\lambda) = 2$, then contact cannot occur in a set χ^{2k}_{con} with $k < r^{wu}$, since r is even. If $\tilde{z}(\tau^-) \in \chi^{2k}_{\mathrm{con}}$, $k < r^{wu}$, and $\tilde{z}(\tau^+) \in \nu^*$, then $\deg(\lambda) \ge 3$.

Proposition 14. *Let $m = n_u = 1$, and $r^{wu} = 2k + 1$, $k \ge 0$. Then $\tilde{z}(\tau^-)$ is an entry state, only if* $\deg(\lambda) \le 1$ *or* $\deg(\lambda) \ge 3$.

Proof. One has $\tilde{z}(\tau^+) \in \nu^*$, hence $z_1^{(i)}(\tau^+) = 0$ for all $i \ge 0$. The fact that $\tilde{z}(\tau^-)$ is an entry state means that $\tilde{z}(\tau^-) \in \chi_{\mathrm{con}}$, so that $z_1(\tau) = 0$, and the first nonzero $z_1^{(i)}(\tau^-)$ is either > 0 or < 0. The higher order derivatives satisfy $z_1^{(2k)}(\tau^-) \ge 0$ and $z_1^{(2k+1)}(\tau^-) \le 0$. It follows that at $t = \tau$ the measures $\mathrm{d}\mathcal{J}_i$ in (32) satisfy $\mathrm{d}\mathcal{J}_{2k} \ge 0$ and $\mathrm{d}\mathcal{J}_{2k+1} = 0$. Proposition 13 says that if $\deg(\lambda) = 2$ at τ (i.e. $\mathrm{d}\mathcal{J}_r > 0$), then r^{wu} is even. A first situation is when $z_1^{(i)}(\tau^-) \ne 0$ for some $i < r - 1$. Depending on the values of the derivatives it is possible that $\mathrm{d}\mathcal{J}_{2k} > 0$ for $k < r^{wu}$ and that $\mathrm{d}\mathcal{J}_r = 0$ and $\deg(\lambda) \ge 3$. A second situation is when the contact is hypertangential (see Definition 3), so that $\mathrm{d}\mathcal{J}_i = 0$ for all $1 \le i \le r$ (in this last case the trajectory may either detach from the constraint, or stay on it with $g_r(t) > 0$ on some nonzero time interval, where $g_r(\cdot)$ is in (32)). Then $\deg(\lambda) \le 1$, i.e. λ is a function. ■

As an example, let us consider $r^{wu} = 3$. If $z_1(\tau^-) = 0$, $z_2(\tau^-) < 0$, and $z_3(\tau^-) = z_3(\tau^-) = z_3(\tau^-) = z_3(\tau^-) > 0$, i.e. $\tilde{z}(\tau^-) \in \chi^0_{\mathrm{con}}$, then $\mathrm{d}\mathcal{J}_2 = z_2(\tau^+) - z_2(\tau^-) = -z_2(\tau^-) > 0$ while all other $\mathrm{d}\mathcal{J}_i$ have no atom at τ. Thus $\deg(\lambda) = 6$. Another type of approach is detailed in Example 2, where this time $\deg(\lambda) = 3$. It immediately follows from Proposition 14 that if λ is restricted to be a measure, then systems with $r^{wu} = 2k + 1$ possess optimal trajectories such that $\tilde{z}$ and consequently $\tilde{x}$ are time continuous, because $\deg(\lambda) \le 1$. It is important to recall that there may exist trajectories of (4) with $x(\cdot)$ and $B^{\mathrm{T}}\eta(\cdot)$ continuous functions, and with $\deg(\lambda) \ge 3$, see Proposition 6 and Example 2. Once $\tilde{z}(\tau^+) \in \nu^*$, then the existence of a boundary arc relies upon the well-posedness of a LCP whose solution is $\lambda(t)$, $t > \tau$, see Section 5. But if $r^{wu} = 2k$, $k \ge 1$, nothing hampers an optimal trajectory to possess left accumulations of state jumps, when λ is a measure. In any case, left accumulations at an entry state can exist in the state variables $\{z_i\}(\cdot)$ with $i \le r - 1$.

Proposition 15. *Let $r^{wu} = 1$, $n_u = m = 1$. Assume that $\bar{x}_1$ is reachable from $\bar{x}_0$ with $\mathbf{U} = \mathbf{U}^0_{\mathcal{J}}$. Then the optimal trajectory and controller are time-continuous functions on* $[0, T_1]$.

Proof. The reachability assumption is fundamental, as Example 7 in the appendix shows. Since it holds one can search for optimal inputs which are functions, see (15). From Proposition 14 it follows that $\deg(\lambda)$ is either ≤ 1 or ≥ 3. From (14) it follows that $\deg(\lambda) \ge 3 \Rightarrow \deg(z_2) \ge 2$. This is in contradiction with the choice of admissible inputs $\mathbf{U}$. Thus $\deg(\lambda) \le 1$ so that both $x(\cdot)$ and $\eta(\cdot)$ are time-continuous on $[0, T_1]$. ■

This result is consistent with the fact that the HOSP solutions are in $\mathcal{T}_1([0, T_1])$ for $r = 2$, which implies that $\deg(\tilde{z}) \le 2$, hence $\deg(\lambda) \le 3$. As Example 7 demonstrates, a necessary condition for the framework that is developed in this paper around the HOSP, is that the zero dynamics $\dot{\xi}(t) = A_\xi\xi(t) + B_\xi z_1(t)$ be reachable from $\bar{\xi}_0$ to $\bar{\xi}_1$ with z_1 a function of time.

In [26,23], it is stated that if at a contact time $u^{r^{wu}-1}(\tau^+) \ne u^{r^{wu}-1}(\tau^-)$ with all the smaller order derivatives continuous, then odd-relative degree systems do not have a boundary arc. This optimal input jump condition implies that $\deg(\lambda) = 2$, as can be easily seen from the $\tilde{z}$-dynamics. From Proposition 14 it follows that $\deg(\lambda) \le 1$. However the conclusion that there is no boundary arc is premature, see Section 5.

Remark 8. These results are of the same nature as results in [8,23,26,33]. For instance, compiling Corollary 6 and Proposition 13, one finds that the sum of the order of the first discontinuous derivative of $u(\cdot)$ and of the relative degree r^{wu}, is always odd at an entry time when $\deg(\lambda) \le 2$ (see [32, Theorem 5.1, Corollary 5.2] for similar results). However our study is more general as it relies on a systematic and intrinsic to (A, B, C) partitioning of $bd\Phi$ as well as structural properties of the system (4). In contrast, the condition which allows one to prove a similar result as Proposition 13, namely [26, equ. (81)], does not involve the Markov parameter $M^{(r)}$, but is based on suitable Taylor expansions of $\frac{\partial H}{\partial u}$ and $w(\cdot)$. The partitioning of $bd\Phi$ also contains constraint qualifications of the form: there exist x and u such that $(Ax+Bu)^{\mathrm{T}}C^{\mathrm{T}} < 0$ for all $x \in bd\Phi$ (which is nothing else but $z_1 \in \chi^1_{\mathrm{con}}$). Interestingly enough, the work in [23] already defined sets similar to those in Lemma 5, but the developments remained at an embryonic stage. It is sometimes argued that boundary (or constrained) arcs are impossible for odd r^{wu} [40]. Proposition 13, Corollary 7 and Proposition 14, and the material in Section 5 show that this is more subtle. See also [8, Remark 4.11] which presents an example which proves that boundary arcs may exist for any r^{wu}. Notice that [33] study the problem with input constraints, so that the necessary conditions are not as in (4) since $u(t) = -K\mathrm{sgn}(B^{\mathrm{T}}\eta(t))$ for some K. The necessary conditions can still be written under a complementarity framework since the sign function lends

itself to a complementarity formulation:

$$\begin{aligned}
&\begin{pmatrix}\dot{x}(t)\\ \dot{\eta}(t)\end{pmatrix} = \begin{pmatrix}A & 0\\ Q & -A^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}x(t)\\ \eta(t)\end{pmatrix} + \begin{pmatrix}0\\ -C^{\mathrm{T}}\end{pmatrix}\lambda(t) + \begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}\begin{pmatrix}\lambda_1(t)\\ \lambda_2(t)\end{pmatrix}\\
&0 \le \begin{pmatrix}w(t) = Cx(t) + D\\ w_1(t)\\ w_2(t)\end{pmatrix} \perp \begin{pmatrix}\lambda(t)\\ \lambda_1(t)\\ \lambda_2(t)\end{pmatrix} \ge 0\\
&\lambda_1(t) + \lambda_2(t) = 1
\end{aligned} \tag{55}$$

where $w_1(\cdot)$ and $-w_2(\cdot)$ are the positive and negative parts of $B^{\mathrm{T}}\eta(\cdot)$, respectively. System (55) possesses a non uniform vector relative degree, since relay systems lend themselves to a description as relative degree 0 system [14]. The HOSP is presented in [1] for $r \ge 1$, and should be extended to $r = 0$ in order to encompass systems as (55). It is finally noteworthy that when input constraints are considered, then the material in Section 3.5 is meaningless, unless these constraints are considered almost everywhere except at contact or entry times.

The following concerns the complementarity conditions of Proposition 2 for an entry state.

Corollary 8. *Let $\tilde{z}(\tau^-) \in \chi_{\mathrm{con}}^{2k+1}$ and $\tilde{z}(\tau^+) \in \nu^*$, $k \le \frac{r+1}{2}$. Then $0 \le \{z_i\}(\tau^+) \perp \mathrm{d}\nu_i(\{\tau\}) \ge 0$ for all $1 \le i \le 2k+2$. Moreover $\mathrm{d}\mathcal{J}_{2k+2} > 0$ if $k \le r^{wu} - 1$ and r^{wu} must be even if $k = r^{wu} - 1$.*

Proof. Using the tangent cones definition in Section 3.1, the definition of the sets $\chi_{\mathrm{con}}^{2k+1}$ and ν^*, and the material of Remark 4, it follows that $T_{\Phi}^{i}(\{z_1\}(t^-), \ldots, \{z_i\}(t^-)) = \mathbb{R}^+$ and $\partial\psi_{T_{\Phi}^{i-1}(\{z_1\}(t^-),\ldots,\{z_{i-1}\}(t^-))}(\{z_i\}(t^+)) = \mathbb{R}^-$ for all $1 \le i \le 2k+2$ (recall that $z_1^{(2k+1)}(\cdot) = z_{2k+2}(\cdot)$). The first result follows from (37). The second statement follows from (30) and (31), and using Lemma 1 and Proposition 13. ■

4.3. Exit times and states

Since the results which can be obtained for exit times are similar to those obtained for entry times, only one proposition is given in this section.

