Numer Algor (2006) 42: 281–307
DOI 10.1007/s11075-006-9030-x

# Transformation of high order linear differential-algebraic systems to first order

**Volker Mehrmann · Chunchao Shi**

Received: 9 February 2006 / Accepted: 24 April 2006 /
Published online: 28 September 2006


**Abstract** We study general nonsquare linear systems of differential-algebraic systems of arbitrary order. We analyze the classical procedure of turning the system into a first order system and demonstrate that this approach may lead to different solvability results and smoothness requirements. We present several examples that demonstrate this phenomenon and then derive existence and uniqueness results for differential-algebraic systems of arbitrary order and index. We use these results to identify exactly those variables for which the order reduction to first order does not lead to extra smoothness requirements and demonstrate the effects of this new formulation with a numerical example.

**Keywords** differential-algebraic equation · high order system · order reduction · index reduction · strangeness index

**AMS(MOS) subject classification** 65L80 · 65L05 · 34A30

## 1. Introduction

We study general linear $l$th order systems of Differential-Algebraic Equations (DAEs) with variable coefficients

$$A_l(t)x^{(l)}(t) + A_{l-1}(t)x^{(l-1)}(t) + \cdots + A_0(t)x(t) = f(t), \tag{1}$$

Dedicated to Richard S. Varga on the occasion of his 77th birthday.

V. Mehrmann (✉)
Institut für Mathematik, MA 4–5 TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany
e-mail: mehrmann@math.tu-berlin.de.

C. Shi
12 Tenterden Crescent, Kents Hill, Milton Keynes, MK7 6HH, UK

in a real interval $\mathbb{I} \subset \mathbb{R}$, together with initial conditions

$$x(t_0) = x_0^{[0]}, \ldots, x^{(l-2)}(t_0) = x_0^{[l-2]}, x^{(l-1)}(t_0) = x_0^{[l-1]}, t_0 \in \mathbb{I}. \tag{2}$$

Here, the coefficients satisfy $A_i(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$, $i = 0, 1, \ldots, l$, $A_l(t) \not\equiv 0$, $x(t)$ is an unknown vector-valued function in $\mathcal{C}(\mathbb{I}, \mathbb{C}^n)$, and the right-hand side $f(t)$ is a given vector-valued function in $\mathcal{C}^k(\mathbb{I}, \mathbb{C}^m)$, where $\mathcal{C}^k(\mathbb{I}, \mathbb{C}^{m,n})$, $k \in \mathbb{N}_0$, denotes the set of all $k$-times continuously differentiable matrix-valued functions from the real interval $\mathbb{I}$ to the complex vector space $\mathbb{C}^{m,n}$ and $k$ is sufficiently large. In the following we will refer to DAEs with order $l$ greater than 1 simply as *high order* systems.

DAEs play a key role in the modeling and simulation of constrained dynamical systems in many applications. Such systems have been intensively studied, theoretically as well as numerically, in the past three decades. For a systematic and comprehensive exposition of important aspects regarding the theory, the numerical treatment and many applications of first order DAEs, see e.g. [3–5, 8, 12, 15, 16, 20, 28] and the references therein.

Linear high order DAEs arise from linearizations of general nonlinear high order DAEs of the form

$$F(t, x, \dot{x}, \ldots, x^{(l)}) = 0 \tag{3}$$

around reference solutions. Typical applications where second order DAEs arise naturally are multi-body systems, see [8, 28] or models of electrical circuits [13, 14].

Usually, in the classical theory of ordinary differential equations, high order systems are turned into first order systems by introducing new variables for the derivatives up to order $l-1$. There is no unique way of performing this transformation, and only recently for the case of constant coefficients (in the representation of matrix polynomials) a systematic theory for transformation to first order has been derived [23]. It has been indicated there, but also in several other publications, see [1, 7, 29], that the classical textbook approach of turning high order systems into first order form has to be performed with great care, since it may lead to substantial mathematical difficulties, in particular for DAEs.

Before we discuss this issue further, we have to mention the solvability concepts that we use here. Usually when we write down a differential equation where the $l$th derivative of the unknown variable $x$ occurs, then we require the solution to be $l$ times differentiable. On the other hand, if the leading term $A_l(t)$ is singular, then for functions in the kernel of $A_l(t)$ less smoothness is sufficient. One can use this fact to consider weaker solvability concepts that require only the minimum necessary differentiability. To analyze the exact degree of smoothness required for differential-algebraic systems is an active research topic, see [2, 24, 25], in this paper we will focus on the order reduction.

To motivate the analysis that we present below and to illustrate the potential difficulties, consider the following examples.

**Example 1.** The simple example of a mathematical pendulum

$$m\ddot{x} + 2x\lambda = 0,$$
$$m\ddot{y} + 2y\lambda + mg = 0,$$
$$x^2 + y^2 - l^2 = 0,$$

is a particular case of a multi-body system. In such systems it is common practice, [8] to derive a first order formulation by just introducing new variables $v_1 = \dot{x}_1$ and $v_2 = \dot{y}_2$ but not the derivative of $\lambda$. In this way an unnecessary differentiability requirement for the Lagrange multiplier $\lambda$ is avoided.

The order reduction described in Example 1 is a special case of the general procedure that we will derive below. If derivatives of the whole vector were introduced and if the degree of differentiability of the right-hand side $f(t)$ in a high order system is limited, then this classical transformation to first order may lead to different smoothness requirements for the inhomogeneity.

**Example 2.** Consider the linear second order constant coefficient DAE

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ddot{x}(t) + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \dot{x}(t) + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) = f(t), \quad t \in \mathbb{I}, \tag{4}$$

where $x(t) = [x_1(t), x_2(t)]^T$, and $f(t) = [f_1(t), f_2(t)]^T$. System (4) has the unique solution

$$\begin{aligned} x_1(t) &= f_2(t), \\ x_2(t) &= f_1(t) - \dot{f}_2(t) - \ddot{f}_2(t). \end{aligned} \tag{5}$$

Here it is sufficient to require that $x_1$ is twice differentiable, while for $x_2$ continuity is enough. In view of this, the minimum smoothness requirement for the inhomogeniety $f$ is that $f_1$ is continuous and $f_2$ is twice continuously differentiable. Using the classical transformation to first order

$$v(t) = [v_1(t), v_2(t)]^T = [\dot{x}_1(t), \dot{x}_2(t)]^T, \; y(t) = [v_1(t), v_2(t), x_1(t), x_2(t)]^T,$$

we need

$$\begin{aligned} v_1(t) &= \dot{f}_2(t), \\ v_2(t) &= \dot{f}_1(t) - \ddot{f}_2(t) - f_2^{(3)}(t) \end{aligned} \tag{6}$$

and thus, $f_2$ has to be three times continuously differentiable. If, however, we only introduce $v_1 = \dot{x}_1$, then no extra differentiability requirements are needed.

We see from Example 2 that the classical approach of introducing the derivative of the complete unknown function $x(t)$ may lead to higher smoothness requirements for the inhomogeneity. The example also shows that by introducing only some new variables this difficulty can be circumvented. But exactly which linear combinations of variables can be introduced in a general high order system to obtain a first order system with the same smoothness requirements, at first look is not clear.

**Example 3.** Consider the linear second order constant coefficient DAE

$$\begin{bmatrix} I & t+1 \\ t & t^2+t \end{bmatrix} \ddot{x} + \begin{bmatrix} 0 & 2 \\ 0 & 2t \end{bmatrix} \dot{x} + \begin{bmatrix} 1 & t \\ 1+t & 1+t+t^2 \end{bmatrix} x = f(t), \quad t \in \mathbb{I}, \tag{7}$$

where $x = [x_1, x_2]^T$, and $f(t) = [f_1(t), f_2(t)]^T$. System (7) has the unique solution (leaving off arguments)

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (1+t+t^2) f_1 - t f_2 + (2+2t) \dot{f}_1 - (1+t) \ddot{f}_2 + (t+t^2) \ddot{f}_1 \\ f_2 - (t+1) f_1 - 2 \dot{f}_1 - t \ddot{f}_1 + \ddot{f}_2 \end{bmatrix},$$

i.e., $f_1$, $f_2$ have to be four times continuously differentiable. Here, by introducing the new variable $v = \dot{x}_1 + (1+t)\dot{x}_2$ we obtain a first order form for which $f_1$, $f_2$ only have to be three times continuously differentiable.

The analysis that we describe below will present an equivalent formulation of the differential-algebraic equation from which a first order form can be constructed that minimizes the smoothness requirements.

Another difficulty that arises in practical numerical methods for the solution of high order systems is that the system may be badly scaled and one also has disturbances and perturbations in the data (see the concept of perturbation index in [16]), so that the transformation to first order may lead to very different solutions in the perturbed system.

**Example 4.** Consider the second order system

$$\epsilon_1 \ddot{x}(t) + \epsilon_2 \dot{x}(t) + \epsilon_3 x(t) = \epsilon_4 f(t), \quad t \in \mathbb{I}, \tag{8}$$

with coefficients $\epsilon_i$, $i = 1, \ldots, 4$ of absolute value close or smaller than the machine precision and $f$ of norm approximately 1. If we transform (8) to first order in the classical way by introducing

$$y(t) = [y_1(t), y_2(t)]^T := [\dot{x}(t), x(t)]^T,$$

then we obtain the system

$$\begin{bmatrix} \epsilon_1 & 0 \\ 0 & 1 \end{bmatrix} \dot{y}(t) + \begin{bmatrix} \epsilon_2 & \epsilon_3 \\ -1 & 0 \end{bmatrix} y(t) = \begin{bmatrix} \epsilon_4 f(t) \\ 0 \end{bmatrix}. \tag{9}$$

For different values of the $\epsilon_i$, in finite precision arithmetic, we may decide that the system (9) has a unique solution, no solutions at all or is actually a underdetermined system.

It is well known that for differential-algebraic equations of differentiation index higher than 1 problems may arise in the numerical solution. These difficulties can result e.g. in an order reduction or even divergence, or a drift from the hidden manifolds [3, 8, 16] as well other instabilities [1, 29]. The classical transformation to first order form may add to these problems and these problems can be partially avoided by a direct discretization of the second order formulation. An extreme example that demonstrates this effect was given in [1, 29].

**Example 5.**

$$\begin{bmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{bmatrix} = 2\begin{bmatrix} y_2 \\ -y_1 \end{bmatrix} + \lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$
$$0 = y_1^2 + y_2^2 - 1.$$

The classical first order representation is

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$
$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = 2\begin{bmatrix} y_2 \\ -y_1 \end{bmatrix} + \lambda \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$
$$0 = y_1^2 + y_2^2 - 1.$$

If one discretizes this first order system with the implicit Euler method then one obtains the absolute errors given in table 1. If this problem is discretized directly in the second order form or if one only introduces new variables for the derivatives $\dot{y}_1$ and $\dot{y}_2$ as is common in the classical methods for multibody systems, then the implicit Euler and also other BDF methods perform fine.

