TECHNISCHE UNIVERSITÄT BERLIN

Differential-algebraic Riccati Decoupling for Linear-quadratic Optimal Control Problems for Semi-explicit Index-2 DAEs

Jan Heiland

Preprint 2012/06

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin
http://www.math.tu-berlin.de/preprints

Preprint 2012/06 March 2012

Abstract

We investigate existence and structure of solutions to quadratic control problems with semi-explicit differential algebraic constraints. By means of an equivalent index-1 formulation we identify conditions for the unique existence of optimal solutions. Knowing of the existence of an optimal input we provide a representation of the associated feedback-law via a Riccati-like decoupling that is formulated for the original index-2 equations.

Keywords Optimal Control, DAEs, Differential Riccati Equation, Euler-Lagrange equations
AMS subject classification 34H05, 49J15

1 Introduction

A well-posed formulation of a control system for differential-algebraic equations (DAEs) has to relate the algebraic constraints to the controls, since it may happen that some components of the input are not free. As can be seen from the simple example

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad x_1(0) = 0, \tag{1}$$

an arbitrarily defined input can make the model inconsistent, as in (1) the existence of solutions is only possible if u_2 is either zero or a function of the state x_1 or x_2. On the other hand the definition of the input can change the characteristics of the model. With the assignment $u_2 := \dot{x}_2$, System (1) can be interpreted as an ODE for x_1 and x_2. Assigning $u_2 := x_1$ the system can be reformulated to give only algebraic equations for x_1 and x_2. A general approach to this issue bases on the behavior formulation, cf. [23], which considers the control problem as an underdetermined system in the augmented variable $z := [x, u]$. For the behavior system one can identify free components of z that are then defined as the new controls, cf. [18]. This approach is natural, since if the chosen controls are not free variables in the behavior formulation, then the problem is ill-posed.

If it comes to applications, however, one often cannot freely redefine the controls and variables since they are prescribed by the physical setup. In this case one can only hope that the problem is well-posed and if necessary consider a remodeling that ensures well-posedness in terms of the original inputs. We will investigate optimal control problems, where the control acts only in the differential part of the differential-algebraic equations such that the inputs will always be free variables.

In optimal control one tackles the problem of determining an input u such that

$$\mathcal{J}(x, u) \to \min \text{ subject to } F(t, \dot{x}, x, u) = 0,$$

where x is the state of the system, $\mathcal{J}$ is a cost functional and F stands for the constraints given by the state equations. For the solution of optimal control problems there are basically two approaches, the value function and the variational approach, cf. [26].

In order to extend these approaches that are well-understood for ODE constrained optimization problems to DAE constraints, one has cope with the *strangeness* of the DAEs. The strangeness describes to which extent differential and algebraic equations are tied together and is quantified by means of various index concepts [27, Ch. 1]. Generally spoken, in strangeness-free or index-1 DAEs, algebraic and differential equations are well separated and higher indices mean higher interlocking.

As in the ODE case the value function approach leads to a feedback-law via the solution of the corresponding Hamilton-Jacobi-Bellman (HJB) equations. These equations were formulated and investigated for nonlinear strangeness-free DAEs in [25], by expanding the DAEs into power series and solving the linear part. With the assumption of impulsive controllability HJB-like equations for DAEs were formulated in [29].

If one uses a variational or Lagrange-multiplier approach one ends up with variants of the so-called Euler-Lagrange equations. The structure of these equations suggests that the solution to optimal control problems with DAE constraints is a function of the state, i.e. a feedback control. Unlike the ODE case, for DAE constraints the existence and uniqueness of solutions to the involved adjoint equations and thus to the optimality system is in general not guaranteed, cf. [4, 6, 10, 21] and in particular [16] for the linear quadratic case. To provide necessary and sufficient conditions for the existence of optimal controls one can for example exploit the special structure of semi-explicit equations, cf. [10–12], or consider linear DAEs with properly stated leading term, cf. [3–5, 7, 21]. The special case of Riccati-feedback solutions was investigated in [22]. The more general way to regularize the DAEs and formulate the conditions for the resulting strangeness-free system is taken in [19]. In [17, 18] conditions and procedures for the construction of state feedbacks are presented such that the behavior system is strangeness-free.

The formulations of the optimality conditions in the references listed above to the Lagrange approach to DAE constrained optimal control involve index reduction procedures, except from the contributions in [3, 4]. However, from the numerical point of view, a formulation in the unreduced equations is preferable in some respects. First of all, the original equations contain those constraints that are of importance for the computed solutions rather than the so called hidden constraints that appear in index reduction procedures. Second, most index reduction techniques use projectors or implicit function which may be expensive to compute. Thus for an efficient implementation an index reduction, if necessary, should be tailored to the solution of the specific problem rather than to the derivation of theoretical results.

In our approach we make use of the structure of an optimality system that is stated in the original variables without any explicit or implicit index reduction. The obtained results regarding optimality of the solutions to these Euler-Lagrange equations are basically the same or even less general than those obtained in [3, 4]. However, we use the special structure to prove the existence of a solution and its representation via a matrix differential-algebraic Riccati equation. The specific structure and the specific Riccati ansatz overcomes the peculiarities that were found and illustrated in [16]. Also the extension to particular nonlinearities as e.g. in the Navier-Stokes equations is already set up in the basic results.

This manuscript is structured as follows. In Section 2 we explain what is meant by optimal control of semi-explicit DAEs. Section 3 introduces the special type of optimization problems that is considered and states basic results that are useful to formulate necessary conditions for consistency and regularity of candidate solutions. Section 4 contains the main results regarding existence of solutions to the linear problems with appropriately chosen quadratic costs. There we also show that an optimal input is given via a feedback-law which can be explicitly computed via the solution of a differential-algebraic matrix Riccati equation. The last Section 7 points out the basic differences of our results to the results in [24] and discusses a possible extension to control constraint problems.

2 Optimal Control of Semi-explicit DAEs

We consider the optimal control problem of finding an input u such that the cost functional

$$\mathcal{J}(v,p,u) = \mathcal{M}(v(T)) + \int_0^T \mathcal{K}(v,p,u)\ \mathrm{d}t. \tag{2}$$

is minimal, where v and p are solutions of DAEs of the form

$$M(t)\dot{v} - f(t,v,p,u) = 0, \quad v(0) = v^0, \tag{3a}$$
$$g(t,v,p,u) = 0, \tag{3b}$$

formulated on a time interval $\mathbb{I} = (0,T]$. We will always assume that $M(t)$ is invertible, so that the system can be formally brought into a form, in which the time derivative of the differential variables is explicitly given.

The associated formally derived Euler-Lagrange equations, cf. [19], are then given by

$$M\dot{v} - f(t,v,p,u) = 0, \quad v(0) = v^0, \tag{4a}$$
$$g(t,v,p,u) = 0, \tag{4b}$$
$$-\tfrac{d}{dt}M^T\lambda_1 - f_v^T\lambda_1 + g_v^T\lambda_2 + \mathcal{K}_v^T = 0, \quad M^T\lambda_1(T) = -\mathcal{M}_v^T(v(T)), \tag{4c}$$
$$-f_p^T\lambda_1 + g_p^T\lambda_2 + \mathcal{K}_p^T = 0, \quad \mathcal{M}_p = 0, \tag{4d}$$
$$\mathcal{K}_u - f_u^T\lambda_1 + g_u^T\lambda_2 = 0, \tag{4e}$$

where the arguments in the functions have been dropped. Note that the Euler-Lagrange equations are only formally stated, since for the derivation one already assumes the existence of the Lagrange-multipliers λ_1 and λ_2.

3 Semi-explicit Semi-linear DAEs of Index 2

In this section we consider a special case of System (3) of the form:

$$M(t)\dot{v} - A(t,v)v - J_1(t)^T p - B_1(t)u = f_v, \quad v(0) = v^0 \in \mathbb{R}^{n_v}, \tag{5a}$$
$$-J_2(t)v - B_2(t)u = f_p. \tag{5b}$$

For several semi-explicit systems of the form (5) the differentiation index, cf. [9], is defined as follows:

Definition 3.1. A semi-explicit DAE as given by (3) is of differentiation index k if it takes $k-1$ differentiations in t of the algebraic constraints (3b) to determine the algebraic variable p in terms of the differential variable v.

In order to guarantee existence of solutions $(v,p) \in \mathcal{C}^1(\mathbb{I},\mathbb{R}^{n_v}) \times \mathcal{C}(\mathbb{I},\mathbb{R}^{n_p})$ of (5), we make the following assumption:

Assumption 3.2. For the DAE (5) with coefficients $M, A(\cdot,v) \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_v,n_v})$, $J_1, J_2 \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_v,n_p})$, $B_1 \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_u,n_v})$, $B_2 \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_u,n_u})$, right-hand sides $f_v \in (\mathbb{I},\mathbb{R}^{n_v})$, $f_p \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_p})$, an initial condition $v_0 \in \mathbb{R}^{n_v}$ and inputs $u \in \mathcal{C}(\mathbb{I},\mathbb{R}^{n_u})$ we assume

(A1) differentiation index 2, i.e.
$S := J_2 M^{-1} J_1^T$ is invertible,

(A2) sufficient regularity of the data and the input, i.e.
f_p, B_2u, $M^{-1}J_1^TS^{-1}$ and J_2 are differentiable and

(A3) consistency of the data and the input, i.e.
$J_2(0)v(0) = f_p(0) - B_2(0)$.

We will not investigate existence of solutions here. However, the following Theorem 3.3 gives a solution representation by means of the inherent ODE that will be used to ensure existence and uniqueness in the linear case of System (5).