Proposition 16. *Let $\deg(\lambda) \le 2$. Then the connection between a boundary arc and an interior arc has to occur at an exit state in $\chi_{\mathrm{rel}}^{2k+1}$ with $k \ge r^{wu} - 1$.*

We omit the proof which relies on the same arguments as above proofs. Similar results concerning costate and input regularity conditions can be derived for the exit times, and are not given here for the sake of brevity. The following result applies to a contact state that is not an entry state, but a grazing state, and is a direct consequence of Proposition 2.

Corollary 9. *Let $\tilde{z}(\tau^-) \in \chi_{\mathrm{con}}^{2k+1}$ and $\tilde{z}(\tau^+) \in \chi_{\mathrm{rel}}^{2k+1}$, $k \le \frac{r+1}{2}$. Then the following CP is satisfied:*

$$0 \le \{z_{2k+1}\}(\tau^+) \perp -\mathrm{d}\nu_{2k+1}(\{\tau\}) \in \partial\psi_{T_{\Phi}^{2k}(\{z_1\}(\tau^-),\ldots,\{z_{2k}\}(\tau^-))}\left(\{z_{2k+1}\}(\tau^+)\right).$$

Corollaries 8 and 9 and Proposition 2 show that the additional constraints which may be imposed at contact times, are taken care of with the multipliers $\mathrm{d}\nu_i$, which in turn satisfy a special set of complementarity conditions, not present in (4). It is worth noting that the complementarity conditions satisfied by the multipliers $\mathrm{d}\nu_i$ depend both on the pre and post-contact states. This generalizes Nonsmooth Mechanics in the framework of which the percussion value (i.e. the magnitude of the Dirac measure at an impact, see Definition 1.2 in [11]) depends on both the pre and post-impact velocities. To the best of our knowledge this is introduced for the first time in the context of optimal control with inequality state constraints.

Remark 9. The introduction of the measure multipliers $\mathrm{d}\nu_i$, $1 \le i \le r$, generalises the case $r^{wu} = 1$ treated in [5,29] where a multiplier associated to $z_2(t)$ is introduced (denoted as $\dot{\mu}(t)$ in these papers, and which corresponds in our framework to the function $g_2(t)$). The conditions given for instance in [29, equ. (11b) (12b)] are a particular case of the conditions in Proposition 2 and especially (34) with $r = 2$. The inclusion (34) also includes [30, equ. (25)], which is stated directly along boundary intervals (t_k, t_{k+1}) and $\{t_k\}_{k\ge 0}$ is assumed to be a finite set in [30]. Let us consider the minimisation problem of Proposition 8, with $I(u)$ in (52). The measures $\mathrm{d}\nu_i$, $1 \le i \le r$, represent the Lagrange multipliers associated with the (unknown) contact times $t_k \in [0, T_1]$. As noted above, Proposition 2 shows that the function part $g_i(\cdot)$ of the measures $\mathrm{d}\nu_i$, $1 \le i \le r-1$, is zero a.e. Only the atomic part $\mathrm{d}\mathcal{J}_i$ may take positive values.

4.4. Multivariable systems

Let us assume that $m = n_u \geq 2$. The multivariable case differs in many points from the monovariable case. This section only aims at pointing out some of these discrepancies. We consider $C \in \mathbb{R}^{m\times n}$, $D = 0$ and $u \in \mathbb{R}^m$, with $m \geq 2$. We suppose that the system (A, B, C) has a uniform relative degree r^{wu} so that the z-dynamics can be written [48]. Then from Lemma 1 the same holds for $(\tilde{A}, \tilde{B}, \tilde{C})$ with relative degree $r = 2r^{wu}$. One has $mr^{wu} \leq n$ (equivalently $mr \leq 2n$), $z_i = (z_i^1, \ldots, z_i^m) \in \mathbb{R}^m$, $1 \leq i \leq r^{wu}$ (notice that $z_i = CA^jx$ for all $1 \leq j \leq r^{wu}$), $\xi \in \mathbb{R}^{n-mr^{wu}}$, $\tilde{\xi} \in \mathbb{R}^{2n-mr}$, $CA^{r^{wu}-1}B \in \mathbb{R}^{m\times m}$, $\tilde{C}\tilde{A}^{r-1}\tilde{B} \in \mathbb{R}^{m\times m}$ are full-rank matrices. Similar, more general, canonical forms can be derived (see e.g. [49,50]) for multivariable systems, but we will content ourselves with the one presented in [48]. One notices that if $mr^{wu} = n$ then there is no zero dynamics in the z-canonical form as defined above.

Let us introduce the following sets and functions, defined for each $i \in \{1, \ldots, m\}$:

$$r_{C_i} : bd\Phi \to \mathbb{N}$$

with

$$r_{C_i}(x) = \min\{j \in \mathbb{N} \mid w_i^{(j)}(x, u) \neq 0, u \in \mathbf{U}\} = \min\{r^{wu};\ \inf\{i \in \mathbb{N} \mid CA^ix \neq 0\}\},$$

and

$$\nu_{C_i} = \{x \in bd\Phi_i \mid r_{C_i}(x) = r^{wu}\} = \{x \in \mathrm{Ker}(C_i) \mid C_iA^jx = 0 \text{ for all } 1 \leq j \leq r^{wu}\}.$$

where $bd\,\Phi_i = \{x \in \mathbb{R}^n \mid C_ix = 0\}$. Notice that one has $bd\Phi = \Phi \cap \left\{\bigcup_{i=1}^m \{x \in \mathbb{R}^n \mid C_ix = 0\}\right\}$. When $m = 1$ and assumptions (i)–(iv) hold, one has $\nu_C = \nu^*$ [18, Lemma A.1]. However in the multivariable case one has

$$\nu_C = \left(\bigcup_{i=1}^m \nu_{C_i}\right) \cap bd\Phi$$

and

$$\nu^* = \bigcap_{i=1}^m \nu_{C_i}.$$

Example 6. Let us consider $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$, $R = I_2$, $Q = I_4$. Then $r^{wu} = 2$ (i.e. the vector relative degree is $(2\ \ 2)^{\mathrm{T}}$) and the leading Markov parameter of $(\tilde{A}, \tilde{B}, \tilde{C})$ is $M^{(4)} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} > 0$. The $\tilde{z}$-dynamics is (in the $\tilde{z}$ coordinates, one has $z_1 = (w_1w_2)^{\mathrm{T}} = (x_1x_1 + x_3)^{\mathrm{T}}$)

$$\begin{cases} \dot{z}_1^1(t) = z_2^1(t) \\ \dot{z}_2^1(t) = z_3^1(t) \\ \dot{z}_3^1(t) = z_4^1(t) \\ \dot{z}_4^1(t) = \eta_2(t) + x_1(t) + \lambda_1 + \lambda_2 \\ w_1(t) = z_1^1(t) \end{cases} \qquad \begin{cases} \dot{z}_1^2(t) = z_2^2(t) \\ \dot{z}_2^2(t) = z_3^2(t) \\ \dot{z}_3^2(t) = z_4^2(t) \\ \dot{z}_4^2(t) = \eta_2(t) - x_3(t) + \eta_4(t) - x_1(t) + \lambda_1 + 2\lambda_2 \\ w_2(t) = z_1^2(t) \end{cases} \tag{56}$$

with $z^1 = (z_1^1, z_2^1, z_3^1, z_4^1)^{\mathrm{T}}$, $z^2 = (z_1^2, z_2^2, z_3^2, z_4^2)^{\mathrm{T}}$ (these two vectors are needed to define the tangent cone as a product, see Remark 3), $z_i = (z_i^1 z_i^2)^{\mathrm{T}}$, $1 \leq i \leq 4$. For the system (A, B, C) one finds:

$$r_{C_1}(x) = 1 \quad \text{for all } x \text{ with } x_2 \neq 0, \qquad r_{C_1}(x) = 2 \quad \text{for all } x \text{ with } x_2 = 0,$$
$$r_{C_2}(x) = 1 \quad \text{for all } x \text{ with } x_2 + x_4 \neq 0, \qquad r_{C_1}(x) = 2 \quad \text{for all } x \text{ with } x_2 + x_4 = 0,$$
$$\nu_{C_1} = \{(0, 0, x_3, x_4)^{\mathrm{T}} \in \mathbb{R}^4\}, \qquad \nu_{C_2} = \{(x_1, x_2, -x_1, -x_2)^{\mathrm{T}} \in \mathbb{R}^4\}.$$

Consequently

$$\nu^* = \nu_{C_1} \cap \nu_{C_2} = \{0\}.$$

Theorem 1 on existence and uniqueness of solutions on $\mathbb{R}^+$ of (17)–(19) continues to hold in the multivariable case when $M^{(r)}$ is a nonsingular symmetric **M**-matrix [39, p. 227]. This secures that $M^{(r)}$ is positive definite and that all its entries $M_{ij}^{(r)} \geq 0$ for all $1 \leq i \leq m$.

4.4.1. Extension of Proposition 10

One has $\nu_{C_i} = \{x \in \mathrm{Ker}(C_i) \text{ and } x \in \mathrm{Ker}(C_iA^j) \text{ for all } 1 \leq j \leq r^{wu}\}$. Thus for all $z \in \nu_{C_i}$ one has $\bar{z}^i = (z_1^i, \ldots, z_r^i)^{\mathrm{T}} = 0^r$, $1 \leq i \leq m$, and for all $z \in \nu^*$ one has $\bar{z} = (z_1^{\mathrm{T}}, \ldots, z_r^{\mathrm{T}})^{\mathrm{T}} = 0^{rm}$. Thus if $mr = n$ one has that $\nu^* = \{0\}$ as in the monovariable case. We can conclude that in this case the extension of Proposition 10 to the multivariable case is straightforward. It is however noteworthy that assuming $m = n_u$ and a uniform relative degree is quite stringent. Canonical forms that are extensions of the z-dynamics for rectangular multivariable systems ($m \neq n_u$) or for square systems with non-uniform relative degree exist [49,50]. The study in [18]extends to the rectangular case. It is noteworthy that even if the system (A, B, C) is rectangular, the necessary condition system $(\tilde{A}, \tilde{B}, \tilde{C})$ is always square, but it may not possess a uniform relative degree. Keeping in mind that the HOSP is extended in [1] only to the square uniform relative degree case, the non-uniform relative degree and rectangular cases are rich sources of problems that are left for future work.

