Pure and Applied Functional Analysis
Volume 4, Number 2, 2019, 363–387
Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

NECESSARY AND SUFFICIENT CONDITIONS FOR OPTIMAL CONTROL PROBLEMS SUBJECT TO HESSENBERG DIFFERENTIAL ALGEBRAIC EQUATIONS OF ARBITRARY INDEX AND MIXED CONTROL-STATE CONSTRAINTS

BJOERN MARTENS AND MATTHIAS GERDTS

Abstract. The paper studies optimal control problems subject to mixed control-state constraints and differential-algebraic equations (DAEs) with Hessenberg structure of arbitrary index. We derive necessary optimality conditions in terms of a local minimum principle and establish second order sufficient conditions. The latter involves a coercivity condition and a Riccati equation that needs to have a bounded solution.

1. Introduction

The paper aims to establish necessary and sufficient conditions for a class of nonlinear optimal control problems subject to mixed control-state constraints and differential-algebraic equations (DAEs) of arbitrary index. The DAE is supposed to have Hessenberg structure, which frequently occurs in mechanical engineering and path following problems. While the investigation of necessary and sufficient conditions is well-established for optimal control problems subject to ordinary differential equations (ODEs), only a few results are available for DAEs, especially for higher index DAEs and for nonlinear problems. Up to the knowledge of the authors, a systematic treatment of Hessenberg DAEs of arbitrary index in terms of necessary and especially sufficient conditions has not been performed as yet. While the authors are not aware of sufficient conditions for optimal control problems subject to nonlinear higher index DAEs, there exist some results regarding sufficiency for linear-quadratic DAE optimal control problems, compare [23, 24, 12, 14, 1].

Necessary conditions have been studied for semi-explicit index-1 DAEs with set constraints on the controls in [25] and with pure state and mixed control-state constraints in [4], for semi-explicit index-2 DAEs in [7], for Hessenberg-type DAEs up to index 3 DAEs in [29], and for arbitrarily structured DAEs in [13]. An overview on necessary conditions with further references can be found in the survey paper [9].

In contrast to optimal control problems subject to DAEs there is a comparatively rich literature for ODEs. First and second order sufficient optimality conditions for

2010 *Mathematics Subject Classification.* 49J15 and 49K15 and 65L80.

Key words and phrases. Optimal control and Differential algebraic equation and Hessenberg form and necessary conditions and second order sufficient conditions and Riccati equation and Legendre Clebsch.

infinite dimensional programming problems are discussed in [21]. The results are applied to optimal control problems subject to ODEs with pure state constraints. A coercivity condition for the second derivative of the Lagrangian is derived, taking the two-norm discrepancy into account. Furthermore they present sufficient conditions of Riccati-type. [20] analyze first and second-order necessary as well as sufficient optimality conditions for infinite-dimensional programming problems and utilize the results to derive second-order sufficient conditions for optimal control problems subject to ODEs with mixed control-state and pure state constraints using the Hamilton-Jacobi inequality. They also introduce a Legendre-Clebsch and Riccati-type condition. In [31] first and second-order sufficient conditions for optimal control problems with control constraints are obtained. Second-order sufficient conditions for cone-constrained optimization problems are acquired by [16]. They apply the results to optimal control problems for affine systems subject to state-space constraints. [5] make use of three different norms to obtain second order sufficient conditions for infinite dimensional optimization problems, which are utilized to derive sufficient conditions for optimal control problems with endpoint constraints, and equality and inequality constraints on the controls. In [22] second order sufficient conditions for optimal control problems subject to state and control constraints are discussed. Second order sufficient conditions for optimal control problems subject to mixed control-state constraints as well as pure state constraints of order one are analyzed by [19]. [2] investigate second-order conditions for optimal control problems with pure state constraints of arbitrary order and mixed control-state constraints.

In this paper we aim to extend some of the results for ODEs to DAE optimal control problems. In particular we discuss necessary and second-order sufficient conditions for optimal control problems subject to DAEs in Hessenberg form of index $k \in \mathbb{N}$. Herein, the derivation of sufficient conditions is a new contribution up to the knowledge of the authors, while the statement of necessary conditions builds on existing results and formalizes it for problems with arbitrary index. We consider the following optimal control problem **(OCP)** on a fixed and compact time interval $[t_0, t_f]$:

Minimize

$$\varphi\left(x_1(t_f), x_2(t_f), \ldots, x_{k-1}(t_f)\right)$$

with respect to

$$x_1 \in W_{1,\infty}^{n_{x_1}}\left([t_0,t_f]\right), x_2 \in W_{2,\infty}^{n_{x_2}}\left([t_0,t_f]\right), \ldots, x_{k-1} \in W_{k-1,\infty}^{n_{x_{k-1}}}\left([t_0,t_f]\right),$$
$$y \in L_\infty^{n_y}\left([t_0,t_f]\right), u \in L_\infty^{n_u}\left([t_0,t_f]\right)$$

subject to the Hessenberg DAE

$$\dot{x}_1(t) = f_1\left(x_1(t), x_2(t), \ldots, x_{k-1}(t), y(t), u(t)\right), \tag{1.1}$$
$$\dot{x}_2(t) = f_2\left(x_1(t), x_2(t), \ldots, x_{k-1}(t)\right), \tag{1.2}$$
$$\dot{x}_3(t) = f_3\left(x_2(t), x_3(t), \ldots, x_{k-1}(t)\right), \tag{1.3}$$
$$\vdots$$

$$\dot{x}_{k-1}(t) = f_{k-1}\left(x_{k-2}(t), x_{k-1}(t)\right), \tag{1.4}$$

$$0 = g\left(x_{k-1}(t)\right), \tag{1.5}$$

the initial conditions

$$\begin{aligned} 0 &= D_1\left(x_1(t_0) - x_1^0\right), \\ 0 &= D_2\left(x_2(t_0) - x_2^0\right), \\ &\vdots \\ 0 &= D_{k-1}\left(x_{k-1}(t_0) - x_{k-1}^0\right), \end{aligned}$$

and the mixed control-state constraints

$$c\left(x_1(t), x_2(t), \ldots, x_{k-1}(t), y(t), u(t)\right) \leq 0.$$

Herein, $n_{\boldsymbol{x}} := \sum_{i=1}^{k-1} n_{x_i}$ is the dimension of the differential state vector function $\boldsymbol{x} = (x_1, x_2, \ldots, x_{k-1})^\top$, y is the algebraic state vector function, and u is the control vector function. The DAE consists of the differential equations (1.1)-(1.4) and the algebraic constraint (1.5), which can be interpreted as a pure state constraint that is active for every $t \in [t_0, t_f]$. The particular structure of the DAE (1.1)-(1.5) is called Hessenberg structure. Please note the smoothness properties of the differential state components $x_j \in W_{j,\infty}^{n_{x_j}}([t_0, t_f]$, $j = 1, \ldots, k-1$, where the Banach spaces $W_{j,\infty}^{n_{x_j}}([t_0, t_f]$ will be defined in Section 2.

The functions

$$\begin{aligned} \varphi &: \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}, \\ f_1 &: \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_{x_1}}, \\ f_2 &: \mathbb{R}^{n_{\boldsymbol{x}}} \to \mathbb{R}^{n_{x_2}}, \\ f_3 &: \mathbb{R}^{n_{\boldsymbol{x}} - n_{x_1}} \to \mathbb{R}^{n_{x_3}}, \\ f_4 &: \mathbb{R}^{n_{\boldsymbol{x}} - n_{x_1} - n_{x_2}} \to \mathbb{R}^{n_{x_4}}, \\ &\vdots \\ f_{k-1} &: \mathbb{R}^{n_{\boldsymbol{x}} - \sum_{i=1}^{k-3} n_{x_i}} \to \mathbb{R}^{n_{x_{k-1}}}, \\ g &: \mathbb{R}^{n_{x_{k-1}}} \to \mathbb{R}^{n_y}, \\ c &: \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}, \end{aligned}$$

and the matrices

$$D_i \in \mathbb{R}^{\left(n_{x_i} - n_y\right) \times n_{x_i}}, \quad i = 1, \ldots, k-1,$$

are supposed to be given. It is well known for DAEs of type (1.1)-(1.5) that not only the algebraic constraint in (1.5) has to be satisfied, but also its derivatives with respect to time up to order $(k-1)$ impose so-called hidden constraints. These hidden constraints are obtained by repeated differentiation of (1.5) and substitution of (1.1)-(1.4). To this end we introduce the following notation for the hidden constraints, which read as follows:

$$
\begin{aligned}
0 &= g\left(x_{k-1}(t)\right) \\
&=: g_{k-1}\left(x_{k-1}(t)\right) \\
0 &= \frac{d}{dt}\, g_{k-1}\left(x_{k-1}(t)\right) = g'_{k-1,x_{k-1}}\left(x_{k-1}(t)\right)\, f_{k-1}\left(x_{k-2}(t), x_{k-1}(t)\right) \\
&=: g_{k-2}\left(x_{k-2}(t), x_{k-1}(t)\right) \\
0 &= \frac{d}{dt}\, g_{k-2}\left(x_{k-2}(t), x_{k-1}(t)\right) \\
&= g'_{k-2,x_{k-2}}\left(x_{k-2}(t), x_{k-1}(t)\right) f_{k-2}\left(x_{k-3}(t), x_{k-2}(t), x_{k-1}(t)\right) \\
&\qquad + g'_{k-2,x_{k-1}}\left(x_{k-2}(t), x_{k-1}(t)\right) f_{k-1}\left(x_{k-2}(t), x_{k-1}(t)\right) \\
&=: g_{k-3}\left(x_{k-3}(t), x_{k-2}(t), x_{k-1}(t)\right) \\
&\vdots \\
0 &= \frac{d}{dt}\, g_{1}\left(x_{1}(t), x_{2}(t), \ldots, x_{k-1}(t)\right) \\
&= \sum_{j=2}^{k-1} g'_{1,x_j}\left(x_{1}(t), x_{2}(t), \ldots, x_{k-1}(t)\right) f_j(x_{j-1}(t), \ldots, x_{k-1}(t)) \\
&\qquad + g'_{1,x_1}\left(x_{1}(t), x_{2}(t), \ldots, x_{k-1}(t)\right) f_1(x_1(t), \ldots, x_{k-1}(t), y(t), u(t)) \\
&=: g_0\left(x_{1}(t), x_{2}(t), \ldots, x_{k-1}(t), y(t), u(t)\right).
\end{aligned}
$$

This paper is organized as follows: In Section 2 we introduce important notations and assumptions. Necessary conditions for **(OCP)** are derived in Section 3. Section 4 deals with sufficient conditions for **(OCP)** and an illustrative example is presented in Section 5.