Theorem 3.3. *Each solution (v,p) of (5) can be represented as $(v_0 + Lv, p)$, where*

$$Lv = -M^{-1}J_1^TS^{-1}[B_2u + f_p], \tag{6a}$$

$$p = -L^-[M^{-1}A(Lv + v_0)[Lv + v_0] + M^{-1}B_1u + M^{-1}f_v] + L^-[\dot{v_0} + \dot{Lv}], \tag{6b}$$

and $v_0 := [I - L]v$ solves the ODE

$$\begin{aligned} \dot{v}_0 - \left[\tfrac{d}{dt}(I-L) + [I-L]M^{-1}A(Lv+v_0)\right][Lv+v_0] &= [I-L][M^{-1}B_1u + M^{-1}f_v], \\ v_0(0) &= [I-L]v^0. \end{aligned} \tag{6c}$$

with $L := M^{-1}J_1^TS^{-1}J_2$ and $L^- := S^{-1}J_2$.

Proof of Theorem 3.3. Rewriting and abbreviating (5) by

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} - \begin{bmatrix} M^{-1}A(v) & M^{-1}J_1^T \\ J_2 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} = \begin{bmatrix} M^{-1}(B_1u + f_v) \\ B_2u + f_p \end{bmatrix} \quad \text{and} \quad \mathcal{E}\dot{x} - \mathcal{A}(x)x = q, \tag{7}$$

respectively, with $x := (v,p)$, we compute the operator chain as described e.g. in [7] but with a slightly different notation. It will turn out that the nonlinear part in $\mathcal{A}$ does not interfere with the definitions of the projections and subspaces such that the linear theory of [7] is indeed applicable here. The operator chain is given by

$$\mathcal{E}_0 := \mathcal{E} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{A}_0 := \mathcal{A} = \begin{bmatrix} M^{-1}A(v) & M^{-1}J_1^T \\ J_2 & 0 \end{bmatrix}, \tag{8a}$$

$$\mathscr{Q}_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \text{ (projector onto } \ker \mathcal{E}_0), \quad \mathscr{P}_0 = I - \mathscr{Q}_0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \tag{8b}$$

$$\mathcal{E}_1 = \mathcal{E}_0 + \mathcal{A}_0\mathscr{Q}_0 = \begin{bmatrix} I & M^{-1}J_1^T \\ 0 & 0 \end{bmatrix}, \quad \mathcal{A}_1 = \mathcal{A}_0\mathscr{P}_0 = \begin{bmatrix} M^{-1}A(v) & 0 \\ J_2 & 0 \end{bmatrix}, \tag{8c}$$

$$\mathscr{Q}_1 = \begin{bmatrix} M^{-1}J_1^T(J_2M^{-1}J_1^T)^{-1}J_2 & 0 \\ -(J_2M^{-1}J_1^T)^{-1}J_2 & 0 \end{bmatrix} =: \begin{bmatrix} L & 0 \\ -L^- & 0 \end{bmatrix}, \quad \mathscr{P}_1 = \begin{bmatrix} I-L & 0 \\ L^- & I \end{bmatrix}. \tag{8d}$$

$$\mathscr{Q}_1 = \begin{bmatrix} M^{-1}J_1^TS^{-1}J_2 & 0 \\ -S^{-1}J_2 & 0 \end{bmatrix} = \begin{bmatrix} L & 0 \\ -L^- & 0 \end{bmatrix}, \quad \mathscr{P}_1 = \begin{bmatrix} I-L & 0 \\ L^- & I \end{bmatrix}, \tag{8e}$$

$$\mathcal{E}_2 = \begin{bmatrix} I + M^{-1}AL & M^{-1}J_1^T \\ J_2 & 0 \end{bmatrix}. \tag{8f}$$

With the projector $L = M^{-1}J_1^TS^{-1}J_2$ which satisfies

$$L^2 = L, \quad J_2L = J_2, \quad LM^{-1}J_1^T = M^{-1}J_1^T \quad \text{and} \quad L^-L = L^-,$$

one can verify that

$$\mathcal{E}_2^{-1} = \begin{bmatrix} I-L & [I-[I-L]M^{-1}A]M^{-1}J_1^TS^{-1} \\ L^- & -[I+L^-M^{-1}AM^{-1}J_1^T]S^{-1} \end{bmatrix} \tag{9}$$

for any $A(v)$. Having scaled the state equations (7) by $\mathcal{E}_2^{-1}$ we obtain the equivalent system

$$\begin{bmatrix} I-L & 0 \\ L^- & 0 \end{bmatrix} \begin{bmatrix} \dot v \\ \dot p \end{bmatrix} - \Big[\begin{bmatrix} [I-L]M^{-1}A[I-L] & 0 \\ L^-M^{-1}A[I-L] & 0 \end{bmatrix} + \begin{bmatrix} L & 0 \\ -L^- & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \Big] \begin{bmatrix} v \\ p \end{bmatrix} = \mathcal{E}_2^{-1} \begin{bmatrix} M^{-1}B_1u + M^{-1}f_v \\ B_2u + f_p \end{bmatrix}. \tag{10}$$

Having applied the projectors $\mathscr{Q}_1$, $\mathscr{Q}_0\mathscr{P}_1$ and $\mathscr{P}_0\mathscr{P}_1$ from (8) to (10) we obtain the three subsystems

$$-\begin{bmatrix} L & 0 \\ -L^- & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} = \mathscr{Q}_1\mathcal{E}_2^{-1}q = \begin{bmatrix} M^{-1}J_1^TS^{-1}[B_2u+f_p] \\ S^{-1}[B_2u+f_p] \end{bmatrix}, \tag{11a}$$

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ L^- & 0 \end{bmatrix} \begin{bmatrix} \dot v \\ \dot p \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ L^-M^{-1}A[I-L] & I \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} &= \mathscr{Q}_0\mathscr{P}_1\mathcal{E}_2^{-1}q \\ &= \begin{bmatrix} 0 \\ L^-[M^{-1}B_1u + M^{-1}f_v - M^{-1}AM^{-1}J_1^TS^{-1}[B_2u+f_p]] \end{bmatrix} \end{aligned} \tag{11b}$$

and

$$\begin{aligned} \begin{bmatrix} I-L & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot v \\ \dot p \end{bmatrix} - \begin{bmatrix} [I-L]M^{-1}A[I-L] & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} &= \mathscr{P}_0\mathscr{P}_1\mathcal{E}_2^{-1}q \\ &= \begin{bmatrix} [I-L][M^{-1}B_1u + M^{-1}f_v - M^{-1}AM^{-1}J_1^TS^{-1}[B_2u+f_p]] \\ 0 \end{bmatrix}, \end{aligned} \tag{11c}$$

respectively. Since $\mathscr{Q}_1 + \mathscr{P}_0\mathscr{P}_1 + \mathscr{Q}_0\mathscr{P}_1 = I$, equations (11) contain all information of (10) and vice versa. We decompose $v = v_0 + Lv$, where $v_0 := [I-L]v$ so that from (11a) we can deduce that

$$Lv = -M^{-1}J_1^TS^{-1}[B_2u+f_p] \tag{12}$$

and that Lv is differentiable by assumption. With $\dot v = \dot{Lv} + \dot v_0$, equation (11b) gives

$$p = -L^-[M^{-1}A(Lv+v_0)[Lv+v_0] + M^{-1}B_1u + M^{-1}f_v] + L^-[\dot v_0 + \dot{Lv}], \tag{13}$$

while (11c) defines the inherent ODE for $v_0 := [I-L]v$ via

$$\dot v_0 - \big[\tfrac{d}{dt}(I-L) + [I-L]M^{-1}A(Lv+v_0)\big][Lv+v_0] = [I-L][M^{-1}B_1u + M^{-1}f_v], \quad v_0(0) = [I-L]v^0. \tag{14}$$

Note the necessity of the consistency condition (A3) in Assumption 3.2, since by (12) the condition

$$J_2v(0) = J_2[Lv(0) + [I-L]v(0)] = J_2Lv(0) = -B_2u(0) - f_p(0),$$

must hold and note, that an initial condition for p would have to fulfill (13) at $t = 0$. □

Remark 3.4. In the setting of the Navier-Stokes equations, the projector L realizes the discrete Helmholtz-decomposition that splits a vector field into a divergence free part and a part that can be expressed as the gradient of a scalar potential, cf. [13, Cor. 3.4]. If J_2 is the discrete divergence operator then the decomposition $v = Lv + [I-L]v =: Lv + v_0$ delivers that $J_2v_0 = 0$ and Lv is in the range of $M^{-1}J_1^T$, which is the discrete gradient operator in many discretization schemes. The matrix L^- is a generalized left inverse of $M^{-1}J_1^T$ and can be seen as the operator that maps the potential field $Lv = M^{-1}J_1^T\rho$ onto its potential ρ. Accordingly, (6b) is the discrete Pressure-Poisson equation, cf. [14].

Corollary 3.5. *If $B_2 = 0$, then the solutions of* (5) *do not depend on the time derivative of the input. The condition $B_2 = 0$ is also necessary for the existence of solutions for all smooth but not differentiable inputs.*

The first fact of Corollary 3.5 follows from the representation of the solution as given in Theorem 3.3. For the converse direction, one concludes that the solution component $\frac{d}{dt}(B_2 u)$ can only exist for all smooth u if $B_2 = 0$.

For the results of the next sections we will always require $B_2 = 0$ which by Corollary 3.5 is necessary and sufficient for the admissibility of inputs that are not differentiable. This is what in [3, 4] and [24] is also assumed and referred to as *causality*.

