Questions concerning differential-algebraic operators: Toward a reliable direct numerical treatment of differential-algebraic equations

Michael Hanke\textsuperscript{a}, Roswitha März\textsuperscript{b}

\textsuperscript{a}KTH Royal Institute of Technology, School of Engineering Sciences, 10044 Stockholm, Sweden
\textsuperscript{b}Humboldt University of Berlin, Institute of Mathematics, D-10099 Berlin, Germany

Abstract

The nature of so-called differential-algebraic operators and their approximations is constitutive for the direct treatment of higher-index differential-algebraic equations. We treat first-order differential-algebraic operators in detail and contribute to justify the overdetermined polynomial collocation applied to higher-index differential-algebraic equations. Besides, we discuss several practical aspects concerning higher-order differential-algebraic operators and the associated equations.

Keywords: differential-algebraic operator, essentially ill-posed problem, least-squares problem, polynomial collocation, nonclosed-range operator, higher index, first-order differential-algebraic equation, higher-order differential-algebraic equation

1. Introduction

To a large extend, in the framework of numerical analysis, differential-algebraic equations (DAEs), in particular higher-index ones, are recognized as special ordinary differential equations (ODEs), and, accordingly, they are treated by means of derivative-array systems and an involved or preceded expensive index-reduction.

In contrast, the experiments and theoretical contributions reportet in \cite{3, 10, 9} give rise to the conjecture that next to the existing derivative-array based methods there is further potential toward a reliable direct numerical treatment of DAEs. The main aim of this note is to fill the gap between the theoretical convergence results for least-squares collocation methods \cite{10, 9} and its practical realization. Moreover, we will explore the relevant scope concerning higher-order DAEs.

Recap well-known facts concerning first-order ordinary differential operators: Let $B(t) \in \mathbb{R}^{m,m}$ be continuous. The initial value problem (IVP)

$$x'(t) + B(t)x(t) = g(t), \quad t \in [a, b], \quad x(a) = d \in \mathbb{R}^m,$$

Email addresses: hanke@nada.kth.se (Michael Hanke), maerz.math.hu-berlin.de (Roswitha März)
has the unique solution

\[ x(t) = X(t, a)d + (Vg)(t), \quad t \in [a, b], \quad d \in \mathbb{R}^m, \quad (Vg)(t) = \int_a^t X(t, s)g(s)ds. \]

Introduce the operators \( T x := x' + Bx, \quad T_{IC}x := x(a), \quad \mathcal{T}x := (T x, T_{IC}x) \), such that the IVP is represented by the operator equation \( \mathcal{T}x = (g, d) \) and the inverse of \( \mathcal{T} \) is given by

\[ \mathcal{T}^{-1}(g, d)(t) = X(t, a)d + (Vg)(t), \quad t \in [a, b]. \] (1)

As it is well known, these operators have useful properties making ODE problems easily accessible for the numerical treatment: \( T : C^1([a, b], \mathbb{R}^m) \to C([a, b], \mathbb{R}^m) \) is bounded and surjective, \( \dim \ker T = m \), and hence, \( T \) is fredholm, and \( \mathcal{T} : C^1([a, b], \mathbb{R}^m) \to C([a, b], \mathbb{R}^m) \times \mathbb{R}^m \) is bounded and bijective, thus a homeomorphism. The same properties persist in the Hilbert space setting \( T : H^1((a, b), \mathbb{R}^m) \to L^2((a, b), \mathbb{R}^m), \quad \mathcal{T} : H^1((a, b), \mathbb{R}^m) \to L^2((a, b), \mathbb{R}^m) \times \mathbb{R}^m \). Such properties will survive for index one differential-algebraic equations in an appropriately modified version. The situation in the higher index case is much more involved.

The paper is organized as follows: Section 2 deals with a small first-order index-three example and foreshadows the potential of overdetermined polynomial collocation. Section 3 is devoted to features of regular first-order arbitrary-index differential-algebraic operators acting in reasonable Hilbert spaces. We turn back to the overdetermined polynomial collocation in Section 4 now considered by operators representing finite-dimensional approximations and provide with the main Theorem 4.3 sufficient conditions justifying overdetermined collocation. Section 5 surveys higher-order differential-algebraic operators in this context.

Below, though using different norms we mark the related norms by extra tags merely on those places where confusions are actually imminent.
List of some symbols and abbreviations

\( \mathbb{R}, \mathbb{N} \) \hspace{1cm} sets of real and natural numbers

\( \mathcal{L}(X,Y) \) \hspace{1cm} for linear spaces \( X, Y \) the space of linear operators

\( C(I,X) \) \hspace{1cm} space of \( k \)-times continuously differentiable functions mapping \( I \) into \( X \)

\( L^2 := \mathcal{L}^2((a,b), \mathbb{R}^m) \) \hspace{1cm} Lebesgue space of functions mapping \( (a,b) \) into \( \mathbb{R}^m \)

\( H^k := H^k((a,b), \mathbb{R}^m) \) \hspace{1cm} Sobolev space of functions mapping \( (a,b) \) into \( \mathbb{R}^m \)

\( K^* \) \hspace{1cm} adjoint of \( K \)

\( K^- \) \hspace{1cm} generalized inverse of \( K \), \( KK^-K = K \), \( K^-KK^- = K^- \)

\( K^+ \) \hspace{1cm} Moore-Penrose inverse of \( K \)

\( \ker K \) \hspace{1cm} nullspace (kernel) of \( K \)

\( \text{im } K \) \hspace{1cm} image (range) of \( K \)

\( \langle \cdot, \cdot \rangle \) \hspace{1cm} scalar product in \( \mathbb{R}^m \)

\( \langle \cdot, \cdot \rangle \) \hspace{1cm} scalar product in function spaces

\( | \cdot | \) \hspace{1cm} Euclidian vector norm and spectral norm of matrices

\( \| \cdot \| \) \hspace{1cm} norm in function spaces, operator norm

\( \oplus \) \hspace{1cm} topological direct sum

DAE \hspace{1cm} differential-algebraic equation

ODE \hspace{1cm} ordinary differential equation

IVP, BVP \hspace{1cm} initial value problem, boundary value problem

2. A symptomatic example

The DAE

\[
x'_2(t) + x_1(t) = g_1(t),
\]

\[
\eta x'_2(t) + x'_3(t) + (\eta + 1)x_2(t) = g_2(t),
\]

\[
\eta x_2(t) + x_3(t) = g_3(t), \quad t \in [0, 1],
\]

has index 3 uniformly for every \( \eta \in \mathbb{R} \). To each sufficiently smooth \( y \) there exists exactly one solution \( x \). The DAE is somewhat snaky, so that step by step integration methods generate waste unless an a priori or a posteriori incorporated regularization via derivative array systems is incorporated, e.g., [12, Example 8.5].

Here we set \( \eta = -2 \) and determine \( g_1, g_2, g_3 \) such that the solution becomes

\[
x_1(t) = e^{-t} \sin t, \quad x_2(t) = e^{-2t} \sin t, \quad x_3(t) = e^{-t} \cos t.
\]

We set \( N \geq 1 \) and approximate the solution components \( x_2 \) and \( x_3 \) by continuously connected piecewise polynomials of degree \( N \) and the component \( x_1 \) by possibly discontinuous piecewise polynomials of degree \( N - 1 \) on uniform partitions of the interval \([0, 1]\) with stepsize \( h = 1/n \). On each subinterval we choose \( M \) collocation points.
Table 1: Componentwise maximal error, polynomial degree $N = 3$

| $n$ | Standard collocation | Least-squares collocation |
|-----|-----------------------|---------------------------|
|     | $i = 1$ | $i = 2$ | $i = 3$ | $i = 1$ | $i = 2$ | $i = 3$ |
| 20  | 5.56e+006 | 3.03e+004 | 5.99e+004 | 2.09e-04 | 1.10e-06 | 2.18e-06 |
| 40  | 1.55e+017 | 4.23e+014 | 8.41e+014 | 5.03e-05 | 1.31e-07 | 2.65e-07 |
| 80  | 5.70e+038 | 7.76e+035 | 1.55e+036 | 1.23e-05 | 1.60e-08 | 3.20e-08 |
| 160 | 3.36e+082 | 2.29e+079 | 4.57e+079 | 3.06e-06 | 1.98e-09 | 4.00e-09 |
| 320 | 4.93e+170 | 1.68e+167 | 3.35e+167 | 7.68e-07 | 2.50e-10 | 5.00e-10 |

Table 2: Error ($\|e_1\|_{L^2} + \|e_2\|_{H^1} + \|e_3\|_{H^1}$) of the collocation solution for $N = 3$

| $n$ | $M = 2N + 1$ uniform points | $M = N + 1$ Gaussian points |
|-----|-------------------------------|-------------------------------|
|     | error | order | error | order |
| 10  | 6.31e-4 | 6.46e-4 |
| 20  | 1.44e-4 | 2.1   |
| 40  | 3.47e-5 | 3.47e-5 |
| 80  | 8.53e-6 | 8.53e-6 |
| 160 | 2.12e-6 | 2.12e-6 |
| 320 | 5.27e-7 | 5.27e-7 |

The classical or standard collocation procedures use $M = N$ collocation points per subinterval (cf. [13]). This results in $3nN$ equations to determine the $3nN + 2$ parameters of the unknown collocation solution. In order to obtain a unique collocation solution, we pose additionally two consistent initial condition.