2. Preliminaries

We denote the Banach-space, which consists of all essentially bounded functions $v : [t_0, t_f] \to \mathbb{R}^n$ by $\left(L^n_\infty\left([t_0, t_f]\right), \|\cdot\|_\infty\right)$ and the Banach-space, which consists of functions $v : [t_0, t_f] \to \mathbb{R}^n$ with bounded derivatives up to order $m \in \mathbb{N}$ in $L^n_p\left([t_0, t_f]\right)$ by $\left(W^n_{m,p}\left([t_0, t_f]\right), \|\cdot\|_{m,p}\right)$, with
$\|v\|_{m,p} := \max_{i=1,\ldots,m}\left\{\|v^{(i)}\|_p\right\}$, where $v^{(i)}$ is the i-th derivative of v and $p = 2, \infty$. For a Banach-space $\left(X, \|\cdot\|_X\right)$ we denote the dual-space by $X^\star$.
Furthermore we make use of the following abbreviations:

$$
\boldsymbol{x}^{(i)}(\cdot) := \begin{pmatrix} x_i(\cdot) \\ x_{i+1}(\cdot) \\ \vdots \\ x_{k-1}(\cdot) \end{pmatrix}, \qquad i = 1, \ldots, k-1, \quad \boldsymbol{x}^0 := \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_{k-1}^0 \end{pmatrix},
$$

$$
\boldsymbol{x}(\cdot) := \boldsymbol{x}^{(1)}(\cdot), \quad \boldsymbol{z}(\cdot) := \begin{pmatrix} \boldsymbol{x}(\cdot) \\ y(\cdot) \\ u(\cdot) \end{pmatrix},
$$

$$
\begin{aligned}
X &:= W^{n_{x_1}}_{1,\infty}\left([t_0,t_f]\right) \times W^{n_{x_2}}_{2,\infty}\left([t_0,t_f]\right) \times \ldots \times W^{n_{x_{k-1}}}_{k-1,\infty}\left([t_0,t_f]\right),\\
\hat{X} &:= W^{n_{x_1}}_{1,2}\left([t_0,t_f]\right) \times W^{n_{x_2}}_{2,2}\left([t_0,t_f]\right) \times \ldots \times W^{n_{x_{k-1}}}_{k-1,2}\left([t_0,t_f]\right),\\
Y &:= L^{n_{x_1}}_{\infty}\left([t_0,t_f]\right) \times W^{n_{x_2}}_{1,\infty}\left([t_0,t_f]\right) \times \ldots \times W^{n_{x_{k-1}}}_{k-2,\infty}\left([t_0,t_f]\right),\\
D &:= \begin{pmatrix} D_1 & 0 & \cdots & 0\\ 0 & D_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & D_{k-1}\end{pmatrix},
\end{aligned}
$$

$$
\boldsymbol{f}\left(\boldsymbol{z}(\cdot)\right) := \begin{pmatrix} f_1\left(\boldsymbol{z}(\cdot)\right)\\ f_2\left(\boldsymbol{x}^{(1)}(\cdot)\right)\\ f_3\left(\boldsymbol{x}^{(2)}(\cdot)\right)\\ \vdots\\ f_{k-1}\left(\boldsymbol{x}^{(k-2)}(\cdot)\right)\end{pmatrix}, \quad \boldsymbol{g}\left(\boldsymbol{x}(\cdot)\right) := \begin{pmatrix} g_1\left(\boldsymbol{x}^{(1)}(\cdot)\right)\\ g_2\left(\boldsymbol{x}^{(2)}(\cdot)\right)\\ g_3\left(\boldsymbol{x}^{(3)}(\cdot)\right)\\ \vdots\\ g_{k-1}\left(\boldsymbol{x}^{(k-1)}(\cdot)\right)\end{pmatrix}.
$$

and equip the space X and $\hat{X}$ with the norms

$$
\begin{aligned}
\|\boldsymbol{x}\|_X &:= \max_{i=1,\ldots,k-1}\left\{\|x_i\|_{i,\infty}\right\},\\
\|\boldsymbol{x}\|_{\hat{X}} &:= \max_{i=1,\ldots,k-1}\left\{\|x_i\|_{i,2}\right\}.
\end{aligned}
$$

Throughout this paper we assume the following:

(A1) The optimal control problem **(OCP)** has a weak local minimizer

$$
(\hat{\boldsymbol{x}}, \hat{y}, \hat{u}) \in X \times L^{n_y}_{\infty}\left([t_0,t_f]\right) \times L^{n_u}_{\infty}\left([t_0,t_f]\right).
$$

(A2) The initial value is consistent, i.e., the matrix

$$
E := \begin{pmatrix} \boldsymbol{g}'_{\boldsymbol{x}}\left(\boldsymbol{x}(t_0)\right)\\ D\end{pmatrix}
$$

has full rank.

(A3) The DAE has index k, i.e., the Jacobian $g'_{0,y}(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t))$ is non-singular almost everywhere on $[t_0, t_f]$ and the inverse is essentially bounded.

(A4) The functions f_i are $i+1$ times Fréchet differentiable and the respective derivatives are continuous in all arguments. c is twice Fréchet differentiable and the respective derivatives are continuous in all arguments and g is $k+1$ times Fréchet differentiable and the respective derivatives are continuous in all arguments.

We denote partial derivatives with a subscript and derivatives at an optimal point as functions of time t with squared brackets, e.g.,
$f'_{1,x_1}[t] := \frac{\partial f_1(\hat{\boldsymbol{x}}(t),\hat{y}(t),\hat{u}(t))}{\partial x_1}$. Furthermore we denote the pseudo-inverse of a matrix $A \in \mathbb{R}^{n\times m}$ and $rank\,(A) = n$ by $A^{+} := A^{\top}\left(A\,A^{\top}\right)^{-1}$.

3. Necessary conditions

To derive a local minimum principle for **(OCP)** we consider the optimal control problem as an infinite optimization problem and apply first order necessary conditions of Fritz John type. **(OCP)** is equivalent with the infinite optimization problem

$$\textbf{(P)}: \qquad \text{Minimize} \quad J(\boldsymbol{z}) \quad \text{w.r.t.} \quad \boldsymbol{z} \in Z \quad \text{s.t.} \quad H(\boldsymbol{z}) = 0,\ G(\boldsymbol{z}) \in K,$$

where

$$\begin{aligned}
Z &:= X \times L_\infty^{n_y}([t_0,t_f]) \times L_\infty^{n_u}([t_0,t_f]),\\
\hat{Z} &:= \hat{X} \times L_2^{n_y}([t_0,t_f]) \times L_2^{n_u}([t_0,t_f]),\\
V &:= Y \times W_{k-1,\infty}^{n_y}([t_0,t_f]) \times \mathbb{R}^{n_{\boldsymbol{x}}-(k-1)\,n_y},\\
W &:= L_\infty^{n_c}([t_0,t_f]),\\
J &: Z \to \mathbb{R}, \quad H: Z \to V, \quad G: Z \to W\\
J(\boldsymbol{z}) &:= \varphi\left(\boldsymbol{x}(t_f)\right),\\
H(\boldsymbol{z}) &:= \begin{pmatrix} \boldsymbol{f}\left(\boldsymbol{z}(\cdot)\right) - \dot{\boldsymbol{x}}(\cdot) \\ g_{k-1}\left(\boldsymbol{x}^{(k-1)}(\cdot)\right) \\ -D\left(\boldsymbol{x}(t_0) - \boldsymbol{x}^0\right) \end{pmatrix},\\
G(\boldsymbol{z}) &:= -c(\boldsymbol{z}(\cdot)),\\
K &:= \{\vartheta \in W \mid \vartheta(t) \geq 0 \text{ a.e. in } [t_0,t_f]\}.
\end{aligned}$$

and equip the spaces Z and $\hat{Z}$ with the norms

$$\begin{aligned}
\|\boldsymbol{z}\|_Z &:= \max\left\{\|\boldsymbol{x}\|_X, \|y\|_\infty \|u\|_\infty\right\},\\
\|\boldsymbol{z}\|_{\hat{Z}} &:= \max\left\{\|\boldsymbol{x}\|_{\hat{X}}, \|y\|_2 \|u\|_2\right\}.
\end{aligned}$$

Note that the mappings J, H, and G are Fréchet differentiable, if Assumption **(A4)** holds. Moreover, if assumptions **(A1)** - **(A4)** hold, then $H'(\hat{\boldsymbol{z}}) : Z \to V$ is a surjective operator, which can be proven similarly to [8, Lemma 3.1.4] and by using similar arguments as in Equations (3.5) - (3.14) below.

Under these assumptions, the first order necessary Fritz John conditions hold, compare [8, Theorem 2.3.24], and yield the existence of non-trivial multipliers

$$\ell_0 \geq 0, \quad \boldsymbol{\lambda}^\star \in V^\star, \quad \eta^\star \in W^\star$$

such that

$$\begin{aligned}
&0 = \ell_0\, J'(\hat{\boldsymbol{z}})(\boldsymbol{z}) - \boldsymbol{\lambda}^\star\left(H'(\hat{\boldsymbol{z}})(\boldsymbol{z})\right) - \eta^\star\left(G'(\hat{\boldsymbol{z}})(\boldsymbol{z})\right) \qquad (3.1)\\
&0 = \eta^\star\left(G(\hat{\boldsymbol{z}})\right), \quad \eta^\star \in K^+
\end{aligned}$$

holds for all $\boldsymbol{z} \in Z$, where K^+ is the positive dual cone of K. We define

$$\boldsymbol{\lambda_f}^\star := \left(\lambda_{f_1}^\star, \lambda_{f_2}^\star, \ldots, \lambda_{f_{k-1}}^\star\right), \quad \boldsymbol{\lambda}^\star := \left(\boldsymbol{\lambda_f}^\star, \lambda_g^\star, \sigma^\star\right).$$

For every $\boldsymbol{x} \in X$, every $y \in L_\infty^{n_y}([t_0,t_f])$, and every $u \in L_\infty^{n_u}([t_0,t_f])$ the variational equation (3.1) yields

$$0 = \boldsymbol{\lambda_f}^\star\left(\dot{\boldsymbol{x}}(\cdot) - \boldsymbol{f}'_{\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right) - \lambda_g^\star\left(g'_{k-1,\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right) + \eta^\star\left(c'_{\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right), \qquad (3.2)$$

$$+ \sigma^\top D\, \boldsymbol{x}(t_0) + \ell_0\, \varphi'\left(\hat{\boldsymbol{x}}(t_f)\right) \boldsymbol{x}(t_f)$$

$$0 = -\boldsymbol{\lambda_f}^\star \left(\boldsymbol{f}'_y[\cdot]\, y(\cdot)\right) + \eta^\star \left(c'_y[\cdot]\, y(\cdot)\right), \tag{3.3}$$

$$0 = -\boldsymbol{\lambda_f}^\star \left(\boldsymbol{f}'_u[\cdot]\, u(\cdot)\right) + \eta^\star \left(c'_u[\cdot]\, u(\cdot)\right). \tag{3.4}$$

We intend to derive explicit representations of the multipliers $\boldsymbol{\lambda_f}^\star, \lambda_g^\star$ and $\eta^\star$. For arbitrary $\boldsymbol{h} \in Y$, $h_k \in W^{n_y}_{k-1,\infty}\left([t_0, t_f]\right)$ and $h_{k+1} \in L^{n_c}_\infty\left([t_0, t_f]\right)$ we consider the linear system

$$\dot{\boldsymbol{x}}(t) = A_f(t)\, \boldsymbol{x}(t) + B_f(t)\, y(t) + C_f(t)\, u(t) + \boldsymbol{h}(t), \tag{3.5}$$

$$0 = g'_{k-1,\boldsymbol{x}}[t]\, \boldsymbol{x}(t) + h_k(t), \tag{3.6}$$

$$h_{k+1}(t) = A_c(t)\, \boldsymbol{x}(t) + B_c(t)\, y(t) + C_c(t)\, u(t), \tag{3.7}$$

$$0 = D\, \boldsymbol{x}(t_0), \tag{3.8}$$

where

$$A_f(t) := \boldsymbol{f}'_{\boldsymbol{x}}[t], \quad B_f(t) := \boldsymbol{f}'_y[t], \quad C_f(t) := \boldsymbol{f}'_u[t],$$
$$A_c(t) := c'_{\boldsymbol{x}}[t], \quad B_c(t) := c'_y[t], \quad C_c(t) := c'_u[t],$$
$$\boldsymbol{h}(t) := \begin{pmatrix} h_1(t) \\ h_2(t) \\ \vdots \\ h_{k-1}(t) \end{pmatrix}.$$

