Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

TuB09.2

Transformation of optimal control problems with a state constraint avoiding interior boundary conditions

Knut Graichen, Nicolas Petit, and Andreas Kugi

Abstract— A well–known problem in constrained optimal control is the presence of interior boundary conditions for constrained arcs which require a–priori knowledge of the optimal solution. This paper presents a saturation function method to transform an optimal control problem (OCP) with a state constraint into an unconstrained OCP in new coordinates. The approach allows a tangential entry and exit of constrained arcs without involving interior boundary conditions. An additional regularization term is used in the new OCP to avoid singular arcs which correspond to constrained arcs in the original OCP. Interestingly, the continuity order of the saturation function plays an important role for the existence of bounded trajectories which represent inverse images of the optimal constrained solution in the original coordinates.

I. INTRODUCTION

In general, numerical methods to solve optimal control problems (OCPs) can roughly be divided in two different classes. In *direct methods*, the OCP is discretized to obtain a finite–dimensional parameter optimization problem, see e.g. [1]–[7]. Well–known advantages of the direct approaches are the good numerical robustness with respect to the initial guess as well as the efficient handling of constraints.

On the other hand, *indirect approaches* are based on the calculus of variations and require the solution of a two–point boundary value problem (BVP), see e.g. [8]. Indirect methods are known to show a fast numerical convergence in the neighborhood of the optimal solution and to deliver highly accurate solutions, which makes them particularly attractive in aerospace industries [9]–[12]. However, the handling of constraints via Pontryagin's maximum principle [13] is in general non–trivial, since the overall structure of the BVP depends on the sequence between singular/nonsingular and unconstrained/constrained arcs, which requires a–priori knowledge of the optimal solution structure.

This paper proposes a saturation function approach to circumvent these problems and applies the method to a single–input optimal control problem (called OCP_x) with a single state constraint. In the first step, the system is transformed into the input–output normal form by using the state constraint as linearizing output. An equivalent OCP_y is derived in these coordinates, where the state constraint appears at top of the normal form cascade. The constraint is then incorporated by a saturation function and successive differentiation along the normal form cascade. These substitutions propagate through the internal dynamics and eventually define a new dynamics.

K. Graichen and A. Kugi are with Automation and Control Institute, Vienna University of Technology, A–1140 Vienna, Austria {graichen, kugi}@acin.tuwien.ac.at
N. Petit is with Centre Automatique et Systèmes, Ecole des Mines de Paris, 75272 Paris, France nicolas.petit@ensmp.fr

As a consequence, a new unconstrained $\mathrm{OCP}_\xi^\varepsilon$ is derived, which automatically satisfies the state constraint. An additional regularization term with parameter ε is introduced in $\mathrm{OCP}_\xi^\varepsilon$ to account for singular arcs, which correspond to constrained arcs in OCP_y and the original OCP_x. The unconstrained $\mathrm{OCP}_\xi^\varepsilon$ can be readily treated in the calculus of variations without the necessity of interior boundary conditions. Nevertheless, the saturation function guarantees a tangential entry and exit of constrained arcs.

An interesting property of $\mathrm{OCP}_\xi^\varepsilon$ is that the order of continuity of the saturation function plays an important role concerning the existence of bounded trajectories as inverse images of the optimal solution. More precisely, the continuity order of the saturation function and of the optimal constrained trajectory have to match which is a necessary condition for the existence of series expansions along the unconstrained arcs. An example is used to illustrate the approach and the influence of the saturation function continuity on the convergence of the trajectories for decreasing values of the regularization parameter ε.

The approach is presented for a single–input OCP_x with a single state constraint to rigorously investigate its properties. However, the concept is extendable to a multiple–input state–constrained OCPs, where the state constraints have a well–defined vector relative degree.

II. PROBLEM STATEMENT

This section states the optimal control problem OCP_x and transforms it into an equivalent OCP_y in input–output normal form, which is the basis for the subsequent sections.

A. Optimal control problem OCP_x

Consider the constrained optimal control problem OCP_x:

$$
\begin{cases}
\min\ J(u) := \varphi(x(T)) + \displaystyle\int_0^T L(x,u,t)\,\mathrm{d}t & (1)\\[2mm]
\text{s.t.}\ \dot{x} = f(x,u),\quad x \in \mathbb{R}^n,\quad u \in \mathbb{R} & (2)\\[2mm]
\quad x(0) = x_0,\quad \chi(x(T)) = 0 & (3)\\[2mm]
\quad c(x) \in [c^-, c^+] & (4)
\end{cases}
$$

The nonlinear single–input system (2) with $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is assumed to yield a unique state x for each input u, such that the cost (1) with $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $L : \mathbb{R}^n \times \mathbb{R} \times [0,T] \to \mathbb{R}$ can be stated as the functional $J(u)$. The final conditions $\chi : \mathbb{R}^n \to \mathbb{R}^l$ in (3) are of dimension $l \leq n$. The state constraint function $c : \mathbb{R} \to [c^-, c^+]$ as well as φ, L, f, χ, and c are sufficiently smooth. The final time T of OCP_x is fixed for the sake of simplicity. In the following, we assume the existence of a unique optimal solution (u^*, x^*) with

$$
J^* = \min J(u),\quad u^* = \operatorname{argmin} J(u). \tag{5}
$$

B. Normal form representation

Following [14], [15], the relative degree $r \leq n$ of the constraint function $c(x)$ at a point $x = x^0$ is defined by

$$\frac{\partial}{\partial u} L_f^i c(x) = 0, \quad i = 1, \ldots, r-1, \quad \frac{\partial}{\partial u} L_f^r c(x) \neq 0,$$

where L_f denotes the Lie–derivative along f. We assume that the relative degree r is well–defined in a sufficiently large neighborhood of x^0. The constraint function $c(x)$ can be used as (partially) linearizing output to derive a change of coordinates

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} \theta_y(x) \\ \theta_z(x) \end{pmatrix} = \theta(x) \tag{6}$$

with $y = (y_1, \ldots, y_r)^\mathsf{T}$ and $\theta_y = (\theta_1, \ldots, \theta_r)^\mathsf{T}$ defined by

$$\theta_1(x) = c(x), \quad \theta_i(x) = L_f^{i-1} c(x), \quad i = 2, \ldots, r. \tag{7}$$

The additional coordinates $z = \theta_z(x) \in \mathbb{R}^{n-r}$ are necessary to complete the transformation (6) if $r < n$. In the new coordinates (6), OCP_x is transformed into an equivalent OCP_y in input–output normal form:

$$\begin{cases} \min \quad \bar{J}(u) := \bar{\varphi}(y(T), z(T)) + \int_0^T \bar{L}(y, z, u, t) \, \mathrm{d}t & (8) \\ \text{s.t.} \quad \dot{y}_i = y_{i+1}, \quad i = 1, \ldots, r-1 & (9) \\ \qquad \dot{y}_r = \alpha(y, z, u) & (10) \\ \qquad \dot{z} = \beta(y, z) & (11) \\ \qquad y(0) = y_0, \quad z(0) = z_0, \quad \bar{\chi}(y(T), z(T)) = 0 & (12) \\ \qquad y_1 \in [c^-, c^+] & (13) \end{cases}$$

where the initial conditions $y_0 = \theta_y(x_0)$, $z_0 = \theta_z(x_0)$, and $\bar{\varphi} = \varphi \circ \theta^{-1}$, $\bar{L} = L \circ \theta^{-1}$, $\bar{\chi} = \chi \circ \theta^{-1}$ follow from OCP_x with the change of coordinates (6). [1]

The integrators (9)–(10) with $\alpha = L_f^r c(x) \circ \theta^{-1}$ form the input–output dynamics, where the transformed constraint (13) appears at the top of the cascade. The second part (11) of the dynamics with $\beta_i = L_f \theta_{z,i}(x) \circ \theta^{-1}$, $i = 1, \ldots, n-r$ are the internal dynamics of the normal form.