In contrast, choosing $M > N$ leads to an overdetermined collocation system which can be treated by a least-squares solver. Corresponding first experiments are reported in [10] with $M = 2N + 1$. Table 1 shows the componentwise maximal error in both versions: The standard collocation generates waste as expected, however, the overdetermined least-squares collocation provides surprisingly nice results.

Further experiments (cf. [10, 9]) give rise to the conjecture that a much smaller number $M \geq N + 1$ will do in general and that the special position of the collocation points does not matter. Tables 2 and 3 are quoted from [9] to this effect. At this place it should be noted that the overdetermined collocation method is treated in [10, 9] against the background of a least-squares problem in Hilbert spaces. For this reason, the errors are now measured in $L^2$ and $H^1$ norms. The theoretically confirmed convergence order is $N - \mu + 1 = 1$, but we observe order 2. To date there is no theoretical recognition of this impressive, nice behavior!

Table 3 indicates the further interesting observation that even though no convergence is proved for $N = 1$ the numerical approximations remain bounded and seem to converge with order 0.4.

The bounded operator $T : \{ x \in C([0, 1], \mathbb{R}^3) : x_2, x_3 \in C^1([0, 1], \mathbb{R}) \} \to C([0, 1], \mathbb{R}^3)$ associated
Table 3: Error \((\|e_1\|_{I^p}^p + \|e_2\|_{I^p}^p + \|e_3\|_{I^p}^p)^{1/p}\) of the collocation solution for \(N = 1\)

| \(n\) | \(M = 3\) uniform points | \(M = 2\) Gaussian points |
|---|---|---|
| 10 | 5.65e-1 | 5.65e-1 |
| 20 | 3.93e-1 | 3.93e-1 | 0.5 |
| 40 | 2.49e-1 | 2.49e-1 | 0.7 |
| 80 | 1.85e-1 | 1.85e-1 | 0.4 |
| 160 | 1.42e-1 | 1.42e-1 | 0.4 |
| 320 | 1.12e-1 | 1.12e-1 | 0.3 |

with our test DAE reads in detail

\[
T x = \begin{bmatrix}
x_2' + x_1 \\
t \eta x_2' + x_2' + (\eta + 1)x_2 \\
t \eta x_2' + x_3
\end{bmatrix}, \quad x_1 \in C([0, 1], \mathbb{R}), \quad x_2, x_3 \in C^1([0, 1], \mathbb{R}),
\]

and one immediately checks that

\[
\ker T = \{0\}, \quad \text{im} T = \{g \in C([0, 1], \mathbb{R}^3) : g_3 \in C^1([0, 1], \mathbb{R}), \ g_2 - g_3' \in C^1([0, 1], \mathbb{R})\},
\]

further \(T = T\), since \(T\) itself is injective, and hence, no initial or boundary conditions are allowed. The inverse operator

\[
T^{-1} g = \begin{bmatrix}
g_1 - (g_2 - g_3')' \\
g_2 - g_3' \\
g_3 - \eta(g_2 - g_3')
\end{bmatrix}, \quad g \in \text{im} T.
\]

is unbounded in this setting since \(\text{im} T\) is a nonclosed subset in \(C([0, 1], \mathbb{R}^3)\). This is a fundamental contrast to the case of regular ODEs. Now we do not have closed range and fredholm properties, and the main ingredient of the inverse of \(T\) is not a nice Volterra operator but a higher-order differential operator. We have

\[
T^{-1} g = \begin{bmatrix}
g_1 - g_2' + g_3'' \\
g_2 - g_3' \\
g_3' - \eta(g_2 - g_3')
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} g'' + \begin{bmatrix}
0 & -1 & 0 \\
0 & 0 & -1 \\
0 & 0 & \eta
\end{bmatrix} g' + \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & -\eta & 1
\end{bmatrix} g, 
\]

\(g \in \{v \in C([0, 1], \mathbb{R}^3) : v_2 \in C^1([0, 1], \mathbb{R}), v_3 \in C^2([0, 1], \mathbb{R})\} \subset \text{im} T\).

Analogous situations arise also, e.g., in the settings

\[
T : C^1([0, 1], \mathbb{R}^3) \to C([0, 1], \mathbb{R}^3),
\]

\[
T : \{x \in L^2((0, 1), \mathbb{R}^3) : x_2, x_3 \in H^1([0, 1], \mathbb{R})\} \to L^2((0, 1), \mathbb{R}^3),
\]

\[
T : H^1((0, 1), \mathbb{R}^3) \to L^2((0, 1), \mathbb{R}^3),
\]

with their natural norms, \[14\].
3. Regular first-order DA operators in a Hilbert space setting

3.1. Setting

We begin this part by considering the operator, henceforth called a DA operator,

\[ \hat{T} \in \mathcal{L}(X, Y), \quad \hat{T} x := Ex' + Fx, \quad x \in \text{dom} \hat{T} := C^1([a, b], \mathbb{R}^m) \subseteq X, \]

so that the operator equation \( \hat{T} x = g \) represents the so-called standard form DAE \( Ex' + Fx = g \).

The coefficient functions \( E, F : [a, b] \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m) \) are at least continuous. The function spaces \( X \) and \( Y \) will be specified later.

We suppose a nontrivial leading coefficient \( E \) the nullspace of which is a \( C^1 \)-subspace varying in \( \mathbb{R}^m \). We are looking for a Hilbert-space setting with a bounded DA operator.

First we try \( X = Y = L^2((a, b), \mathbb{R}^m) \), with the usual norms. In this setting, \( \hat{T} \) is unbounded, but densely defined and closable, see [14]. The closure of \( \hat{T} \), \( T : \text{dom} T \subseteq X \rightarrow Y \), is densely defined and closed, but also unbounded. We apply the usual graph-norm approach: The space \( X_T := \text{dom} T \) equipped with graph-norm,

\[ \| x \|_T = \| x \| + \| T x \|, \quad x \in X_T := \text{dom} T, \]

is complete and \( T : X_T \rightarrow Y \) is a bounded operator. How can we specify \( \text{dom} T \) and \( T \)? To answer this question we need so-called proper factorizations of \( E \).

**Definition 3.1.** The factorization \( E = AD \) is called proper, if \( 0 < k \leq m \), \( A : [a, b] \rightarrow \mathcal{L}(\mathbb{R}^k, \mathbb{R}^m) \) is continuous, \( D : [a, b] \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k) \) is continuously differentiable, and

\[ \ker A(t) \oplus \text{im} D(t) = \mathbb{R}^k, \quad t \in [a, b]. \]

There are many possible proper factorizations. We fix an arbitrary one \( E = AD \), put \( B = F - AD' \), and observe that

\[ \hat{T} x = Ex' + Fx = ADx' + Fx = A(Dx)' + (F - AD')x = A(Dx)' + Bx, \quad \text{for all} \quad x \in \text{dom} \hat{T}. \]

Observe that \( A(Dx)' + Bx = g \) is a DAE with so-called properly stated leading term. This indicates how the closure \( T \) as well as \( \text{dom} T \) look like.

**Theorem 3.2.** (i) If \( D : [a, b] \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k) \) is continuously differentiable and has constant rank, then the function space

\[ H^1_D((a, b), \mathbb{R}^m) = \{ x \in L^2((a, b), \mathbb{R}^m) : Dx \in H^1((a, b), \mathbb{R}^k) \} \]

equipped with the inner product

\[ (x, \bar{x})_{H^1_D} = (x, \bar{x})_{L^2} + ((Dx)', (D\bar{x})')_{L^2} \]

is a Hilbert space.
(ii) For each proper factorization \( E = AD \), it results that
\[
\text{dom } T = H^1_D((a, b), \mathbb{R}^m), \quad T x = A(Dx)' + Bx, \quad x \in \text{dom } T.
\]

(iii) The norm \( \| \cdot \|_{H^1_D} \) is equivalent to the graph-norm \( \| \cdot \|_T \).

**Proof.** (i) is an immediate consequence of \([14, \text{Lemma 6.9}]\).

(ii) The closure \( T = T^{**} \) is provided by means of the biadjoint of \( T \) in \([14, \text{Theorem 3.1 (3)}] \).

(iii) Let \( x \in \text{im } T \) be given. From \( A(Dx)' = T x - Bx \) we obtain \( A^{-1}A(Dx)' = A^{-1}T x - A^{-1}Bx \). Owing to the proper factorization we may deduce the representation \( (Dx)' = A^{-1}T x + ((A^{-1}A)'D - A^{-1}B)x \). This yields
\[
\| (Dx)' \|_{L^2} \leq c_1(\| x \|_{L^2} + \| T x \|_{L^2}),
\]
\[
\| x \|_{H^1_D}^2 = \| x \|_{L^2}^2 + \| (Dx)' \|_{L^2}^2 \leq \| x \|_{L^2}^2 + c_1^2(\| x \|_{L^2} + \| T x \|_{L^2})^2 \leq (1 + c_1^2)(\| x \|_{L^2} + \| T x \|_{L^2})^2.
\]

On the other hand we have simply \( \| x \|_T = \| x \|_{L^2} + \| T x \|_{L^2} \leq (1 + \| T \|)\| x \|_{H^1_D} \).

We emphasize that \( \text{dom } T \) and \( T \) remain invariant under proper refactorizations \( AD = \bar{A}D \) since the closure of a densely defined closable operator is unique.