4 Linear Quadratic Optimal Control

We start with investigating a linearized version of (5), i.e. $A(t, v) = A(t)$, and a quadratic cost functional of type

$$\mathcal{J}(v,p,u) = \frac{1}{2}\begin{bmatrix} v - v^* \\ p - p^* \end{bmatrix}^T \begin{bmatrix} V_1 & V_{12} \\ V_{21} & V_2 \end{bmatrix} \begin{bmatrix} v - v^* \\ p - p^* \end{bmatrix}\Bigg|_{t=T} + \\ + \frac{1}{2}\int_0^T \begin{bmatrix} v - v^* \\ p - p^* \\ u \end{bmatrix}^T \begin{bmatrix} W_1 & W_{12} & S_{vu} \\ W_{21} & W_2 & S_{pu} \\ S_{uv} & S_{up} & R \end{bmatrix} \begin{bmatrix} v - v^* \\ p - p^* \\ u \end{bmatrix} \, \mathrm{d}t,$$

with symmetric positive semi-definite weighting matrices $\begin{bmatrix} V_1 & V_{12} \\ V_{21} & V_2 \end{bmatrix}$ and $\begin{bmatrix} W_1 & W_{12} & S_{vu} \\ W_{21} & W_2 & S_{pu} \\ S_{uv} & S_{up} & R \end{bmatrix}$, that is appropriate for driving the system into a desired state (v^*, p^*). The stated results will then give the basis for the treatment of the semi-linear problems. In this setting the formal Euler-Lagrange equations (4e) are given by

$$M\dot{v} - Av - J_1^T p - B_1 u = f_p, \quad v(0) = v^0 \tag{15a}$$

$$-J_2 v - B_2 u = f_g \tag{15b}$$

$$-\tfrac{d}{dt}(M^T\lambda_1) - A^T\lambda_1 - J_2^T\lambda_2 + W_1(v - v^*) + W_{12}(p - p^*) + S_{vu}u = 0, \tag{15c}$$

$$M^T\lambda_1(T) = -V_1(v - v^*)\Big|_{t=T} - V_{12}(p - p^*)\Big|_{t=T},$$

$$-J_1\lambda_1 + W_{21}(v - v^*) + W_2(p - p^*) + S_{pu}u = 0, \tag{15d}$$

$$0 = V_{21}(v - v^*)\Big|_{t=T} + V_2(p - p^*)\Big|_{t=T},$$

$$-B_1^T\lambda_1 - B_2^T\lambda_2 + S_{uv}(v - v^*) + S_{up}(p - p^*) + Ru = 0. \tag{15e}$$

If there exists a solution, then system (15) provides necessary and sufficient conditions for an optimal input u, cf. [5]. Thus, the task is now to identify under which conditions the Euler-Lagrange equations possess a unique solution.

Since we consider state solutions $(v, p) \in \mathcal{C}^1 \times \mathcal{C}$ and inputs $u \in \mathcal{C}$ candidate solutions of (15) must not contain $\dot{u}$ or $\dot{p}$. Thus, by Corollary 3.5 it is necessary that

$$B_2 = 0, \quad W_2 = 0 \quad \text{and} \quad S_{pu} = S_{up}^T = 0. \tag{16}$$

The other possibility that solutions of DAEs fail to exist, is the inconsistency of the initial data. For the state equations (15a), consistency is equivalent to $-J_2 v^0 = f_g(0)$. Since the terminal condition in (15e) must hold for any trajectories of v and p, we conclude that $V_2 = 0$ and $V_{21} = V_{12}^T = 0$ is necessary. Furthermore, by combining (15d) and the terminal condition for λ_1, we find that V_1 and W_{21} must be such that

$$-J_1\lambda_1(T) = -W_{21}(v(T) - v^*(T)) = J_1 M^{-T} V_1(v(T) - v^*(T)).$$

We will ensure this condition by requiring

$$J_1 M^{-T} V_1 = 0, \quad \text{and} \quad W_{21} = W_{12}^T = 0. \tag{17}$$

The latter condition means that V_1 acts only on the dynamical part of v as it is given by (6c). Note that these conditions are equivalent to the assumptions that were made in [3]. The next Lemma will show that the assumptions for smooth solutions (16), that were derived for the single equations (15a-b) and (15c-d), are necessary also for the coupled system. In particular, we will confirm that if $B_2 = 0$, then there is no hope for a lower index of the optimality system, that may weaken the regularity conditions.

Lemma 4.1. *The Euler-Lagrange equation system* (15) *is of differentiation index* $\nu_d = 1$ *if and only if*

$$\begin{bmatrix} B_2R^{-1}B_2^T & -B_2R^{-1}S_{up} \\ -S_{pu}R^{-1}B_2^T & -W_2 + S_{pu}R^{-1}S_{up} \end{bmatrix} \text{ is invertible.}$$

In particular, if B_2 *does not have full row rank, then* (15) *has differentiation index* $\nu_d \geq 2$.

Assuming (16)*, then* (15) *is a semi-explicit DAE of differentiation index* $\nu_d = 2$.

Proof of Lemma 4.1.

Remark 4.2. In the following derivations we will make use of the time derivative of M. This will make the notation clearer but is by no means necessary for the existence of smooth multipliers. In fact, $M \in \mathcal{C}$ is enough, since one can scale the state equations by M^{-1} or use the transformed multiplier $M^T\lambda_1$.

We use the invertibility of R to express u via

$$u = R^{-1}(B_1^T\lambda_1 + B_2^T\lambda_2 - S_{uv}(v - v^*) - S_{up}(p - p^*))$$

and write (15) in matrix vector form

$$\begin{bmatrix} 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 \\ -M^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot\lambda_1 \\ \dot\lambda_2 \\ \dot v \\ \dot p \end{bmatrix} -$$
$$- \begin{bmatrix} B_1R^{-1}B_1^T & B_1R^{-1}B_2^T & A - B_1R^{-1}S_{uv} & J_1^T - B_1R^{-1}S_{up} \\ B_2R^{-1}B_1^T & B_2R^{-1}B_2^T & J_2 - B_2R^{-1}S_{uv} & -B_2R^{-1}S_{up} \\ A^T + \dot M^T - S_{vu}R^{-1}B_1^T & J_2^T - S_{vu}R^{-1}B_2^T & -W_1 + S_{vu}R^{-1}S_{uv} & -W_{12} + S_{uv}R^{-1}S_{up} \\ J_1 - S_{pu}R^{-1}B_1^T & -S_{pu}R^{-1}B_2^T & -W_{21} + S_{pu}R^{-1}S_{uv} & -W_2 + S_{up}R^{-1}S_{up} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix}$$
$$= \begin{bmatrix} f_v + B_1R^{-1}(S_{uv}v^* + S_{up}p^*) \\ f_p + B_2(S_{uv}v^* + S_{up}p^*) \\ W_1v^* + W_{12}p^* - S_{vu}R^{-1}(S_{uv}v^* + S_{up}p^*) \\ W_2p^* + W_{21}v^* - S_{pu}R^{-1}(S_{uv}v^* + S_{up}p^*) \end{bmatrix} \tag{18}$$

If now the submatrix $\mathscr{H}_{22} := \begin{bmatrix} B_2R^{-1}B_2^T & -B_2R^{-1}S_{up} \\ -S_{pu}R^{-1}B_2^T & -W_2 + S_{pu}R^{-1}S_{up} \end{bmatrix}$ is invertible, then one can solve algebraically for λ_2 and p in (18) and would end up with an boundary value problem for λ_1 and v, which is the characterization of a DAE of differentiation index $\nu_d = 1$. If, however, B_2 does not have full row rank, then $\mathscr{H}_{22}$ is singular and (18) has index $\nu_d \geq 2$.

Assuming now that (16) holds, the corresponding terms in (18) vanish, i.e.

$$\begin{bmatrix} 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 \\ -M^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot\lambda_1 \\ \dot\lambda_2 \\ \dot v \\ \dot p \end{bmatrix} - \begin{bmatrix} B_1R^{-1}B_1^T & 0 & A - B_1R^{-1}S_{uv} & J_1^T \\ 0 & 0 & J_2 & 0 \\ A^T + \dot M^T - S_{vu}R^{-1}B_1^T & J_2^T & -W_1 + S_{vu}R^{-1}S_{uv} & -W_{12} \\ J_1 & 0 & -W_{21} & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix}$$
$$= \begin{bmatrix} f_v + B_1R^{-1}S_{uv}v^* \\ f_p \\ W_1v^* + W_{12}p^* - S_{vu}R^{-1}S_{uv}v^* \\ W_{21}v^* \end{bmatrix}. \tag{19}$$

By inverting the mass matrices and permuting the rows and the columns, System (19) can be brought into the form of (5). Then the differentiation index 2 property, cf. Definition 3.1 follows from

$$\begin{bmatrix} 0 & J_2 \\ J_1 & -W_{21} \end{bmatrix} \begin{bmatrix} 0 & M \\ -M^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 & J_1^T \\ J_2^T & -W_{12} \end{bmatrix} = \begin{bmatrix} 0 & J_2M^{-1}J_1^T \\ -J_1M^{-T}J_2^T & -W_{21}M^{-1}J_1^T - J_1M^{-T}W_{12} \end{bmatrix}$$

is invertible by the assumptions on (5). □

Assuming further that $W_{21} = 0$, cf. (17), we can write the system as

$$\begin{bmatrix} 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 \\ -M^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot\lambda_1 \\ \dot\lambda_2 \\ \dot v \\ \dot p \end{bmatrix} - \begin{bmatrix} G & 0 & F & J_1^T \\ 0 & 0 & J_2 & 0 \\ F^T + \dot M^T & J_2^T & H & 0 \\ J_1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix} = \begin{bmatrix} f_v + B_1R^{-1}S_{uv}v^* \\ f_p \\ W_1v^* - S_{vu}R^{-1}S_{uv}v^* \\ 0 \end{bmatrix}, \tag{20a}$$

$$v(0) = v^0 \quad \text{and} \quad M^T\lambda_1(T) = -V_1[v(T) - v^*(T)], \tag{20b}$$

with $F := A - B_1R^{-1}S_{uv}$, symmetric matrices $G := B_1R^{-1}B_1^T$ and $H := -W_1 + S_{vu}R^{-1}S_{uv}$ and

$$u = R^{-1}[B_1^T\lambda_1 - S_{uv}[v - v^*]]. \tag{20c}$$

5 Optimal Control for Output Trajectory Tracking

We briefly recall a common strategy for modifying the cost-functional, such that the framework introduced in Chapter 4 for tracking a state trajectory can be used to track a desired output. A comprehensive introduction of these principles can be found in [15]. Note that we still consider state-feedback. This is different from output-feedback which computes the control based on incomplete information on the state and which requires the design of state-estimators (observers), see [15].

