Convergence of Collocation Schemes for Boundary Value Problems in Nonlinear Index 1 DAEs with a Singular Point

A. Dick, O. Koch, R. März, and E. Weinmüller

January 11, 2021

Abstract

We analyze the convergence behavior of collocation schemes applied to approximate solutions of BVPs in nonlinear index 1 DAEs, which exhibit a critical point at the left boundary. Such a critical point of the DAE causes a singularity in the inherent nonlinear ODE system. In particular, we focus on the case when the inherent ODE system is singular with a *singularity of the first kind* and apply polynomial collocation to the *original DAE system*. We show that for a certain class of well-posed boundary value problems in DAEs having a sufficiently smooth solution, the global error of the collocation scheme converges uniformly with the so-called stage order. Due to the singularity, superconvergence at the mesh points does not hold in general. The theoretical results are supported by numerical experiments.

1 Introduction

Motivated by numerous important applications from physics [6, 7, 20, 36], chemistry [10, 34, 35], mechanics [9], ecology [27, 32], or economy [11, 13, 21], a lot of interest and effort has been put into the development of efficient numerical algorithms for the approximate solution of BVPs in explicit ODEs exhibiting singularities. Such problems are often given in the form

$$x'(t) = \frac{1}{t^\alpha} M(t)x(t) + h(x(t), t), \quad t \in (0, 1], \tag{1a}$$

$$B_0 x(0) + B_1 x(1) = \beta, \quad x \in C[0, 1], \tag{1b}$$

where $\alpha \geq 1$, x is an m–dimensional real function, M is a smooth $m \times m$ matrix function, h is a smooth function mapping into $\mathbb{R}^m$, and $\beta \in \mathbb{R}^p$, $B_0, B_1 \in \mathbb{R}^{p \times m}$. For $\alpha = 1$ the problem is called singular with a singularity of the first kind, for $\alpha > 1$ it is essentially singular (singularity of the second kind).
Research activities in related fields, like the computation of connecting orbits in dynamical systems [33], or singular Sturm-Liouville problems [4, 12], also benefit from new findings for problems of the form (1). Both the analysis and the numerical treatment of this problem class considerably differ from the regular case [16, 17, 18, 19, 25].

Here, we focus on the case $\alpha = 1$. Depending on the spectrum of $M(0)$, we generally encounter unbounded contributions to the solution manifold, such that $x \in C(0, 1]$. However, irrespective of the eigenvalues of $M(0)$, by posing proper homogeneous initial conditions equivalent to $M(0)x(0) = 0$, we can ensure $x \in C[0, 1]$. The system is then augmented by two-point boundary conditions to define a locally unique solution. If all the eigenvalues of $M(0)$ have negative real parts or are equal to zero, the boundary value problem can equivalently be posed as an initial value problem, making a shooting approach possible for both theoretical and practical purposes. We focus in our analysis exactly on the above subclass of BVPs which is important for many applications. For this subclass the smoothness of x depends only on the smoothness of the problem data. For $M \in C^{p+1}$ and $h \in C^p$, the solution satisfies $x \in C^{p+1}[0, 1]$. Conversely, if $M(0)$

has eigenvalues with positive real parts along with eigenvalues with negative real parts, then there is no equivalent initial value problem and in this case a shooting approach is not viable. For more details see [16] and the discussion of (23) and (24) in Section 2.

The first attempt to extend collocation techniques developed in the context of singular explicit ODEs to DAEs was discussed for linear index 1 DAEs in [26]. Our decision to use polynomial collocation was motivated by its advantageous convergence properties for (1), while in the presence of a singularity other high order methods show order reductions and become inefficient. Convergence of collocation schemes applied to solve (1) with a singularity of the first kind and spectrum of $M(0)$ defining the subclass we specified above for linear and nonlinear BVPs was analyzed in [17] and [2], respectively. These results have been generalized to arbitrary spectrum of $M(0)$ in [25]. It turns out that for k general interior collocation points, both the uniform convergence order and the order at the mesh points is equal to k. For the distribution of the collocation points such that superconvergence is observed for regular problems with smooth solutions, e.g. Gaussian points or an odd number of equidistant nodes, for problems with a singularity of the first kind the convergence order generally is $k+1$, both uniformly in t and at the mesh points. The usual high-order superconvergence at the mesh points does not hold in general for singular problems [17].

