ASC Report No. 20/2007

Defect-based A-posteriori Error Estimation for Index-1 DAEs

Winfried Auzinger, Herbert Lehner, Ewa Weinmüller

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-00-1

Most recent ASC Reports

19/2007 *Georg Kitzhofer, Othmar Koch, Ewa Weinmüller*
Pathfollowing for Essentially Singular Boundary Value Problems with Application to the Complex Ginzburg-Landau Equation

18/2007 *Svatoslav Staněk, Gernot Pulverer, Ewa B. Weinmüller*
Analysis and Numerical Simulation of Positive and Dead Core Solutions of Singular Two-point Boundary Value Problems

17/2007 *Thorsten Sickenberger, Ewa Weinmüller, Renate Winkler*
Local Error Estimates for Moderately Smooth Problems: Part II – SDEs and SDAEs with Small Noise

16/2007 *Thorsten Sickenberger, Ewa Weinmüller, Renate Winkler*
Local Error Estimates for Moderately Smooth ODEs and DAEs

15/2007 *Markus Brunk, Ansgar Jüngel*
Simulation of Thermal Effects in Optoelectronic Devices Using Coupled Energy-transport and Circuit Models

14/2007 *Jose Antonio Carrillo, Maria Pia Gualdani, Ansgar Jüngel*
Convergence of an Entropic Semi-discretization for Nonlinear Fokker-Planck Equations in R^d

13/2007 *Ansgar Jüngel, Stefan Krause, Paola Pietra*
A Hierarchy of Diffuse Higher-order Moment Equations for Semiconductors

12/2007 *Markus Brunk, Ansgar Jüngel*
Numerical Coupling of Electric Circuit Equations and Energy-transport Models for Semiconductors

11/2007 *Markus Brunk, Ansgar Jüngel*
Numerical Simulation of Thermal Effects in Coupled Optoelectronic Device-circuit Systems

10/2007 *Ansgar Jüngel, Ingrid Violet*
First-Order Entropies for the Derrida-Lebowitz-Speer-Spohn Equation

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-00-1

ASC TU WIEN

Defect-based a posteriori error estimation for index-1 DAEs

W. Auzinger[*], H. Lehner[†], E. Weinmüller[‡§]

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstrasse 8–10/101
A-1040 Wien, Austria

Abstract

A computationally efficient a posteriori error estimator is introduced and analyzed for collocation solutions to linear index-1 DAEs with properly stated leading term. The procedure is based on a modified defect correction principle, extending an established technique from the ODE context to the DAE case. We prove that the resulting error estimate is asymptotically correct, and illustrate the method by means of a numerical example.

To keep the presentation reasonably self-contained, we also briefly review some arguments from the literature on DAEs concerning the decoupling of the problem and its discretization, which is essential for our analysis.

Subject Classification (AMS): 65L80, 65B05.

Key words: Differential algebraic equations; collocation; a posteriori error estimation; defect correction.

1 Introduction

We consider linear systems of DAEs of index 1,

$$\boldsymbol{A}(t)(\boldsymbol{D}(t)\boldsymbol{x}(t))' + \boldsymbol{B}(t)\boldsymbol{x}(t) = \boldsymbol{g}(t), \quad t \in [a,b], \tag{1.1}$$

with appropriately smooth data $\boldsymbol{A}(t) \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{n}}$, $\boldsymbol{D}(t) \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{m}}$, $\boldsymbol{B}(t) \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{m}}$. In our analysis, we assume that (1.1) is well-posed as an initial value problem. Furthermore we assume

$$\boldsymbol{m} > \boldsymbol{n}, \quad \text{and} \quad \ker \boldsymbol{A}(t) = \{\boldsymbol{0}\}, \quad \operatorname{im} \boldsymbol{D}(t) = \mathbb{R}^{\boldsymbol{n}}, \quad t \in [a,b]. \tag{1.2}$$

Remark. (1.2) implies that $(\boldsymbol{AD})(t) \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{m}}$ is singular, with $\operatorname{rank}(\boldsymbol{AD})(t) \equiv \boldsymbol{n}$. The assumptions can be weakened in a way where $\boldsymbol{A}(t)$ and $\boldsymbol{D}(t)$ themselves do not need to have full, but constant rank, replacing the requirement in (1.2) by $\ker \boldsymbol{A}(t) \oplus \operatorname{im} \boldsymbol{D}(t) = \mathbb{R}^{\boldsymbol{n}}$, cf. e.g. [14]. However, we restrict our discussion to the standard case (1.2). □

The above assumptions mean that the system (1.1) is *properly stated*, cf. e.g. [6, 7, 8, 11, 13, 14], which implies the existence of a unique, natural decoupling into an *inherent ODE* and associated algebraic components.

[*] w.auzinger@tuwien.ac.at
[†] h.lehner@tele2.at
[‡] e.weinmueller@tuwien.ac.at
[§] Supported by the Austrian Science Fund Project P 17253

We assume that the system has *tractability index 1*, with a regular inherent ODE; see Section 5 for the technical details of this index definition.

In this paper, the restriction to linear problems is reasonable because it enables a precise discussion of the underlying ideas with moderate technical effort. In particular, the focus is on the effective design and analysis of an asymptotically correct a posteriori error estimator for collocation solutions to (1.1), with a uniform, 'black box' treatment of the differential and algebraic components and an appropriate handling of the case where $\boldsymbol{D}(t)$ is not constant. The generalization of the method and its analysis for nonlinear and/or higher index problems as well as problems with a singular inherent ODE will be considered separately.