We use the notation of Section 4 and introduce a linear operator C, that extracts an output or observation y from the state v. To steer the system (5) into a state v, such that an observation $y = Cv$ approaches a given output trajectory y^* one may consider minimizing the cost-functional

$$\mathcal{J}(v,u) = \frac{1}{2}\left[Cv - y^*\right]^T \tilde{V} \left[Cv - y^*\right]\Big|_{t=T} + \frac{1}{2}\int_0^T \left[Cv - y^*\right]^T \tilde{W} \left[Cv - y^*\right] + u^T Ru \,\mathrm{d}t, \tag{21}$$

subject to the state equations (5). For ease of notation we do not consider cross terms like $u^T S[Cv - y^*]$ in the cost-functional, as one can use standard techniques, cf. [15], to find equivalent formulations of the linear-quadratic optimal control problem, that do not contain cross terms.

In order to transform the observation tracking problem (21) into a state tracking problem, one has to find a $\tilde{v}^*$ such that $C\tilde{v}^* = y^*$. Then (21) may be written as

$$\mathcal{J}(v,u) = \frac{1}{2}\left[v - \tilde{v}^*\right]^T C^T V_y C \left[v - \tilde{v}^*\right]\Big|_{t=T} + \frac{1}{2}\int_0^T \left[v - \tilde{v}^*\right]^T C^T W_y C \left[v - \tilde{v}^*\right] + u^T Ru \,\mathrm{d}t,$$

with matrices V_y, W_y designed to weight an output, or as

$$\mathcal{J}(v,u) = \frac{1}{2}\left[v - \tilde{v}^*\right]^T V \left[v - \tilde{v}^*\right]\Big|_{t=T} + \frac{1}{2}\int_0^T \left[v - \tilde{v}^*\right]^T W \left[v - \tilde{v}^*\right] + u^T Ru \,\mathrm{d}t,$$

with matrices V, W suited to measure the state.

In many applications the output operator C is in $\mathbb{R}^{n_l,n_v}$ with $n_l \ll n_v$. In this case and, if, as it should be, C has full rank, the one can compute the left inverse $C^- := C^T[CC^T]^{-1}$ and obtain a $\tilde{v}^*$ via $\tilde{v}^* = C^- y^*$.

6 Existence and Representations of Optimal Solutions

One outcome of the proof of Lemma 4.1 is that under the assumption (16) the considered Euler-Lagrange equations are in the form (5). Therefore, one may apply Theorem 3.3 to identify the inherent ODE (14). If then the data is consistent, one may use the theory for ODEs to state the existence of solutions to the obtained linear boundary value problem, cf. [2, Thm. 3.26]. However, the reformulation as used in Theorem 3.3 will not preserve the symmetry of (20) and thus make it more difficult to investigate whether the boundary values admit the existence of a solution. The following theorem makes use of a reformulation that preserves the structure such that the existence of a solution can be obtained via a standard differential Riccati equation.

Lemma 6.1. *Consider the semi-explicit linear DAE of index 2*

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} - \begin{bmatrix} A & J_1^T \\ J_2 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u = \begin{bmatrix} f_v \\ f_p \end{bmatrix}, \quad v(0) = v^0 \tag{22}$$

and a cost functional

$$\mathcal{J}(v,u) = \frac{1}{2}\left[v - v^*\right]^T V_1 \left[v - v^*\right]\Big|_{t=T} + \frac{1}{2}\int_0^T \begin{bmatrix} v - v^* \\ u \end{bmatrix}^T \begin{bmatrix} W_1 & S_{vu} \\ S_{uv} & R \end{bmatrix} \begin{bmatrix} v - v^* \\ u \end{bmatrix} \,dt, \tag{23}$$

which does not act onto the algebraic variable p and with symmetric positive semi-definite weighting matrices and R symmetric positive definite. Define the matrix functions $F := A - B_1 R^{-1} S_{uv}$, $G := B_1 R^{-1} B_1^T$ and $H := -W_1 + S_{vu} R^{-1} S_{uv}$.

1.) If R is invertible, then each solution $(v, p, \lambda_1, \lambda_2)$ of the associated Euler-Lagrange equations as given by (20) has a representation with $(v, p) = (v_0 + Lv, p)$ and $(M^T\lambda_1, \lambda_2) = (\lambda_0 + L^T M^T \lambda_1, \lambda_2)$ given by the decoupled system

$$\begin{aligned}
Lv &= M^{-1} J_1^T S^{-1} f_p, && \text{(24a)}\\
L^T M^T \lambda_1 &= 0, && \text{(24b)}\\
\lambda_2 &= S^{-T} J_1 M^{-T} [H[Lv + v_0] + [W_1 - S_{vu} R^{-1} S_{uv}] v^* - F M^{-T} \lambda_0], && \text{(24c)}\\
p &= -L^- M^{-1} [F[Lv + v_0] + f_v + B_1 R^{-1} S_{uv} v^*] - L^- G[L^T [H[Lv + v_0] + \\
&\quad + [W_1 - S_{vu} R^{-1} S_{uv}] v^*] + \lambda_0] - S^{-1} \dot{f}_p, && \text{(24d)}
\end{aligned}$$

and

$$\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} \dot{\lambda}_0 \\ \dot{v}_0 \end{bmatrix} - \begin{bmatrix} G_0 & F_0 \\ F_0^T & H_0 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} [I - L] M^{-1} [f_v + B_1 R^{-1} S_{uv} v^* - F M^{-1} J_1^T S^{-1} f_p] \\ [I - L^T][W_1 v^* - S_{vu} R^{-1} S_{uv} v^* + H M^{-1} J_1^T S^{-1} f_p] \end{bmatrix}, \tag{24e}$$

$$v_0(0) = [I - L] v^0 \quad \text{and} \quad \lambda_0(T) = -[I - L^T] V_1 [v(T) - v^*(T)], \tag{24f}$$

where $F_0 := \frac{d}{dt}([I - L]) + [I - L] M^{-1} F [I - L]$, $G_0 = G_0^T := [I - L] M^{-1} G M^{-T} [I - L^T]$, $H_0 = H_0^T := [I - L^T] H [I - L]$ and L, L^- and S as defined in Theorem 3.3.

2.) If in addition

$$J_2 v^0 = f_p(0) \quad \text{and} \quad J_1 M^{-T} V_1 = 0, \tag{25}$$

then the Euler-Lagrange equations (20) possess a unique solution.

3.) If in addition f_v, f_p and v^* are zero, then the solution of (20) can be decoupled via

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix}, \tag{26}$$

where $X_1 = X_1^T$ and X_2 fulfill the differential-algebraic matrix Riccati equation

$$\begin{aligned}
\tfrac{d}{dt} M^T X_1 M + M^T X_1 F + F^T X_1 M + M^T X_1 G X_1 M + H + M^T X_2^T J_2 + J_2^T X_2 M = 0,&\\
M^T X_1(T) M = -V_1,& \quad \text{(27a)}\\
M^T J_1 X_1 = 0 \quad \text{and} \quad J_1 X_1 M = 0.& \quad \text{(27b)}
\end{aligned}$$

Equations (27) uniquely define a symmetric positive semi-definite X_1.

4.) If, however, f_v, f_p and v^* are not identically zero, then the solution of (20) decouples via

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \tag{28}$$

where X_1 and X_2 are given by a solution of (27) and (w_1, w_2) is the unique solution of

$$\begin{aligned}
\begin{bmatrix} M^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{w}_1 \\ \dot{w}_2 \end{bmatrix} - \begin{bmatrix} M^T X_1 G + F^T & J_2^T \\ J_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= \begin{bmatrix} f_{\lambda_1} + M^T X_1 \tilde{f}_v + M^T X_2 f_p \\ f_{\lambda_2} \end{bmatrix}, && \text{(29)}\\
M^T w_1(T) &= -V_1 v^*(T),
\end{aligned}$$

with $\tilde{f}_v := f_v + B_1 R^{-1} S_{uv} v^*$, $f_{\lambda_1} := [W_1 - S_{vu} R^{-1} S_{uv}] v^*$ and $f_{\lambda_2} = 0$.