4.4.2. Extension of Proposition 13

As noted in [18], contact can occur on a portion of $\mathrm{bd}\Phi$, such that for instance $x(\tau^-) \in \chi_{\mathrm{con},1}^{2k} \cap \nu_2^*$. Then the optimal trajectory has a contact time with $bd\Phi_1$ but may have an entry time with $\mathrm{bd}\Phi_2$. Let $n_u = m$. If $x(\tau^-) \in \chi_{\mathrm{con},1}^{2k+1} \cap \chi_{\mathrm{con},2}^{2k+5}$ and $2k + 1 < r^{wu} < 2k + 5$, one optimal input will be distributional while the other one will be a function of time at τ. It may also happen that $\chi_{\mathrm{rel},j}^i \notin \mathrm{bd}\Phi$ for some constraint j and some $i \geq 1$ (see e.g. [19, Example 6.6.1]). An optimal trajectory with a boundary arc on the constraint j will possess an exit state with an optimal controller with reduced regularity, depending on i. Thus the qualitative behaviour of the optimal trajectory that is made above is still possible, but is more involved. It is however noteworthy that the HOSP state reinitialisation mapping in (30) and (31) still works, and brings an answer to the questioning about the collision map definition in [18, Section VI]. In the multivariable case ($m \geq 2$, uniform relative degree r, definite Markov parameter $M^{(r)}$), the conclusions of Proposition 13 need to be refined. For instance, if contact occurs in $\chi_{\mathrm{con}}^{2r^{wu}-1} = \times_{i=1}^m \chi_{\mathrm{con},i}^{2r_i^{wu}-1} \cap bd\Phi$, which means that all $z_{r,i}(\tau^-) < 0$, and if $z(\tau^+) \in \nu^* = \cap_{i=1}^m \nu_{C_i i}$, then r^{wu} does not necessarily have to be even. It has to be even if the leading Markov parameter $M^{(r)}$ is diagonal. But in general if $M^{(r)}$ is negative definite (without being diagonal), then the reasoning in the proof of Proposition 13 may fail as $(M^{(r)})^{-1}(-z_r(\tau^-))$ may have positive components despite $M^{(r)} < 0$ and all components of $-z_r(\tau^-)$ are positive.

Let us recall that a $m \times m$ strictly semimonotone matrix [17, Definition 3.3.9] (or **E**-matrix) satisfies $[v \in (\mathbb{R}^+)^m, v \neq 0] \Rightarrow [v_i > 0 \text{ and } (M^{(r)}v)_i > 0 \text{ for some } i]$. This allows us to propose an extension of Proposition 13.

Proposition 17.
- (a) *Let $M^{(r)}$ be a nonsingular symmetric* **M***-matrix. Then it is possible that contact occurs in the set $\chi_{\mathrm{con}}^{2r^{wu}-1}$, and with $z(\tau^+) \in \nu^*$. This implies that r^{wu} is even.*
- (b) *Let $M^{(r)}$ be an* **E***-matrix. Then it is possible that contact occurs in the set $\chi_{\mathrm{con}}^{2r^{wu}-1}$, and with $\bar{\tilde{z}}(\tau^+) \in \nu^*$. If $-M^{(r)}$ is an* **E***-matrix, this is impossible.*

Proof. One first notices that contact occurring in $\chi_{\mathrm{con}}^{2r^{wu}-1}$ implies that $\deg(\lambda) \leq 2$, i.e. λ is a measure. Moreover, we recall that $\nu^* = \{0\}$. (a) From the property of the inverse of a nonsingular symmetric **M**-matrix the result follows. The second result is true since a nonsingular symmetric **M**-matrix is positive definite, and using Lemma 1. (b) From [17, Theorem 3.10.7 (b) (f)], we get that if $M^{(r)}$ is an **E**-matrix, then $[v \geq 0] \Rightarrow [M^{(r)}v \geq 0]$. The algebraic dynamics are $-z_r(\tau^-) = M^{(r)}\lambda_\tau$, where $\lambda_\tau \geq 0$ is the magnitude of $\mathrm{d}\mathcal{J}_r$ at τ. Since contact occurs in the set $\chi_{\mathrm{con}}^{2r^{wu}-1}$ one has that $-z_r(\tau^-) > 0$. It is therefore possible that the equality be satisfied for some $1 \leq i \leq m$. $-M^{(r)}$ being an **E**-matrix means that whatever $v \in (\mathbb{R}^+)^m$, $v \neq 0$, there exists an index i such that $v_i > 0$ and $M^{(r)}v_i < 0$. This is a contradiction to $0 < -z_r(\tau^-) = M^{(r)}\lambda_\tau > 0$. ■

It is not possible in general to relate the **E-**property to positive or semi positive definiteness. Consequently, we cannot relate it neither to the fact that r^{wu} is even or odd, except if $m = 1$. Many other generalisations would be possible, like the one indicated above when the leading Markov parameter is diagonal. The considered classes of matrices should all possess some kind of "no sign reversing" property. We do not tackle them here for the sake of briefness. From a general point of view, the contact, touch and exit problems (with λ being a measure) may be formulated as follows:

Contact [Touch] (Exit) Problem: Find $\mathrm{d}\nu_r$ such that

$$\mathrm{d}z_r = \tilde{C}\tilde{A}^r\tilde{W}^{-1}z(\tau)\mathrm{d}t + M^{(r)}\mathrm{d}\nu_r$$

$$\mathrm{d}\nu_r \in -M^{(r)}\ \partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\ldots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+))$$

$$z_r(\tau^-) \in \chi_{\mathrm{con}}\ [\chi_{\mathrm{con}}]\ (\nu^*)$$

$$z_r(\tau^+) \in \nu^*\ [\chi_{\mathrm{rel}}]\ (\chi_{\mathrm{rel}})$$

where τ generically denotes a contact [touch] (exit) time. This can be easily generalised to the case when λ is not a measure, using the measure differential formalism (22)–(24). Proposition 17 gives some conditions under which this problem possesses or not a solution.

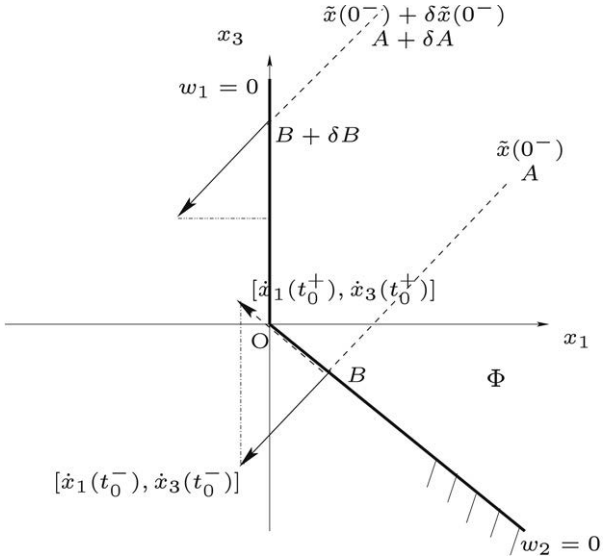

Fig. 1. Sensitivity of the BVP solutions w.r.t. initial data $\bar{x}_0$.

4.4.3. Continuous dependence of BVP solutions

It is well known from Mechanics that solutions of differential inclusions like the sweeping process, may be discontinuous with respect to initial conditions when $m \geq 2$ [31,42]. If ones embeds (4) into the HOSP, and proves that the solution is optimal with respect to an extended integral action, then the optimal pair $(x(\cdot), u(\cdot))$ may be very sensitive to the BVP data. It is noteworthy that this phenomenon may already exist when λ is a measure, i.e. when the HOSP is a measure differential inclusion, and should therefore be considered as an important feature. We shall say that the BVP in (4) is continuous with respect to data if small perturbations on the data $(\bar{x}_0, \bar{x}_1, T_1)$, result in a small perturbation of the solution $\tilde{z}$. More formally, let the sequences $\{\bar{x}_{0,n}\}$, $\{\bar{x}_{1,n}\}$, and $\{T_{1,n}\}$ converge in $\mathbb{R}$ towards $\bar{x}_0$, $\bar{x}_1$, and T_1 respectively. Let us consider the solutions of the BVP (4) as being elements in $\mathcal{F}_\infty([0, T_1], \mathbb{R})$, and solutions of the measure differential formalism in (22)–(24) (or (25) and (26)). Let us denote them as $\tilde{z}_n(\cdot)$ for the corresponding data. Then weak (resp. uniform) continuity in the data holds if and only if $\{\tilde{z}_n(\cdot)\}_{n\geq 0}$ converges weakly (resp. uniformly) towards $\tilde{z}(\cdot)$. The analysis which follows is essentially qualitative and illustrates why continuity w.r.t. the data may fail. Let us consider the same system as in Example 6.