Introducing $\boldsymbol{u}(t) = \boldsymbol{D}(t)\boldsymbol{x}(t)$ as a separate variable, we obtain the dilated system

$$\begin{aligned} \boldsymbol{A}(t)\boldsymbol{u}'(t) + \boldsymbol{B}(t)\boldsymbol{x}(t) &= \boldsymbol{g}(t), \\ \boldsymbol{u}(t) - \boldsymbol{D}(t)\boldsymbol{x}(t) &= \boldsymbol{0}, \end{aligned} \tag{1.3}$$

which is equivalent to (1.1). In this formulation, it is obvious what initial or boundary conditions lead to a well-posed problem. For the IVP case, only $\boldsymbol{x}(0)$ can be prescribed, while $\boldsymbol{u}(0)$ is uniquely fixed by the algebraic relation $\boldsymbol{u}(0) = \boldsymbol{D}(0)\boldsymbol{x}(0)$. Note that (1.3) can again be written in the original form (1.1),

$$\hat{\boldsymbol{A}}(t)(\hat{\boldsymbol{D}}(t)\hat{\boldsymbol{x}}(t))' + \hat{\boldsymbol{B}}(t)\hat{\boldsymbol{x}}(t) = \hat{\boldsymbol{g}}(t), \quad t \in [a,b], \tag{1.4}$$

with

$$\begin{aligned} \hat{\boldsymbol{A}}(t) &= \begin{pmatrix} \boldsymbol{A}(t) \\ \boldsymbol{0}_{n\times n} \end{pmatrix}, \quad \hat{\boldsymbol{D}}(t) = \begin{pmatrix} \boldsymbol{0}_{n\times m} & \boldsymbol{I}_{n\times n} \end{pmatrix}, \\ \hat{\boldsymbol{B}}(t) &= \begin{pmatrix} \boldsymbol{B}(t) & \boldsymbol{0}_{m\times n} \\ \boldsymbol{D}(t) & -\boldsymbol{I}_{n\times n} \end{pmatrix}, \quad \hat{\boldsymbol{g}}(t) = \begin{pmatrix} \boldsymbol{g}(t) \\ \boldsymbol{0}_n \end{pmatrix}, \quad \hat{\boldsymbol{x}}(t) = \begin{pmatrix} \boldsymbol{x}(t) \\ \boldsymbol{u}(t) \end{pmatrix}, \end{aligned} \tag{1.5}$$

where, in particular, $\hat{\boldsymbol{D}} := \hat{\boldsymbol{D}}(t) \in \mathbb{R}^{\hat{\boldsymbol{n}}\times\hat{\boldsymbol{m}}}$ is now a *constant* matrix, $\hat{\boldsymbol{A}}(t) \in \mathbb{R}^{\hat{\boldsymbol{m}}\times\hat{\boldsymbol{n}}}, \hat{\boldsymbol{B}}(t) \in \mathbb{R}^{\hat{\boldsymbol{m}}\times\hat{\boldsymbol{m}}}, \hat{\boldsymbol{m}} = \boldsymbol{m}+\boldsymbol{n}$, $\hat{\boldsymbol{n}} = \boldsymbol{n}$. The analogue of (1.2),

$$\hat{\boldsymbol{m}} > \hat{\boldsymbol{n}}, \quad \text{and} \quad \ker \hat{\boldsymbol{A}}(t) = \{\boldsymbol{0}\}, \quad \operatorname{im} \hat{\boldsymbol{D}}(t) = \mathbb{R}^{\hat{\boldsymbol{n}}} \tag{1.6}$$

is also satisfied.

In the sequel, we will use the symbols A, B, D, x, g, m, n as a generic denotation for either $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{x}, \boldsymbol{g}, \boldsymbol{m}, \boldsymbol{n}$ or $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{D}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{g}}, \hat{\boldsymbol{m}}, \hat{\boldsymbol{n}}$ wherever the respective statements apply to both sets of variables, under the assumption that D is constant. The same convention applies to further introduced symbols like G, Q, M, p. Applying this convention, we may write (1.1) or (1.4), respectively, as

$$A(t)(Dx(t))' + B(t)x(t) = g(t), \quad t \in [a,b], \quad \text{with} \quad D = \text{const.} \tag{1.7}$$

In other words: (1.7) represents the original system (1.1) if $\boldsymbol{D}(t) = \boldsymbol{D} = \text{const}$. In general, (1.7) is to be identified with (1.4) where $\hat{\boldsymbol{D}}(t) = \hat{\boldsymbol{D}} = \text{const.}$ by construction. Actually, this latter property is the motivation for rewriting the system in the form (1.4), especially in view of favorable application of collocation methods, see Sections 2 and 5 for a further discussion. For $\boldsymbol{D}(t) = \text{const.}$, the two interpretations of (1.7) are equivalent, also in respect to numerical approximation.

For later reference, let $x^*(t)$ denote the exact solution of (1.7).

2 Collocation methods

We consider collocation methods applied to the system (1.7), with $D = \text{const}$. The collocation solution is a continuous piecewise polynomial function $p(t)$ of degree $\le s$ which satisfies the given initial condition and obeys, in a pointwise sense, the DAE (1.7) at the collocation points

$$t_{ij} := \tau_i + c_j h_i, \tag{2.1}$$

where

$$\begin{gathered}\tau_0 = a, \quad h_i := \tau_{i+1} - \tau_i > 0, \quad \tau_N = b, \\ 0 < c_1 < \ldots c_j < \ldots < c_s = 1,\end{gathered} \tag{2.2}$$

for $i = 0 \ldots N-1,\ j = 1 \ldots s$. Note, in particular, that $c_s = 1$ is essential for our analysis. This means that the method is *stiffly accurate*, an property which in a natural way ensures stability of the scheme, cf. e.g. [8] for a more detailed discussion.

We also assume that s is even, which will be necessary to guarantee the asymptotic correctness of our error estimator to be defined in Section 4. We also denote $h := \max_{i=0\ldots N-1} h_i$.

The complete system of collocation equations reads

$$A(t_{ij})(Dp)'(t_{ij}) + B(t_{ij})p(t_{ij}) = g(t_{ij}), \quad i = 0 \ldots N-1,\ j = 1 \ldots s, \tag{2.3}$$

where $p(t)$ is represented by a polynomial $p_i(t)$ of degree $\leq s$ on each subinterval $[\tau_i, \tau_{i+1}]$, together with the continuity relations

$$p_{i-1}(\tau_i) = p_i(\tau_i), \quad i = 0 \ldots N-1. \tag{2.4}$$

For later use we also define

$$c_0 := 0, \qquad t_{i0} := \tau_i,\ \ i = 0 \ldots N-1. \tag{2.5}$$

Remark. If we identify (1.7) with the DAE system rewritten in dilated form (1.3), the collocation conditions (2.3) are equivalent to

$$\begin{aligned}\boldsymbol{A}(t_{ij})\boldsymbol{q}'(t_{ij}) + \boldsymbol{B}(t_{ij})\boldsymbol{p}(t_{ij}) &= \boldsymbol{g}(t_{ij}), \\ \boldsymbol{q}(t_{ij}) - \boldsymbol{D}(t_{ij})\boldsymbol{p}(t_{ij}) &= \boldsymbol{0},\end{aligned} \tag{2.6}$$

together with the continuity conditions for $\boldsymbol{p}$ and $\boldsymbol{q}$ as in (2.4), where $p(t) = \hat{\boldsymbol{p}}(t) = (\boldsymbol{p}(t), \boldsymbol{q}(t))^T$, and $\boldsymbol{p}(t)$ and $\boldsymbol{q}(t)$ are piecewise polynomial approximations for $\boldsymbol{x}(t)$ and $\boldsymbol{u}(t)$, respectively. Note that for $\boldsymbol{D}(t) \not\equiv$ const. we have $\boldsymbol{D}(t)\boldsymbol{p}_i(t) \not\equiv \boldsymbol{q}_i(t)$, and therefore the collocation scheme (2.6) cannot be interpreted as collocation applied directly to (1.1) because, in general,

$$\boldsymbol{A}(t_{ij})(\boldsymbol{D}\boldsymbol{p}_i)'(t_{ij}) + \boldsymbol{B}(t_{ij})\boldsymbol{p}(t_{ij}) \neq \boldsymbol{g}(t_{ij}). \tag{2.7}$$