Proof of Lemma 6.1. ad *1.)* In order to make the self-adjoint structure obvious, cf. [20], we assume the existence of $\dot{M}$ for a moment. We abbreviate the right-hand side by $\tilde{f}$, so that the associated Euler-Lagrange system reads

$$\begin{bmatrix} 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 \\ -M^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \dot{v} \\ \dot{p} \end{bmatrix} - \begin{bmatrix} G & 0 & F & J_1^T \\ 0 & 0 & J_2 & 0 \\ F^T + \dot{M}^T & J_2^T & H & 0 \\ J_1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix} = \tilde{f}, \tag{30a}$$

$$v(0) = v^0 \quad \text{and} \quad M^T\lambda_1(T) = -V_1[v(T) - v^*(T)], \tag{30b}$$

with $F := A - B_1R^{-1}S_{uv}$, and symmetric $G := B_1R^{-1}B_1^T$ and $H := -W_1 + S_{vu}R^{-1}S_{uv}$, cf. (20).

In order to preserve the symmetry, only congruence transformations should be applied. Thus a scaling of the equations has to be followed by a variable transformation. To start with, we perform a scaling by $S_1 := \mathtt{diag}(M^{-1}, I, I, I)$ and a variable transformation by S_1^{-T} such that (30a) now reads:

$$\begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 \\ -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \frac{d}{dt} \begin{bmatrix} M^T\lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix} - \begin{bmatrix} M^{-1}GM^{-T} & 0 & M^{-1}F & M^{-1}J_1^T \\ 0 & 0 & J_2 & 0 \\ F^TM^{-T} & J_2^T & H & 0 \\ J_1M^{-T} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} M^T\lambda_1 \\ \lambda_2 \\ v \\ p \end{bmatrix} = S_1\tilde{f}.$$

Thus, one can see that the existence of $\dot{M}$ is indeed not necessary for the proper statement of the Euler-Lagrange equations.

In accordance to (10) we then congruently transform the system by

$$S_2 := \begin{bmatrix} \mathcal{E}_2^{-1} & & \\ & I & \\ & & I \end{bmatrix} = \begin{bmatrix} [I-L] & [I - [I-L]M^{-1}F]M^{-1}J_1^TS^{-1} & 0 & 0 \\ L^- & -[I + L^-M^{-1}FM^{-1}J_1^TS^{-1} & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix},$$

where $\mathcal{E}_2 = \begin{bmatrix} I + M^{-1}FL & M^{-1}J_1^T \\ J_2 & 0 \end{bmatrix}$ as defined in (8) with the inverse given in (9). The summand that comes from the time-dependency in the variable transformation S_2^T is given by

$$S_2 \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 \\ -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \dot{S}_2^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{d}{dt}(I - L^T) & -\dot{L}^{T-} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

With this we get the scaled and transformed system

$$\begin{bmatrix} 0 & 0 & I-L & 0 \\ 0 & 0 & L^- & 0 \\ -[I-L^T] & -L^{T-} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\tilde{\lambda}}_1 \\ \dot{\tilde{\lambda}}_2 \\ \dot{v} \\ \dot{p} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{d}{dt}(I - L^T) & -\dot{L}^{T-} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ v \\ p \end{bmatrix}$$

$$- \begin{bmatrix} [I-L]M^{-1}GM^{-T}[I-L^T] & M^{-1}GM^{-T}L^{T-} & [I-L]M^{-1}F[I-L] + L & 0 \\ L^-M^{-1}GM^{-T} & 0 & L^-M^{-1}F[I-L] - L^- & I \\ [I-L^T]F^TM^{-T}[I-L^T] + L^T & [I-L^T]F^TM^{-T}L^{T-} - L^{T-} & H & 0 \\ 0 & I & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ v \\ p \end{bmatrix} = \tilde{f}, \tag{31}$$

with the transformed state and scaled right hand side

$$\begin{bmatrix}\tilde\lambda_1\\ \tilde\lambda_2\\ v\\ p\end{bmatrix} := S^{-T}\begin{bmatrix}M^T\lambda_1\\ \lambda_2\\ v\\ p\end{bmatrix} = \begin{bmatrix}[I+L^TF^TM^{-T}]M^T\lambda_1 + J_2^T\lambda_2\\ J_1\lambda_1\\ v\\ p\end{bmatrix}$$

$$\text{and}\quad \tilde f := S_2S_1\begin{bmatrix}f_v + B_1R^{-1}S_{uv}v^*\\ f_p\\ W_1v^* - S_{vu}R^{-1}S_{uv}v^*\\ 0\end{bmatrix},$$

respectively. From the last line in (31) we find that $\tilde\lambda_2 = 0$ so that we can rewrite the equations for $(\tilde\lambda_1, v, p)$ as

$$\begin{bmatrix}I-L & 0\\ L^- & 0\end{bmatrix}\begin{bmatrix}\dot v\\ \dot p\end{bmatrix} - \begin{bmatrix}[I-L]M^{-1}GM^{-T}[I-L^T] & 0\\ L^- & 0\end{bmatrix}\begin{bmatrix}\tilde\lambda_1\\ \tilde\lambda_2\end{bmatrix} - \begin{bmatrix}[I-L]M^{-1}F[I-L]+L & 0\\ L^-M^{-1}F[I-L]-L^- & 0\end{bmatrix}\begin{bmatrix}v\\ p\end{bmatrix} = \mathcal{E}_2^{-1}M^{-1}\begin{bmatrix}f_v + B_1R^{-1}S_{uv}v^*\\ f_p\end{bmatrix} \tag{32a}$$

and

$$-\tfrac{d}{dt}([I-L^T]\tilde\lambda_1) - [[I-L^T]F^TM^{-T}[I-L^T] + L^T]\tilde\lambda_1 - Hv = [W_1 - S_{vu}R^{-1}S_{uv}]v^*. \tag{32b}$$

Analogously to (11) we apply the projectors

$$\mathscr{Q}_1 = \begin{bmatrix}L & 0\\ -L^- & 0\end{bmatrix},\quad \mathscr{Q}_0\mathscr{P}_1 = \begin{bmatrix}0 & 0\\ L^- & I\end{bmatrix}\quad\text{and}\quad \mathscr{P}_0\mathscr{P}_1 = \begin{bmatrix}I-L & 0\\ 0 & 0\end{bmatrix}$$

to (32a) to obtain the three subsystems

$$-\begin{bmatrix}L & 0\\ -L^- & 0\end{bmatrix}\begin{bmatrix}v\\ p\end{bmatrix} = \begin{bmatrix}M^{-1}J_1^TS^{-1}f_p\\ S^{-1}f_p]\end{bmatrix}, \tag{33a}$$

$$\begin{bmatrix}0 & 0\\ L^- & 0\end{bmatrix}\begin{bmatrix}\dot v\\ \dot p\end{bmatrix} - \begin{bmatrix}0 & 0\\ L^-M^{-1}GM^{-T} & 0\end{bmatrix}\begin{bmatrix}\tilde\lambda_1\\ \tilde\lambda_2\end{bmatrix} - \begin{bmatrix}0 & 0\\ L^-M^{-1}F[I-L] & I\end{bmatrix}\begin{bmatrix}v\\ p\end{bmatrix} = \begin{bmatrix}0\\ L^-M^{-1}[f_v + B_1R^{-1}S_{uv}v^* - FM^{-1}J_1^TS^{-1}f_p]\end{bmatrix} \tag{33b}$$

and

$$\begin{bmatrix}I-L & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\dot v\\ \dot p\end{bmatrix} - \begin{bmatrix}[I-L]M^{-1}GM^{-T}[I-L^T] & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\tilde\lambda_1\\ \tilde\lambda_2\end{bmatrix} - \begin{bmatrix}[I-L]M^{-1}F[I-L] & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}v\\ p\end{bmatrix} = \begin{bmatrix}[I-L]M^{-1}[f_v + B_1R^{-1}S_{uv}v^* - FM^{-1}J_1^TS^{-1}f_p]\\ 0\end{bmatrix}, \tag{33c}$$

respectively. Having used the projector property of $[I-L^T]$ to obtain the relation

$$\tfrac{d}{dt}([I-L^T]\tilde\lambda_1) = \tfrac{d}{dt}([I-L^T]\tilde\lambda_1) - L^T\tfrac{d}{dt}([I-L^T]\tilde\lambda_1) + \tfrac{d}{dt}(I-L^T)[I-L^T]\lambda_1$$

we split (32b) into the two subsystems

$$L^T \tfrac{d}{dt}([I - L^T]\tilde{\lambda}_1) \; - L^T\tilde{\lambda}_1 - L^T H v = L^T[W_1 - S_{vu}R^{-1}S_{uv}]v^* \quad (34a)$$

and

$$\begin{aligned} &-\tfrac{d}{dt}([I - L^T]\tilde{\lambda}_1) - \tfrac{d}{dt}(I - L^T)[I - L^T]\tilde{\lambda}_1 - && (34b)\\ &\quad -[I - L^T]F^T M^{-T}[I - L^T]\tilde{\lambda}_1 - [I - L^T]Hv = [I - L^T][W_1 - S_{vu}R^{-1}S_{uv}]v^*, && (34c) \end{aligned}$$

respectively. If we then define $v_0 := [I - L]v$ and $\tilde{\lambda}_0 := [I - L^T]\tilde{\lambda}_1$ and decompose $\tilde{\lambda}_1 = \tilde{\lambda}_0 + L^T\tilde{\lambda}_1$ and $v = v_0 + Lv$ we find that (33a-b) and (34a) define algebraic relations for

$$\begin{aligned} Lv &= M^{-1}J_1^T S^{-1} f_p, && (35a)\\ L^T\tilde{\lambda}_1 &= -L^T[HLv + Hv_0 + [W_1 - S_{vu}R^{-1}S_{uv}]v^*] + L^T\dot{\tilde{\lambda}}_0 && (35b) \end{aligned}$$