3.2. Regular DA operators

We adopt the regularity notion in \([12, \text{Sections 2.4.2 and 2.7}]\). Regularity of a linear DAE is solely a matter of its coefficients \( E, F \) and \( A, D, B \), respectively. The associated coefficient pair \( (E, F) \) is regular on \([a, b]\) exactly if one (equivalently each) associated proper triple \( (A, D, B) \) resulting from a proper factorization is regular on \([a, b]\). This is in full accordance with the fact that the closure \( T \) does not depend on the special factorization.

**Definition 3.3.** The DA operators \( \hat{T} \) and \( T \) are regular with tractability index \( \mu \in \mathbb{N} \) and characteristic values \( 0 < r_0 \leq \cdots \leq r_{\mu-1} < r_\mu = m \), if the associated DAE is regular on \([a, b]\) with these characteristics.

A regular DA operator is said to be fine, if the coefficients \( A, D, B \) are smooth enough for the existence of completely decoupling projectors.

Definition 3.3 is consistent with \([9, \text{Definition 2.1}]\). The operator \( \hat{T} \) and its closure \( T \) share their index and characteristic values. Recall that \( l := m - \sum_{j=0}^{\mu-1} (m - r_j) \geq 0 \) is the dynamical degree of freedom of the DAE. \( l = 0 \) may happen, see, e.g., the example DAE in Section 2.2.

A regular on \([a, b]\) DAE is associated with the so-called **canonical projector function**
\[
\Pi_{can} \in C([a, b], \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)), \quad \Pi_{can}(t)^2 = \Pi_{can}(t), \quad \text{rank } \Pi_{can}(t) = l, \quad t \in [a, b],
\]
which can be seen as generalization of the spectral projection for regular matrix pencils. Each IVP for the matrix-valued function \( X \),
\[
A(DX)' + BX = 0, \quad X(c) = \Pi_{can}(c), \quad c \in [a, b],
\]
\footnote{Note that this is at the same time consistent with \([12, \text{Theorem 2.21}]\).}
is uniquely solvable. We refer to [12, Section 2.4], [13, Section 2.2], [14, Section 4.2] for details, but note that, owing to the continuity of the coefficients, the solution of the above IVP, \(X(\cdot, c)\) is well-defined and continuous with continuously differentiable part \(DX(\cdot, c)\). \(X(t, c)\) is called \textit{maximal-size fundamental solution matrix normalized at c}. It has constant rank \(l < m\), and it may happen that \(l = 0\), see Section[2]. This is in contrast to the case of regular ODEs.

In the following we focus on bounded DA operators given in their natural Hilbert spaces,

\[
T : H^1_D(I, \mathbb{R}^m) \to L^2(I, \mathbb{R}^m), \quad T x := A(Dx)' + Bx, \quad x \in H^1_D(I, \mathbb{R}^m), \quad I := (a, b).
\]

If \(T\) is regular, then

\[
dim \ker T = l.
\]

Introduce the operator \(T_{IC}\) to capture initial conditions as well as the composed operator \(T\) by

\[
T_{IC} : H^1_D(I, \mathbb{R}^m) \to \mathbb{R}^l, \quad T_{IC} x = G_a x(a), \quad x \in H^1_D(I, \mathbb{R}^m),
\]

\[
T : H^1_D(I, \mathbb{R}^m) \to L^2(I, \mathbb{R}^m) \times \mathbb{R}^l, \quad T x = (T x, T_{IC} x), \quad x \in H^1_D(I, \mathbb{R}^m).
\]

Thereby we suppose that \(G_a \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^l)\) and \(\ker D(a) \subseteq \ker G_a\). The latter condition ensures the relation \(G_a = G_a D(a)^+ D(a)\). Then owing to the continuous embedding \(H^1 \hookrightarrow C\) the operator \(T_{IC}\) is well-defined and bounded.

Next we adopt the notion of accurately stated initial condition [13, Definition 2.3] accordingly.

**Definition 3.4.** The operator \(T_{IC}\) is accurately stated if \(\im T = \im T \times \mathbb{R}^l\) and the composed operator \(T\) is injective.

Owing to [13, Corollary 2.2], if \(T\) is regular, then \(T_{IC}\) is accurately stated exactly if \(\ker G_a \cap \im H_{can}(a) = \{0\}\).

Supposing \(T\) to be fine and \(T_{IV}\) to be accurately stated, the inverse of the composed operator \(T\) can be represented as (e.g., [13, Section 2.2], [12, Section 2.6]).

\[
T^{-1}(g, d)(t) = \underbrace{X(t, a)G_a^{-1}d}_{\in \im H_{can}(t)} + \underbrace{(V g)(t)}_{\in \ker H_{can}(t)} + \underbrace{(\mathcal{D} g)(t)}_{\in \im H_{can}(t)}, \quad (g, d) \in \im T \times \mathbb{R}^l,
\]

(2)

\[
(V g)(t) = \int_a^t X(t, s)G_{\mu}(s)^{-1} g(s) ds,
\]

(3)

\[
\mathcal{D} g = v_0 + v_1 + \cdots + v_{\mu-1},
\]

(4)

in which the functions \(v_i\) are successively explicitly determined by simple multiplications with certain matrix coefficients, by differentiation of \(D v_j, \quad j = i + 1, \ldots, \mu - 1\), and subsequent linear
combinations,

\[ v_{\mu-1} = L_{\mu-1}g, \]
\[ v_{\mu-2} = L_{\mu-2}g - N_{\mu-2,\mu-1}(Dv_{\mu-1})', \]
\[ \vdots \]
\[ v_1 = L_1g - \sum_{s=2}^{\mu-1} N_{1,s}(Dv_s)' - \sum_{s=3}^{\mu-1} M_{1,s}v_s, \]
\[ v_0 = L_0g - \sum_{s=1}^{\mu-1} N_{0,s}(Dv_s)' - \sum_{s=2}^{\mu-1} M_{0,s}v_s. \]

The coefficients \( N_{i,k}, M_{i,k}, L_i \), and \( G_\mu \) are at least continuous. They are fully determined by \( A, D, B \) via a sequence of admissible matrix functions corresponding to a complete decoupling.

With the representation (2) of the inverse \( T^{-1} \) we intend to emphasize, on the one hand, the partial resemblance to (1). On the other hand, the second term \( Dg \) emerge for DA operators only. It is a differential operator, with may be higher order, see Section 2. If \( T \) has index \( \mu > 1 \) then \( T^{-1} \) includes derivatives up to order \( \mu - 1 \).

The representation (2) shows that the inverse of the composed operator actually decomposes into two parts. The first “good” part is close to (1), i.e., the case of regular ODEs. This part may disappear. The second part is always present and representative for DA operators. Unfortunately, the canonical projector function which separates the parts is practically available in a few special cases only.

We close this subsection by quoting further relevant results from [9, Section 2]

**Theorem 3.5.** Let \( T \) be fine with index \( \mu \in \mathbb{N} \) and \( T_{IC} \) be accurately stated.

(i) If \( \mu = 1 \), then \( \text{im} \ T = L^2(I, \mathbb{R}^m) \), and the composed operator \( T \) is a homeomorphism.

(ii) If \( \mu > 1 \), then \( \text{im} \ T \subset L^2(I, \mathbb{R}^m) \) is nonclosed and the inverse \( T^{-1} \) as well as the Moore-Penrose inverse \( T^+ \) are no longer continuous. Then the equation \( Tx = (g, d) \) is essentially ill-posed in Tikhonov’s sense.

4. Justification of the overdetermined polynomial collocation

In this section we deal with regular higher-index DA operators \( T \), the related composed operator \( T \) and their approximations \( R_{\pi,M}TU_\pi \) and \( R_{\pi,M}TU_\pi \). On the background of the corresponding properties we provide new sufficient convergence conditions for the overdetermined polynomial collocation.

4.1. Basic technicalities

We consider the linear IVP or BVP,

\[ A(t)(Dx)'(t) + B(t)x(t) = g(t), \quad t \in [a, b], \]
\[ G_\alpha x(a) + G_\beta x(b) = d, \]
with the constant matrix \( D = [I \ 0] \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k) \), \( \text{rank} \, D = k \), and at least continuous matrix coefficients \( A : [a, b] \to \mathcal{L}(\mathbb{R}^k, \mathbb{R}^m) \), \( B : [a, b] \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m) \). The DAE (5) is supposed to be fine in the sense of [12, Section 2.6], with tractability index \( \mu \in \mathbb{N} \) and dynamical degree of freedom \( l \leq k \). Recall that \( \mu > 1 \) necessarily implies \( l < k \). The matrices \( G_a, G_b \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^l) \) are supposed to satisfy the conditions

\[
\ker D \subseteq \ker G_a, \quad \ker D \subseteq \ker G_b.
\]

Condition (6) is further supposed to be accurately stated in the sense of [13, Definition 2.3]), so that the problems

\[
A(t)(Dx)'(t) + B(t)x(t) = 0, \quad t \in [a, b], \quad G_a(x(a)) + G_b(x(b)) = d,
\]

are uniquely solvable for each \( d \in \mathbb{R}^l \). In particular, the homogeneous linear BVP has the trivial solution only.

The function \( g \) is assumed to be admissible, so that the DAE (5) is solvable. Then the BVP (5), (6) has exactly one solution \( x_* \) to be approximated later on.

