Necessary and sufficient conditions in the optimal control for general nonlinear differential-algebraic equations *

Peter Kunkel † Volker Mehrmann ‡

September 9, 2006

Abstract

We study optimal control problems for general unstructured nonlinear differential-algebraic equations of arbitrary index. In particular, we derive necessary conditions in the case of linear-quadratic control problems and extend them to the general nonlinear case. We also present a Pontryagin maximum principle for general unstructured nonlinear DAEs in the case of restricted controls. Moreover, we discuss the numerical solution of the resulting two-point boundary value problems and present a numerical example.

Keywords: nonlinear differential-algebraic equations, optimal control, maximum principle, solvability, necessary optimality conditions, sufficient optimality conditions, behavior approach, strangeness index, regularization

AMS(MOS) subject classification: 93C10, 93C15, 65L80, 49K15, 34H05

1 Introduction

We study optimal control problems

$$\mathcal{J}(x,u) = \mathcal{M}(x(\overline{t})) + \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, x(t), u(t))\ dt = \min! \tag{1}$$

subject to a constraint given by an initial value problem associated with a nonlinear system of differential-algebraic equations (*descriptor system*) consisting of

$$F(t, x, u, \dot{x}) = 0 \tag{2}$$

and

$$x(\underline{t}) = \underline{t}. \tag{3}$$

We assume that $F \in C^0(\mathbb{I} \times \mathbb{D}_x \times \mathbb{D}_u \times \mathbb{D}_{\dot{x}}, \mathbb{R}^m)$ is sufficiently smooth, that $\mathbb{I} = [\underline{t}, \overline{t}] \subseteq \mathbb{R}$ is a (compact) interval, and that $\mathbb{D}_x, \mathbb{D}_{\dot{x}} \subseteq \mathbb{R}^n$, $\mathbb{D}_u \subseteq \mathbb{R}^l$ are open sets.

*Supported through the Research-in-Pairs Program at Mathematisches Forschungsinstitut Oberwolfach.

†Mathematisches Institut, Universität Leipzig, Augustusplatz 10–11, D-04109 Leipzig, Germany.

‡Institut für Mathematik, MA 4-5, Technische Universität Berlin, D-10623 Berlin, Germany. Supported by *Deutsche Forschungsgemeinschaft*, through MATHEON, the DFG Research Center "Mathematics for Key Technologies" in Berlin.

Throughout the paper we will frequently make use of the *behavior representation* (see [57]) of the control problem, i.e., we combine (x, u) into one *generalized state vector* z and then study the optimization problem

$$\mathcal{J}(z) = \mathcal{M}(z(\overline{t})) + \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, z(t)) \; dt = \min! \tag{4}$$

subject to the constraint

$$F(t, z, \dot{z}) = 0, \tag{5}$$

and the initial condition

$$z(\underline{t}) = \underline{z}. \tag{6}$$

Optimal control problems like (1)–(3) arise in the control of mechanical multibody systems [25, 28], electrical circuits [30, 31], chemical engineering [22, 23] or heterogeneous systems, where different models are coupled together [55].

The theory of optimal control problems for ordinary differential equations is well established since the middle of the 20th century, see, e.g., [8, 26, 33, 34, 36, 62] and the references therein. For systems where the constraint is a differential-algebraic equation, the situation is much more difficult and the existing literature is more recent. First results, mainly for special cases such as linear constant coefficient systems or semi-explicit systems of index 1, were obtained in [7, 20, 50, 53, 54, 56, 61].

A major difficulty in deriving adjoint equations, optimality systems or even a maximum principle for general higher index DAEs is that for the potential candidates of adjoint equations and optimality systems, existence and uniqueness of solutions cannot be guaranteed, see [1, 2, 4, 21, 49, 56, 60] for examples and discussion of the difficulties.

Due to these difficulties, the standard approach to deal with optimal control problems for DAEs is to first perform regularization and index reduction via feedback or differentiation. Conditions when such transformations exist have been studied in [10, 11, 13, 14] and in their most general form in [41, 43]. Some of these results were reproduced and extended in a different setting in the recent work of [1, 2, 49].

There also exist some papers that derive optimality conditions for specially structured higher index systems directly. For semi-explicit systems of index 1 a general maximum principle was proved in [56] and extended to systems up to differentiation index 3 in [60]. Further results for index 2 systems are presented in [28], for multibody systems in [12, 27] and for DAEs with properly stated leading term in [1, 2, 3, 5, 6, 49].

In the present paper we take a more general approach and discuss general unstructured linear and nonlinear systems of arbitrary index. We follow the strangeness index concept, see [42], and consider the system in a behavior setting as a general over- or underdetermined differential-algebraic system. For this behavior system a derivative array, see [17], is formed and from this array, a reduced control problem is determined that has the same solution set (in the behavior setting) but is essentially index one. Based on this reduced system then the optimality conditions are derived.

The paper is organized as follows. We first give a brief survey of the theoretical results on the strangeness index concept that will be needed in Section 2 and recall some general functional analytic results on optimization in Banach spaces. Then we derive necessary optimality conditions for optimal control problems subject to general linear and nonlinear unstructured DAEs in Section 3. Furthermore, a Pontryagin maximum principle for general DAEs will be presented. In Section 4 we describe another formulation of the optimality boundary value problem that can be used in the context of numerical methods. We finally present a numerical example and give some conclusions in Section 5.

2 Preliminaries

In this section we will introduce some notation and recall some results on differential-algebraic equations and on optimization in Banach spaces. Throughout the paper we assume that all functions are sufficiently smooth, i.e., sufficiently often continuously differentiable.

2.1 Notation

We will make frequent use of the *Moore-Penrose pseudoinverse* of a matrix valued function $A\colon \mathbb{I} \to \mathbb{R}^{m,n}$, which is the unique matrix function $A^+\colon \mathbb{I} \to \mathbb{R}^{n,m}$ that satisfies the four Penrose axioms

$$AA^+A = A, \quad A^+AA^+ = A^+, \quad (AA^+)^T = AA^+, \quad (A^+A)^T = A^+A \tag{7}$$

pointwise, see, e.g. [19]. Note that if $A \in C^k(\mathbb{I}, \mathbb{R}^{m,n})$ and has constant rank on $\mathbb{I}$ then $A^+ \in C^k(\mathbb{I}, \mathbb{R}^{n,m})$.

In the context of restricted control values we must allow for bounded (with respect to the L_∞-norm) and up to a finite number of points continuous control functions. We denote the set of all these functions on the interval $\mathbb{I}$ by $L_\infty^c(\mathbb{I}, \mathbb{R}^l)$.

2.2 DAE theory

The theory of differential-algebraic equations has changed significantly in the last 20 years, see [9, 29, 59, 42]. We recall some necessary concepts and follow [42] in notation and style of presentation.

When studying control problems like (2) one can essentially distinguish two viewpoints. Either one takes the behavior approach and considers the optimization problem (4) subject to (5). For this underdetermined system one can study existence and uniqueness of solutions. In this setting, feedbacks are just standard equivalence transformations. If one carries out a transformation to canonical or condensed form, then index reduction and regularization via feedback follow directly, see [42].

If it is clear that the variables u really describe input variables and the variables x states, as is often the case in practice, then one has to distinguish whether solutions exist for all controls in a given input set $\mathbb{U}$ or whether there exist controls at all for which the system is solvable. To discuss these questions we consider the following solution concept.

Definition 1 *Consider system (2) with a given fixed input function u that is sufficiently smooth. A function $x : \mathbb{I} \to \mathbb{R}^n$ is called a* solution *of (2) if $x \in C^1(\mathbb{I}, \mathbb{R}^n)$ and x satisfies (2) pointwise. It is called a* solution of the initial value problem *(2)–(3) if x is a solution of (2) and satisfies (3). An initial condition (3) is called* consistent *if the corresponding initial value problem has at least one solution.*

It is possible to weaken this solution concept [37, 45, 52, 58]. In particular, it will turn out that it is necessary to slightly weaken this solution concept in order to define underlying Banach space operators with appropriate properties. But to do so, we first must introduce some additional theory. Note, however, that under the assumption of sufficient smoothness we will always be in the case of Definition 1.

Definition 2 *A control problem of the form (2) with a given set of controls* $\mathbb{U}$ *is called* consistent *(with* $\mathbb{U}$*) if there exists an input function* $u \in \mathbb{U}$ *for which there exists a solution* x *in the sense of Definiton 1.*

It is called regular *(locally with respect to a given solution* $(\hat{x}, \hat{u})$ *of (2)) if it has a unique solution for every sufficiently smooth input function* u *in a neighborhood of* $\hat{u}$ *and every initial value in a neighborhood of* $\hat{x}(\underline{t})$ *that is consistent for the system with input function* u*.*

In order to analyze the properties of the system, in [40] for the square nonlinear case, in [43] for the rectangular linear case, and in [41] for the general over- and underdetermined case, hypotheses have been formulated which lead to an index concept, the so-called *strangeness index*. See [42] for a detailed derivation and analysis of this concept.

To summarize the strangeness index concept, we consider the constraint system in the form (5). As in [40], we introduce a nonlinear derivative array, see also [16, 18], of the form

$$F_\ell(t, z, \dot{z}, \ldots, z^{(\ell+1)}) = 0, \tag{8}$$

which stacks the original equation and all its derivatives up to level ℓ in one large system, i.e.,

$$F_\ell(t, z, \dot{z}, \ldots, z^{(\ell+1)}) = \begin{bmatrix} F(t, z, \dot{z}) \\ \frac{d}{dt} F(t, z, \dot{z}) \\ \vdots \\ \frac{d^\ell}{dt^\ell} F(t, z, \dot{z}) \end{bmatrix}. \tag{9}$$

Partial derivatives of F_ℓ with respect to selected variables p from $(t, z, \dot{z}, \ldots, z^{(\ell+1)})$ are denoted by $F_{\ell;p}$, e.g.,

$$F_{\ell;z} = \tfrac{\partial}{\partial z} F_\ell, \quad F_{\ell;\dot{z},\ldots,z^{(\ell+1)}} = [\, \tfrac{\partial}{\partial \dot{z}} F_\ell \ \cdots \ \tfrac{\partial}{\partial z^{(\ell+1)}} F_\ell \,].$$

A corresponding notation is also used for partial derivatives of other functions.

In order to analyze existence and uniqueness of solutions we need to introduce the solution set of the nonlinear algebraic equation associated derivative array F_μ for some integer μ. We denote this solution set by

$$\mathbb{L}_\mu = \{ z_\mu \in \mathbb{I} \times \mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n \mid F_\mu(z_\mu) = 0 \}. \tag{10}$$

Then we make the following hypothesis, see [42].

Hypothesis 1 *Consider the general system of nonlinear differential-algebraic equations (2). There exist integers* μ*,* r*,* a*,* d*, and* v *such that* $\mathbb{L}_\mu$ *is not empty and such that for every* $z_\mu^0 = (t_0, z_0, \dot{z}_0, \ldots, z_0^{(\mu+1)}) \in \mathbb{L}_\mu$ *there exists a (sufficiently small) neighborhood in which the following properties hold:*

1. *The set* $\mathbb{L}_\mu \subseteq \mathbb{R}^{(\mu+2)n+1}$ *forms a manifold of dimension* $(\mu+2)n + 1 - r$*.*
2. *We have* $\operatorname{rank} F_{\mu;z,\dot{z},\ldots,z^{(\mu+1)}} = r$ *on* $\mathbb{L}_\mu$*.*
3. *We have* $\operatorname{corank} F_{\mu;z,\dot{z},\ldots,z^{(\mu+1)}} - \operatorname{corank} F_{\mu-1;z,\dot{z},\ldots,z^{(\mu)}} = v$ *on* $\mathbb{L}_\mu$*, where the corank is the dimension of the corange and the convention is used that* $\operatorname{corank} F_{-1;z} = 0$*.*

4. *We have* $\operatorname{rank} F_{\mu;\dot z,\ldots,z^{(\mu+1)}} = r - a$ *on* $\mathbb{L}_\mu$ *such that there exist smooth full rank matrix functions* Z_2 *and* T_2 *of size* $(\mu+1)m \times a$ *and* $n \times (n-a)$*, respectively, satisfying*

$$Z_2^T F_{\mu;\dot z,\ldots,z^{(\mu+1)}} = 0, \quad \operatorname{rank} Z_2^T F_{\mu;z} = a, \quad Z_2^T F_{\mu;z} T_2 = 0 \tag{11}$$

on $\mathbb{L}_\mu$.

5. *We have* $\operatorname{rank} F_{\dot z} T_2 = d = m - a - v$ *on* $\mathbb{L}_\mu$ *such that there exists a smooth full rank matrix function* Z_1 *of size* $n \times d$ *satisfying* $\operatorname{rank} Z_1^T F_{\dot z} T_2 = d$.

As in [40, 42], we call the smallest possible μ for which Hypothesis 1 is valid the *strangeness index* of (5). Systems with vanishing strangeness index are called *strangeness-free.* The strangeness index is closely related to the differentiation index, see [9], but it should be observed that it allows over- and underdetermined systems and the counting is different, since in the strangeness index concept ordinary differential equations and purely algebraic equations both have $\mu = 0$. See [42] for a detailed analysis of the relationship between different index concepts.

It has been shown in [41] that Hypothesis 1 implies locally (via the implicit function theorem) the existence of a *reduced system* given by

$$\begin{array}{ll} \text{(a)} & \hat F_1(t, z_1, z_2, z_3, \dot z_1, \dot z_2, \dot z_3) = 0, \\ \text{(b)} & \hat F_2(t, z_1, z_2, z_3) = 0, \end{array} \tag{12}$$

with $\hat F_1 = Z_1^T F$, where $(z_1, z_2, z_3) \in \mathbb{R}^d \times \mathbb{R}^{n-a-d} \times \mathbb{R}^a$ is a suitable splitting of the unknown z. Part 4 of Hypothesis 1 garantees that equation (12b) can be solved for z_3 according to $z_3 = \mathcal{R}(t, z_1, z_2)$. Eliminating z_3 and $\dot z_3$ in (12a) with the help of this relation and its derivative then leads to

$$\hat F_1(t, z_1, z_2, \mathcal{R}(t, z_1, z_2), \dot z_1, \dot z_2, \mathcal{R}_t(t, z_1, z_2) + \mathcal{R}_{z_1}(t, z_1, z_2)\dot z_1 + \mathcal{R}_{z_2}(t, z_1, z_2)\dot z_2) = 0.$$

By part 5 of Hypothesis 1 we may assume without loss of generality that this system can (locally) be solved for $\dot z_1$ leading to the system

$$\begin{aligned} \dot z_1 &= \mathcal{L}(t, z_1, z_2, \dot z_2), \\ z_3 &= \mathcal{R}(t, z_1, z_2). \end{aligned} \tag{13}$$

Obviously, in this system, interpreted as a DAE, $z_2 \in C^1(\mathbb{I}, \mathbb{R}^{n-a-d})$ can be chosen arbitrarily (at least when staying in the domain of definition of $\mathcal{R}$ and $\mathcal{L}$), while the resulting system has locally a unique solution for z_1 and z_3, provided a consistent initial condition is given. This means that z_2 can be interpreted as a control. We summarize these observations in the following theorem, see [41, 42].

Theorem 3 *Let* F *in (2) be sufficiently smooth and satisfy Hypothesis 1 with* μ*,* a*,* d*,* v*. Then every sufficiently smooth solution of (5) also solves the reduced problems (12) and (13) consisting of* d *differential and* a *algebraic equations.*

Under some further assumptions, the converse of Theorem 3 holds as well, see again [41, 42].

Theorem 4 *Let F in (2) be sufficiently smooth and satisfy Hypothesis 1 with μ, a, d, v and $\mu+1$ (replacing μ), a, d, v. Let $z^0_{\mu+1} \in \mathbb{L}_{\mu+1}$ be given and let the parameterization of $\mathbb{L}_{\mu+1}$ include $\dot{x}_2$. Then, for every function $z_2 \in C^1(\mathbb{I}, \mathbb{R}^{n-a-d})$ with $z_2(t_0) = z_{2,0}$, $\dot{z}_2(t_0) = \dot{z}_{2,0}$, the reduced DAEs (12) and (13) have unique solutions z_1 and z_3 satisfying $z_1(t_0) = z_{1,0}$. Moreover, the so obtained function $z = (z_1, z_2, z_3)$ locally solves the original problem.*

The quantity v in Theorem 3, which has not been addressed yet, measures the number of equations in the original system that give rise to trivial equations $0 = 0$, i. e., it counts the number of redundancies in the system. Together with a and d it gives a complete classification of the m equations into d differential equations, a algebraic equations and v trivial equations. Of course, trivial equations can be simply removed without altering the solution set.

If the variable z is a combined vector of states and controls, then, since (12) consists of original variables, these can again be split into parts stemming from x and from u. It has been shown in [41, 43], see also [42], how this system then can be treated in the control context concerning solvability, regularizability, and model consistency.

Theorem 5 *Suppose that the control problem (2) in the form (5) satisfies Hypothesis 1 with μ, a, d, v and assume that $d + a = n$. Then there (locally) exists a state feedback $u = K(t, x)$ satisfying the initial condition*

$$u(\underline{t}) = \underline{u} = K(\underline{t}, \underline{x}), \quad \dot{u}(\underline{t}) = \underline{\dot{u}} = K_t(\underline{t}, \underline{x}) + K_x(\underline{t}, \underline{x})\underline{\dot{x}} \tag{14}$$

such that the resulting closed loop reduced problem is regular and strangeness-free.

Corollary 6 *Suppose that the control problem (2) in the form (5) satisfies Hypothesis 1 with μ, a, d, v and with $\mu+1$ (replacing μ), a, d, v and assume that $d+a = n$. Furthermore, let u be a control in the sense that u and $\dot{u}$ can be chosen as part of the parametrization of $\mathbb{L}_{\mu+1}$ at $z^0_{\mu+1} \in \mathbb{L}_{\mu+1}$. Let $u = K(t, x)$ be a state feedback which satisfies the initial conditions (14) and yields a regular and strangeness-free closed loop reduced system. Then, the closed loop reduced problem has a unique solution satisfying the initial values given by $z^0_{\mu+1}$. Moreover, this solution locally solves the closed loop problem*

$$F(t, x, K(t, x), \dot{x}) = 0.$$

Similar results are given in [41, 42] for output control problems, but in this paper we restrict our attention to optimal control problems without output equation.

Note that due to the application of the implicit function theorem, the above results are only valid locally. For linear problems, the local results automatically hold globally. In the general nonlinear case, however, when global constructions are needed as in the present case, a more detailed analysis is required. We will present corresponding results in Section 3.4.