and

$$\begin{aligned} p = &-L^- M^{-1}[F[Lv + v_0] + f_v + B_1 R^{-1}S_{uv}v^*] - L^- M^{-1}GM^{-T}[L^T\tilde{\lambda}_1 + \tilde{\lambda}_0] + \\ &+ L^-[\dot{v}_0 + \tfrac{d}{dt}(Lv)], \quad (35c) \end{aligned}$$

while $\tilde{\lambda}_0$ and v_0 are defined by the coupled ODEs given by (33c) and (34b):

$$\begin{aligned} &-\dot{\tilde{\lambda}}_0 - \left[\tfrac{d}{dt}(I - L^T) + [I - L^T]F^T M^{-T}[I - L^T]\right]\tilde{\lambda}_0 - [I - L^T]H[I - L]v_0 = && (36a)\\ &\qquad [I - L^T][W_1 v^* - S_{vu}R^{-1}S_{uv}v^* + HM^{-1}J_1^T S^{-1} f_p] \end{aligned}$$

and

$$\begin{aligned} &\dot{v}_0 - [I - L]M^{-1}GM^{-T}[I - L^T]\tilde{\lambda}_0 - \left[\tfrac{d}{dt}([I - L]) + [I - L]M^{-1}F[I - L]\right]v_0 = && (36b)\\ &\qquad [I - L]M^{-1}[f_v + B_1 R^{-1}S_{uv}v^* - FM^{-1}J_1^T S^{-1} f_p]. && (36c) \end{aligned}$$

Note that we have used the projector property $[I - L] = [I - L]^2$ to keep the symmetry in (36) obvious.

In view of expressing the obtained relations in terms of the original variables (λ_1, λ_2) we observe that

$$\tilde{\lambda}_0 = [I - L^T]\tilde{\lambda}_1 = [I - L^T][M^T\lambda_1 + L^T F^T \lambda_1 + J_2^T\lambda_2] = [I - L^T]M^T\lambda_1 =: \lambda_0.$$

From $\lambda_2 = J_1\lambda_1 = 0$ we confer

$$L^T M^T \lambda_1 = J_2^T S^{-T} J_1 \lambda_1 = 0.$$

For λ_2 we can make use of $L^T\tilde{\lambda}_1 = L^T[I + L^T F M^{-T}]M^T\lambda_1 + L^T J_2^T\lambda_2 = L^T F M^{-T}\lambda_0 + J_2^T\lambda_2$, relation (35a) and the regularity of $S^T = J_1 M^{-1} J_2^T$ to get

$$\begin{aligned} \lambda_2 &= S^{-T}J_1 M^{-T}[L^T H[Lv + v_0] + [W_1 - S_{vu}R^{-1}S_{uv}]v^* - L^T F M^{-T}\lambda_0]\\ &= S^{-T}J_1 M^{-T}[H[Lv + v_0] + [W_1 - S_{vu}R^{-1}S_{uv}]v^* - FM^{-T}\lambda_0]. \quad (37) \end{aligned}$$

Similarly one can express the equation for p in terms of (λ_1, λ_2) which completes the derivation of Equations (24).

ad *2.)*

First, we show that for any v^0 and $[I - L^T]V_1$ symmetric positive semi-definite the decoupled system (24) has a unique solution $(v_0, Lv, p, \lambda_0, L^T M^T \lambda_1, \lambda_2)$. Second, we confer that under the consistency conditions (25) the solution of (24) provides a solution of the Euler-Lagrange equations (20). Finally, by *1.)* every solution of (20) has a representation in (24), such that in summary the Euler-Lagrange equations must possess a unique solution.

We first consider in (24e-f) the case with a zero right hand side and $v^* = 0$. With the Riccati ansatz $\lambda_0 = X_0(t)v_0(t)$ this equation can be rewritten as the differential matrix Riccati equation

$$\dot{X}_0 = -X_0 G_0 X_0 - X_0 F_0 - F_0^T X_0 - H_0, \quad X_0(T) = -[I - L^T]V_1, \tag{38}$$

which has a unique solution, cf. [1, Thm. 4.1.6], since $[I - L^T]V_1, G_0$ and $-H_0 = W_1 - S_{uv}R^{-1}S_{vu}$ as the Schur complement of a symmetric positive semi-definite matrix are symmetric positive semi-definite. With this X_0 we get v_0 and λ_0 as the solution of $\dot{v}_0 - [G_0 X_0 + F_0]v_0 = 0, v_0(0) = [I - L]v^0$ and $\lambda_0 = X_0 v_0$, respectively.

One can show that if there exists a solution to (24e-f) with a zero right hand side, then it is unique. This is equivalent to the fact that the linear part of the affine boundary conditions are stated such, that (24e-f) with $[I - L^T]V_1$ symmetric positive semi-definite, has a unique solution, cf. [2, Thm. 3.26], for any continuous right hand side and any $v^* \neq 0$.

By construction a solution of (20) uniquely defines a solution to (24). The converse is true if and only if the algebraic variables fulfill the initial and terminal conditions, i.e.,

$$Lv(0) = Lv_0 = M^{-1} J_1^T S^{-1} J_2 v_0 \tag{39a}$$

$$\text{and} \quad L^T M^T \lambda_1(T) = -L^T V_1(v(T) - v^*(T)) = -J_2^T S^{-T} J_1 M^{-T} V_1(v(T) - v^*(T)). \tag{39b}$$

By (24a) we have that $Lv(0) = M^{-1} J_1^T S^{-1} f_p(0)$ such that $J_2 v(0) = f_p(0)$ is necessary and sufficient for (39a). By (24b) we have that $L^T M^T \lambda_1 = 0$ such that $J_1 M^{-T} V_1 = 0$ is sufficient but not necessarily necessary for (39b). Note, however, that in this case we can infer that

$$J_1^T M^{-T} V_1 = 0 \quad \Rightarrow \quad V_1 M^{-1} J_1 = 0 \quad \Rightarrow \quad V_1 L = 0 \quad \Rightarrow \quad V_1 v = V_1 [I - L]v$$

such that in (24f) $[I - L^T]V_1$ can be replaced by $[I - L^T]V_1[I - L]$. Thus, condition (25) implies the symmetry in the terminal condition that was sufficient for the existence of X_0 in (38).

ad *3.)*

With the ansatz

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} \tag{40}$$

we obtain that

$$\frac{d}{dt}\left(\begin{bmatrix} M^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}\right) = \begin{bmatrix} \frac{d}{dt}(M^T X_1 M) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} + \begin{bmatrix} M^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix}. \tag{41}$$

In (41) we replace $\begin{bmatrix} M^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$ and $\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix}$ via the relations given in (20) and every occurrence of $\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$ by the ansatz (40) to obtain

$$\mathscr{X} \begin{bmatrix} v \\ p \end{bmatrix} = 0, \quad \text{where } \mathscr{X} := \begin{bmatrix} \begin{matrix} \frac{d}{dt}(M^T X_1 M) + F^T X_1 M + M^T X_1 F + \\ M^T X_1 G X_1 M + H + J_2^T X_2 M + M^T X_2^T J_2 \end{matrix} & M^T X_1 J_1^T \\ J_1 X_1 M & 0 \end{bmatrix}. \tag{42}$$

Since $\mathscr{X}\begin{bmatrix} v \\ p \end{bmatrix} = 0$ must hold for every state trajectory, one requires $\mathscr{X} = 0$ which gives three equations for X_1 and X_2:

$$\tfrac{d}{dt}(M^T X_1 M) + F^T X_1 M + M^T X_1 F + M^T X_1 G X_1 M + H + J_2^T X_2 M + M^T X_2^T J_2 = 0,$$
$$M^T X_1(T) M = -V_1, \qquad (43a)$$
$$M^T X_1 J_1^T = 0 \quad \text{and} \quad J_1 X_1 M = 0. \qquad (43b)$$

The terminal condition in (43a) is defined via (20b) and (26):

$$M^T \lambda_1(T) = M^T X_1(T) M v(T) = -V_1 v(T) \quad \Rightarrow \quad M^T X_1(T) M = -V_1.$$

To show that (43) has a solution we consider Equation (43a) in the transformed variables $X := -M^T X_1 M$ and $Y := X_2 M$ for $t \in [0, T]$:

$$-\dot{X} - F^T M^{-T} X - X M^{-1} F + X M^{-1} G^{-T} M^{-T} X + H + J_2^T Y + Y^T J_2 = 0, \quad X(T) = V_1. \quad (44)$$

By means of the projector $L := M^{-1} J_1^T [J_2 M^{-1} J_1^T]^{-1} J_2$ we write $X = [L^T + [I - L^T]] X [[I - L] + L]$. From (43b) we obtain that $L^T X = XL = 0$ and thus X is completely defined via $X_0 := [I - L^T] X [I - L]$. Applying $[I - L^T]$ and $[I - L]$ to (44) from the left and the right, respectively, we get a standard differential Riccati equation

$$-\dot{X}_0 - \big[[I - L^T] F^T M^{-T} - \tfrac{d}{dt}(I - L^T)\big] X_0 - X_0 \big[M^{-1} F [I - L] - \tfrac{d}{dt}(I - L)\big] +$$
$$+ X_0 M^{-1} G M^{-T} X_0 + [I - L^T] H [I - L] = 0,$$
$$X_0(T) = [I - L^T] V_1 [I - L], \qquad (45)$$

which has a unique and symmetric positive semi-definite solution, cf. [1, Thm. 4.1.6], since V_1, G and $-H$ are symmetric positive semi-definite. Again, the consistency condition (25) ensures that $X_0(0)$ also satisfies the initial condition and the algebraic constraints in (43). Since $L^T X = 0$ and $XL = 0$, we have $X_1 = -M^{-T} X M^{-1}$ is unique and symmetric negative semi-definite.