The diffeomorphism (6) represents a (bijective) state transformation, such that OCP_y has a unique optimal solution (u^*, y^*, z^*) with

$$J^* = \min \bar{J}(u), \quad u^* = \operatorname{argmin} \bar{J}(u), \tag{14}$$

which is related to OCP_x via $y^* = \theta_y(x^*)$ and $z^* = \theta_z(x^*)$.

C. Constrained arcs

Along constrained arcs, where the optimal trajectory $y_1^*(t)$ stays at c^- or c^+ for a non–vanishing time interval $t \in [t_{in}, t_{out}]$, the following conditions have to hold:

$$y_1^*(t) = c^\pm, \quad y_i^*(t) = 0, \quad i = 2, \ldots, r, \quad \dot{y}_r^*(t) = 0, \tag{15}$$

see Fig. 1. The common approach to account for state constraints in the calculus of variations is to impose the conditions (15) at the entry point t_{in} (or exit point t_{out}) to ensure the tangential entry in the constrained arc, see,

---

[1] The notation "$\circ$" is consistently used throughout the text as substitution rule to replace a specific argument of a function $p(\cdot, v, \cdot)$ by a given transformation $v = q(w)$, i.e. $p(\cdot, q(w), \cdot) = p(\cdot, v, \cdot) \circ q$.
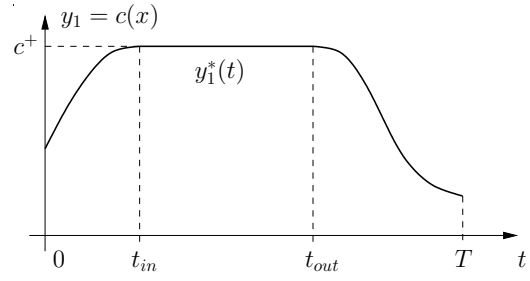

Fig. 1. Optimal trajectory $y_1^*(t)$ with active state constraint $y_1^* = c(x^*) = c^+$ for $t \in [t_{in}, t_{out}]$.

e.g., [8], [9]. The derivation of the optimality conditions then leads to a two–point boundary value problem (BVP) with additional interior boundary conditions at t_{in} or t_{out}.

In general, the overall structure of the BVP depends on the number and sequence of constrained and unconstrained arcs. Hence, the derivation of the BVP and the interior boundary conditions requires a–priori knowledge of the structure of the optimal solution (y^*, z^*, u^*).

Remark 1: *An important attribute is the order of continuity of $y_1^* \in \mathcal{C}^\rho$, which denotes how many times $y_1^*(t)$ can be continuously differentiated. The conditions (15) imply $\rho \geq r-1$ to ensure that all states y_i^*, $i = 1, \ldots, r$ are continuous over the junction points t_{in} and t_{out}. The explicit value can be characterized (under certain assumptions) with respect to the order r of the state constraint [16], [17].*

III. SATURATION FUNCTION APPROACH

In the following, a saturation function approach is presented, which circumvents the interior boundary conditions for the state constraint. The method systematically incorporates the constraint (13) within a new system representation by using a (sufficiently smooth) saturation function and differentiating along the normal form cascade (9)–(10). This method results in an unconstrained $\mathrm{OCP}_\xi^\varepsilon$ in new coordinates, where an additional regularization term with parameter ε accounts for singular arcs, which correspond to constrained arcs in the previous constrained OCP_y.

A. New system representation

In the first step, the state constraint $y_1 \in [c^-, c^+]$ is replaced by a saturation function

$$y_1 = \psi(\xi_1) \in [c^-, c^+] \tag{16}$$

with $\psi : \mathbb{R} \to [c^-, c^+]$ and the new unconstrained variable $\xi_1 \in \mathbb{R}$, see Figure 2. For $\xi_1 \in (\xi_1^-, \xi_1^+)$, the saturation function $\psi(\xi_1)$ stays inside the constraints (c^-, c^+) and is strictly monotonically increasing, i.e. $\frac{\mathrm{d}\psi}{\mathrm{d}\xi_1} > 0$. For $\xi_1 \leq \xi_1^-$ and $\xi_1 \geq \xi_1^+$, $\psi(\xi_1)$ is in saturation either at c^- or c^+ with vanishing derivatives $\frac{\mathrm{d}^i \psi}{\mathrm{d}\xi_1^i} = 0$, $i > 0$. As stated in Remark 2 at the end of this section, the saturation function $\psi(\xi_1) \in \mathcal{C}^\sigma$ has to be sufficiently smooth, where σ denotes the order of continuity of $\psi(\xi_1)$. The construction of $\psi(\xi_1)$ as used in this article is given in Appendix I.

In order to substitute the remaining coordinates $y_2, \ldots, y_r$, Eq. (16) is successively differentiated and new states ξ_{i+1} are introduced for the derivatives $\dot{\xi}_i = \xi_{i+1}$. When $\dot{y}_r$ is reached,
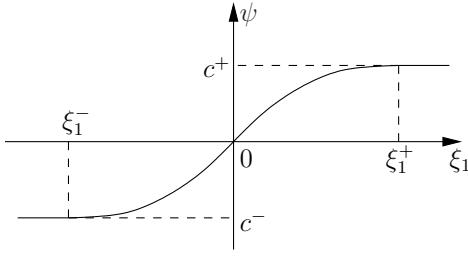
Fig. 2. Saturation function $\psi(\xi_1)$ with constraints $c^{\pm}$ and state ξ_1.

a new input $\tilde{u}$ is used for the final derivative $\dot{\xi}_r = \tilde{u}$. For instance, if the state constraint $c(x)$ is of order $r = 2$, the differentiation of (16) leads to

$$\begin{aligned}
y_2 &= \psi' \dot{\xi}_1, & \dot{\xi}_1 &= \xi_2, \\
\dot{y}_2 &= \psi'' \xi_2^2 + \psi' \dot{\xi}_2, & \dot{\xi}_2 &= \tilde{u},
\end{aligned} \tag{17}$$

with the compact notation $\psi' = \frac{\mathrm{d}\psi}{\mathrm{d}\xi_1}$ and $\psi'' = \frac{\mathrm{d}^2\psi}{\mathrm{d}\xi_1^2}$.

In summary, a new normal form cascade

$$\dot{\xi}_i = \xi_{i+1}, \quad i = 1, \ldots, r-1, \quad \dot{\xi}_r = \tilde{u}, \tag{18}$$

is generated with the states $\xi = (\xi_1, \ldots, \xi_r)^{\mathsf{T}}$ and the new input $\tilde{u}$. The coordinates $y = (y_1, \ldots, y_r)^{\mathsf{T}}$ and $\dot{y}_r$ are related to ξ and $\tilde{u}$ by

$$\begin{aligned}
y_1 &= h_1(\xi_1) & &= \psi(\xi_1), & (19\mathrm{a}) \\
y_i &= h_i(\xi_1, \ldots, \xi_i) = \gamma_i(\xi_1, \ldots, \xi_{i-1}) + \psi' \xi_i, & & & (19\mathrm{b}) \\
& & & i = 2, \ldots, r, \\
\dot{y}_r &= h_{r+1}(\xi, \tilde{u}) & &= \gamma_{r+1}(\xi) + \psi' \tilde{u}, & (19\mathrm{c})
\end{aligned}$$

where the relation between y and ξ can be written in the vector notation

$$y = h(\xi) = (h_1(\xi_1), \ldots, h_r(\xi))^{\mathsf{T}} \tag{20}$$

with $h : \mathbb{R}^r \to [c^-, c^+] \times \mathbb{R}^{r-1}$. The nonlinear terms γ_i in (19) follow from the successive application of the product and chain rule during the differentiations. For instance, (17) shows that $\gamma_2(\xi_1) = 0$ and $\gamma_3(\xi_1, \xi_2) = \psi'' \xi_2^2$. In the general case, the terms γ_i exhibit the particular structure

$$\gamma_i(\xi_1, \ldots, \xi_{i-1}) = \sum_{j=2}^{i-1} \frac{\mathrm{d}^j\psi}{\mathrm{d}\xi_1^j} \gamma_{i,j}(\xi_2, \ldots, \xi_{i-1}), \tag{21}$$

which can be proven by induction. The structure of the terms γ_i also shows that the relations (19b,c) evaluate to zero if $\psi(\xi_1)$ is in saturation with $\xi_1 \le \xi_1^-$ or $\xi_1 \ge \xi_1^+$ and exactly vanishing derivatives $\frac{\mathrm{d}^i\psi}{\mathrm{d}\xi_1^i} = 0$, $i \ge 1$, see Figure 2. Hence, the conditions (15) are automatically satisfied along constrained arcs.