2.3 Optimization in Banach spaces

We recall some results from general optimization theory, see, e.g., [63]. For this consider the optimization problem

$$\mathcal{J}(z) = \min! \tag{15}$$

subject to the constraint

$$\mathcal{F}(z) = 0, \tag{16}$$

where

$$\mathcal{J} : \mathbb{D} \to \mathbb{R}, \quad \mathcal{F} : \mathbb{D} \to \mathbb{Y}, \quad \mathbb{D} \subseteq \mathbb{Z} \text{ open},$$

with real Banach spaces $\mathbb{Z}, \mathbb{Y}$. Let, furthermore,

$$z^* \in \mathbb{M} = \{z \in \mathbb{D} \mid \mathcal{F}(z) = 0\}.$$

Then we have the following theorem which is due to [51].

Theorem 7 *Let $\mathcal{J}$ be Fréchet differentiable in z^* and let $\mathcal{F}$ be a submersion in z^*, i.e., let $\mathcal{F}$ be Fréchet differentiable in a neighborhood of z^* with Fréchet derivative $D\mathcal{F}(z^*) : \mathbb{Z} \to \mathbb{Y}$ surjective and* kernel $D\mathcal{F}(z^*)$ *continuously projectable.*

If z^* is a local minimum of (15), then there exists a unique Λ in the dual space $\mathbb{Y}^*$ of $\mathbb{Y}$ with

$$D\mathcal{J}(z^*)\Delta z + \Lambda(D\mathcal{F}(z^*)\Delta z) = 0 \quad \text{for all } \Delta z \in \mathbb{Z}. \tag{17}$$

The functional Λ in Theorem 7 is called the *Lagrange multiplier* associated with the constraint (16).

In general we are interested in function representations of the Lagrange multiplier functional Λ. Such representations are obtained by the following theorem.

Theorem 8 *Let $\mathbb{Y} = C^0(\mathbb{I}, \mathbb{R}^m) \times \mathbb{V}$ with a vector space $\mathbb{V} \subseteq \mathbb{R}^m$ and let $(\lambda, \gamma) \in \mathbb{Y}$. Then*

$$\Lambda(g, r) = \int_{\underline{t}}^{\overline{t}} \lambda(t)^T g(t)\ dt + \gamma^T r$$

defines a linear form $\Lambda \in \mathbb{Y}^*$, which conversely uniquely determines $(\lambda, \gamma) \in \mathbb{Y}$.

A sufficient condition that guarantees that also the minimum is unique is given by the following theorem, which, e.g., covers linear-quadratic control problems with positive definite reduced Hessian.

Theorem 9 *Suppose that $\mathcal{F} : \mathbb{Z} \to \mathbb{Y}$ is affine linear and that $\mathcal{J} : \mathbb{Z} \to \mathbb{R}$ is strictly convex on $\mathbb{M}$, i.e.,*

$$\mathcal{J}(\alpha z_1 + (1-\alpha)z_2) < \alpha\mathcal{J}(z_1) + (1-\alpha)\mathcal{J}(z_2)$$
$$\text{for all } z_1, z_2 \in \mathbb{M} \text{ with } z_1 \neq z_2 \text{ and for all } \alpha \in (0,1),$$

then the optimization problem (15) subject to (16) has a unique minimum.

For the analysis and solution of optimal control problems subject to constraints given by differential-algebraic equations we will have to carry out changes of variables and linear or nonlinear feedbacks. To see how these effect the minimization problem, consider a local diffeomorphism $\phi : \mathbb{Z} \to \mathbb{Z}$ in a neighborhood of $\tilde{z}^*$ with $z^* = \phi(\tilde{z}^*)$. If we transform the optimization problem (15) and the constraint (16) to the new variable $\tilde{z}$ via $z = \phi(\tilde{z})$, then we obtain the transformed optimization problem

$$\tilde{\mathcal{J}}(\tilde{z}) = \min!$$

subject to the constraint

$$\tilde{\mathcal{F}}(\tilde{z}) = 0,$$

where

$$\tilde{\mathcal{J}}(\tilde{z}) = \mathcal{J}(\phi(\tilde{z})), \quad \tilde{\mathcal{F}}(\tilde{z}) = \mathcal{F}(\phi(\tilde{z})).$$

If $\tilde{z}^*$ satisfies the necessary condition (17) in the form

$$D\tilde{\mathcal{J}}(\tilde{z}^*)\Delta\tilde{z} + \Lambda(D\tilde{\mathcal{F}}(\tilde{z}^*)\Delta\tilde{z}) = 0 \;\text{ for all } \Delta\tilde{z} \in \mathbb{Z},$$

then

$$D\mathcal{J}(\phi(\tilde{z}^*))D\phi(\tilde{z}^*)\Delta\tilde{z} + \Lambda(D\mathcal{F}(\phi(\tilde{z}^*))D\phi(\tilde{z}^*)\Delta\tilde{z}) = 0 \quad \text{for all } \Delta\tilde{z} \in \mathbb{Z}.$$

With $\Delta z = D\phi(\tilde{z}^*)\Delta\tilde{z}$ we then have

$$D\mathcal{J}(z^*)\Delta z + \Lambda(D\mathcal{F}(z^*)\Delta z) = 0 \;\text{ for all } \Delta z \in \mathbb{Z},$$

and thus, z satisfies the necessary condition (17) for the optimization problem (15) subject to (16).

If the controls are restricted by $u(t) \in \mathcal{U}$, then we must admit bang-bang controls. In this case the optimal solution is obtained via versions of the Pontryagin maximum principle.

For the *Bolza problem* to determine $(\overline{t}, x, u) \in \mathbb{R} \times C^0(\mathbb{I}, \mathbb{R}^n) \times L^c_\infty(\mathbb{I}, \mathbb{R}^l)$ as a solution of

$$\mathcal{J}(x,u) = \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, x(t), u(t))\; dt = \min! \tag{18}$$

subject to

$$\begin{aligned} x(t) &= \underline{x} + \textstyle\int_{\underline{t}}^{t} f(s, x(s), u(s))\; ds \\ 0 &= h(\overline{t}, x(\overline{t})),\;\; h \in C(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n), \\ u(t) &\in \mathcal{U} \subset \mathbb{R}^l \;\text{ for all } t \in \mathbb{I}, \end{aligned} \tag{19}$$

one has the following theorem.

Theorem 10 *If $(\overline{t}, x^*, u^*)$ is a local solution of the Bolza problem (18) subject to (19), then there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_n$, which do not all vanish simultaneously, $\alpha_0 \geq 0$, and a multiplier $\lambda \in C^0(\mathbb{I}, \mathbb{R}^n)$ such that, with $H(t, x, u, \lambda, \alpha_0) = \lambda^T f(t, x, u) - \alpha_0 \mathcal{K}(t, x, u)$ and $\alpha = (\alpha_1, \cdots, \alpha_n)^T$,*

$$\begin{aligned} H(t, x^*(t), u^*(t), \lambda(t), \alpha_0) &= \max\nolimits_{u \in \mathcal{U}} H(t, x^*(t), u, \lambda(t), \alpha_0), \\ \dot{\lambda}(t) &= -\nabla_x H(t, x^*(t), u^*(t), \lambda(t), \alpha_0), \\ \dot{x}^*(t) &= \nabla_\lambda H(t, x^*(t), u^*(t), \lambda(t), \alpha_0), \\ \lambda(\overline{t}) &= -\alpha^T \nabla_{x(\overline{t})} h(\overline{t}, x^*(\overline{t})), \end{aligned} \tag{20}$$

for all $t \in \mathbb{I}$, where u^* is continuous.

3 Necessary conditions

In this section we will derive necessary optimality conditions for the minimization of (1) subject to (2). We first start with the special case of a linear-quadratic optimal control problem and then extend the results to the general case.

3.1 Linear-quadratic optimal control problems

The linear-quadratic optimal control problem for differential-algebraic equations has been well studied for constant coefficient systems, see [53] and the references therein, and variable coefficient problems in [39]. These results are based on the idea to first use index reduction and feedback regularization to transform the problem into a regular, strangeness-free problem and then to use the analysis for this case. Recently, in [1, 2, 4, 49] this problem was studied again in a different setting for some restricted classes of linear DAEs with small index.

Here we study the general case of unstructured linear-quadratic optimal problems, i.e., we study the cost functional

$$\mathcal{J}(x,u) = \frac{1}{2}x(\overline{t})^T M x(\overline{t}) + \frac{1}{2}\int_{\underline{t}}^{\overline{t}} (x^T W x + 2x^T S u + u^T R u)\ dt, \tag{21}$$

with $W \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $S \in C^0(\mathbb{I}, \mathbb{R}^{n,l})$, $R \in C^0(\mathbb{I}, \mathbb{R}^{l,l})$, and we assume, furthermore, that W and R are pointwise symmetric and also that $M \in \mathbb{R}^{n,n}$ is symmetric. As constraint we consider the initial value problem for a general linear differential-algebraic equations with variable coefficients of the form

$$E\dot{x} = Ax + Bu + f, \quad x(\underline{t}) = \underline{x}, \tag{22}$$

with $E \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $A \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $B \in C^0(\mathbb{I}, \mathbb{R}^{n,l})$, $f \in C^0(\mathbb{I}, \mathbb{R}^n)$, and $\underline{x} \in \mathbb{R}^n$. Our goal is to determine optimal controls $u \in \mathbb{U} = C^0(\mathbb{I}, \mathbb{R}^l)$.

We could discuss a more general situation and allow the coefficient functions E, A to be non-square but, as it has been shown in [41], this case can always be transformed to the regular square case. In order to avoid unnecessary technicalities we therefore restrict ourselves here to the regular square case.

For a better readability of the more complicated formulas we omit here and in the following the argument t of the involved coefficient functions.

In the case of linear ordinary differential equations, corresponding to the case $E(t) = I$ in (22), the initial value problem has a unique solution $x \in C^1(\mathbb{I}, \mathbb{R}^n)$ for every $u \in \mathbb{U}$, every $f \in C^0(\mathbb{I}, \mathbb{R}^n)$, and every initial value $\underline{x} \in \mathbb{R}^n$. In contrast to this, in the case of differential-algebraic equations, where $E(t)$ may be singular, the equation is not necessarily (uniquely) solvable for any $u \in \mathbb{U}$ and also the initial conditons may be restricted, see [42]. Furthermore, it will be necessary to consider solutions $x \in \mathbb{X}$, where $\mathbb{X}$ usually is a larger space than $C^1(\mathbb{I}, \mathbb{R}^n)$.

For our analysis we consider the system in behavior form (5)

$$\mathcal{E}\dot{z} = \mathcal{A}z + f, \tag{23}$$

with

$$\mathcal{E} = [\,E\ \ 0\,], \quad \mathcal{A} = [\,A\ \ B\,].$$

Its associated derivative array is given by

$$M_\ell(t)\dot{z}_\ell = N_\ell(t)z_\ell + g_\ell(t), \tag{24}$$

where

$$\begin{aligned}
(M_\ell)_{i,j} &= \tbinom{i}{j}\mathcal{E}^{(i-j)} - \tbinom{i}{j+1}\mathcal{A}^{(i-j-1)}, \; i,j = 0,\dots,\ell,\\
(N_\ell)_{i,j} &= \begin{cases} \mathcal{A}^{(i)} & \text{for } i = 0,\dots,\ell, \; j = 0,\\ 0 & \text{otherwise,}\end{cases}\\
(z_\ell)_j &= z^{(j)}, \; j = 0,\dots,\ell,\\
(g_\ell)_i &= f^{(i)}, \; i = 0,\dots,\ell.
\end{aligned}$$

We assume that this system has a well defined strangeness index μ according to Hypothesis 1 and, furthermore, as we have already stated before, we assume that there is no consistency condition for the inhomogeneities, i.e., $v = 0$. Under the assumptions of Theorem 4, the initial value problem (22) is equivalent (in the sense that it has the same set of solutions) to the *reduced system*

$$\hat{E}\dot{x} = \hat{A}x + \hat{B}u + \hat{f}, \quad x(\underline{t}) = \underline{x}, \tag{25}$$

where

$$\hat{E} = \begin{bmatrix} \hat{E}_1 \\ 0 \end{bmatrix}, \quad \hat{A} = \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix}, \quad \hat{f} = \begin{bmatrix} \hat{f}_1 \\ \hat{f}_2 \end{bmatrix} \tag{26}$$

with

$$\begin{aligned}
\hat{E}_1 = Z_1^T E, \quad [\, \hat{A}_1 \;\; \hat{B}_1 \,] &= Z_1^T [\, A \;\; B \,], & \hat{f}_1 &= Z_1^T f,\\
[\, \hat{A}_2 \;\; \hat{B}_2 \,] &= Z_2^T N_\mu [\, I_{n+l} \;\; 0 \;\; \cdots \;\; 0]^T, & \hat{f}_2 &= Z_2^T g_\mu.
\end{aligned}$$

By construction, the reduced system (25) is strangeness-free. In particular, the matrix function $\hat{E}_1$ has full row rank d and $[\, \hat{A}_2 T_2' \;\; \hat{B}_2 \,]$ has full row rank a with a matrix function T_2' satisfying $\hat{E}_1 T_2' = 0$ and $T_2'^T T_2' = I_a$. Due to the fact that the solution set has not changed, one can consider the minimization of (21) subject to (25) instead of (22). Unfortunately, (25) still may not be solvable for all $u \in \mathbb{U}$. But, since $[\, \hat{A}_2 T_2' \;\; \hat{B}_2 \,]$ has full row rank, it follows from Theorem 5, see also [43], that there exists a linear feedback

$$u = Kx + w, \tag{27}$$

with $K \in C^0(\mathbb{I}, \mathbb{R}^{l,n})$ such that in the closed loop system

$$\hat{E}\dot{x} = (\hat{A} + \hat{B}K)x + \hat{B}w + \hat{f}, \quad x(\underline{t}) = \underline{x}, \tag{28}$$

the matrix function $(\hat{A}_2 + \hat{B}_2 K)T_2'$ is pointwise nonsingular, implying that the DAE in (28) is regular and strangeness-free for every given $w \in \mathbb{U}$.

If we insert the feedback (27) in (25), then we obtain an optimization problem for the variables x, w instead of x, u, and according to the analysis in Section 2.3, these problems and the solutions are directly transferable to each other. For this reason we may in the following assume w.l.o.g. that the differential-algebraic system (22) is regular and strangeness-free as a free system without control, i.e., when $u = 0$.

Under these assumptions it is then known, see, e.g., [42], that there exist $P \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$ and $Q \in C^1(\mathbb{I}, \mathbb{R}^{n,n})$ pointwise orthogonal such that

$$\begin{aligned}
&\tilde{E} = PEQ = \begin{bmatrix} E_{1,1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{A} = PAQ - PE\dot{Q} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix},\\
&\tilde{B} = PB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad \tilde{f} = Pf = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad x = Q\tilde{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \underline{x} = Q\underline{\tilde{x}} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix},
\end{aligned} \tag{29}$$

with $E_{1,1} \in C(\mathbb{I}, \mathbb{R}^{d,d})$ and $A_{2,2} \in C(\mathbb{I}, \mathbb{R}^{a,a})$ pointwise nonsingular. To get solvability of (22) for arbitrary $u \in \mathbb{U}$ and $f \in C(\mathbb{I}, \mathbb{R}^n)$, in view of

$$E\dot{x} = EE^+E\dot{x} = E\tfrac{d}{dt}(E^+Ex) - E\tfrac{d}{dt}(E^+E)x,$$

we have to interpret (22) as

$$E\tfrac{d}{dt}(E^+Ex) = (A + E\tfrac{d}{dt}(E^+E))x + Bu + f, \quad (E^+Ex)(\underline{t}) = \underline{x}, \tag{30}$$

which allows the larger solution space, see [38],

$$\mathbb{X} = C^1_{E^+E}(\mathbb{I}, \mathbb{R}^n) = \left\{x \in C^0(\mathbb{I}, \mathbb{R}^n) \mid E^+Ex \in C^1(\mathbb{I}, \mathbb{R}^n)\right\} \tag{31}$$

equipped with the norm

$$\|x\|_{\mathbb{X}} = \|x\|_{C^0} + \|\tfrac{d}{dt}(E^+Ex)\|_{C^0}. \tag{32}$$

One should note that the choice of the initial value $\underline{x}$ is restricted by the requirement in (30).