Let us now consider a solution $\tilde{x}(\cdot)$ of the BVP (4), with initial condition $(\bar{x}_0, \eta(0^-))$ (point A on Fig. 1). Let us assume that $\tilde{x}(\cdot)$ hits the boundary $w_2 = 0$ at a time $t_0 < T_1$ in a neighbourhood of the origin (point B on Fig. 1), with the following data: $(z_1^1(t_0^-), z_1^2(t_0^-)) = (\bar{z}_1^1, 0)$, $\bar{z}_1^1 > 0$, $\dot{w}_2(t_0^-) = z_2^2(t_0^-) = \dot{x}_1(t_0^-) + \dot{x}_3(t_0^-) < 0$, $\dot{x}_1(t_0^-) < 0$, $\dot{x}_3(t_0^-) < 0$ and $\dot{x}_1(t_0^-) = \dot{x}_3(t_0^-)$ (a normal contact), $z_2^2(t_0^-) < 0$, $z_3^2(t_0^-) = 0$, $z_4^2(t_0^-) < 0$, whereas $z_2^1(t_0^-) \in \mathbb{R}$, $z_3^1(t_0^-) \in \mathbb{R}$, $z_4^1(t_0^-) \in \mathbb{R}$. This way we assume that the contact is made in the set $\chi^1_{\text{con},2}$. The tangent cones can be computed as indicated in Section 3.1 (see in particular Remarks 3 and 4), and we find $T^0_\Phi = \Phi = \mathbb{R}^+ \times \mathbb{R}^+$, $T^1_\Phi(z_1(t_0^-)) = T_\Phi(t_0^-) = \mathbb{R} \times \mathbb{R}^+$, $T^2_\Phi(z_1(t_0^-), z_2(t_0^-)) = T_\mathbb{R}(z_2^1(t_0^-)) \times T_\mathbb{R}(z_2^2(t_0^-)) = \mathbb{R} \times \mathbb{R}^+$, $T^3_\Phi(z_1(t_0^-), z_2(t_0^-), z_2(t_0^-)) = \mathbb{R} \times \mathbb{R}^+$. It follows from (30) and (31) that $z_1^1(t_0^+) = z_1^1(t_0^-)$, $z_1^2(t_0^+) = z_1^2(t_0^-)$, $z_2^2(t_0^+) = 0$, $z_2^1(t_0^+) = z_2^1(t_0^-)$, $z_3^1(t_0^+) = z_3^1(t_0^-)$, $z_3^2(t_0^+) = 0$. Though $M^{(4)}$ does not satisfy the condition in Remark 5 (i), there are two possibilities: $z_4^1(t_0^+) \in \mathbb{R}$ and $z_4^2(t_0^+) = 0$ or $z_4^1(t_0^+) \in \mathbb{R}$ and $z_4^2(t_0^+) > 0$. Since it is assumed that B is in a neighbourhood of the origin, and since the trajectory is analytic in a right neighborhood of t_0, there exists a junction time $t_1 > t_0$ with contact either at the origin (the trajectory "slides" on the boundary $w_2 = 0$) or with the boundary $w_1 = 0$ at some point C that is also in a neighbourhood of the origin. In any case the trajectory will either continue to evolve along the boundary $w_1 = 0$, or hit it and leave it with a large tangential velocity. If the optimal trajectory detaches at B and hits the boundary $w_1 = 0$ at t_1 at some point C, then the "velocity" is reinitialised to $\dot{x}_1(t_1^+) = 0$ and $\dot{x}_3(t_1^+) = \dot{x}_1(t_1^-) + \dot{x}_3(t_1^-)$. Since we can assume that the trajectory evolves arbitrarily close to the corner at the origin, it follows that $\dot{x}_3(t_1^+)$ has a magnitude that is of the same order as that of $\dot{x}_1(t_0^+) = \dot{x}_1(t_0^-)$ (because $z_2^1(t_0^+) = z_2^1(t_0^-)$). Therefore given any $\epsilon > 0$, it is possible to initialise the trajectory in A with a large enough $|\dot{x}_1(t_0^-)|$, so that at a time $t_1 + \delta t_1 > t_1 > t_0$, it enters the domain $D_\epsilon = \{0 \leq x_1 \leq \epsilon, x_3 \geq \epsilon\}$. The same reasoning for the trajectory starting at $A + \delta A$ can be done to conclude that it enters $\bar{D}_\epsilon = \{x_1 \geq \epsilon, 0 \leq x_1 + x_3 \leq \epsilon\}$ after a finite time. It is noteworthy that $\delta\tilde{x}_0$ may be arbitrarily small while ϵ in D_ϵ and $\bar{D}_\epsilon$ keeps its value. Whatever the behaviour at the corner may be, the continuous dependence on data as defined above cannot hold.

We conclude that provided the initial data satisfies some magnitude constraint, the unperturbed optimal trajectory starting at $(\bar{x}_0, \eta(0))$ reaches an end-point $(\bar{x}_1, \eta_1)$ in a time T_1, but the perturbed trajectory will not reach a neighbourhood of $(\bar{x}_1, \eta_1)$ in a time $T_1 + \delta T_1$. In order that it does, the BVP solution starting at $\bar{x}_0 + \delta\bar{x}_0$ has to be initialized with a costate

that is not a small perturbation of $\eta(0)$. Therefore the discontinuous dependence behaviour of the IVP solution, may create a jump in the initial condition of the BVP solution. Consequently, the optimal controller will also suffer from a "discontinuity".

5. Boundary arcs, relative degree $r^{wu} = 2k + 1, k \geq 0$

In the previous section, nothing was said about the influence of the sign of the leading Markov parameter on the behaviour of optimal trajectories on $bd\Phi$ (i.e. on boundary arcs in ν^*). We now study this aspect, i.e. we study what happens when a solution of the system in (4) has a boundary arc. From Lemma 1 one sees that systems in (2) with $r^{wu} = 2k + 1, k \geq 0$ yield $(\tilde{A}, \tilde{B}, \tilde{C})$ with a leading Markov parameter that is always (semi) negative definite. Let us investigate the consequences on the well-posedness of the dynamics on boundary arcs, i.e. those portions of optimal trajectories that satisfy $Cx(t) + D = 0$ for $t \in (\tau, \tau + \epsilon) \subseteq [0, T_1], \epsilon > 0$. On $(\tau, \tau + \epsilon)$ one has $u(t) = R^{-1}B^{\mathrm{T}}\eta(t)$ and $\dot{\eta}(t) = Qx(t) - A^{\mathrm{T}}\eta(t) - C^{\mathrm{T}}\lambda(t)$, with $\lambda(t)$ the solution of the LCP in (38) (equal to $g_r(t)$ in this case, where $g_r(\cdot)$ is in (32) and Proposition 2). One notices that the complementarity relations $0 \leq \{z_i\}(t^+) \perp g_i(t) \geq 0$ are trivially satisfied for all $1 \leq i \leq r - 1$ (see (33)).

Remark 10. Following [12, Sections 3.10, 3.11] let us define the Hamiltonian function $\bar{H}(x, u, \eta, \alpha) = H(x, u, \eta) + \alpha^{\mathrm{T}} z_{r^{wu}+1}$. On a boundary arc one has $z_{r^{wu}+1}(t) = \dot{z}_{r^{wu}}(t) = CA^{r^{wu}}x(t) + CA^{r^{wu}-1}Bu(t) = 0$. Thus we deduce that $u(t) = -(CA^{r^{wu}-1}B)^{-1}CA^{r^{wu}}x(t)$ (and consequently $B^{\mathrm{T}}\eta(t) = -B^{\mathrm{T}}(CA^{r^{wu}-1}B)^{-1}CA^{r^{wu}}x(t)$). Let us now compute the multiplier $\alpha(\cdot)$ from $\frac{\partial \bar{H}}{\partial u} = 0$. One finds $\alpha(t) = (-1)^{r^{wu}+1}[M^{(r)}]^{-1}(CA^{r^{wu}}x(t) + CA^{r^{wu}-1}BB^{\mathrm{T}}\eta(t))$. Inserting the value for $B^{\mathrm{T}}\eta(t)$ one finds that $\alpha(t) = 0$ which is consistent with (33) since $\alpha(\cdot) = g_{r^{wu}+1}(\cdot)$, the multiplier associated with the coordinate $z_{r^{wu}+1}$.

Let us examine the case $r^{wu} = 1$. On boundary arcs one has $w(t) = \tilde{C}\tilde{x}(t) + D = 0$ and $\dot{w}(t) = \tilde{C}\tilde{A}\tilde{x}(t) = 0$, since $r = 2$, and $\ddot{w}(t) = \tilde{C}\tilde{A}^2\tilde{x}(t) + M^{(2)}\lambda(t) = 0$. The fact that the trajectory keeps on evolving on $bd\Phi$ or detaches from $bd\Phi$ is monitored by the LCP: $0 \leq \lambda(t) \perp \ddot{w}(t) = \tilde{C}\tilde{A}^2\tilde{x}(t) + M^{(2)}\lambda(t) \geq 0$. The fact that the dynamics on boundary arcs are well-posed or not is closely linked to the well-posedness of this LCP, which holds on the interval $(\tau, \tau + \epsilon)$. The following is true since $M^{(2)} < 0$:

Lemma 6. *Let $m = 1$. Consider the LCP: $0 \leq \lambda(t) \perp \tilde{C}\tilde{A}^2\tilde{x}(t) + M^{(2)}\lambda(t) \geq 0$. Then:*

- *If $\tilde{C}\tilde{A}^2\tilde{x}(t) < 0$, the LCP has no solution,*
- *If $\tilde{C}\tilde{A}^2\tilde{x}(t) = 0$, the LCP has one solution $\lambda(t) = 0$,*
- *If $\tilde{C}\tilde{A}^2\tilde{x}(t) > 0$, the LCP has two solutions $\lambda(t) = 0$ and $\lambda(t) = -(M^{(2)})^{-1}\tilde{C}\tilde{A}^2\tilde{x}(t)$.*

Let the state satisfy $\tilde{C}\tilde{A}^2\tilde{x}(\tau^+) > 0$ at an entry time $\tau \in [0, T_1)$. The last item shows that there may not be uniqueness of the solution to the optimal control problem, if the optimal trajectory grazes the constraint boundary $bd\Phi$. Either the trajectory "detaches" from $bd\Phi$ ($\lambda(t) = 0 \Rightarrow \ddot{w}(t) > 0$), or remains on $bd\Phi$ ($\lambda(t) > 0 \Rightarrow \ddot{w}(t) = 0$). This is not surprising, since the leading Markov parameter is negative. The problem loses its convexity and the system in (4) is no longer well-posed as an IVP. It seems that little attention has been paid in the literature to the fact that the multiplier λ is the solution of an LCP on boundary arcs. The boundary arc input satisfies $\dot{w}(t) = CAx(t) + CBu(t)$, however this may not be the optimal controller. We shall see later what may happen when $\tilde{C}\tilde{A}^2\tilde{x}(\tau) < 0$ at an entry time τ. Lemma 6 can be extended to the general case $r^{wu} = 2k + 1$ and $m \geq 1$, where the LCP in (38) has to be studied. In view of Propositions 13 and 14 and Corollary 7, there are severe restrictions for the existence of entry times for systems with odd r^{wu}. The problem of interest here is to investigate what may happen after a contact time.

Proposition 18. *Let $r^{wu} = 2k + 1, k \geq 0$, and let $m = 1$. Let $x(\tau^+) \in bd\Phi$. Then:*

- *If $\tilde{C}\tilde{A}^r\tilde{x}(\tau^+) < 0$, the trajectory leaves Φ in a right neighbourhood of τ.*
- *If $\tilde{C}\tilde{A}^r\tilde{x}(\tau^+) = 0$, τ may be an entry time followed by a grazing trajectory.*
- *If $\tilde{C}\tilde{A}^r\tilde{x}(\tau^+) > 0$, τ may be either a touch time ($\lambda(\tau^+) = 0$) or an entry time ($\lambda(\tau^+) > 0$).*

Consequently, optimal trajectories satisfy $\tilde{C}\tilde{A}^r x(\tau^+) \geq 0$.