The result of the differentiation procedure is that the new normal form dynamics (18) with state ξ and new input $\tilde{u}$ eventually replaces the input–output dynamics (9)–(10) and the constraint (13). The original input u can be recovered from (10) by substituting y and $\dot{y}_r$ by the single expressions (19):

$$u = h_u(\xi, z, \tilde{u}) = \alpha^{-1}(h(\xi), z, h_{r+1}(\xi, \tilde{u})). \tag{22}$$

Note that the inverse function α^{-1} exists due to the (assumed) well–defined relative degree.

Remark 2: *Equation* (21) *shows that the final coordinate y_r in* (19b) *involves derivatives of $\psi(\xi_1)$ up to order $r-1$. This implies that the order of continuity of $\psi(\xi_1) \in \mathcal{C}^\sigma$ must satisfy $\sigma \ge r-1$ to ensure continuity of $y_r(t)$ at the junction points t_{in} and t_{out} to constrained arcs.*

B. New regularized optimal control problem $\mathrm{OCP}_\xi^\varepsilon$

The new normal form (18) replaces the input–output dynamics (9)–(10), whereby the new states ξ also propagate through the internal dynamics (11) via the relations (20), cf. (27). The saturation function $y = \psi(\xi_1)$ thereby ensures that the constraint (13) is automatically satisfied.

In the new coordinates ξ, the cost functional (8) transforms to

$$\tilde{J}(\tilde{u}) := \tilde{\varphi}(\xi(T), z(T)) + \int_0^T \tilde{L}(\xi, z, \tilde{u}, t)\, \mathrm{d}t \tag{23}$$

with $\tilde{\varphi} = \bar{\varphi} \circ h$ and $\tilde{L} = \bar{L} \circ h \circ h_u$. However, special attention has to be paid to constrained arcs. As will be shown in more detail in Section IV, a constrained arc of OCP_y corresponds to a singular arc behavior in the new variables $(\xi, \tilde{u})$. This effect is compensated by adding a regularization term with the parameter ε to the cost (23), which leads to the regularized and unconstrained optimal control problem $\mathrm{OCP}_\xi^\varepsilon$:

$$\begin{cases}
\min \;\; P(\tilde{u}, \varepsilon) := \tilde{J}(\tilde{u}) + \varepsilon \int_0^T \tilde{u}^2\, \mathrm{d}t & (24) \\
\text{s.t.} \;\; \dot{\xi}_i = \xi_{i+1}, \quad i = 1, \ldots, r-1, & (25) \\
\quad\;\; \dot{\xi}_r = \tilde{u} & (26) \\
\quad\;\; \dot{z} = \tilde{\beta}(\xi, z) & (27) \\
\quad\;\; \xi(0) = \xi_0, \;\; z(0) = z_0, \;\; \tilde{\chi}(\xi(T), z(T)) = 0 & (28)
\end{cases}$$

where $\tilde{\beta} = \beta \circ h$ and $\tilde{\chi} = \bar{\chi} \circ h$ follow from OCP_y. To determine the initial conditions ξ_0, the relations (19a)–(19b) are solved for $\xi_1, \ldots, \xi_r$:

$$\begin{aligned}
\xi_1 &= h_1^{-1}(y_1) & &= \psi^{-1}(y_1) \\
\xi_i &= h_i^{-1}(\xi_1, \ldots, \xi_{i-1}, y_i) = \frac{y_i - \gamma_i(\xi_1, \ldots, \xi_{i-1})}{\psi'(\xi_1)} & & (29\mathrm{a})
\end{aligned}$$

with $i = 2, \ldots, r$. The single expressions (29a) are successively evaluated to obtain the initial values ξ_0 in (28), which can be written in the vector notation

$$\xi_0 = h^{-1}(y_0). \tag{29b}$$

Naturally, the first element $y_{1,0}$ of y_0 in (12) must strictly satisfy the constraint $y_{1,0} \in (c^-, c^+)$ to calculate the inverse saturation function ψ^{-1} and to ensure $\psi' > 0$ in (29a).

In practice, the new optimal control problem $\mathrm{OCP}_\xi^\varepsilon$ has to be solved successively with decreasing values of the regularization parameter $\varepsilon \to 0$.[2] The normal form coordinates y and finally the original input u and state x follow from the relations (20), (22), and the inverse transformation $x = \theta^{-1}(y, z)$ of (6).

[2]At this point, we assume solvability of $\mathrm{OCP}_\xi^\varepsilon$ for each $\varepsilon > 0$.

IV. OPTIMALITY CONDITIONS AND CONSTRAINED ARCS

This section derives the necessary optimality conditions, which can be used to solve $\text{OCP}_\xi^\varepsilon$. In particular, the structure of the optimality conditions reveals the behavior of $\text{OCP}_\xi^\varepsilon$ along constrained arcs and the influence of the regularization parameter ε.

A. Necessary optimality conditions

The necessary optimality conditions for $\text{OCP}_\xi^\varepsilon$ follow from the classical calculus of variations, see, e.g., [8]. Define the Hamiltonian

$$H(\xi, z, \lambda, \tilde{u}, t) = \tilde{L}(\xi, z, \tilde{u}, t) + \varepsilon \tilde{u}^2 + \sum_{i=1}^{r-1} \lambda_{\xi,i}\, \xi_{i+1} + \lambda_{\xi,r}\, \tilde{u} + \lambda_z^\mathsf{T}\, \tilde{\beta}(\xi, z)$$

with the adjoint states $\lambda_\xi = (\lambda_{\xi,1}, \ldots, \lambda_{\xi,r})^\mathsf{T}$ and $\lambda_z \in \mathbb{R}^{n-r}$ comprised in $\lambda^\mathsf{T} = (\lambda_\xi^\mathsf{T}, \lambda_z^\mathsf{T})$. Then, the optimal control $\tilde{u}$ has to satisfy

$$\frac{\partial H}{\partial \tilde{u}} = \frac{\partial \tilde{L}}{\partial \tilde{u}} + 2\varepsilon \tilde{u} + \lambda_{\xi,r} = 0. \qquad (30)$$

The partial derivative $\partial \tilde{L} / \partial \tilde{u}$ can be expressed in more detail

$$\frac{\partial \tilde{L}}{\partial \tilde{u}} = \left[\frac{\partial L}{\partial u} \circ h \circ h_u\right]\frac{\partial h_u}{\partial \tilde{u}} = \left[\frac{\partial L}{\partial u}\frac{\partial \alpha^{-1}}{\partial \dot{y}_r} \circ h \circ h_u\right]\psi'(\xi_1) \qquad (31)$$

using the relations (20) and (22). In addition, the adjoint system for λ is defined by $\dot{\lambda}_{\xi,i} = -\partial H/\partial \xi_i$, $i = 1, \ldots, r$ and $\dot{\lambda}_z^\mathsf{T} = -\partial H/\partial z$:

$$\dot{\lambda}_{\xi,1} = -\frac{\partial \tilde{L}}{\partial \xi_1} - \lambda_z^\mathsf{T}\frac{\partial \tilde{\beta}}{\partial \xi_1} \qquad (32a)$$

$$\dot{\lambda}_{\xi,i} = -\frac{\partial \tilde{L}}{\partial \xi_i} - \lambda_{\xi,i-1} - \lambda_z^\mathsf{T}\frac{\partial \tilde{\beta}}{\partial \xi_i}, \quad i = 2, \ldots, r \qquad (32b)$$