Following [38], we can use in (16) the constraint function

$$\mathcal{F} : \mathbb{X} \to \mathbb{Y} = C^0(\mathbb{I}, \mathbb{R}^n) \times \operatorname{range} E^+(\underline{t})E(\underline{t})$$

given by

$$\mathcal{F}(x) = \left(E\tfrac{d}{dt}(E^+Ex) - (A + E\tfrac{d}{dt}(E^+E))x - Bu - f, (E^+Ex)(\underline{t}) - \underline{x}\right).$$

Then from (30) we obtain

$$\begin{aligned}&PEQQ^T\tfrac{d}{dt}(QQ^TE^+P^TPEQQ^Tx)\\ &\quad= \left(PAQ + PEQQ^T\tfrac{d}{dt}(QQ^TE^+P^TPEQQ^T)Q\right)Q^Tx + PBu + Pf,\end{aligned}$$

or equivalently

$$\begin{aligned}&\tilde{E}Q^T\tfrac{d}{dt}(Q\tilde{E}^+\tilde{E}\tilde{x})\\ &\quad= \left(\tilde{A} + PP^T\tilde{E}Q^T\dot{Q} + \tilde{E}Q^T\tfrac{d}{dt}(Q\tilde{E}^+\tilde{E}Q^T)Q\right)\tilde{x} + \tilde{B}u + \tilde{f}.\end{aligned}$$

Using the product rule and cancelling equal terms on both sides we obtain

$$\begin{aligned}&\tilde{E}Q^TQ\tfrac{d}{dt}(\tilde{E}^+\tilde{E}\tilde{x})\\ &\quad= \left(\tilde{A} + \tilde{E}Q^T\dot{Q} + \tilde{E}\tfrac{d}{dt}(\tilde{E}^+\tilde{E}) + \tilde{E}\tilde{E}^+\tilde{E}\dot{Q}^TQ\right)\tilde{x} + \tilde{B}u + \tilde{f}.\end{aligned}$$

Since by definition $\tilde{E}\tilde{E}^+\tilde{E} = \tilde{E}$ and $\dot{Q}^TQ + Q^T\dot{Q} = 0$, we then obtain

$$\tilde{E}\tfrac{d}{dt}(\tilde{E}^+\tilde{E}\tilde{x}) = \left(\tilde{A} + \tilde{E}\tfrac{d}{dt}(\tilde{E}^+\tilde{E})\right)\tilde{x} + \tilde{B}u + \tilde{f}, \quad (\tilde{E}^+\tilde{E}\tilde{x})(\underline{t}) = \underline{\tilde{x}}, \tag{33}$$

i.e., (30) transforms covariantly with pointwise orthogonal P and Q. If we partition P and Q conformably to (29) as

$$P = \begin{bmatrix} Z'^T \\ Z^T \end{bmatrix}, \quad Q = [\,T' \;\; T\,],$$

then $Z^T E = 0$, $ET = 0$, and we can write (33) as

$$\begin{bmatrix} E_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad \begin{bmatrix} x_1(\underline{t}) \\ 0 \end{bmatrix} = \begin{bmatrix} \underline{x}_1 \\ 0 \end{bmatrix}.$$

Since $A_{2,2}$ is pointwise nonsingular, this system is uniquely solvable for arbitrary continuous functions u, f_1, and f_2, and for any $\underline{x}_1$, with solution components satisfying

$$x_1 \in C^1(\mathbb{I}, \mathbb{R}^d), \quad x_2 \in C^0(\mathbb{I}, \mathbb{R}^a)$$

such that

$$x = Q\tilde{x} = [\, T' \;\, T \,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{X}.$$

In particular, this construction defines a solution operator of the form

$$\mathcal{S} : \mathbb{U} \times \mathbb{Y} \to \mathbb{X}, \;\; (u, f, \underline{x}) \mapsto x. \tag{34}$$

The Fréchet derivative $D\mathcal{F}(z)$ of $\mathcal{F}$ at $z \in \mathbb{Z} = \mathbb{X} \times \mathbb{U}$ is given by

$$D\mathcal{F}(z)\Delta z = \left(E\tfrac{d}{dt}(E^+ E\Delta x) - (A + E\tfrac{d}{dt}(E^+ E))\Delta x - B\Delta u, (E^+ E\Delta x)(\underline{t}) \right).$$

For $(g, r) \in \mathbb{Y}$, the equation $D\mathcal{F}(z) = (g, r)$ then takes the form

$$E\tfrac{d}{dt}(E^+ E\Delta x) - (A + E\tfrac{d}{dt}(E^+ E))\Delta x - B\Delta u = g, \quad (E^+ E\Delta x)(\underline{t}) = r.$$

A possible solution is given by $u = 0$ and $\Delta x = \mathcal{S}(0, g, r)$, hence $D\mathcal{F}(z)$ is surjective. Moreover, the kernel is given by

$$\begin{aligned} &\operatorname{kernel}(D\mathcal{F}(z)) \\ &\quad = \{(\Delta x, \Delta u) \mid E\tfrac{d}{dt}(E^+ E\Delta x) - (A + E\tfrac{d}{dt}(E^+ E))\Delta x - B\Delta u = 0, \;\, (E^+ E\Delta x)(\underline{t}) = 0\} \\ &\quad = \{(\Delta x, \Delta u) \mid \Delta x = \mathcal{S}(\Delta u, 0, 0), \;\, \Delta u \in \mathbb{U}\} \subseteq \mathbb{X} \times \mathbb{U}. \end{aligned}$$

Observe that $\operatorname{kernel}(D\mathcal{F}(z))$ is parameterized with respect to Δu and that

$$\mathcal{P}(z) = \mathcal{P}(x, u) = (\mathcal{S}(u, 0, 0), u)$$

defines a projection $\mathcal{P} : \mathbb{Z} \to \mathbb{Z}$ onto $\operatorname{kernel}(D\mathcal{F}(z))$. Here,

$$\|\, (\mathcal{S}(u, 0, 0), u) \,\|_{\mathbb{Z}} = \|\mathcal{S}(u, 0, 0)\|_{\mathbb{X}} + \|u\|_{\mathbb{U}}, \;\text{ and } \|\mathcal{S}(u, 0, 0)\|_{\mathbb{X}} = \|x\|_{\mathbb{X}},$$

where x is the solution of the homogeneous problem

$$E\tfrac{d}{dt}(E^+ Ex) - (A + E\tfrac{d}{dt}(E^+ E))x - Bu = 0, \;\, (E^+ Ex)(\underline{t}) = 0. \tag{35}$$

Replacing again $x = Q\tilde{x}$ as in (29), we can write (35) as

$$\begin{bmatrix} E_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \quad x_1(\underline{t}) = 0,$$

or equivalently

$$E_{1,1}\dot{x}_1 = (A_{1,1} - A_{1,2}A_{2,2}^{-1}A_{2,1})x_1 + (B_1 - A_{1,2}A_{2,2}^{-1}B_2)u, \quad x_1(\underline{t}) = 0, \tag{36}$$

$$x_2 = -A_{2,2}^{-1}(A_{2,1}x_1 + B_2 u). \tag{37}$$

The variation of the constant formula for the ODE in (36) yields the estimate $\|x_1\|_{C^0} + \|\dot{x}_1\|_{C^0} \leq c_1\|u\|_{\mathbb{U}}$, with a constant c_1, and thus $\|x_2\|_{C^0} \leq c_2\|u\|_{\mathbb{U}}$ with a constant c_2. Altogether, using (32) we then get the estimate

$$\begin{aligned}\|x\|_{\mathbb{X}} &= \|x\|_{C^0} + \|\tfrac{d}{dt}(E^+Ex)\|_{C^0} = \|Q\tilde{x}\|_{C^0} + \|\tfrac{d}{dt}(E^+ET'x_1)\|_{C^0} \\ &= \|Q\tilde{x}\|_{C^0} + \|\tfrac{d}{dt}(E^+ET')x_1 + (E^+ET')\dot{x}_1)\|_{C^0} \leq c_3\|u\|_{\mathbb{U}},\end{aligned}$$

with a constant c_3. With this we have shown that $\mathcal{P}$ is continuous and thus kernel$(D\mathcal{F}(z))$ is continuously projectable. Hence, we can apply Theorem 7 and obtain the existence of a unique Lagrange multiplier $\Lambda \in \mathbb{Y}^*$. To determine Λ, we make the ansatz

$$\Lambda(g, r) = \int_{\underline{t}}^{\overline{t}} \lambda^T g \; dt + \gamma^T r. \tag{38}$$

Using the cost function (21) we have

$$D\mathcal{J}(z)\Delta z = x(\overline{t})^T M \Delta x(\overline{t}) + \int_{\underline{t}}^{\overline{t}} (x^T W \Delta x + x^T S \Delta u + u^T S^T \Delta x + u^T R \Delta u) \; dt,$$

and in a local minimum $z = (x, u)$ we obtain that for all $(\Delta x, \Delta u) \in \mathbb{X} \times \mathbb{U}$ the relationship

$$\begin{aligned}0 = x(\overline{t})^T M \Delta x(\overline{t}) &+ \int_{\underline{t}}^{\overline{t}} (x^T W \Delta x + x^T S \Delta u + u^T S^T \Delta x + u^T R \Delta u) \; dt \\ &+ \int_{\underline{t}}^{\overline{t}} \lambda^T \left(E\tfrac{d}{dt}(E^+E\Delta x) - (A + E\tfrac{d}{dt}(E^+E))\Delta x - B\Delta u \right) \; dt + \gamma^T(E^+E\Delta x)(\underline{t}))\end{aligned} \tag{39}$$

has to hold. If $\lambda \in C^1_{E^+E}(\mathbb{I}, \mathbb{R}^n)$, then, using the fact that $E = EE^+E = (EE^+)^T E$, we have by partial integration

$$\begin{aligned}\int_{\underline{t}}^{\overline{t}} \lambda^T E \tfrac{d}{dt}(E^+E\Delta x) \; dt &= \int_{\underline{t}}^{\overline{t}} \lambda^T (EE^+)^T E \tfrac{d}{dt}(E^+E\Delta x) \; dt \\ &= \int_{\underline{t}}^{\overline{t}} (EE^+\lambda)^T E \tfrac{d}{dt}(E^+E\Delta x) \; dt \\ &= \lambda^T EE^+E\Delta x \Big|_{\underline{t}}^{\overline{t}} - \int_{\underline{t}}^{\overline{t}} \tfrac{d}{dt}\left[(EE^+\lambda)^T E\right](E^+E\Delta x) \; dt \\ &= \lambda^T E\Delta x \Big|_{\underline{t}}^{\overline{t}} - \int_{\underline{t}}^{\overline{t}} \left[\tfrac{d}{dt}(EE^+\lambda)^T E + (EE^+\lambda)^T \dot{E}\right](E^+E\Delta x) \; dt \\ &= \lambda^T E\Delta x \Big|_{\underline{t}}^{\overline{t}} - \int_{\underline{t}}^{\overline{t}} \left[\tfrac{d}{dt}(EE^+\lambda)^T E\Delta x + (EE^+\lambda)^T \dot{E}E^+E\Delta x\right] \; dt.\end{aligned}$$

Therefore, we can rewrite (39) as

$$\begin{aligned}0 = &\int_{\underline{t}}^{\overline{t}} \left(x^T W + u^T S^T - \tfrac{d}{dt}(EE^+\lambda)^T E - (EE^+\lambda)^T \dot{E}E^+E - \lambda^T A - \lambda^T E\tfrac{d}{dt}(E^+E) \right) \Delta x \; dt \\ &+ \int_{\underline{t}}^{\overline{t}} (x^T S + u^T R - \lambda^T B)\Delta u \; dt \\ &+ x(\overline{t})^T M \Delta x(\overline{t}) + \lambda^T(\overline{t})E(\overline{t})\Delta x(\overline{t}) - \lambda^T(\underline{t})E(\underline{t})\Delta x(\underline{t}) + \gamma^T(E^+E\Delta x)(\underline{t}).\end{aligned}$$

If we first choose $\Delta x = 0$ and vary over all $\Delta u \in \mathbb{U}$, then we obtain the necessary *optimality condition*

$$S^T x + Ru - B^T \lambda = 0. \tag{40}$$

Varying then over all $\Delta x \in \mathbb{X}$ with $\Delta x(\underline{t}) = \Delta x(\overline{t}) = 0$, we obtain the *adjoint equation*

$$Wx + Su - E^T \tfrac{d}{dt}(EE^+\lambda) - E^+E\dot{E}^T EE^+\lambda - A^T\lambda - \tfrac{d}{dt}(E^+E)E^T\lambda = 0. \tag{41}$$

Varying finally over $\Delta x(\underline{t}) \in \mathbb{R}^n$ and $\Delta x(\overline{t}) \in \mathbb{R}^n$, respectively, yields the initial condition

$$(E^+(\underline{t})E(\underline{t}))^T\gamma = E^T(\underline{t})\lambda(\underline{t}), \quad \text{i.e.,} \quad \gamma = E(\underline{t})^T\lambda(\underline{t}) \tag{42}$$

and the end condition

$$Mx(\overline{t}) + E(\overline{t})^T\lambda(\overline{t}) = 0, \tag{43}$$

respectively.

Observe that the condition (43) can only be satisfied when $Mx(\overline{t}) \in \operatorname{cokernel} E(\overline{t})$. This extra requirement for the cost term involving the final state was observed already for constant coefficient systems in [53] and in a different setting in [49]. If this condition on M holds, then from (43) we obtain $\lambda(\overline{t}) = -E^+(\overline{t})^T Mx(\overline{t})$.

Using the identity

$$EE^+\dot{E}E^+E + E\tfrac{d}{dt}(E^+E) = EE^+(\dot{E}E^+E + E\tfrac{d}{dt}(E^+E)) = EE^+\tfrac{d}{dt}(EE^+E) = EE^+\dot{E},$$

we obtain the initial value problem for the adjoint equation in the form

$$E^T\tfrac{d}{dt}(EE^+\lambda) = Wx + Su - (A + EE^+\dot{E})^T\lambda, \quad (EE^+\lambda)(\overline{t}) = -E^+(\overline{t})^T Mx(\overline{t}). \tag{44}$$

As we had to interpret (22) in the form (30) for the correct choice of the spaces, (44) is the correct interpretation of the problem

$$\tfrac{d}{dt}(E^T\lambda) = Wx + Su - A^T\lambda, \quad \lambda(\overline{t}) = -E^+(\overline{t})^T Mx(\overline{t}). \tag{45}$$

Note again that these re-interpretations are not crucial when the coefficient functions are sufficiently smooth. The formulation (45) now suggests the following definition.

Definition 11 *Let (E, A) be a pair of matrix functions with $E \in C^1(\mathbb{I}, \mathbb{R}^{n,n})$ and $A \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$. The pair $(E^T, -(A + \dot{E})^T)$ is called the* adjoint pair *of (E, A).*

The notion "adjoint pair" is not only justified by the above construction but also by the following property.

Theorem 12 *Let (E, A) have the adjoint pair $(E^T, -(A+\dot{E})^T)$. Then $(E^T, -(A+\dot{E})^T)$ has an adjoint pair which is given by (E, A).*

Proof. Obviously we have $E^T \in C^1(\mathbb{I}, \mathbb{R}^{n,n})$ and $-(A + \dot{E})^T \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$. Hence, the pair $(E^T, -(A + \dot{E})^T)$ has an adjoint pair given by $((E^T)^T, -[-(A + \dot{E})^T + \dot{E}^T]^T) = (E, A)$. $\square$

It is possible to show that if a pair of matrix functions has a well-defined differentiation index ν then its adjoint pair also has a well-defined differentiation index ν. Since we do not need this result in the course of this paper we omit a proof of this observation. A more important property of the adjoint pair especially for the treatment of concrete problems is its behavior under equivalence transformations. For this, let $P,Q\in C^1(\mathbb{I},\mathbb{R}^{n,n})$ be pointwise nonsingular and let $\tilde{E}=PEQ$ and $\tilde{A}=PAQ-PE\dot{Q}$. Assuming that (E,A) possesses an adjoint pair, we see that $(\tilde{E},\tilde{A})$ possesses an adjoint pair as well which is given by

$$
\begin{aligned}
&(\tilde{E}^T,-(\tilde{A}+\dot{\tilde{E}})^T)\\
&\quad=(Q^TE^TP^T,-(Q^TA^TP^T-\dot{Q}^TE^TP^T+\dot{Q}^TE^TP^T+Q^T\dot{E}^TP^T+Q^TE^T\dot{P}^T))\\
&\quad=(Q^TE^TP^T,-Q^T(A+\dot{E})^TP^T-Q^TE^T\dot{P}^T).
\end{aligned}
$$

The latter representation then states that the adjoint pair of the transformed pair is equivalent to the adjoint pair of the original pair. Hence, we are always allowed to transform a given pair into a suitable form before we build the adjoint pair.

Returning to the adjoint equation and the optimality condition, we will study the action of the special equivalence transformtions of (29) on these equations. Using that $(EE^+)^T=EE^+$, we obtain for (44) the transformed system

$$
\begin{aligned}
&QE^TP^TP\tfrac{d}{dt}(P^TPEQQ^TE^+P^TP\lambda)\\
&\quad=Q^TWQQ^Tx+Q^TSu-(Q^TA^TP^T+Q^T\dot{E}P^TPEQQ^TE^+P^T)P\lambda.
\end{aligned}
$$

Setting

$$
\tilde{W}=Q^TWQ,\quad \tilde{S}=Q^TS,\quad \tilde{\lambda}=P\lambda,\quad \tilde{M}=Q(\bar{t})^TMQ(\bar{t}),
$$

we obtain

$$
\begin{aligned}
&\tilde{E}P\tfrac{d}{dt}(P^T\tilde{E}\tilde{E}^+\tilde{\lambda})\\
&\quad=\tilde{W}\tilde{x}+\tilde{S}u-\Big(\tilde{A}^T+\dot{Q}^TQ\tilde{E}^T+Q^T(Q\tilde{E}^T\dot{P}+Q\dot{\tilde{E}}^TP+\dot{Q}\tilde{E}^TP)P^T\tilde{E}\tilde{E}^+\Big)\tilde{\lambda}
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
&\tilde{E}P\dot{P}^T\tilde{E}\tilde{E}^+\tilde{\lambda}+\tilde{E}\tfrac{d}{dt}(\tilde{E}\tilde{E}^+\tilde{\lambda})\\
&\quad=\tilde{W}\tilde{x}+\tilde{S}u-\Big(\tilde{A}^T+\dot{Q}^TQ\tilde{E}^T+\tilde{E}^T\dot{P}P^T\tilde{E}\tilde{E}^++\dot{\tilde{E}}^T\tilde{E}\tilde{E}^++Q^T\dot{Q}\tilde{E}^T\tilde{E}\tilde{E}^+\Big)\tilde{\lambda}.
\end{aligned}
$$

Using the orthogonality of P,Q, which implies that $\dot{Q}^TQ+Q\dot{Q}=0$ and $\dot{P}^TP+P\dot{P}=0$, we obtain

$$
\tilde{E}\tfrac{d}{dt}(\tilde{E}\tilde{E}^+\tilde{\lambda})=\tilde{W}\tilde{x}+\tilde{S}u-(\tilde{A}+\tilde{E}\tilde{E}^+\dot{\tilde{E}})^T\tilde{\lambda}.
$$

For the initial condition we obtain accordingly

$$
\begin{aligned}
(\tilde{E}\tilde{E}^+\tilde{\lambda})(\bar{t})&=(PEQQ^TE^+P^TP\lambda)(\bar{t})=(PEE^+\lambda)(\bar{t})\\
&=-P(\bar{t})E^+(\bar{t})^TQ(\bar{t})Q(\bar{t})^TMQ(\bar{t})Q(\bar{t})^Tx(\bar{t})=-\tilde{E}^+(\bar{t})^T\tilde{M}\tilde{x}(\bar{t}).
\end{aligned}
$$

Thus, we have shown that (44) transforms covariantly and that we may consider (44) in the condensed form associated with (29). Setting (with comformable partitioning)

$$
\tilde{\lambda}=\begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix},\quad \tilde{W}=\begin{bmatrix}W_{1,1} & W_{1,2}\\ W_{2,1} & W_{2,2}\end{bmatrix},\quad \tilde{S}=\begin{bmatrix}S_1\\ S_2\end{bmatrix},\quad \tilde{M}=\begin{bmatrix}M_{1,1} & M_{1,2}\\ M_{2,1} & M_{2,2}\end{bmatrix},\tag{46}
$$

we obtain the system

$$\begin{aligned}
E_{1,1}^T\dot\lambda_1 &= W_{1,1}x_1 + W_{1,2}x_2 + S_1 u - (A_{1,1}+\dot E_{1,1})^T\lambda_1 - A_{2,1}^T\lambda_2,\\
&\qquad\qquad \lambda_1(\bar t) = -E_{1,1}^{-T}(\bar t)(M_{1,1}x_1(\bar t) + M_{1,2}x_2(\bar t)),\\
0 &= W_{2,1}x_1 + W_{2,2}x_2 + S_2 u - A_{1,2}^T\lambda_1 - A_{2,2}^T\lambda_2.
\end{aligned}$$

We immediately see that as a differential-algebraic equation in λ this system is again regular and strangeness-free. In particular, since $A_{2,2}$ is pointwise nonsingular, this system yields a unique solution $\lambda \in C^1_{EE+}(\mathbb{I},\mathbb{R}^n)$ for every $(x,u)\in\mathbb{Z}$.