The open domain MATLAB code `bvpsuite` has been designed to solve general implicit systems of ODEs which may have arbitrary order including zero. In particular, algebraic constraints are permitted and therefore, DAEs are in the scope of the code. In [24, 26] numerical experiments and comparisons with existing software can be found. We stress that in the present paper, we only use `bvpsuite` executed on uniform grids in order to illustrate the convergence order of the involved collocation schemes. We do not attempt to compare the available software for DAE systems here. As a matter of fact, `bvpsuite` has been used to work out the related conjectures before proving them. We recall that the dependable performance of the code is ensured by a strict analysis only for singular problems of the form (1) with $\alpha \leq 1$. For $\alpha > 1$ the convergence theory of collocation applied to (1) is an open and extremely challenging question. The additional difficulty in the case of DAEs is due to their *implicit nature*.

Much progress has been made concerning DAE theory and applications, but there are still many questions left open. In particular, the numerical treatment of critical points and singularities is just emerging. Encouraged by the positive results for the linear case [26], we approach here singular nonlinear index 1 DAE systems.

DAEs *with properly stated leading term* were introduced and studied for example in [5, 14, 30]. This enables a proper and natural description of the involved solution derivatives. To this end, one considers DAEs written in the form

$$f((D(t)x(t))', x(t), t) = 0, \quad t \in [a, b]. \tag{2}$$

One of the advantages of this precise description of the problem structure is that there exists an *inherent explicit regular ODE* uniquely determined by the problem data [14, 15]. Under mild assumptions, DAEs in standard form can be reformulated to have properly stated leading terms. For DAEs with properly stated leading terms arising in applications, see [14].

In [31], *linear DAEs* with properly stated leading term and type 0-critical points as well as type $1A$-critical points have been analyzed. This means that after decoupling the system using the matrix chain technique developed in [31] into the differential and algebraic components, the related inherent ODE exhibits a singularity of the first or second kind. The singularities discussed here are the counterparts to the 0-critical points for *nonlinear DAEs*.

Recall that according to [26], for linear systems of DAEs with a singularity of the first kind and appropriately smooth problem data, the stage order of the collocation scheme is retained, provided that the so-called canonical projector remains bounded. This means that the global error of the collocation scheme with k

collocation points is $O(h^k)$ uniformly in t, where h denotes the uniform mesh width. We observe order reductions if the canonical projector becomes unbounded. In this article we will formulate the respective convergence results for BVPs in nonlinear index 1 DAEs with a singularity of the first kind and bounded canonical projector.

Clearly, the convergence results for polynomial collocation at interior collocation points presented here also hold for index 1 DAEs without singularities. We stress that we aim to analyze known collocation schemes applied *directly to the DAE system in its original formulation.* For the proposed scheme no pre-handling is necessary. The projection decoupling is used merely in the analysis, but not in context of the numerical schemes. In contrast to our approach a completely different strategy has been proposed in [28], where a certain transformation and index reduction procedure is incorporated into the collocation scheme. During this procedure the derivative free equations are separated from those including solution derivatives and then Gaussian type schemes are applied to the dynamical part of the system and Lobatto type collocation is used for the algebraic constraints. The strategy can be applied to regular DAE systems of arbitrary index — the higher the index, the more smoothness is required. Moreover, the schemes are complex and require adaptation measures in the course of computation. For a further discussion on numerical approaches to *regular DAEs*, see [26] and [29].

As stressed in [28], constant rank conditions (see Hypothesis 2.1 in [28]) are of central importance for the proposed method. These conditions do not hold in presence of a singularity. At present we are also not aware of a constructive way to treat *singular higher index DAEs.* As shown here, the above class of singular index 1 DAEs can be successfully treated by transparent standard collocation schemes applied to the original DAE formulation — without any transformation or reduction. However, the analysis of the problem is far from being trivial. In this context the commutativity of the discretization and the decoupling is crucial (which is not valid when one separates the collocation schemes, as it is done in [28]).

The paper is structured in the following way. In Section 2, we describe the problem setting and show how the analytical system can be decoupled into the differential and algebraic components. The problem data is given in such a way that the inherent ODE exhibits a singularity of the first kind at $t = 0$. Collocation methods are introduced and their *convergence behavior at collocation points* is analyzed in Section 3, while their *uniform convergence* is discussed in Section 4. The analytical results are illustrated by numerical experiments in Section 5. In Section 6, we construct a special class of quasi-linear DAE systems. Here, the aim is to precisely specify *sufficient* conditions in terms of the original problem data leading to the decoupled system from Section 2.