Proof. The proof follows from (38), Lemmas 6 and 1. ■

Let us notice that it is possible that $\bar{\tilde{z}}(\tau^-) = 0$, which implies from (30) and (31) that $\bar{\tilde{z}}(\tau^+) = 0$, but that $\tilde{C}\tilde{A}^r x(\tau^+) < 0$, because of the value of the zero-dynamics state $\xi(\tau)$. In other words, odd-r^{wu} systems can possess entry times with $\deg(\lambda) \leq 1$ (see Proposition 14). However $\xi(\tau)$ has to be such that $\tilde{C}\tilde{A}^r x(\tau^+) \geq 0$ for a boundary arc to exist on the right of τ. Clearly for an even-r^{wu} system, the trajectory can be kept in Φ as there always exists a multiplier $g_r(\tau^+)$ solution of the LCP (38). The zero dynamics plays a major role in the well-posedness of the LCP and whether or not odd-r^{wu} systems possess boundary arcs. For instance if $r^{wu} = n$, the first item of Proposition 18 becomes irrelevant at an hypertangential contact state, because necessarily $\tilde{C}\tilde{A}^r x(\tau^+) = 0$.

Corollary 10. *Let $r^{wu} = m = n_u = 1$. If an entry time τ exists on $[0, T_1]$, then uniqueness of the solution of the necessary conditions system holds locally if and only if the trajectory along the boundary arc is grazing, that is $(CA^2 + CBB^{\mathrm{T}}Q)x(t) + (CABB^{\mathrm{T}} - CBB^{\mathrm{T}}A^{\mathrm{T}})\eta(t) = 0$ for all $t \in [\tau, \tau + \epsilon]$ where $\tau + \epsilon$ is the exit time.*

Proof. Follows from Proposition 18 and the calculation of $\tilde{C}\tilde{A}^2$. ■

Definition 4. A boundary arc with entry time τ and exit time $\tau + \epsilon$ is well-posed if the LCP: $0 \leq w^{(r)}(t) \perp \lambda(t) \geq 0$ possesses at least one solution for all $t \in (\tau, \tau + \epsilon)$.

On a boundary arc $\bar{\tilde{z}}^{\mathrm{T}} = (z_1, z_2, \ldots, z_r) = (0, \ldots, 0)$. Definition 4 applies to systems with $m \geq 1$. The next corollary concerns scalar systems as $\dot{x}(t) = ax(t) + bu(t)$, $w(t) = cx(t) + d$, $Q = R = 1$. Then the Hamiltonian dynamics are $\begin{pmatrix} \dot{x}(t) \\ \dot{\eta}(t) \end{pmatrix} = \begin{pmatrix} a & b^2 \\ 1 & -a \end{pmatrix} \begin{pmatrix} x(t) \\ \eta(t) \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \end{pmatrix} \lambda$. The LCP to be solved on boundary arcs is: $0 \leq \lambda \perp -b^2c^2\lambda - d(a^2 + b^2) \geq 0$.

Corollary 11. **(i)** *If $r^{wu} = n = 1$ and $cx_0 + d > 0$, $cx_1 + d = 0$, then the dynamics on boundary arcs are well-posed if and only if $d \leq 0$. Moreover $bd\Phi = \{x = -\frac{d}{c}\}$ is attained only at $t = T_1$ on the optimal trajectory.* **(ii)** *Let us now consider $n \geq 2$, $D = 0$ and $m = 1$ in (1). If $r^{wu} < n$, then the well-posedness of the dynamics along boundary arcs implies that $\tilde{C}\tilde{A}^2\tilde{W}^{-1}\begin{pmatrix} 0 \\ \xi(t) \end{pmatrix} \geq 0$, $\forall\, t \in (\tau, \tau + \epsilon)$.*

Proof. Since **(ii)** is an easy consequence of Lemma 6, we just prove **(i)**. Integrating the dynamics without constraint, one finds the optimal trajectory on $[0, T_1]$:

$$x(t) = \frac{x_0 \exp(-\sqrt{a^2 + b^2}T_1) - x_1}{\exp(-\sqrt{a^2 + b^2}T_1) - \exp(\sqrt{a^2 + b^2}T_1)} \exp(\sqrt{a^2 + b^2}t) + \frac{x_1 - x_0 \exp(\sqrt{a^2 + b^2}T_1)}{\exp(-\sqrt{a^2 + b^2}T_1) - \exp(\sqrt{a^2 + b^2}T_1)} \exp(-\sqrt{a^2 + b^2}t). \tag{57}$$

The optimal control can be calculated with $\eta(t) = \frac{1}{b^2}(\dot{x}(t) - ax(t))$ and $u(t) = b\eta(t)$. It can be checked that $x(t) > -\frac{d}{c}$ for all $t \in [0, T_1)$. Consequently the boundary $\partial\Phi$ is attained only at $t = T_1$, and T_1 is a *contact* time. The LCP for λ on $bd\Phi$ is given by: $0 \leq \lambda \perp -d(a^2 + b^2) - c^2b^2\lambda \geq 0$. We deduce that if $-d < 0$ the LCP has no solution, if $d = 0$ then $\lambda = 0$ and if $-d > 0$ then two solutions $\lambda = 0$ and $\lambda = -\frac{d(a^2+b^2)}{c^2b^2}$ are possible. However the behaviour at T_1 has no consequence on the value of the action integral $I(u)$ since the optimal control and state trajectory are Lebesgue measurable functions. ■

The last condition in Corollary 11 can be written also as $\tilde{C}\tilde{A}^2\tilde{W}^{-1}\begin{pmatrix} 0 \\ \xi(\tau)\exp(\tilde{A}_\xi(t-\tau)) \end{pmatrix} \geq 0$. The matrix $\tilde{A}_\xi$ is characterised from the zero dynamics of (A, B, C), see Lemma 2. In case (i) one sees that $\nu^* = \{-\frac{d}{c}\}$ and $u^* = \frac{ad}{cb}$. From Proposition 9 boundary arcs and entry states belong to ν^*, which is consistent with Corollary 11(i).

Let us now consider the planar system $A = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $C = (c \;\; d)$, $D = e$. This system has transfer function $G(s) = \frac{c+ds}{s(s-b)-a}$, relative degree $r^{wu} = 1$ and $CB = d \neq 0$. Therefore $M^{(2)} = -d^2 < 0$. Let us take $Q = \mathrm{diag}(q_1, q_2)$. With this system we associate its optimality conditions as in (4) and we check that $\tilde{C}\tilde{A}\tilde{B} = -d^2$. We get $H_{w\lambda}(s) = \frac{(c+ds)(c-ds)}{(s^2+bs-a)(s^2-bs-a)-s^2+1}$.

Lemma 7. *Assume that $-ad^2 + c^2 + bcd \neq 0$ and that $c \neq 0$. The above planar system has a well-posed boundary arc on the time interval $[\tau, \tau + \epsilon) \subseteq [0, T_1]$ if and only if*

$$\begin{cases} \dot{\eta}_1(t) = \alpha\eta_2(t) - c\eta_1(t) + \delta \\ \dot{\eta}_2(t) = -(1 + d)\eta_1(t) + \beta\eta_2(t) \\ \eta(\tau) = \begin{pmatrix} e_3^{\mathrm{T}} \\ e_4^{\mathrm{T}} \end{pmatrix} \exp(\tilde{A}\tau)\tilde{x}(0) \\ \gamma\eta_2(t) - d\eta_1(t) + \epsilon \geq 0 \quad \textit{for all } t \in [\tau, \tau + \epsilon) \end{cases} \tag{58}$$

where $\alpha = \frac{q_1d^2+a^2d^2-ac^2-abcd}{-ad^2+c^2+bcd} - \frac{c}{d}\frac{abd^3-cd^2(q_2+a+b^2)+c^3}{ad^2-c^2-bcd}$, $\beta = \frac{-q_2dc+abd^2-bc^2-b^2cd+abd^3-cd^2(q_2+a+b^2)+c^3}{-ad^2+c^2+bcd}$, $\gamma = \frac{abd^3-cd^2(q_2+a+b^2)+c^3}{-ad^2+c^2+bcd}$, $\epsilon = -e\frac{a}{c}(c + bd)[ad^2(-ad^2 + c^2 + bcd) + 1]$, $\delta = -q_1\frac{ad^2e}{c}(-ad^2 + c^2 + bcd) - q_1\frac{e}{c}$.

Proof. On the boundary arc one has $\tilde{C}\tilde{x}(t) + D = 0$ and $\tilde{C}\tilde{A}\tilde{x}(t) = 0$, which respectively yield $cx_1(t) + dx_2(t) + e = 0$ and $adx_1(t) + (c + bd)x_2(t) + d\eta_2(t) = 0$. We deduce that $x_2(t) = \frac{-cd}{-ad^2+c^2+bcd}\eta_2(t) + ade(-ad^2 + c^2 + bcd)$, and $x_1(t) = \frac{d^2}{-ad^2+c^2+bcd}\eta_2(t) - \frac{ad^2e}{c}(-ad^2 + c^2 + bcd) - \frac{e}{c}$. Moreover from the values of $\tilde{A}$ and $\tilde{C}$ we get that

$$\begin{cases} \dot{\eta}_1(t) = q_1x_1(t) - a\eta_2(t) - c\lambda(t) \\ \dot{\eta}_2(t) = q_2x_2(t) - \eta_1(t) - b\eta_2(t) - d\lambda(t) \\ 0 \leq cx_1(t) + dx_2(t) + e \perp \lambda(t) \geq 0. \end{cases} \tag{59}$$

Along the boundary arc we have from Lemma 6: $\lambda(t) = \frac{1}{d^2}\tilde{C}\tilde{A}^2\tilde{x}(t)$ provided $\tilde{C}\tilde{A}^2\tilde{x}(t) \geq 0$. Now $\tilde{C}\tilde{A}^2\tilde{x}(t) = a(c+bd)x_1(t)+(cb+d(q_2+a+b^2))x_2(t)-d\eta_1(t)+c\eta_2(t)$. Combining the above values we get that $\lambda(t) = \frac{abd^3-cd^2(q_2+a+b^2)+c^3}{-d^2(ad^2-c^2-bcd)}\eta_2(t)-\frac{1}{d}\eta_1(t)$. Injecting all the calculated values for $x_1(t), x_2(t)$ and $\lambda(t)$ into (59), the result follows. The conditions are sufficient but also necessary for if $\gamma\eta_2(t) - \mathrm{d}\eta_1(t) + \epsilon < 0$ the boundary arc does not exist at time t. ■

Since the system in (58) is linear, of order two, it is quite possible in practice to test whether or not (58) holds. This provides conditions on the entry values $\eta_1(\tau)$ and $\eta_2(\tau)$, such that a boundary arc is well-posed or not. Clearly Lemma 7 does not answer the question whether or not the optimal trajectory possesses boundary arcs. Notice that $\tilde{x}(0)$ is not known as $\eta(0)$ is an unknown of the BVP.