On the other hand, (2.6) can be interpreted as as a well-established implicit Runge-Kutta (IRK) method for (1.1), see Appendix A. Therefore, the convergence results from [7] for IRK methods apply. These results are based on a natural decoupling of the problem based on its properly stated formulation and an analogous decoupling of the discrete scheme. The analysis of our error estimator in Section 5 will also be essentially based on such a decoupling procedure. □

In the sequel we occasionally use the subscript denotation $[\cdots]_{ij}$ for any scalar, vector or matrix valued function as a shortcut for evaluation at the grid point t_{ij}.

3 Defect-based error estimation

A posteriori error estimation in ODEs based on the defect correction principle is an old idea originally due to Zadunaisky [16] and further developed by Stetter [15]. In the context of regular and singular ODEs, this approach was refined and analyzed in [3, 4] and implemented in [2]. In particular, for a special realization of the defect, an efficient, asymptotically correct error estimator, the QDeC estimator, was designed in [3] for collocation solutions on arbitrary grids. These ideas will now be extended to the DAE context. As will be seen below, this is not straightforward because of the coupling between differential and algebraic components. A detailed analysis of the proposed estimator is given for the linear index-1 case (1.1).

In abstract notation, the basic structure of a defect-based error estimator can be described as follows: Consider a numerical solution $x_\Delta \approx x^*$ for a problem

$$F(x(t)) = 0, \quad t \in [a, b], \tag{3.1}$$

on a grid Δ. Define the *defect* $d = d(t)$ by interpolating x_Δ by a continuous piecewise polynomial function $p(t)$ of degree $\leq s$ and substituting $p(t)$ into (3.1),

$$d(t) := F(p(t)), \quad t \in [a, b]. \tag{3.2}$$

Obviously, $p(t)$ is the exact solution to a *neighboring problem*

$$F(x(t)) = d(t) \tag{3.3}$$

related to the original problem (3.1). Now use a procedure of low effort (typically a low order scheme), the so-called *auxiliary scheme* $\tilde{F}$, to obtain approximate discrete solutions $\tilde{x}_\Delta$ and $\tilde{x}_\Delta^{\mathrm{def}}$ for both the original and neighboring problems on the grid Δ, i.e. $\tilde{F}(\tilde{x}_\Delta) = 0$ and $\tilde{F}(\tilde{x}_\Delta^{\mathrm{def}}) = d_\Delta$, where d_Δ is an appropriate restriction of $d(t)$ to the grid Δ.

Since (3.1) and (3.3) differ only by the (presumably) small defect d, we expect that

$$\epsilon_\Delta := \tilde{x}_\Delta^{\mathrm{def}} - \tilde{x}_\Delta \tag{3.4}$$

is a good estimate for the global error

$$e := x_\Delta - x^*. \tag{3.5}$$

In other terms,

$$\begin{aligned} e_\Delta = x_\Delta - x_\Delta^* &\approx F^{-1}(d) - F^{-1}(0) \\ &\approx \tilde{F}^{-1}(d_\Delta) - \tilde{F}^{-1}(0) = \tilde{x}_\Delta^{\mathrm{def}} - \tilde{x}_\Delta = \epsilon_\Delta. \end{aligned} \tag{3.6}$$

This is exactly the procedure originally proposed in [15]. However, in concrete applications, the auxiliary scheme $\tilde{F}$ and a suitable representation for the defect d_Δ have to be carefully chosen. In particular, in [3] collocation for the ODE case was considered. For $\tilde{F}$ chosen as the backward Euler scheme, it was shown that a modified version of the pointwise defect (3.2) has to be used in order to obtain an asymptotically correct estimator for the error of a given collocation approximation $p(t)$ playing the role of x_Δ. In the following section this approach (the 'QDeC estimator') is described in more detail and will be extended to the DAE case.

For linear problems, the procedure can be realized in a simpler way, where ϵ_Δ is computed by a single application of the auxiliary scheme,

$$\tilde{F}\epsilon_\Delta = d_\Delta. \tag{3.7}$$

This version will be used in Section 4; see also [5] for a further discussion.

4 The QDeC estimator for DAEs

Now we apply the procedure described in Section 3 to the linear DAE (1.7). In addition to the collocation method introduced in Section 2, we use a scheme of backward Euler type over the collocation nodes as an auxiliary method. Let $h_{ij} := t_{ij} - t_{i,j-1}$ and consider the grid function ϵ_{ij} satisfying the auxiliary scheme

$$A(t_{ij})\,\frac{D\epsilon_{ij} - D\epsilon_{i,j-1}}{h_{ij}} + B(t_{ij})\epsilon_{ij} = \bar{d}_{ij}, \tag{4.1}$$

with homogeneous initial condition $\epsilon_{0,0} = 0$. This is nothing but (3.7), where the backward Euler scheme plays the role of $\tilde{F}$.

According to (3.2), the straightforward, classical way to define the defect $\bar{d}_{ij}$ would be to substitute $p(t)$ into (1.7) in the pointwise sense,

$$d(t) := A(t)(Dp)'(t) + B(t)p(t) - g(t), \quad t \in [a,b], \tag{4.2}$$

and using the pointwise defect $\bar{d}_{ij} := d(t_{ij})$ in (4.1). However, as has been pointed out in [3] in the ODE context, this procedure does not lead to successful results. For collocation this is obvious: Since, by definition of the collocation solution (2.3), the defect $d(t_{ij})$ vanishes at each point t_{ij} ($i = 0 \dots N-1$, $j = 1 \dots s$) which enters the backward Euler scheme, the error estimate $\epsilon(t_{ij})$ would always be zero.