Motivated by the described difficulties, we will investigate the analytical properties of general high order linear DAEs by extending the algebraic techniques derived for first order systems in [18, 19] to general nonsquare high order systems. We show, in particular, how we can transform a high order system to first order without changing the smoothness requirements for the inhomogeneity?

The paper is organized as follows. In Section 2 we discuss equivalence transformations for tuples of matrix functions and we recall some results on condensed forms for pairs of matrix functions from [18, 19]. In Section 3 we then present condensed forms for triples of matrix functions that are associated with systems of second order DAEs. In Section 4, the results for second order systems are then extended to general high order systems. We finish with a numerical example and some conclusions.

**Table 1** Absolute errors of implicit Euler method for variable $\lambda$.

| Step number | Stepsize $\times 10^{-3}$ | $\lvert\lambda(t_j) - \lambda_j\rvert$ implicit Euler |
|---|---|---|
| 1 | 1.000 | **2.0080** |
| 2 | 1.000 | 0.0080 |
| 3 | 0.200 | **8.0303** |
| 4 | 0.040 | **8.0348** |
| 5 | 0.008 | **8.0357** |
| 6 | 0.008 | 0.0001 |
| 7 | 0.016 | **1.0047** |
| 8 | 0.032 | **1.0048** |
| 9 | 0.064 | **1.0052** |
| 10 | 0.064 | 0.0006 |

## 2. Preliminaries

In this section we discuss different types of equivalence relations and associated condensed forms and we recall some results for pairs of matrix functions.

It is well-known that the nature of the solutions of linear first order constant coefficient DAEs

$$A_1\dot{x}(t) + A_0x(t) = f(t), \quad A_0, A_1 \in \mathbb{C}^{m,n}$$

can be determined by the algebraic properties of the corresponding matrix pencil $\lambda A_1 + A_0$, which follow from the *Kronecker canonical form* for matrix pencils under the equivalence transformation

$$\lambda(PA_1Q) + (PA_0Q), \tag{10}$$

where $P \in \mathbb{C}^{m,m}$ and $Q \in \mathbb{C}^{n,n}$ are any nonsingular matrices, see, e.g. [3, 9].

Unfortunately, these results do not carry over to high order systems, since it is a well-known open problem [11, 31] to find a corresponding canonical form for matrix polynomials of degree higher than 1.

But for the analysis of the solvability of linear first order DAEs, the complete information from these canonical forms is not necessary. To obtain solvability results, we will extend results of [18, 19] for matrix pencils to matrix tuples $(A_l, \ldots, A_1, A_0)$ and tuples $(A_l(t), \ldots, A_1(t), A_0(t))$ of matrix functions. For this we need three types of equivalence relations. The first one generalizes the classical equivalence for matrix pairs to matrix tuples.

**Definition 6.** Two tuples of matrices $(A_l, \ldots, A_1, A_0)$ and $(B_l, \ldots, B_1, B_0)$, $A_i, B_i \in \mathbb{C}^{m,n}$, $i = 0, 1, \ldots, l$, $l \in \mathbb{N}_0$, are called *strongly equivalent* if there exist nonsingular matrices $P \in \mathbb{C}^{m,m}$ and $Q \in \mathbb{C}^{n,n}$ such that

$$B_i = PA_iQ, \quad i = 0, 1, \ldots, l. \tag{11}$$

We then write

$$(A_l, \ldots, A_1, A_0) \sim (B_l, \ldots, B_1, B_0).$$

In the case of $l$th order systems, i.e. for tuples of matrix valued functions, we have the following definition of *global equivalence*, which results from scaling the equation by a nonsingular matrix $P(t)$ and carrying out a change of basis with a nonsingular matrix $Q(t)$.

**Definition 7.** Two tuples of matrix-valued functions $(A_l(t), \ldots, A_1(t), A_0(t))$ and $(B_l(t), \ldots, B_1(t), B_0(t))$ with $A_i(t), B_i(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$, $i = 0, 1, \ldots, l$ are called *glob-*

*ally equivalent* if there exist pointwise nonsingular matrix-valued functions $P(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,m})$ and $Q(t) \in \mathcal{C}^l(\mathbb{I}, \mathbb{C}^{n,n})$ such that

$$
\begin{aligned}
&[B_l(t), \ldots, B_0(t)] \\
&= P(t)[A_l(t), \ldots, A_0(t)] \begin{bmatrix} Q(t) & \binom{l}{1} \frac{d}{dt} Q(t) & \cdots & \cdots & \binom{l}{1} \frac{d^l}{dt^l} Q(t) \\ & Q(t) & \binom{l-1}{1} \frac{d}{dt} Q(t) & \cdots & \binom{l-1}{l-1} \frac{d^{l-1}}{dt^{l-1}} Q(t) \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \binom{l}{1} \frac{d}{dt} Q(t) \\ & & & & Q(t) \end{bmatrix},
\end{aligned} \tag{12}
$$

$i, j \in \mathbb{N}$, $i \leq j$. If this it is clear from the context, then we still write $(A_l(t), \ldots, A_1(t), A_0(t)) \sim (B_l(t), \ldots, B_1(t), B_0(t))$.

It is well-known, see e.g. [10, 18], that strong equivalence is not the associated local (in the neighborhood of a fixed point $\hat{t}$) version of global equivalence for variable coefficient systems. Since due to the Theorem of Hermite interpolation, see e.g. [30], for any nonsingular matrices $\hat{P} \in \mathbb{C}^{m,m}$, $\hat{Q} \in \mathbb{C}^{n,n}$ and any matrices $R_1, \ldots, R_l \in \mathbb{C}^{n,n}$, we can always find pointwise nonsingular matrix functions $P(t)$ and $Q(t)$ such that $P(\hat{t}) = \hat{P}$, $Q(\hat{t}) = \hat{Q}$ and $\frac{d^i}{dt^i} Q(t) = R_i$, $i = 1, \ldots, l$, we obtain the following definition.

**Definition 8.** Two tuples of matrices $(A_l, \ldots, A_1, A_0)$ and $(B_l, \ldots, B_1, B_0)$ with $A_i, B_i \in \mathbb{C}^{m,n}$, $i = 0, 1, \ldots, l$ are called *locally equivalent* if there exist matrices $P \in \mathbb{C}^{m,m}$ nonsingular, $Q \in \mathbb{C}^{n,n}$ nonsingular and $R_1, \ldots, R_l \in \mathbb{C}^{n,n}$ such that

$$
\begin{aligned}
&[B_l, \ldots, B_0] \\
&= P[A_l, \ldots, A_0] \begin{bmatrix} Q & \binom{l}{1} R_1 & \cdots & \cdots & \binom{l}{1} R_1 \\ & Q & \binom{l-1}{1} R_1 & \cdots & \binom{l-1}{l-1} R_{l-1} \\ & & \ddots & \ddots & \vdots \\ & & & Q & \binom{l}{1} R_1 \\ & & & & Q \end{bmatrix}.
\end{aligned} \tag{13}
$$

As already suggested by the definitions, the relations (11), (12) and (13) are equivalence relations, see Appendix A of [27] for detailed proofs. The canonical form for pairs of matrices or matrix functions under local and global equivalence has been derived in [18, 19]. For completeness and to compare with the results in the high order case we recall these results.

**Theorem 9.** [18, 19] Let $A_0, A_1 \in \mathbb{C}^{m,n}$ and let the columns of

$$\begin{array}{ll} (a) & T \text{ form a basis of kernel } A_1, \\ (b) & Z \text{ form a basis of corange } A_1 = \text{kernel } A_1^H, \\ (c) & T' \text{ form a basis of cokernel } A_1 = \text{range } A_1^H, \\ (d) & V \text{ form a basis of corange}(Z^H A_0 T). \end{array}$$

Then the quantities (with the convention that the rank of an empty matrix is 0)

$$\begin{array}{lll} (a) & r = \text{rank } A_1 & \text{(rank)} \\ (b) & a = \text{rank}(Z^H A_0 T) & \text{(algebraic part)} \\ (c) & s = \text{rank}(V^H Z^H A_0 T') & \text{(strangeness)} \\ (d) & d = r - s & \text{(differential part)} \\ (e) & u = n - r - a & \text{(undetermined part)} \\ (f) & v = m - d - a - s & \text{(vanishing part)} \end{array}$$

are invariant under local equivalence (13) and $(A_1, A_0)$ is locally equivalent to the canonical form

$$\left( \begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) \begin{array}{l} s \\ d \\ a, \\ s \\ v \end{array} \tag{14}$$

where the last block column in both matrices has width $u$.

A slight modification of this result leads to the following condensed form for pairs of matrices under strong equivalence.

**Theorem 10.** Let $A_0, A_1 \in \mathbb{C}^{m,n}$, and let the columns of

$$\begin{array}{ll} (a) & Z_1 \in \mathbb{C}^{m,m-r} \text{ form a basis for kernel}(A_1^T), \\ (b) & Z_2 \in \mathbb{C}^{n,n-r} \text{ form a basis for kernel}(A_1). \end{array} \tag{15}$$

Then, the matrix pair $(A_1, A_0)$ is strongly equivalent to a matrix pair of the form

$$\left( \begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{12} & 0 & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right) \begin{array}{l} s \\ d \\ a, \\ s \\ v \end{array} \tag{16}$$

where $s, d, a, u, v \in \mathbb{N}_0$, the last block column has size $u$, and the quantities

$$\begin{array}{ll} (a) & r = \text{rank}(A_1) \\ (b) & a = \text{rank}(Z_1^T A_0 Z_2) \\ (c) & s = \text{rank}(Z_1^T A_0) - a \\ (d) & d = r - s \\ (e) & u = n - r - a \\ (f) & v = m - r - a - s \end{array} \tag{17}$$

are invariant under the strong equivalence relation (11).

*Proof.* The proof is similar to the corresponding result in [19], just leaving out all the elimination steps that are not allowed in strong equivalence, see Appendix B in [27] for details. □

For matrix functions we may apply the local equivalence globally in a neighborhood of a fixed point $\hat{t}$.

**Theorem 11.** [18, 19] Let $A_1, A_0 \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$ be sufficiently smooth and suppose that

$$r(t) \equiv r, \quad a(t) \equiv a, \quad s(t) \equiv s \tag{18}$$

for the local characteristic values of $(A_1(t), A_0(t))$ in the neighborhood of a fixed point $\hat{t} \in \mathbb{I}$. Then, $(A_1(t), A_0(t))$ is globally equivalent to the condensed form

$$\begin{array}{cc} \left(\begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{1,2}(t) & 0 & A_{1,4}(t) \\ 0 & 0 & 0 & A_{2,4}(t) \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) & \begin{array}{l} s \\ d \\ a\,. \\ s \\ v \end{array} \end{array} \tag{19}$$

All entries $A_{i,j}(t)$ are matrix functions on $\mathbb{I}$ and the last block column in both matrices has size $u = n - s - d - a$.