6. Numerical solution of the BVP

The objective is to numerically solve the extremisation problem in Proposition 8, using the time-stepping scheme developed in [1] which approximates the solutions of the measure differential formalism in (22)–(24), whose solutions are functions of time. Though no complete convergence result has been presented yet for this time-stepping algorithm, it is shown in [1, Section 5] that it possesses strong properties. Convergence is proved in specific examples. Numerically solving the minimisation problem in (52) is equivalent to solving the MBVP in Section 3.6. Let us outline the way the MBVP may be discretised. For details and subtleties on the time-stepping scheme that is presented next, the reader is referred to [1, Section 5]. Since the time-stepping scheme is based on the $\tilde{z}$-dynamics, let us first rewrite the MBVP as:

$$\begin{cases} -\mathrm{d}\tilde{z} + \tilde{W}\tilde{A}\tilde{W}^{-1}\tilde{z}(t)\mathrm{d}t + \tilde{\mathcal{M}}\mathrm{d}\nu_{\tilde{z}} = 0 \\ \mathrm{d}\nu_{\tilde{z}} = (\mathrm{d}\nu_1, \ldots, \mathrm{d}\nu_r, 0_{2n-r})^{\mathrm{T}} \\ \mathrm{d}\nu_i \in -\partial\psi_{T_\Phi^{i-1}(z_1(t^-),\ldots,z_{i-1}(t^-))}(z_i(t^+)) \quad \text{for all } 1 \leq i \leq r \\ G(\tilde{z}(0), \tilde{z}(T_1)) = 0 \end{cases}$$

for some function $G(\cdot,\cdot)$ representing the boundary constraints, and where $\tilde{z} = \tilde{W}\tilde{x}$ and $\tilde{\mathcal{M}}$ is defined in Section 3.5. Let us define a finite partition $0 = \tau_0 < \tau_1 < \cdots < \tau_k < \cdots < \tau_N = T_1$ of $[0, T_1]$. For the sake of simplicity, we choose $h = \tau_{k+1} - \tau_k > 0$ so that $T_1 = Nh$. The approximation of $\tilde{z}(\tau_k)$ is denoted as $\tilde{z}_k$. Then the MBVP is discretised as follows:

Discrete MBVP :

$$\tilde{z}_{k+1} - \tilde{z}_k = h\tilde{W}\tilde{A}\tilde{W}^{-1}\tilde{z}_{k+1} + \tilde{\mathcal{M}}\mu_{k+1}, \quad 0 \leq k \leq N$$

$$\mu_{k+1} = (\mu_{1,k+1}, \ldots, \mu_{r,k+1}, 0_{2n-r})^{\mathrm{T}}$$

$$\mu_{i,k+1} \in -\partial\psi_{T_\Phi^{i-1}(z_{1,k},\ldots,z_{i-1,k})}(z_{i,k+1}), \quad 1 \leq i \leq r$$

$$G(\tilde{z}_0, \tilde{z}_N) = 0$$

It is crucial to note that the primary variables in this algorithm are the quantities $\mathrm{d}\tilde{z}((\tau_k, \tau_{k+1}])$ and the discretised multipliers $\mu_{i,k}$, which do not intend to approximate the Dirac measures at atoms of the measures $\mathrm{d}\nu_i$, but their magnitude, i.e. $\mu_{i,k+1} = \mathrm{d}\nu_i((\tau_k, \tau_{k+1}])$. They are therefore always bounded quantities. Moreover this time-stepping scheme is not an implicit Euler scheme, though it looks like [2]. This time-stepping scheme approximates the solutions of the measure differential formalism, which are functions in the set $\mathcal{F}_\infty(\mathbb{R}^+; \mathbb{R})$. The practical implementation of the algorithm is detailed in [1, Section 5.4]. The next step is to construct a multiple shooting algorithm starting from the Discrete MBVP. This will be the object of a future work dedicated to the numerical aspects of this optimal control problem.

7. Conclusions

This paper concerns the optimal control of linear invariant systems, with state inequality constraints. This is a major topic, which has continuously attracted the interest of researchers for more than 70 years. The main contributions are the following. The Bolyantskii–Pontryagin necessary conditions are embedded into a distributional differential inclusion framework (the higher order Moreau's sweeping process [1]) which allows one to clearly understand their dynamics and provides a clear formalism for the derivation of an extended action to be minimised (allowing for costate and/or inputs and/or system's state, to belong to a specific set of Schwarz' distributions). The powerful geometric tools introduced by ten Dam in the context of unilaterally constrained controlled dynamical systems [18,19], are used to improve the qualitative analysis of optimal trajectories and controllers, both in mono- and multivariable cases. They allow us to generalise results concerning the specificity of odd relative degree systems (one of the results being that odd relative degree systems can possess boundary arcs with entry times such that the multiplier associated to the unilateral constraint is a distribution of degree ≥ 3; then the action to be minimised has to be extended to incorporate the state and/or costate jumps). The theory

of complementarity problems is shown to be quite useful in order to better understand the behaviour of optimal path on the boundary of the admissible domain. Most of the tools which are used in this paper (distributional inclusion, ten Dam's geometrical study, complementarity problems, the role played by the zero dynamics) have never been introduced before in the context of optimal control with unilateral state constraints. Let us point out some open issues which might be the object of future works:

- Extend the qualitative study of trajectories at junction times to nonlinear systems (ten Dam's geometric approach extends to a class of nonlinear systems).
- Extend the study to rectangular or non-uniform relative degree systems.
- Design a multiple shooting algorithm based on the higher order sweeping process time-stepping discretization.

Notation and nomenclature: contact time and state (state on the boundary coming from an interior arc), entry time and state (contact state followed by a boundary arc), touch time and state (contact state coming from an interior arc and followed by an interior arc), junction time and state (entry, contact, exit, touch time and state), CP (complementarity problem), LCP (linear CP), ψ_K (the indicator function of a set K [25, p. 82]), $\partial\psi_K$ (the subdifferential of the indicator of a convex set K [22, p. 67]), supp(T) (the support of a distribution T [21, p. 142]), $x \succ 0$ (x is not zero and the first nonzero element of the vector x is positive), $x \succcurlyeq 0$ (the first nonzero element of the vector x is non negative), similarly for $\prec$ and $\preccurlyeq$, $\sigma_f(t) = f(t^+) - f(t^-)$ the jump of the function $f(\cdot)$ at time t, BV (bounded variation), RCLBV (right continuous of local bounded variation), RCSLBV (right continuous of special local bounded variation), BVP (boundary value problem), IVP (initial value problem), HOSP (higher order sweeping process), δ_t the Dirac measure at t, $0^m = (0, \dots, 0)^{\mathrm{T}} \in \mathbb{R}^m$, $0_m = (0^m)^{\mathrm{T}}$, I_n is the n-identity matrix, $0_{n\times n}$ is the zero $n \times n$ matrix. Contact times are denoted as τ (when only one contact is analysed) or t_k when a sequence of contact times is examined, $E_0(h)$ denotes the set of jumps of the function $h(\cdot)$. e_i is the i-th unit vector of $\mathbb{R}^n$; a function in C_0^∞ is infinitely differentiable and with compact support; continuous functions $\mathbb{R} \to \mathbb{R}$ are in $C^0[\mathbb{R};\mathbb{R}]$. $D\cdot$ denotes the distributional derivative, and $\partial\cdot$ may denote the subdifferential of a convex function, or the partial derivative, depending on the context. $bd\Phi$ denotes the boundary of Φ, i.e. $bd(\Phi) = \{x \mid Cx + D = 0\}$, Int($\Phi$) is its interior, i.e. $\mathrm{Int}(\Phi) = \{x \mid Cx + D > 0\}$.

Appendix A. Some mathematical definitions

Given $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^n$, the problem of finding $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ satisfying

$$y = Ax + B \geq 0, \quad x \geq 0, x^{\mathrm{T}}y = 0 \tag{60}$$

is called a Linear Complementarity Problem (LCP). It can be equivalently written as

$$0 \leq x \perp y = Ax + B \geq 0. \tag{61}$$

Roughly put, the LCP has a unique solution x^* whatever B if and only if A satisfies some positivity conditions, see [17]. Positive definiteness of A is sufficient. The so-called **P**-property of matrices is necessary and sufficient for the LCP to have a unique solution for any B. The variables x and y in (61) satisfy a CP.

The next notions may be found in [1,37,31] Let I denote a non-degenerate real interval (not empty nor reduced to a singleton).

• By $z \in BV(I;\mathbb{R}^n)$ it is meant that z is a $\mathbb{R}^n$-valued function of bounded variation if there exists a constant $C > 0$ such that for all finite sequences $t_0 < t_1 < \cdots < t_N$ (N arbitrary) of points of I, we have

$$\sum_{i=1}^{N} \|z(t_i) - z(t_{i-1})\| \leq C.$$

Let J be a subinterval of I. The real number

$$\mathrm{var}(z, J) := \sup \sum_{i=1}^{N} \|z(t_i) - z(t_{i-1})\|,$$

where the supremum is taken with respect to all the finite sequences $t_0 < t_1 < \cdots < t_N$ (N arbitrary) of points of J, is called the variation of z in J.

Any BV function has a countable set of discontinuity points and is almost everywhere differentiable. A BV function defined on $[a, b] \subset I$ possesses left-limits in $]a, b]$ and right-limits in $[a, b[$. Moreover, the functions $t \mapsto z(t^+) := \lim_{s\to t, s>t} z(s)$ and $t \mapsto z(t^-) := \lim_{s\to t, s<t} z(s)$ are both BV functions.

• We denote by LBV($I;\mathbb{R}^n$) the space of functions of locally bounded variation, i.e. of bounded variation on every compact subinterval of I.