If $(x,u)\in\mathbb{Z}$ is a local minimum, then from (43) and (44) we can determine Lagrange multipliers $\lambda \in C^1_{EE+}(\mathbb{I},\mathbb{R}^n)$ and $\gamma \in \operatorname{cokernel} E(\bar t)$. It is, however, not clear, whether this λ also satisfies the optimality condition (40).

Suppose this is not the case, i.e., for the given (x,u,λ) we have

$$S^T x + Ru - B^T\lambda \neq 0. \tag{47}$$

Then there exists $\Delta u \in \mathbb{U}$ with

$$\int_{\underline{t}}^{\bar t} (x^T S + u^T R - \lambda^T B)\Delta u \; dt \neq 0.$$

Using this Δu, we have a unique $\Delta x \in \mathbb{X}$ satisfying

$$E\tfrac{d}{dt}(E^+E\Delta x) = (A + E\tfrac{d}{dt}(E^+E))\Delta x + B\Delta u, \quad (E^+E\Delta x)(\underline{t}) = 0,$$

which implies that for $z + \varepsilon\Delta z = (x,u) + \varepsilon(\Delta x, \Delta u)$, we have $\mathcal{F}(z+\varepsilon\Delta z) = 0$ and

$$\begin{aligned}
&\mathcal{J}(z+\epsilon\Delta z) - \mathcal{J}(z)\\
&\quad = \varepsilon\Big[x(\bar t)^T M\Delta x(\bar t) + \int_{\underline{t}}^{\bar t}(x^T W\Delta x + x^T S\Delta u + u^T S^T\Delta x + u^T R\Delta u)\; dt\Big] + \mathcal{O}(\varepsilon^2)\\
&\quad = \varepsilon\Big[x(\bar t)^T M\Delta x(\bar t) + \int_{\underline{t}}^{\bar t}\big((x^T W + u^T S^T)\Delta x + (x^T S + u^T R)\Delta u\big)\; dt\Big] + \mathcal{O}(\varepsilon^2)\\
&\quad = \varepsilon\Big[x(\bar t)^T M\Delta x(\bar t) + \int_{\underline{t}}^{\bar t}\Big(\tfrac{d}{dt}(EE^+\lambda)^T E + \lambda^T(A + EE^+\dot E)\Big)\Delta x\; dt\\
&\qquad\quad + \int_{\underline{t}}^{\bar t}(x^T S + u^T R)\Delta u\; dt\Big] + \mathcal{O}(\varepsilon^2)\\
&\quad = \varepsilon\Big[x(\bar t)^T M\Delta x(\bar t) + (\lambda^T E\Delta x)\Big|_{\underline{t}}^{\bar t} - \int_{\underline{t}}^{\bar t}(EE^+\lambda)^T\dot E E^+E\Delta x\; dt - \int_{\underline{t}}^{\bar t}\lambda^T E\tfrac{d}{dt}(E^+E\Delta x)\; dt\\
&\qquad\quad + \int_{\underline{t}}^{\bar t}\lambda^T(A + EE^+\dot E)\Delta x\; dt + \int_{\underline{t}}^{\bar t}(x^T S + u^T R)\Delta u\; dt\Big] + \mathcal{O}(\varepsilon^2)\\
&\quad = \varepsilon\Big[\int_{\underline{t}}^{\bar t}\lambda^T\Big((A + EE^+\dot E)\Delta x - E\tfrac{d}{dt}(E^+E\Delta x) - EE^+\dot E E^+E\Delta x\Big)\; dt\\
&\qquad\quad + \int_{\underline{t}}^{\bar t}(x^T S + u^T R)\Delta u\; dt\Big] + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

$$= \varepsilon\Big[\int_{\underline{t}}^{\overline{t}} \lambda^T \Big((A + EE^+\dot{E}) - (A + E\tfrac{d}{dt}(E^+E)) - EE^+\dot{E}E^+E\Big)\,\Delta x\; dt$$

$$+ \int_{\underline{t}}^{\overline{t}} (x^T S + u^T R - \lambda^T B)\Delta u\; dt\Big] + \mathcal{O}(\varepsilon^2)$$

$$= \varepsilon\Big[\int_{\underline{t}}^{\overline{t}} (x^T S + u^T R - \lambda^T B)\Delta u\; dt\Big] + \mathcal{O}(\varepsilon^2)$$

Since ε can take any positive and negative value, it follows that z was not a local minimum. Hence, the vector λ defined by (44) must satisfy (40).

It thus follows that the functional that is defined via (38), (44) and $\gamma = E(\underline{t})^T\lambda(\underline{t})$ as in (42) has the property (17) and is, therefore, the desired Lagrange multiplier. Furthermore, it is then clear that $(z, \lambda) = (x, u, \lambda)$ is a local minimum of the unconstrained optimization problem

$$\begin{aligned} \hat{\mathcal{J}}(z,\lambda) &= \mathcal{J}(z) + \Lambda(\mathcal{F}(z)) \\ &= \tfrac{1}{2}x(\overline{t})^T M x(\overline{t}) + \tfrac{1}{2}\int_{\underline{t}}^{\overline{t}}(x^T W x + 2x^T S u + u^T R u)\; dt \\ &\quad + \int_{\underline{t}}^{\overline{t}} \lambda^T (E(\tfrac{d}{dt}(E^+Ex) - (A + E\tfrac{d}{dt}(E^+E))x - Bu - f)\; dt \\ &\quad + \gamma^T\,((E^+Ex)(\underline{t}) - \underline{x}) = \min! \end{aligned} \tag{48}$$

In summary, we have proved the following theorem.

Theorem 13 *Consider the optimal control problem (21) subject to (22) with a consistent initial condition. Suppose that (22) is strangeness-free as a behavior system and that* $Mx(\overline{t}) \in$ cokernel $E(\overline{t})$.

If $(x, u) \in \mathbb{X} \times \mathbb{U}$ *is a solution to this optimal control problem, then there exists a Lagrange multiplier function* $\lambda \in C^1_{E^+E}(\mathbb{I}, \mathbb{R}^n)$*, such that* (x, λ, u) *satisfy the* optimality boundary value problem

$$\begin{array}{ll} \text{(a)} & E\tfrac{d}{dt}(E^+Ex) = (A + E\tfrac{d}{dt}(E^+E))x + Bu + f,\ (E^+Ex)(\underline{t}) = \underline{x}, \\ \text{(b)} & E^T\tfrac{d}{dt}(EE^+\lambda) = Wx + Su - (A + EE^+\dot{E})^T\lambda,\ (EE^+\lambda)(\overline{t}) = -E^+(\overline{t})^T Mx(\overline{t}), \\ \text{(c)} & 0 = S^T x + Ru - B^T\lambda. \end{array} \tag{49}$$

It should be noted again that the assumption in Theorem 13 that the system (22) is regular and strangeness-free in the behavior formulation is not a restriction, since we can always assume that we have already obtained the reduced system (25) which has this property. The same is true for the requirement of consistent initial conditions, which are easily obtained from the reduced system. The third assumption can be easily guaranteed as well, since usually the weight on the final state is something that is chosen independently of the model.

An important question for the numerical computation of optimal controls is when the optimality system (49) is regular and strangeness-free and whether the strangeness index of (49) is related to the strangeness index of the original system. For other index concepts like the tractability index this question has been discussed in [5, 3, 6, 49].

Theorem 14 *The DAE in (49) is regular and strangeness-free if and only if*

$$\hat{R} = \begin{bmatrix} 0 & A_{2,2} & B_2 \\ A_{2,2}^T & W_{2,2} & S_2 \\ B_2^T & S_2^T & R \end{bmatrix} \tag{50}$$

is pointwise nonsingular, where we used the notation of (29).

Proof. Consider the reduced system (25) associated with the DAE (22) and derive the boundary value problem (49) from this reduced system. If we carry out the change of basis with orthogonal transformations leading to the normal form (29), then we obtain the transformed boundary value problem

$$\begin{array}{ll} \text{(a)} & E_{1,1}\dot{x}_1 = A_{1,1}x_1 + A_{1,2}x_2 + B_1u + f_1,\ x_1(\underline{t}) = \underline{x}_1 \\ \text{(b)} & \qquad 0 = A_{2,1}x_1 + A_{2,2}x_2 + B_2u + f_2, \\ \text{(c)} & E_{1,1}^T\dot{\lambda}_1 = W_{1,1}x_1 + W_{1,2}x_2 + S_1u - (A_{1,1} + \dot{E}_{1,1})^T\lambda_1 - A_{2,1}^T\lambda_2, \\ & \qquad \lambda_1(\overline{t}) = -E_{1,1}(\overline{t})^{-T}M_{1,1}x_1(\overline{t}), \\ \text{(d)} & \qquad 0 = W_{2,1}x_1 + W_{2,2}x_2 + S_2u - A_{1,2}^T\lambda_1 - A_{2,2}^T\lambda_2, \\ \text{(e)} & \qquad 0 = S_1^Tx_1 + S_2^Tx_2 + Ru - B_1^T\lambda_1 - B_2^T\lambda_2. \end{array} \tag{51}$$

We can rewrite (51) in a symmetrized way as

$$\begin{aligned} &\left[\begin{array}{cc|ccc} 0 & E_{1,1} & 0 & 0 & 0 \\ -E_{1,1}^T & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{c} -\dot{\lambda}_1 \\ \dot{x}_1 \\ \hline -\dot{\lambda}_2 \\ \dot{x}_2 \\ \dot{u} \end{array}\right] \\ &= \left[\begin{array}{cc|ccc} 0 & A_{1,1} & 0 & A_{1,2} & B_1 \\ (A_{1,1} + \dot{E}_{1,1})^T & W_{1,1} & A_{2,1}^T & W_{2,1}^T & S_1 \\ \hline 0 & A_{2,1} & 0 & A_{2,2} & B_2 \\ A_{1,2}^T & W_{2,1} & A_{2,2}^T & W_{2,2} & S_2 \\ B_1^T & S_1^T & B_2^T & S_2^T & R \end{array}\right] \left[\begin{array}{c} -\lambda_1 \\ x_1 \\ \hline -\lambda_2 \\ x_2 \\ u \end{array}\right] + \left[\begin{array}{c} f_1 \\ 0 \\ \hline f_2 \\ 0 \\ 0 \end{array}\right]. \end{aligned} \tag{52}$$

Obviously this DAE is regular and strangeness-free if and only if the symmetric matrix function $\hat{R}$ is pointwise nonsingular. □

If (22) itself is regular and strangeness-free as a free system with $u = 0$, then $A_{2,2}$ is pointwise nonsingular. In our analysis we have shown that this property can always be achieved, but note that we do not need that $A_{2,2}$ is pointwise nonsingular to obtain a regular and strangeness-free optimality system (49).

On the other hand for $\hat{R}$ to be pointwise nonsingular, it is clearly necessary that $[\,A_{2,2}\ B_2\,]$ has pointwise full row rank. This condition is equivalent to the condition that the behavior system (23) belonging to the reduced problem satisfies Hypothesis 1 with $\mu = 0$ and $v = 0$, see [43] for a detailed discussion of this issue and also for an extension of these results to the case of control systems with output equations.

Example 15 An example of a control problem of the form (22) that is not directly strangeness-free in the behavior setting is discussed in [2, p. 50]. This linear-quadratic control problem has the coefficients

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad f = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad R = 1,$$

and the initial condition $x_1(0) = \alpha,\ x_2(0) = 0$. A possible reduced system (25) is given by

$$\hat{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \hat{A} = A, \quad \hat{B} = B.$$

Observe that the corresponding free system of this reduced problem (i.e. with $u = 0$) itself is regular and strangeness-free. It follows that the adjoint equation and the optimality condition are given by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \dot{\lambda}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix},$$

$$0 = -\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} + u,$$

respectively, with the end condition $\lambda_1(\overline{t}) = -x_1(\overline{t})$.

We obtain that the matrix function $\hat{R}$ in (50) given by

$$\hat{R} = \left[\begin{array}{cc|cc|c} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 0 & 1 \end{array}\right]$$

is pointwise nonsingular, and hence the boundary value problem (49) is regular and strangeness-free. Moreover, it has a unique solution which is given by

$$x_1 = \alpha(1 - \frac{t}{2+\overline{t}}), \quad x_2 = \lambda_3 = 0, \quad x_3 = u = -\lambda_2 = -\frac{\alpha}{2+\overline{t}}, \quad \lambda_1 = -\frac{2\alpha}{2+\overline{t}}.$$

Example 16 In [49] the optimal control problem to minimize

$$\mathcal{J}(x,u) = \int_0^{\overline{t}} (x_1(t)^2 + u(t)^2)\ dt$$

subject to

$$\frac{d}{dt}\left(\begin{bmatrix} 0 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad x_2(0) = x_{2,0}$$

is discussed. Obviously, x_1 does not enter the DAE and therefore rather plays the role of a control than of a state. Consequently, the corresponding free system is not regular. Rewriting the system as

$$\begin{bmatrix} 0 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad x_2(0) = x_{2,0},$$

and analyzing this system in our described framework, we first of all observe that this system possesses a strangeness index and that it is even regular and strangeness-free as a behavior system. A possible reduced system (25) is given by

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \quad x_2(0) = x_{2,0}.$$

Also here the corresponding free system is not regular although it is strangeness-free. Moreover, we can read off

$$\hat{R} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

which is obviously pointwise nonsingular. Hence, the boundary value problem (49) is regular and strangeness-free.

In view of Definition 11 together with (45) one may be tempted to drop the assumptions of Theorem 13 and to consider directly the *formal optimality boundary value problem* given by

$$\begin{array}{ll} \text{(a)} & E\dot{x} = Ax + Bu + f, \ x(\underline{t}) = \underline{x} \\ \text{(b)} & \frac{d}{dt}(E^T\lambda) = Wx + Su - A^T\lambda, \ (E^T\lambda)(\overline{t}) = -Mx(\overline{t}), \\ \text{(b)} & 0 = S^T x + Ru - B^T\lambda. \end{array} \tag{53}$$

But it was already observed in [2, 49, 53] that it is in general not correct to just consider this system. First of all, as we have shown, the cost matrix M for the final state has to be in the correct cokernel, since otherwise the initial value problem may not be solvable due to a wrong number of conditions. An example for this is given in [2, 49]. A further difficulty arises from the fact that the formal adjoint equation (53b) may be a high index equation in the variable λ and thus extra differentiability conditions may arise which may not be satisfied, see the following example.

Example 17 Consider the problem

$$\mathcal{J}(x,u) = \frac{1}{2}\int_0^1 (x_1(t)^2 + u(t)^2)dt = \min!$$

subject to the differential-algebraic system

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

The reduced system (25) in this case is the purely algebraic equation

$$0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u + \begin{bmatrix} f_1 + \dot{f}_2 \\ f_2 \end{bmatrix}.$$

The associated adjoint equation (44) is then

$$0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},$$

and no initial conditions are necessary. The optimality condition (40) is given by

$$0 = u - \lambda_1.$$

A simple calculation yields the optimal solution

$$x_1 = u = \lambda_1 = -\tfrac{1}{2}(f_1 + \dot{f}_2), \quad x_2 = -f_2, \quad \lambda_2 = 0.$$

If, however, we consider the formal adjoint equation (53b) given by

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \dot\lambda_1 \\ \dot\lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, \quad \lambda_1(1) = 0$$

together with the optimality condition (53c), then we obtain that

$$x_1 = u = \lambda_1 = -\tfrac{1}{2}(f_1 + \dot f_2), \quad x_2 = -f_2, \quad \lambda_2 = -\tfrac{1}{2}(\dot f_1 + \ddot f_2)$$

without using the initial condition $\lambda_1(1) = 0$. Depending on the data, this initial condition may be consistent or not. In view of the correct solution it is obvious that this initial condition should not be present. But this cannot be seen from (53). Moreover, the determination of λ_2 requires more smoothness of the inhomogeneity than in (49).

As we have demonstrated by Example 17, difficulties may arise by working with the formal adjoint equations. In particular, they may not be solvable due to additional initial conditions or due to lack of smoothness. If, however, the cost functional is positive semidefinite, then one can show that any solution of the formal optimality system yields a minimum and thus constitutes a sufficient condition. This was, e.g., shown for ODE optimal control in [15], for linear constant coefficient DAEs in [53], and in a specific setting for linear DAEs with variable coefficients in [2]. The general result is given by the following theorem.