In slight variation of the procedure introduced in [3] we now define a modified defect via the integral means

$$\bar{d}_{ij} := \sum_{k=0}^{s} \alpha_{jk}\, d(t_{ik}) = \frac{1}{h_{ij}} \int_{t_{i,j-1}}^{t_{ij}} d(t)dt + \mathcal{O}(h^{s+1}), \tag{4.3}$$

for $i = 0 \dots N-1,\ j = 1 \dots s$, where the α_{jk} are quadrature coefficients for the integral means in (4.3), i.e.

$$\alpha_{jk} = \frac{1}{c_j - c_{j-1}} \int_{c_{j-1}}^{c_j} L_k(t)dt, \quad j = 1 \dots s,\ k = 0 \dots s, \tag{4.4}$$

with the Lagrange polynomials L_k of degree s, such that $L_k(c_j) = \delta_{jk}$. Note that, in contrast to collocation at s nodes in each subinterval excluding the left endpoint $t_{i0} = \tau_i + c_0 h_i$, we now include the additional node $c_0 = 0$ (see (2.5)) for the polynomial quadrature defining (4.3).

Remark. As explained in [3], the modified defect (4.3) is related to the residual of the p_{ij} w.r.t. a higher order Runge-Kutta scheme of collocation type including the nodes t_{i0}. Originally, the motivation for defining the $\bar{d}_{ij}$ in this way comes from the ODE context, see [3]. Actually, due to (2.3) the sum in (4.3) reduces to one term,

$$\bar{d}_{ij} = \alpha_{j0}\, d(t_{i0}), \tag{4.5}$$

at least if we ignore numerical errors in the computation of $p(t)$.

In the DAE case, where the algebraic relations are satisfied exactly at the collocation nodes, see (2.6), such a weighting of pointwise defect values seems to be suspicious at first sight. However, in the following theorem, which extends the results from [3, 4], we show that the outcome is an asymptotically correct error estimate also in the DAE case. The essential observation is that a separate handling of differential and algebraic system components is not necessary, which lets us expect that the procedure will also be successfully applicable to problems with nonlinear coupling. □

We are now in the position to state our main result; the proof is given in Section 5.

Theorem 1. *While the global error of the collocation method* (2.3) *is of order* h^s*, i.e.*

$$e(t) = p(t) - x^*(t) = \mathcal{O}(h^s), \tag{4.6}$$

the error estimate of the global error (4.6) *based on the modified defect* (4.3) *and the auxiliary scheme* (4.1) *is asymptotically correct, i.e.*

$$\epsilon_{ij} - e(t_{ij}) = \mathcal{O}(h^{s+1}). \tag{4.7}$$

5 Analysis of the error estimator

The analysis given below is based on a decoupling argument and an associated inherent ODE according to the ideas from [6, 7, 8] et al. Note that the assumption $D = \text{const.}$ is important here: It guarantees the problem to be *numerically well-formulated*, i.e. the original problem and the discrete schemes decouple in a parallel way, which is essential to ensure stable integration (cf. e.g. [8]).

In particular, concerning the stability and convergence of the collocation scheme (2.3) and the auxiliary scheme (4.1), we may simply refer to the respective results from the DAE literature, e.g. [6, 7, 8]: For the numerically well-formulated problem at hand together with the stiffly accurate schemes, these arguments are straightforward and can be directly based on conventional stability arguments in the ODE sense, applied to our problem after appropriate decoupling. In particular, the last step of the proof is based on the stability of the backward Euler scheme, which is is obvious and standard in the present context.

The following proof of Theorem 1, i.e. of (4.7), is organized in several steps. In the first step of the proof we also specify the precise assumptions concerning the index of the DAE considered, namely that it has tractability index 1. For the respective technical details required from theory of DAEs, we follow [14] and references therein, explicating details only where appropriate. In the last two steps, the use of the modified defect (4.3) is seen to be essential; the argument does not remain valid if $\bar{d}_{ij}$ is replaced by $d(t_{ij})$.

Proof of Theorem 1.

- *Decoupling of the index-1 DAE.*

  Let $Q \in \mathbb{R}^{m\times m}$ be any projector onto $\ker D$ and define

  $$G(t) := A(t)D + B(t)Q \in \mathbb{R}^{m\times m}. \tag{5.1}$$

  We assume that the system (1.7), with $D = \text{const.}$, satisfies assumptions as in (1.2). Furthermore we assume that the system has *tractability index 1*, which, by definition means that $G(t)$ is invertible (in contrast to $A(t)D$).

  Let $D^- \in \mathbb{R}^{m\times n}$ be a generalized reflexive inverse of D (cf. [11, 14, 17]) such that $D^-D = I_{m\times m} - Q$ and $DD^- = I_{n\times n}$. Furthermore we denote

  $$M(t) := G^{-1}(t)B(t)D^- \in \mathbb{R}^{m\times n}. \tag{5.2}$$

  With these definitions, we have

  $$\begin{pmatrix} DG^{-1} \\ QG^{-1} \end{pmatrix} \cdot \begin{pmatrix} AD & B & -I_{m\times m} \end{pmatrix} = \begin{pmatrix} D & DMD & -DG^{-1} \\ 0 & Q+QMD & -QG^{-1} \end{pmatrix}. \tag{5.3}$$

  Applying this identity decouples (1.7) into a pure ODE for $Dx(t)$ and a pure algebraic equation expressing $Qx(t)$ in terms of $Dx(t)$,

  $$(Dx)'(t) + DM(t)Dx(t) = DG^{-1}(t)g(t), \tag{5.4a}$$
  $$Qx(t) + QM(t)Dx(t) = QG^{-1}(t)g(t). \tag{5.4b}$$

  The solution of (1.7) can then be represented as

  $$x(t) = (I_{m\times m} - Q)x(t) + Qx(t) = D^-(Dx)(t) + (Qx)(t). \tag{5.5}$$

  Note that Dx and Qx can be seen as separate variables, as Q projects onto $\ker D$.

Remark. In this decoupling procedure, Qx plays the role of the algebraic solution component. The projector Q is in a sense arbitrary, but this is not essential since the definition of the tractability index and the coefficient matrices DM, DG^{-1}, $Q+QMD$ and QG^{-1} do not actually depend on the choice of Q. In particular, the *inherent ODE* (5.4a) for the differential component Dx is uniquely determined by the problem data.