• We denote by RCLBV($I;\mathbb{R}^n$) the space of right-continuous functions of locally bounded variation. It is known that if $z \in RC$LBV($I;\mathbb{R}^n$) and $[a, b]$ denotes a compact subinterval of I, then z can be represented in the form:

$$z(t) = \mathcal{J}_z(t) + [z](t) + \zeta_z(t), \quad \forall t \in [a, b],$$

where $\mathcal{J}_z$ is a jump function, $[z]$ is an absolutely continuous function and ζ_z is a singular function. Here $\mathcal{J}_z$ is a jump function in the sense that $\mathcal{J}_z$ is right-continuous and given any $\varepsilon > 0$, there exist finitely many points of discontinuity $t_1, \ldots, t_N$ of $\mathcal{J}_z$ such that $\sum_{i=1}^{N} \|\mathcal{J}_z(t_i) - \mathcal{J}_z(t_i^-)\| + \varepsilon > \mathrm{var}(\mathcal{J}_z, [a, b])$, $[z]$ is an absolutely continuous function in the sense that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^{N} \|[z](\beta_i) - [z](\alpha_i)\| < \varepsilon$, for any collection of disjoint subintervals $]\alpha_i, \beta_i] \subset [a, b] (1 \le i \le N)$ such that $\sum_{i=1}^{N} (\beta_i - \alpha_i) < \delta$, and ζ_z is a singular function in the sense that ζ_z is a continuous and bounded variation function on $[a, b]$ such that $\dot{\zeta}_z = 0$ almost everywhere on $[a, b]$.

• By $z \in \mathrm{RCSLBV}(I; \mathbb{R}^n)$ it is meant that z is a right-continuous function of special locally bounded variation, i.e. z is of bounded variation, and can be written as the sum of a jump function and an absolutely continuous function on every compact subinterval of I. So, if $z \in \mathrm{RCSLBV}(I; \mathbb{R}^n)$ then

$$z = [z] + \mathcal{J}_z \tag{62}$$

where $[z]$ is a locally absolutely continuous function called the absolutely continuous component of z and $\mathcal{J}_z$ is uniquely defined up to a constant by

$$\mathcal{J}_z(t) = \sum_{t \ge t_n} z(t_n^+) - z(t_n^-) = \sum_{t \ge t_n} z(t_n) - z(t_n^-) \tag{63}$$

where $t_1, t_2, \ldots, t_n, \ldots$ denote the countably many points of discontinuity of z in I.

Stieltjes (or differential) measure. Let $z \in \mathrm{LBV}(I; \mathbb{R}^n)$ be given. We denote by $\mathrm{d}z$ the Stieltjes measure generated by z. Recall that for $a \le b$, $a, b \in I$ one has $\mathrm{d}z([a, b]) = z(b^+) - z(a^-)$, $\mathrm{d}z([a, b[) = z(b^-) - z(a^-)$, $\mathrm{d}z(]a, b]) = z(b^+) - z(a^+)$, $\mathrm{d}z(]a, b[) = z(b^-) - z(a^+)$. In particular, we have $\mathrm{d}z(\{a\}) = z(a^+) - z(a^-)$.

Appendix B. Example of a function in $\mathcal{F}_\infty$ and distributions in $\mathcal{T}_n$

This is taken from [1]. Set $I = (0, +\infty)$ and let $z : I \to \mathbb{R}$ be the function given by

$$z(t) = |\sin(t)|, \quad \forall t \ge 0.$$

It is clear that

$$\hat{z}^{(0)} := z \in \mathrm{RCSLBV}(I; \mathbb{R})$$

since $z(\cdot)$ is Lipschitz-continuous. Then we obtain

$$\hat{z}^{(1)}(t) := \frac{\mathrm{d}^+}{\mathrm{d}t}[\hat{z}^{(0)}](t) = \frac{\mathrm{d}^+ z}{\mathrm{d}t}(t) = \cos(t - k\pi) \quad \text{if } t \in [k\pi, (k+1)\pi), (k \in \mathbb{N}).$$

We see that $E_0(\hat{z}^{(1)}) = \{k\pi; k \in \mathbb{N} \setminus \{0\}\}$ and

$$\hat{z}^{(1)}(\cdot) = [\hat{z}^{(1)}](\cdot) + J(\cdot),$$

where

$$[\hat{z}^{(1)}](t) = -2k + \cos(t - k\pi) \quad \text{if } t \in [k\pi, (k+1)\pi), (k \in \mathbb{N})$$

and

$$J(t) = 2k \quad \text{if } t \in [k\pi, (k+1)\pi), (k \in \mathbb{N}).$$

Thus

$$\hat{z}^{(1)}(\cdot) \in \mathrm{RCSLBV}(I; \mathbb{R}).$$

Then

$$\hat{z}^{(2)}(t) := \frac{\mathrm{d}^+}{\mathrm{d}t}[\hat{z}^{(1)}](t) = -|\sin(t)|$$

so that

$$\hat{z}^{(2)}(\cdot) \in \mathrm{RCSLBV}(I; \mathbb{R}).$$

And so on, we see that

$$\hat{z}^{(k)}(t) = \begin{cases} (-1)^m \hat{z}^{(0)}(t) & \text{if } k = 2m \\ (-1)^m \hat{z}^{(1)}(t) & \text{if } k = 2m+1, \end{cases} \qquad (m \in \mathbb{N}),$$

so that $\hat{z}^{(k)} \in \mathrm{RCSLBV}(I; \mathbb{R})$, $\forall k \in \mathbb{N}$, and thus

$$z(\cdot) \in \mathcal{F}_\infty(I; \mathbb{R}).$$

Let us now consider the distribution T defined by

$$\langle T, \varphi\rangle = \int_{-\infty}^{\infty} |\sin(t)|\varphi(t)\mathrm{d}t, \quad \forall\, \varphi \in C_0^\infty(I).$$

Then for a given function $\varphi \in C_0^\infty(I)$, we see that:

$$\begin{aligned}
\langle DT, \varphi\rangle &= \int_{-\infty}^{\infty} \hat{z}^{(1)}(t)\varphi(t)\mathrm{d}t = \sum_{k\in\mathbb{N}\cap\mathrm{supp}\{\varphi\}} \int_{k\pi}^{(k+1)\pi} \cos(t-k\pi)\varphi(t)\mathrm{d}t,\\
\langle D^2T, \varphi\rangle &= \int_{-\infty}^{\infty} \hat{z}^{(2)}(t)\varphi(t)\mathrm{d}t + \sum_{k\in\mathbb{N}\setminus\{0\}\cap\mathrm{supp}\{\varphi\}} (\hat{z}^{(1)}(k\pi^+) - \hat{z}^{(1)}(k\pi^-))\langle\delta_{k\pi}, \varphi\rangle\\
&= -\int_{-\infty}^{\infty} |\sin(t)|\varphi(t)\mathrm{d}t + 2\sum_{k\in\mathbb{N}\setminus\{0\}\cap\mathrm{supp}\{\varphi\}} \langle\delta_{k\pi}, \varphi\rangle,
\end{aligned}$$

and so on. We have:

$$\begin{aligned}
&T \equiv \{T\} = \hat{z}^{(0)} = |\sin(.)|, \deg(T) = 0,\\
&DT \equiv \{T^{(1)}\} = \{DT\} = \hat{z}^{(1)} = \cos(.-k\pi) \quad \text{on } [k\pi, (k+1)\pi)\ (k\in\mathbb{N}), \deg(DT) = 1,\\
&D^2T \equiv \langle\langle D^2T\rangle\rangle = -|\sin(.)| + 2\sum_{k\in\mathbb{N}\setminus\{0\}} \delta_{k\pi}, \deg(D^2T) = 2,\\
&\{T^{(2)}\} = \{D^2T\} = \hat{z}^{(2)} = -|\sin(.)|,
\end{aligned}$$

and

$$d\langle\langle D^2T\rangle\rangle = d\hat{z}^{(1)}.$$

Appendix C. Hints on the controllability of (2)

Example 7. Let us consider the system

$$\begin{cases} \dot{x}_1(t) = x_2(t)\\ \dot{x}_2(t) = u(t)\\ x_2(t) \geq 0. \end{cases} \tag{64}$$

Clearly if $\mathbf{U}$ is restricted to the set of bounded functions of time, then (A, B) is not controllable in the set $\Phi = \{x \mid x_2 \geq 0\}$. However if distributions of degree 3 (derivatives of Dirac measures in the sense of Schwartz' distributions) are allowed in $\mathbf{U}$, then controllability holds (but, the quadratic term in u in (1) is meaningless for such inputs). Indeed negative jumps in $x_1(\cdot)$ are then possible to reach a state $x(T_1)$ from $x(0)$ with $x_1(T_1) < x_1(0)$, $T_1 > 0$. In particular $x(T_1) = 0$ belongs to the reachable space from $x(0^-) > 0$. If the constraint is replaced by $x_1(t) + d \geq 0$, $d \in \mathbb{R}$, then controllability holds in Φ with $\mathbf{U}$ a set of functions. One sees that reachability does not hold in the set of inputs $\mathbf{U}^1_{\mathcal{J}}$ (see Definition 2) because the zero dynamics state $\xi = x_1$ needs to jump and the degree of the input (as a distribution in $\mathbf{U}^1_{\mathcal{J}}$) is ≤ 2. This is not permitted by the HOSP framework.

References

[1] V. Acary, B. Brogliato, D. Goeleven, Higher order Moreau's sweeping process: Mathematical formulation and numerical simulation, Math. Program. A 113 (2008) 133–217.
[2] V. Acary, B. Brogliato, Numerical Methods for Nonsmooth Dynamical Systems. Applications in Mechanics and Electronics, in: Lecture Notes in Applied and Computational Mechanics, vol. 35, Springer Verlag, Heidelberg, 2008.
[3] K.M. Anstreicher, Generation of feasible descent directions in continuous time linear programming, Systems Optimization Laboratory, Dept. of Operations Research, Stanford university, Technical Report SOL 83-18, September 1983.
[4] A.V. Arutyunov, S.M. Aseev, State constraints in optimal control. The degeneracy phenomenon, Systems Control Lett. 26 (1995) 267–273.
[5] D. Augustin, H. Maurer, Second order sufficient conditions and sensitivity analysis for the optimal control of a container crane under state constraints, Optimization 49 (2001) 351–368.
[6] R. Bellman, Bottleneck problems and dynamical programming, Proc. Natl. Acad. Sci. 39 (1953) 947–951.
[7] P. Berkmann, H.J. Pesch, Abort landing in windshear: Optimal control problem with third-order state constraint and varied switching structure, J. Optim. Theory Appl. 85 (1) (1995) 21–57.
[8] J.F. Bonnans, A. Hermant, Well-posedness of the shooting algorithm for state constrained optimal control problems with a single constraint and control, SIAM J. Control Optim. 46 (4) (2007) 1398–1430.
[9] B. Bonnard, L. Faubourg, G. Launay, E. Trélat, Optimal control with state constraints and the space shuttle re-entry problem, J. Dynam. Control Systems 9 (2) (2003) 155–200.
[10] B. Bonnard, E. Trélat, Une approche géométrique de contrôle optimal de l'arc atmosphérique de la navette spatiale, ESAIM Control Optim. Calc. Var. 7 (2002) 179–222.
[11] B. Brogliato, Nonsmooth Mechanics, 2nd ed., Springer Verlag, London, 1999. Erratum and addendum at: http://www.inrialpes.fr/bipop/people/brogliato/erratum.html.