2 Problem Specification and Analytical Results

In this section we recapitulate the analytical results for the nonlinear boundary value problem for a system of DAEs given in the following form:

$$f((D(t)x(t))', x(t), t) = 0, \quad t \in (0, 1], \tag{3a}$$

$$B_0 D(0)x(0) + B_1 D(1)x(1) = \beta, \tag{3b}$$

where $f(y, x, t) \in \mathbb{R}^m$, $D(t) \in \mathbb{R}^{n\times m}$, $y \in \mathbb{R}^n$, $x \in \mathcal{D}$, with $\mathcal{D} \subseteq \mathbb{R}^m$ open, $t \in [0, 1]$, $n \leq m$. The data f, f_y, f_x, D are assumed to be at least continuous on their definition domains. Moreover, we require that

$$\ker f_y(y, x, t) = 0, \quad (y, x, t) \in \mathbb{R}^n \times \mathcal{D} \times (0, 1], \tag{4}$$

$$\mathcal{R}(D(t)) = \mathbb{R}^n, \quad t \in [0, 1]. \tag{5}$$

Conditions (4) and (5) mean that the matrix $D(t)$ has constant full row rank n on the closed interval while $f_y(y, x, t)$ has full column rank n on $\mathbb{R}^n \times \mathcal{D} \times (0, 1]$, respectively. At $t = 0$ the matrix $f_y(y, x, t)$ may undergo

a rank drop. The structure (4) and (5) means that the system (3a) has a properly stated leading term on $\mathbb{R}^n \times \mathcal{D} \times (0,1]$, cf. [14]. We consider solutions in the function space

$$C_D^1([0,1],\mathbb{R}^m) := \{x \in C([0,1],\mathbb{R}^m) : Dx \in C^1([0,1],\mathbb{R}^n)\}.$$

This setting includes classical singular boundary value problems of the form (1), where $\alpha = 1$, $m = n$, $D(t) = I$, $f(y,x,t) = ty - M(t)x - th(x,t)$. In this paper, we are interested in $n < m$.
The structure of the boundary conditions given in (3b) which are necessary and sufficient for (3) to be well-posed, will be specified later.

We now define

$$N_0(t) := \ker D(t), \quad t \in [0,1], \tag{6}$$

and note that owing to the properties of the leading term, cf. (4), (5),

$$\ker f_y(y,x,t)D(t) = N_0(t), \quad (y,x,t) \in \mathbb{R}^n \times \mathcal{D} \times (0,1].$$

Let us denote by Q_0 a continuous pointwise projector function onto $\ker D$, $Q_0(t)^2 = Q_0(t)$, $\mathcal{R}(Q_0(t)) = \ker D(t)$, $t \in [0,1]$, and let $P_0(t) := I - Q_0(t)$. Here, we could choose $P_0(t)$ and $Q_0(t)$ to be orthogonal projectors. We point out, however, that the choice of the projectors does not matter for the following investigations.

Moreover, let us define

$$G_0(y,x,t) := f_y(y,x,t)D(t), \quad (y,x,t) \in \mathbb{R}^n \times \mathcal{D} \times [0,1], \tag{7}$$
$$G_1(y,x,t) := G_0(y,x,t) + f_x(y,x,t)Q_0(t), \quad (y,x,t) \in \mathbb{R}^n \times \mathcal{D} \times [0,1]. \tag{8}$$

In the following we discuss DAEs (3a) which are regular with tractability index 1 on $\mathbb{R}^n \times \mathcal{D} \times (0,1]$. Consequently, $G_1(y,x,t)$ is nonsingular on $\mathbb{R}^n \times \mathcal{D} \times (0,1]$, see [14, 15]. However, we permit a singular behavior of $G_1(y,x,t)$ for $t \to 0$, causing a singularity of the first kind in the associated *inherent ODE*. To this end, we assume that

$$tG_1(y,x,t)^{-1} \tag{9}$$

has a continuous extension on $\mathbb{R}^n \times \mathcal{D} \times [0,1]$.
We introduce finally the pointwise generalized inverse D^- of D uniquely defined by the following requirements:

$$D^-DD^- = D^-, \;\; DD^-D = D, \;\; DD^- = I, \;\; D^-D = P_0 \tag{10}$$

which need to hold pointwise in $[0,1]$. Note that D^- is also continuous on $[0,1]$. For an illustration of the above structural properties of the DAE system, in particular for the involved projection functions, see the example in Section 5.