Theorem 11 is not yet sufficient to explain the solution behavior of linear differential-algebraic systems of first order. For this, consider the system of differential-algebraic equations that corresponds to (19) (in the transformed variables). We get

$$\begin{aligned} &(a)\ \dot{x}_1(t) = A_{12}(t)x_2(t) + A_{14}(t)x_4(t) + f_1(t) \\ &(b)\ \dot{x}_2(t) = A_{24}(t)x_4(t) + f_2(t) \\ &(c)\ \ 0 = x_3(t) + f_3(t) \\ &(d)\ \ 0 = x_1(t) + f_4(t) \\ &(e)\ \ 0 = f_5(t). \end{aligned} \tag{20}$$

Here, we can insert the derivative of (20d) in (20a), which then becomes an algebraic equation. This corresponds to passing from (19) to:

$$\left(\begin{bmatrix} I_s & 0 & 0 & 0 \\ 0 & I_d & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & A_{12}(t) & 0 & A_{14}(t) \\ 0 & 0 & 0 & A_{24}(t) \\ 0 & 0 & I_a & 0 \\ I_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right), \tag{21}$$

for which we again compute characteristic values $r, a, s, d, u, v$.

Applying this step inductively, one obtains an inductive definition of a sequence of pairs of matrix functions $(A_1^i(t), A_0^i(t)),\ i \in \mathbb{N}_0$, where $(A_0^0(t), A_1^0(t)) = (A_1(t), A_0(t))$ and $(A_1^{i+1}(t), A_0^{i+1}(t))$ is derived from $(A_1^i(t), A_0^i(t))$ by bringing it into the form (19) and passing then to (21). Here we must assume (18) for every occurring pair of matrices. For pairs of matrix functions and also in the nonlinear case these assumptions can be significantly relaxed, see [20].

Connected with this iterative process, one then has sequences $r_i(t) \equiv r_i$, $a_i(t) \equiv a_i$, $s_i(t) \equiv s_i$, $d_i(t) \equiv d_i$, $u_i(t) \equiv u_i$, $v_i(t) \equiv v_i$, which are characteristic for the given pair $(A_1(t), A_0(t))$, that is, they do not depend on the specific way they are obtained. Note that this process stops after finitely many (say $\mu(t) \equiv \mu$) steps with $s_i = 0$. The quantity $\mu$ is called the *strangeness index* of the pencil $(A_1(t), A_0(t))$.

In the next section we generalize these results to the case of triples of matrix valued functions arising from second order systems.

## 3. Condensed forms for triples of matrices and matrix functions

We begin the analysis with systems of linear second order DAEs with constant coefficients

$$A_{pt2}\ddot{x}(t) + A_1\dot{x}(t) + A_0x(t) = f(t), t \in \mathbb{I}, \tag{22}$$

with $A_2, A_1, A_0 \in \mathbb{C}^{m,n}$, $f(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^m)$ sufficiently smooth, together with initial conditions

$$x(t_0) = x_0^{[0]}, \dot{x}(t_0) = x_0^{[1]}, \quad x_0^{[0]}, x_0^{[1]} \in \mathbb{C}^n. \tag{23}$$

Similarly as in the case of first order systems, the behavior of the system (22) (and the initial value problem (22)–(23)) depends on the properties of the quadratic matrix polynomial

$$A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0.$$

If we apply a strong equivalence transformation with nonsingular matrices $P \in \mathbb{C}^{m,m}$ and $Q \in \mathbb{C}^{n,n}$ then we obtain a transformed quadratic matrix polynomial

$$\hat{A}(\lambda) = \lambda^2\hat{A}_2 + \lambda\hat{A}_1 + \hat{A}_0 := \lambda^2(PA_2Q) + \lambda(PA_1Q) + (PA_0Q). \tag{24}$$

As we have already mentioned, it is a well-known open problem [11, 31] to find a canonical form for quadratic matrix polynomials (24) under strong equivalence. However, as for first order systems, we do not need the complete canonical form to understand the solution behavior of the corresponding DAE.

**Theorem 12.** Let $A_2, A_1, A_0 \in \mathbb{C}^{m,n}$. Then, $(A_2, A_1, A_0)$ is strongly equivalent to a matrix triple $(\hat{A}_2, \hat{A}_1, \hat{A}_0)$ of the following form

$$\left( \begin{bmatrix} I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{s^{(0,2)}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{d^{(2)}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \right.$$

$$\begin{bmatrix} 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & * & I_{s^{(0,1)}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{d^{(1)}} & 0 & 0 \\ I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{25}$$

$$\left. \begin{bmatrix} 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & I_a & 0 \\ 0 & 0 & 0 & 0 & I_{s^{(0,1)}} & 0 & 0 & 0 \\ 0 & 0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 \\ I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) \begin{array}{l} s^{(0,1,2)} \\ s^{(1,2)} \\ s^{(0,2)} \\ d^{(2)} \\ s^{(0,1)} \\ d^{(1)} \\ s^{(0,1,2)} \\ s^{(1,2)} \\ a \\ s^{(0,1)} \\ s^{(0,2)} \\ s^{(0,1,2)} \\ v \end{array}$$

where $s^{(0,1,2)}, s^{(1,2)}, s^{(0,2)}, s^{(0,1)}, d^{(2)}, d^{(1)}, a, v$ and $u$ (the size of the last block column) are in $\mathbb{N}_0$, and the entries marked with * are blocks of matrices which are not specified.

*Proof.* The proof is a straighforward but rather technical extension of the matrix pencil case, it is given in Appendix C of [27]. □

Each of the integer quantities in Theorem 12 has an expression in terms of dimensions of column spaces or ranks of matrices, and is invariant under strong equivalence as the next lemma shows.

**Lemma 13.** Let $A_2, A_1, A_0 \in \mathbb{C}^{m,n}$ and let the columns of

$$\begin{aligned}
&(a)\ \ Z_1 \text{ form a basis for kernel}(A_2^T),\\
&(b)\ \ Z_2 \text{ form a basis for kernel}(A_2),\\
&(c)\ \ Z_3 \text{ form a basis for kernel}(A_2^T) \cap \text{kernel}(A_1^T),\\
&(d)\ \ Z_4 \text{ form a basis for kernel}(A_2) \cap \text{kernel}(Z_1^T A_1).
\end{aligned} \tag{26}$$

Then the quantities

$$\begin{aligned}
&(a)\ \ r = \text{rank}(A_2) && \text{(rank of } A_2)\\
&(b)\ \ a = \text{rank}(Z_3^T A_1 Z_4) && \text{(algebraic part)}\\
&(c)\ \ s^{(0,1,2)} = \dim\big(\mathcal{R}(A_2^T) \cap \mathcal{R}(A_1^T Z_1) \\
&\qquad\qquad \cap \mathcal{R}(A_0^T Z_3)\big) && \text{(strangeness due to } A_2, A_1, A_0)\\
&(d)\ \ s^{(0,1)} = \text{rank}(Z_3^T A_0 Z_2) - a && \text{(strangeness due to } A_1, A_0)\\
&(e)\ \ d^{(1)} = \text{rank}(Z_1^T A_1 Z_2) - s^{(0,1)} && \text{(1st order differential part)}\\
&(f)\ \ s^{(1,2)} = \text{rank}(Z_1^T A_1) - s^{(0,1,2)} \\
&\qquad\qquad -s^{(0,1)} - d^{(1)} && \text{(strangeness due to } A_2, A_1)\\
&(g)\ \ s^{(0,2)} = \text{rank}(Z_3^T A_0) - a \\
&\qquad\qquad -s^{(0,1,2)} - s^{(0,1)} && \text{(strangeness due to } A_2, A_0)\\
&(h)\ \ d^{(2)} = r - s^{(0,1,2)} - s^{(1,2)} - s^{(0,2)} && \text{(2nd order differential part)}\\
&(i)\ \ v = m - r - 2s^{(0,1)} - d^{(1)} - 2s^{(0,1,2)} \\
&\qquad\qquad -s^{(1,2)} - a - s^{(0,2)} && \text{(vanishing equations)}\\
&(j)\ \ u = n - r - s^{(0,1)} - d^{(1)} - a && \text{(undetermined part)}
\end{aligned} \tag{27}$$

are invariant under the strong equivalence relation (11) and $(A_2, A_1, A_0)$ is strongly equivalent to the form (25).

*Proof.* The proof follows in a similar fashion as the proof of strong equivalence for matrix pairs.

*Step 1.* First, we show that the quantities in (27) are well-defined with respect to the choices of the bases in (26). We take $a = \text{rank}(Z_3^T A_0 Z_4)$ as an example. Every change of basis can be represented by

$$\tilde{Z}_3 = Z_3 Q_1, \quad \tilde{Z}_4 = Z_4 Q_2$$

with nonsingular matrices $Q_1, Q_2$. From

$$\text{rank}(\tilde{Z}_3^T A_0 \tilde{Z}_4) = \text{rank}(Q_1^T Z_3^T A_0 Z_4 Q_2) = \text{rank}(Z_3^T K Z_4),$$

it then follows that $\text{rank}(Z_1^T A_1 Z_2)$ is well-defined. Similarly, we can prove that the other quantities in (27) are also well-defined.

*Step 2.* Next, we show that the quantities in (27) are invariant under strong equivalence. Here, we just take $s^{(0,1,2)}$ as an example. Let $(A_2, A_1, A_0)$ and $(\tilde{A}_2, \tilde{A}_1, \tilde{A}_0)$ be strongly equivalent, i.e. there exist nonsingular matrices $P$ and $Q$, such that

$$\tilde{A}_2 = PA_2Q, \quad \tilde{A}_1 = PA_1Q, \quad \tilde{A}_0 = PA_0Q. \tag{28}$$

Let the columns of $\tilde{Z}_1$ form a basis for kernel$(\tilde{A}_2^T)$, and let the columns of $\tilde{Z}_3$ form a basis for kernel$(\tilde{A}_2^T) \cap$ kernel$(\tilde{A}_1^T)$. Then, from (28) it follows that the columns of $Z_1 := P^T\tilde{Z}_1$ form a basis for kernel$(A_2^T)$, and the columns of $Z_3 := P^T\tilde{Z}_3$ form a basis for kernel$(A_2^T) \cap$ kernel$(A_1^T)$. Thus, the invariance of $s^{(0,1,2)}$ follows from

$$\begin{aligned}
\tilde{s}^{(0,1,2)} &= \dim\Big(\mathcal{R}\big(\tilde{A}_2^T\big) \cap \mathcal{R}\big(\tilde{A}_1^T\tilde{Z}_1\big) \cap \mathcal{R}\big(\tilde{A}_0^T\tilde{Z}_3\big)\Big) \\
&= \dim\Big(\mathcal{R}\big(Q^TA_2^TP^T\big) \cap \mathcal{R}\big(Q^TA_1^TP^T\tilde{Z}_1\big) \cap \mathcal{R}\big(Q^TA_0^TP^T\tilde{Z}_3\big)\Big) \\
&= \dim\Big(\mathcal{R}\big(A_2^TP^T\big) \cap \mathcal{R}\big(A_1^TP^T\tilde{Z}_1\big) \cap \mathcal{R}\big(A_0^TP^T\tilde{Z}_3\big)\Big) \\
&= \dim\big(\mathcal{R}\big(A_2^T\big) \cap \mathcal{R}\big(A_1^TZ_1\big) \cap \mathcal{R}\big(A_0^TZ_3\big)\big) \\
&= s^{(0,1,2)}.
\end{aligned}$$

Similarly, the invariance of the other quantities in (27) can be proved.