$$\dot{\lambda}_z^\mathsf{T} = -\frac{\partial \tilde{L}}{\partial z} - \lambda_z^\mathsf{T}\frac{\partial \tilde{\beta}}{\partial z} \qquad (32c)$$

with the terminal conditions

$$\lambda_\xi^\mathsf{T}(T) = \frac{\partial \tilde{\varphi}}{\partial \xi} + \nu^\mathsf{T}\frac{\partial \tilde{\chi}}{\partial \xi}, \quad \lambda_z^\mathsf{T}(T) = \frac{\partial \tilde{\varphi}}{\partial z} + \nu^\mathsf{T}\frac{\partial \tilde{\chi}}{\partial z} \qquad (33)$$

and the additional multipliers $\nu \in \mathbb{R}^l$. Using the transformations (20) and (22), the partial derivative of $\tilde{L}$ and $\tilde{\beta}$ become

$$\frac{\partial \tilde{L}}{\partial \xi_i} = \sum_{j=1}^{r}\left[\frac{\partial \bar{L}}{\partial y_j} \circ h \circ h_u\right]\frac{\partial h_j}{\partial \xi_i} + \left[\frac{\partial \bar{L}}{\partial u} \circ h \circ h_u\right]\frac{\partial h_u}{\partial \xi_i} \qquad (34a)$$

$$\frac{\partial \tilde{\beta}}{\partial \xi_i} = \sum_{j=1}^{r}\left[\frac{\partial \beta}{\partial y_j} \circ h\right]\frac{\partial h_j}{\partial \xi_i}. \qquad (34b)$$

The differential equations and boundary conditions (25)–(28) and (32)–(33) for the states (ξ, z) and $\lambda^\mathsf{T} = (\lambda_\xi^\mathsf{T}, \lambda_z^\mathsf{T})$ together with the algebraic equation (30) for $\tilde{u}$ defines a two–point boundary value problem (BVP), which has to be solved numerically to obtain the input $\tilde{u}$, the states (ξ, z, λ) and the multipliers ν. Note in particular that no interior boundary conditions are involved to account for the state constraint (13) due to its inherent incorporation by the saturation function $y_1 = \psi(\xi_1)$.

B. Behavior on constrained arcs

To illustrate the constrained arc behavior, we consider the case that one constrained subinterval exists for $\text{OCP}_\xi^\varepsilon$ with a given $\varepsilon > 0$, where $y_1 = \psi(\xi_1)$ stays at one of the constraints (13) for a non–vanishing time interval $t \in [t_{in}, t_{out}]$:

$$y_1(t) = \psi(\xi_1(t)) = c^\pm, \quad \xi_1(t) \gtrless \xi_1^\pm, \quad t \in [t_{in}, t_{out}]. \qquad (35)$$

Along the constrained arc, $\psi(\xi_1)$ is in saturation with $\xi_1 \geq \xi_1^+$ or $\xi_1 \leq \xi_1^-$, respectively, and the derivatives vanish, i.e. $\mathrm{d}^i\psi/\mathrm{d}\xi_1^i = 0$, $i \geq 1$, see Fig. 2. Hence, the terms in (31) and (34) evaluate to zero, cf. (19) and (21), which reduces the optimality conditions (30) and (32a)–(32b) to

$$\lambda_{\xi,r} + 2\varepsilon \tilde{u} = 0, \qquad (36)$$

$$\dot{\lambda}_{\xi,1} = 0, \quad \dot{\lambda}_{\xi,i} = -\lambda_{\xi,i-1}, \quad i = 2, \ldots, r. \qquad (37)$$

for $t \in [t_{in}, t_{out}]$. The reduced optimality condition $\partial H/\partial \tilde{u} = 0$ in (36) shows that the regularization term in the cost (24) is necessary to still be able to compute $\tilde{u}$. For $\varepsilon = 0$, $\text{OCP}_\xi^\varepsilon$ would be singular (i.e. $\partial^2 H/\partial \tilde{u}^2 = 0$) along constrained arcs. Moreover, note that the adjoint cascade (37) has the opposite direction to the normal form dynamics (25)–(26) for ξ.

Based on the simplified optimality conditions (36)–(37), the following lemma clarifies the structure of ξ and $\tilde{u}$ on constrained arcs:

Lemma 1: Along constrained subintervals $t \in [t_{in}, t_{out}]$, $\xi = (\xi_1, \ldots, \xi_r)^\mathsf{T}$ and $\tilde{u}$ are uniquely determined by time polynomial functions

$$\xi(t) = p(\xi(t_{in}^-), \xi(t_{out}^+), t), \qquad (38a)$$

$$\tilde{u}(t) = q(\xi(t_{in}^-), \xi(t_{out}^+), t), \quad t \in [t_{in}, t_{out}] \qquad (38b)$$

with respect to the junction values of the unconstrained arcs:

$$\xi(t_{in}^-) = \lim_{t\uparrow t_{in}} \xi(t), \quad \xi(t_{out}^+) = \lim_{t\downarrow t_{out}} \xi(t). \qquad (39)$$

Proof: The normal form dynamics (25)–(26) together with the simplified optimality conditions (36)–(37) can be written in the higher–order form

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r}\xi_1 = \xi_1^{(r)} = \tilde{u}, \quad \lambda_{\xi,r}^{(r)} = 0, \quad \lambda_{\xi,r} + 2\varepsilon \tilde{u} = 0. \qquad (40)$$

Using $\tilde{u} = -\lambda_{\xi,r}/2\varepsilon$ in $\xi_1^{(r)} = -\lambda_{\xi,r}/2\varepsilon$ and further differentiation leads to $\xi_1^{(2r)} = -\lambda_{\xi,r}^{(r)}/2\varepsilon$, which evaluates to $\xi_1^{(2r)} = 0$ for all $\varepsilon > 0$ due to $\lambda_{\xi,r}^{(r)} = 0$ in (40). Hence, $\xi_1^{(2r)}$ can be re–integrated $2r$–times, which yields a time polynomial of the form

$$\xi_1(t) = \sum_{i=0}^{2r-1} e_i\, t^i, \quad t \in [t_{in}, t_{out}]. \qquad (41)$$

The $2r$ coefficients e_i uniquely follow from the set of linear equations stemming from the $2r$ junction conditions $\xi(t_{in}^-) = \xi(t_{in})$ and $\xi(t_{out}^+) = \xi(t_{out})$. Hence, (41) can be written in the form $\xi_1(t) = p_1(\xi(t_{in}^-), \xi(t_{out}^+), t)$. The overall vector function $p = (p_1, \ldots, p_r)^\mathsf{T}$ and q in (38) follow from differentiation: $p_{i+1} = \frac{\mathrm{d}^i}{\mathrm{d}t^i}p_1$, $i = 1, \ldots, r-1$ and $q = \frac{\mathrm{d}^r}{\mathrm{d}t^r}p_1$. ∎

The polynomials p in (38a) smoothly connect the entry and exit points $\xi(t_{in}^-)$ and $\xi(t_{out}^+)$ to the unconstrained arcs. An interesting property thereby is that p is independent of the regularization parameter ε. However, it depends indirectly on ε due to the values of $\xi(t_{in}^-)$ and $\xi(t_{out}^+)$.

V. CORRESPONDENCE TO OCP$_y$ FOR $\varepsilon \to 0$

This section considers the limit problem $\varepsilon \to 0$ for OCP$_\xi^\varepsilon$ and the existence of bounded trajectories $(\tilde{u}^*, \xi^*, z^*)$ as inverse images of the optimal constrained solution (u^*, y^*, z^*) of OCP$_y$. We consider the case addressed in Sec. IV-B of one constrained subinterval. As will be shown, the order of continuity ρ of the saturation function $\psi(\xi_1) \in \mathcal{C}^\rho$ plays an important role in this context.