Following the ideas of [10, 9] we represent the BVP (5), (6) as operator equation \( T \, x = y := (g, d) \) in Hilbert spaces by introducing the spaces

\[
L^2 := L^2((a, b), \mathbb{R}^m), \quad H^1_D := H^1_D((a, b), \mathbb{R}^m) := \{x \in L^2 : Dx \in H^1((a, b), \mathbb{R}^k)\},
\]

equipped with the inner products

\[
(x, \bar{x})_{L^2} = \int_a^b \langle x(t), \bar{x}(t) \rangle \, dt, \quad x, \bar{x} \in L^2, \quad (x, \bar{x})_{H^1_D} = (x, \bar{x})_{L^2} + \langle (Dx)', (\bar{Dx})' \rangle_{L^2((a, b), \mathbb{R}^k)}, \quad x, \bar{x} \in H^1_D,
\]

and operators

\[
T : H^1_D \to L^2, \quad Tx = A(Dx)' + Bx, \quad x \in H^1_D,
\]

\[
T_{BC} : H^1_D \to \mathbb{R}^l, \quad T_{BC}x = G_a(x(a)) + G_b(x(b)), \quad x \in H^1_D,
\]

\[
\mathcal{T} : H^1_D \to L^2 \times \mathbb{R}^l := Y, \quad \mathcal{T}x = (Tx, T_{BC}x), \quad x \in H^1_D.
\]

The operator \( T_{BC} \) is well defined and bounded owing to condition (7) and the continuous embedding \( H^1((a, b), \mathbb{R}^k) \hookrightarrow C([a, b], \mathbb{R}^k) \). Then, the DA operator \( T \) as well as the composed operator \( \mathcal{T} \) are obviously bounded. Moreover, \( \mathcal{T} \) is injective and \( \text{im} \, \mathcal{T} = \text{im} \, T \times \mathbb{R}^l \). At this place let us emphasize again that we focus our interest on higher-index DAEs, \( \mu \geq 2 \), but then \( \text{im} \, T \) is a nonclosed subset of \( L^2 \) and \( \mathcal{T}^{-1} \) is an unbounded operator, cf. Section 3 also [10, 9, 14].

Given the partition

\[
\pi : a = t_0 < t_1 < \cdots < t_n = b,
\]

with stepsizes \( h_j = t_j - t_{j-1} \), minimal stepsize \( h_{\pi, \text{min}} \), and maximal stepsize \( h_{\pi} \), we denote by \( C_\pi = C_\pi([a, b], \mathbb{R}^m) \) the space of piecewise continuous functions having breakpoints merely at the mesh points. Note that the supremum-norm \( \| \cdot \|_{L^\infty} \) is well-defined for the elements of \( C_\pi \).
Next we fix a number $N \geq 1$ and introduce the space of ansatz functions to approximate $x$, by piecewise polynomial functions,

$$X_\pi = \{ x \in C_\pi([a, b], \mathbb{R}^m) : Dx \in C([a, b], \mathbb{R}^k), \}
$$

$$x_{\kappa}|_{t_{j-1}, t_j} \in \Psi_N, \quad \kappa = 1, \ldots, k, \quad x_{\kappa}|_{t_{j-1}, t_j} \in \Psi_{N-1}, \quad \kappa = k + 1, \ldots, m, \quad j = 1, \ldots, n. \tag{10}$$

The finite-dimensional space $X_\pi$ is a closed subspace of $H_D^1$, and the latter decomposes into the topological sum $X_\pi \oplus X_\kappa^1 = H_D^1$. We agree upon that

$$U_\pi : H_D^1 \rightarrow H_D^1 \quad \text{denotes the orthoprojection operator of } H_D^1 \text{ onto } X_\pi.$$

For later reference, the following norm in $X_\pi$ will be needed:

$$\|x\|_{C_D^1} = \|x\|_\infty + \|(Dx)'\|_\infty, \quad x \in X_\pi.$$

The ansatz space $X_\pi$ has dimension $nNm + k$. Choosing values

$$0 < \tau_1 < \cdots < \tau_M < 1$$

we specify $M \geq N + 1$ collocation points per subinterval, i.e.,

$$t_{ji} = t_{j-1} + \tau_j h_j, \quad i = 1, \ldots, M, \quad j = 1, \ldots, n,$$

and are then confronted with the overdetermined collocation system of $nMm + l > nMm + k$ equations for providing an approximation $x \in X_\pi$, namely,

$$A(t_{ji})(Dx)'(t_{ji}) + B(t_{ji}) x(t_{ji}) - g(t_{ji}) = 0, \quad i = 1, \ldots, M, \quad j = 1, \ldots, n, \tag{11}$$

$$G_a x(t_0) + G_b x(t_n) - d = 0. \tag{12}$$

As a matter of course, the choice $M > N$ goes along with an overdetermined system comprising more equations than unknowns. This is different from standard collocation methods for ODEs and index-1 DAEs, e.g., [13]. Here we treat the overdetermined collocation system in a least-squares sense.

Let $R_{\pi,M} : C_\pi([a, b], \mathbb{R}^m) \rightarrow C_\pi([a, b], \mathbb{R}^m)$ denote the restriction operator which assigns to $w \in C_\pi([a, b], \mathbb{R}^m)$ the piecewise polynomial $R_{\pi,M} w \in C_\pi([a, b], \mathbb{R}^m)$ of degree less than or equal to $M - 1$ such that the interpolation conditions,

$$(R_{\pi,M} w)(t_{ji}) = w(t_{ji}), \quad i = 1, \cdots, M, \quad j = 1, \cdots, n,$$

are satisfied. We also assign to $w \in C_\pi([a, b], \mathbb{R}^m)$ the vector $W \in \mathbb{R}^{mMn}$,

$$W = \begin{bmatrix} W_1 \\ \vdots \\ W_n \end{bmatrix} \in \mathbb{R}^{mMn}, \quad W_j = \left( \frac{h_j}{M} \right) \begin{bmatrix} w(t_{j1}) \\ \vdots \\ w(t_{jM}) \end{bmatrix} \in \mathbb{R}^{mM}.$$
which yields (cf. [10, Subsection 2.3])

\[ \|R_{\pi,M}w\|_{L^2}^2 = W^T L W, \quad w \in C_\pi([a, b], \mathbb{R}^m), \]

with the matrix \( L \) being positive definite, symmetric and independent of \( h_\pi \). There are further constants \( \kappa_l, \kappa_u > 0 \) such that

\[ \kappa_l |W|^2 \leq W^T L W \leq \kappa_u |W|^2, \quad W \in \mathbb{R}^{mn}. \]

If \( w \in C_\pi([a, b], \mathbb{R}^m) \) is of class \( C^M \) on each subinterval of the partition \( \pi \), then

\[ |w(t) - (R_{\pi,M}w)(t)| \leq \frac{1}{M!} \|w^{(M)}\|_{\infty} h_\pi^M, \quad t \in [a, b]. \]

Additionally, we introduce the restriction operator \( R_{\pi,M} : C_\pi([a, b], \mathbb{R}^m) \times \mathbb{R}^l \rightarrow C_\pi([a, b], \mathbb{R}^m) \times \mathbb{R}^l \) by

\[ R_{\pi,M}y = (R_{\pi,M}g, d), \quad y = (g, d) \in C_\pi([a, b], \mathbb{R}^m) \times \mathbb{R}^l. \]

The overdetermined least-squares collocation means now that we seek an element \( \tilde{x}_\pi \in X_\pi \) minimizing the functional

\[ \psi_{\pi,M}(x) = \|R_{\pi,M}(T x - y)\|_{L^2(\mathbb{R}^l)}^2 = \|R_{\pi,M}(T x - g)\|_{L^2}^2 + |T_{BC} x - d|^2, \quad x \in X_\pi. \]

With \( w = T x - g \) we may represent

\[ \psi_{\pi,M}(x) = W^T L W + |T_{BC} x - d|^2, \quad x \in X_\pi, \]

which reveals that by minimizing \( \psi_{\pi,M}(x) \) subject to \( x \in X_\pi \) we actually provide a least-squares solution of the collocation system (11),(12). The mathematics behind is closely related to special properties of the restriction operator \( R_{\pi,M} \) on the one hand, but on the other hand, to the problem to minimize the functional

\[ \psi(x) = \|T x - y\|_{L^2(\mathbb{R}^l)}^2 = \|T x - g\|_{L^2}^2 + |T_{BC} x - d|^2, \quad x \in X_\pi, \]

for which (16) serves as approximation.