*Step 3.* Finally, we show that the quantities in the equivalent form (25) of $(A_2, A_1, A_0)$ are identical with those defined in (27). Let $P \in \mathbb{C}^{m,m}$, $Q \in \mathbb{C}^{n,n}$ be nonsingular matrices such that

$$(\hat{A}_2, \hat{A}_1, \hat{A}_0) = (PA_2Q, PA_1Q, PA_0Q),$$

where $(\hat{A}_2, \hat{A}_1, \hat{A}_0)$ is of the form (25). Furthermore, let $P$ and $Q$ be partitioned as $P := \big[P_1^T, P_2^T, \ldots, P_{13}^T\big]^T$ and $Q := \big[Q_1, Q_2, \ldots, Q_8\big]$ conformably with (25), respectively. Then, by (25), we have

$$\begin{aligned}
\big[P_5^T, \ldots, P_{13}^T\big]^T A_2 &= 0, \\
A_2\big[Q_5, \ldots, Q_8\big] &= 0, \\
\big[P_9^T, \ldots, P_{13}^T\big]^T A_1 &= 0, \\
\big[P_5^T, \ldots, P_{13}^T\big]^T A_1\big[Q_7, Q_8\big] &= 0,
\end{aligned}$$

namely the columns of $\big[P_5^T, \ldots, P_{13}^T\big]$ form a basis of kernel$(A_2^T)$, the columns of $\big[Q_5, \ldots, Q_8\big]$ form a basis of kernel$(A_2)$, the columns of $\big[P_9^T, \ldots, P_{13}^T\big]$ form a basis of kernel$(A_2^T) \cap$ kernel$(A_1^T)$, and the columns of $\big[Q_7, \ldots, Q_8\big]$ form a basis of kernel$(A_2) \cap$ kernel$\left(\big[P_5^T, \ldots, P_{13}^T\big]\right)^T A_1)$. Observing that, by (25),

$$\begin{bmatrix} P_9 \\ \vdots \\ P_{13} \end{bmatrix} A_0\big[Q_7, Q_8\big] = \begin{bmatrix} I_a & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

we have $a = \operatorname{rank}(Z_3^TA_0Z_4)$. Similarly, we can prove that the other quantities in the equivalent form (25) are equal to those defined in (27). □

Next, we consider linear second order DAEs with variable coefficients

$$A_2(t)\ddot{x}(t) + A_1(t)\dot{x}(t) + A_0(t)x(t) = f(t), \quad t \in \mathbb{I}, \tag{29}$$

where $A_2(t), A_1(t), A_0(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$, $f(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^m)$.

**Lemma 14.** Under the assumptions of Theorem 12, the quantities in (27) are invariant under the local equivalence relation (13) (l=2) and the triple $(A_2, A_1, A_0)$ is locally equivalent to the form (25).

*Proof.* Since the strong equivalence relation (11) is a special case of the local equivalence relation (13) by setting $R_i = 0$, $i = 1, \ldots, 2$, it follows form Theorem 12 that $(A_2, A_1, A_0)$ is locally equivalent to the form (25). It remains to show that the quantities in (27) are invariant under the local equivalence relation (13). Consider $s^{(0,1,2)}$ and let $(A_2, A_1, A_0)$ and $(\tilde{A}_2, \tilde{A}_1, \tilde{A}_0)$ be locally equivalent. Let the columns of $\tilde{Z}_1$ form a basis for kernel$(\tilde{A}_2^T)$, and let the columns of $\tilde{Z}_3$ form a basis for kernel$(\tilde{A}_2^T) \cap$ kernel$(\tilde{A}_1^T)$. Then, from (13) it follows that the columns of $Z_1 := P^T\tilde{Z}_1$ form a basis for kernel$(A_2^T)$. Since for any $z \in \tilde{Z}_3$, and any matrix $R_1$ of appropriate size

$$Q^T A_2^T P^T z = 0, \quad Q^T A_1^T P^T z + 2R_1^T A_2^T P^T z = 0,$$

if and only if

$$A_2^T P^T z = 0, \quad A_1^T P^T z = 0,$$

it follows that the columns of $Z_3 := P^T\tilde{Z}_3$ form a basis for kernel$(A_2^T) \cap$ kernel$(A_1^T)$. Thus, the invariance of $s^{(0,1,2)}$ follows from

$$\begin{aligned}
\tilde{s}^{(0,1,2)} &= \dim\Big(\mathcal{R}\big(\tilde{A}_2^T\big) \cap \mathcal{R}\big(\tilde{A}_1^T\tilde{Z}_1\big) \cap \mathcal{R}\big(\tilde{A}_0^T\tilde{Z}_3\big)\Big) \\
&= \dim\Big(\mathcal{R}\big(Q^T A_1^T P^T\big) \cap \mathcal{R}\big(Q^T A_0^T P^T \tilde{Z}_1 + 2R_1^T A_2^T P^T \tilde{Z}_1\big) \\
&\qquad \cap\mathcal{R}\big(Q^T A_0^T P^T \tilde{Z}_3 + R_1^T A_1^T P^T \tilde{Z}_3 + R_2^T A_2^T P^T \tilde{Z}_3\big)\Big) \\
&= \dim\Big(\mathcal{R}\big(Q^T A_2^T P^T\big) \cap \mathcal{R}\big(Q^T A_1^T P^T \tilde{Z}_1\big) \cap \mathcal{R}\big(Q^T A_0^T P^T \tilde{Z}_3\big)\Big) \\
&= \dim\Big(\mathcal{R}\big(A_2^T P^T\big) \cap \mathcal{R}\big(A_1^T P^T \tilde{Z}_1\big) \cap \mathcal{R}\big(A_0^T P^T \tilde{Z}_3\big)\Big) \\
&= \dim\big(\mathcal{R}\big(A_2^T\big) \cap \mathcal{R}\big(A_1^T Z_1\big) \cap \mathcal{R}\big(A_0^T Z_3\big)\big) \\
&= s^{(0,1,2)}.
\end{aligned}$$

Similarly, the invariance of the other quantities in (27) can be proved. □

For triples $(A_2(t), A_1(t), A_0(t))$ of matrix-valued functions we can then calculate, based on Lemma 14, the characteristic quantities in (27) for $(A_2(\hat{t}), A_1(\hat{t}), A_0(\hat{t}))$ at any fixed value $\hat{t} \in \mathbb{I}$. We obtain functions

$$r, a, s^{(0,1,2)}, s^{(0,1)}, d^{(1)}, s^{(1,2)}, s^{(0,2)}, d^{(2)}, u, v : \mathbb{I} \to \mathbb{N}_0.$$

and we assume that in $\mathbb{I}$:

$$\begin{aligned}
&r(t) \equiv r, a(t) \equiv a, s^{(0,1,2)}(t) \equiv s^{(0,1,2)}, s^{(0,1)}(t) \equiv s^{(0,1)}, \\
&d^{(1)}(t) \equiv d^{(1)}, s^{(1,2)}(t) \equiv s^{(1,2)}, s^{(0,2)}(t) \equiv s^{(0,2)},
\end{aligned} \tag{30}$$

By (27) and (30), it then follows that $d^{(2)}(t), u(t), v(t)$ are also constant on $\mathbb{I}$ and we obtain the following global condensed form. (For convenience of expression, we drop the subscripts of the blocks and omit the argument $t$ unless they are needed for clarification.)

**Lemma 15.** Let $A_2(t), A_1(t), A_0(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$ be sufficiently smooth, and suppose that the conditions (30) hold for the local characteristic values of $(A_2(t), A_1(t), A_0(t))$. Then, $(A_2(t), A_1(t), A_0(t))$ is globally equivalent to a matrix-valued triple $(\hat{A}_2(t), \hat{A}_1(t), \hat{A}_0(t))$ of the condensed form

$$
\left(
\begin{bmatrix}
I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & I_{s^{(0,2)}} & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & I_{d^{(2)}} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
0 & 0 & * & * & 0 & 0 & * & *\\
0 & 0 & * & * & 0 & 0 & * & *\\
0 & 0 & * & * & 0 & 0 & * & *\\
0 & 0 & * & * & 0 & 0 & * & *\\
0 & 0 & 0 & * & I_{s^{(0,1)}} & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & I_{d^{(1)}} & 0 & 0\\
I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & * & 0 & * & 0 & * & 0 & *\\
0 & 0 & 0 & 0 & 0 & 0 & I_a & 0\\
0 & 0 & 0 & 0 & I_{s^{(0,1)}} & 0 & 0 & 0\\
0 & 0 & I_{s^{(0,2)}} & 0 & 0 & 0 & 0 & 0\\
I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\right)
\begin{array}{l}
s^{(0,1,2)}\\ s^{(1,2)}\\ s^{(0,2)}\\ d^{(2)}\\ s^{(0,1)}\\ d^{(1)}\\ s^{(0,1,2)}\\ s^{(1,2)}\\ a\\ s^{(0,1)}\\ s^{(0,2)}\\ s^{(0,1,2)}\\ v
\end{array}
\qquad (31)
$$

All blocks in (31) are again matrix-valued functions on $\mathbb{I}$.