By differentiating the algebraic equation (3.6) $k-1$ times we obtain the linear system

$$\dot{\boldsymbol{x}}(t) = A_f(t)\, \boldsymbol{x}(t) + B_f(t)\, y(t) + C_f(t)\, u(t) + \boldsymbol{h}(t) \tag{3.9}$$

$$0 = A_g(t)\, \boldsymbol{x}(t) + B_g(t)\, y(t) + C_g(t)\, u(t) + q\left(\boldsymbol{h}(t)\right) + \frac{d^{k-1}}{dt^{k-1}}\, h_k(t) \tag{3.10}$$

$$h_{k+1}(t) = A_c(t)\, \boldsymbol{x}(t) + B_c(t)\, y(t) + C_c(t)\, u(t) \tag{3.11}$$

$$0 = E\, \boldsymbol{x}(t_0) + \boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) + \boldsymbol{p}_0\left(h_k(t_0)\right) \tag{3.12}$$

where

$$\begin{pmatrix} q_1\left(\boldsymbol{h}(t)\right) \\ q_2\left(\boldsymbol{h}(t)\right) \\ \vdots \\ q_{k-1}\left(\boldsymbol{h}(t)\right) \end{pmatrix} := \boldsymbol{g}'_{\boldsymbol{x}}\left(\hat{\boldsymbol{x}}(t)\right) \boldsymbol{h}(t), \quad q\left(\boldsymbol{h}(t)\right) := \sum_{i=1}^{k-1} \frac{d^{i-1}}{dt^{i-1}}\, q_i\left(\boldsymbol{h}(t)\right),$$

$$\boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) := \begin{pmatrix} \sum_{i=1}^{k-2} \left(\frac{d^{i-1}}{dt^{i-1}}\, q_{i+1}\left(\boldsymbol{h}(t)\right)\right)\Big|_{t=t_0} \\ \sum_{i=1}^{k-3} \left(\frac{d^{i-1}}{dt^{i-1}}\, q_{i+2}\left(\boldsymbol{h}(t)\right)\right)\Big|_{t=t_0} \\ \vdots \\ \sum_{i=1}^{1} \left(\frac{d^{i-1}}{dt^{i-1}}\, q_{i+k-2}\left(\boldsymbol{h}(t)\right)\right)\Big|_{t=t_0} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\boldsymbol{p}_0\left(h_k(t_0)\right) := \begin{pmatrix} \left(\frac{d^{k-2}}{dt^{k-2}}\, h_k(t)\right)\Big|_{t=t_0} \\ \left(\frac{d^{k-3}}{dt^{k-3}}\, h_k(t)\right)\Big|_{t=t_0} \\ \vdots \\ \left(\frac{d}{dt}\, h_k(t)\right)\big|_{t=t_0} \\ h_k(t_0) \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$A_g(t) := g'_{0,\boldsymbol{x}}[t], \quad B_g(t) := g'_{0,y}[t], \quad C_g(t) := g'_{0,u}[t]$$

For convenience we assume $rank \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix} = n_y + n_c$ for almost every $t \in [t_0, t_f]$ and the pseudo-inverse $\begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+$ is supposed to be essentially bounded. Later we will introduce a weaker condition, because this condition is often too strong, e.g. it does not hold for some box constraints. With this assumption we are able to solve (3.10), (3.11) for (y, u):

$$\begin{pmatrix} y(t) \\ u(t) \end{pmatrix} = \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+ \cdot \left(\begin{pmatrix} -q\left(\boldsymbol{h}(t)\right) - \frac{d^{k-1}}{dt^{k-1}}\, h_k(t) \\ h_{k+1}(t) \end{pmatrix} - \begin{pmatrix} A_g(t) \\ A_c(t) \end{pmatrix} \boldsymbol{x}(t) \right).$$

Inserting into (3.9) gives us the linear ODE

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= A(t)\,\boldsymbol{x}(t) + \tilde{\boldsymbol{h}}(t), \\ 0 &= E\,\boldsymbol{x}(t_0) + \boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) + \boldsymbol{p}_0\left(h_k(t_0)\right), \end{aligned} \tag{3.13}$$

with

$$A(t) := A_f(t) - (B_f(t), C_f(t)) \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+ \begin{pmatrix} A_g(t) \\ A_c(t) \end{pmatrix},$$

$$\tilde{\boldsymbol{h}}(t) := \boldsymbol{h}(t) + (B_f(t), C_f(t))$$

$$\cdot \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+ \begin{pmatrix} -q\left(\boldsymbol{h}(t)\right) - \frac{d^{k-1}}{dt^{k-1}} h_k(t) \\ h_{k+1}(t) \end{pmatrix}.$$

We denote the solution of the initial value problem

$$\dot{\Phi}(t) = A(t)\,\Phi(t), \quad \Phi(t_0) = I$$

by $\Phi_A(\cdot)$. The solution of (3.13) then reads as follows

$$\boldsymbol{x}(t) = \Phi_A(t)\,\left(-E^{-1}\,\left(\boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) + \boldsymbol{p}_0\left(h_k(t_0)\right)\right)\right) + \Phi_A(t)\int_{t_0}^{t} \Phi_A(\tau)^{-1}\,\tilde{\boldsymbol{h}}(\tau)\,d\tau. \tag{3.14}$$

Please note that E is non-singular owing to assumption **(A2)**. Adding (3.2) - (3.4) and exploiting (3.5) - (3.8), and (3.14) yields

$$\begin{aligned}
0 &= \boldsymbol{\lambda_f}^\star\left(\dot{\boldsymbol{x}}(\cdot) - A_f(\cdot)\,\boldsymbol{x}(\cdot) - B_f(\cdot)\,y(\cdot) - C_f(\cdot)\,u(\cdot)\right) \\
&\quad - \lambda_g^\star\left(g'_{k-1,\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right) + \sigma^\top D\,\boldsymbol{x}(t_0) + \ell_0\,\varphi'[t_f]\,\boldsymbol{x}(t_f) \\
&\quad + \eta^\star\left(A_c(\cdot)\,\boldsymbol{x}(\cdot) + B_c(\cdot)\,y(\cdot) + C_c(\cdot)\,u(\cdot)\right) \\
&= \boldsymbol{\lambda_f}^\star\left(\boldsymbol{h}(\cdot)\right) + \lambda_g^\star\left(h_k(\cdot)\right) + \eta^\star\left(h_{k+1}(\cdot)\right) \\
&\quad + \sigma^\top D\,\boldsymbol{x}(t_0) + \ell_0\,\varphi'[t_f]\,\boldsymbol{x}(t_f) \\
&= \boldsymbol{\lambda_f}^\star\left(\boldsymbol{h}(\cdot)\right) + \lambda_g^\star\left(h_k(\cdot)\right) + \eta^\star\left(h_{k+1}(\cdot)\right) + \sigma^\top D\,\boldsymbol{x}(t_0) \\
&\quad + \ell_0\,\varphi'[t_f]\,\Phi_A(t_f)\,\left(-E^{-1}\,\left(\boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) + \boldsymbol{p}_0\left(h_k(t_0)\right)\right)\right) \\
&\quad + \ell_0\,\varphi'[t_f]\,\Phi_A(t_f)\int_{t_0}^{t_f} \Phi_A(\tau)^{-1}\,\tilde{\boldsymbol{h}}(\tau)\,d\tau.
\end{aligned} \tag{3.15}$$

We introduce the notations

$$\begin{aligned}
\boldsymbol{\xi}^\top &:= -\ell_0\,\varphi'[t_f]\,\Phi_A(t_f)\,E^{-1}, \\
\boldsymbol{\lambda_f}(t)^\top &:= \ell_0\,\varphi'[t_f]\,\Phi_A(t_f)\,\Phi_A(t)^{-1}, \\
\lambda_g(t) &:= -\boldsymbol{\lambda_f}(t)^\top\,\left(B_f(t), C_f(t)\right) \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+ \begin{pmatrix} I \\ 0 \end{pmatrix}, \\
\eta(t) &:= \boldsymbol{\lambda_f}(t)^\top\,\left(B_f(t), C_f(t)\right) \begin{pmatrix} B_g(t) & C_g(t) \\ B_c(t) & C_c(t) \end{pmatrix}^+ \begin{pmatrix} 0 \\ I \end{pmatrix}
\end{aligned}$$

and insert these expressions into (3.15) to get

$$\begin{aligned}
&\boldsymbol{\lambda_f}^\star\left(\boldsymbol{h}(\cdot)\right) + \lambda_g^\star\left(h_k(\cdot)\right) + \eta^\star\left(h_{k+1}(\cdot)\right) = \\
&\qquad - \boldsymbol{\xi}^\top\,\boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right) - \int_{t_0}^{t_f} \boldsymbol{\lambda_f}(t)^\top\,\boldsymbol{h}(t)\,dt - \int_{t_0}^{t_f} \lambda_g(t)^\top\,q\left(\boldsymbol{h}(t)\right)\,dt \\
&\qquad - \boldsymbol{\xi}^\top\,\boldsymbol{p}_0\left(h_k(t_0)\right) - \int_{t_0}^{t_f} \lambda_g(t)^\top\,\frac{d^{k-1}}{dt^{k-1}}\,h_k(t)\,dt \\
&\qquad + \int_{t_0}^{t_f} \eta(t)^\top\,h_{k+1}(t)\,dt.
\end{aligned}$$

Thus, from this we obtain the following explicit representation of the multipliers:

$$\boldsymbol{\lambda_f}^{\star}\left(\boldsymbol{h}(\cdot)\right)=-\boldsymbol{\xi}^{\top}\boldsymbol{q}_0\left(\boldsymbol{h}(t_0)\right)-\int_{t_0}^{t_f}\boldsymbol{\lambda_f}(t)^{\top}\boldsymbol{h}(t)\,dt,\\ -\int_{t_0}^{t_f}\lambda_g(t)^{\top}q\left(\boldsymbol{h}(t)\right)\,dt, \tag{3.16}$$

$$\lambda_g^{\star}\left(h_k(\cdot)\right)=-\boldsymbol{\xi}^{\top}\boldsymbol{p}_0\left(h_k(t_0)\right)-\int_{t_0}^{t_f}\lambda_g(t)^{\top}\frac{d^{k-1}}{dt^{k-1}}h_k(t)\,dt \tag{3.17}$$

$$\eta^{\star}\left(h_{k+1}(\cdot)\right)=\int_{t_0}^{t_f}\eta(t)^{\top}h_{k+1}(t)\,dt \tag{3.18}$$

for arbitrary $\boldsymbol{h}\in Y$, $h_k\in W^{n_y}_{k-1,\infty}\left([t_0,t_f]\right)$ and $h_{k+1}\in L^{n_c}_{\infty}\left([t_0,t_f]\right)$. We define the (augmented) Hamilton function by

$$\mathcal{H}:\mathbb{R}^{n_{\boldsymbol{x}}}\times\mathbb{R}^{n_y}\times\mathbb{R}^{n_u}\times\mathbb{R}^{n_{\boldsymbol{x}}}\times\mathbb{R}^{n_y}\times\mathbb{R}^{n_c}\to\mathbb{R},$$
$$\mathcal{H}\left(\boldsymbol{x},y,u,\boldsymbol{\lambda_f},\lambda_g,\eta\right):=\boldsymbol{\lambda_f}^{\top}\boldsymbol{f}\left(\boldsymbol{x},y,u\right)+\lambda_g^{\top}g_0\left(\boldsymbol{x},y,u\right)+\eta^{\top}c\left(\boldsymbol{x},y,u\right).$$