This decoupling transformation does not necessarily require the data function $D(t)$ to be constant, it requires only $\operatorname{im} D(t)$ to be constant, cf. [7, 8]. Due to (1.2), this is the case for $\boldsymbol{D}(t)$. If $D(t)$ actually depends on t, so may $Q = Q(t)$. □

Remark. Under assumption (1.2), $\hat{\boldsymbol{G}}(t) = \hat{\boldsymbol{A}}(t)\hat{\boldsymbol{D}} + \hat{\boldsymbol{B}}(t)\hat{\boldsymbol{Q}}$ is invertible iff $\boldsymbol{G}(t) = \boldsymbol{A}(t)\boldsymbol{D} + \boldsymbol{B}(t)\boldsymbol{Q}$ is invertible, which can be seen as follows: Since the particular choice for $\hat{\boldsymbol{Q}}$ is irrelevant, we consider the natural orthogonal projector

$$\hat{\boldsymbol{Q}} = \begin{pmatrix} \boldsymbol{I_{m\times m}} & \boldsymbol{0_{m\times n}} \\ \boldsymbol{0_{n\times m}} & \boldsymbol{0_{n\times n}} \end{pmatrix}, \quad \text{yielding} \quad \hat{\boldsymbol{G}} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{A} \\ \boldsymbol{D} & \boldsymbol{0_{n\times n}} \end{pmatrix}. \tag{5.6}$$

Now, making use of the identities $\operatorname{im} \boldsymbol{Q} = \ker \boldsymbol{D}$, $\boldsymbol{D}^-\boldsymbol{D} = \boldsymbol{I_{m\times m}} - \boldsymbol{Q}$ and $\boldsymbol{D}\boldsymbol{D}^- = \boldsymbol{I_{n\times n}}$, we observe

$$\hat{\boldsymbol{G}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{Q} & \boldsymbol{D}^- \\ \boldsymbol{D} & \boldsymbol{0_{n\times n}} \end{pmatrix}}_{=:\,\hat{\boldsymbol{T}}} = \underbrace{\begin{pmatrix} \boldsymbol{G} & \boldsymbol{B}\boldsymbol{D}^- \\ \boldsymbol{0_{n\times m}} & \boldsymbol{I_{n\times n}} \end{pmatrix}}_{=:\,\hat{\boldsymbol{H}}}, \tag{5.7}$$

where $\hat{\boldsymbol{T}}$ is invertible (note that $\hat{\boldsymbol{T}}^2 = \hat{\boldsymbol{I}}$). This shows $\operatorname{rank} \hat{\boldsymbol{G}} = \operatorname{rank} \hat{\boldsymbol{H}}$; furthermore, $\hat{\boldsymbol{H}}$ is invertible iff $\boldsymbol{G}$ is invertible. □

- *Decoupling of the discrete equations; a priori estimate for $De(t)$ and $Qe(t)$.*

  The same decoupling transformation (5.3) which led us to (5.4) can be applied to the collocation equations (2.3), yielding

$$Dp'(t_{ij}) + DM(t_{ij})Dp(t_{ij}) = DG^{-1}(t_{ij})g(t_{ij}), \tag{5.8a}$$
$$Qp(t_{ij}) + QM(t_{ij})Dp(t_{ij}) = QG^{-1}(t_{ij})g(t_{ij}). \tag{5.8b}$$

  for $i = 0 \ldots N-1,\ j = 1 \ldots s$. Together with (5.4), which is satisfied by $x = x^*$, we also have

$$De'(t_{ij}) + DM(t_{ij})De(t_{ij}) = 0, \tag{5.9a}$$
$$Qe(t_{ij}) + QM(t_{ij})De(t_{ij}) = 0, \tag{5.9b}$$

  for $i = 0 \ldots N-1,\ j = 1 \ldots s$, for the global error $e(t) = p(t) - x^*(t)$ of the collocation solution. The convergence order of the error in the differential component and its derivative,

$$De(t) = Dp(t) - Dx^*(t) = \mathcal{O}(h^s), \quad De'(t) = Dp'(t) - Dx^{*\prime}(t) = \mathcal{O}(h^s), \tag{5.10}$$

  follow from the general stability and convergence theory of collocation for ODEs, see [1].

  Equation (5.8b) says that for each i, the polynomial $Qp_i(t)$ of degree $\le s$ interpolates $QG^{-1}(t)g(t) - QM(t)Dp(t)$ at t_{ij} for $j = 1 \ldots s$, and interpolation at $t_{i0} = \tau_i = t_{i-1,s}$ follows from the continuity condition (2.4). From the standard estimate for the polynomial interpolation error we have

$$Qp(t) + QM(t)Dp(t) - QG(t)^{-1}g(t) = \mathcal{O}(h^s). \tag{5.11}$$

Subtracting (5.4b) from (5.11) and using (5.10) yields

$$Qe(t) = -QM(t)De(t) + \mathcal{O}(h^s) = \mathcal{O}(h^s). \tag{5.12}$$

The decoupling transformation (5.3) applies to the auxiliary scheme (4.1) as well, resulting in

$$\begin{aligned} \frac{D\epsilon_{ij} - D\epsilon_{i,j-1}}{h_{ij}} + DM(t_{ij})D\epsilon_{ij} &= DG^{-1}(t_{ij})\bar{d}_{ij}, \\ Q\epsilon_{ij} + QM(t_{ij})D\epsilon_{ij} &= QG^{-1}(t_{ij})\bar{d}_{ij}. \end{aligned} \tag{5.13}$$

- *Decoupling of the defect.*

  Decoupling the defect definition (4.2) according to (5.3) leads to

$$DG^{-1}(t)d(t) = Dp'(t) + DM(t)Dp(t) - DG^{-1}(t)g(t), \tag{5.14a}$$
$$QG^{-1}(t)d(t) = Qp(t) + QM(t)Dp(t) - QG^{-1}(t)g(t), \tag{5.14b}$$

  and subtracting (5.4) (with $x = x^*$) from (5.14) yields the following representation for the components of the pointwise defect:

$$DG^{-1}(t)d(t) = De'(t) + DM(t)De(t), \tag{5.15a}$$
$$QG^{-1}(t)d(t) = Qe(t) + QM(t)De(t). \tag{5.15b}$$

  Note that due to (5.8) we have $DG^{-1}(t_{ij})d(t_{ij}) = 0$ and $QG^{-1}(t_{ij})d(t_{ij}) = 0$ for $i = 0 \dots N-1$, $j = 1 \dots s$. Moreover, $QG^{-1}(t_{i0})d(t_{i0}) = 0$ (cf. the remark following (5.10)) – this is the point in the proof where assumption $c_s = 1$ is essential.