It is well known that properties of the linearized problem are crucial in the analysis of the nonlinear setting. Therefore, we assume that a solution $x_\star \in C_D^1([0,1],\mathbb{R}^m)$ of (3) exists and introduce the linearization of (3a) as

$$A_\star(t)(D(t)z(t))' + B_\star(t)z(t) = 0, \quad t \in (0,1], \tag{11}$$

where

$$A_\star(t) := f_y((D(t)x_\star(t))', x_\star(t), t), \;\; B_\star(t) := f_x((D(t)x_\star(t))', x_\star(t), t), \quad t \in [0,1].$$

Since the matrix

$$G_{\star 1}(t) := A_\star(t)D(t) + B_\star(t)Q_0(t) = G_1((D(t)x_\star(t))', x_\star(t), t)$$

is nonsingular for $t \in (0,1]$, the linear DAE (11) is regular with tractability index 1 on the interval $(0,1]$. It was demonstrated in [5] that with the above assumptions the solutions of the DAE (11), see also [26], can be decoupled on $(0,1]$ into the *differential components* Dz and the *algebraic components* $Q_0 z$. While Dz satisfies the explicit inherent ODE,

$$(D(t)z(t))' + \underbrace{D(t)G_{\star 1}(t)^{-1}B_\star(t)D(t)^-}_{=:-\frac{1}{t}M_\star(t)} D(t)z(t) = 0, \quad t \in (0,1], \tag{12}$$

the algebraic components are given by

$$Q_0(t)z(t) = -Q_0(t)G_{\star 1}(t)^{-1}B_\star(t)D(t)^- D(t)z(t), \tag{13}$$

and the solutions of (11) can be expressed as

$$z(t) = D(t)^- D(t)z(t) + Q_0(t)z(t) = (I - Q_0(t)G_{\star 1}(t)^{-1}B_\star(t))D(t)^- D(t)z(t), \quad t \in (0,1]. \tag{14}$$

We note that $M_\star$ is continuous on $[0,1]$.

Analogously to the theory of explicit ODEs [22], we say that the solution $x_\star$ of boundary value problem (3) is *isolated* if and only if the linearization of (3),

$$A_\star(t)(D(t)z(t))' + B_\star(t)z(t) = 0, \quad t \in (0,1], \tag{15a}$$
$$B_0 D(0)z(0) + B_1 D(1)z(1) = 0, \tag{15b}$$

has only the trivial solution. In this case the boundary value problem (3) is said to be *well-posed*.

We now turn back to the nonlinear DAE (3a) and decouple it into the inherent ODE and the algebraic constraints. We first introduce the notation

$$u_\star(t) := D(t)x_\star(t), \;\; w_\star(t) := Q_0(t)x_\star(t) + D(t)^-(D(t)x_\star(t))', \quad t \in [0,1]$$

as well as the function

$$F(w,u,t) := f(D(t)w, D(t)^- u + Q_0(t)w, t),$$

where $w \in \mathbb{R}^m$, $u \in \mathbb{R}^n$, $t \in (0,1]$ are such that $D(t)^- u + Q_0(t)w \in \mathcal{D}$.

We observe that $F(w_\star(t), u_\star(t), t) = 0$, $t \in [0,1]$ and

$$F_w(w_\star(t), u_\star(t), t) = f_y((D(t)x_\star(t))', x_\star(t), t)D(t) + f_x((D(t)x_\star(t))', x_\star(t), t)Q_0(t) = G_{\star 1}(t), \quad t \in [0,1].$$

$G_{\star 1}(t)$ is nonsingular for $t > 0$. Consequently, it follows from the Implicit Function Theorem that for $t \in (0,1]$, the equation $F(w,u,t) = 0$ is locally equivalent to $w = \omega(u,t)$, where the function ω together with its partial derivative ω_u are continuous and

$$\omega(u_\star(t), t) = w_\star(t), \;\; \omega_u(u_\star(t), t) = -G_{\star 1}(t)^{-1}B_\star(t)D(t)^-, \quad t \in (0,1], \tag{16}$$

for details see [14].