Next we investigate the variational equations (3.2) - (3.4) by using the representation of the multipliers (3.16) - (3.18). For (3.2) we get

$$\begin{aligned}
0&=\boldsymbol{\lambda_f}^{\star}\left(\dot{\boldsymbol{x}}(\cdot)-\boldsymbol{f}'_{\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right)-\lambda_g^{\star}\left(g'_{k-1,\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right)+\eta^{\star}\left(c'_{\boldsymbol{x}}[\cdot]\,\boldsymbol{x}(\cdot)\right)\\
&\quad+\sigma^{\top}D\,\boldsymbol{x}(t_0)+\ell_0\,\varphi'\left(\hat{\boldsymbol{x}}(t_f)\right)\boldsymbol{x}(t_f)\\
&=\boldsymbol{\xi}^{\top}\boldsymbol{q}_0\left(\boldsymbol{f}'_{\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0)-\dot{\boldsymbol{x}}(t_0)\right)+\int_{t_0}^{t_f}\boldsymbol{\lambda_f}(t)^{\top}\left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t)-\dot{\boldsymbol{x}}(t)\right)dt\\
&\quad+\int_{t_0}^{t_f}\lambda_g(t)^{\top}\sum_{i=1}^{k-1}\frac{d^{i-1}}{dt^{i-1}}\left[g'_{i,\boldsymbol{x}}[t]\,\left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t)-\dot{\boldsymbol{x}}(t)\right)\right]dt\\
&\quad+\boldsymbol{\xi}^{\top}\boldsymbol{p}_0\left(g'_{k-1,\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0)\right)+\int_{t_0}^{t_f}\lambda_g(t)^{\top}\frac{d^{k-1}}{dt^{k-1}}\left(g'_{k-1,\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\right)dt\\
&\quad+\int_{t_0}^{t_f}\eta(t)^{\top}c'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\,dt+\sigma^{\top}D\,\boldsymbol{x}(t_0)+\ell_0\,\varphi'\left(\hat{\boldsymbol{x}}(t_f)\right)\boldsymbol{x}(t_f)\\
&=-\boldsymbol{\xi}^{\top}\left(\boldsymbol{q}_0\left(\boldsymbol{f}'_{\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0)-\dot{\boldsymbol{x}}(t_0)\right)+\boldsymbol{p}_0\left(g'_{k-1,\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0)\right)\right)\\
&\quad+\boldsymbol{\lambda_f}(t_0)^{\top}\boldsymbol{x}(t_0)-\boldsymbol{\lambda_f}(t_f)^{\top}\boldsymbol{x}(t_f)\\
&\quad+\int_{t_0}^{t_f}\boldsymbol{\lambda_f}(t)^{\top}\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t)+\dot{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t)^{\top}\boldsymbol{x}(t)\,dt\\
&\quad+\int_{t_0}^{t_f}\lambda_g(t)^{\top}g'_{0,\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\,dt+\int_{t_0}^{t_f}\eta(t)^{\top}c'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\,dt\\
&\quad+\sigma^{\top}D\,\boldsymbol{x}(t_0)+\ell_0\,\varphi'\left(\hat{\boldsymbol{x}}(t_f)\right)\boldsymbol{x}(t_f)\\
&=\int_{t_0}^{t_f}\left(\dot{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t)^{\top}+\mathcal{H}'_{\boldsymbol{x}}[t]\right)\boldsymbol{x}(t)\,dt+\left(\ell_0\,\varphi'\left(\hat{\boldsymbol{x}}(t_f)\right)-\boldsymbol{\lambda_f}(t_f)^{\top}\right)\boldsymbol{x}(t_f)\\
&\quad+\left(\sigma^{\top}D+\boldsymbol{\lambda_f}(t_0)^{\top}+\boldsymbol{\xi}^{\top}\begin{pmatrix}\boldsymbol{g}'_{\boldsymbol{x}}[t_0]\\0\end{pmatrix}\right)\boldsymbol{x}(t_0),
\end{aligned}$$

where we exploited the following equality

$$\begin{aligned}
&\sum_{i=1}^{k-1} \frac{d^{i-1}}{dt^{i-1}} \left[g'_{i,\boldsymbol{x}}[t] \ \left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\right)\right] + \frac{d^{k-1}}{dt^{k-1}} \ \left(g'_{k-1,\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\right) \\
&\quad = \sum_{i=1}^{k-2} \frac{d^{i-1}}{dt^{i-1}} \left[g'_{i,\boldsymbol{x}}[t] \ \left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\right)\right] \\
&\qquad + \left(\frac{d^{k-1}}{dt^{k-1}} \ \left(g'_{k-1,\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\right)\right) + \frac{d^{k-2}}{dt^{k-2}} \left[g'_{k-1,\boldsymbol{x}}[t] \ \left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\right)\right] \\
&\quad = \sum_{i=1}^{k-2} \frac{d^{i-1}}{dt^{i-1}} \left[g'_{i,\boldsymbol{x}}[t] \ \left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\right)\right] \\
&\qquad + \frac{d^{k-2}}{dt^{k-2}} \left[\left(\frac{d}{dt} \ \left(g'_{k-1,\boldsymbol{x}}[t]\right) + g'_{k-1,\boldsymbol{x}}[t]\,\boldsymbol{f}'_{\boldsymbol{x}}[t]\right) \ \boldsymbol{x}(t)\right] \\
&\quad = \sum_{i=1}^{k-2} \frac{d^{i-1}}{dt^{i-1}} \left[g'_{i,\boldsymbol{x}}[t] \ \left(\boldsymbol{f}'_{\boldsymbol{x}}[t]\,\boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\right)\right] + \frac{d^{k-2}}{dt^{k-2}} \ \left[g'_{k-2,\boldsymbol{x}}[t]\,\boldsymbol{x}(t)\right] \\
&\quad \vdots \\
&\quad = \left(\frac{d}{dt} \ \left(g'_{1,\boldsymbol{x}}[t]\right) + g'_{1,\boldsymbol{x}}[t]\,\boldsymbol{f}'_{\boldsymbol{x}}[t]\right) \ \boldsymbol{x}(t) \\
&\quad = g'_{0,\boldsymbol{x}}[t]\,\boldsymbol{x}(t).
\end{aligned}$$

Similarly we obtain

$$\boldsymbol{q}_0 \left(\boldsymbol{f}'_{\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0) - \dot{\boldsymbol{x}}(t_0)\right) + \boldsymbol{p}_0 \left(g'_{k-1,\boldsymbol{x}}[t_0]\,\boldsymbol{x}(t_0)\right) = \begin{pmatrix} \boldsymbol{g}'_{\boldsymbol{x}}[t_0] \\ 0 \end{pmatrix} \boldsymbol{x}(t_0).$$

With this (3.3) gives us

$$\begin{aligned}
0 &= -\boldsymbol{\lambda}_{\boldsymbol{f}}{}^{\star} \left(\boldsymbol{f}'_y[\cdot]\,y(\cdot)\right) + \eta^{\star} \left(c'_y[\cdot]\,y(\cdot)\right) \\
&= \boldsymbol{\xi}^{\top}\,\boldsymbol{q}_0 \left(\boldsymbol{f}'_y[t_0]\,y(t_0)\right) + \int_{t_0}^{t_f} \ \boldsymbol{\lambda}_{\boldsymbol{f}}(t)^{\top}\,\boldsymbol{f}'_y[t]\,y(t)\,dt \\
&\quad + \int_{t_0}^{t_f} \ \lambda_g(t)^{\top}\,q \left(\boldsymbol{f}'_y[t]\,y(t)\right)\,dt + \int_{t_0}^{t_f} \ \eta(t)^{\top}\,c'_y[t]\,y(t)\,dt \\
&= \int_{t_0}^{t_f} \ \boldsymbol{\lambda}_{\boldsymbol{f}}(t)^{\top}\,\boldsymbol{f}'_y[t]\,y(t)\,dt + \int_{t_0}^{t_f} \ \lambda_g(t)^{\top}\,g_{1,\boldsymbol{x}}[t]\,\boldsymbol{f}'_y[t]\,y(t)\,dt \\
&\quad + \int_{t_0}^{t_f} \ \eta(t)^{\top}\,c'_y[t]\,y(t)\,dt \\
&= \int_{t_0}^{t_f} \ \boldsymbol{\lambda}_{\boldsymbol{f}}(t)^{\top}\,\boldsymbol{f}'_y[t]\,y(t)\,dt + \int_{t_0}^{t_f} \ \lambda_g(t)^{\top}\,g_{0,y}[t]\,y(t)\,dt \\
&\quad + \int_{t_0}^{t_f} \ \eta(t)^{\top}\,c'_y[t]\,y(t)\,dt
\end{aligned}$$

$$= \int_{t_0}^{t_f} \mathcal{H}'_y[t]\, y(t)\, dt$$

and analog for (3.4) we get

$$\begin{aligned} 0 &= -\boldsymbol{\lambda_f}^\star \left(\boldsymbol{f}'_u[\cdot]\, u(\cdot)\right) + \eta^\star \left(c'_u[\cdot]\, u(\cdot)\right) \\ &= \int_{t_0}^{t_f} \mathcal{H}'_u[t]\, u(t)\, dt. \end{aligned}$$

Using a variational Lemma (see [8, p.115f]) we proved the following local minimum principle, which is the first main result of this paper.

Theorem 3.1. *Let Assumptions **(A1)-(A4)** be satisfied, and let rank* $\begin{pmatrix} g_{0,y}[t] & g_{0,u}[t] \\ c'_y[t] & c'_u[t] \end{pmatrix} = n_y + n_c$ *almost everywhere in* $[t_0, t_f]$ *and let the pseudo-inverse* $\begin{pmatrix} g_{0,y}[t] & g_{0,u}[t] \\ c'_y[t] & c'_u[t] \end{pmatrix}^+$ *be essentially bounded. Then there exist multipliers*

$$\ell_0 \in \mathbb{R}, \quad \boldsymbol{\lambda_f} \in W^{n_{\boldsymbol{x}}}_{1,\infty}\left([t_0, t_f]\right), \quad \lambda_g \in L^{n_y}_\infty\left([t_0, t_f]\right), \quad \eta \in L^{n_c}_\infty\left([t_0, t_f]\right), \quad \boldsymbol{\lambda_0} \in \mathbb{R}^{n_{\boldsymbol{x}}}$$

such that the following conditions hold:

(i) $\ell_0 \geq 0$, $(\ell_0, \boldsymbol{\lambda_f}, \lambda_g, \eta, \boldsymbol{\lambda_0}) \neq 0$

(ii) *Adjoint DAE: For almost every* $t \in [t_0, t_f]$ *we have*

$$\begin{aligned} \dot{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t) &= -\mathcal{H}'_{\boldsymbol{x}}\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \boldsymbol{\lambda_f}(t), \lambda_g(t), \eta(t)\right)^\top, \\ 0 &= \mathcal{H}'_y\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \boldsymbol{\lambda_f}(t), \lambda_g(t), \eta(t)\right)^\top. \end{aligned}$$

(iii) *Transversality conditions:*

$$\begin{aligned} \boldsymbol{\lambda_f}(t_0)^\top &= -\boldsymbol{\lambda_0}^\top E, \\ \boldsymbol{\lambda_f}(t_f)^\top &= \ell_0\, \varphi'\left(\hat{\boldsymbol{x}}(t_f)\right). \end{aligned}$$

(iv) *Stationarity of the Hamilton function: For almost every* $t \in [t_0, t_f]$ *we have*

$$0 = \mathcal{H}'_u\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t), \boldsymbol{\lambda_f}(t), \lambda_g(t), \eta(t)\right)^\top.$$

(v) *Complementarity conditions: For almost every* $t \in [t_0, t_f]$ *we have*

$$\eta(t) \geq 0 \quad \text{and} \quad \eta(t)^\top c\left(\hat{\boldsymbol{x}}(t), \hat{y}(t), \hat{u}(t)\right) = 0.$$

Remark 3.2. As noted before, the rank condition

$$rank \begin{pmatrix} g_{0,y}[t] & g_{0,u}[t] \\ c'_y[t] & c'_u[t] \end{pmatrix} = n_y + n_c$$

is often too strong. By interpreting the algebraic equation $g_0[t] = 0$ as active constraints, one can show similarly to [18], that it is sufficient to assume there exist $\varrho > 0$ and $\alpha > 0$ such that

$$\left\| \begin{pmatrix} B_g(t) & C_g(t) \\ B_c^\varrho(t) & C_c^\varrho(t) \end{pmatrix}^\top d \right\| \geq \alpha\, \|d\| \quad \text{for all } d \in \mathbb{R}^{n_y + i_\varrho(t)} \quad \text{and a.e. } t \in [t_0, t_f]$$

where

$$I := \{1,2,\dots,n_c\}\,, \quad I_\varrho(t) := \{i \in I \mid c_i[t] \geq -\varrho\}\,, \quad i_\varrho(t) := |I_\varrho(t)|\,,$$
$$B_c^\varrho(t) := \left[c'_{i,y}[t]\right]_{i\in I_\varrho(t)}, \quad C_c^\varrho(t) := \left[c'_{i,u}[t]\right]_{i\in I_\varrho(t)}.$$

4. Sufficient conditions

In this section we derive second order sufficient conditions for **(OCP)** in form of a Riccati equation. To this end we assume the following:

(A5) There exists a KKT point

$$\left(\hat{\boldsymbol{x}}, \hat{y}, \hat{u}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}, \hat{\ell}_0\right),$$

which satisfies the local minimum principle in Theorem 3.1 with $\hat{\ell}_0 = 1$.