[12] A.E. Bryson Jr, Y.C. Ho, Applied Optimal Control; Optimization, Estimation, and Control, Taylor and Francis, 1975, Revised printing.
[13] G. Buttazzo, D. Percivale, On the approximation of the elastic bounce problem on Riemannian manifolds, J. Differential Equations 47 (1983) 227–245.
[14] M.K. Çamlibel, Complementarity methods in the analysis of piecewise linear dynamical systems, Ph.D. Thesis ISBN: 90-5668-079 X, Tilburg University, NL, 2001.
[15] D. Cobb, Descriptor variable systems and optimal state regulation, IEEE Trans. Automat. Control 28 (5) (1983) 601–611.
[16] D. Cobb, C.J. Wang, A characterization of bounded-input bounded-output stability for linear time-varying systems with distributional inputs, SIAM J. Control Optim. 42 (4) (2003) 1222–1243.
[17] R.W. Cottle, J.S. Pang, R.E. Stone, The Linear Complementarity Problem, Academic Press, Computer Science and Scientific Computing, 1992.
[18] A.A. ten Dam, K.F. Dwarshuis, J.C. Willems, The contact problem for linear continuous-time dynamical systems: A geometric approach, IEEE Trans. Automat. Control 42 (4) (1997) 458–472.
[19] A.A. ten Dam, Unilaterally constrained dynamical systems, Ph.D. Thesis, Rijsuniversiteit Groningen, NL. Available at http://irs.ub.rug.nl/ppn/159407869, 1997.
[20] V.A. Dykhta, Optimal pulse control in models of economics and quatum electronics, Autom. Remote Control 60 (11) (1999) 1603–1613.
[21] J. Campos Ferreira, Introduction to the Theory of Distributions, in: Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 87, Addison Wesley Longman Limited, 1997.
[22] D. Goeleven, D. Motreanu, Y. Dumont, M. Rochdi, Variational and Hemivariational Inequalities: Theory, Methods and Applications; Volume I: Unilateral Analysis and Unilateral Mechanics, in: Nonconvex Optimization and its Applications, Kluwer Academic Publishers, 2003.
[23] W.E. Hamilton Jr, On nonexistence of boundary arcs in control problems with bounded state variables, IEEE Trans. Automat. Control 17 (3) (1972) 338–343.
[24] R.F. Hartl, S.P. Sethi, R.G. Vickson, A survey of the maximum principles for optimal control problems with state constraints, SIAM Rev. 37 (2) (1995) 181–218.
[25] J.B. Hiriart-Urruty, C. Lemaréchal, Fundamentals of Convex Analysis, Springer Grundlehren Text Editions, Berlin, 2000.
[26] D.H. Jacobson, M.M. Lele, J.L. Speyer, New necessary conditions of optimality for control problems with state-variable inequality constraints, J. Math. Anal. Appl. 35 (1971) 255–284.
[27] D. Kinzebulatov, Systems with distributions and viability theorem, J. Math. Anal. Appl. 331 (2007) 1046–1067.
[28] P. Lancaster, M. Tismenetsky, The Theory of Matrices, 2nd ed., Academic Press, London, 1985.
[29] K. Malanovski, H. Maurer, S. Pickenhain, Second-order sufficient conditions for state-constrained optimal control problems, J. Optim. Theory Appl. 123 (3) (2004) 595–617.
[30] K. Malanovski, H. Maurer, Sensitivity analysis fo optimal control problems subject to higher order state constraints, Ann. Oper. Res. 101 (2001) 43–73.
[31] M.D.P. Monteiro Marques, Differential Inclusions in Nonsmooth Mechanical Problems, Shocks and Dry Friction, in: Progress in Nonlinear Differential Equations and Their Applications, Birkhauser, 1993.
[32] H. Maurer, On optimal control problems with bounded state variables and control appearing linearly, SIAM J. Control Optim. 15 (3) (1977) 345–362.
[33] J.P. McDanell, W.F. Powers, Necessary conditions for joning singular and nonsingular subarcs, SIAM J. Control Optim. 9 (1971) 161–173.
[34] B.M. Miller, E.Y. Rubinovich, Impulsive Control in Continuous and Discrete-Continuous Systems, Kluwer Academic Publishers, Dordrecht, 2003.
[35] J.J. Moreau, An expression of classical mechanics, Ann. Inst. H. Poincaré Anal. Non Linéaire 6 (Suppl. 1–48) (1989).
[36] J.J. Moreau, Unilateral contact and dry friction in finite freedom dynamic, in: J.J. Moreau, P.D. Panagiotopoulos (Eds.), CISM Courses and Lectures No. 302 International Centre for Mechanical Sciences, Springer-Verlag, 1988, pp. 1–82.
[37] J.J. Moreau, Bounded variation in time, in: J.J. Moreau, P.D. Panagiotopoulos, G. Strang (Eds.), Topics in Nonsmooth Dynamics, Birkhauser, Basel, 1988, pp. 1–74.
[38] J.M. Murray, Existence theorems for optimal control and calculus of variations problems where the state can jump, SIAM J. Control Optim. 24 (3) (1986) 412–438.
[39] K.G. Murty, Linear complementarity, linear and nonlinear programming, Internet edition. http://www-personal.umich.edu/~murty/, University of Michigan, 1997.
[40] H.J. Oberle, W. Grimm, BNDSCO, A program for the numerical solution of optimal control problems, Report no 515 der DFVLR, Deutsche Forschungs und Versuchsanstalt für Luft und Raumfahrt e. V., 1989. Hamburger Beiträge zur Angewandten Mathematik, 2001.
[41] Y. Orlov, Schwartz' distributions in nonlinear setting: Applications to differential equations, filtering and optimal control, Mathematical Problems in Engineering 8 (4–5) (2002) 367–387.
[42] L. Paoli, A numerical scheme for impact problems with inelastic shocks: A convergence result in the multi-constraint case, in: C.C. Baniotopoulos (Ed.), Nonsmooth/Nonconvex Mechanics with Applications in Engineering, (Proc. of the Int. Conf. in memoriam of P.D. Panagiotopoulos, 5–6 July 2002), Editions Ziti, Thessaloniki, Greece, 2002, pp. 269–274.
[43] F.L. Pereira, G.N. Silva, Necessary conditions of optimality for vector-valued impulsive control problems, Systems Control Lett. 40 (2000) 205–215.
[44] A.F. Perold, Fundamentals of a continuous time simplex method, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, Technical Report SOL 78–26, December 1978.
[45] M.D.R. De Pinho, M.M.A. Ferreira, F.A.C.C. Fontes, An Euler–Lagrange inclusion for optimal control problems with state constraints, J. Dynam. Control Systems 8 (1) (2002) 23–45.
[46] P. Rinaldi, H. Schättler, An optimal control problem with state space constraints arising in the design of bipolar transistors, in: Proc. 41st IEEE Conf. on Decision and Control, Las Vegas, Nevada, USA, December, 2002, pp. 4722–4727.
[47] H. Robbins, Junction phenomena for optimal control with state variable inequality of third order, J. Optim. Theory Appl. 31 (1980) 85–99.
[48] P. Sannuti, Direct singular perturbation analysis of high-gain and cheap control problems, Automatica 19 (1) (1983) 41–51.
[49] P. Sannuti, H.S. Wason, Multipe time-scale decomposition in cheap control problems — Singular control, IEEE Trans. Automat. Control 30 (7) (1985) 633–644.
[50] P. Sannuti, A. Saberi, Special coordinate basis for multivariable linear systems — Finite and infinite zero structure, suqaring down and decoupling, Int. J. Control 45 (5) (1987) 1655–1704.
[51] P. Sattayatham, Strongly nonlinear impulsive evolution equations and optimal control, Nonlinear Anal. 57 (2004) 1005–1020.
[52] A.J. van der Schaft, J.M. Schumacher, An Introduction to Hybrid Dynamical Systems, in: Springer Verlag Lecture Notes in Control and Information Sciences, vol. 251, 2000.
[53] J.M. Schumacher, Complementarity systems in optimization, Math. Program. Ser. B 101 (2004) 263–295.
[54] G.N. Silva, R.B. Vinter, Necessary conditions for optimal impulsive control problems, SIAM J. Control Optim. 35 (6) (1997) 1829–1846.
[55] E.D. Sontag, Mathematical Control Theory; Deterministic Finite Dimensional Systems, 2nd ed., in: Springer Texts in Applied Mathematics, vol. 6, 1998.
[56] A. Steindl, W. Steiner, H. Troger, Optimal control of a retrieval of a tethered subsatellite, in: G. Rega, F. Vestroni (Eds.), IUTAM Symp. on Chaotic Dynamics and Control of Systems and Processes in Mechanics, in: Solid Mechanics and its Applications, Springer, Dordrecht, 2005, pp. 441–450.
[57] O. von Stryk, R. Burlisch, Direct and indirect methods for trajectory optimization, Ann. Oper. Res. 37 (1992) 357–373.
[58] F.A. Valentine, The problem of Lagrange with differential inequalities as added side conditions, in: Contributions to the Calculus of Variations, Chicago University Press, Chicago, 1937, pp. 407–448.
[59] R. Vinter, Optimal Control, in: Systems and Control: Foundations and Applications, Birkhauser, Boston, 2000.
[60] H. Wen, D.P. Jin, H.Y. Hu, Optimal feedback control of the deployment of a tethered subsatellite subject to perturbations, Nonlinear Dynam. 1 (4) (2008) 501–514.