Remark: Note that for the linear case the function ω can be easily specified. We have

$$f(y,x,t) := A(t)y + B(t)x - g(t)$$

and

$$G_1(t) = A(t)D(t) + B(t)Q_0(t).$$

Consequently,

$$F(w,u,t) = A(t)D(t)w + B(t)(D(t)^{-}u + Q_0(t)w(t)) - g(t) = G_1(t)w + B(t)D(t)^{-}u - g(t),$$

and hence

$$\omega(u,t) = -G_1(t)^{-1}(B(t)D(t)^{-}u - g(t)), \quad u \in \mathbb{R}^n, \quad t \in (0,1].$$

We now use our so-called *decoupling function* $\omega : \mathcal{D}_\omega \times (0,1] \to \mathbb{R}^m$, where $\mathcal{D}_\omega \subseteq \mathbb{R}^n$ is an open set, to specify the inherent ODE associated with the nonlinear DAE (3a). Let $x(\cdot)$ be any solution of the DAE (3a) defined on $(0,1]$ and let us define

$$u(t) := D(t)x(t), \;\; w(t) := Q_0(t)x(t) + D(t)^{-}(D(t)x(t))',$$

such that

$$D(t)w(t) = (D(t)x(t))', \;\; Q_0(t)w(t) = Q_0(t)x(t), \;\; P_0(t)x(t) = D(t)^{-}D(t)x(t) = D(t)^{-}u(t), \tag{17}$$

and

$$x(t) = D(t)^{-}u(t) + Q_0(t)w(t), \quad t \in (0,1]. \tag{18}$$

With the above notation, the original DAE (3a) can be rewritten as

$$f(D(t)w(t), D(t)^{-}u(t) + Q_0(t)w(t), t) = 0.$$

In case that the domain $\mathcal{D}_\omega$ is sufficiently large, the solution can be represented in the following form:

$$x(t) = D(t)^{-}u(t) + Q_0(t)\omega(u(t),t), \quad t \in (0,1], \tag{19}$$

where u satisfies the inherent ODE,

$$u'(t) = D(t)\omega(u(t),t), \quad t \in (0,1]. \tag{20}$$

In order to apply the standard analysis for singular boundary value problems, cf. [16] and [25], we assume that the decoupling function ω satisfies

$$D(t)\omega(u,t) = \frac{1}{t}M(t)u + g(u,t), \quad u \in \mathcal{D}_\omega, \quad t \in (0,1], \tag{21}$$

where the $n \times n$ matrix function M and the function g are appropriately smooth on $[0,1]$ and $\mathcal{D}_\omega \times [0,1]$, respectively. Later on, in Section 6, for a large class of quasi-linear DAEs, we derive sufficient conditions for (21) to hold in terms of the original data of the DAE system.
The inherent ODE (20) is now augmented by the boundary conditions (3b),

$$B_0u(0) + B_1u(1) = \beta. \tag{22}$$

This yields the following boundary value problem:

$$u'(t) = \frac{1}{t}M(t)u(t) + g(u(t),t), \quad t \in (0,1], \tag{23a}$$

$$B_0u(0) + B_1u(1) = \beta. \tag{23b}$$

The linearization of the above boundary value problem reads, cf. (12),

$$\zeta'(t) = D(t)\omega_u(u_\star(t),t)\zeta(t) = \frac{1}{t}M_\star(t)\zeta(t), \quad t \in (0,1], \tag{24a}$$

$$B_0\zeta(0) + B_1\zeta(1) = 0. \tag{24b}$$

Note that the linear boundary value problem (15), for a system of DAEs, has only the trivial solution exactly when this is the case for the related boundary value problem for the inherent ODE (24), due to the solution representation (14). This means that the nonlinear boundary value problem (3), for a system of DAEs, is well-posed exactly when this holds for the related boundary value problem for the nonlinear inherent ODE (23). Hence, we now specify the necessary and sufficient conditions for the linear ODE problem (24) to have only the trivial solution. First of all, it is clear that for the matrices B_0, B_1 in (24b) and (3b), $B_0, B_1 \in \mathbb{R}^{n\times m}$ has to be true. It was shown in [16] that the form of the boundary conditions (24b) which guarantee that (24) has only the trivial solution depends on the spectral properties of the coefficient matrix $M_\star(0)$. Note that (21) implies

$$M_\star(t) = M(t) + tg_u(u_\star(t), t), \quad t \in (0, 1],$$

and therefore

$$M_\star(0) = M(0).$$

To avoid fundamental modes of (24a) which have the form $\cos(\sigma \ln(t)) + \mathrm{i}\sin(\sigma \ln(t))$, we assume that zero is the only eigenvalue of $M(0)$ on the imaginary axis.
Now, let R_+ denote the projection onto the invariant subspace which is associated with eigenvalues of $M(0)$ which have strictly positive real parts. Let R be a projection onto the kernel of $M(0)$. Finally, define

$$U := R_+ + R, \quad V := I - U, \tag{25}$$

where I denotes the identity matrix in $\mathbb{R}^n$. In [16] it was shown that the boundary value problem (24) is well-posed if and only if the boundary conditions (24b) can equivalently be written as

$$V\zeta(0) = 0, \tag{26a}$$
$$R_+\zeta(1) = 0, \tag{26b}$$
$$R\zeta(0) = 0, \text{ or } R\zeta(1) = 0. \tag{26c}$$

The homogeneous initial conditions specified in (26a) are necessary and sufficient for $\zeta \in C[0, 1]$. As explained before, in the further analysis we restrict ourselves to the case where $R_+ = 0$.