*Proof.* The rather technical and lengthy proof of Lemma 15 is a straightforward generalization of the proof for pairs of matrix functions. It is given in Appendix D of [27]. □

It should be noted that the equivalence transformations in Lemma 15 constitute a one-to-one relationship between the solution sets of the associated differential-algebraic systems. Note also that the block in position (5, 4) of $\hat{A}^1(t)$ in (31) satisfies $A^1_{5,4}(t) \equiv 0$, whereas the corresponding block $A^1_{5,4}$ in (25) may be a nonzero matrix, which is the major difference between condensed forms (31) and (25). This difference is due to the global equivalence which allows to eliminate $A^1_{5,4}(t)$ by solving an initial value problem for a linear ordinary differential equation. If we consider the associated system of DAEs

$$\hat{A}_2(t)\ddot{y}(t) + \hat{A}_1(t)\dot{y}(t) + \hat{A}_0(t)y(t) = \hat{f}(t), \tag{32}$$

then we obtain the equations

$$\begin{aligned}
&(a)\quad \ddot{y}_1(t) + \textstyle\sum_{i=3,4,7,8} A^1_{1,i}(t)\dot{y}_i(t) + \sum_{i=2,4,6,8} A^0_{1,i}(t)y_i(t) = \hat{f}_1(t),\\
&(b)\quad \ddot{y}_2(t) + \textstyle\sum_{i=3,4,7,8} A^1_{2,i}(t)\dot{y}_i(t) + \sum_{i=2,4,6,8} A^0_{2,i}(t)y_i(t) = \hat{f}_2(t),\\
&(c)\quad \ddot{y}_3(t) + \textstyle\sum_{i=3,4,7,8} A^1_{3,i}(t)\dot{y}_i(t) + \sum_{i=2,4,6,8} A^0_{3,i}(t)y_i(t) = \hat{f}_3(t),\\
&(d)\quad \ddot{y}_4(t) + \textstyle\sum_{i=3,4,7,8} A^1_{4,i}(t)\dot{y}_i(t) + \sum_{i=2,4,6,8} A^0_{4,i}(t)y_i(t) = \hat{f}_4(t),\\
&(e)\quad \dot{y}_5(t) + \textstyle\sum_{i=2,4,6,8} A^0_{5,i}(t)y_i(t) = \hat{f}_5(t),\\
&(f)\quad \dot{y}_6(t) + \textstyle\sum_{i=2,4,6,8} A^0_{6,i}(t)y_i(t) = \hat{f}_6(t),\\
&(g)\quad \dot{y}_1(t) + \textstyle\sum_{i=2,4,6,8} A^0_{7,i}(t)y_i(t) = \hat{f}_7(t),\\
&(h)\quad \dot{y}_2(t) + \textstyle\sum_{i=2,4,6,8} A^0_{8,i}(t)y_i(t) = \hat{f}_8(t),\\
&(i)\quad y_7(t) = \hat{f}_9(t),\\
&(j)\quad y_5(t) = \hat{f}_{10}(t),\\
&(k)\quad y_3(t) = \hat{f}_{11}(t),\\
&(l)\quad y_1(t) = \hat{f}_{12}(t),\\
&(m)\ 0 = \hat{f}_{13}(t).
\end{aligned} \tag{33}$$

Similar to the case of matrix pairs, we

1. insert the derivative of Equation (33-l) in (33-g) to eliminate $\dot{y}_1(t)$;
2. insert the second derivative of Equation (33-l) in (33-a) to eliminate $\ddot{y}_1(t)$;
3. insert the second derivative of Equation (33-k) in (33-c) to eliminate $\ddot{y}_3(t)$;
4. insert the derivative of Equation (33-j) in (33-e) to eliminate $\dot{y}_5(t)$;
5. insert the derivative of Equation (33-h) in (33-b) to eliminate $\ddot{y}_2(t)$; and
6. insert the derivative or if possible only parts of Equation (33-i) in (33-a)–(33-d) to eliminate possibly existent $\dot{y}_7(t)$.

These steps correspond to transforming the systems (32) into an equivalent second order system of DAEs

$$A_2^{\langle 1\rangle}(t)\ddot{y}(t) + A_1^{\langle 1\rangle}(t)\dot{y}(t) + A_0^{\langle 1\rangle}(t)y(t) = f^{\langle 1\rangle}(t), \tag{34}$$

with $(A_2^{\langle 1\rangle}(t), A_1^{\langle 1\rangle}(t), A_0^{\langle 1\rangle}(t); f^{\langle 1\rangle})$ being of the following form

$$\left(\begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{d^{(2)}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right.$$

$$\begin{bmatrix} 0 & 0 & * & * & 0 & 0 & 0 & * \\ 0 & * & * & * & 0 & * & 0 & * \\ 0 & 0 & * & * & 0 & 0 & 0 & * \\ 0 & 0 & * & * & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{d^{(1)}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{35}$$

$$\left.\begin{bmatrix} 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & I_a & 0 \\ 0 & 0 & 0 & 0 & I_{s^{(0,1)}} & 0 & 0 & 0 \\ 0 & 0 & I_{s^{(1,2)}} & 0 & 0 & 0 & 0 & 0 \\ I_{s^{(0,1,2)}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \begin{bmatrix} \hat{f}_1(t) - \ddot{\hat{f}}_{12}(t) \\ \hat{f}_2(t) - \dot{\hat{f}}_8(t) \\ \hat{f}_3(t) - \ddot{\hat{f}}_{11}(t) \\ \hat{f}_4(t) \\ \hat{f}_5(t) - \dot{\hat{f}}_{10}(t) \\ \hat{f}_6(t) \\ \hat{f}_7(t) - \dot{\hat{f}}_{12}(t) \\ \hat{f}_8(t) \\ \hat{f}_9(t) \\ \hat{f}_{10}(t) \\ \hat{f}_{11}(t) \\ \hat{f}_{12}(t) \\ \hat{f}_{13}(t) \end{bmatrix}\right)$$

It should be noted that the procedure of using derivatives of some equations to eliminate coefficients in other equations only involves the absolutely necessary derivatives of the right-hand side $\hat{f}(t)$. Moreover, after the transformation from the system (32) to the system (34), the solution sets of the two systems are still the same.

We then continue the *index reduction procedure* as follows. For the triple $(A_2^{\langle 1\rangle}(t), A_1^{\langle 1\rangle}(t), A_0^{\langle 1\rangle}(t))$ in (35), we can again transform to the condensed form (31), and then apply Steps (1)–(6) to pass it to the form (35). Proceeding inductively, we obtain a sequence of triples of matrix functions (matrices) $(A_2^{\langle i\rangle}(t), A_1^{\langle i\rangle}(t), A_0^{\langle i\rangle}(t))$, $i \in \mathbb{N}_0$, where $(A_2^{\langle 0\rangle}(t), A_1^{\langle 0\rangle}(t), A_0^{\langle 0\rangle}(t)) = (A_2(t), A_1(t), A_0(t))$ and $(A_2^{\langle i+1\rangle}(t), A_1^{\langle i+1\rangle}(t), A_0^{\langle i+1\rangle}(t))$ is derived from $(A_2^{\langle i\rangle}(t), A_1^{\langle i\rangle}(t), A_0^{\langle i\rangle}(t))$ by bringing it into the form (31) and then applying Steps (1)–(6) again. In the $j$th step we assume that

$$\begin{aligned} &s_j^{(0,1,2)}(t), s_j^{(1,2)}(t), s_j^{(0,2)}(t), d_j^{(2)}(t), s_j^{(0,1)}(t), d_j^{(1)}(t), \\ &a_j(t), v_j(t), u_j(t) \text{ are constant in } \mathbb{I}. \end{aligned} \tag{36}$$

Comparing $\hat{A}_2(t)$ in (31) with $A_2^{\langle 1\rangle}(t)$ in (35), we have

$$\begin{aligned} \operatorname{rank}(A_2^{\langle 1\rangle}(t)) &= \operatorname{rank}(\hat{A}_2(t)) - s_{\langle 0\rangle}^{(0,1,2)} - s_{\langle 0\rangle}^{(0,2)} - s_{\langle 0\rangle}^{(1,2)} \\ &= \operatorname{rank}(A_2^{\langle 0\rangle}(t)) - s_{\langle 0\rangle}^{(0,1,2)} - s_{\langle 0\rangle}^{(0,2)} - s_{\langle 0\rangle}^{(1,2)}, \end{aligned} \tag{37}$$

where $s_{\langle 0\rangle}^{(0,1,2)}$, $s_{\langle 0\rangle}^{(0,2)}$, and $s_{\langle 0\rangle}^{(1,2)}$ denote the strangeness due to $A_2^{\langle 0\rangle}(t), A_1^{\langle 0\rangle}(t), A_0^{\langle 0\rangle}(t)$, the strangeness due to $A_2^{\langle 0\rangle}(t), A_0^{\langle 0\rangle}(t)$, and the strangeness due to $A_2^{\langle 0\rangle}(t), A_1^{\langle 0\rangle}(t)$, respectively. Since after the differentiation-and-elimination Step 4, Equation (33-j) becomes an uncoupled purely algebraic equation, it follows that

$$\operatorname{rank}(A_0) \geq a_{\langle 1\rangle} \geq \left(a_{\langle 0\rangle} + s_{\langle 0\rangle}^{(0,1)}\right), \tag{38}$$

where $a_{\langle 1\rangle}$, $a_{\langle 0\rangle}$, and $s_{\langle 0\rangle}^{(0,1)}$ denotes the size of the algebraic part of $(A_2^{\langle 1\rangle}, A_1^{\langle 1\rangle}, A_0^{\langle 1\rangle})$, the size of the algebraic part of $(A_2^{\langle 0\rangle}, A_1^{\langle 0\rangle}, A_0^{\langle 0\rangle})$, and the strangeness due to $A_1^{\langle 0\rangle}, A_0^{\langle 0\rangle}$, respectively. Hence, the relations in (37) and (38) guarantee that after a finite number (say $\hat{\mu}(t) \equiv \hat{\mu}$) of steps, the strangeness $s_{\langle\hat{\mu}\rangle}^{(0,1,2)}$ due to $A_2(t), A_1(t), A_0(t)$, the strangeness $s_{\langle\hat{\mu}\rangle}^{(0,2)}$ due to $A_2(t), A_0(t)$, the strangeness $s_{\langle\hat{\mu}\rangle}^{(0,1)}$ due to $A_1(t), A_0(t)$, and the strangeness $s_{\langle\hat{\mu}\rangle}^{(1,2)}$ due to $A_2(t), A_1(t)$ must vanish.

It can be shown in a similar way as for the case of first order systems in [18] that under the given assumptions the quantities $\hat{\mu}$, $d_{\hat{\mu}}^{(2)}(t) \equiv d_{\hat{\mu}}^{(2)}$, $d_{\hat{\mu}}^{(1)}(t) \equiv d_{\hat{\mu}}^{(1)}$, and $a_{\hat{\mu}}(t) \equiv a_{\hat{\mu}}$ are invariant, i.e. they are independent of the order of the transformations that are performed.

We call $\hat{\mu}$ the *strangeness index* of the second order system of DAEs and we call the final equivalent second order system of DAEs *strangeness-free*.