A. Unconstrained arcs and continuity of $\psi(\xi_1) \in \mathcal{C}^\sigma$

In the first step, consider the unconstrained time intervals $[0, t_{in}]$ and $(t_{out}, T]$. Since $\psi(\xi_1)$ is not in saturation, the states $\xi^*(t)$ and input $\tilde{u}^*(t)$ can be uniquely determined in terms of the optimal solution (u^*, y^*, z^*) of OCP$_y$

$$\xi^*(t) = h^{-1}(y^*(t)), \quad \tilde{u}^*(t) = h_{\tilde{u}}(y^*(t), z^*(t), u^*(t)), \quad (42)$$

where h^{-1} is defined in (29a). The function $h_{\tilde{u}}$ follows from inverting (19c) with (10)

$$h_{\tilde{u}}(y, z, u) = \frac{\alpha(y, z, u) - \gamma_{r+1}(h^{-1}(y))}{\psi'(\psi^{-1}(y_1))}. \quad (43)$$

Note that $\psi' > 0$ along unconstrained arcs, such that the inverse relations (42) and (43) are well–defined.

However, a closer look is required at the limit points t_{in}^- and t_{out}^+ to the time interval $[t_{in}, t_{out}]$ of the constrained arc. An important influence has the order of continuity σ of the saturation function $\psi(\xi_1) \in \mathcal{C}^\sigma$, which can be related to the continuity of the optimal constrained solution $y_1^*(t) \in \mathcal{C}^\rho$ (see Remark 1):

Lemma 2: *Assume there exist power series which equal $y_1^*(t)$ on the time intervals $[0, t_{in}]$ and $[t_{out}, T]$. Then, the matching condition*

$$\sigma = \rho \quad (44)$$

between $y_1^*(t) \in \mathcal{C}^\rho$ and $\psi(\xi_1) \in \mathcal{C}^\sigma$ is a necessary condition for the existence of power series of $\xi^*(t)$ on $[0, t_{in}]$, $[t_{out}, T]$.

In particular, the existence of series formulations of $\xi^*(t)$ implies boundedness of the junction values of the states

$$\xi^*(t_{in}^-) = \lim_{t \uparrow t_{in}} h^{-1}(y^*(t)), \quad \xi^*(t_{out}^+) = \lim_{t \downarrow t_{out}} h^{-1}(y^*(t)) \quad (45a)$$

and of the new input

$$\tilde{u}^*(t_{in}^-) = \lim_{t \uparrow t_{in}} h_{\tilde{u}}(y^*(t), z^*(t), u^*(t)),$$
$$\tilde{u}^*(t_{out}^+) = \lim_{t \downarrow t_{out}} h_{\tilde{u}}(y^*(t), z^*(t), u^*(t)). \quad (45b)$$

The proof is given in Appendix II. Note that $\sigma = \rho$ is only a necessary and not sufficient condition and that the statement in Lemma 2 concerns series formulations of $y_1^*(t)$ and $\xi_1^*(t)$. However, in certain cases where a finite time series of $y_1^*(t)$ exists, the boundedness of $\xi^*(t_{in}^-)$ and $\xi^*(t_{out}^+)$ can explicitly be shown for $\sigma = \rho$ (see Remark 4 in Appendix II and the example in Section VI).

B. Inverse images of optimal solution to OCP$_y$

The considerations for the unconstrained and constrained arcs in Sections IV-B and V-A can be pieced together to state the following result:

Proposition 1: *Under the assumption that the bounded junction values (45) exist, there are bounded trajectories $(\xi^*, z^*, \tilde{u}^*)$ with*

$$\xi^*(t) = \begin{cases} h^{-1}(y^*(t)) & \text{for } t \in [0, t_{in}), (t_{out}, T] \\ p(\xi^*(t_{in}^-), \xi^*(t_{out}^+), t) & \text{for } t \in [t_{in}, t_{out}] \end{cases} \quad (46)$$

$$\tilde{u}^*(t) = \begin{cases} h_{\tilde{u}}(y^*(t), z^*(t), u^*(t)) & \text{for } t \in [0, t_{in}), (t_{out}, T] \\ q(\xi^*(t_{in}^-), \xi^*(t_{out}^+), t) & \text{for } t \in [t_{in}, t_{out}] \end{cases} \quad (47)$$

which satisfy the system equations (25)–(27) by definition and are inverse images of the optimal solution (y^*, z^*, u^*) to OCP$_y$ via the relations (20), (22):

$$y^* = h(\xi^*), \quad u^* = h_u(\xi^*, z^*, \tilde{u}^*). \quad (48)$$

Moreover, the regularized cost (24) corresponds to the optimal cost (8) of OCP$_y$ according to

$$\tilde{J}(\tilde{u}^*) = \bar{J}(u^*) = J^*, \quad \lim_{\varepsilon \to 0} \varepsilon \int_0^T (\tilde{u}^*)^2 \, \mathrm{d}t = 0. \quad (49)$$

Proof: The trajectories of ξ^* and u^* are uniquely determined over the constrained and unconstrained arcs, see (38) and (42). The relations (48) are an immediate consequence and lead to the optimal cost $\tilde{J}(\tilde{u}^*) = \bar{J}(h_u(\xi^*, z^*, \tilde{u}^*)) = J^*$. Moreover, the penalty term in (24) vanishes for $\varepsilon \to 0$ due to the boundedness of $\tilde{u}^*$. ∎

Proposition 1 only considers the inverse images to the optimal solution (y^*, z^*, u^*) of OCP$_y$ and does not imply convergence. To prove that the solution of OCP$_\xi^\varepsilon$ actually converges towards $(\xi^*, z^*, \tilde{u}^*)$ for a decreasing sequence $\{\varepsilon^k\}$ would require additional convexity and solvability assumptions, see, e.g., [18].

Remark 3: *In practice, it may be difficult to meet the condition (44) without a–priori information of the optimal solution $y_1^*(t) \in \mathcal{C}^\rho$ (also see Remark 1). If instead the saturation function $\psi(\xi_1) \in \mathcal{C}^\sigma$ is chosen with $\sigma > \rho$, the trajectory $(\xi^*, \tilde{u}^*)$ may not exist due to unboundedness of the junction values (45). However, this is not necessarily a restriction, since the decreasing sequence $\{\varepsilon^k\}$ for the successive solution of OCP$_\xi^\varepsilon$ is usually stopped at a certain step k when the corresponding solution is sufficiently close to the optimal one, see Figure 4.*

VI. EXAMPLE

To illustrate the saturation function approach and the influence of the continuity of the saturation function $\psi(\xi_1) \in \mathcal{C}^\sigma$, consider the well–known double–integrator example [8]

$$\begin{cases} \min \quad J(u) := \frac{1}{2} \int_0^1 u^2 \, \mathrm{d}t & (50) \\ \text{s.t.} \quad \dot{y}_1 = y_2, \quad \dot{y}_2 = u, & (51) \\ \quad y(0) = (0, 1)^\mathsf{T}, \quad y(1) = (0, -1)^\mathsf{T} & (52) \\ \quad y_1 \le c^+ & (53) \end{cases}$$
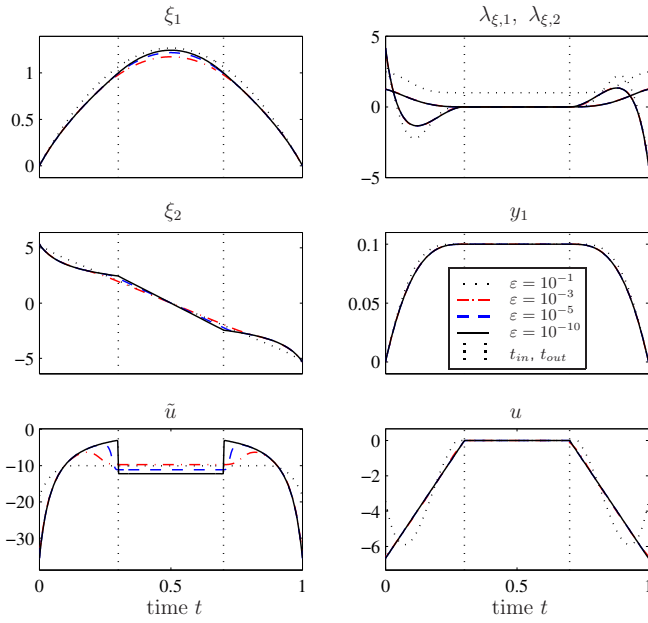
Fig. 3. Simulation results for $\psi(\xi_1) \in \mathcal{C}^2$ with decreasing ε and bounded junction points at t_{in}^- and t_{out}^+.