Theorem 18 *Consider the optimal control problem (21) subject to (22) with a consistent initial condition and suppose that in the cost functional (21) we have that*

$$\begin{bmatrix} W & S \\ S^T & R \end{bmatrix}, \; M$$

are (pointwise) positive semidefinite. If (x^*, u^*, λ) satisfies the formal optimality system (53), then for any (x, u) satisfying (22) we have

$$\mathcal{J}(x, u) \geq \mathcal{J}(x^*, u^*).$$

Proof. We consider the function

$$\Phi(s) = \mathcal{J}((1-s)x^* + sx, (1-s)u^* + su)$$

and show that $\Phi(s)$ has a minimum at $s = 0$. We have

$$\begin{aligned}
\Phi(s) = \frac{1}{2} \int_{\underline{t}}^{\overline{t}} \Bigg(& (1-s)^2 \begin{bmatrix} x^* \\ u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x^* \\ u^* \end{bmatrix} \\
& + 2s(1-s) \begin{bmatrix} x^* \\ u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
& + s^2 \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \Bigg) \, dt \\
& + \frac{1}{2} \Big((1-s)^2 x^{*T} M x^* + 2s(1-s) x^{*T} M x + s^2 x^T M x \Big) \Big|_{t=\overline{t}},
\end{aligned}$$

and

$$\frac{d}{ds}\Phi(0) = \int_{\underline{t}}^{\overline{t}} \Bigg(\begin{bmatrix} x^* \\ u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - \begin{bmatrix} x^* \\ u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x^* \\ u^* \end{bmatrix} \Bigg) \, dt + \Big(x^{*T} M x - x^{*T} M x^* \Big)\Big|_{t=\overline{t}}.$$

If we consider (53b) for (x^*, u^*) and multiply from the left by x^{*T}, then we obtain

$$-x^{*T} E^T \dot{\lambda} - x^{*T} \dot{E}^T \lambda + x^{*T} W x^* + x^{*T} S u^* - x^{*T} A^T \lambda = 0.$$

Inserting the transpose of (53a) yields

$$-x^{*T} E^T \dot{\lambda} - x^{*T} \dot{E}^T \lambda + x^{*T} W x^* + x^{*T} S u^* - \dot{x}^{*T} E^T \lambda + u^{*T} B^T \lambda + f^T \lambda = 0.$$

Finally, inserting (53c) gives

$$\begin{aligned} -\frac{d}{dt} \left(x^{*T} E^T \lambda \right) &+ x^{*T} W x^* + 2 x^{*T} S u^* + u^{*T} R u^* + f^T \lambda \\ &= -\frac{d}{dt} \left(x^{*T} E^T \lambda \right) + \begin{bmatrix} x^* \\ u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x^* \\ u^* \end{bmatrix} + f^T \lambda = 0. \end{aligned}$$

Analogously, we obtain for (x, u) the equation

$$-\frac{d}{dt} \left(x^T E^T \lambda \right) + \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x^* \\ u^* \end{bmatrix} + f^T \lambda = 0.$$

Thus, we obtain that

$$\begin{aligned} \frac{d}{ds}\Phi(0) &= \int_{\underline{t}}^{\overline{t}} \left(\tfrac{d}{dt} \left(x^T E^T \lambda \right) - \tfrac{d}{dt} \left(x^{*T} E^T \lambda \right) \right) \, dt + \Big(x^{*T} M (x - x^*) \Big)\Big|_{t=\overline{t}} \\ &= \Big((x - x^*)^T E^T \lambda \Big)\Big|_{\underline{t}}^{\overline{t}} + \Big((x - x^*)^T M x^* \Big)\Big|_{t=\overline{t}} = 0, \end{aligned}$$

since $x(\underline{t}) = x^*(\underline{t})$ and $(E^T \lambda)(\overline{t}) = -M x^*(\overline{t})$. Due to the positive semidefiniteness of the cost functional we have

$$\frac{d^2}{ds^2}\Phi(0) = \int_{\underline{t}}^{\overline{t}} \begin{bmatrix} x - x^* \\ u - u^* \end{bmatrix}^T \begin{bmatrix} W & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x - x^* \\ u - u^* \end{bmatrix} \, dt + \Big((x - x^*)^T M (x - x^*) \Big)\Big|_{t=\overline{t}} \geq 0$$

and this implies that Φ has a minimum at $s = 0$, which may, however, not be unique. □

We can summarize the results of this section as follows. The necessary optimality condition for the optimal control problem (21) subject to (22) is given by (49) and not by the formal optimality system (53). If, however, (53) has a solution, then it corresponds to a minimum of the optimal control problem. If no index reduction is performed, then a necessary condition for the DAE in (21) to be regular and strangeness-free is that the DAE (22) itself is regular and strangeness-free as a behavior system.

3.2 Differential-algebraic Riccati equations

One of the classical approaches to solve boundary value problems arising in the linear-quadratic optimal control problem of ordinary differential equations is the use of Riccati differential equations. This approach has also been studied in the case of differential-algebraic equations, see [7, 39, 53], and it has been observed in [39] that the Riccati approach is not always possible. If, however, some further conditions hold, then the Riccati approach can be carried out. For linear differential-algebraic systems with constant coefficients this has been studied in detail in [53] and for variable coefficient systems in a different setting in [24] for special cases. We present here the general case.

Let us first consider the optimality boundary value problem (49) in its symmetrized normal form (52). If $\hat{R}$ is pointwise nonsingular, then

$$\begin{bmatrix} -\lambda_2 \\ x_2 \\ u \end{bmatrix} = -\hat{R}^{-1}\left(\begin{bmatrix} 0 & A_{2,1} \\ A_{1,2}^T & W_{2,1} \\ B_1^T & S_1^T \end{bmatrix} \begin{bmatrix} -\lambda_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} f_2 \\ 0 \\ 0 \end{bmatrix} \right). \tag{54}$$

The remaining equations can be written as

$$\begin{bmatrix} E_{1,1}\dot{x}_1 \\ \frac{d}{dt}((-E_{1,1}^T)(-\lambda_1)) \end{bmatrix} = \begin{bmatrix} 0 & A_{1,1} \\ A_{1,1}^T & W_{1,1} \end{bmatrix} \begin{bmatrix} -\lambda_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} 0 & A_{1,2} & B_1 \\ A_{2,1}^T & W_{2,1}^T & S_1 \end{bmatrix} \begin{bmatrix} -\lambda_2 \\ x_2 \\ u \end{bmatrix} + \begin{bmatrix} f_1 \\ 0 \end{bmatrix}. \tag{55}$$

Inserting (54) and defining

$$\begin{array}{ll}
\text{(a)} & F_1 = E_{1,1}^{-1}\left(A_{1,1} - [\,0\ \ A_{1,2}\ \ B_1\,]\hat{R}^{-1}[\,A_{2,1}^T\ \ W_{2,1}^T\ \ S_1\,]^T\right), \\
\text{(b)} & G_1 = E_{1,1}^{-1}[\,0\ \ A_{1,2}\ \ B_1\,]\hat{R}^{-1}[\,0\ \ A_{1,2}\ \ B_1\,]^T E_{1,1}^{-T}, \\
\text{(c)} & H_1 = W_{1,1} - [\,A_{2,1}^T\ \ W_{2,1}^T\ \ S_1\,]\hat{R}^{-1}[\,A_{2,1}^T\ \ W_{2,1}^T\ \ S_1\,]^T, \\
\text{(d)} & g_1 = E_{1,1}^{-1}\left(f_1 - [\,0\ \ A_{1,2}\ \ B_1\,]\hat{R}^{-1}[\,f_2^T\ \ 0\ \ 0\,]^T\right), \\
\text{(e)} & h_1 = -[\,A_{2,1}^T\ \ W_{2,1}^T\ \ S_1\,]\hat{R}^{-1}[\,f_2^T\ \ 0\ \ 0\,]^T,
\end{array}$$

we obtain the boundary value problem with *Hamiltonian structure* given by

$$\begin{array}{ll}
\text{(a)} & \dot{x}_1 = F_1 x_1 + G_1(E_{1,1}^T\lambda_1) + g_1,\ \ x_1(\underline{t}) = \underline{x}_1, \\
\text{(b)} & \frac{d}{dt}(E_{1,1}^T\lambda_1) = H_1 x_1 - F_1^T(E_{1,1}^T\lambda_1) + h_1,\ \ (E_{1,1}^T\lambda_1)(\overline{t}) = -M_{1,1}x_1(\overline{t}).
\end{array} \tag{56}$$

Making the ansatz

$$E_{1,1}^T\lambda_1 = X_{1,1}x_1 + v_1, \tag{57}$$

and using its derivative

$$\tfrac{d}{dt}(E_{1,1}^T\lambda_1) = \dot{X}_{1,1}x_1 + X_{1,1}\dot{x}_1 + \dot{v}_1,$$

the Hamiltonian boundary value problem (56) yields

$$\dot{X}_{1,1}x_1 + X_{1,1}(F_1x_1 + G_1(X_{1,1}x_1 + v_1) + g_1) + \dot{v}_1 = H_1x_1 - F_1^T(X_{1,1}x_1 + v_1) + h_1,$$

or

$$\begin{aligned}
&\left(\dot{X}_{1,1} + X_{1,1}F_1 + F_1^TX_{1,1} + X_{1,1}G_1X_{1,1} - H_1\right)x_1 \\
&\qquad + \left(\dot{v}_1 + X_{1,1}G_1v_1 + F_1^Tv_1 + X_{1,1}g_1 - h_1\right) = 0.
\end{aligned}$$

Thus, we can solve the two initial value problems

$$\dot{X}_{1,1} + X_{1,1}F_1 + F_1^T X_{1,1} + X_{1,1}G_1X_{1,1} - H_1 = 0,\ \ X_{1,1}(\bar{t}) = -M_{1,1}, \tag{58}$$

and

$$\dot{v}_1 + X_{1,1}G_1v_1 + F_1^T v_1 + X_{1,1}g_1 - h_1 = 0,\ \ v_1(\bar{t}) = 0, \tag{59}$$

to obtain $X_{1,1}$ and v_1 and to decouple the solution to (56).

In this way we have obtained a Riccati approach for the dynamic part of the system. Ideally, however, we would like to have a Riccati approach directly for the boundary value problem associated with (33), (44), and (40) in the original data, without carrying out the change of bases and going to normal form. If we make a similar ansatz for the general situation, i.e., $\lambda = Xx + v$, then we face the problem that neither the whole x nor the whole λ may be differentiable. To accomodate for the appropriate solution spaces, we therefore make the modified ansatz

$$\begin{array}{ll} \text{(a)} & \lambda = XEx + v = XEE^+Ex + v, \\ \text{(b)} & \frac{d}{dt}(EE^+\lambda) = \frac{d}{dt}(EE^+X)Ex + (EE^+X)\dot{E}E^+Ex \\ & \qquad\qquad + (EE^+X)E\frac{d}{dt}(E^+Ex) + \frac{d}{dt}(E^+Ev), \end{array} \tag{60}$$

where

$$X \in C^1_{EE^+}(\mathbb{I}, \mathbb{R}^{n,n}), \quad v \in C^1_{EE^+}(\mathbb{I}, \mathbb{R}^n). \tag{61}$$

In this way we have obtained an ansatz that fits to the solution spaces for x and λ. The disadvantage of this approach, however, is that $X(I - EE^+)$ now can be chosen arbitrarily. Using again the transformation to normal form (29) and that

$$P\lambda = PXP^TPEQQ^Tx + Pv,$$

we obtain

$$\tilde{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \tilde{X}\tilde{E}\tilde{x} + \tilde{v},$$

with

$$\tilde{X} = PXP^T = \begin{bmatrix} \tilde{X}_{1,1} & \tilde{X}_{1,2} \\ \tilde{X}_{2,1} & \tilde{X}_{2,2} \end{bmatrix}, \quad \tilde{v} = \begin{bmatrix} \tilde{v}_1 \\ \tilde{v}_2 \end{bmatrix}.$$

Comparing with (57) we obtain

$$X_{1,1} = E_{1,1}^T\tilde{X}_{1,1}E_{1,1}, \quad v_1 = E_{1,1}^T\tilde{v}_1.$$

In particular, we obtain that $\tilde{X}_{1,1}$ and $\tilde{v}_1$ are continuously differentiable under the assumption that (58) is solvable on the interval $\mathbb{I}$. Furthermore, $\tilde{X}_{1,1}$ is pointwise symmetric. From (54) we then obtain

$$\lambda_2 = [\,I\ \ 0\ \ 0\,]\hat{R}^{-1}\left(\begin{bmatrix} 0 & A_{2,1} \\ A_{1,2}^T & W_{2,1} \\ B_1^T & S_1^T \end{bmatrix}\begin{bmatrix} -(\tilde{X}_{1,1}E_{1,1}x_1 + \tilde{v}_1) \\ x_1 \end{bmatrix} + \begin{bmatrix} f_2 \\ 0 \\ 0 \end{bmatrix}\right) = \tilde{X}_{2,1}E_{1,1}x_1 + \tilde{v}_2,$$

with

$$\tilde{X}_{2,1}E_{1,1} = [\,I\ \ 0\ \ 0\,]\hat{R}^{-1}\begin{bmatrix} 0 & A_{2,1} \\ A_{1,2}^T & W_{2,1} \\ B_1^T & S_1^T \end{bmatrix}\begin{bmatrix} -\tilde{X}_{1,1}E_{1,1} \\ I \end{bmatrix},$$

and

$$\tilde{v}_2 = [\,I\ \ 0\ \ 0\,]\hat{R}^{-1}\begin{bmatrix} f_2 \\ -A_{1,2}^T\tilde{v}_1 \\ -B_1^T\tilde{v}_1 \end{bmatrix}.$$

If we assume that R itself is pointwise nonsingular (which corresponds to the assumption that all controls are weighted in the cost functional), then from (40) we obtain that

$$u = R^{-1}(B^T\lambda - S^T x)$$

and thus from (44) and (35) we obtain

$$\begin{aligned} E^T\tfrac{d}{dt}&(EE^+X)Ex + E^T(EE^+X)\dot{E}E^+Ex \\ &+ E^T(EE^+X)\left(Ax + E\tfrac{d}{dt}(E^+E)x + BR^{-1}B^T(XEx+v) - BR^{-1}S^Tx + f\right) \\ &+ E^T\tfrac{d}{dt}(EE^+v) \\ =&\ Wx + SR^{-1}B^T(XEx+v) - SR^{-1}S^Tx - (A+EE^+\dot{E})^T(XEx+v), \end{aligned}$$

or

$$\begin{aligned} \Big(&E^T\tfrac{d}{dt}(EE^+X)E + E^T(EE^+X)\dot{E}E^+E + E^T(EE^+X)E\tfrac{d}{dt}(E^+E) \\ &+ \dot{E}^T(EE^+X)EE^+E + E^TXA + E^TXBR^{-1}B^TXE + A^TXE \\ &- E^TXBR^{-1}S^T - SR^{-1}B^TXE + SR^{-1}S^T - W\Big)x \\ &+ \left(\tfrac{d}{dt}(EE^+v) + E^TXBR^{-1}B^Tv + E^TXf - SR^{-1}B^Tv + A^Tv + \dot{E}^T(EE^+v)\right) = 0. \end{aligned}$$

Introducing the notation

$$\begin{array}{ll} \text{(a)} & F = A - BR^{-1}S^T, \\ \text{(b)} & G = BR^{-1}B^T, \\ \text{(c)} & H = W - SR^{-1}S^T, \end{array} \tag{62}$$

we obtain

$$\begin{aligned} &\left(\tfrac{d}{dt}(E^T(EE^+X)E(E^+E)) + E^TXF + F^TXE + E^TXGXE - H\right)x \\ &\qquad + \left(E^T\tfrac{d}{dt}(EE^+v) + \dot{E}^T(EE^+v) + E^TXGv + F^Tv + E^TXf\right) = 0, \end{aligned}$$

which yields the two initial value problems

$$\tfrac{d}{dt}(E^TXE) + E^TXF + F^TXE + E^TXGXE - H = 0,\ (E^TXE)(\overline{t}) = -M, \tag{63}$$

and

$$\tfrac{d}{dt}(E^Tv) + E^TXGv + F^Tv + E^TXf = 0,\ (E^Tv)(\overline{t}) = 0. \tag{64}$$

Note that we must have $M = E(\overline{t})^T\tilde{M}E(\overline{t})$ with suitable $\tilde{M}$ and $H = E^T\tilde{H}E$ with suitable $\tilde{H}$ as necessary condition for the solvability of (63). Note also that (as already in the case of ODEs) the optimality boundary value problem (49) may be solvable, whereas (63) does not allow for a solution on the whole interval $\mathbb{I}$.

The analysis in this section shows that we can obtain a Riccati approach if the system (22) is strangeness-free in the behavior setting and R is invertible.

3.3 A modified cost functional

In the previous sections we have derived necessary conditions for linear-quadratic control problem and studied how these can be solved. In particular, we have seen that extra conditions on the cost functional have to hold for the optimality system or the associated Riccati equation to have a solution.

Since the cost functional is often a matter of choice one could modify it to reduce the requirements. A simple modification is the following cost functional, see e.g. [35] in the case of constant coefficients,

$$\mathcal{J}(x,u) = \frac{1}{2}x(\overline{t})^T \tilde{M} x(\overline{t}) + \frac{1}{2}\int_{\underline{t}}^{\overline{t}} (x^T \tilde{W} x + 2x^T \tilde{S} u + u^T R u)\ dt, \tag{65}$$

with $\tilde{M} = E(\overline{t})^T M E(\overline{t})$, $\tilde{W} = E^T W E$, and $\tilde{S} = E^T S$.

Assuming again that the original system (22) is strangeness-free as a behavior system, the same analysis as before leads to the modified optimality boundary value problem

$$\begin{array}{ll} \text{(a)} & E\frac{d}{dt}(E^+Ex) = (A + E\frac{d}{dt}(E^+E))x + Bu + f,\ (E^+Ex)(\underline{t}) = \underline{x}, \\ \text{(b)} & E^T\frac{d}{dt}(EE^+\lambda) = E^TWEx + E^TSu - (A + EE^+\dot{E})\lambda, \\ & \qquad\qquad (EE^+\lambda)(\overline{t}) = -E^+(\overline{t})^T E(\overline{t})^T M E(\overline{t}) x(\overline{t}), \\ \text{(c)} & 0 = S^TEx + Ru - B^T\lambda. \end{array} \tag{66}$$

Considering the conditions that guarantee that the optimality system is again strangeness-free, we obtain the following corollary.

Corollary 19 *Consider the optimal control problem to minimize (65) subject to (22) and assume that (22) is strangeness-free as a free system (with $u = 0$). Then the optimality system (66) is strangeness-free if and only if R is nonsingular.*

Proof. Consider the system (22) in the normal form (29). By assumption, we have that $A_{2,2}$ is invertible and in the transformed cost functional (46) we obtain that $\tilde{S}_2 = 0$ and $\tilde{W}_{2,2} = 0$. The modified matrix $\hat{R}$ then takes the form

$$\hat{R} = \begin{bmatrix} 0 & A_{2,2} & B_2 \\ A_{2,2}^T & 0 & 0 \\ B_2^T & 0 & R \end{bmatrix},$$

which is clearly pointwise nonsingular if and only if R is nonsingular. □

The Riccati approach also changes when we use the modified cost functional. In particular, we obtain

$$\begin{array}{ll} \text{(a)} & \tilde{F} = A - BR^{-1}\tilde{S}^T, \\ \text{(b)} & \tilde{G} = G = BR^{-1}B^T, \\ \text{(b)} & \tilde{H} = \tilde{W} - \tilde{S}R^{-1}\tilde{S}^T, \end{array} \tag{67}$$

In this case one obtains the two initial value problems

$$\tfrac{d}{dt}(E^TXE) + E^TX\tilde{F} + \tilde{F}^TXE + E^TX\tilde{G}XE - \tilde{H} = 0,\ (E^TXE)(\overline{t}) = -\tilde{M}, \tag{68}$$

and

$$\tfrac{d}{dt}(E^Tv) + E^TX\tilde{G}v + \tilde{F}^Tv + E^TXf = 0,\ (E^Tv)(\overline{t}) = 0. \tag{69}$$

Observe that the necessary conditions for solvability as stated in the end of Section 3.2 for (63) are now trivially fulfilled.