Application of $[I - L^T]$ from the left and L from the right to (44) gives

$$-X_0 \dot{L} - X_0 M^{-1} F L + [I - L^T] H L = -[I - L^T] Y^T J_2 L = -[I - L^T] Y^T J_2,$$

which is uniquely solvable for $[I - L^T] Y^T$. The projected equation obtained via L^T and $[I - L]$ is the transpose of the above equation and bears no additional information.

Finally, one can determine $L^T M^T X_2^T$ from the projection of (44) onto the range of L^T and L which reads

$$L^T H L + L^T Y^T J_2 L + L^T J_2^T Y L = 0. \qquad (46)$$

With $J_2 L = J_2$, we find that (46) is of the form $[YL]^T J_2 + J_2^T [YL] = -L^T H L$ that was investigated in [8]. With $L^- := M^{-1} J_1^T [J_2 M^{-1} J_1^T]^{-1}$ being a generalized inverse to J_2 we obtain the projectors $P_1 := L^- J_2 = L$ and $P_2 := J_2 L^- = I$ and the existence of solutions to (46) follows by [8, Thm. 1], since $L^T H L$ is symmetric and $[I - P_1]^T L^T H L [I - P_1] = 0$.

The general solution to (46) is given by

$$YL = \frac{1}{2} [J_1 M^{-T} J_2^T]^{-1} J_1 M^{-T} H L + Z J_2,$$

where Z is an arbitrary matrix with $Z^T = -Z$. Thus the existence of $M^T X_1$ and $M^T X_2^T = Y^T = [I - L^T] Y^T + L^T Y^T$ and therefore X_1 and X_2 is proved.

By construction, with X_1 and X_2 as determined above, the solution of

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} - \left(\begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} F & J_1^T \\ J_2 & 0 \end{bmatrix} \right) \begin{bmatrix} v \\ p \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad v(0) = v^0,$$

and

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix}$$

gives the solution of (20) with a zero right-hand side.

ad *4.)*

Proceeding analogously to the first steps for part *3.)*, but with the affine linear ansatz (28) instead of the linear (26), we come to the expression

$$\mathscr{X} \begin{bmatrix} v \\ p \end{bmatrix} + \tfrac{d}{dt} \left(\begin{bmatrix} M^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right) - \begin{bmatrix} M^T X_1 G + F^T & J_2^T \\ J_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} f_{\lambda_1} + M^T X_1 \tilde{f}_v + M^T X_2 f_p \\ f_{\lambda_2} \end{bmatrix},$$
$$M^T w_1(T) = -V_1 v^*(T), \tag{48}$$

where $\mathscr{X}$ is as in (42). Again, the requirement $\mathscr{X} = 0$ uniquely defines X_1 and $X_2 =: \tilde{X}_2 + Z J_2 M^{-1}$ up to an arbitrary skew-symmetric matrix Z. Anticipating the proof of Remark 6.2 we write $M^T X_2^T f_p = M^T \tilde{X}_2^T f_p - J_2^T Z f_p$ and define $\tilde{w}_2 := w_2 - Z f_p$. With this Equation (48) gives a system for $(w_1, \tilde{w}_2)$:

$$\begin{bmatrix} \frac{d}{dt}(M^T w_1) \\ 0 \end{bmatrix} - \begin{bmatrix} M^T X_1 G + F^T & J_2^T \\ J_1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \tilde{w}_2 \end{bmatrix} = \begin{bmatrix} f_{\lambda_1} + M^T X_1 \tilde{f}_v + M^T X_2 f_p \\ f_{\lambda_2} \end{bmatrix},$$
$$M^T w_1(T) = -V_1 v^*(T), \tag{49}$$

which is of type (5). Since by (25) the terminal condition is consistent, system (49) has a unique solution $(w_1, \tilde{w}_2)$. In particular, w_1 is independent of $Z f_p$, cf. (12) and (14). For the solution w_2 of (48) we have $w_2 = \tilde{w}_2 + Z f_p$. Thus the existence of the functions used for the ansatz (28) is shown.

Thus, by construction we have that the ansatz (28) leads to the solution of the Euler-Lagrange equations (20) via the decoupled system

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} - \left(\begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} F & J_1^T \\ J_2 & 0 \end{bmatrix} \right) \begin{bmatrix} v \\ p \end{bmatrix} = \begin{bmatrix} f_v + B_1 R^{-1} S_{uv} v^* + G w_1 \\ f_p \end{bmatrix},$$
$$v(0) = v^0,$$

and

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

□

Remark 6.2. The solution of (27) is unique up to an additive term $Z J_2 M^{-1}$ in X_2, with an arbitrary matrix Z, that fulfills $Z^T = -Z$. However, this does not contradict the unique solvability of the Euler-Lagrange equations, since λ_1 and λ_2 as defined via (26) or (28) are independent of any choice of Z.

Proof of Remark 6.2. As shown in the proof of Lemma 6.1, the matrix functions X_1 and X_2 for the generalized Riccati ansatz in (26) and (28) are uniquely defined up to an additive constant ZJ_2M^{-1} in X_2. This nonuniqueness is also found in the affine part $w_2 = \tilde{w}_2 + Zf_p$, where $\tilde{w}_2$ is independent of Z, as is w_1. Since a solution v must satisfy the algebraic condition in (22), i.e. $J_2v = -f_p$, one obtains that

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} X_1 & X_2^T \\ X_2 & 0 \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} X_1Mv + w_1 \\ \tilde{X}_2Mv + ZJ_2v + \tilde{w}_2 + Zf_p \end{bmatrix} = \begin{bmatrix} X_1Mv + w_1 \\ \tilde{X}_2Mv + \tilde{w}_2 \end{bmatrix}$$

is uniquely defined independent of any choice of Z. □

Corollary 6.3. *The matrix X_1 from the Riccati-decoupling* (26) *that defines the feedback-law for the unreduced system* (22) *also constitutes the feedback-law for the dynamical part of the reduced system (24e-f). In particular one has $X_0 = M^TX_1M$, where X_0 decouples (24e) via $v_0 = X_0\lambda_0$.*

Proof. Since (38) and (45) coincide up to the sign of the unknown one confers that $X_0 = -X = M^TX_1M$. □

In view of optimal control, the above results can be summarized as follows. To obtain an optimal input u for (22) with respect to a cost functional of type (23) it is sufficient to have a solution of the associated Euler-Lagrange equations (20), cf. [5]. By Lemma 6.1 it follows that for the considered state equations and cost functionals this solution exists, that it is unique and that it can be obtained via the separation ansatz (28). Finally one obtains an optimal u via expression (20c). Thus, we have proved the following theorem:

Theorem 6.4. *Let $T > 0$, $\mathbb{I} = (0, T]$ a time interval, $n_u, n_v, n_p \in \mathbb{N}$, $n_v > n_p$, $M \in \mathcal{C}(\mathbb{I}, \mathbb{R}^{n_v,n_v})$ pointwise invertible, $A \in \mathcal{C}(\mathbb{I}, \mathbb{R}^{n_v,n_v})$, and let $J_1, J_2 \in \mathcal{C}(\mathbb{I}, \mathbb{R}^{n_p,n_v})$ such that $S := J_2M^{-1}J_1^T$ is invertible and that $L := M^{-1}J_1^TS^{-1}J_2$ is differentiable. Let $W_1, V_1 \in \mathbb{R}^{n_v,n_v}$ symmetric positive semi-definite, $S_{uv} = S_{vu}^T \in \mathbb{R}^{n_u,n_v}$ an arbitrary matrix and let $R \in \mathbb{R}^{n_u,n_u}$ symmetric positive definite.*

For a given $v^* \in \mathcal{C}^1(\mathbb{I}, \mathbb{R}^{n_v})$ consider the linear-quadratic optimal control problem of finding $u \in \mathcal{C}(\mathbb{I}, \mathbb{R}^{n_u})$ such that

$$\mathcal{J}(v, u) = \frac{1}{2}\left[v - v^*\right]^T V_1 \left[v - v^*\right]\bigg|_{t=T} + \frac{1}{2}\int_0^T \begin{bmatrix} v - v^* \\ u \end{bmatrix}^T \begin{bmatrix} W_1 & S_{vu} \\ S_{uv} & R \end{bmatrix} \begin{bmatrix} v - v^* \\ u \end{bmatrix} \, dt,$$

is minimal, where v on $\mathbb{I}$ satisfies the state equations

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{p} \end{bmatrix} - \begin{bmatrix} A & J_1^T \\ J_2 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u = \begin{bmatrix} f_v \\ f_p \end{bmatrix}, \quad v(0) = v^0.$$

If $f_v \in \mathcal{C}(\mathbb{I}, \mathbb{R}^{n_v})$, $f_p \in \mathcal{C}^1(\mathbb{I}, \mathbb{R}^{n_p})$ and if

$$J_2v^0 = f_p(0) \quad \text{and} \quad J_1M^{-T}V_1 = 0,$$

then the optimal control problem is solvable and an optimal input u is given via the feedback-law

$$u = R^{-1}[B_1^T[X_1Mv + w_1] - S_{uv}(v - v^*)],$$

where $X_1 = X_1^T$, negative semi-definite, and w_1 are the unique solutions of

$$\begin{aligned} \tfrac{d}{dt}M^TX_1M + F^TX_1M + M^TX_1F + M^TX_1GX_1M + H + J_2^TX_2M + M^TX_2^TJ_2 = 0,& \\ M^TX_1(T)M = -V_1,& \\ J_1X_1M = 0,& \end{aligned}$$

and

$$\begin{aligned}\tfrac{d}{dt}(M^T w_1) - [M^T X_1 G + F^T]w_1 - J_2^T w_2 &= f_{\lambda_1} + M^T X_1 \tilde{f}_v + M^T X_2 f_p,\\ M^T w_1(T) &= -V_1 v^*(T),\\ J_1 w_1 &= 0,\end{aligned}$$

respectively, where $F := A - B_1R^{-1}S_{uv}$*,* $G := B_1R^{-1}B_1^T$ *and* $H := -W_1 + S_{vu}R^{-1}S_{uv}$ *and with* $\tilde{f}_v := f_v + B_1R^{-1}S_{uv}v^*$ *and* $f_{\lambda_1} := [W_1 - S_{vu}R^{-1}S_{uv}]v^*$*.*

7 Discussion and Outlook

The results of Section 4 for the linear case can serve as a basis for the numerical treatment of the semi-linear equations (3) that sooner or later will carry out linearizations. Another crucial point is the incorporation of constraints for the control as well as lower regularity.