Fig. 4. Simulation results for $\psi(\xi_1) \in \mathcal{C}^3$ with decreasing ε and unboundedness of the junction points at t_{in}^- and t_{out}^+ for $\varepsilon \to 0$.

with the state constraint (53) of order $r = 2$. OCP_y has the analytic solution (for $c^+ \leq 1/6$)

$$y_1^*(t) = \begin{cases} c^+ - c^+\left(1 - \frac{t}{3c^+}\right)^3 & \text{if } t \in [0, 3c^+) \\ c^+ & \text{if } t \in [3c^+, 1 - 3c^+] \\ c^+ - c^+\left(1 - \frac{1-t}{3c^+}\right)^3 & \text{if } t \in (1 - 3c^+, 1], \end{cases} \quad (54)$$

which shows that the state constraint (53) is active on the subinterval $t \in [3c^+, 1 - 3c^+]$. The second optimal state y_2^* and optimal input u^* follow from differentiating (54) and yield the optimal cost $J^* = J(u^*) = 4/(9c^+)$.

The dynamics (51) are already given in the normal form (9)–(10) without the internal dynamics (11). The constrained state y_1 is replaced by the saturation function $y_1 = \psi(\xi_1) \in [c^-, c^+]$ with symmetric constraints $c^\pm$. Additional differentiation according to (17) then leads to the relations

$$y_1 = \psi(\xi_1), \quad y_2 = \psi'\xi_2, \quad u = \psi''\xi_2^2 + \psi'\tilde{u} \quad (55)$$

with the new dynamics $\dot{\xi}_1 = \xi_2$, $\dot{\xi}_2 = \tilde{u}$ and the new input $\tilde{u}$. In summary, the regularized $\text{OCP}_\xi^\varepsilon$ (24)–(28) reads

$$\begin{cases} \min \ P(\tilde{u}, \varepsilon) := \frac{1}{2}\int_0^1 (\psi''\xi_2^2 + \psi'\tilde{u})^2 + \varepsilon\tilde{u}^2 \, dt & (56) \\ \text{s.t.} \ \dot{\xi}_1 = \xi_2, \quad \dot{\xi}_2 = \tilde{u} & (57) \\ \quad \xi(0) = \left(0, \frac{1}{\psi'(0)}\right)^\mathsf{T}, \quad \xi(1) = \left(0, \frac{-1}{\psi'(0)}\right)^\mathsf{T}. & (58) \end{cases}$$

The boundary conditions (58) follow from the inverse relations (29a) with $\xi_1(0) = \psi^{-1}(y_1(0))$ and $\xi_2(0) = y_2(0)/\psi'(\xi_1(0))$. These relations simplify to (58) by using symmetric values $\xi_1^\pm$ and $c^\pm$ for $\psi(\xi_1)$, such that $\psi(\xi_1)$ passes through the origin, see Fig. 2 and Appendix I.

With the Hamiltonian

$$H(\xi, \lambda_\xi, \tilde{u}, t) = (\psi''\xi_2^2 + \psi'\tilde{u})^2 + \varepsilon\tilde{u}^2 + \lambda_{\xi,1}\xi_2 + \lambda_{\xi,2}\tilde{u},$$
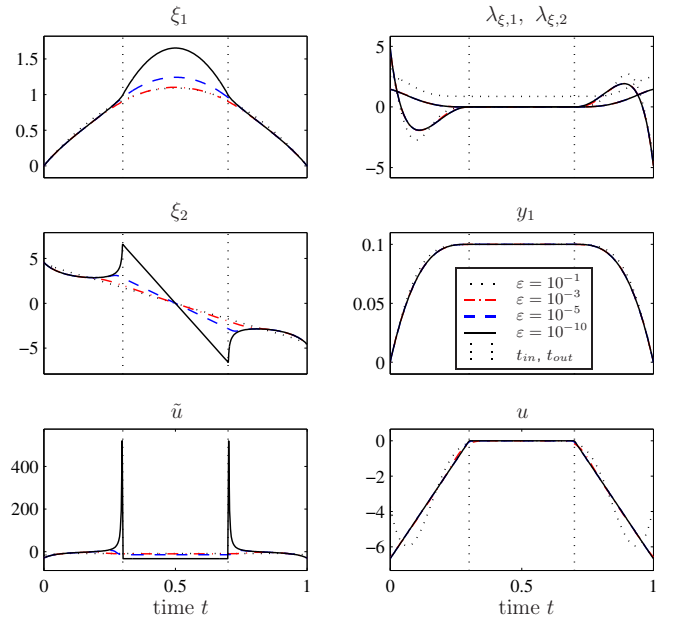
the optimality conditions (30)–(33) become

$$0 = (\psi''\xi_2^2 + \psi'\tilde{u})\psi' + 2\varepsilon\tilde{u} + \lambda_{\xi,2} \quad (59)$$

$$\begin{aligned} \dot{\lambda}_{\xi,1} &= -(\psi''\xi_2^2 + \psi'\tilde{u})(\psi'''\xi_2^2 + \psi''\tilde{u}), \quad \lambda_{\xi,1}(1) = \nu_1 \\ \dot{\lambda}_{\xi,2} &= -(\psi''\xi_2^2 + \psi'\tilde{u})2\psi''\xi_2 - \lambda_{\xi,1}, \quad \lambda_{\xi,2}(1) = \nu_2 \end{aligned} \quad (60)$$

with the adjoint states $\lambda_\xi = (\lambda_{\xi,1}, \lambda_{\xi,2})^\mathsf{T}$ and the multipliers $\nu = (\nu_1, \nu_2)^\mathsf{T}$.

It can be verified by differentiating (54), that the optimal solution $y_1^*(t)$ is twice continuously differentiable, i.e. $y_1^*(t) \in \mathcal{C}^2$. According to condition (44), the order of continuity of the saturation function $\psi(\xi_1) \in \mathcal{C}^\sigma$ is chosen as $\sigma = 2$ for the boundedness of the junction values (45). [3]

The BVP (57)–(60) with the additional algebraic equation (59) is solved in MATLAB with a modified version of the collocation–based BVP solver bvp4c, which is applicable to BVPs of (index 1) differential–algebraic equations, see, e.g., [19] for a short description. The initial guess for $\xi(t)$ is a linear interpolation between the boundary conditions (58) on a uniform time mesh $t_k \in [0, 1]$ with 30 points. For $\tilde{u}$ and λ_ξ, the initial guess is simply zero over $t_k \in [0, 1]$. Note that the boundary conditions for λ_ξ in (60) can be considered as being free and are therefore omitted (together with ν) in the numerical solution. The regularized BVP (57)–(60) is successively solved with the collocation solver and automatic mesh refinement for a decreasing sequence $\{\varepsilon^k\}$ from 10^0 to 10^{-10} in 11 steps.

Figure 3 shows the simulation results for the constraint $y_1^* \leq c^+ = 0.1$ and the saturation function $\psi(\xi_1) \in \mathcal{C}^2$ in (61) with $\xi_1^\pm = \pm 1$. For decreasing values of ε, $(\xi_1, \xi_2, \tilde{u})$ approach the bounded trajectories $(\xi_1^*, \xi_2^*, \tilde{u}^*)$, as defined

---

[3]Note that (44) is a sufficient condition for the existence of bounded junction values (45), since $y_1^*(t)$ in (54) is a finite time polynomial, see Remark 4 in Appendix II.

in (46)–(47). The original variables (y_1, y_2, u) in (55) accordingly approach the optimal trajectories following from (54).