4.2. The operators \( R_{\pi,M}TU_\pi \) and \( R_{\pi,M}TU_\pi \)

We begin this section by providing useful norm inequalities. Regarding convergence properties for \( h_\pi \) tending to zero we have in mind sequences of partitions. In favor for an easier reading we drop an extra labeling, but we thoroughly assure that the indicated constants do not depend on the partitions and stepsizes \( h_\pi \) in fact. We allow partitions \( \pi \) having quotients \( h_\pi/h_{\pi,\min} \leq r \), with a global bound \( 1 \leq r < \infty \). A consequence of [7, Theorem 3.2.6] is that there exists a constant \( c_K = c_K(r) \) such that

\[ \|z\|_{\infty} \leq c_K h_\pi^{-\frac{3}{2}} \|z\|_{L^2}, \]

for all functions \( z \in C_\pi([a, b], \mathbb{R}^r) \) being a polynomial of degree less than or equal to \( K \) on each subinterval of the partition \( \pi \).

\[ ^2 \text{The entries of } L \text{ are fully determined by the corresponding } M \text{ Lagrangian basis polynomials, thus, by } M \text{ and } \tau_1, \ldots, \tau_M. \]
Lemma 4.1. There is a constant \( \kappa > 0 \) such that

\[
kh_{\pi} \|x\|_{C_D}^2 \leq \|x\|_{H_D^1}^2 \leq (b - a) \|x\|_{C_D}^2, \quad x \in X_\pi.
\]

Proof. The inequality \( \|x\|_{H_D^1}^2 \leq (b - a) \|x\|_{C_D}^2 \) is evident for all \( x \in X_\pi \). On the other hand, owing to (19) each arbitrary \( x \in X_\pi \) satisfies

\[
\|x\|_{C_D}^2 \leq c_N h_{\pi}^{-1} \|x\|_{L^2}^2, \quad \|(Dx)'\|_{C_D}^2 \leq c_{N-1} h_{\pi}^{-1} \|(Dx)'\|_{L^2}^2,
\]

and finally

\[
\|x\|_{C_D}^2 \leq 2 \|x\|_{C_D}^2 + 2 \|(Dx)'\|_{C_D}^2 \leq 2(c_N^2 + c_{N-1}^2) h_{\pi}^{-1} \|x\|_{H_D^1}^2 =: \frac{1}{\kappa} h_{\pi}^{-1} \|x\|_{H_D^1}^2.
\]

\( \square \)

If \( A \) and \( B \) are constant matrices, then \( T \) is piecewise polynomial with degree less than or equal to \( N \) for \( x \in X_\pi \). Owing to \( M \geq N + 1 \) this leads to

\[
R_{\pi,M} T U_{\pi} x = T U_{\pi} x, \quad R_{\pi,M} T U_{\pi} x = T U_{\pi} x, \quad x \in H_D^1.
\]

(20)

In general, the operators \( R_{\pi,M} T U_{\pi} \) and thus \( R_{\pi,M} T U_{\pi} \) are well-defined on \( H_D^1 \) since \( T U_{\pi} x \) belongs to \( C_{\pi} \) for all \( x \in H_D^1 \).

Proposition 4.2. Let the DA operator \( T \) be fine with index \( \mu \in \mathbb{N} \) and \( T_{BC} \) be accurately stated. Let \( M \geq N + 1 \) and let the entries of \( A \) and \( B \) be of class \( C^M \). Then the following assertions are valid for all sufficiently fine partitions \( \pi \):

(i) There is a constant \( C_{AB1} \) such that

\[
\|R_{\pi,M} T U_{\pi} x\|_{L^2} \leq C_{AB1} \|x\|_{H_D^1}, \quad x \in H_D^1,
\]

\[
\|R_{\pi,M} T U_{\pi} x\|_{L^2 \times \mathbb{R}'} \leq (C_{AB1}^2 + \|T_{BC}\|^2)^{\frac{1}{2}} \|x\|_{H_D^1}, \quad x \in H_D^1.
\]

(ii) There is a constant \( C_{AB2} \) such that

\[
\|R_{\pi,M} T U_{\pi} x - T U_{\pi} x\|_{L^2} \leq C_{AB2} h_{\pi}^{M - \frac{1}{2}} \|x\|_{H_D^1}, \quad x \in H_D^1,
\]

\[
\|R_{\pi,M} T U_{\pi} x - T U_{\pi} x\|_{L^2 \times \mathbb{R}'} \leq C_{AB2} h_{\pi}^{M - \frac{1}{2}} \|x\|_{H_D^1}, \quad x \in H_D^1.
\]

(iii) If additionally \( M \geq N + \mu \), then there is a constant \( C \) such that

\[
\ker R_{\pi,M} T U_{\pi} = \ker T U_{\pi} = \ker U_{\pi}, \quad \text{and} \quad \|(R_{\pi,M} T U_{\pi})^\dagger\| \leq C h_{\pi}^{-1-\mu}.
\]

(iv) If the entries of \( A \) and \( B \) are polynomials of degree less than or equal to \( N_{AB} \) and \( M \geq N + 1 + N_{AB} \), then (20) and (22) remain valid, too.
Proof. (i) We choose a number $K \geq 1$, $K + N \leq M$, $K$ interpolations nodes $0 < \tilde{\tau}_1 < \cdots < \tilde{\tau}_K < 1$ and provide piecewise entrywise polynomial approximations $\tilde{\Phi}$ and $\tilde{\Phi}$ of degree $K - 1$ such that

\[ \tilde{\Phi}(t_{j-1} + \tilde{\tau}_i h_{j-1}) = A(t_{j-1} + \tilde{\tau}_i h_{j-1}), \quad \tilde{\Phi}(t_{j-1} + \tilde{\tau}_i h_{j-1}) = B(t_{j-1} + \tilde{\tau}_i h_{j}), \quad i = 1, \ldots, K, \quad j = 1, \ldots, n. \]

Then there are constants $c_1, c_2$ such that

\[ ||\tilde{\Phi} - A||_\infty \leq c_1 h^K_n, \quad ||\tilde{\Phi} - A||_\infty \leq c_1 h^K_n \]

\[ ||\tilde{\Phi}||_\infty \leq ||A||_\infty + ||\tilde{\Phi} - A||_\infty \leq ||A||_\infty + c_1 h^K_n \leq c_2, \quad ||\tilde{\Phi}||_\infty \leq ||B||_\infty + ||\tilde{\Phi} - B||_\infty \leq ||B||_\infty + c_1 h^K_n \leq c_2. \]

The auxiliary operator $\tilde{T} : H_D^1 \to L^2$

\[ \tilde{T} x = \tilde{\Phi}(Dx)' + \tilde{B} x, \quad x \in H_D^1, \]

is bounded, $||\tilde{T}|| \leq \sqrt{2} c_2$ and owing to $M \geq N + K$ it holds that $R_{\pi, \tilde{T}} U_{\tilde{T}} x = \tilde{T} U_{\tilde{T}} x$ for all $x \in H_D^1$. Next, to each arbitrary $x \in X_\pi$ we form $v := (\tilde{T} - T) x \in C_\pi$ and the vector $V \in \mathbb{R}^{mMn}$,

\[ V = \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix} \in \mathbb{R}^{mMn}, \quad V_j = \begin{bmatrix} h_j M \\ v(t_{j1}) \\ \vdots \\ v(t_{jM}) \end{bmatrix} \in \mathbb{R}^{mM}, \]

which yields (cf. (13))

\[ ||R_{\pi, \tilde{T}} V||_{L^2}^2 = V^T L V \leq \kappa_v ||V||^2 = \kappa_v \sum_{j=1}^n \frac{h_j}{M} \sum_{i=1}^M |v(t_{ji})|^2 \leq \kappa_v (b - a)||v||^2, \]

On the other hand, regarding Lemma 4.1 and $x = U_{\pi} x$ we estimate

\[ ||v||_\infty \leq c_1 h^K_n ||(Dx)'||_\infty + c_1 h^K_n ||x||_\infty = c_1 h^K_n ||x||_{C^1_D} = c_1 h^K_n ||U_{\pi} x||_{C^1_D} \leq c_1 h^K_n \frac{1}{\sqrt{K}} ||x||_{H^1_D}. \]

For each arbitrary $x \in H_D^1$ it follows that

\[ ||R_{\pi, \tilde{T}} U_{\pi} x||_{L^2} \leq ||R_{\pi, \tilde{T}} U_{\pi} x||_{L^2} + ||R_{\pi, \tilde{T}} U_{\pi} x||_{L^2} = ||R_{\pi, \tilde{T}} V||_{L^2} + \sqrt{2} c_2 ||x||_{H^1_D}, \]

which verifies the inequality (21) with a suitable bound $C_{AB1}$. Then it also results that

\[ ||R_{\pi, \tilde{T}} U_{\pi} x||_{L^2}^2 = ||R_{\pi, \tilde{T}} U_{\pi} x||_{L^2}^2 + ||T_{BC} U_{\pi} x||^2 \leq (C_{AB1}^2 + ||T_{BC}||^2)||x||_{H^1_D}^2, \]

and the assertion is verified.

(ii) To each arbitrary $x \in X_\pi$ we set $w = A(Dx)' + Bx$ and derive on each subinterval of the partition $\pi$ that

\[ w' = A(Dx)'' + A'(Dx)' + Bx' + B'x, \]

\[ \vdots \]