In the DAE analysis, the so called canonical projector along $\ker D(t)$, see [15],

$$P_{can}(y, x, t) := I - Q_0(t)G_1(y, x, t)^{-1} f_x(y, x, t)$$

plays an important role. In particular, the linear homogeneous BVP for a system of DAEs (15) has the solution

$$z(t) = P_{\star\, can}(t)D(t)^{-}\zeta(t),$$

where ζ solves the BVP (24), and

$$P_{\star\, can}(t) := P_{can}((D(t)x_\star(t))', x_\star(t), t).$$

By definition, problem (3) is well-posed if z vanishes identically. In turn, z vanishes identically, if ζ does. Later, when dealing with the convergence of collocation approximations, we require the canonical projector to have a continuous extension for $t \to 0$ to avoid respective order reductions, see [26].
We now take a closer look at the canonical projector function and characterize a class of DAEs equipped with a canonical projector remaining bounded for $t \to 0$. To this end, we introduce the subspace

$$S(y, x, t) := \{z \in \mathbb{R}^m : f_x(y, x, t)z \in \mathcal{R}(f_y(y, x, t))\} = \ker W(y, x, t)f_x(y, x, t), \quad y \in \mathbb{R}^n,\ x \in \mathcal{D},\ t \in [0, 1],$$

where $W(y, x, t)$ denotes the orthogonal projector onto $\mathcal{R}(f_y(y, x, t))^\perp$. Since $f_y(y, x, t)$ has constant rank n for $t > 0$, $W(y, x, t)$ depends continuously on its arguments for $t > 0$. Often, in the given DAE the derivative

free equations are separated, cf. (35), and therefore, we simply have $W(y,x,t) = \mathrm{diag}(I,0)$. In general, we assume that W has a continuous extension W^{ext} for $t \to 0$, such that for $t > 0$

$$W^{ext}(y,x,t) = W(y,x,t).$$

We stress that, due to a rank drop of $f_y(y,x,t)$ at $t = 0$, in general $W^{ext}(y,x,0) \neq W(y,x,0)$, but $W^{ext}(y,x,0)f_y(y,x,0) = 0$. The projector function W^{ext} has the constant rank $m-n$. Owing to the given index 1 property, the subspaces S and N_0 are transversal for $t > 0$, see [15], i.e.

$$S(y,x,t) \oplus N_0(t) = \mathbb{R}^m, \quad y \in \mathbb{R}^n,\, x \in \mathcal{D},\, t \in (0,1],$$

and $P_{can}(y,x,t)$ projects onto $S(y,x,t)$ along N_0.

Let us now introduce

$$S^{ext}(y,x,t) := \ker W^{ext}(y,x,t)f_x(y,x,t), \quad y \in \mathbb{R}^n,\, x \in \mathcal{D},\, t \in [0,1],$$

such that for $t > 0$,

$$S^{ext}(y,x,t) \oplus N_0(t) = \mathbb{R}^m, \quad \operatorname{rank} W^{ext}(y,x,t)f_x(y,x,t) = m-n.$$

This transversality condition remains valid at $t = 0$, that is

$$S^{ext}(y,x,0) \oplus N_0(0) = \mathbb{R}^m, \tag{27}$$

if and only if $W^{ext}(y,x,0)f_x(y,x,0)$ has rank $m-n$ and the subspaces $S^{ext}(y,x,0)$ and $N_0(0)$ intersect trivially. Therefore, the condition

$$\operatorname{rank} \begin{pmatrix} W^{ext}(y,x,0)f_x(y,x,0) \\ D(0) \end{pmatrix} = m \tag{28}$$

indicates the decomposition (27). Let us assume that (27) holds, then the projector function P^{ext}_{can} onto S^{ext} along N_0 is uniquely determined and continuous for $t \geq 0$. Since P_{can} coincides with P^{ext}_{can} for $t > 0$, P_{can} remains bounded for $t \to 0$. Now, it is clear that condition (27) ensures the boundedness of the canonical projector function.