Figure 4 presents the corresponding results for a three times differentiable saturation function $\psi(\xi_1) \in \mathcal{C}^3$. The peaks in ξ_2 and $\tilde{u}$ clearly indicate the unboundedness at the junction points t_{in}^- and t_{out}^+ for $\varepsilon \to 0$. Nevertheless, the original states (y_1, y_2) and input u approach the optimal solution in (54) for $\varepsilon \to 0$ (also see Remark 3).

In addition, Figure 5 shows the convergence of the cost $\tilde{J}(\tilde{u})$ to the optimal value $J^* = 4/(9c^+)$ for decreasing ε and the two saturation functions $\psi(\xi_1) \in \mathcal{C}^2$ and $\psi(\xi_1) \in \mathcal{C}^3$. For $\psi(\xi_1) \in \mathcal{C}^2$, $\tilde{J}(\tilde{u})$ converges faster to J^* due to the boundedness of $\tilde{u}$, which results in a smaller value of the integrated regularization term $\varepsilon \tilde{u}^2$ in (56).
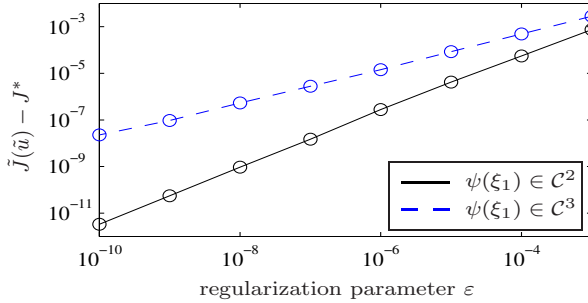

Fig. 5. Distance of $\tilde{J}(\tilde{u})$ (without regularization term) to optimal cost J^* for decreasing ε using the saturation functions $\psi(\xi_1) \in \mathcal{C}^2$ and $\psi(\xi_1) \in \mathcal{C}^3$.

VII. CONCLUSIONS

The scope of the saturation function approach is to systematically incorporate the state constraint of an optimal control problem (OCP) within a new system dynamics. This transformation results in an unconstrained OCP in new coordinates with an additional regularization term, which has to be successively reduced during the numerical solution. Important is the fact that the saturation function allows a tangential entry and exit of constrained arcs, although no interior boundary conditions are involved.

An interesting result for the new OCP is that it plays an important role for the existence of bounded trajectories if the saturation function has the same order of continuity as the optimal trajectory of the state constraint. This condition can be used to appropriately choose the continuity of the saturation function. If this matching condition is not satisfied, the trajectories in the new coordinates may become unbounded in the limit $\varepsilon \to 0$. However, this is not necessarily a drawback, since in practice the numerical solution with decreasing ε is stopped when the corresponding trajectories in the original coordinates are sufficiently close to the optimal ones.

Numerical studies have shown a remarkable numerical robustness of the approach due to the inherent unconstrained nature of the new OCP. For instance, the example in this paper can readily be solved for both considered saturation functions by decreasing ε from 10^0 to 10^{-10} in one or two steps.

The approach has been presented for the single–input case with one state constraint. However, the results can be extended to the multiple input case and multiple state

constraints with well–defined vector relative degree. Then, the constraints can be used as linearizing outputs to derive the input–output normal form as the basis for the saturation function approach.

APPENDIX I – CHOICE OF SATURATION FUNCTION

The saturation function $\psi(\xi_1)$ shown in Figure 2 can be constructed in three parts

$$\psi(\xi_1) = \begin{cases} \psi_0(\xi_1) & \text{if } \xi_1 \in (\xi_1^-, \xi_1^+) \\ c^{\pm} & \text{if } \xi_1 \gtrless \xi_1^{\pm} \end{cases} \tag{61a}$$

The function $\psi_0(\xi_1)$ has to connect the limits $c^{\pm}$ sufficiently smoothly to ensure the continuity of $\psi(\xi_1) \in \mathcal{C}^\sigma$. This leads to the $2(1 + \sigma)$ boundary conditions

$$\psi_0(\xi_1^{\pm}) = c^{\pm}, \quad \left.\frac{\mathrm{d}^i \psi_0}{\mathrm{d}\xi_1^i}\right|_{\xi_1 = \xi_1^{\pm}} = 0, \quad i = 1, \ldots, \sigma.$$

For instance, $\psi_0(\xi_1)$ can be constructed with the polynomial

$$\psi_0(\xi_1) = c^+ - (c^+ - c^-) \sum_{i=\sigma+1}^{2\sigma+1} b_i \left(\frac{\xi_1^+ - \xi_1}{\xi_1^+ - \xi_1^-}\right)^i \tag{61b}$$

for $\xi_1 \in [\xi_1^-, \xi_1^+]$, where the coefficients are determined according to [20], [21]

$$b_i = \frac{(-1)^{i-\sigma-1}(2\sigma + 1)!}{i \cdot \sigma!(i - \sigma - 1)!(2\sigma + 1 - i)!}. \tag{61c}$$

APPENDIX II – PROOF OF LEMMA 2

To prove Lemma 2, we restrict our considerations to the entry point t_{in}^- into the (upper) constrained arc, as shown in Fig. 1. In the first step, let us assume condition (44) holds, and consider series expressions for $y_1^*(t)$ and $\xi_1^*(t)$ to the left of t_{in} to match coefficients. Based on these results, we show by contradiction that for $\sigma > \rho$ or $\sigma < \rho$, there exists no power series formulation for $\xi_1^*(t)$ which corresponds to $y_1^*(t)$.

A series expression of $y_1^*(t)$ around $y_1^*(t_{in}) = c^+$ can be written as

$$y_1^*(t_{in} - \tau) = c^+ - \sum_{i=\rho+1}^{\infty} a_i \tau^i \tag{62}$$

with $a_i = 0$, $i = 0, 1, \ldots, \rho$ and $a_{\rho+1} \neq 0$ due to the tangential entry in $y_1^*(t_{in}) = c^+$ and the continuity of $y_1^*(t) \in \mathcal{C}^\rho$ with $\rho \geq r - 1$, see (15) and Remark 1.

For $\tau \geq 0$ (i.e. $t \leq t_{in}$), the saturation function (61a) is determined by the polynomial $\psi_0(\xi_1)$ in (61b). According to (44), we set $\sigma = \rho$, and write $\psi(\xi_1) = \psi_0(\xi_1)$ as

$$\psi(\xi_1^+ - \Delta\xi_1) = c^+ - \sum_{i=\rho+1}^{2\rho+1} \bar{b}_i \Delta\xi_1^i \tag{63}$$

with $\bar{b}_i = b_i(c^+ - c^-)/(\xi_1^+ - \xi_i^-)^i$. Moreover, we consider a power series of $\xi_1^*(t_{in} - \tau) = \xi_1^+ - \Delta\xi_1(\tau)$ with

$$\Delta\xi_1(\tau) = \sum_{i=1}^{\infty} c_i \tau^i. \tag{64}$$

In order to determine the coefficients c_i, the series (64) is substituted in (63):

$$\psi(\xi_1^+ - \Delta\xi(\tau))$$

$$= c^+ - \bar{b}_{\rho+1}\left[c_1\tau + \mathcal{O}(\tau^2)\right]^{\rho+1} - \sum_{\rho+2}^{2\rho+1} \bar{b}_i \Delta\xi_1^i(\tau) \quad (65a)$$

$$= c^+ - \sum_{i=\rho+1}^{\infty} d_i \tau^i. \quad (65b)$$

Comparing the first coefficient $d_1 = \bar{b}_{\rho+1}c_1^{\rho+1}$ with $a_{\rho+1}$ in (62) yields c_1 of the series (64):

$$c_1 = \sqrt[\rho+1]{a_{\rho+1}/\bar{b}_{\rho+1}}. \quad (66)$$

The calculation of the further coefficients c_i, $i = 2, \ldots, r$ requires a closer look at the structure behind the coefficients d_i in (65b). In order to determine the lowest orders of τ where the single c_i appears, the terms $c_1\tau$ and $c_i\tau^i$ in (65a) are rearranged

$$\psi(\xi_1^+ - \Delta\xi(\tau)) = c^+ - \quad (67)$$

$$\bar{b}_{\rho+1}\Big[c_1\tau + c_i\tau^i + \sum_{j=2}^{i-1} c_j\tau^j + \mathcal{O}(\tau^{i+1})\Big]^{\rho+1} + \sum_{j=\rho+2}^{2\rho+1} \bar{b}_j \Delta\xi_1^j(\tau)$$

and $(c_1\tau + c_i\tau^i)^{\rho+1}$ is expanded using the binomial theorem:

$$\big(c_1\tau + c_i\tau^i\big)^{\rho+1} = c_1^{\rho+1}\tau^{\rho+1} + (\rho+1)c_1^\rho c_i\tau^{\rho+i} + \mathcal{O}(\tau^{\rho+2i-1}).$$