\[ w^{(M)} = A(Dx)^{(M+1)} + \cdots + A^{(M)}(Dx)' + Bx^{(M)} + \cdots + B^{(M)} x. \]
Since \((Dx)^{(N+1)}\) and \(x^{(N+1)}\) vanish identically, we obtain the inequality
\[
|w^{(M)}(t)| \leq c_3(\|x\|_{C^1_D} + \|x'\|_{C^1_D} + \cdots + \|x^{(N)}\|_{C^1_D}),
\]
with a constant \(c_3\) being determined by the coefficients \(A\) and \(B\), and their involved derivatives. By Lemma 4.1 this yields
\[
\|w^{(M)}\|_{\infty} \leq c_3\frac{1}{\sqrt{k}}h_\pi^{-\frac{1}{2}}(\|x\|_{H^1_D} + \|x'\|_{H^1_D} + \cdots + \|x^{(N)}\|_{H^1_D}).
\]
Owing to [9, Lemma 4.2] it follows that
\[
\|w^{(M)}\|_{\infty} \leq c_3\frac{1}{\sqrt{k}}h_\pi^{-\frac{1}{2}}(\|x\|_{H^1_D} + h^{-1}_\pi r \sqrt{C_1^\pi}\|x\|_{H^1_D} + \cdots + h^{-N}_\pi r^N \sqrt{C_N^\pi}\|x\|_{H^1_D}).
\]
We emphasize that the \(C^*_i\) are also constants independent of the partition and stepsize. Then there is a constant \(c_4\) such that
\[
\|w^{(M)}\|_{\infty} \leq c_4h^{-N-\frac{1}{2}}\|x\|_{H^1_D}.
\]
Finally we arrive at
\[
\|R_{\pi,M}TU_\pi x - TU_\pi x\|_{L^2} = \|R_{\pi,M}w - w\|_{L^2} \leq \frac{\sqrt{b - a}}{M!}h_\pi^{-M}\|w^{(M)}\|_{\infty} \leq \frac{\sqrt{b - a}}{M!}c_4h^{-N-\frac{1}{2}}\|x\|_{H^1_D}
\]
and further
\[
\|R_{\pi,M}TU_\pi x - TU_\pi x\|_{L^2} \leq \frac{\sqrt{b - a}}{M!}c_4h^{-M-\frac{1}{2}}\|x\|_{H^1_D}.
\]
(iii) The Moore-Penrose inverse of \(TU_\pi\) satisfies, by [9, Theorem 4.1], the inequality
\[
\|(TU_\pi)^+\| \leq \frac{\mu h^{-\mu-1}_\pi}{c_\gamma},
\]
with a positive constant \(c_\gamma\).
Denote by \(V_\pi\) and \(V_{\pi,M}\) the orthoprojectors of \(L^2 \times \mathbb{R}^l\) onto \(\text{im}\,TU_\pi\) and \(\text{im}\,R_{\pi,M}TU_\pi\), respectively. Assertion (ii) implies now
\[
\|V_\pi R_{\pi,M}TU_\pi - TU_\pi\| = \|V_\pi(R_{\pi,M}TU_\pi - TU_\pi)\| \\
\leq \|R_{\pi,M}TU_\pi - TU_\pi\| \leq C_{AB}h_\pi^{-M-\frac{1}{2}},
\]
and further, for sufficiently fine partitions,
\[
\|(TU_\pi)^+\| \leq \|V_\pi R_{\pi,M}TU_\pi - TU_\pi\| \leq C_{AB}h^{-M-\frac{1}{2}+\mu}_\pi \leq \frac{1}{2}.
\]
Next we represent
\[ \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi = \mathcal{V}_\pi (\mathcal{R}_{\pi, M} T U_\pi - T U_\pi) + T U_\pi. \]

The subspace \( \text{im} \mathfrak{A} = \text{im} T U_\pi \) has finite dimension, \( \text{im} \mathfrak{E} \subseteq \text{im} \mathfrak{A} \). Furthermore, it holds that \( \ker \mathfrak{A} = \ker T U_\pi = \ker U_\pi \) and \( \ker \mathfrak{E} \supseteq \ker U_\pi \). By [8, Lemma A.2] it follows that
\[
\ker \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi = \ker \mathfrak{A}, \quad \| (\mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi)^+ \| \leq \frac{\| \mathfrak{A}^+ \|}{1 - \| \mathfrak{A}^+ \| \| \mathfrak{E} \|} \leq 2 \| \mathfrak{A}^+ \|. 
\]

Now the inclusions
\[
\ker U_\pi \subseteq \ker \mathcal{R}_{\pi, M} T U_\pi \subseteq \ker \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi = \ker U_\pi 
\]
hold and, hence, \( \ker \mathcal{R}_{\pi, M} T U_\pi = \ker U_\pi \). In the end we compute
\[
(\mathcal{R}_{\pi, M} T U_\pi)^+ = (\mathcal{R}_{\pi, M} T U_\pi - \mathcal{R}_{\pi, M} T U_\pi)^+ + U_\pi (\mathcal{R}_{\pi, M} T U_\pi)^+ \\
= (\mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi)^+ \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi (\mathcal{R}_{\pi, M} T U_\pi)^+ + \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi (\mathcal{R}_{\pi, M} T U_\pi)^+ \mathcal{V}_\pi \mathcal{R}_{\pi, M} T U_\pi,
\]
and hence \( \| (\mathcal{R}_{\pi, M} T U_\pi)^+ \| \leq 2 \| \mathfrak{A}^+ \| \).

(iv) This assertion is a direct consequence of the fact that \( T U_\pi x \) is piecewise polynomial of degree less than or equal to \( N + N_{AB} \).

4.3. Error estimations
Recall that \( x_\pi \) denotes the sought solution, i.e., \( T x_\pi = y \) for given \( y = (g, d) \in \text{im} T \times \mathbb{R}^l \). As described in Subsection 4.1 the overdetermined polynomial collocation actually means that we generate the minimizer of the functional \( \phi_{\pi, M} \), cf. (17), that is,
\[
\bar{x}_\pi = \arg\min_{x \in X_\pi} \{ \| \mathcal{R}_{\pi, M}(T U_\pi x - y) \|_{L^2 \times \mathbb{R}^l}^2 : x \in X_\pi \} = (\mathcal{R}_{\pi, M} T U_\pi)^+ \mathcal{R}_{\pi, M} y,
\]
to approximate \( x_\ast \). Now we provide a corresponding error estimate.
Suppose that the solution is smooth, \( x_\ast \in C^N([a, b], \mathbb{R}^m) \), \( D x_\ast \in C^{N+1}([a, b], \mathbb{R}^m) \). With the \( N \) nodes \( 0 < \tau_{s_1} < \cdots < \tau_{s_N} < 1 \), the interpolating function \( p_s \in X_\pi \) uniquely defined by
\[
p_s(t_{j-1} + \tau_s h_j) = x_\ast(t_{j-1} + \tau_s h_j), \quad i = 1, \ldots, N, \quad j = 1, \ldots, n, \quad D p_s(t_0) = D x_\ast(t_0),
\]
satisfies the inequalities
\[
\| p_s - x_\ast \|_{C^0_\mathfrak{D}} \leq c_s h_\pi^N, \\
\| p_s - x_\ast \|_{H^1_\mathfrak{B}} \leq \sqrt{b-a} c_s h_\pi^N =: c_{a} h_\pi^N, \\
\| U_\pi x_\ast - x_\ast \|_{H^1_\mathfrak{D}} \leq \| p_s - x_\ast \|_{H^1_\mathfrak{B}} \leq c_{a} h_\pi^N,
\]
in which the constant \( c_s \) is determined by \( x_\ast \). Next, owing to Proposition 4.2(i), we may estimate
\[
\| \mathcal{R}_{\pi, M}(T p_s - T U_\pi x_\ast) \| = \| \mathcal{R}_{\pi, M}(T U_\pi p_s - T U_\pi x_\ast) \| = \| \mathcal{R}_{\pi, M} T U_\pi (p_s - x_\ast) \| \\
\leq (C_{AB}^2 + \| T_{BC} \|^2)^{\frac{1}{2}} \| p_s - x_\ast \|_{H^1_\mathfrak{B}} \leq (C_{AB}^2 + \| T_{BC} \|^2)^{\frac{1}{2}} c_{a} h_\pi^N.
\]
Denoting \( w_\pi = T(x_\pi - p_\pi) \in C_\pi \) and using the Lagrange basis polynomials we further derive

\[
\| R_{\pi,M} w_\pi \|_\infty \leq \max_{j=1,...,n} \max_{j_i \neq j} \sum_{i=1}^{M} |l_j(t)| \| w_\pi \|_\infty = \max_{j=1,...,n} \max_{j_i \neq j} \sum_{i=1}^{M} \left| \frac{T - \tau_j}{T_i - \tau_j} \right| \| w_\pi \|_\infty =: C_L \| w_\pi \|_\infty.
\]

Because of

\[
\| w_\pi \|_\infty \leq \max \{ \| A \|_\infty, \| B \|_\infty \} \| x_\pi - p_\pi \|_{C_\pi} \leq \max \{ \| A \|_\infty, \| B \|_\infty \} \ c_\pi h_\pi^N
\]

we obtain

\[
\| R_{\pi,M} T(x_\pi - p_\pi) \|_\infty = \| R_{\pi,M} w_\pi \|_\infty \leq C_L \max \{ \| A \|_\infty, \| B \|_\infty \} \ c_\pi h_\pi^N,
\]

such that

\[
\| R_{\pi,M} T(x_\pi - p_\pi) \|_{L_2^{\infty} C^1} = \| R_{\pi,M} T(x_\pi - p_\pi) \|_{L_2}^2 + |T_{BC}(x_\pi - p_\pi)|^2
\]

\[
\leq (b-a) C_L^2 \max \{ \| A \|_\infty^2, \| B \|_\infty^2 \} c_\pi^2 h_\pi^{2N} + \| T_{BC} \|_{C_\pi} c_\pi^2 h_\pi^{2N} =: C_R h_\pi^{2N}.
\]