This assumption holds if some constraint qualifications are satisfied (see e.g. [8, p.148]). Furthermore we introduce the Lagrange function

$$\mathcal{L}\left(\boldsymbol{x}, y, u, \boldsymbol{\lambda}_{\boldsymbol{f}}, \lambda_g, \eta, \boldsymbol{\lambda}_{\boldsymbol{0}}\right) := \varphi\left(\boldsymbol{x}(t_f)\right) + \boldsymbol{\lambda}_{\boldsymbol{0}}^\top \begin{pmatrix} \boldsymbol{g}\left(\boldsymbol{x}(t_0)\right) \\ D\left(\boldsymbol{x}(t_0) - \boldsymbol{x}^0\right) \end{pmatrix}$$
$$+ \int_{t_0}^{t_f} \mathcal{H}\left(\boldsymbol{x}(t), y(t), u(t), \boldsymbol{\lambda}_{\boldsymbol{f}}(t), \lambda_g(t), \eta(t)\right) - \boldsymbol{\lambda}_{\boldsymbol{f}}(t)^\top \dot{\boldsymbol{x}}(t)\, dt.$$

Our goal is to find sufficient conditions such that the coercivity condition

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) \geq \gamma\, \|\boldsymbol{z}\|^2_{\hat{Z}}$$

holds for some $\gamma > 0$. To this end we introduce the set of active indices

$$I_a(t) := \{i \in I \mid c_i[t] = 0\}\,, \quad i_a(t) := |I_a(t)|$$

and the set of indices of those active constraints where the strict complementarity condition holds

$$I_+(t) := \{i \in I_a(t) \mid \hat{\eta}_i(t) > 0\}\,.$$

We introduce the following abbreviations

$$A_c^a(t) := \left[c'_{i,\boldsymbol{x}}[t]\right]_{i\in I_a(t)}, \quad B_c^a(t) := \left[c'_{i,y}[t]\right]_{i\in I_a(t)}, \quad C_c^a(t) := \left[c'_{i,u}[t]\right]_{i\in I_a(t)},$$
$$A_c^+(t) := \left[c'_{i,\boldsymbol{x}}[t]\right]_{i\in I_+(t)}, \quad B_c^+(t) := \left[c'_{i,y}[t]\right]_{i\in I_+(t)}, \quad C_c^+(t) := \left[c'_{i,u}[t]\right]_{i\in I_+(t)}.$$

Furthermore we assume

(A6) There exists some $\alpha > 0$ such that

$$\begin{pmatrix} B_g(t) & C_g(t) \\ B_c^a(t) & C_c^a(t) \end{pmatrix} \begin{pmatrix} B_g(t) & C_g(t) \\ B_c^a(t) & C_c^a(t) \end{pmatrix}^\top \geq \alpha\, I_{n_y + i_a(t)} \quad \text{for a.e. } t \in [t_0, t_f].$$

(A7) The strengthened Legendre Clesch condition is satisfied, namely there exists some $\tilde{\beta} > 0$ such that

$$\left(d_y^\top, d_u^\top\right) \begin{pmatrix} \mathcal{H}''_{yy}[t] & \mathcal{H}''_{yu}[t] \\ \mathcal{H}''_{uy}[t] & \mathcal{H}''_{uu}[t] \end{pmatrix} \begin{pmatrix} d_y \\ d_u \end{pmatrix} \geq \tilde{\beta} \left\| \begin{pmatrix} d_y \\ d_u \end{pmatrix} \right\|^2$$

for almost every $t \in [t_0, t_f]$ and for all

$$\begin{pmatrix} d_y \\ d_u \end{pmatrix} \in \ker \begin{pmatrix} B_g(t) & C_g(t) \\ B_c^+(t) & C_c^+(t) \end{pmatrix}.$$

We define the function

$$\kappa\left(\boldsymbol{x}_0, \boldsymbol{x}_f, \boldsymbol{\lambda_0}\right) := \varphi\left(\boldsymbol{x}_f\right) + \boldsymbol{\lambda_0}^\top \begin{pmatrix} \boldsymbol{g}\left(\boldsymbol{x}_0\right) \\ D\left(\boldsymbol{x}_0 - \boldsymbol{x}^0\right) \end{pmatrix}.$$

The next step is to derive a suitable Riccati equation. To this end we consider the following parametric mathematical program depending on the parameter $\zeta = (\boldsymbol{x}, \boldsymbol{\lambda_f}) \in \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}}$:

(MP(ζ))

$$\begin{aligned} \text{Minimize} \quad & \boldsymbol{\lambda_f}^\top \boldsymbol{f}\left(\boldsymbol{x}, y, u\right) \\ \text{w.r.t.} \quad & \begin{pmatrix} y \\ u \end{pmatrix} \in \mathbb{R}^{n_y + n_u} \\ \text{s.t.} \quad & g_0\left(\boldsymbol{x}, y, u\right) = 0, \\ & c\left(\boldsymbol{x}, y, u\right) \leq 0. \end{aligned}$$

For $\hat{\zeta}(t) = \left(\hat{\boldsymbol{x}}(t), \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}(t)\right)$ the problem **(MP($\hat{\zeta}(t)$))** has the solution and multipliers $\left(\hat{y}(t), \hat{u}(t), \hat{\lambda}_g(t), \hat{\eta}(t)\right)$ by virtue of **(A6)** and **(A7)**. Using the sensitivity result from [28] we get the following: There exists a $\rho > 0$ such that for all $t \in [t_0, t_f]$ and every

$$\zeta \in \left\{ \zeta \in \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}} \,\middle|\, \left\| \zeta - \hat{\zeta}(t) \right\| \leq \rho \right\}$$

there exists a locally unique solution $(y(\zeta), u(\zeta))$ of **(MP(ζ))** and unique associated multipliers $(\lambda_g(\zeta), \eta^a(\zeta)) \in \mathbb{R}^{n_y} \times \mathbb{R}^{i_a(t)}$. Furthermore $(y(\zeta), u(\zeta))$ and $(\lambda_g(\zeta), \eta^a(\zeta))$ are Fréchet differentiable functions with respect to ζ at $\hat{\zeta}(t)$. For arbitrary increment $d \in \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}}$ the differentials

$$\begin{pmatrix} y_\zeta(\hat{\zeta}(t)) \\ u_\zeta(\hat{\zeta}(t)) \end{pmatrix} d \in \mathbb{R}^{n_y + n_u}, \qquad \begin{pmatrix} \lambda_{g,\zeta}(\hat{\zeta}(t)) \\ \eta^a_\zeta(\hat{\zeta}(t)) \end{pmatrix} d \in \mathbb{R}^{n_y + i_a(t)}$$

are given as the solution and the associated multipliers of the following linear quadratic mathematical program:

(LQR$_d(t)$)

$$\begin{aligned} \text{Min} \quad & \frac{1}{2}\left(v^\top, w^\top\right) \begin{pmatrix} \mathcal{H}''_{yy}[t] & \mathcal{H}''_{yu}[t] \\ \mathcal{H}''_{uy}[t] & \mathcal{H}''_{uu}[t] \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} \\ & + \left(v^\top, w^\top\right) \begin{pmatrix} \mathcal{H}''_{yx}[t] & B_f(t)^\top \\ \mathcal{H}''_{ux}[t] & C_f(t)^\top \end{pmatrix} d \\ \text{w.r.t.} \quad & \begin{pmatrix} v \\ w \end{pmatrix} \in \mathbb{R}^{n_y + n_u} \end{aligned}$$

$$\text{s.t.} \quad \begin{pmatrix} B_g(t) & C_g(t) \\ B_c^a(t) & C_c^a(t) \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} + \begin{pmatrix} A_g(t) & 0 \\ A_c^a(t) & 0 \end{pmatrix} d = 0.$$

The necessary conditions yield the solution and multipliers

$$\begin{pmatrix} y_\zeta(\hat{\zeta}(t)) \\ u_\zeta(\hat{\zeta}(t)) \\ \lambda_{g,\zeta}(\hat{\zeta}(t)) \\ \eta_\zeta^a(\hat{\zeta}(t)) \end{pmatrix} d = -R^a(t)^{-1} \, (S^a(t), K^a(t)) \, d,$$

where

$$R^a(t) := \begin{pmatrix} \mathcal{H}''_{yy}[t] & \mathcal{H}''_{yu}[t] & B_g(t)^\top & B_c^a(t)^\top \\ \mathcal{H}''_{uy}[t] & \mathcal{H}''_{uu}[t] & C_g(t)^\top & C_c^a(t)^\top \\ B_g(t) & C_g(t) & 0 & 0 \\ B_c^a(t) & C_c^a(t) & 0 & 0 \end{pmatrix},$$

$$S^a(t) := \begin{pmatrix} \mathcal{H}''_{y\boldsymbol{x}}[t] \\ \mathcal{H}''_{u\boldsymbol{x}}[t] \\ A_g(t) \\ A_c^a(t) \end{pmatrix}, \quad K^a(t) := \begin{pmatrix} B_f(t)^\top \\ C_f(t)^\top \\ 0 \\ 0 \end{pmatrix}$$

with $R^a(t) \in \mathbb{R}^{(n_y+n_u+n_y+i_a(t))\times(n_y+n_u+n_y+i_a(t))}$, $S^a(t) \in \mathbb{R}^{(n_y+n_u+n_y+i_a(t))\times n_{\boldsymbol{x}}}$ and $K^a(t) \in \mathbb{R}^{(n_y+n_u+n_y+i_a(t))\times n_{\boldsymbol{x}}}$. Since $d \in \mathbb{R}^{n_{\boldsymbol{x}}} \times \mathbb{R}^{n_{\boldsymbol{x}}}$ was arbitrarily chosen we obtain the partial derivatives

$$\begin{pmatrix} y_{\boldsymbol{x}}(\hat{\zeta}(t)) \\ u_{\boldsymbol{x}}(\hat{\zeta}(t)) \\ \lambda_{g,\boldsymbol{x}}(\hat{\zeta}(t)) \\ \eta_{\boldsymbol{x}}^a(\hat{\zeta}(t)) \end{pmatrix} = -R^a(t)^{-1} \, S^a(t),$$

$$\begin{pmatrix} y_{\boldsymbol{\lambda_f}}(\hat{\zeta}(t)) \\ u_{\boldsymbol{\lambda_f}}(\hat{\zeta}(t)) \\ \lambda_{g,\boldsymbol{\lambda_f}}(\hat{\zeta}(t)) \\ \eta_{\boldsymbol{\lambda_f}}^a(\hat{\zeta}(t)) \end{pmatrix} = -R^a(t)^{-1} \, K^a(t).$$