We stress that problem (3) may be well-posed with a bounded smooth solution, although the canonical projector is unbounded for $t \to 0$. The following example illustrates this fact, see [26] for details,

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} (x_1(t) - x_2(t))' + \begin{pmatrix} 2 & 0 \\ 0 & t+2 \end{pmatrix} x(t) = 0, \quad x_1(0) - x_2(0) = 0. \tag{29}$$

The inherent ODE of the above problem reads

$$u'(t) = -\frac{2t+4}{t}u(t), \quad u(t) = x_1(t) - x_2(t). \tag{30}$$

We have $M(0) = -4$, $R_+ = R = 0$, $V = 1$, and the boundary condition is equivalent to $u(0) = x_1(0) - x_2(0) = 0$. Therefore, problem (29) is well-posed. The general solution of the ODE (30) is $u(t) = ce^{-2t}t^{-4}$ which implies that only the trivial solution satisfies the condition $u(0) = 0$. This condition corresponds to (26a). All other solutions of the inherent ODE grow unboundedly for $t \to 0$. The canonical projector, see [26, Section 3.1],

$$P_{can}(t) = I - Q_{can}(t) = \begin{pmatrix} 1 + \frac{2}{t} & -\frac{2+t}{t} \\ \\ \frac{2}{t} & 1 - \frac{2+t}{t} \end{pmatrix}$$

is unbounded for $t \to 0$. Observe that

$$N_0(t) = \operatorname{span} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad W(t) = W^{ext}(t) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad S(t) = S^{ext}(t) = \operatorname{span} \begin{pmatrix} 1 + \frac{t}{2} \\ 1 \end{pmatrix}$$

for $t \in [0,1]$, and consequently, $S^{ext}(0)$ and $N_0(0)$ are no longer transversal. In the terminology of [31], the DAE (29) shows $t = 0$ to be a type $1A$-critical point. Those points are excluded from our analysis by requiring the canonical projector to be bounded. Consequently, in the sequel, we deal with type 0-critical points.

In the next section, we apply polynomial collocation to approximate solutions of (3) by means of an enlarged system,

$$f(u'(t), x(t), t) = 0, \tag{31a}$$
$$D(t)x(t) - u(t) = 0, \quad t \in (0,1], \tag{31b}$$

which can be brought into the form

$$\widehat{f}((\widehat{D}(t)\widehat{x}(t))', \widehat{x}(t), t) = 0, \quad t \in (0,1], \tag{32}$$

where $\widehat{x}(t) = (x(t), u(t))^T$ and

$$\widehat{f}(y, \widehat{x}, t) := \begin{pmatrix} f(y, x, t) \\ D(t)x - u \end{pmatrix}, \quad \widehat{D}(t) = \begin{pmatrix} 0 & I \end{pmatrix} =: \widehat{D}, \quad \widehat{x} = \begin{pmatrix} x \\ u \end{pmatrix} \in \mathcal{D} \times \mathbb{R}^n, \quad y \in \mathbb{R}^n.$$

Problem (32) is a regular DAE system with properly stated leading term and tractability index 1 for $t > 0$. To see this, note that $\widehat{D}(t)$ is constant and therefore we define the related matrices $\widehat{G}_0(t)$, $\widehat{Q}_0$, and $\widehat{G}_1(t)$ as

$$\widehat{G}_0(y, \widehat{x}, t) := \begin{pmatrix} 0 & f_y(y, x, t) \\ 0 & 0 \end{pmatrix}, \quad \widehat{Q}_0 := \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\widehat{G}_1(y, \widehat{x}, t) := \widehat{G}_0(y, \widehat{x}, t) + \begin{pmatrix} f_x(y, x, t) & 0 \\ D(t) & -I \end{pmatrix} \widehat{Q}_0 = \begin{pmatrix} f_x(y, x, t) & f_y(y, x, t) \\ D(t) & 0 \end{pmatrix},$$

respectively. Moreover,

$$\ker \widehat{G}_1 = \{z \in \mathbb{R}^{m+n},\, z = (z_1, z_2)^T;\, z_1 = Q_0 w,\, z_2 = Dw,\, w \in \ker G_1\}$$

for $t > 0$, which means that $\widehat{G}_1$ is nonsingular for $t > 0$. The enlarged DAE (32) has exactly the same inherent ODE as the original DAE (3a).