Proposition 4.2(iii) implies, for \( M \geq N + \mu \), that \( U_\pi = (R_{\pi,M} T U_\pi)^{+} R_{\pi,M} T U_\pi \), which gives rise to the error representation

\[
\tilde{x}_\pi - U_{\pi} x_\pi = (R_{\pi,M} T U_\pi)^{+} R_{\pi,M} y - (R_{\pi,M} T U_\pi)^{+} R_{\pi,M} T U_{\pi} x_\pi
\]

\[
= (R_{\pi,M} T U_\pi)^{+} R_{\pi,M} T x_\pi - (R_{\pi,M} T U_\pi)^{+} R_{\pi,M} T U_{\pi} x_\pi
\]

\[
= (R_{\pi,M} T U_\pi)^{+} (R_{\pi,M} T (x_\pi - p_\pi) - R_{\pi,M} T U_\pi (x_\pi - p_\pi)),
\]

so that

\[
\| \tilde{x}_\pi - U_{\pi} x_\pi \|_{H_D^{\mu}} \leq \| (R_{\pi,M} T U_\pi)^{+} \| (\| R_{\pi,M} T (x_\pi - p_\pi) \| + \| R_{\pi,M} T U_\pi (x_\pi - p_\pi) \|)
\]

\[
\leq C R_{\pi}^{1-\mu} (\sqrt{C_R} h_\pi^{N} + (C_{AB}^2 + \| T_{BC} \|_{C_\pi}^2 c_\pi^2 h_\pi^{2N}) =: C_\pi h_\pi^{N-\mu+1}.
\]

At the end we arrive at

\[
\| \tilde{x}_\pi - x_\pi \|_{H_D^{\mu}} \leq \| \tilde{x}_\pi - U_{\pi} x_\pi \|_{H_D^{\mu}} + \| x_\pi - U_{\pi} x_\pi \|_{H_D^{\mu}} \leq (C_\pi + c_\pi) h_\pi^{N-\mu+1}.
\]

We summarize the result in the following theorem.

**Theorem 4.3.** Let the DA operator \( T \) be fine with index \( \mu \in \mathbb{N} \) and \( T_{BC} \) be accurately stated. Let \( M \geq N + \mu \) and let the entries of \( A \) and \( B \) be of class \( C^M \). Let \( g \in \text{im} \ T \), and the solution \( x_\pi = T^{-1}(g, d) \) be of class \( C^N \) with \( Dx_\pi \) of class \( C^{N+1} \). Then there is a constant \( \tilde{C} \) such that the error estimation

\[
\| \tilde{x}_\pi - x_\pi \|_{H_D^{\mu}} \leq \tilde{C} h_\pi^{N-\mu+1}
\]

is valid for all sufficiently fine partitions \( \pi \).
In contrast to Theorem 4.3, the earlier error estimation from [9, Theorem 3.1 (a)] is given for the least-squares approximation \( x_\pi \) (cf. (18)),

\[
x_\pi = \arg\min\{ \phi(x) : x \in X_\pi \} = \arg\min\{ \| T U_\pi x - y \|_{L^2_{x \in \mathbb{R}^l}}^2 : x \in X_\pi \} = (T U_\pi)^+ y.
\]

Then, supposing the entries of \( A \) and \( B \) to be polynomials of degree less than or equal to \( N_{AB} \) and letting \( M \geq N + 1 + N_{AB} \) the estimation (23) is derived by a different technique ([9, Theorem 5.1(a)]). In the present context of operator properties this means that then the operators \( R_{\pi, M} T U_\pi \) and \( T U_\pi \) coincide, see Proposition 4.2.

So far we know only sufficient convergence and order conditions. The question concerning an appropriate or even optimal choice of \( N \geq 1 \) and \( M \geq N + 1 \) remains open. Of course, smaller \( N \) and \( M \) are associated with less computational effort. So far, in experiments \( M = N + 1 \) works quite well. In general, the practical performance is much better than we can substantiate till now. Much further analysis is needed.

5. Higher-order DA operators

General linear order-\( s \) DA operators have the form (e.g., [6, 4, 15])

\[
\hat{T} x = E x^{(s)} + \cdots + E_1 x' + E_0 x, \quad x \in \text{dom } \hat{T} = C([a, b], \mathbb{R}^m),
\]

with at least continuous matrix-coefficients \( E_i \) and a singular leading coefficient \( E_s \). So far the consolidated knowledge of higher-order DAEs and the related operators is rather poor. Naturally the class of higher-order DA operators is much more complex than the class of first-order ones, nevertheless the numerical treatment of the corresponding DAEs might be often easier than expected. In particular, systems of ODEs of mixed order, which can be handled by traditional approved software packages such as COLNEW, COLDAE and BVPSUITE ([11, 12, 11]), actually apply to higher-order DAEs directly. For instance, the simple system

\[
\begin{align*}
x_1'' + x_1 &= g_1, \\
x_2' + x_1 + x_2 &= g_2,
\end{align*}
\]

corresponds to the DA operator \( \hat{T} : \text{dom } \hat{T} = C^2([a, b], \mathbb{R}^2) \subset C([a, b], \mathbb{R}^2) \rightarrow C([a, b], \mathbb{R}^2) \)

\[
\hat{T} x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x'' + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x' + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x, \quad x \in \text{dom } \hat{T},
\]

and its extension \( T \), with \( \text{dom } T = \{ z \in C([a, b], \mathbb{R}^2) : z_1 \in C^2([a, b], \mathbb{R}), z_2 \in C^1([a, b], \mathbb{R}) \} \)

\[
T x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} [ \begin{bmatrix} 1 & 0 \end{bmatrix} x'' + \begin{bmatrix} 0 \\ 0 \end{bmatrix} [ \begin{bmatrix} 0 & 1 \end{bmatrix} x'] + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} x, \quad x \in \text{dom } T.
\]

One has \( \text{im } T = C([a, b], \mathbb{R}^2) \) and \( \text{dim ker } T = 3 \), and hence \( T \) is fredholm, so that IVPs and BVPs can be stated in a well-posed way. It arises the question which further DA operators allow
a reliable direct treatment of the corresponding DAE. For the time being we are not able to present a general answer. Below, we survey certain related aspects.

In the early paper [3] higher-order Hessenberg DAEs arising from higher-order ODEs subject to constraints are introduced and analyzed with respect to their perturbation index. Written in the form (24), the associated operators are

\[ \hat{T}x = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} x^{(s)} + \sum_{i=1}^{s-1} \begin{bmatrix} E_{i,11} & 0 \\ E_{i,21} & 0 \end{bmatrix} x^{(i)} + \begin{bmatrix} E_{0,11} & E_{0,12} \\ E_{0,21} & E_{0,22} \end{bmatrix} x, \quad x \in \text{dom } \hat{T} = C^s([a, b], \mathbb{R}^m) \tag{25} \]

and, with properly involved derivatives by the additional matrix \( D = [I \ 0] \), \( \text{rank } D = m_1 \),

\[ TX = \begin{bmatrix} I \\ 0 \end{bmatrix} (DX)^{(s)} + \sum_{i=1}^{s-1} \begin{bmatrix} E_{i,11} \\ E_{i,21} \end{bmatrix} (DX)^{(i)} + \begin{bmatrix} E_{0,11} & E_{0,12} \\ E_{0,21} & E_{0,22} \end{bmatrix} x, \quad x \in \text{dom } T = \{ x \in C([a, b], \mathbb{R}^m) : Dx \in C^s([a, b], \mathbb{R}^m) \}. \tag{26} \]

We set \( X_T := \text{dom } T \) and introduce the norm \( ||x|| = ||x||_\infty + ||Dx||_{H^s}, x \in X_T \), so that \( T : X_T \rightarrow C([a, b], \mathbb{R}^m) \) is bounded.

If \( E_{0,22} \) remains nonsingular, then it is evident that

\[ \text{im } T = C([a, b], \mathbb{R}^m), \quad \text{dim ker } T = s m_1, \]

i.e., \( T \) is also fredholm. Therefore the operator \( T \) is marked as index-1 operator independently of \( s \geq 1 \). The corresponding IVPs and BVPs are well-posed and can be treated even by standard polynomial collocation. By introducing new variables \( v_i = (DX)^i, \ i = 1, \ldots, s-1 \), one obtains an associated first-order formulation of the DAE in the variables \( v_1, \ldots, v_{s-1}, x \), which has the same index 1.

If \( E_{0,22} \) vanishes identically, but \( E_{s-1,21}E_{0,12} \) remains nonsingular, then

\[ \text{im } T = \{ g \in C([a, b], \mathbb{R}^{m_1+m_2}) : E_{0,12}(E_{s-1,21}E_{0,12})^{-1}g_2 \in C^s([a, b], \mathbb{R}^{m_1}) \}, \]

and \( T \) is marked as index-2 operator independently of \( s \geq 1 \). Although \( \text{im } T \) is a nonclosed subset in \( C([a, b], \mathbb{R}^m) \), utilizing the special problem structure and incorporating special projections, corresponding BVPs with accurately stated boundary conditions can be directly treated by COLDAE, see [3]. Again, substituting the derivatives \( (DX)^{(i)} \) by new variables in the DAE leads to a first-order index-2 system which can be solved by overdetermined least-squares collocation, too.

If \( E_{0,22} \) as well as \( E_{1,21}, \ldots, E_{s-1,21} \) vanish identically, but \( E_{0,21}E_{0,12} \) remains nonsingular, then the operator \( T \) has index \( s + 1 \), and

\[ \text{im } T = \{ g \in C([a, b], \mathbb{R}^{m_1+m_2}) : E_{0,12}(E_{0,21}E_{0,12})^{-1}g_2 \in C^s([a, b], \mathbb{R}^{m_1}) \}. \]

Also here, substituting the derivatives \( (DX)^{(i)} \) by new variables leads to an index-(\( s + 1 \)) first-order DAE, and the latter can be treated by overdetermined least-squares collocation.
The index notion used in [3] is the perturbation index $\mu \in \mathbb{N}$ of a suitable first-order formulation. In the operator context we consider the extension $T$ of $\hat{T} : \text{dom}\, \hat{T} \subset C([a, b], \mathbb{R}^m) \to C([a, b], \mathbb{R}^m)$, turn then to the bounded operator $T : X_T \to C([a, b], \mathbb{R}^m)$ and show that the elements $g$ of $\text{im}\, T$ are involved therein together with parts of derivatives up to order $\mu - 1$, and $\mu$ is the smallest such number.