Next we consider the following linear quadratic optimal control problem:

$$\begin{aligned} \text{Min} \quad & \tfrac{1}{2} \int_{t_0}^{t_f} \boldsymbol{z}(t)^\top \, \mathcal{H}''_{\boldsymbol{zz}}[t] \, \boldsymbol{z}(t) \, dt \\ & + \tfrac{1}{2} \left(\boldsymbol{x}(t_f)^\top \, \kappa''_{\boldsymbol{x}_f \boldsymbol{x}_f} \, \boldsymbol{x}(t_f) + \boldsymbol{x}(t_0)^\top \, \kappa''_{\boldsymbol{x}_0 \boldsymbol{x}_0} \, \boldsymbol{x}(t_0) \right) \\ \text{w.r.t.} \quad & \boldsymbol{z} \in Z \\ \text{s.t.} \quad & \dot{\boldsymbol{x}}(t) = A_f(t) \, \boldsymbol{x}(t) + B_f(t) \, y(t) + C_f(t) \, u(t), \\ & 0 = A_g(t) \, \boldsymbol{x}(t) + B_g(t) \, y(t) + C_g(t) \, u(t), \\ & 0 = A_c^a(t) \, \boldsymbol{x}(t) + B_c^a(t) \, y(t) + C_c^a(t) \, u(t), \\ & 0 = E \, \boldsymbol{x}(t_0). \end{aligned}$$

Since the matrix E is non-singular by assumption **(A2)**, we can deduce that $\boldsymbol{x}(t_0) = 0$ and therefore consider the equivalent problem

(LQOCP)

$$\begin{array}{lrcl} \text{Min} & \multicolumn{3}{l}{\frac{1}{2}\int_{t_0}^{t_f} \boldsymbol{z}(t)^\top \mathcal{H}''_{\boldsymbol{zz}}[t]\,\boldsymbol{z}(t)\,dt + \frac{1}{2}\boldsymbol{x}(t_f)^\top \kappa''_{\boldsymbol{x}_f\boldsymbol{x}_f}\,\boldsymbol{x}(t_f)} \\ \text{w.r.t.} & \multicolumn{3}{l}{\boldsymbol{z} \in Z} \\ \text{s.t.} & \dot{\boldsymbol{x}}(t) &=& A_f(t)\,\boldsymbol{x}(t) + B_f(t)\,y(t) + C_f(t)\,u(t), \\ & 0 &=& A_g(t)\,\boldsymbol{x}(t) + B_g(t)\,y(t) + C_g(t)\,u(t), \\ & 0 &=& A_c^a(t)\,\boldsymbol{x}(t) + B_c^a(t)\,y(t) + C_c^a(t)\,u(t), \\ & 0 &=& \boldsymbol{x}(t_0). \end{array}$$

Our goal is to derive sufficient conditions under which the objective function of **(LQOCP)** is coercive on the feasible set. According to [26, ch.III] the associated Riccati equation to **(LQOCP)** looks as follows:

$$\begin{aligned} \dot{P}(t) = {} & -P(t)\,A_f(t) - A_f(t)^\top P(t) - \mathcal{H}''_{\boldsymbol{xx}}[t] \\ & + \left(P(t)\,K^a(t)^\top + S^a(t)^\top\right) R^a(t)^{-1}\left(K^a(t)\,P(t) + S^a(t)\right), \end{aligned} \tag{4.1}$$

$$P(t_f) = \kappa''_{\boldsymbol{x}_f\boldsymbol{x}_f}, \tag{4.2}$$

where $P(t) \in \mathbb{R}^{n_{\boldsymbol{x}} \times n_{\boldsymbol{x}}}$ is symmetric. To prove that there exists a $\gamma > 0$ such that the coercivity condition

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) \geq \gamma\,\|\boldsymbol{z}\|^2_{\hat{Z}} \tag{4.3}$$

holds for all feasible $\boldsymbol{z} = (\boldsymbol{x}, y, u)$, we assume the following:

(A8) The Riccati equation (4.1), (4.2) has a bounded solution $P(\cdot)$.

$\boldsymbol{z} = (\boldsymbol{x}, y, u)$ satisfies the linear differential equation

$$0 = \dot{\boldsymbol{x}}(t) - A_f(t)\,\boldsymbol{x}(t) - B_f(t)\,y(t) - C_f(t)\,u(t).$$

Multiplying by $-2\,\boldsymbol{x}(t)^\top P(t)$ from the left, using integration by parts, exploiting the differential equations for $\boldsymbol{x}$ and P, and rearranging terms yields

$$\begin{aligned} 0 &= 2\int_{t_0}^{t_f} \boldsymbol{x}(t)^\top P(t)\,\left(A_f(t)\,\boldsymbol{x}(t) + B_f(t)\,y(t) + C_f(t)\,u(t) - \dot{\boldsymbol{x}}(t)\right)\,dt \\ &= \int_{t_0}^{t_f} \boldsymbol{x}(t)^\top P(t)\,A_f(t)\,\boldsymbol{x}(t) + \boldsymbol{x}(t)^\top A_f(t)^\top P(t)\,\boldsymbol{x}(t) \\ &\quad + 2\,\boldsymbol{x}(t)^\top P(t)\,\left(B_f(t)\,y(t) + C_f(t)\,u(t)\right) + \boldsymbol{x}(t)^\top \dot{P}(t)\,\boldsymbol{x}(t)\,dt \\ &\quad - \boldsymbol{x}(t_f)^\top \kappa''_{\boldsymbol{x}_f\boldsymbol{x}_f}\,\boldsymbol{x}(t_f) \\ &= \int_{t_0}^{t_f} 2\,\boldsymbol{x}(t)^\top P(t)\,\left(B_f(t)\,y(t) + C_f(t)\,u(t)\right) - \boldsymbol{x}(t)^\top \mathcal{H}''_{\boldsymbol{xx}}[t]\,\boldsymbol{x}(t) \\ &\quad + \boldsymbol{x}(t)^\top \left(P(t)\,K^a(t)^\top + S^a(t)^\top\right) R^a(t)^{-1}\left(K^a(t)\,P(t) + S^a(t)\right)\boldsymbol{x}(t)\,dt \\ &\quad - \boldsymbol{x}(t_f)^\top \kappa''_{\boldsymbol{x}_f\boldsymbol{x}_f}\,\boldsymbol{x}(t_f). \end{aligned}$$

Adding this equality to

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) = \int_{t_0}^{t_f} \boldsymbol{z}(t)^\top \mathcal{H}''_{\boldsymbol{zz}}[t]\, \boldsymbol{z}(t)\, dt + \boldsymbol{x}(t_f)^\top \kappa''_{\boldsymbol{x}_f \boldsymbol{x}_f}\, \boldsymbol{x}(t_f)$$

and performing some lengthly but basic calculations yields

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) = \int_{t_0}^{t_f} v^a(t)^\top R^a(t)\, v^a(t)\, dt,$$

where

$$v^a(t) := R^a(t)^{-1} \left(K^a(t)\, P(t) + S^a(t)\right) \boldsymbol{x}(t) + \begin{pmatrix} y(t) \\ u(t) \\ 0 \\ 0 \end{pmatrix}.$$

Next we want to prove, that

$$v^a(t) \in \ker \begin{pmatrix} B_g(t) & C_g(t) & 0 & 0 \\ B_c^a(t) & C_c^a(t) & 0 & 0 \end{pmatrix}. \tag{4.4}$$

We consider the equalities

$$0 = \begin{pmatrix} A_g(t) & B_g(t) & C_g(t) \\ A_c^a(t) & B_c^a(t) & C_c^a(t) \end{pmatrix} \begin{pmatrix} \boldsymbol{x}(t) \\ y(t) \\ u(t) \end{pmatrix}$$

and

$$\begin{pmatrix} B_g(t) & C_g(t) & 0 & 0 \\ B_c^(t) & C_c^a(t) & 0 & 0 \end{pmatrix} R^\alpha(t)^{-1} = \begin{pmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix}.$$

With this we get

$$\begin{aligned} &\begin{pmatrix} B_g(t) & C_g(t) & 0 & 0 \\ B_c^a(t) & C_c^a(t) & 0 & 0 \end{pmatrix} v^a(t) \\ &\quad = \begin{pmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \left(K^a(t)\, P(t) + S^a(t)\right) \boldsymbol{x}(t) \\ &\qquad + \begin{pmatrix} B_g(t) & C_g(t) \\ B_c^a(t) & C_c^a(t) \end{pmatrix} \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} \\ &\quad = \begin{pmatrix} A_g(t) & B_g(t) & C_g(t) \\ A_c^a(t) & B_c^a(t) & C_c^a(t) \end{pmatrix} \begin{pmatrix} \boldsymbol{x}(t) \\ y(t) \\ u(t) \end{pmatrix} \\ &\quad = 0. \end{aligned}$$

(4.4) implies

$$v^a(t)^\top R^a(t)\, v^a(t) \geq 0,$$

which gives us

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) \geq 0.$$

Now we choose a sufficiently small $\beta > 0$ such that assumption **(A7)** is satisfied and the Riccati-equation, where we replace $R^a(t)$ by

$$R^a_\beta(t) := R^a(t) - \beta \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

has a bounded solution $P_\beta(\cdot)$, which is possible by an extension of a basic result for Volterra integral equations, see [30, p. 103]. Taking the same steps as before we obtain

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) - \beta\, \|(y,u)\|_2^2 = \int_{t_0}^{t_f} v^a_\beta(t)^\top R^a_\beta(t)\, v^a_\beta(t)\, dt \geq 0$$

with

$$v^a_\beta(t) := R^a_\beta(t)^{-1}\left(K^a(t)\, P_\beta(t) + S^a(t)\right) \boldsymbol{x}(t) + \begin{pmatrix} y(t) \\ u(t) \\ 0 \\ 0 \end{pmatrix}$$

which gives us the inequality

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\boldsymbol{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) \geq \beta\, \|(y,u)\|_2^2 .$$

[11, Lemma 3.3.4] prove the following inequality

$$\left\|\Phi_{A_f}(t)\, \Phi_{A_f}(s)^{-1}\right\| \leq e^{\|A_f\|_\infty (t-s)},$$

which we exploit to get a bound for $\boldsymbol{x}$, namely

$$\begin{aligned}
\|\boldsymbol{x}\|_2^2 &= \int_{t_0}^{t_f} \left\| \int_{t_0}^{t} \Phi_{A_f}(t)\, \Phi_{A_f}(\tau)^{-1} \left(B_f(\tau), C_f(\tau)\right) \begin{pmatrix} y(\tau) \\ u(\tau) \end{pmatrix} d\tau \right\|^2 dt \\
&\leq \int_{t_0}^{t_f} \int_{t_0}^{t} \left\| \Phi_{A_f}(t)\, \Phi_{A_f}(\tau)^{-1} \left(B_f(\tau), C_f(\tau)\right) \begin{pmatrix} y(\tau) \\ u(\tau) \end{pmatrix} \right\|^2 d\tau\, dt \\
&\leq \int_{t_0}^{t_f} \int_{t_0}^{t_f} \left(e^{\|A_f\|_\infty (t-t_0)}\, \|(B_f, C_f)\|_\infty \left\| \begin{pmatrix} y(\tau) \\ u(\tau) \end{pmatrix} \right\| \right)^2 d\tau\, dt \\
&\leq \frac{\|(B_f, C_f)\|_\infty^2}{2\, \|A_f\|_\infty} \left(e^{2\, \|A_f\|_\infty (t_f - t_0)} - 1 \right) \left\| \begin{pmatrix} y \\ u \end{pmatrix} \right\|_2^2 .
\end{aligned}$$