**Theorem 16.** Consider the system (29), suppose that (36) holds in every step of the reduction procedure and let $\hat{\mu}$ be the resulting strangeness index of (29). If

$f(t) \in \mathcal{C}^{\hat{\mu}}(\mathbb{I}, \mathbb{C}^m)$, then system (29) is equivalent (in the sense that there is a one-to-one correspondence between the solution sets) to a system of second order differential-algebraic equations of the form

$$
\begin{aligned}
&(a)\ \ \ddot{\tilde{x}}_1(t) + \tilde{A}^1_{1,1}(t)\dot{\tilde{x}}_1(t) + \tilde{A}^1_{1,4}(t)\dot{\tilde{x}}_4(t) \\
&\qquad\qquad + \tilde{A}^0_{1,1}(t)\tilde{x}_1(t) + \tilde{A}^0_{1,2}(t)\tilde{x}_2(t) + \tilde{A}^0_{1,4}(t)\tilde{x}_4(t) = \tilde{f}_1(t), \\
&(b)\ \ \dot{\tilde{x}}_2(t) + \tilde{A}^0_{2,1}(t)\tilde{x}_1(t) + \tilde{A}^0_{2,2}(t)\tilde{x}_2(t) + \tilde{A}^0_{2,4}(t)\tilde{x}_4(t) = \tilde{f}_2(t), \\
&(c)\ \ \tilde{x}_3(t) = \tilde{f}_3(t), \\
&(d)\ \ 0 = \tilde{f}_4(t),
\end{aligned}
\tag{39}
$$

where the inhomogeneity $\tilde{f}(t) := \left[\tilde{f}_1^T(t), \ldots, \tilde{f}_4^T(t)\right]^T$ is determined by $f^{(0)}(t), \ldots, f^{(\hat{\mu})}(t)$. In particular, $d^{(2)}_{\hat{\mu}}(t) \equiv d^{(2)}_{\hat{\mu}}$, $d^{(1)}_{\hat{\mu}}(t) \equiv d^{(1)}_{\hat{\mu}}$, and $a_{\hat{\mu}}(t) \equiv a_{\hat{\mu}}$ are the number of second order differential, first order differential, and algebraic components of the unknown $\tilde{x}(t) := \left[\tilde{x}_1^T(t), \ldots, \tilde{x}_4^T(t)\right]^T$ in (39-a), (39-b), and (39-c), respectively, while $u_{\hat{\mu}}(t) \equiv u_{\hat{\mu}}$ is the dimension of the undetermined vector $\tilde{x}_4(t)$ in (39-a) and (39-b), and $v_{\hat{\mu}}(t) \equiv v_{\hat{\mu}}$ is the number of conditions in (39-d).

*Proof.* Inductively transforming $(A_2(t), A_1(t), A_0(t))$ to the condensed form (31), and then applying Steps (1)–(6) until $s^{(0,1,2)}_{\hat{\mu}} = s^{(1,2)}_{\hat{\mu}} = s^{(0,2)}_{\hat{\mu}} = s^{(0,1)}_{\hat{\mu}} = 0$ yields a triple of matrix functions $(\tilde{A}_2(t), \tilde{A}_1(t), \tilde{A}_0(t))$ of the form

$$
\left(
\begin{bmatrix} I_{d^2_{\hat{\mu}}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},
\begin{bmatrix} \tilde{A}^1_{1,1}(t) & 0 & 0 & \tilde{A}^1_{1,4}(t) \\ 0 & I_{d^1_{\hat{\mu}}} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},
\right.
$$
$$
\left.
\begin{bmatrix} \tilde{A}^0_{1,1}(t) & \tilde{A}^0_{1,2}(t) & 0 & \tilde{A}^0_{1,4}(t) \\ \tilde{A}^0_{2,1}(t) & \tilde{A}^0_{2,2}(t) & 0 & \tilde{A}^0_{2,4}(t) \\ 0 & 0 & I_{a_{\hat{\mu}}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\right),
$$

with block size $v_{\hat{\mu}}$ for the last block row and $u_{\hat{\mu}}$ for the last block column. We know that the transformation from $(A_2(t), A_1(t), A_0(t))$ to $(\hat{A}_2(t), \hat{A}_1(t), \hat{A}_0(t))$ in the condensed form (31) establishes a one-to-one correspondence between the solution sets of the two corresponding systems of DAEs. Hence, for any solution $x(t)$ of the system (29) (if existent), there exists a solution of the system (39) such that

$$x(t) = \tilde{Q}(t)\tilde{x}(t),$$

where $\tilde{Q}(t)$ is a nonsingular matrix function, and *vice versa.* □

*Remark 17.* It should be noted that the maximal number of derivatives of $f$ that is needed to obtain any of the components $\tilde{f}_i$ in (39) is $2\hat{\mu}$. Therefore, generically the minimal smoothness requirement for $f$ that we need in order to obtain a continuous solution is that the inhomogeneity is $2\hat{\mu}$ times differentiable.

*Remark 18.* In order to derive the condensed form (31), we have to assume the regularity condition (36) in each step of the index reduction procedure. This seems a rather strong assumption. But it follows from a result for pairs of matrix functions in ([6], Ch. 10) that for a closed interval $\mathbb{I}$ and sufficiently smooth coefficients functions $A_i(t)$, there exist open intervals $\mathbb{I}_j$, $j \in \mathbb{N}$, such that

$$\overline{\bigcup_{j\in\mathbb{N}} \mathbb{I}_j} = \mathbb{I}, \quad \mathbb{I}_i \cap \mathbb{I}_j = \emptyset \text{ for } i \neq j, \tag{40}$$

where (36) holds, i.e., the strangeness index $\hat{\mu}(t)$ is defined on a dense subset of $\mathbb{I}$. Furthermore, as for linear first order systems, see [20], these constant rank requirements can be significantly relaxed also for higher order systems, when one is using derivative arrays. This topic is currently under investigation.

**Example 19.** Consider again the system (4) of second order DAEs. By the described index reduction procedure, system (4) can be equivalently transformed to the following strangeness-free system

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \ddot{\tilde{x}}(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \dot{\tilde{x}}(t) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tilde{x}(t) = \tilde{f}(t),$$

where $\tilde{x}(t) = [x_2(t), x_1(t)]^T$, and $\tilde{f}(t) = [f_1(t) - \dot{f}_2(t) - \ddot{f}_2(t), f_2(t)]^T$. Hence, by Theorem 16, the strangeness index of system (4) is $\hat{\mu} = 2$.

In contrast to this, the classical first order formulation can be equivalently transformed to the strangeness-free system

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \dot{\tilde{y}}(t) + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tilde{y}(t) = \begin{bmatrix} f_2(t) \\ f_1(t) - \dot{f}_2(t) - \ddot{f}_2(t) \\ \dot{f}_2(t) \\ \dot{f}_1(t) - \ddot{f}_2(t) - f_2^{(3)}(t) \end{bmatrix},$$

where $\tilde{y}(t) = [x_1(t), x_2(t), \dot{x}_1(t), \dot{x}_2(t)]^T$. By Theorem 16, the strangeness index of the first order version is $\mu = 3$, which is larger by 1 than the strangeness index of the original second order system (4).

The most important consequence of Theorem 16 is the identification of those second derivatives of variables that can be replaced to obtain a first order system without changing the smoothness requirements and increasing the index.

**Corollary 20.** Under the assumptions of Theorem 16, let $\hat{\mu}$ be the strangeness index of the matrix triple associated with the system (29) and let $f(t) \in \mathcal{C}^{\hat{\mu}}(\mathbb{I}, \mathbb{C}^m)$. Then, the solution set of system (29) is in one-to-one correspondence (without further smoothness requirements) to the partial solution set given by the components $\tilde{x}_1(t), \ldots, \tilde{x}_4(t)$ of the system of first order differential-algebraic equations

$$\begin{aligned}
&(a)\ \dot{\tilde{x}}_5(t) + \tilde{A}^1_{1,1}\dot{\tilde{x}}_1(t) + \tilde{A}^1_{1,4}\dot{\tilde{x}}_4(t) \\
&\qquad\qquad + \tilde{A}^0_{1,1}\tilde{x}_1(t) + \tilde{A}^0_{1,2}\tilde{x}_2(t) + \tilde{A}^0_{1,4}\tilde{x}_4(t) = \tilde{f}_1(t), \\
&(b)\ \dot{\tilde{x}}_2(t) + \tilde{A}^0_{2,1}\tilde{x}_1(t) + \tilde{A}^0_{2,2}\tilde{x}_2(t) + \tilde{A}^0_{2,4}\tilde{x}_4(t) = \tilde{f}_2(t), \\
&(c)\ \tilde{x}_3(t) = \tilde{f}_3(t), \\
&(d)\ 0 = \tilde{f}_4(t), \\
&(e)\ \dot{\tilde{x}}_1 = \tilde{x}_5.
\end{aligned} \tag{41}$$

*Proof.* The proof follows immediately from (39) and it is clear that no extra smoothness requirements are needed. □

To demonstrate the effect of obtaining the first order formulation via the described procedure we present the following numerical example.

**Example 21.** Consider the linear second order constant coefficient DAE

$$\begin{bmatrix} 1 & 1 \\ t & t \end{bmatrix} \ddot{x}(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \dot{x}(t) + \begin{bmatrix} 1 & 0 \\ 1+t & 1 \end{bmatrix} x(t) = f(t), \tag{42}$$

with right-hand side

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t^2\left(\ln(t) - \frac{1}{2}\right) - \frac{1}{2}t^2 \\ e^t \end{bmatrix}$$

on the interval $[-1, 1]$.

System (42) has the solution

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} f_1(t) + 2\dot{f}_1 + t\ddot{f}_1 - \ddot{f}_2(t) \\ f_2(t) - (t+1)f_1(t) - 2\dot{f}_1 - t\ddot{f}_1 + \ddot{f}_2(t) \end{bmatrix}.$$

The classical first order formulation is

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & t & t \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1+t & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ 0 \\ 0 \end{bmatrix} \tag{43}$$

while the first order system obtained by only introducing $v_1 = \dot{x}_1 + \dot{x}_2$ is given by

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & t \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 1+t & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ 0 \end{bmatrix} \tag{44}$$

We have solved the strangeness-free first order systems (43) and (44) using the MATLAB implementation [22] of the code GELDA [21] (with BDF methods as integrator) on the interval $[-1, 1]$ with tolerances $RTOL = ATOL = 10^{-5}$. The results for (44) are presented in figure 1. For (43) GELDA properly detects a singularity at the point $t = 0$ and therefore the integrator is stopped. To compare the results we have then solved (43) with GELDA also on the interval $[0.001, 1]$ with the same tolerances, see figure 2. The best results are obtained, however, if we avoid the first order formulation completely as is demonstrated by using the BDF methods of [29] in the implementation of [32] for the second order formulation, see figure 3.

After having derived the results for second order systems, the extension to arbitrary systems of order $l$ is obvious. For completeness we state these results in the next section.