The second term involving c_i and $\tau^{\rho+i}$ is the first one in the series (65a) where the coefficient c_i appears. However, the full evaluation of the series (67) produces further terms in $\tau^{\rho+i}$ depending on the previous coefficients $c_1, \ldots, c_{i-1}$. Hence, the coefficients d_i in (65b) exhibit the structure

$$d_{\rho+i} = d_{\rho+i}^0(\bar{b}, c_1, \ldots, c_{i-1}) + (\rho+1)\bar{b}_{\rho+1}c_1^\rho c_i \quad (68)$$

with $\bar{b} = (\bar{b}_{\rho+1}, \ldots, \bar{b}_{2\rho+1})^\mathsf{T}$. Comparing this expression with the corresponding term $a_{\rho+i}$ in the Taylor series (62) of $y_1^*(t)$ finally leads to the coefficient

$$c_i = \frac{a_{\rho+i} - d_{\rho+i}^0(\bar{b}, c_1, \ldots, c_{i-1})}{(\rho+1)\bar{b}_{\rho+1}c_1^\rho}, \quad i \geq 2. \quad (69)$$

Note that c_i is bounded due to $\bar{b}_{\rho+1}, c_1 \neq 0$, cf. (61c) and (66). Finally, the series (64) can be used to recover the original state $\xi_1^*(t_{in} - \tau) = \xi_1^+ - \Delta\xi_1(\tau)$ and the derivatives $\frac{d^i}{dt^i}\xi_1^*(t) = (-1)^{i+1}\frac{d^i}{d\tau^i}\Delta\xi_1(\tau)$, $i = 1, \ldots, r$. This yields the bounded junction values at $t = t_{in}^-$:

$$\xi_1(t_{in}^-) = \xi_1^+, \quad \xi_i(t_{in}^-) = (-1)^i (i-1)! \, c_{i-1}, \quad i = 2, \ldots, r$$

$$\tilde{u}(t_{in}^-) = (-1)^{r+1} r! \, c_r.$$

Now consider the case that condition (44) were not satisfied and assume e.g. $\sigma = \rho + 1$. Then, the series in (65b) starts with $d_{\rho+2}\tau^{\rho+2}$ and $d_{\rho+1}$ evaluates to zero. This however contradicts the first non–zero term $a_{\rho+1}\tau^{\rho+1}$ with $a_{\rho+1} \neq 0$ of the series (62).

Next, consider the case $\sigma = \rho - 1$, where the series (65b) already starts with $d_\rho\tau^\rho$ and $d_\rho = \bar{b}_\rho c_1^\rho$. Comparison with the corresponding zero term in (62), $0 = a_\rho = \bar{b}_\rho c_1^\rho$, leads to $c_1 = 0$. However, the further coefficients in the series (65b)

have the structure (68) which involves the next c_i multiplied by $c_1 = 0$. As a consequence, c_i cannot be determined from the comparison of the series (65b) with (62).

The conclusion is that in both cases $\sigma < \rho$ and $\sigma > \rho$, there exists no power series of the form (64) which equals the power series (62) of the optimal trajectory $y_1^*(t) = \psi(\xi^*(t)) = \psi(\xi_1^+ - \Delta\xi_1(\tau))$. This proves Lemma 2.

Remark 4: *If the Taylor series (62) for $y_1^*(t)$ has only a finite number of terms and thus reduces to a time polynomial, the series (64) is also finite and represents an exact solution of $\xi_1^*(t)$ over $t \in [0, t_{in}]$. Then, $\sigma = \rho$ is a sufficient condition for the boundedness of the junction values $\xi^*(t_{in}^-)$. This is e.g. the case for the example in Section VI, see Eq. (54).*

References

[1] C. Hargraves and S. Paris, "Direct trajectory optimization using nonlinear programming and collocation", *J. Guid. Contr. Dyn.*, vol. 10, pp. 338–342, 1987.

[2] H. Seywald, "Trajectory optimization based on differential inclusion", *J. Guid. Contr. Dyn.*, vol. 17, no. 3, pp. 480–487, 1994.

[3] S.K. Agrawal and N. Faiz, "Optimization of nonlinear dynamic systems: new efficient method without Lagrange multipliers", *J. Optim. Appl.*, vol. 97, pp. 11–28, 1998.

[4] J. Betts, "Survey of numerical methods for trajectory optimization", *J. Guid. Contr. Dyn.*, vol. 21, pp. 193–207, 1998.

[5] J. Nocedal and S. Wright, *Numerical Optimization*. Springer, New York, 1999.

[6] J. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*. SIAM, Philadelphia, PA, 2001.

[7] R. Bhattacharya, "OPTRAGEN: a Matlab toolbox for optimal trajectory generation", *Proc. 45th IEEE Conf. Dec. Contr. (CDC)*, pp. 6832–6836, San Diego, CA, 2006.

[8] A. Bryson and Y.-C. Ho, *Applied Optimal Control*. Ginn & Company, Waltham, Massachusetts, 1969.

[9] H. Pesch, "A practical guide to the solution of real–life optimal control problems", *Contr. Cybern.*, vol. 23, pp. 7–60, 1994.

[10] H. Kreim, B. Kugelmann, H. Pesch, and M. Breitner, "Minimizing the maximum heating of a reentry space shuttle: an optimal control problem with multiple control constraints", *Opt. Contr. Appl. Meth.*, vol. 17, pp. 45–69, 1996.

[11] A. Bryson, *Dynamic Optimization*. Addison–Wesley, Menlo Park, CA, 1999.

[12] B. Bonnard, L. Faubourg, G. Launay, and E. Trélat, "Optimal control with state constraints and the space shuttle re–rentry problem", *J. Dyn. Contr. Syst.*, vol. 9, pp. 155–199, 2003.

[13] L. Pontryagin, V. Boltyansky, V. Gamkrelidze, and E. Mischenko, *Mathematical Theory of Optimal Processes*. Wiley–Interscience, New York, 1962.

[14] A. Isidori, *Nonlinear Control Systems*, 3rd ed. Springer, 1995.

[15] H. Nijmeijer and A. van der Schaft, *Nonlinear Dynamical Control Systems*. Springer, 1996.

[16] D.H. Jacobson, M.M. Lele, and J.L. Speyer, "New necessary conditions of optimality for control problems with state–variable inequality constraints", *J. Math. Anal. Appl.*, vol. 35, pp. 255–284, 1971.

[17] J.F. Bonnans and A. Hermant, "No–gap second-order optimality conditions for optimal control problems with a single state constraint and control", *Math. Prog. Series B* (accepted).

[18] K. Graichen, N. Petit, "Incorporating constraints into the dynamics of optimal control problems" (submitted), 2008.

[19] K. Graichen, N. Petit, "Constructive methods for initialization and handling mixed state–input constraints in optimal control", *J. Guid. Contr. Dyn.*, scheduled for vol. 31(5), 2008.

[20] A. Piazzi and A. Visioli, "Optimal noncausal set–point regulation of scalar systems", *Automatica*, vol. 37, pp. 121–127, 2001.

[21] K. Graichen, V. Hagenmeyer, and M. Zeitz, "A new approach to inversion–based feedforward control design for nonlinear systems", *Automatica*, vol. 41, pp. 2033–2041, 2005.