3.4 General nonlinear problems

In this section we discuss the general nonlinear optimal control problem to minimize (1) subject to (2). We assume that all describing functions are sufficiently smooth and that the system described in the behavior setting (5) satisfies Hypothesis 1 with $v = 0$.

Let $z \in C^0(\mathbb{I}, \mathbb{R}^{n+l})$ be a potential candidate for a minimum of (1) subject to (2), (3). In particular, let z be part of a (continuous) path

$$(t, z(t), \mathcal{P}(t)) \in \mathbb{L}_{\mu+1} \text{ for all } t \in \mathbb{I}, \tag{70}$$

cp. Theorem 4.34 of [42]. Due to Hypothesis 1 there exist

$$Z_2 \in C^0(\mathbb{I}, \mathbb{R}^{(\mu+1)n,a}), \quad T_2 \in C^0(\mathbb{I}, \mathbb{R}^{n+l,n+l-a}), \quad Z_1 \in C^0(\mathbb{I}, \mathbb{R}^{n,d}),$$

with the described properties. Let

$$Z_2' \in C^0(\mathbb{I}, \mathbb{R}^{(\mu+1)n,(\mu+1)n-a}), \quad T_2' \in C^0(\mathbb{I}, \mathbb{R}^{n+l,a}), \quad Z_1' \in C^0(\mathbb{I}, \mathbb{R}^{n,n-d}),$$

be such that

$$[\, Z_2' \;\; Z_2 \,], \quad [\, T_2' \;\; T_2 \,], \quad [\, Z_1' \;\; Z_1 \,]$$

are pointwise orthogonal. Furthermore, there exist

$$T_1 \in C^0(\mathbb{I}, \mathbb{R}^{(\mu+1)(n+l),(\mu+1)l+a}), \quad T_1' \in C^0(\mathbb{I}, \mathbb{R}^{(\mu+1)(n+l),(\mu+1)n-a})$$

such that

$$[\, T_1' \;\; T_1 \,]$$

is pointwise orthogonal and

$$Z_2'(t)^T F_{\mu;\dot{z},\ldots,z^{(\mu+1)}}(t, z(t), \mathcal{P}(t)) T_1(t) = 0 \text{ for all } t \in \mathbb{I}.$$

If we define a function $\mathcal{H}$ via

$$\mathcal{H}(t, z, p, \phi) = \begin{bmatrix} F_\mu(t, z, p) + Z_2(t)\phi \\ T_1(t)^T(p - \mathcal{P}(t)) \end{bmatrix}, \tag{71}$$

then

$$\begin{array}{ll} \text{(a)} & \mathcal{H}(t, z(t), \mathcal{P}(t), 0) = 0, \\ \text{(b)} & \mathcal{H}_{p,\phi}(t, z(t), \mathcal{P}(t), 0) = \begin{bmatrix} F_{\mu;\dot{z},\ldots,z^{(\mu+1)}}(t, z(t), \mathcal{P}(t)) & Z_2(t) \\ T_1(t)^T & 0 \end{bmatrix}. \end{array} \tag{72}$$

By construction $\mathcal{H}_{p,\phi}(t, z(t), \mathcal{P}(t), 0)$ is nonsingular for all $t \in \mathbb{I}$ and thus we can locally solve for p and ϕ as

$$\phi = \hat{F}_2(t, z), \quad p = \hat{\mathcal{P}}(t, z).$$

We have, in particular, that

$$\hat{F}_2(t, z(t)) = 0, \quad \hat{\mathcal{P}}(t, z(t)) = \mathcal{P}(t)$$

and

$$F_\mu(t, z(t), \hat{\mathcal{P}}(t, z)) + Z_2(t)\hat{F}_2(t, z) = 0 \text{ for all } (t, z),$$

and hence

$$F_{\mu;z} + F_{\mu;\dot{z},\ldots,z^{(\mu+1)}}\hat{\mathcal{P}}_z + Z_2\hat{F}_{2;z} = 0,$$

which implies

$$\hat{F}_{2;z}(t,z(t)) = -Z_2(t)^T F_{\mu;z}(t,z(t),\mathcal{P}(t)),$$

i.e., $\hat{F}_{2;z}$ has full row rank along $(t,z(t))$. The equation

$$\hat{F}_2(t,z) = 0 \tag{73}$$

thus is just the requirement that z satisfies at time t all constraints that are contained in (2).

With the change of variables

$$z = T_2 z_1 + T_2' z_2, \quad z_1 = T_2^T z, \quad z_2 = T_2'^T z$$

equation (73) turns into

$$\hat{F}_2(t, T_2(t)z_1 + T_2'(t)z_2) = 0. \tag{74}$$

If we set $z_1(t) = T_2^T(t)z(t)$, $z_2(t) = T_2'(t)^T z(t)$ then it follows that for all $t \in \mathbb{I}$

(a) $\hat{F}_2(t, T_2(t)z_1(t) + T_2'(t)z_2(t)) = 0$,
(b) $\hat{F}_{2;z}(t,z(t))T_2'(t)$ is nonsingular.

Thus, we can solve (74) for z_2 as

$$z_2 = \mathcal{R}(t,z_1) \tag{75}$$

and we have

$$z_2(t) = \mathcal{R}(t,z_1(t)) \text{ for all } t \in \mathbb{I}. \tag{76}$$

Since Hypothesis 1 also holds for the transformed system

$$\tilde{F}(t,z_1,z_2,\dot{z}_1,\dot{z}_2) = F(t, T_2z_1 + T_2'z_2, \tfrac{d}{dt}(T_2z_1 + T_2'z_2)), \tag{77}$$

we obtain from (70) a path

$$(t, z_1(t), z_2(t), \tilde{\mathcal{P}}(t)) \in \tilde{\mathbb{L}}_{\mu+1} \text{ for all } t \in \mathbb{I}, \tag{78}$$

where $\tilde{\mathbb{L}}_{\mu+1}$ is the solution set associated with (77). Besides (76) we have

$$p_2(t) = \mathcal{R}_t(t,z_1(t)) + \mathcal{R}_{z_1}(t,z_1(t))p_1(t), \tag{79}$$

where we use the partition

$$[\, I_{n+l} \;\; 0 \;\; \cdots \;\; 0 \,]\tilde{\mathcal{P}} = \begin{bmatrix} p_1(t) \\ p_2(t) \end{bmatrix},$$

compare the proof of Theorem 4.34 in [42]. From (77) and (78) we then obtain

$$\begin{aligned} &\tilde{F}(t,z_1(t),z_2(t),p_1(t),p_2(t)) \\ &\quad = F(t, T_2(t)z_1(t) + T_2'(t)\dot{z}_2(t), \dot{T}_2(t)z_1(t) + T_2(t)p_1(t) + \dot{T}_2'(t)z_2(t) + T_2'(t)p_2(t)) = 0, \\ &\qquad \text{for all } t \in \mathbb{I}, \end{aligned} \tag{80}$$

in which we can eliminate z_2, p_2 via (75) and (79), respectively. If we define

$$\begin{aligned}\tilde{F}_1(t,z_1,p_1) = Z_1(t)^T F(t, &T_2(t)z_1 + T_2'(t)\mathcal{R}(t,z_1),\\ &\dot{T}_2(t)z_1 + T_2(t)p_1(t) + \dot{T}_2'(t)\mathcal{R}(t,z_1) + T_2'(t)(\mathcal{R}_t(t,z_1) + \mathcal{R}_{z_1}(t,z_1)p_1)),\end{aligned}$$

then $(t, z_1(t), p_1(t))$ solves $\tilde{F}_1(t, z_1, p_1) = 0$.

Furthermore,

$$\tilde{F}_{1;p_1}(t, z_1(t), p_1(t)) = Z_1(t)^T F_{\dot{z}}(t, z(t), p(t))(T_2(t) + T_2'(t)\mathcal{R}_{z_1}(t, z_1(t))), \tag{81}$$

where $[\, I_{n+l} \;\; 0 \;\; \cdots \;\; 0 \,]\mathcal{P} = p$.

To determine $\mathcal{R}_{z_1}(t, z_1(t))$, one observes that from

$$\hat{F}_2(t, T_2(t)z_1(t) + T_2'(t)\mathcal{R}_{z_1}(t, z_1(t))) = 0 \text{ for all } t \in \mathbb{I},$$

it follows that

$$\hat{F}_{2;z}(t, z(t))(T_2(t) + T_2'(t)\mathcal{R}_{z_1}(t, z_1(t))) = 0 \text{ for all } t \in \mathbb{I}$$

and hence, using (73) we obtain

$$Z_2(t)^T F_{\mu;z}(t, z(t), \mathcal{P}(t))(T_2(t) + T_2'(t)\mathcal{R}_{z_1}(t, z_1(t))) = 0 \text{ for all } t \in \mathbb{I}.$$

By the construction of Z_2, T_2, and T_2', we immediatly obtain that

$$\mathcal{R}_{z_1}(t, z_1(t)) = 0 \text{ for all } t \in \mathbb{I}$$

and that $\tilde{F}_{1;p_1}(t, z_1(t), p_1(t))$ has full row rank for all $t \in \mathbb{I}$. Thus, there exists a pointwise orthogonal matrix function $[\, V' \;\; V \,] \in C^0(\mathbb{I}, \mathbb{R}^{d+l,d+l})$, with

$$Z_1(t)^T F_{\dot{z}}(t, z(t), p(t))T_2(t)[\, V'(t) \;\; V(t) \,] = [\, \Sigma(t) \;\; 0 \,], \tag{82}$$

with pointwise nonsingular Σ. Making a change of variables

$$z_1 = V'z_3 + Vz_4, \quad z_3 = V'^T z_1, \quad z_4 = V^T z_1, \tag{83}$$

and introducing

$$p_3 = \dot{V}'^T z_1 + V'^T p_1, \quad p_4 = \dot{V}^T z_1 + V^T p_1,$$

gives

$$p_1 = \dot{V}'z_3 + V'p_3 + \dot{V}z_4 + Vp_4,$$

and we obtain

$$\tilde{F}_1(t, V'(t)z_3(t) + V(t)z_4(t), \dot{V}'(t)z_3(t) + V'(t)p_3(t) + \dot{V}(t)z_4(t) + V(t)p_4(t)) = 0 \text{ for all } t \in \mathbb{I}.$$

If we define

$$\mathcal{H}(t, z_3, z_4, p_3, p_4) = \tilde{F}_1(t, V'(t)z_3 + V(t)z_4, \dot{V}'(t)z_3 + V'(t)p_3 + \dot{V}(t)z_4 + V(t)p_4),$$

then it follows that

(a) $\mathcal{H}(t, z_3(t), z_4(t), p_3(t), p_4(t)) = 0$,
(b) $\mathcal{H}_{p_3}(t, z_3(t), z_4(t), p_3(t), p_4(t)) = Z_1^T(t)F_{\dot{z}}(t, z(t), p(t))T_2(t)V'(t)$,

and we can solve for p_3 according to

$$p_3 = \mathcal{L}(t, z_3, z_4, p_4). \tag{84}$$

If we insert (83) into (75) we, moreover, obtain

$$z_2 = \mathcal{R}(t, V'(t)z_3 + V(t)z_4). \tag{85}$$

Note that by construction p_3 and p_4 represent the derivatives of z_3 and z_4, respectively. If we require that z_3 and z_4 are continuously differentiable and that $\mathcal{P}$ satisfies $p_3(t) = \dot{z}_3(t)$ and $p_4(t) = \dot{z}_4(t)$ for all $t \in \mathbb{I}$, then we notice that $z_4 \in C^1(\mathbb{I}, \mathbb{R}^l)$ plays the role of a control in the sense that one can choose it freely in $C^1(\mathbb{I}, \mathbb{R}^l)$ and with an appropriate initial condition $z_3(\underline{t})$ we obtain a unique solution of the ODE $\dot{z}_3 = \mathcal{L}(t, z_3, z_4, \dot{z}_4)$ corresponding to (84). Setting then $(x_1, x_2, u) = (z_3, z_2, z_4)$ we can rewrite (84), (85) as

$$\begin{array}{ll} \text{(a)} & \dot{x}_1 = \mathcal{L}(t, x_1, u, \dot{u}), \\ \text{(b)} & x_2 = \mathcal{R}(t, x_1, u), \end{array} \tag{86}$$

where we have used the same notation as in (85) for the function $\mathcal{R}$ in the renamed variables.

The appearance of $\dot{u}$ in (86) and the implied higher smoothness requirement for u cannot be avoided in the general case. However, the structure of the problem often implies that actually $\dot{u}$ is not present in (86), see e.g. [42, Remark 4.36]. We therefore use

$$\begin{array}{ll} \text{(a)} & \dot{x}_1 = \mathcal{L}(t, x_1, u), \\ \text{(b)} & x_2 = \mathcal{R}(t, x_1, u) \end{array} \tag{87}$$

instead of (86) and allow u to be only continuous.

If we transform the cost function correspondingly, then the optimal control problem (1) changes to

$$\mathcal{J}(x_1, x_2, u) = \mathcal{M}(x_1(\overline{t}), x_2(\overline{t})) + \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, x_1, x_2, u)\; dt = \min! \tag{88}$$

subject to (87) with initial condition $x_1(\underline{t}) = \underline{x}_1$.

Let $z = (x_1, x_2, u)$ be a (local) solution of this optimal control problem, where

$$x_1 \in C^1(\mathbb{I}, \mathbb{R}^d), \quad x_2 \in C^0(\mathbb{I}, \mathbb{R}^a), \quad u \in C^0(\mathbb{I}, \mathbb{R}^l).$$

Then (according to the standard theory for control problems with algebraic and differential constraints), there exist Lagrange multipliers

$$\lambda_1 \in C^1(\mathbb{I}, \mathbb{R}^d), \quad \lambda_2 \in C^0(\mathbb{I}, \mathbb{R}^a), \quad \gamma \in \mathbb{R}^d$$

such that $(x_1, x_2, u, \lambda_1, \lambda_2, \gamma)$ solves the unconstrained problem

$$\begin{aligned} &\mathcal{M}(x_1(\overline{t}), x_2(\overline{t})) + \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, x_1, x_2, u)\; dt + \gamma^T(x_1(\underline{t}) - \underline{x}_1) \\ &\qquad + \int_{\underline{t}}^{\overline{t}} \lambda_1^T(\dot{x}_1 - \mathcal{L}(t, x_1, u))\; dt + \int_{\underline{t}}^{\overline{t}} \lambda_2^T(x_2 - \mathcal{R}(t, x_1, u))\; dt = \min! \end{aligned} \tag{89}$$

in $\mathbb{W} = \mathbb{Z} \times \mathbb{Y}$ with

$$\mathbb{Z} = C^1(\mathbb{I}, \mathbb{R}^d) \times C^0(\mathbb{I}, \mathbb{R}^a) \times C^0(\mathbb{I}, \mathbb{R}^l), \quad \mathbb{Y} = C^1(\mathbb{I}, \mathbb{R}^d) \times C^0(\mathbb{I}, \mathbb{R}^a) \times \mathbb{R}^d.$$

From (89) we obtain the necessary condition

$$\begin{aligned}
&\mathcal{M}_{x_1}(x_1(\overline{t}), x_2(\overline{t}))\Delta x_1(\overline{t}) + \mathcal{M}_{x_2}(x_1(\overline{t}), x_2(\overline{t}))\Delta x_2(\overline{t}) + \\
&\qquad + \int_{\underline{t}}^{\overline{t}} (\mathcal{K}_{x_1}(t, x_1, x_2, u)\Delta x_1 + \mathcal{K}_{x_2}(t, x_1, x_2, u)\Delta x_2 + \mathcal{K}_u(t, x_1, x_2, u)\Delta u)\; dt \\
&\qquad + \int_{\underline{t}}^{\overline{t}} \lambda_1^T (\Delta \dot{x}_1 - \mathcal{L}_{x_1}(t, x_1, u)\Delta x_1 - \mathcal{L}_u(t, x_1, u)\Delta u)\; dt \\
&\qquad + \int_{\underline{t}}^{\overline{t}} \Delta\lambda_1^T (\dot{x}_1 - \mathcal{L}_{x_1}(t, x_1, u))\; dt \\
&\qquad + \int_{\underline{t}}^{\overline{t}} \lambda_2^T (\Delta x_2 - \mathcal{R}_{x_1}(t, x_1, u)\Delta x_1 - \mathcal{R}_u(t, x_1, u)\Delta u)\; dt \\
&\qquad + \gamma^T \Delta x_1(\underline{t}) + \Delta\gamma^T (x_1(\underline{t}) - \underline{x}_1) = 0
\end{aligned}$$

for all $(\Delta x_1, \Delta x_2, \Delta u, \Delta\lambda_1, \Delta\lambda_2, \Delta\gamma) \in \mathbb{W}$. Variation over $\Delta\lambda_1$, $\Delta\lambda_2$, and $\Delta\gamma_1$ as usual reproduces the constraints. Moreover, comparing with the linear case we only have to perform the replacements

$$\begin{array}{ll}
\text{(a)} & x_1(\overline{t})^T M \leftarrow [\, \mathcal{M}_{x_1}(x_1(\overline{t}), x_2(\overline{t}))\;\; \mathcal{M}_{x_2}(x_1(\overline{t}), x_2(\overline{t}))\,], \\
\text{(b)} & x^T W + u^T S \leftarrow [\, \mathcal{K}_{x_1}(t, x_1, x_2, u)\;\; \mathcal{K}_{x_2}(t, x_1, x_2, u)\,], \\
\text{(c)} & x^T S^T + u^T R \leftarrow \mathcal{K}_u(t, x_1, x_2, u), \\
\text{(d)} & E \leftarrow \begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix},\; A \leftarrow \begin{bmatrix} \mathcal{L}_{x_1}(t, x_1, u) & 0 \\ \mathcal{R}_{x_1}(t, x_1, u) & -I \end{bmatrix},\; B \leftarrow \begin{bmatrix} \mathcal{L}_u(t, x_1, u) \\ \mathcal{R}_u(t, x_1, u) \end{bmatrix},
\end{array} \tag{90}$$

in (39).