The latter has been investigated in [24] for a large class of optimal control problems of type (2) subject to semi-explicit DAEs (3) and in particular the cases that are considered here. The formulation of a maximum principle in [24] bases on an equivalent index 1 representation of the state equations that can be formally obtained for the semi-explicit case. In the index 2 case it is given via

$$M\dot{v} - f(t,v,p,u) = 0, \quad v(0) = v^0, \tag{51a}$$

$$G(t,v,p,u) = 0, \tag{51b}$$

where the function $G(t,v,p,u) := \dot{g} + gM^{-1}f$ allows an implicit function representation of $p = F(v,u)$. The maximum principle in [24] states that a solution (v,p,u) of the associated optimal control problem with a cost-functional as in (2) and pointwise constraints on u is a maximizer for a specifically chosen Hamilton function, i.e. (v,p,u) solves

$$\max_{(p,u)\in\mathcal{D}(t,v)} \mathcal{H}(t,p,u;v,\lambda) = \max_{(p,u)\in\mathcal{D}(t,v)} \lambda_1^T f(t,v,p,u) - \mathcal{K}(t,v,p,u), \tag{52}$$

where

$$\mathcal{D}(t,v) := \big\{(p,u):\; u \text{ is admissible and } G(t,v,p,u) = 0\big\},$$

and where λ_1 solves the adjoint equation of the reduced system (51). In order to compare it to our result, we assume for the moment that u is unconstrained and that the first variation, cf. [28], of f and $\mathcal{K}$ with respect to u exists. With the implicit function theorem we write $p = F(v,u)$ and the constrained maximization problem in (52) becomes unconstrained:

$$\max_u \mathcal{H}(t,F(v,u),u;v,\lambda) = \max_u \lambda_1^T \tilde{f}(t,v,u) - \tilde{\mathcal{K}}(t,v,u),$$

with the reduced functions $\tilde{f}$ and $\tilde{\mathcal{K}}$. Now a candidate solution may be obtained by setting the first variation to zero, i.e. to find an u such that

$$\tilde{f}_u^T \lambda_1 - \tilde{\mathcal{K}}_u^T = 0. \tag{53}$$

Thus Equation (53) replaces the specification of the maximum in (52) and together with the state equations (51) and its adjoint equations it is equivalent to the Euler-Lagrange as given in (4e) for the index reduced system and a cost-functional that can be reduced to act only on u and the dynamic variable v. This means that under the assumptions that were necessary for the derivation of our results the maximum principle may also provide a system for the computation of an optimal input, however, on the base of an index-1 formulation.

The above reasoning also shows that the results of [24] are not the natural extension of our results to input constrained problems. Therefore, a next step will be to find a maximum principle that is given in terms of the unreduced system.

In view of solving the optimal control problem numerically it may be worth investigating, whether the Riccati decoupling can be exploited for efficient numerical routines.

The type of system considered here was chosen to fit spatially discretized PDEs as the Navier-Stokes Equations. For a system-theoretical insight, one may consider similar manipulations on the original infinite-dimensional system.

References

[1] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank. *Matrix Riccati equations in control and systems theory.* Birkhäuser, Basel, 2003.

[2] U. M. Ascher, R. M. Mattheij, and R. D. Russell. *Numerical solution of boundary value problems for ordinary differential equations.* SIAM, Philadelphia, 1995.

[3] A. Backes. Optimale Steuerung der linearen DAE im Fall Index 2. Preprint 2003-04, Institut für Mathematik, Humboldt-Universität zu Berlin, 2003.

[4] A. Backes. *Extremalbedingungen für Optimierungs-Probleme mit Algebro-Differentialgleichungen.* PhD thesis, Institut für Mathematik, Humboldt-Universität zu Berlin, 2006.

[5] K. Balla, G. A. Kurina, and R. März. Index criteria for differential algebraic equations arising from linear-quadratic optimal control problems. *J. Dyn. Control Syst.*, 12(3):289–311, 2006.

[6] K. Balla and V. H. Linh. Adjoint pairs of differential-algebraic equations and Hamiltonian systems. *Appl. Numer. Math.*, 53(2-4):131–148, 2005.

[7] K. Balla and R. März. A unified approach to linear differential algebraic equations and their adjoints. *Z. Anal. Anwendungen*, 21(3):783–802, 2002.

[8] H. Braden. The equations $A^TX \pm X^TA = B$. *SIAM J. Matrix Anal. Appl.*, 20(2):295–302, 1998.

[9] S. L. Campbell and C. W. Gear. The index of general nonlinear DAEs. *Numer. Math.*, 72(2):173–196, 1995.

[10] M. d. R. de Pinho and R. Vinter. Necessary conditions for optimal control problems involving nonlinear differential algebraic equations. *J. Math. Anal. Appl.*, 212(2):493–516, 1997.

[11] M. Gerdts. Optimal control and real-time optimization of mechanical multi-body systems. *Z. Angew. Math. Mech.*, 83(10):705–719, 2003. ECMI-Workshop "Numerical Methods in Multibody Dynamics" (Bad Herrenalb, 2001).

[12] M. Gerdts. Local minimum principle for optimal control problems subject to differential-algebraic equations of index two. *J. Optim. Theory Appl.*, 130(3):443–462, 2006.

[13] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations. Theory and algorithms. (Extended version of the 1979 publ.).* Springer, Berlin, 1986.

[14] P. M. Gresho and R. L. Sani. *Incompressible flow and the finite element method. Vol. 2: Isothermal laminar flow.* Wiley, Chichester, UK, 2000.

[15] J. a. P. Hespanha. *Linear systems theory.* Princeton University Press, Princeton, NJ, 2009.

[16] P. Kunkel and V. Mehrmann. The linear quadratic optimal control problem for linear descriptor systems with variable coefficients. *Math. Control Signals Syst.*, 10(3):247–264, 1997.

[17] P. Kunkel and V. Mehrmann. Analysis of over- and underdetermined nonlinear differential-algebraic systems with application to nonlinear control problems. *Math. Control Signals Syst.*, 14(3):233–256, 2001.

[18] P. Kunkel and V. Mehrmann. *Differential-algebraic equations. Analysis and numerical solution.* European Mathematical Society Publishing House, Zürich, 2006.

[19] P. Kunkel and V. Mehrmann. Optimal control for unstructured nonlinear differential-algebraic equations of arbitrary index. *Math. Control Signals Syst.*, 20(3):227–269, 2008.

[20] P. Kunkel, V. Mehrmann, and L. Scholz. Self-adjoint differential-algebraic equations. Preprint 2011-13, TU Berlin, Institut für Mathematik, July 2011. 07.07.2011, submitted for publication.

[21] G. A. Kurina and R. März. On linear-quadratic optimal control problems for time-varying descriptor systems. *SIAM J. Control Optim.*, 42(6):2062–2077, 2004.

[22] G. A. Kurina and R. März. Feedback solutions of optimal control problems with DAE constraints. *SIAM J. Control Optim.*, 46(4):1277–1298, 2007.

[23] J. W. Polderman and J. C. Willems. *Introduction to mathematical systems theory. A behavioral approach.* Springer, New York, 1997.

[24] T. Roubiček and M. Valášek. Optimal control of causal differential-algebraic systems. *J. Math. Anal. Appl.*, 269(2):616–641, 2002.

[25] J. Sjöberg. *Optimal Control and Model Reduction of Nonlinear DAE Models.* PhD thesis, Linköpings Universitet, 2008.

[26] E. D. Sontag. *Mathematical control theory. Deterministic finite dimensional systems. 2nd ed.* Springer, New York, 1998.

[27] C. Tischendorf. Coupled systems of differential algebraic and partial differential equations in circuit and device simulation. modeling and numerical analysis. Habilitationsschrift, Institut für Mathematik, Humboldt-Universität zu Berlin, 2004.

[28] F. Tröltzsch. *Optimal control of partial differential equations. Theory, procedures, and applications. (Optimale Steuerung partieller Differentialgleichungen. Theorie, Verfahren und Anwendungen).* Vieweg, Wiesbaden, 2005.

[29] H. Xu and K. Mizukami. Hamilton-Jacobi equation for descriptor systems. *Syst. Control Lett.*, 21(4):321–327, 1993.