We define the constant $\delta_0 := \frac{\|(B_f,C_f)\|_\infty^2}{2\,\|A_f\|_\infty} \left(e^{2\,\|A_f\|_\infty (t_f - t_0)} - 1 \right)$. For $\dot{\boldsymbol{x}}$ we obtain

$$\begin{aligned}
\|\dot{\boldsymbol{x}}\|_2^2 &= \int_{t_0}^{t_f} \left\| A_f(t)\, \boldsymbol{x}(t) + \left(B_f(t), C_f(t)\right) \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} \right\|^2 dt \\
&\leq \left(\|A_f\|_\infty\, \|\boldsymbol{x}\|_2 + \|(B_f, C_f)\|_\infty \left\| \begin{pmatrix} y \\ u \end{pmatrix} \right\|_2 \right)^2 \\
&\leq \left(\|A_f\|_\infty\, \sqrt{\delta_0} + \|(B_f, C_f)\|_\infty \right)^2 \left\| \begin{pmatrix} y \\ u \end{pmatrix} \right\|_2^2
\end{aligned}$$

and we define the constant $\delta_1 := \left(\|A_f\|_\infty \sqrt{\delta_0} + \|(B_f, C_f)\|_\infty\right)^2$. We introduce the notation

$$\boldsymbol{f}^{(i)}\left(\boldsymbol{x}^{(i-1)}(\cdot)\right) := \begin{pmatrix} f_i\left(\boldsymbol{x}^{(i-1)}(\cdot)\right) \\ f_{i+1}\left(\boldsymbol{x}^{(i)}(\cdot)\right) \\ \vdots \\ f_{k-1}\left(\boldsymbol{x}^{(k-2)}(\cdot)\right) \end{pmatrix}, \quad A_f^{(i)}(t) := \frac{\partial \boldsymbol{f}^{(i)}}{\partial \boldsymbol{x}^{(i-1)}}[t].$$

The derivative of $\boldsymbol{x}^{(2)}$ with respect to t then reads as follows

$$\dot{\boldsymbol{x}}^{(2)}(t) = A_f^{(2)}(t)\, \boldsymbol{x}^{(1)}(t).$$

Derivation with respect to t yields

$$\ddot{\boldsymbol{x}}^{(2)}(t) = \left(\frac{d}{dt} A_f^{(2)}(t)\right) \boldsymbol{x}^{(1)}(t) + A_f^{(2)}(t)\, \dot{\boldsymbol{x}}^{(1)}(t).$$

so we get the estimate

$$\begin{aligned} \left\|\ddot{\boldsymbol{x}}^{(2)}\right\|_2^2 &= \int_{t_0}^{t_f} \left\| \left(\frac{d}{dt} A_f^{(2)}(t)\right) \boldsymbol{x}^{(1)}(t) + A_f^{(2)}(t)\, \dot{\boldsymbol{x}}^{(1)}(t) \right\|^2 dt \\ &\le \left(\left\| \left(\frac{d}{dt} A_f^{(2)}(\cdot)\right) \right\|_\infty \left\|\boldsymbol{x}^{(1)}\right\|_2 + \left\|A_f^{(2)}\right\|_\infty \left\|\dot{\boldsymbol{x}}^{(1)}\right\|_2 \right)^2 \\ &\le \left(\left\| \left(\frac{d}{dt} A_f^{(2)}(\cdot)\right) \right\|_\infty \sqrt{\delta_0} + \left\|A_f^{(2)}\right\|_\infty \sqrt{\delta_1} \right)^2 \left\| \begin{pmatrix} y \\ u \end{pmatrix} \right\|_2^2 \end{aligned}$$

and we define $\delta_2 := \left(\left\| \left(\frac{d}{dt} A_f^{(2)}(\cdot)\right) \right\|_\infty \sqrt{\delta_0} + \left\|A_f^{(2)}\right\|_\infty \sqrt{\delta_1} \right)^2$. Similarly we get an upper bound δ_3 for the third derivative of $\boldsymbol{x}^{(3)}$, a bound δ_4 for the fourth derivative of $\boldsymbol{x}^{(4)}$ up to δ_{k-1} for the $(k-1)$-st derivative $\boldsymbol{x}^{(k-1)}$. Altogether we have the bound

$$\|\boldsymbol{z}\|_{\hat{Z}}^2 \le \delta \left\| \begin{pmatrix} y \\ u \end{pmatrix} \right\|_2^2$$

with $\delta := \max\{\delta_0, \delta_1, \delta_2, \dots, \delta_{k-1}, 1\}$, which finally gives us the coercivity condition

$$\mathcal{L}''_{\boldsymbol{zz}}\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}.\hat{\boldsymbol{\lambda}}_{\mathbf{0}}\right)(\boldsymbol{z}, \boldsymbol{z}) \ge \beta \, \|(y,u)\|_2^2 \ge \frac{\beta}{\delta} \, \|\boldsymbol{z}\|_{\hat{Z}}^2 .$$

We thus proved the following sufficient condition for **(OCP)**, which is the second main result of this paper:

Theorem 4.1. *Let the conditions **(A2)** - **(A8)** for a KKT point* $\left(\hat{\boldsymbol{z}}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \hat{\lambda}_g, \hat{\eta}, \hat{\boldsymbol{\lambda}}_{\mathbf{0}}\right)$ *be satisfied. Then there exist* $\chi > 0$ *and* $\epsilon > 0$ *such that*

$$J(\boldsymbol{z}) \ge J(\hat{\boldsymbol{z}}) + \chi \, \|\boldsymbol{z} - \hat{\boldsymbol{z}}\|_{\hat{Z}}^2 \tag{4.5}$$

for all feasible $\boldsymbol{z} \in \{\boldsymbol{v} \in Z \,|\, \|\boldsymbol{v} - \hat{\boldsymbol{z}}\|_Z < \epsilon\}$.

Proof. Since the coercivity condition (4.3) holds, according to [21, Thm. 3.5] there exist $\chi > 0$ and $\epsilon > 0$ such that (4.5) holds for every feasible $z \in \{\boldsymbol{v} \in Z \,|\, \|\boldsymbol{v} - \hat{\boldsymbol{z}}\|_Z < \epsilon\}$. ☐

5. Example

We consider the following index 3 problem equivalent to the Minimum-Energy Problem (see e.g.: [8, p. 4]):

Minimize

$$\frac{1}{2}\,x_1(1)^\top \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_1(1) + (0,3,1)\; x_1(1)$$

with respect to

$$x_1 \in W^3_{1,\infty}\left([0,1]\right), x_2 \in W^2_{2,\infty}\left([0,1]\right), y \in L^1_\infty\left([0,1]\right), u \in L^1_\infty\left([0,1]\right)$$

subject to

$$\begin{aligned} \dot{x}_1(t) &= \begin{pmatrix} u(t)-y(t) \\ u(t) \\ \frac{1}{2}\,u(t)^2 \end{pmatrix}, \\ \dot{x}_2(t) &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} x_1(t) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x_2(t), \\ 0 &= (0,1)\; x_2(t) - 1, \end{aligned}$$

and the initial conditions

$$\begin{aligned} 0 &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(x_1(0) - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right), \\ 0 &= (1,0) \left(x_2(0) - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

We obtain the hidden constraints by differentiating the algebraic equation twice with respect to time t, namely

$$\begin{aligned} &0 = g_1\left(x_1(t), x_2(t)\right) = (1,0,0)\; x_1(t) + (1,0)\; x_2(t), \\ &0 = g_0\left(x_1(t), x_2(t), y(t), u(t)\right) = (0,-1,0)\; x_1(t) - y(t) + u(t). \end{aligned}$$

Assumptions **(A2)** and **(A3)** are satisfied, since $g'_{0,y}\left(x_1, x_2, y, u\right) = -1$ holds for every $(x_1, x_2, y, u) \in \mathbb{R}^3 \times \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R}$ and the matrix

$$E = \begin{pmatrix} \boldsymbol{g_x}\left(\boldsymbol{x}(0)\right) \\ D \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

has full rank. To derive the necessary conditions for this problem we introduce the Hamilton function

$$\begin{aligned} &\mathcal{H}\left(x_1, x_2, y, u, \lambda_{f,1}, \lambda_{f,2}, \lambda_g\right) \\ &\quad = \lambda_{f,1}^\top \begin{pmatrix} u-y \\ u \\ \frac{1}{2}\,u^2 \end{pmatrix} + \lambda_{f,2}^\top \left(\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x_2 \right) \end{aligned}$$

$$+ \lambda_g^\top \left((0,-1,0)\ x_1 - y + u\right).$$

This yields the adjoint DAE

$$\begin{aligned}
\dot{\lambda}_{f,1}(t) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \lambda_{f,2}(t) + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \lambda_g(t), \\
\dot{\lambda}_{f,2}(t) &= \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \lambda_{f,2}(t), \\
0 &= (-1,0,0)\ \lambda_{f,1}(t) - \lambda_g(t),
\end{aligned}$$

the transversality conditions

$$\begin{aligned}
\begin{pmatrix} \lambda_{f,1}(0) \\ \lambda_{f,2}(0) \end{pmatrix} &= -E^\top \begin{pmatrix} \lambda_{0,1} \\ \lambda_{0,2} \end{pmatrix} \\
\lambda_{f,1}(1) &= \ell_0 \left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \hat{x}_1(1) + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \right), \\
\lambda_{f,2}(1) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix},
\end{aligned}$$

and the stationarity of the Hamilton function

$$0 = (1,1,\hat{u}(t))\ \lambda_{f,1}(t) + \lambda_g(t).$$

For $\ell_0 = 1$ we get the solution

$$\begin{aligned}
&\hat{x}_1(t) = \begin{pmatrix} -t^2 + t \\ -2\,t + 1 \\ 2\,t \end{pmatrix}, \quad \hat{x}_2(t) = \begin{pmatrix} t^2 - t \\ 1 \end{pmatrix}, \\
&\hat{y}(t) = 2\,t - 3, \quad \hat{u}(t) = -2, \\
&\hat{\lambda}_{f,1}(t) = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \hat{\lambda}_{f,2}(t) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \hat{\lambda}_g(t) = 0, \\
&\hat{\lambda}_{0,1} = \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}, \quad \hat{\lambda}_{0,2} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.
\end{aligned}$$

Next we want to confirm that this KKT-point is a (local) minimizer for our problem, by proving, that the coercivity condition (4.3) holds. To that end we consider the second derivative of the Lagrange function at the KKT-point

$$\begin{aligned}
&\mathcal{L}''_{\boldsymbol{zz}} \left(\hat{x}_1 \hat{x}_2, \hat{\lambda}_{f,1}, \hat{\lambda}_{f,2}, \hat{\lambda}_g, \hat{\lambda}_{0,1}, \hat{\lambda}_{0,2}\right) (\boldsymbol{z}, \boldsymbol{z}) \\
&\quad = x_1(1)^\top \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_1(1) + \int_0^1 \|u(t)\|^2\ dt
\end{aligned}$$

and the linearized system

$$\begin{aligned}
\dot{x}_1 &= \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} y(t) + \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} u(t), \\
\dot{x}_2 &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} x_1(t) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x_2(t), \\
0 &= (0, -1, 0)\; x_1(t) - y(t) + u(t), \\
x_1(0) &= 0, \\
x_2(0) &= 0.
\end{aligned}$$