  The analogous identities for the modified defect (4.3) read

$$DG^{-1}(t_{ij})\bar{d}_{ij} = DG^{-1}(t_{ij}) \sum_{k=0}^{s} \alpha_{jk}\, d(t_{ik}), \tag{5.16a}$$
$$QG^{-1}(t_{ij})\bar{d}_{ij} = QG^{-1}(t_{ij}) \sum_{k=0}^{s} \alpha_{jk}\, d(t_{ik}). \tag{5.16b}$$

- *Difference equation for the error $\delta_{ij} := \epsilon_{ij} - e(t_{ij})$ of the error estimate.*

  Now we express the difference quotients of the exact and estimated errors in terms of quadratures. Integration of $De'(t)$ using (5.15a) gives

$$\begin{aligned} \frac{De_{ij} - De_{i,j-1}}{h_{ij}} &= \frac{1}{h_{ij}} \int_{t_{i,j-1}}^{t_{ij}} De'(t)dt \\ &= \frac{1}{h_{ij}} \int_{t_{i,j-1}}^{t_{ij}} (-DM(t)De(t) + DG^{-1}(t)d(t))dt \\ &= \sum_{k=0}^{s} \alpha_{jk} \left(-DM(t_{ik})De(t_{ik}) + DG^{-1}(t_{ik})d(t_{ik})\right) + \mathcal{O}(h^{s+1}). \end{aligned} \tag{5.17}$$

  Furthermore, substituting (5.16) into (5.13) shows

$$\frac{D\epsilon_{ij} - D\epsilon_{i,j-1}}{h_{ij}} + DM(t_{ij})D\epsilon_{ij} = DG^{-1}(t_{ij}) \sum_{k=0}^{s} \alpha_{jk}\, d(t_{ik}). \tag{5.18}$$

Subtracting (5.17) from (5.18) we obtain a system of difference equations of backward Euler type, together with homogeneous initial conditions, for the error $\delta_{ij} := \epsilon_{ij} - e(t_{ij})$ of the error estimate:

$$\frac{D\delta_{ij} - D\delta_{i,j-1}}{h_{ij}} + DM(t_{ij})D\delta_{ij} = \sum_{k=0}^{s} \alpha_{jk} \left(DM(t_{ik})De(t_{ik}) - DM(t_{ij})De(t_{ij}) \right) \tag{5.19a}$$

$$- \sum_{k=0}^{s} \alpha_{jk} \left(DG^{-1}(t_{ik}) - DG^{-1}(t_{ij}) \right) d(t_{ik}) + \mathcal{O}(h^{s+1}). \tag{5.19b}$$

On the right hand side of (5.19a) we have used $\sum_{k=0}^{s} \alpha_{jk} = 1$.

- *Estimation of δ_{ij}.*

  Given the stability of the backward Euler scheme, it now remains to estimate the inhomogeneities in (5.19a), (5.19b) by $\mathcal{O}(h^{s+1})$ to prove that $D\delta_{ij} = \mathcal{O}(h^{s+1})$. For (5.19a), this estimate follows by Taylor expansion of $DM(t)De(t)$ about $t = t_{ij}$ together with $DM(t) = \mathcal{O}(h)$, $\frac{d}{dt}DM(t) = \mathcal{O}(h)$ and $De(t) = \mathcal{O}(h^s)$, $De'(t) = \mathcal{O}(h^s)$. For (5.19b) we find $DG^{-1}(t_{ij}) - DG^{-1}(t_{ik}) = \mathcal{O}(h)$, and $d(t_{ik}) = \mathcal{O}(h^s)$, where the latter estimate is established in the following way: ¿From (5.15a),

  $$DG^{-1}(t_{ij})d(t_{ij}) = \mathcal{O}(h^s), \tag{5.20}$$

  while from (5.8b) and (5.14b),

  $$QG^{-1}(t_{ij})d(t_{ij}) = 0. \tag{5.21}$$

  Hence,

  $$\begin{aligned} d(t_{ij}) &= G(t_{ij})(G^{-1}(t_{ij})d(t_{ij})) \\ &= G(D^- DG^{-1}(t_{ij})d(t_{ij}) + QG^{-1}(t_{ij})d(t_{ij})) = \mathcal{O}(h^s). \end{aligned} \tag{5.22}$$

  So far we have shown that $D\delta_{ij} = \mathcal{O}(h^{s+1})$. To estimate δ_{ij}, we use (5.9) and (5.13),

  $$Q\delta_{ij} = -QMD\delta_{ij} + QG^{-1}(t_{ij})\bar{d}_{ij}. \tag{5.23}$$

  But

  $$\begin{aligned} QG^{-1}(t_{ij})\bar{d}_{ij} &= \sum_{k=0}^{s} \alpha_{jk}\, QG^{-1}(t_{ij})d(t_{ik}) \\ &= \sum_{k=0}^{s} \alpha_{jk} \left(QG^{-1}(t_{ij}) - QG^{-1}(t_{ik}) \right) d(t_{ik}) + \sum_{k=0}^{s} \alpha_{jk}\, QG^{-1}(t_{ik})d(t_{ik}), \end{aligned} \tag{5.24}$$

  where the first sum on the right hand side can be estimated as $\mathcal{O}(h^{s+1})$ analogously to (5.19b), and the second sum vanishes because all $QG^{-1}(t_{ik})d(t_{ik})$ vanish according to (5.21).

  Thus,

  $$\delta_{ij} = D^- D\delta_{ij} + Q\delta_{ij} = \mathcal{O}(h^{s+1}), \tag{5.25}$$

  which completes the proof of (4.7). □

6 Numerical example

We consider the initial value problem

$$\begin{pmatrix} \mathrm{e}^t \\ \mathrm{e}^t \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \end{pmatrix} x'(t) + \begin{pmatrix} \mathrm{e}^t(1+\cos^2 t) & \cos^2 t \\ \mathrm{e}^t(-1+\cos^2 t) & -\cos^2 t) \end{pmatrix} x(t) = \begin{pmatrix} \sin^2 t(1-\cos t) - \sin t \\ \sin^2 t(-1-\cos t) - \sin t \end{pmatrix}, \tag{6.1}$$

on $[a, b] = [0, 1]$ with initial condition $x(0) = (1, -1)^T$. We use a realization of our method in MATLAB, based on collocation at equidistant points with $s = 4$, on $N = 2, 4, 8, 16, 32$ subintervals of length $1/N$. In the following tables, the asymptotical order $\epsilon - e = \mathcal{O}(h^{s+1})$ is clearly visible; see also Figure 1.