In this way, we obtain the boundary value problem of necessary optimality conditions

$$\begin{array}{ll}
\text{(a)} & \dot{x}_1 = \mathcal{L}(t, x_1, u),\;\; x_1(\underline{t}) = \underline{x}_1, \\
\text{(b)} & x_2 = \mathcal{R}(t, x_1, u), \\
\text{(c)} & \dot{\lambda}_1 = \mathcal{K}_{x_1}(t, x_1, x_2, u)^T - \mathcal{L}_{x_1}(t, x_1, x_2, u)^T \lambda_1 - \mathcal{R}_{x_1}(t, x_1, u)^T \lambda_1, \\
& \qquad\qquad \lambda_1(\overline{t}) = -\mathcal{M}_{x_1}(x_1(\overline{t}), x_2(\overline{t}))^T \\
\text{(d)} & 0 = \mathcal{K}_{x_2}(t, x_1, x_2, u)^T + \lambda_2, \\
\text{(e)} & 0 = \mathcal{K}_u(t, x_1, x_2, u)^T - \mathcal{L}_u(t, x_1, u)^T \lambda_1 - \mathcal{R}_u(t, x_1, u)^T \lambda_2, \\
\text{(f)} & \gamma = \lambda_1(\underline{t})
\end{array} \tag{91}$$

proving the following result.

Theorem 20 *Let z be a local solution of (4) subject to (5) and (6) in the sense that the transformed $(x_1, x_2, u) \in \mathbb{Z}$ is a local solution of (88) subject to (87) and $x_1(\underline{t}) = \underline{x}_1$. Then there exist unique Lagrange multipliers $(\lambda_1, \lambda_2, \gamma) \in \mathbb{Z}$ such that $(x_1, x_2, u, \lambda_1, \lambda_2, \gamma)$ solves the boundary value problem (91).*

Remark 21 The preceding result can be generalized to constraints that additionally contain end conditions, i.e., conditions on parts of $x(\overline{t})$. The observation is the same as for ODEs. In particular, for every additional scalar end condition we lose one scalar condition on values $\lambda(\overline{t})$.

3.5 A Maximum Principle for general DAEs

If we, furthermore, allow the input functions to be constrained, then according to (18), (19) we have the problem

$$\mathcal{J}(x,u) = \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t,x(t),u(t))\ dt = \min! \tag{92}$$

subject to

$$\begin{aligned} 0 &= F(t,x,u,\dot{x}),\\ 0 &= h(\overline{t},x(\overline{t})),\ \ h \in C(\mathbb{R}\times\mathbb{R}^n,\mathbb{R}^n),\\ u(t) &\in \mathcal{U} \subset \mathbb{R}^l \ \text{ for all } t \in \mathbb{I}. \end{aligned} \tag{93}$$

Since the control u ist restricted we must (at least) allow for an optimal control u that has a finite number of jumps, i.e., $u \in L^c_\infty(\mathbb{I},\mathbb{R}^l)$. In view of (87) we must then also allow that the algebraic variables x_2 in a reduced formulation possess jumps at the same locations. Moreover, we must say in what sense the differential-algebraic equation in the constraint is to be satisfied when we allow for jumps in the input.

Starting with $z \in L^c_\infty(\mathbb{I},\mathbb{R}^{n+l})$ and thus with a (piecewise continuous) path $(t,z(t),\mathcal{P}(t))$ as a potential candidate for a minimum, Hypothesis 1 will no longer guarantee that we can perform the construction in the beginning of Section 3.4. In particular, we need additional assumptions that yield the necessary smooth transformations and the necessary implications of the implicit function theorem. Note that in many applications these additional properties hold due to the structure of the given problem.

Following the beginning of Section 3.4, we first assume that in spite of the lack of smoothness the projector functions Z_1', Z_1, Z_2', Z_2, T_1', T_1, T_2', T_2 are still at least continuous. Instead of (71) we define

$$\tilde{\mathcal{H}}(t,z,p,\tilde{p},\phi) = \begin{bmatrix} F_\mu(t,z,p) + Z_2(t)\phi \\ T_1(t)^T(p-\tilde{p}) \end{bmatrix} \tag{94}$$

which can locally be solved according to

$$\phi = \tilde{F}_2(t,z,\tilde{p}), \quad p = \tilde{\mathcal{P}}(t,z,\tilde{p}).$$

Here we must assume that at a point where the solution path has a jump the whole jump lies in the domain of the implicitly defined functions. In particular, we then have that

$$\hat{F}_2(t,z(t)) = 0 \text{ for all } t \in \mathbb{I},$$

where $\hat{F}_2(t,z) = \tilde{F}_2(t,z,\mathcal{P}(t))$. Again, $\hat{F}_{2;z}$ has full row rank along $(t,z(t))$ and $\hat{F}_2(t,z) = 0$ represents all the algebraic constraints that are contained in the DAE.

Proceeding in the same way with the following constructions and corresponding assumptions, we arrive at the reduced formulation consisting of (84) and (85) with no p_4 present in (84). Again $z_4 \in L^c_\infty(\mathbb{I},\mathbb{R}^l)$ plays the role of a control. Given an initial value $z_3(\underline{t})$ we require that

$$z_3(t) = z_3(\underline{t}) + \int_{\underline{t}}^{t} p_3(s)\ ds$$

holds for the given path. Renaming the variables as before, we get the reduced problem

$$\begin{aligned} &\text{(a)} \qquad x_1(t) = x_1(\underline{t}) + \textstyle\int_{\underline{t}}^{t} \mathcal{L}(s,x_1(s),u(s))\ ds,\\ &\text{(b)} \qquad x_2 = \mathcal{R}(t,x_1,u), \end{aligned} \tag{95}$$

which reflects that x_1 does not need to be continuously differentiable on the whole interval $\mathbb{I}$.

Transforming $\mathcal{J}$ and h to the new variables, eliminating the variable x_2 with the help of the algebraic constraint, and assuming that the so obtained h does not depend on u (since the quantity $u(\overline{t})$ does not make sense), we get as interpretation of (92) and (93) a problem of the form

$$\hat{\mathcal{J}}(x_1, u) = \int_{\underline{t}}^{\overline{t}} \hat{\mathcal{K}}(t, x_1(t), u(t))\, dt = \min! \tag{96}$$

subject to

$$\begin{aligned} x_1(t) &= x_1(\underline{t}) + \textstyle\int_{\underline{t}}^{t} \mathcal{L}(s, x_1(s), u(s))\, ds, \\ 0 &= \hat{h}(\overline{t}, x_1(\overline{t})),\ \ h \in C(\mathbb{R} \times \mathbb{R}^d, \mathbb{R}^d), \\ u(t) &\in \mathcal{U} \subset \mathbb{R}^l \ \text{ for all } t \in \mathbb{I} \end{aligned} \tag{97}$$

for the determination of an optimal $(\overline{t}, x_1, u) \in \mathbb{R} \times C^0(\mathbb{I}, \mathbb{R}^d) \times L^c_\infty(\mathbb{I}, \mathbb{R}^d))$. Of course, the missing part x_2 is then given by (95b).

Theorem 22 *If $(\overline{t}, x_1^*, u^*)$ is a local solution of the Bolza problem (96) subject to (97), then there exist scalars $\alpha_0, \alpha_1, \ldots, \alpha_d$, which do not all vanish simultaneously, $\alpha_0 \geq 0$, and a multiplier $\lambda \in C^0(\mathbb{I}, \mathbb{R}^d)$ such that, with $H(t, x_1, u, \lambda, \alpha_0) = \lambda^T \mathcal{L}(t, x_1, u) - \alpha_0 \hat{\mathcal{K}}(t, x_1, u)$ and $\alpha = (\alpha_1, \cdots, \alpha_n)^T$,*

$$\begin{aligned} H(t, x_1^*(t), u^*(t), \lambda(t), \alpha_0) &= \max\nolimits_{u \in \mathcal{U}} H(t, x_1^*(t), u, \lambda(t), \alpha_0), \\ \dot{\lambda}(t) &= -\nabla_{x_1} H(t, x_1^*(t), u^*(t), \lambda(t), \alpha_0), \\ \dot{x}_1^*(t) &= \nabla_\lambda H(t, x_1^*(t), u^*(t), \lambda(t), \alpha_0), \\ \lambda(\overline{t}) &= -\alpha^T \nabla_{x_1(\overline{t})} \hat{h}(\overline{t}, x_1^*(\overline{t})), \end{aligned} \tag{98}$$

for all $t \in \mathbb{I}$, where u^* is continuous.

Theorem 22 covers Theorem 20 in the case of a continuous control u when we fix $\overline{t}$ omitting the condition on it described by h in Theorem 22 and when we omit the costs on the final state described by $\mathcal{M}$ in Theorem 20. Of course, we could have formulated generalized versions of both theorems such that this would directly be the case, but we chose the restricted versions because here we concentrate merely on the DAE aspect of the results and not on the most general possible formulations.

As a simple application of Theorem 22 we consider a semi-explicit differential-algebraic equation of index 1 in the constraints. In particular, we consider the problem

$$\mathcal{J}(x_1, x_2, u) = \int_{\underline{t}}^{\overline{t}} \mathcal{K}(t, x_1(t), x_2(t), u(t))\, dt = \min!$$

subject to

$$\begin{aligned} \dot{x}_1 &= f(t, x_1, x_2, u),\ \ x_1(\underline{t}) = \underline{x}_1 \\ 0 &= g(t, x_1, x_2, u), \\ u(t) &\in \mathcal{U} \subset \mathbb{R}^l \ \text{ for all } t \in \mathbb{I}, \end{aligned}$$

with the assumption that g_y is everywhere nonsingular. Given a (local) solution (x_1, x_2, u) the implicit function theorem yields that x_2 is determined in terms of (t, x_1, u) according to

$$x_2 = \mathcal{R}(t, x_1, u),$$

while the differential part reads

$$\dot{x}_1 = \mathcal{L}(t, x_1, u) = f(t, x_1, \mathcal{R}(t, x_1, u), u).$$

Accordingly, we get

$$\hat{\mathcal{K}}(t, x_1, u) = \mathcal{K}(t, x_1, \mathcal{R}(t, x_1, u), u).$$

Thus, the structure of the differential-algebraic equation immediately gives a suitable reformulation fitting to Theorem 22. With

$$H(t, x_1, u, \lambda, \alpha_0) = \lambda^T f(t, x_1, \mathcal{R}(t, x_1, u), u) - \alpha_0 \mathcal{K}(t, x_1, \mathcal{R}(t, x_1, u), u)$$

the essential part of (98) reads (omitting arguments)

$$\begin{aligned} \dot{\lambda}(t) &= -((f_{x_1}^T + \mathcal{R}_{x_1}^T f_{x_2}^T)\lambda + \alpha_0(\mathcal{K}_{x_1}^T + \mathcal{R}_{x_1}^T \mathcal{K}_{x_2}^T)), \\ \dot{x}_1 &= f(t, x_1, \mathcal{R}(t, x_1, u), u). \end{aligned}$$

This corresponds (up to several technical differences) to the results given in [60]. The same applies to the other results of [60] which deal with differential-algebraic equations in Hessenberg form of index 2 and differential-algebraic equations of index 3 that arise in the modeling of multibody systems. The latter case, however, is treated by an index reduction which increases the number of differential components. Hence, one must pay attention to the correct choice of the boundary conditions.

4 Numerical methods for optimal control problems

In this section we discuss the numerical solution of the optimality boundary value problems (49) and (91), respectively. In contrast to the analytical treatment, for the numerical solution we may not just assume that the free system (with $u = 0$) is strangeness-free, since the regularizing feedbacks can only be computed during the integration, nor can we work with implicitly defined functions as contained in (91). Instead we have to work with initial data and possibly their derivatives.

We again first study the case of linear systems with variable coefficients.

4.1 Numerical methods for linear-quadratic optimal control problems

In order to incorporate (if necessary) an index reduction we use the functions as in (25) that can be determined in every time step from the given data (including derivatives of the coefficient functions) but we have to note that the projection functions Z_1^T and Z_2^T are not realized as smooth functions in numerical methods such as the code `GENDA` [44], although such smooth realizations exist. It would be just too expensive to carry the computation of smooth realizations along. Since most numerical integration methods such as Runge-Kutta methods or BDF methods, see [9, 32], are invariant under transformations from the left, the non-smooth realizations yield the same results.

Taking into account that the coefficient functions in (49) are only available through index reduction and assuming sufficient smoothness of the data, we write (49) in terms of (26) as

$$\begin{array}{ll} \text{(a)} & \hat{E}_1 \dot{x} = \hat{A}_1 x + \hat{B}_1 u + \hat{f}_1, \ (\hat{E}_1^+ \hat{E}_1 x)(\underline{t}) = \underline{x} \\ \text{(b)} & 0 = \hat{A}_2 x + \hat{B}_2 u + \hat{f}_2, \\ \text{(c)} & \frac{d}{dt}(\hat{E}_1^T \lambda_1) = W x + S u - \hat{A}_1^T \lambda_1 - \hat{A}_2^T \lambda_2, \\ & \qquad \lambda_1(\overline{t}) = -[\, \hat{E}_1^+(\overline{t})^T \ \ 0 \,] M x(\overline{t}), \\ \text{(d)} & 0 = S^T x + R u - \hat{B}_1^T \lambda_1 - \hat{B}_2^T \lambda_2. \end{array} \tag{99}$$

The missing smoothness of Z_1, Z_2 is not a problem in (99a) and (99b), but as constructed, the unknowns λ_1, λ_2 are in general not smooth if Z_1, Z_2 are non-smooth. If, however, we choose Z_1 and Z_2 to have orthonormal columns, then at least $Z_1 Z_1^T$ and $Z_2 Z_2^T$ are smooth, since they represent orthogonal projections onto the corresponding image spaces. Thus, with

$$\hat{E}_1^T \lambda_1 = E^T Z_1 \lambda_1 = E^T Z_1 Z_1^T Z_1 \lambda_1 = E^T Z_1 Z_1^T \hat{\lambda}_1$$

and $\hat{\lambda}_1 = Z_1 \lambda_1$, we can obtain smooth coefficients for the unknown $\hat{\lambda}_1$. However, we have to add the condition that $\hat{\lambda}_1 \in \operatorname{range} Z_1$ to the system. If we complete Z_1 via Z_1' to a pointwise orthogonal matrix function, then we can express this as

$$Z_1'^T \hat{\lambda}_1 = 0. \tag{100}$$

Note, that here again it plays no role whether Z' is constructed as a smooth function.

Due to

$$\begin{aligned}
[\, \hat{A}_2 \ \ \hat{B}_2 \,]^T \lambda_2 &= [\, I_{n+l} \ 0 \ \cdots \ 0 \,] N_\mu^T Z_2 \lambda_2 \\
&= [\, I_{n+l} \ 0 \ \cdots \ 0 \,] N_\mu^T Z_2 Z_2^T Z_2 \lambda_2 \\
&= [\, I_{n+l} \ 0 \ \cdots \ 0 \,] N_\mu^T Z_2 Z_2^T \hat{\lambda}_2,
\end{aligned}$$

with $\hat{\lambda}_2 = Z_2 \lambda_2$, we can proceed analogously for λ_2. In particular, we complete Z_2 via Z_2' to a pointwise orthogonal matrix function and require

$$Z_2'^T \hat{\lambda}_2 = 0. \tag{101}$$

Adding (100) and (101) to boundary value problem (99) yields the new boundary value problem

$$\begin{array}{ll}
\text{(a)} & \hat{E}_1 \dot{x} = \hat{A}_1 x + \hat{B}_1 u + \hat{f}_1, \ (\hat{E}_1^+ \hat{E}_1 x)(\underline{t}) = \underline{x}, \\
\text{(b)} & 0 = \hat{A}_2 x + \hat{B}_2 u + \hat{f}_2, \\
\text{(c)} & \frac{d}{dt}(E^T Z_1 Z_1^T \hat{\lambda}_1) = W x + S u - A^T \hat{\lambda}_1 - [\, I_n \ 0 \,|\, 0 \ 0 \,|\, \cdots \,|\, 0 \ 0 \,] N_\mu^T \hat{\lambda}_2, \\
 & \qquad (Z_1^T \hat{\lambda}_1)(\overline{t}) = -[\, \hat{E}_1^+(\overline{t})^T \ \ 0 \,] M x(\overline{t}), \\
\text{(d)} & 0 = S^T x + R u - B^T \hat{\lambda}_1 - [\, 0 \ I_l \,|\, 0 \ 0 \,|\, \cdots \,|\, 0 \ 0 \,] N_\mu^T \hat{\lambda}_2, \\
\text{(e)} & 0 = Z_1'^T \hat{\lambda}_1, \\
\text{(f)} & 0 = Z_2'^T \hat{\lambda}_2.
\end{array} \tag{102}$$

As constructed, this boundary value problem now allows for a smooth solution independent of non-smooth realizations of the coefficient functions. The parts which are not explicitly represented in the original data and their derivatives can be obtained from them by the standard index reduction, see [42]. Compare also with the following discussion of the nonlinear case.

Since the coefficient $E^T Z_1 Z_1^T$ is smooth, we can discretize (102) for example with BDF methods or using the boundary value methods introduced in [46, 47, 48]. This is justified by the following observation.