Our goal is to find a constant $\gamma > 0$, such that

$$\mathcal{L}''_{zz}\left(\hat{x}_1, \hat{x}_2, \hat{y}, \hat{u}, \hat{\lambda}_{f,1}, \hat{\lambda}_{f,2}, \hat{\lambda}_g, \hat{\lambda}_0\right)(z, z) \geq \gamma\; \|z\|^2_{\hat{Z}}$$

for every $z = (x_1, x_2, y, u)$ satisfying the linearized system with the norm $\|z\|_{\hat{Z}} := \max\{\|x_1\|_{1,2}, \|x_2\|_{2,2}, \|y\|_2, \|u\|_2\}$.
Solving the algebraic equation for $y(t)$ and inserting into the differential equation gives us the system

$$\begin{aligned}
\dot{x}_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_1(t) + \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} u(t), \\
\dot{x}_2 &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} x_1(t) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x_2(t), \\
x_1(0) &= 0, \\
x_2(0) &= 0.
\end{aligned}$$

The differential equation for x_1 has the solution

$$x_1(t) = \int_0^t \begin{pmatrix} t-\tau \\ 1 \\ -2 \end{pmatrix} u(\tau)\, d\tau$$

and the norm satisfies

$$\begin{aligned}
\|x_1\|_2^2 &= \int_0^1 \left\| \int_0^t \begin{pmatrix} t-\tau \\ 1 \\ -2 \end{pmatrix} u(\tau)\, d\tau \right\|^2 dt \\
&\leq \int_0^1 \int_0^t \left((t-\tau)^2 + 5\right) \|u(\tau)\|^2 \, d\tau\, dt \\
&\leq 6\, \|u\|_2^2.
\end{aligned}$$

For $y(t) = (0, -1, 0)\; x_1(t) + u(t)$ we get the estimate

$$\begin{aligned}
\|y\|_2^2 &\leq \left(\|\,(0, -1, 0)\; x_1\|_2 + \|u\|_2\right)^2 \\
&\leq \left(\|x_1\|_2 + \|u\|_2\right)^2 \\
&\leq 12\, \|u\|_2^2.
\end{aligned}$$

Using the original equation for $\dot{x}_1$ yields

$$\begin{aligned}\|\dot{x}_1\|_2^2 &\le \left(\left\|\begin{pmatrix}-1\\0\\0\end{pmatrix} y\right\|_2 + \left\|\begin{pmatrix}1\\1\\-2\end{pmatrix} u\right\|_2\right)^2\\ &\le \left(\|y\|_2 + \sqrt{6}\,\|u\|_2\right)^2\\ &\le 36\,\|u\|_2^2.\end{aligned}$$

Now we consider the differential equation for x_2, namely

$$\dot{x}_2(t) = \begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix} x_1(t) + \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} x_2(t),$$

which has the solution

$$\begin{aligned}x_2(t) &= \int_0^t \begin{pmatrix}1 & 0\\ t-\tau & 1\end{pmatrix}\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix} x_1(\tau)\,d\tau\\ &= \int_0^t \int_0^\tau \begin{pmatrix}-1\\ 2\,\tau - t - s\end{pmatrix} u(s)\,ds\,d\tau.\end{aligned}$$

We get the estimate

$$\begin{aligned}\|x_2\|_2^2 &= \int_0^1 \left\|\int_0^t \int_0^\tau \begin{pmatrix}-1\\ 2\,\tau - t - s\end{pmatrix} u(s)\,ds\,d\tau\right\|^2 dt\\ &\le 9\,\|u\|_2^2.\end{aligned}$$

Furthermore we have

$$\begin{aligned}\|\dot{x}_2\|_2^2 &\le \left(\left\|\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix} x_1\right\|_2 + \left\|\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} x_2\right\|_2\right)^2\\ &\le \left(\|x_1\|_2 + \|x_2\|_2\right)^2\\ &\le 30\,\|u\|_2^2.\end{aligned}$$

The second derivative of x_2 satisfies the equation

$$\begin{aligned}\ddot{x}_2(t) &= \begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix} \dot{x}_1(t) + \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} \dot{x}_2(t)\\ &= \begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\end{pmatrix} x_1(t) + \begin{pmatrix}0\\ -1\end{pmatrix} y(t) + \begin{pmatrix}-1\\ 1\end{pmatrix} u(t)\end{aligned}$$

so we get

$$\begin{aligned}\|\ddot{x}_2\|_2^2 &\le \left(\left\|\begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\end{pmatrix} x_1\right\| + \left\|\begin{pmatrix}0\\ -1\end{pmatrix} y\right\| + \left\|\begin{pmatrix}-1\\ 1\end{pmatrix} u\right\|\right)^2\\ &\le \left(\|x_1\|_2 + \|y\|_2 + \sqrt{2}\,\|u\|_2\right)^2\\ &\le 54\,\|u\|_2^2.\end{aligned}$$

Summing up we have $\|z\|_{\hat{Z}}^2 \leq 54 \|u\|_2^2$, so $\|u\|_2^2 \geq \frac{1}{54} \|z\|_{\hat{Z}}^2$. Finally we get a lower bound for the second derivative of the Lagrange function, namely

$$\begin{aligned}
&\mathcal{L}''_{zz}\left(\hat{x}_1, \hat{x}_2, \hat{y}, \hat{u}, \hat{\lambda}_{f,1}, \hat{\lambda}_{f,2}, \hat{\lambda}_g, \hat{\lambda}_0\right)(z, z) \\
&= \|(0,1,0)\ x_1(1)\|^2 + \int_0^1 \|u(t)\|^2\, dt \\
&\geq \|u\|_2^2 \geq \frac{1}{54} \|z\|_{\hat{Z}}^2.
\end{aligned}$$

6. Conclusions

The paper establishes first order necessary optimality conditions in terms of a local minimum principle and sufficient conditions for a class of nonlinear DAE optimal control problems subject to mixed control-state constraints and Hessenberg DAEs of arbitrary index. Such DAEs typically occur in mechanical engineering and path planning problems. The application of the presented results to such problems will be the subject of future research. To this end, it would be important to extend the results to problems with pure state constraints. Likewise, a generalization to more general DAEs or even arbitrarily structured DAEs would be desirable.

References

[1] A. Backes, *Extremalbedingungen für Optimierungs-Probleme mit Algebro-Differentialgleichungen*, PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät, Humboldt-Universität Berlin, Berlin, Germany, 2006.

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

[3] F. Clarke, Y. Ledyaev and M. d. R. de Pinho,*An extension of the Schwarzkopf multiplier rule in optimal control*, SIAM J. Control Optim. **49** (2011), 599–610.

[4] A. V. Dmitruk, *Maximum principle for the general optimal control problem with phase and regular mixed constraints*, Comput. Math. Model. **4** (1993), 364–377.

[5] A. L. Dontchev, W. W. Hager, A. B. Poore and Y. Bing, *Optimality, stability, and convergence in nonlinear control*, Appl. Math. Optim. **31** (1995), 297–326.

[6] L. C. Evans, *Partial Differential Equations*, Second Edition, Graduate Studies in Mathematics, vol.19, American Mathematical Society, 2010.

[7] M. Gerdts, *Local Minimum principle for optimal control problems subject to index-two differential-algebraic equations*, J. Optim. Theory Appl. **130** (2006), 443–462.

[8] M. Gerdts, Optimal control of ODEs and DAEs. Walter de Gruyter, Berlin/Boston, 2012.

[9] M. Gerdts, *A survey on optimal control problems with differential-algebraic equations*, in: Surveys in Differential-Algebraic Equations II. Differential-Algebraic Equations Forum, A. Ilchmann, T. Reis T. (eds), Springer, Cham, 2015, pp. 103–161.

[10] W. W. Hager, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim. **17** (1979), 321–338.

[11] D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory I. Modelling, State Space Analysis, Stability and Robustness*, Text in Applied Mathematics, vol.48: Springer Verlag, Berlin 2005.

[12] P. Kunkel and V. Mehrmann, *The linear quadratic optimal control problem for linear descriptor systems with variable coefficients*, Math. Control, Signals, Systems **10** (1997), 247–264.

[13] P. Kunkel and V. Mehrmann, *Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index*, Math. Control, Signals, Systems **20** (2008), 227–269.

[14] G. A. Kurina and R. März, *On linear-quadratic optimal control problems for time-varying descriptor systems*, SIAM J. Control Optim. **42** (2004), 2062–2077.

[15] L. A. Ljusternik and W. I. Sobolew, *Elemente Der Funktionalanalysis, Fünfte Auflage*, Übersetzung der zweiten russischen Auflage von Klaus Fiedler und herausgegeben von Konrad Gröger. Mathematische Lehrbücher und Monographien, I. Ableitung: Mathematische Lehrbücher, Band VIII. Akademie-Verlag, Berlin, 1976.

[16] K. Malanowski, *Second-order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimization problems in Hilbert spaces*, Appl. Math. Optim. **25** (1992), 51–79.

[17] K. Malanowski and H. Maurer, *Sensitivity analysis for parametric control problems with control-state constraints*, Comp. Optim. Appl. **5** (1996), 253–283.

[18] K. Malanowski, *On normality of Lagrange multipliers for state constrained optimal control problems*, Optimization **52** (2003), 75–91.

[19] K. Malanowski, H. Maurer and S. Pickenhain, *Second-order sufficient conditions for state-constrained optimal control problems*, J. Optim. Theory Appl. **123** (2004), 595–617.

[20] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Program. **16** (1979), 98–110.

[21] H. Maurer, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Math. Program. Study **14** (1981), 163–177.

[22] H. Maurer and S. Pickenhain, *Second order sufficient conditions for control problems with mixed control-state constraints*, J. Optim. Theory Appl. **86** (1995), 649–667.

[23] V. Mehrmann, *Existence, uniqueness, and stability of solutions to singular linear quadratic optimal control problems*, Linear Alg. Appl. **121** (1989), 291–331.

[24] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem. Theory and Numerical Solution*, Lecture Notes in Control and Information Sciences, vol. 163, Springer: Berlin, 1991.

[25] M. d. R. de Pinho and R. B. Vinter, *Necessary conditions for optimal control problems involving nonlinear differential algebraic equations*, J. Math. Anal. Appl. **21** (1997), 493–516.

[26] W. T. Reid, *Riccati Differential Equations*, Mathematics in Science and Engineering, vol. 86, Academic Press: New York, 1972.

[27] S. M. Robinson, *Stability theory for systems of inequalities, part II: differentiable nonlinear systems*, SIAM J. Numer. Anal. **13** (1976), 487–513.

[28] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, part III:stability and sensitivity*, Math. Program. Study **30** (1987), 45–66.

[29] T. Roubicek and M. Valasek, *Optimal control of causal differential-algebraic systems*, J. Math. Anal. Appl. **269** (2002), 616–641.

[30] W. Walter, *Gewöhnliche Differentialgleichungen*, Heidelberger Taschenbücher 110: Springer Heidelberg, 1976.

[31] V. Zeidan, *First and second-order sufficient for optimal control and the calculus of variations*, Appl. Math. Optim. **11** (1984), 209–226.

Manuscript received December 8 2017
revised January 25 2018

B. Martens
Institute of Mathematics and Applied Computing, Department of Aerospace Engineering, University of the Federal Armed Forces at Munich, Werner-Heisenberg-Weg 39, 85577 Neubiberg/München, Germany
E-mail address: `bjoern.martens@unibw.de`

M. Gerdts
Institute of Mathematics and Applied Computing, Department of Aerospace Engineering, University of the Federal Armed Forces at Munich, Werner-Heisenberg-Weg 39, 85577 Neubiberg/München, Germany
E-mail address: `matthias.gerdts@unibw.de`