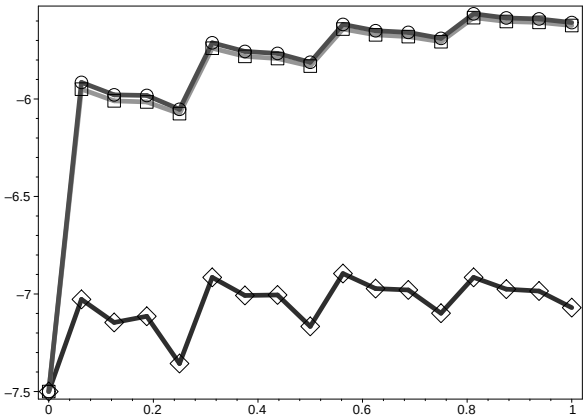

Figure 1: $\log_{10}$ - plot for first solution component, $N = 4$

○ ... error $|e_1(t)| = |p_1(t) - x_1^*(t)|$
□ ... error estimate $|\epsilon_1(t)|$
◇ ... error of error estimate $|\epsilon_1(t) - e_1(t)|$

- First solution component, at $t = 1$:

N	e	ord $_e$	$\epsilon - e$	ord $_{\epsilon-e}$
4	−2.466e−06	3.8	8.513e−08	4.6
8	−1.634e−07	3.9	2.989e−09	4.8
16	−1.051e−08	4.0	9.886e−11	4.9
32	−6.664e−10	4.0	3.180e−12	5.0

- First solution component, maximum absolute values over all collocation points $\in [0, 1]$:

N	e	ord $_e$	$\epsilon - e$	ord $_{\epsilon-e}$
4	2.732e−06	4.0	1.272e−07	5.3
8	1.711e−07	4.0	3.578e−09	5.2
16	1.074e−08	4.0	1.074e−10	5.1
32	6.734e−10	4.0	3.311e−12	5.0

- Second solution component, at $t = 1$:

N	e	ord $_e$	$\epsilon - e$	ord $_{\epsilon-e}$
4	2.906e−05	3.8	−7.927e−07	4.6
8	1.522e−06	3.9	−2.783e−08	4.8
16	9.788e−08	4.0	−9.206e−10	4.9
32	6.205e−09	4.0	−2.961e−12	5.0

A On the interrelation between collocation and Runge-Kutta methods

For ODEs, collocation methods are a special case of implicit Runge-Kutta (IRK) methods, cf. e.g. [10]. For the DAE case the analogous interrelation is explicated in the sequel.

For a linear DAE system (1.1), the following version of a Runge-Kutta method is described and analyzed in [7] (see also [8] for the nonlinear case): Assume that

$$\mathcal{A} = (a_{jk},\ j,k = 1 \dots s) \tag{A.1}$$

is the coefficient matrix in the Butcher array for a given s-stage IRK scheme, with internal nodes $c = (c_1, \dots, c_s)$ and weights $b = (b_1, \dots, b_s)$. As in [7] we assume that the method is stiffly accurate, i.e. $c_s = 1$ (as in (2.2)), $a_{sk} \equiv b_k$ and $\mathcal{A}$ invertible, with

$$\mathcal{A}^{-1} =: \hat{\mathcal{A}} = (\hat{a}_{jk},\ j,k = 1 \dots s)\,. \tag{A.2}$$

For $\tau_{i+1} = \tau_i + h_i$, let $t_{ij} := \tau_i + c_j h_i$ as in Section 2. As usual we denote $\boldsymbol{A}_{ij} := \boldsymbol{A}(t_{ij})$, etc. For notational convenience, let $h := h_i,\ c_0 := 0$ and $t_{i0} := \tau_i$. In [7], an IRK step for (1.1) relating $\boldsymbol{x}_{i+1}$ to $\boldsymbol{x}_i$ is defined via internal approximations stages $\boldsymbol{X}_{ij}$ satisfying the linear system

$$\boldsymbol{A}_{ij}\,\boldsymbol{U}'_{ij} + \boldsymbol{B}_{ij}\boldsymbol{X}_{ij} = \boldsymbol{g}_{ij}, \quad j = 1 \dots s, \tag{A.3}$$

for the unknowns $\boldsymbol{X}_{ij}$, $j = 1 \dots s$. Here, $\boldsymbol{U}'_{ij}$ is the shortcut

$$\boldsymbol{U}'_{ij} := \frac{1}{h}\sum_{k=1}^{s} \hat{a}_{jk}(\boldsymbol{D}_{ik}\boldsymbol{X}_{ik} - \boldsymbol{D}_{i0}\boldsymbol{X}_{i0}), \quad j = 1 \dots s. \tag{A.4}$$

Together with $\boldsymbol{X}_{i0} = \boldsymbol{x}_i$ and with $c_s = 1$, this defines $\boldsymbol{x}_{i+1} := \boldsymbol{X}_{is}$.

Now we additionally assume that the given IRK scheme is equivalent to a collocation scheme in the conventional ODE sense, and we shall demonstrate in which way (A.3), (A.4) can be interpreted as a collocation scheme applied to the given DAE. To this end, we observe that (A.4) can be written as

$$h(\mathcal{A} \otimes \boldsymbol{I}) \cdot \begin{pmatrix} \boldsymbol{U}'_{i1} \\ \vdots \\ \boldsymbol{U}'_{is} \end{pmatrix} = \begin{pmatrix} \boldsymbol{D}_{i1}\boldsymbol{X}_{i1} - \boldsymbol{D}_{i0}\boldsymbol{X}_{i0} \\ \vdots \\ \boldsymbol{D}_{is}\boldsymbol{X}_{is} - \boldsymbol{D}_{i0}\boldsymbol{X}_{i0} \end{pmatrix}, \tag{A.5}$$

or equivalently

$$\boldsymbol{D}_{ij}\boldsymbol{X}_{ij} = \boldsymbol{D}_{i0}\boldsymbol{X}_{i0} + h\sum_{k=1}^{s} a_{jk}\boldsymbol{U}'_{ik}, \quad j = 1 \dots s, \tag{A.6}$$

which has a structure similar to a conventional IRK system.