In contrast to the standard form (24), the version with properly involved derivatives can be seen as source of a reasonable first-order formulation. We conjecture that this idea applies also to further classes of DAEs. To emphasize the capabilities of this idea we use [15, Example] for a demonstration.

**Example 5.1.** We refactorize the leading term of the second order operator $\hat{T}$ given by

\[
\hat{T}x = \begin{bmatrix} 1 & t + 1 & t \\ t & t^2 + t \end{bmatrix} x'' + \begin{bmatrix} 0 & 2 \\ 0 & 2t \end{bmatrix} x' + \begin{bmatrix} 1 & t \\ 1 + t & t^2 + t + 1 \end{bmatrix} x
\]

by

\[
E_2 = \begin{bmatrix} 1 & t + 1 \\ t & t^2 + t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & t + 1 \end{bmatrix} =: A_2 D, \quad Dx'' = (Dx)'' - 2D'x'.
\]

The resulting operator $T$ reads

\[
Tx = A_2(Dx)'' + E_0x = \begin{bmatrix} 1 \\ t \end{bmatrix} (Dx)'' + \begin{bmatrix} 1 \\ 1 + t \end{bmatrix} \begin{bmatrix} t \\ t^2 + t + 1 \end{bmatrix} x,
\]

\[
X_T = \{ x \in C([a, b], \mathbb{R}^2) : Dx \in C^2([a, b], \mathbb{R}) \},
\]

\[
\text{im}\, T = \{ g \in C([a, b], \mathbb{R}^2) : [-t] g \in C^2([a, b], \mathbb{R}) \}.
\]

Introducing the new variable $v = (Dx)' = x_1' + (t + 1)x_2' + x_2$ leads to a first-order formulation with index $\mu = 3$.

A completely different approach to operators (24) and the corresponding DAEs is proposed in [6] in the context of a ring of operators acting on $C^m([a, b], \mathbb{R}^m)$, among them arbitrary-order differential operators of the form (24) with real-analytic coefficients. This serves as background of the following index notion by means of left regularizing operators

**Definition 5.2.** Each operator $Ly = \sum_{i=0}^k F_i y^{(i)}$, $y \in C^k([a, b], \mathbb{R}^m)$, such that the superposition $L \circ \hat{T}$ is a regular ODE of the same order as $\hat{T}$, i.e.,

\[
(L \circ \hat{T})x = \sum_{i=0}^s \bar{E}_i x^{(i)}, \quad x \in C^{s+k}([a, b], \mathbb{R}^m), \quad \bar{E}_s \text{ nonsingular},
\]

is called left regularizing operator of $\hat{T}$. The smallest possible $k$ is said to be the index of $\hat{T}$.  

---

3Not surprisingly, the classical procedure of turning a higher-order ODE into a first-order system applied to a DAE increases the differentiation index and leads to different solvability results and smoothness requirements. For details and examples we refer to [15].

4In [15] merely an index-4 first-order formulation was obtained with $v = x_1' + (t + 1)x_2'$.

5Note that the usual semi-norm family defining the topology of $C^m$ is much too strong to measure practically relevant approximation errors, e.g., [14, Subsection 2.4.2].
Here we denote the index in the sense of Definition 5.2 by $\mu_C$. The construction of left regularizers is closely related to the evaluation of derivative arrays, see [6]. For $s = 1$, constructing a left regularizer is equivalent to providing a so-called completion ODE ([5]) and $\mu_C$ equals the differentiation index. In turn, for regular first-order DAEs the differentiation index equals the perturbation index and also the tractability index as well.

However, for $s > 1$ things are completely different, so that $\mu_C$ is no longer helpful in view of the practical treatment of higher-order DAEs. The following two simple examples allow a first insight. Both examples have the form (25) resp. (26). We compare the perturbation index $\mu$ applied in [3] and $\mu_C$.

**Example 5.3** ($\mu = 1$, $\mu_C = s$).

$$\hat{T} x = \begin{bmatrix} x_1^{(s)} + x_1 \\ x_2 \\ x_2^{(s)} + x_1 \end{bmatrix}, \quad L = \begin{bmatrix} y_1 \\ y_2^{(s)} \end{bmatrix}, \quad (L \circ \hat{T}) x = \begin{bmatrix} x_1^{(s)} + x_1 \\ x_2^{(s)} \\ x_2^{(s)} + x_1 \end{bmatrix},$$

**Example 5.4** ($\mu = s + 1$, $\mu_C = 2s$).

$$\hat{T} x = \begin{bmatrix} x_1^{(s)} + x_3 \\ x_2^{(s)} \\ x_1^{(s)} + x_1 \end{bmatrix}, \quad L = \begin{bmatrix} y_3^{(s)} \\ y_2^{(s)} \\ y_1^{(s)} - y_3^{(2s)} \end{bmatrix}, \quad (L \circ \hat{T}) x = \begin{bmatrix} x_1^{(s)} \\ x_2^{(s)} \\ x_2^{(s)} + x_1 \end{bmatrix},$$

We finish by mentioning that in [15, 16] the given DAE is transformed via derivative arrays to a so-called strangeness-free mixed-order system which then can be handled by standard methods. Under additional quite special conditions, such a system is provided by evaluating involved matrix polynomials in [4].

6. Conclusions

We have explored properties of regular first-order DA operators and their finite-dimensional counterparts associated with the polynomial overdetermined least-squares collocation and provided a new convergence result.

We notice substantial progress in view of the consolidation of the polynomial overdetermined least-squares collocation. Nevertheless, there are essential open questions, e.g., concerning the choice of $N$ and $M$.

Furthermore we have surveyed corresponding results concerning higher-order operators and the direct treatment of higher-order DAEs.

References

[1] U. Ascher and G. Bader. A new basis implementation for a mixed order boundary-value ode solver. *SIAM J. Sci. Statist. Comput.*, 8:483–500, 1987.

[2] U. Ascher and R. Spiteri. Collocation software for boundary-value differential-algebraic equations. *SIAM J. Sci. Comput.*, 15:938–952, 1994.
[3] U.M. Ascher and L.R. Petzold. Projected collocation for higher-order higher-index differential-algebraic equations. *J. Comput. Appl. Math.*, 43:243–259, 1992.

[4] M.V. Bulatov and M.-G. Lee. Application of matrix polynomials to the analysis of linear differential-algebraic equations of higher order. *Differential Equations*, 44(10):1353–1360, 2008.

[5] S.L. Campbell. The numerical solution of higher index linear time varying singular systems of differential equations. *SIAM J. Sci. Statist. Comput.*, 6(2):334–348, 1985.

[6] V.F. Chistyakov. *Algebro-differential operators with a finite-dimensional kernel*. Nauka, Novosibirsk, 1996. (In Russian).

[7] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 40 of *Classics in Applied Mathematics*. SIAM, 2002.

[8] M. Hanke and R. März. A least-squares collocation method for nonlinear higher-index differential-algebraic equations. In preparation, 2018.

[9] M. Hanke, R. März, and C. Tischendorf. Least-squares collocation for higher-index linear differential-algebraic equations: Estimating the stability threshold. *Math. Comp.*, 2018. In print.

[10] M. Hanke, R. März, C. Tischendorf, E. Weinmüller, and S. Wurm. Least-squares collocation for linear higher-index differential-algebraic equations. *J. Comput. Appl. Math.*, 317:403–431, 2017. http://dx.doi.org/10.1016/j.cam.2016.12.017.

[11] G. Kitzhoffer, O. Koch, G. Pulverer, Ch. Simon, and E. Weinmüller. BVPSUITE, a new MATLAB solver for singular implicit boundary value problems. ASC Report 35, Vienna University of Technology, 2009.

[12] R. Lamour, R. März, and C. Tischendorf. *Differential-Algebraic Equations: A Projector Based Analysis*. Differential-Algebraic Equations Forum. Springer-Verlag Berlin Heidelberg New York Dordrecht London, 2013. Series Editors: A. Ilchmann, T. Reis.

[13] R. Lamour, R. März, and E. Weinmüller. *Surveys in Differential-Algebraic Equations III*, chapter Boundary-Value Problems for Differential-Algebraic Equations: A Survey, pages 177–309. Differential-Algebraic Equations Forum. Springer Heidelberg, 2015. ed. by A. Ilchmann and T. Reis.

[14] R. März. *Surveys in Differential-Algebraic Equations II*, chapter Differential-Algebraic Equations from a Functional-Analytic Viewpoint: A Survey, pages 163–285. Differential-Algebraic Equations Forum. Springer Heidelberg, 2015. ed. by A. Ilchmann and T. Reis.

[15] V. Mehrmann and C. Shi. Transformation of high order linear differential-algebraic systems to first order. *Numerical Algorithms*, 42:281–307, 2006.
[16] L. Wunderlich. Numerical treatment of second order differential-algebraic systems. In Proc. Appl. Math. Mech., volume 6, pages 775–776, 2006.