## 4. Linear *l*th order DAEs

As we have seen in Section 3, we can get a condensed form via strong equivalence transformations for matrix triples. Using induction, this form can be extended to

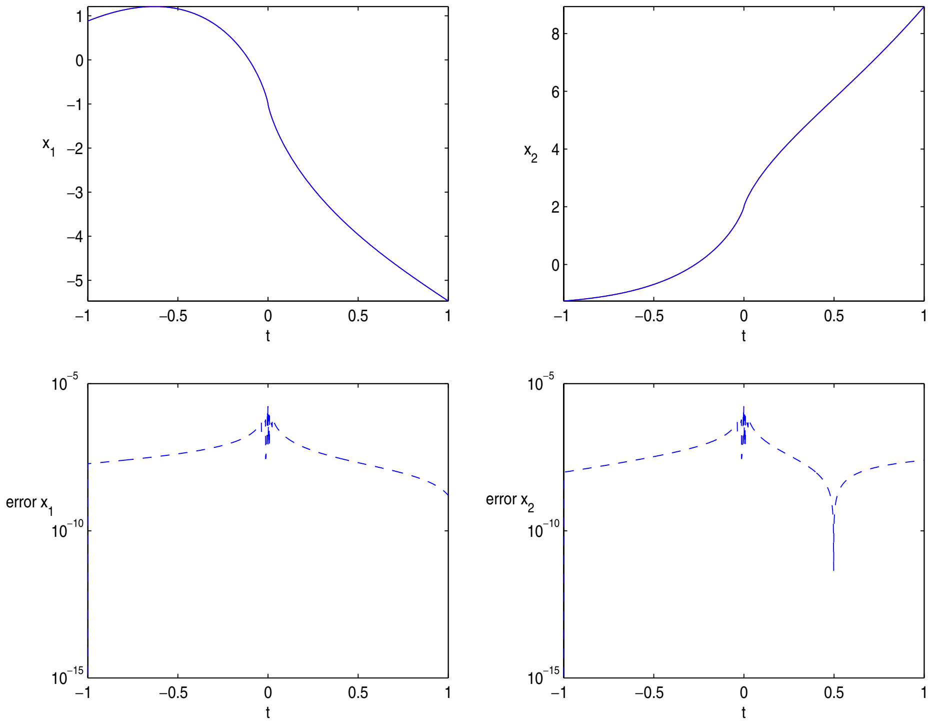


**Figure 1** Solution of strangeness-free system (44) in $[-1, 1]$ together with absolute errors.

$(l+1)$-tuples $(A_l(t), \ldots, A_1(t), A_0(t))$ of matrix functions and we obtain for (1) new sets of quantities, i.e. those that characterize the algebraic part, 1st order, 2nd order, ..., and $l$th order differential parts, and strangeness parts due to each two, each three, ..., each $l$, and $l+1$ matrices out of $A_l(t), \ldots, A_1(t)$, and $A_0(t)$, respectively.

Then, based on the condensed form for $(l+1)$-tuples of matrix functions, we can write down the system of differential-algebraic equations after strong equivalence transformations. Analogous to the treatment of systems of second order, we can design steps of inserting derivatives of some equations into others to decouple those equations that are coupled to each other to reduce it to a simpler but equivalent system, which we can again transform to the condensed form. Inductively, by this procedure we obtain a sequence of $(l+1)$-tuples of matrix functions, and after a finite number $\hat{\mu}$ of steps, which we again call the strangeness index, we obtain a strangeness-free system.

Here we only state the essential results and its main consequences without proof.

**Theorem 22.** Let $f(t) \in \mathcal{C}^{\hat{\mu}}(\mathbb{I}, \mathbb{C}^m)$. Then, under appropriate constant rank assumptions (analogous to those in Equation (36) for second order systems), system (1) is equivalent (in the sense that there is a one-to-one correspondence between the solution sets) to a system of $l$th order differential-algebraic equations

$$A^{\langle l\rangle}(t)\tilde{x}^{(l)}(t) + A^{\langle l-1\rangle}(t)\tilde{x}^{(l-1)}(t) + \cdots + A^{\langle 0\rangle}(t)\tilde{x}(t) = \tilde{f}(t)$$

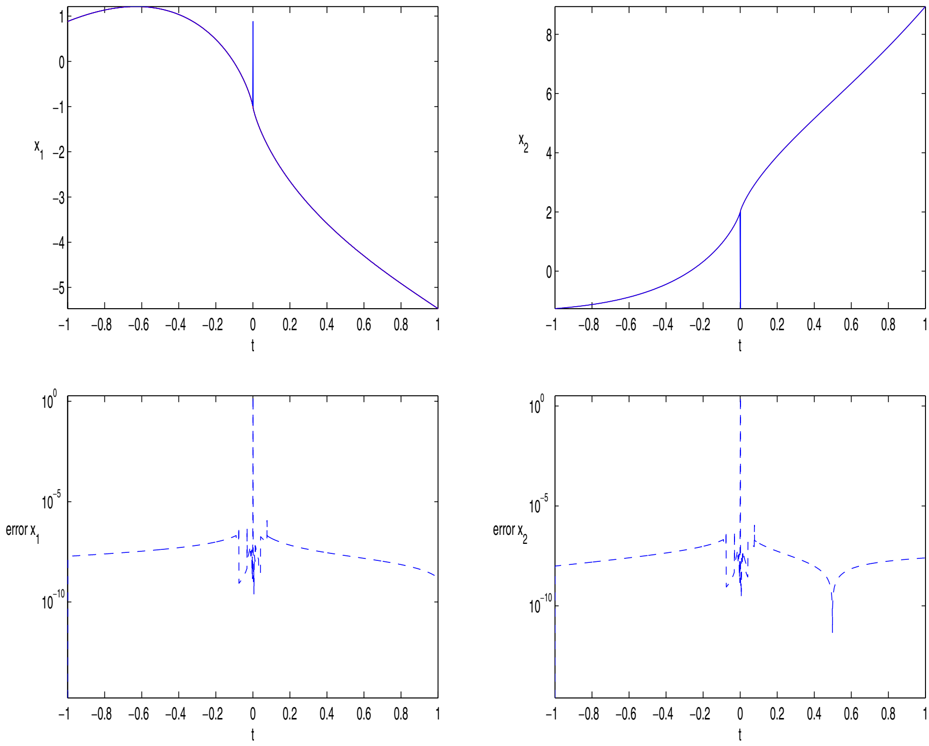


**Figure 2** Solution of system (43) in $[-1, -0.001] \cup [0.001, 1]$ together with absolute errors.

of the form

$$
\begin{array}{ll}
(1) & \tilde{x}_1^{(l)}(t) + \sum_{i=0}^{l-1}\sum_{j=i}^{l-1} A^{\langle i\rangle}_{1,l-j}(t)\tilde{x}^{(i)}_{l-j}(t) \\
& \quad + \sum_{i=0}^{l-1} A^{\langle i\rangle}_{1,l+2}(t)\tilde{x}^{(i)}_{l+2}(t) = \tilde{f}_1(t), \\
(2) & \tilde{x}_2^{(l-1)}(t) + \sum_{i=0}^{l-2}\sum_{j=i}^{l-2} A^{\langle i\rangle}_{2,l-1-j}(t)\tilde{x}^{(i)}_{l-1-j}(t) \\
& \quad + \sum_{i=0}^{l-2}\left(A^{\langle i\rangle}_{2,1}(t)\tilde{x}^{(i)}_1(t) + A^{\langle i\rangle}_{2,l+2}(t)\tilde{x}^{(i)}_{l+2}(t)\right) = \tilde{f}_2(t), \\
\vdots & \qquad \vdots \\
(k) & \tilde{x}_{l-k+1}^{(l-k+1)}(t) + \sum_{i=0}^{l-k}\sum_{j=i}^{l-k} A^{\langle i\rangle}_{k,l-k+1-j}(t)\tilde{x}^{(i)}_{l-k+1-j}(t) \\
& \quad + \sum_{i=0}^{l-k}\left(\sum_{j=1}^{k} A^{\langle i\rangle}_{k,j}(t)\tilde{x}^{(i)}_1(t) + A^{\langle i\rangle}_{k,l+2}(t)\tilde{x}^{(i)}_{l+2}(t)\right) = \tilde{f}_k(t), \\
& \qquad (1 \le k \le l) \\
\vdots & \qquad \vdots \\
(l+1) & \tilde{x}_{l+1}(t) = \tilde{f}_{l+1}(t), \\
(l+2) & 0 = \tilde{f}_{l+2}(t),
\end{array}
\tag{45}
$$

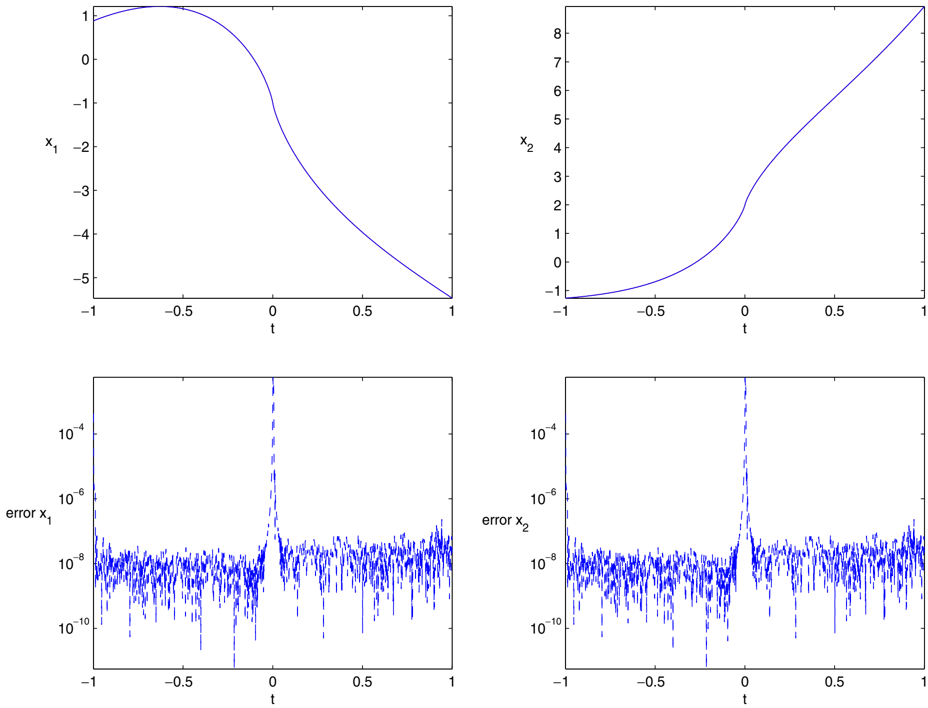


**Figure 3** Solution of system (42) with BDF method together with absolute errors.

where $A^{\langle i\rangle}_{p,q}(t)$, $1 \leq p \leq (l+2)$, $1 \leq q \leq (l+2)$, denotes a sub-matrix of $A^{\langle i\rangle}(t)$, and the inhomogeneity $\tilde{f}(t) := \left[\tilde{f}_1^T(t), \ldots, \tilde{f}_{l+2}^T(t)\right]^T$ is determined by $f^{(0)}(t), \ldots, f^{(\hat{\mu})}(t)$. In particular, $d^{(l)}_{\hat{\mu}}, \ldots, d^{(1)}_{\hat{\mu}}$, and $a_{\hat{\mu}}$ are the number of $l$th order differential, …, first order differential, algebraic components of the unknown $\tilde{x}(t) := \left[\tilde{x}_1^T(t), \ldots, \tilde{x}_{l+2}^T(t)\right]^T$ in (45-1), …, (45-(l+1)), respectively, while $u_{\hat{\mu}}$ is the dimension of the undetermined vector $\tilde{x}_{l+2}(t)$ in (45-1), …, (45-l), and $v_{\hat{\mu}}$ is the number of conditions in (45-(l+2)).