For IRK schemes of collocation type, the coefficients a_{jk} satisfy the simplifying assumption $C(s)$ (cf. [10]),

$$\sum_{k=1}^{s} a_{jk}\, c_k^{\ell-1} = \frac{c_j^{\ell}}{\ell}, \quad j,\ell = 1 \dots s, \tag{A.7}$$

which implies

$$h\sum_{k=1}^{s} a_{jk}\, q(t_{ik}) = \int_{t_{i0}}^{t_{ij}} q(t)\,dt, \quad j = 1 \dots s, \tag{A.8}$$

for any polynomial $q(t)$ of degree $\leq s-1$. Now, let $\boldsymbol{p}_i(t)$ and $\boldsymbol{q}_i(t)$ denote the unique polynomials of degree $\leq s$ interpolating the $\boldsymbol{X}_{ij}$ and $\boldsymbol{D}_{ij}\boldsymbol{X}_{ij}$, respectively, i.e.,

$$\boldsymbol{p}_i(t_{ij}) = \boldsymbol{X}_{ij}, \quad \boldsymbol{q}_i(t_{ij}) = \boldsymbol{D}_{ij}\boldsymbol{X}_{ij}, \quad j = 0 \ldots s. \tag{A.9}$$

Then, due to (A.8) the following identities hold for $j = 1 \ldots s$:

$$\boldsymbol{D}_{ij}\boldsymbol{X}_{ij} = \boldsymbol{q}_i(t_{ij}) = \boldsymbol{q}_i(t_{i0}) + \int_{t_{i0}}^{t_{ij}} \boldsymbol{q}_i'(t)\,dt = \boldsymbol{q}_i(t_{i0}) + h\sum_{j=1}^{s} a_{jk}\,\boldsymbol{q}_i'(t_{ik}) = \boldsymbol{D}_{i0}\boldsymbol{X}_{i0} + h\sum_{k=1}^{s} a_{jk}\,\boldsymbol{q}_i'(t_{ik}), \tag{A.10}$$

or equivalently,

$$h(\mathcal{A}\otimes\boldsymbol{I})\cdot\begin{pmatrix}\boldsymbol{q}_i'(t_{i1})\\ \vdots\\ \boldsymbol{q}_i'(t_{is})\end{pmatrix} = \begin{pmatrix}\boldsymbol{D}_{i1}\boldsymbol{X}_{i1} - \boldsymbol{D}_{i0}\boldsymbol{X}_{i0}\\ \vdots\\ \boldsymbol{D}_{is}\boldsymbol{X}_{is} - \boldsymbol{D}_{i0}\boldsymbol{X}_{i0}\end{pmatrix}. \tag{A.11}$$

Since $\mathcal{A}$ has been assumed to be invertible, (A.5) and (A.11) have unique, identical solutions, i.e.,

$$\boldsymbol{q}_i'(t_{ij}) = \boldsymbol{U}_{ij}', \quad j = 1 \ldots s. \tag{A.12}$$

Now, (A.9) and (A.12) imply that solving the IRK system (A.3),(A.4) is equivalent to determining polynomials $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ of degree $\leq s$ which satisfy $\boldsymbol{p}_i(\tau_i) = \boldsymbol{x}_i,\ \boldsymbol{q}_i(\tau_i) = \boldsymbol{D}(\tau_i)\boldsymbol{x}_i$, and

$$\begin{aligned}\boldsymbol{A}(t_{ij})\boldsymbol{q}'(t_{ij}) + \boldsymbol{B}(t_{ij})\boldsymbol{p}(t_{ij}) &= \boldsymbol{g}(t_{ij}),\\ \boldsymbol{q}(t_{ij}) - \boldsymbol{D}(t_{ij})\boldsymbol{p}(t_{ij}) &= \boldsymbol{0},\end{aligned} \tag{A.13}$$

for $j = 1 \ldots s$, which is exactly (2.6). Starting from (A.13) we can also reverse the above argumentation, ending up with the IRK formulation (A.3),(A.4).

References

[1] U. M. Ascher, R. M. M. Mattheij, R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice Hall, 1988.

[2] W. Auzinger, G. Kneisl, O. Koch, E.Weinmüller, *SBVP 1.0 – A MATLAB solver for singular boundary value problems*, Technical Report ANUM Preprint No.2/02, Vienna University of Technology, 2002.

[3] W. Auzinger, O. Koch, E. Weinmüller, *Efficient collocation schemes for singular boundary value problems*, Numer. Algorithms 31, 5–25, 2002.

[4] W. Auzinger, O. Koch, E. Weinmüller, *Analysis of a new error estimate for collocation methods applied to singular boundary value problems*, SIAM J. Numer. Anal. 42, 2366–2386, 2005.

[5] W. Auzinger, O. Koch, D. Praetorius, E. Weinmüller, *New a posteriori error estimates for singular boundary value problems*, Numer. Algorithms 40, 79–100, 2005.

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

[7] I. Higueras, R. März, *Differential algebraic systems anew*, Appl. Numer. Math. 42, 315–335, 2002.

[8] I. Higueras, R. März, *Differential algebraic equations with properly stated leading term*, Comp. Math. Appl. 48, 215–235, 2004.

[9] I. Higueras, R. März, C. Tischendorf, *Stability preserving integration of index-1 DAEs*, Appl. Numer. Math. 45, 175–200, 2003.

[10] E. Hairer, G. Wanner, S. P. Nørsett, *Solving Ordinary Differential Equations I – Nonstiff Problems*, Springer Series in Computational Mathematics 8, 1987.

[11] P. Kunkel, V. Mehrmann, *Differential-Algebraic Equations – Analysis and Numerical Solution*, European Mathematical Society, 2006.

[12] O. Koch, R. März, D. Praetorius, E. Weinmüller, *Collocation for solving DAEs with singularities*, ASC Report 32/2007, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2007.

[13] R. März, *The index of linear differential algebraic equations with properly stated leading terms*, Results Math. 42, 308–338, 2002.

[14] S. Schulz, *Four Lectures on Differential-Algebraic Equations*, Report Series 497, Dept. of Mathematics, Univ. of Auckland, 2003.

[15] H. J. Stetter, *The defect correction principle and discretization methods*, Numer. Math. 29, 425–443, 1978.

[16] P. E. Zadunaisky, *On the estimation of errors propagated in the numerical integration of ODEs*, Numer. Math. 27, 21–39, 1976.

[17] G. Zielke, *Motivation und Darstellung von verallgemeinerten Matrixinversen*, Beiträge zur Numerischen Mathematik 7, 177–218, 1979.