Lemma 23 *The boundary value problem (102) is regular and strangeness-free iff the boundary value problem (49) is regular and strangeness-free.*

Proof. Consider the coefficients of the boundary value problem given by

$$\mathcal{E} = \begin{bmatrix} \hat{E}_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & E^T Z_1 Z_1^T & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} \hat{A}_1 & \hat{B}_1 & 0 & 0 \\ \hat{A}_2 & \hat{B}_2 & 0 & 0 \\ W & S & -A^T - \frac{d}{dt}(E^T Z_1 Z_1^T) & -\tilde{A}_{3,4} \\ S^T & R & -B^T & -\tilde{A}_{4,4} \\ 0 & 0 & Z_1'^T & 0 \\ 0 & 0 & 0 & Z_2'^T \end{bmatrix}, \tag{103}$$

with

$$\tilde{A}_{3,4} = [\, I_n \;\; 0 \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T, \quad \tilde{A}_{4,4} = [\, 0 \;\; I_l \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T,$$

where obviously $\operatorname{rank} \mathcal{E} = 2d$. With

$$\mathcal{Z} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ I_1 & 0 & 0 & 0 & 0 \\ 0 & T_2' & 0 & 0 & 0 \\ 0 & I_l & 0 & 0 & 0 \\ 0 & 0 & 0 & I_a & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}, \quad \mathcal{T} = \begin{bmatrix} T_2' & 0 & 0 & 0 \\ 0 & I_l & 0 & 0 \\ 0 & 0 & Z_1' & 0 \\ 0 & 0 & 0 & I \end{bmatrix},$$

describing corange and kernel of $\mathcal{E}$, we have $\mathcal{Z}^T \mathcal{E} = 0$, $\mathcal{E}\mathcal{T} = 0$, and thus, the DAE associated with (103) ist regular and strangeness-free if and only if

$$\mathcal{Z}^T \mathcal{A} \mathcal{T} = \begin{bmatrix} \hat{A}_2 T_2' & \hat{B}_2 & 0 & 0 \\ T_2'^T W T_2' & T_2'^T S & -T_2'^T (A^T + \frac{d}{dt}(E^T Z_1 Z_1^T)) Z_2'^T & -T_2'^T \tilde{A}_{3,4} \\ S^T T_2' & R & -B^T Z_1' & -\tilde{A}_{4,4} \\ 0 & 0 & I_a & 0 \\ 0 & 0 & 0 & Z_2'^T \end{bmatrix} \tag{104}$$

is nonsingular. Omitting the block row and block column containing I_a and multiplying the last block column with $[\, Z_2' \;\; Z_2 \,]$, we see that $\mathcal{Z}^T \mathcal{A} \mathcal{T}$ is nonsingular if and only if

$$\begin{bmatrix} \hat{A}_2 T_2' & \hat{B}_2 & 0 & 0 \\ T_2'^T W T_2' & T_2'^T S & -T_2'^T \tilde{A}_{3,4} Z_2' & -T_2'^T \tilde{A}_{3,4} Z_2 \\ S^T T_2' & R & -\tilde{A}_{4,4} Z_2' & -\tilde{A}_{4,4} Z_2 \\ 0 & 0 & I_a & 0 \end{bmatrix},$$

is nonsingular. Since

$$\begin{aligned} T_2'^T \tilde{A}_{3,4} Z_2 &= T_2'^T [\, I_n \;\; 0 \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T Z_2 = T_2'^T \hat{A}_2^T, \\ \tilde{A}_{4,4} Z_2 &= [\, 0 \;\; I_l \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T Z_2 = \hat{B}_2^T, \end{aligned}$$

and observing that

$$A_{2,2} = \hat{A}_2 T_2', \quad B_2 = \hat{B}_2, \quad W_{2,2} = T_2'^T W T_2', \quad S_2^T = S^T T_2',$$

this is the case if and only $\hat{R}$ as in (50) is nonsingular. $\square$

4.2 Numerical methods for nonlinear optimal control problems

The numerical solution of the boundary value problem (91) is approached in a similar way as for the linear case. In order to represent this boundary value problem in the original data we proceed as for the integration of a differential-algebraic equation, see [42]. The differential equations (91a) are represented by the equations $Z_1^T F(t, x, \dot{x}, u) = 0$ with Z_1 defined by Hypothesis 1 and the algebraic equations (91b) are implied by the derivative array $F_\mu(t, x, u, p) = 0$. The remaining equations are defined via linearization and correspond to the equation (102c–d) of the linear case such that in (102c–d) the replacements

$$E \leftarrow F_{\dot{x}}, \quad A \leftarrow -F_x, \quad B \leftarrow -F_u, \quad Mx(\overline{t}) \leftarrow \mathcal{M}_x(x(\overline{t}))^T$$

and

$$N_\mu[\, I_n \;\; 0 \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,]^T \leftarrow -F_{\mu;x}, \quad N_\mu[\, 0 \;\; I_l \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,]^T = -F_{\mu;u}$$

apply. Introducing $\hat{\lambda}_1, \hat{\lambda}_2$ and adding (100) and (101) to the nonlinear boundary value problem we end up with the boundary value problem (omitting arguments)

$$\begin{array}{ll}
\text{(a)} & Z_1^T F = 0, \; (\hat{E}_1^+ \hat{E}_1 x)(\underline{t}) = \underline{x}, \\
\text{(b)} & F_\mu = 0, \\
\text{(c)} & \frac{d}{dt}(F_{\dot{x}}^T Z_1 Z_1^T \hat{\lambda}_1) = \mathcal{K}_x^T + F_x^T \hat{\lambda}_1 + F_{\mu;x} \hat{\lambda}_2, \\
 & \qquad\qquad (Z_1^T \hat{\lambda}_1)(\overline{t}) = -[\, \hat{E}_1^+(\overline{t})^T \;\; 0 \,] \mathcal{M}_x(x\overline{t}), \\
\text{(d)} & 0 = \mathcal{K}_u^T + F_u^T \hat{\lambda}_1 + F_{\mu;u}^T \hat{\lambda}_2, \\
\text{(e)} & 0 = Z_1'^T \hat{\lambda}_1, \\
\text{(f)} & 0 = Z_2'^T \hat{\lambda}_2.
\end{array} \tag{105}$$

Note that we have presented the boundary conditions in terms of the linearization. Of course it is possible to state them in terms of the original data and the involved projections as the other relations. However, the resulting formulas are relatively complicated, since they are directly related to the problem of the consistent initialization of differential-algebraic equations, and we refrain here from presenting these.

4.3 A numerical example

We have discretized the boundary value problem (105) by means of midpoint rule for the differential equations in the state variables and trapezoidal rule for the differential equations in the adjoint variables together with the algebraic constraints at all grid points and simple divided differences for the term $\frac{d}{dt}(F_{\dot{x}}^T Z_1 Z_1^T \hat{\lambda}_1)$.

In order to use numerical differentiation to generate the necessary Jacobians, the relations starting with Z_1^T, $Z_1'^T$, and $Z_2'^T$, respectively were used with smooth projectors $Z_1 Z_1^T$, $Z_1' Z_1'^T$, and $Z_2' Z_2'^T$ instead. Although this introduces a rank deficiency in the Jacobians with respect to the rows, we can expect the Gauß-Newton method to converge at least superlinearly due to the consistency of the solution with the overdetermined relations. Note that this would not be necessary if we generated the Jacobians utilizing the structure of the equations.

The preceding approach was implemented in `FORTRAN` double precision. As one of the test problems we selected a nonlinear optimal control problem for a multibody system taken from [12].

Table 1: Course of the Gauß-Newton iteration

k	$\|\Delta w_k\|_2$
1	0.140D+03
2	0.223D+03
$\vdots$	$\vdots$
16	0.561D+01
17	0.103D+01
18	0.610D-02
19	0.318D-06
20	0.966D-11

A model problem for a motor controlled pendulum to be driven into its equilibrium with minimal costs is given by

$$
\begin{array}{llll}
 & J(x,u) = \displaystyle\int_0^3 u(t)^2\, dt = \min! & & \\
\text{s.t.} & \dot{x}_1 = x_3, & x_1(0) = \frac{1}{2}\sqrt{2}, & g = 9.81 \\
 & \dot{x}_2 = x_4, & x_2(0) = -\frac{1}{2}\sqrt{2}, & \\
 & \dot{x}_3 = -2x_1x_5 + x_2u, & x_3(0) = 0, & \\
 & \dot{x}_4 = -g - 2x_2x_5 - x_1u, & x_4(0) = 0, & \\
 & 0 = x_1^2 + x_2^2 - 1, & x_5(0) = -\frac{1}{2}gx_2(0). &
\end{array}
$$

It is known that the differential-algebraic equation in the constraint satisfies Hypothesis 1 with $\mu = 2$, $a = 3$, $d = 2$, and $v = 0$. Hence, only two scalar initial values are sufficient to describe the initial state. We chose them to be $x_2(0) = -\frac{1}{2}\sqrt{2}$ and $x_3(0) = 0$. Similarly, $x_1(3) = 0$ and $x_3(0) = 0$ are sufficient to describe the equilibrium at the end point. According to Remark 21 no end conditions for the Lagrange multipliers occur. As initial trajectory we took

$$
\begin{array}{lll}
x_1(t) = \frac{1}{2}\sqrt{2} - \frac{1}{6}\sqrt{2}t, & x_3(t) = 0, & \\
x_2(t) = -\sqrt{1 - x_1(t)^2}, & x_4(t) = 0, & x_5(t) = -\frac{1}{2}gx_2(t),
\end{array}
$$

with all other unknowns set to zero on an equidistant grid of 60 intervals. The required tolerance for the Gauß-Newton method was 10^{-7}. See Table 1 for the course of the iteration, where k counts the iterations and $\|\Delta w_k\|_2$ denotes the Euclidian norm of the corresponding Gauß-Newton correction.

In the Gauss-Newton iteration, there is an initial phase with a bad convergence behavior due to a bad initial guess which could be remedied by damping. But in the final phase one easily recognizes quadratic convergence. The obtained final value of the cost function was $J_{\text{opt}} = 3.82$ which is, up to discretization errors, in coincidence with the value given in [12].

Note that the implementation used here is by no means efficient. This would require to incorporate the structure of the problem when setting up the Jacobian and solving the linear subproblems. See [46, 47] for techniques that may be applied.

5 Conclusions

We have presented the optimal control theory for general unstructured linear and nonlinear systems of differential-algebraic equations. We have derived necessary conditions and a maximum principle and have shown how these can be solved numerically. The results are illustrated by a numerical example.

References

[1] A. Backes. A necessary optimality condition for the linear-quadratic DAE control problem. Preprint 2003-16, Institut für Mathematik, Humboldt-Universität zu Berlin, Berlin, Germany, 2003.

[2] A. Backes. *Optimale Steuerung der linearen DAE im Fall Index 2.* Dissertation, Mathematisch-Naturwissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Berlin, Germany, 2006.

[3] K. Balla, G. Kurina, and R. März. Index criteria for differential algebraic equations arising from linear-quadratic optimal control problems. Preprint 2003–14, Institut für Mathematik, Humboldt-Universität zu Berlin, Berlin, Germany, 2003.

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

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

[6] K. Balla and R. März. Linear boundary value problems for differential algebraic equations. *Math. Notes*, 5:3–17, 2004.

[7] D. Bender and A. Laub. The linear quadratic optimal regulator problem for descriptor systems. *IEEE Trans. Automat. Control*, 32:672–688, 1987.

[8] V. Boltyanskii, R. Gamkrelidze, E. Mishenko, and L. S. Pontryagin. *The Mathematical Theory of Optimal Processes.* Interscience, New York, NY, 1962.

[9] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations.* SIAM Publications, Philadelphia, PA, 2nd edition, 1996.

[10] A. Bunse-Gerstner, V. Mehrmann, and N. K. Nichols. Regularization of descriptor systems by derivative and proportional state feedback. *SIAM J. Matr. Anal. Appl.*, 13:46–67, 1992.

[11] A. Bunse-Gerstner, V. Mehrmann, and N. K. Nichols. Regularization of descriptor systems by output feedback. *IEEE Trans. Automat. Control*, 39:1742–1748, 1994.

[12] C. Büskens and M. Gerdts. Numerical solution of optimal problems with DAEs of higher index. In *Proceedings of the Workshop: Optimalsteuerungsprobleme in der Luft und Raumfahrt*, pages 27–38. Sonderforschungsbereich 255: Transatmosphärische Flugsysteme, Hieronymus, München, 2000.

[13] R. Byers, T. Geerts, and V. Mehrmann. Descriptor systems without controllability at infinity. *SIAM J. Cont.*, 35:462–479, 1997.

[14] R. Byers, P. Kunkel, and V. Mehrmann. Regularization of linear descriptor systems with variable coefficients. *SIAM J. Cont.*, 35:117–133, 1997.

[15] S. L. Campbell. *Singular Systems of Differential Equations I.* Pitman, San Francisco, CA, 1980.

[16] S. L. Campbell. Comment on controlling generalized state-space (descriptor) systems. *Internat. J. Control*, 46:2229–2230, 1987.

[17] S. L. Campbell. A general form for solvable linear time varying singular systems of differential equations. *SIAM J. Math. Anal.*, 18:1101–1115, 1987.

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

[19] S. L. Campbell and C. D. Meyer. *Generalized Inverses of Linear Transformations.* Pitman, San Francisco, CA, 1979.

[20] J. D. Cobb. A further interpretation of inconsistent initial conditions in descriptor-variable systems. *IEEE Trans. Automat. Control*, AC-28:920–922, 1983.

[21] E. N. Devdariani and Yu. S. Ledyaev. Maximum principle for implicit control systems. *Appl. Math. Optim.*, 40:79–103, 1999.

[22] M. Diehl, D. B. Leineweber, A. Schäfer, H.G. Bock, and J.P. Schlöder. Optimization of multiple-fraction batch distillation with recycled waste cuts. *AIChE Journal*, 48(12):2869–2874, 2002.

[23] M. Diehl, I. Uslu, R. Findeisen, S. Schwarzkopf, F. Allgöwer, H. G. Bock, T. Bürner, E. D. Gilles, A. Kienle, J. P. Schlöder, and E. Stein. Real-time optimization for large scale processes: Nonlinear model predictive control of a high purity distillation column. In M. Grötschel, S. O. Krumke, and J. Rambau, editors, *Online Optimization of Large Scale Systems: State of the Art*, pages 363–384. Springer, 2001.

[24] H. Döring. Traktabilitätsindex und Eigenschaften von matrix-wertigen Riccati-Typ Algebrodifferentialgleichungen. Diplomarbeit, Institut für Mathematik, Humboldt-Universität zu Berlin, Berlin, Germany, 2005.

[25] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Systems.* Teubner Verlag, Stuttgart, Germany, 1998.

[26] R. Gabasov and F. Kirillova. *The Qualitative Theory of Optimal Processes.* Marcel Dekker, New York, NY, 1976.

[27] M. Gerdts. Optimal control and real-time optimization of mechanical multi-body systems. *Z. Angew. Math. Mech.*, 83:705–719, 2003.

[28] M. Gerdts. Local minimum principle for optimal control problems subject to index two differential algebraic equations systems. Technical report, Fakultät für Mathematik, Universität Hamburg, Hamburg, Germany, 2005.

[29] E. Griepentrog and R. März. *Differential-Algebraic Equations and their Numerical Treatment.* Teubner Verlag, Leipzig, Germany, 1986.

[30] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry I. Mathematical structure and index of network equations. *Surv. Math. Ind.*, 8:97–129, 1999.

[31] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry II. Impact of circuit configurations and parameters. *Surv. Math. Ind.*, 8:131–157, 1999.

[32] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems.* Springer-Verlag, Berlin, Germany, 2nd edition, 1996.

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

[34] A. D. Ioffe and V. M. Tichomirov. *Theorie der Extremalaufgaben.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1979.

[35] V. Ionescu, C. Oara, and M. Weiss. *Generalized Riccati Theory and Robust Control: A Popov Function Approach.* John Wiley and Sons, Chichester, UK, 1999.

[36] A. Kirsch, W. Warth, and J. Werner. *Notwendige Optimalitätsbedingungen und ihre Anwendung.* Springer-Verlag, Berlin, 1978.

[37] P. Kunkel and V. Mehrmann. Generalized inverses of differential-algebraic operators. *SIAM J. Matr. Anal. Appl.*, 17:426–442, 1996.

[38] P. Kunkel and V. Mehrmann. A new class of discretization methods for the solution of linear differential algebraic equations with variable coefficients. *SIAM J. Numer. Anal.*, 33:1941–1961, 1996.

[39] P. Kunkel and V. Mehrmann. The linear quadratic control problem for linear descriptor systems with variable coefficients. *Math. Control, Signals, Sys.*, 10:247–264, 1997.

[40] P. Kunkel and V. Mehrmann. Regular solutions of nonlinear differential-algebraic equations and their numerical determination. *Numer. Math.*, 79:581–600, 1998.

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

[42] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution.* EMS Publishing House, Zürich, Switzerland, 2006.

[43] P. Kunkel, V. Mehrmann, and W. Rath. Analysis and numerical solution of control problems in descriptor form. *Math. Control, Signals, Sys.*, 14:29–61, 2001.

[44] P. Kunkel, V. Mehrmann, W. Rath, and J. Weickert. A new software package for linear differential–algebraic equations. *SIAM J. Sci. Comput.*, 18:115–138, 1997.

[45] P. Kunkel, V. Mehrmann, M. Schmidt, I. Seufer, and A. Steinbrecher. Weak formulations of linear differential-algebraic systems. Technical Report 16, Institut für Mathematik, TU Berlin, Berlin, Germany, 2006.

[46] P. Kunkel, V. Mehrmann, and R. Stöver. Multiple shooting for unstructured nonlinear differential-algebraic equations of arbitrary index. *SIAM J. Numer. Anal.*, 42:2277–2297, 2004.

[47] P. Kunkel, V. Mehrmann, and R. Stöver. Symmetric collocation for unstructured nonlinear differential-algebraic equations of arbitrary index. *Numer. Math.*, 98:277–304, 2004.

[48] P. Kunkel and R. Stöver. Symmetric collocation methods for linear differential-algebraic boundary value problems. *Numer. Math.*, 91:475–501, 2002.

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

[50] J.-Y. Lin and Z.-H. Yang. Optimal control for singular systems. *Internat. J. Control*, 47:1915–1924, 1988.

[51] L. Ljusternik. On constrained extrema of functionals. *Math. Sb.*, 41:390–401, 1934. In Russian.

[52] R. März. Solvability of linear differential algebraic equations with properly stated leading terms. *Res. in Math.*, 45:88–105, 2004.

[53] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem.* Springer-Verlag, Berlin, Germany, 1991.

[54] P. C. Müller. Stability and optimal control of nonlinear descriptor systems. In *Proc. 3rd Int. Symp. Methods and Models in Automation and Robotics (MMAR 96)*, volume 1, pages 17–26, Szcezecin, Poland, 1996. Univ. of Szcezecin.

[55] M. Otter, H. Elmqvist, and S. E. Mattson. Multi-domain modeling with modelica. In Paul Fishwick, editor, *CRC Handbook of Dynamic System Modeling.* CRC Press, 2006. to appear.

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

[57] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioural Approach.* Springer-Verlag, New York, NY, 1998.

[58] P. J. Rabier and W. C. Rheinboldt. Classical and generalized solutions of time-dependent linear differential-algebraic equations. *Lin. Alg. Appl.*, 245:259–293, 1996.

[59] P. J. Rabier and W. C. Rheinboldt. *Theoretical and Numerical Analysis of Differential-Algebraic Equations*, volume VIII of *Handbook of Numerical Analysis.* Elsevier Publications, Amsterdam, The Netherlands, 2002.

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

[61] V. H. Schultz. *Reduced SQP methods for large-scale optimal control problems in DAE with application to path planning problems for satellite mounted robots.* Dissertation, Universität Heidelberg, Interdisz. Zentrum für wissenschaftliches Rechnen, 1996.

[62] R. Vinter. *Optimal Control*. Birkhäuser, Boston, MA, 2000.

[63] E. Zeidler. *Nonlinear Functional Analysis and Its Applications III. Variational Methods and Optimization*. Springer-Verlag, New-York, NY, 1985.