*Proof.* The proof is analogous to the proof of Theorem 16 and follows by induction. □

We immediately have the following results.

**Corollary 23.** Under the assumption of Theorem 22, the following statements hold.

1. The system (1) is solvable if and only if one of the following two cases
   (a) $v_{\hat{\mu}} = 0$.
   (b) If $v_{\hat{\mu}} > 0$, then the $u_{\hat{\mu}}$ functional consistency conditions

$$\tilde{f}_{l+2}(t) = 0$$

are satisfied.

2. If the system (1) is solvable, then it is uniquely solvable without providing any initial condition if and only if the conditions
$$d_{\hat{\mu}}^{(l)} = \cdots = d_{\hat{\mu}}^{(2)} = d_{\hat{\mu}}^{(1)} = u_{\hat{\mu}} = 0$$
hold.
3. If the system (1) is solvable, then initial conditions (2) are consistent if and only if one of the following two cases happens.
   (a) $a_{\hat{\mu}} = 0$.
   (b) If $a_{\hat{\mu}}(t) > 0$, then the $a_{\hat{\mu}}$ conditions
$$\tilde{x}_{l+1}(t_0) = \tilde{f}_{l+1}(t_0),$$
$$\dot{\tilde{x}}_{l+1}(t_0) = \left.\frac{d\tilde{f}_{l+1}(t)}{dt}\right|_{t_0+}, \ldots, \tilde{x}_{l+1}^{(l-1)}(t_0) = \left.\tilde{f}_{l+1}^{(l-1)}(t)\right|_{t_0+}$$
   are implied by (1).
4. If the initial value problem (1)–(2) is solvable, then it is uniquely solvable if and only if
$$u_{\hat{\mu}} = 0$$
holds.

**Corollary 24.** Under the assumptions of Theorem 22, let $\hat{\mu}$ be the strangeness index of the tuple of matrix functions associated with the system (1) and let $f(t) \in \mathcal{C}^{\hat{\mu}}(\mathbb{I}, \mathbb{C}^m)$. Then, the solution set of system (1) is in one-to-one correspondence (without further smoothness requirements) to the partial solution set given by the components $\tilde{x}_1(t), \ldots, \tilde{x}_{l+2}(t)$ of the system of first order differential-algebraic equations that is obtained by replacing in (45) the derivatives $\tilde{x}_{l-k+1}^{(l-k+1)}(t)$ by new variables $v_{l-k+1}$, $k = 1, \ldots, l$.

*Proof.* The proof follows as in the case of matrix triples. □

We also obtain the maximal possible increase in strangeness index.

**Corollary 25.** Under the assumptions of Theorem 22, let $\hat{\mu}$ be the strangeness index of the tuple of matrix functions associated with the system (1) and let $\mu$ be the strangeness index of the first order system obtained by the classical order reduction procedure.

Then, $\mu \leq \hat{\mu} + l - 1$, and equality is possible.

## 5. Conclusions

We have presented the analysis of linear systems of differential-algebraic equations of higher order. We have derived condensed forms for tuples of matrices and tuples of matrix-valued functions which are associated with the systems of constant and variable coefficients, respectively. Based on the condensed forms, we may convert such a system into an equivalent system, from which the behavior with respect to solvability, uniqueness of solutions and consistency of initial conditions can be directly read off.

We have demonstrated that if one turns a higher order problem in the traditional way into a first order system of DAEs, then, to get the solvability and uniqueness of solutions, more smoothness of the right-hand side $f(t)$ may be required. The condensed forms, however, allow to do the transformation to first order without extra smoothness requirements.

Several issues remain open. These include the perturbation theory for higher order systems of DAEs, (see [17] for recent results in the case of constant coefficient systems) in particular how the decision making in the condensed forms influences the transformation to first order (see [26] for first results) as well as the construction of appropriate numerical methods for the treatment of high order, high index differential-algebraic systems, see [29, 32] for first results. Further open questions that are currently under investigation are the construction of systems analogous to those in Theorems 16 and 22 from derivative arrays and the preservation of symmetry structures.

**Acknowledgments** We thank Peter Kunkel for helpful comments and L. Wunderlich for performing the numerical tests. Volker Mehrmann was supported by MATHEON, the DFG Research Center 'Mathematics for key technologies' in Berlin. Chunhao Shi was supported by Deutsche Forschungsgemeinschaft, through Research Grant ME 790/14-1.

## References

[1] Arévalo, C., Lötstedt, P.: Improving the accuracy of bdf methods for index 3 differential-algebraic equations. BIT **35**, 297–308 (1995)
[2] Balla, K., März, R.: A unified approach to linear differential algebraic algebraic equations and their adjoints. ZANAA **21**, 783–802 (2002)
[3] Brenan, K.E., Campbell, S.L., Petzold, L.R.: Numerical Solution of Initial-Value Problems in Differential Algebraic Equations, vol. 14 of Classics in Applied Mathematics, 2nd edn. SIAM, Philadelphia, Pennsylvania, 1996
[4] Campbell, S.L.: Singular Systems of Differential Equations I. Pitman, San Francisco (1980)
[5] Campbell, S.L.: Singular Systems of Differential Equations II. Pitman (1982)
[6] Campbell, S.L., Meyer, C.D.: Generalized Inverses of Linear Transformations. Pitman, San Francisco (1979)
[7] De Boor, C., Kreiss, H.O.: On the condition of the linear systems associated with discretized BVPs of ODEs. SIAM J. Numer. Anal. **23**, 936–939 (1986)
[8] Eich-Soellner, E., Führer, C.: Numerical Methods in Multibody Systems. B.G. Teubner, Stuttgart (1998)
[9] Gantmacher, F.R.: The Theory of Matrices, vol. II. Chelsea, New York, NY (1959)
[10] Gear, C.W., Petzold, L.R.: Differential/algebraic systems and matrix pencils. In: Kågström, B., Ruhe, A. (eds.) Matrix Pencils, vol. 973 of Lecture Notes in Mathematics, pp.75–89. Springer, Berlin Heidelberg New York (1983)
[11] Gohberg, I., Lancaster, P., Rodman, L.: Matrix Polynomials. Academic Press, New York (1982)
[12] Griepentrog, E., März, R.: Differential-Algebraic Equations and Their Numerical Treatment. Teubner Texte zur Mathematik. Teubner-Verlag, Leipzig (1986)
[13] Günther, M., Feldmann, U.: CAD-based electric-circuit modeling in industry. I. Mathematical structure and index of network equations. Surv. Math. Ind. **8**, 97–129 (1999)
[14] Günther, M., Feldmann, U.: CAD-based electric-circuit modeling in industry. II. Impact of circuit configurations and parameters. Surv. Math. Ind. **8**, 131–157 (1999)
[15] Hairer, E., Lubich, C., Roche, M.: The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods. Lecture Notes in Mathematics No. 1409. Springer, Berlin Heidelberg New York (1989)
[16] Hairer, E., Wanner, G.: Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems, 2nd edn. Springer, Berlin Heidelberg New York (1996)

[17] Higham, N.J., Mackey, D.S., Tisseur, F.: The conditioning of linearizations of matrix polynomials. Numerical Analysis Report No. 465, The University of Manchester, School of Mathematics, 2005, to appear in SIAM J. Matrix Analysis (2006)
[18] Kunkel, P., Mehrmann, V.: Canonical forms for linear differential-algebraic equations with variable coefficients. J. Comput. Appl. Math. **56**, 225–259 (1994)
[19] Kunkel, P., Mehrmann, V.: A new look at pencils of matrix valued functions. Linear Algebra Appl. **212/213**, 215–248 (1994)
[20] Kunkel, P., Mehrmann, V.: Differential-Algebraic Equations – Analysis and Numerical Solution. EMS, Zürich (2006)
[21] Kunkel, P., Mehrmann, V., Rath, W., Weickert, J.: A new software package for linear differential-algebraic equations. SIAM J. Sci. Comput. **18**, 115–138 (1997)
[22] Kunkel, P., Mehrmann, V., Seidel, S.: A MATLAB Toolbox for the Numerical Solution of Differential-Algebraic Equations. Preprint 16/2005, Institut für Mathematik, TU Berlin, 2002. http://www.math.tu-berlin.de/preprints/ (2006)
[23] Mackey, D.S., Mackey, N., Mehl, C., Mehrmann, V.: Vector spaces of linearizations for matrix polynomials. Preprint 238, DFG Research Center Matheon, Mathematics for key technologies in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2005. http://www.matheon.de/, to appear in SIAM J. Matrix Analysis (2006)
[24] März, R.: The index of linear differential algebraic equations with properly stated leading terms. Results Math. **42**, 308–338 (2002)
[25] März, R.: Solvability of linear differential algebraic equations with property stated leading terms. Results Math. **45**, 88–105 (2004)
[26] Mattheij, R.M.M., Wijckmans, P.M.E.J.: Sensitivity of solutions of linear DAE to perturbations of the system matrices. Numer. Algor. **19**, 159–171 (1998)
[27] Mehrmann, V., Shi, C.: Analysis of higher order linear differential-algebraic systems. Preprint 2004/17, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG, 2004. URL: http://www.math.tu-berlin.de/preprints/
[28] Rabier, P.J., Rheinboldt, W.C.: Nonholonomic Motion of Rigid Mechanical Systems from a DAE Viewpoint. SIAM, Philadelphia, Pennsylvania 19104-2688, USA (2000)
[29] Sand, J.: On implicit Euler for high-order high-index DAEs. Appl. Numer. Math. **42**, 411–424 (2002)
[30] Stoer, J., Bulirsch, R.: Introduction to Numerical Analysis. Springer, Berlin Heidelberg New York (1993)
[31] Tisseur, F., Meerbergen, K.: A survey of the quadratic eigenvalue problem. SIAM Rev. **43**, 234–286 (2001)
[32] Wunderlich, L.: Numerical Solution of Second Order Differential-Algebraic Equations. Diplomarbeit, TU Berlin. Institut für Mathematik, Berlin, FRG (2004)