3 Collocation Methods – Convergence at Collocation Points

For the theoretical discussion of collocation methods, we define meshes

$$\Delta := (\tau_0, \tau_1, \ldots, \tau_N),$$

and $h_i := \tau_{i+1} - \tau_i,\ i = 0, \ldots, N-1,\ \tau_0 = 0,\ \tau_N = 1$. For reasons of simplicity, we restrict the discussion to equidistant meshes, $h_i = h,\ i = 0, \ldots, N-1$. However, the results also hold for nonuniform meshes which have a limited variation in the stepsizes. For collocation, k distinct points $t_{i,j} := \tau_i + h_i \rho_j,\ j = 1, \ldots, k$, are inserted in each subinterval (τ_i, τ_{i+1}). Since we want to focus on singular problems, we restrict ourselves to interior collocation points, where $\rho_1 > 0$ and $\rho_k < 1$. This avoids replacing the collocation equation at $t_{0,1}$ by an asymptotic relation [8]. A grid with equidistant interior collocation points is illustrated in Figure 1.

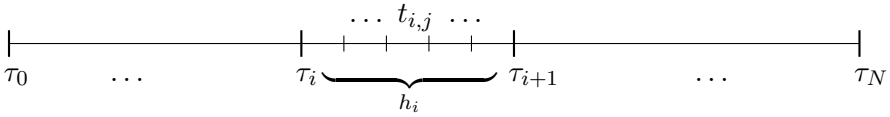

Figure 1: The computational grid

Now, let us denote by $\mathcal{B}_k$ the Banach space of continuous, piecewise polynomial functions $q \in \mathbb{P}_k$ of degree $\leq k,\ k \in \mathbb{N}$, equipped with the maximum norm $\|\cdot\|_\infty$. In the following, we denote by q the vector valued functions from $\mathcal{B}_k$ independently of the number of their components.

By $p \in \mathcal{B}_k$ we denote an approximation to the exact solution $x_\star$ of (3), and by $q \in \mathcal{B}_k$ an approximation to the exact solution $u_\star$ of the inherent ODE (20). As usual, to compute p and q, we set up the collocation equations augmented by the proper number of boundary conditions,

$$f(q'(t_{i,j}), p(t_{i,j}), t_{i,j}) = 0, \tag{33a}$$
$$D(t_{i,j})p(t_{i,j}) - q(t_{i,j}) = 0, \tag{33b}$$
$$B_0\, q(0) + B_1 q(1) = \beta, \tag{33c}$$

where $j = 1, \dots, k$ and $i = 0, \dots, N-1$. By inspection of the number of unknowns ($(k+1)N(n+m)$ polynomial coefficients) and equations ($Nk(n+m)$ collocation conditions, $(N-1)(n+m)$ continuity conditions for p and q, n boundary conditions) we see that m further conditions will be necessary to close the system for the numerical treatment. Clearly, these additional conditions have to be consistent with the original DAEs. Various choices are possible. In general, we may complete the above scheme by the m equations

$$D(0)p(0) - q(0) = 0, \quad W^{ext}(q'(0), p(0), 0) f(q'(0), p(0), 0) = 0. \tag{34}$$

If the DAE (3a) is given with separated derivative free equations, cf. Lemma 8.2,

$$f_1((D(t)x(t))', x(t), t) = 0, \tag{35a}$$
$$f_2(x(t), t) = 0, \tag{35b}$$

where f_1 and f_2 have n and $m-n$ components, respectively, then we can augment the scheme (33) by

$$f_2(p(0), 0) = 0, \quad D(0)p(0) = q(0). \tag{36}$$

In order to show that the collocation scheme for the DAE (33) comprises exactly the same scheme applied to the inherent ODE, we introduce

$$u_{ij} := q(t_{ij}), \quad w_{ij} := D(t_{ij})^- q'(t_{ij}) + Q_0(t_{ij}) p(t_{ij}), \tag{37}$$

and obtain from (33a)

$$f(D(t_{ij})w_{ij}, D(t_{ij})^- u_{ij} + Q_0(t_{ij}) w_{ij}, t_{ij}) = 0. \tag{38}$$

By applying the decoupling function to (38), the relation

$$w_{ij} = \omega(u_{ij}, t_{ij}) = \omega(q(t_{ij}), t_{ij})$$

follows, and hence,

$$q'(t_{ij}) = D(t_{ij})\omega(q(t_{ij}), t_{ij}), \tag{39}$$

and

$$p(t_{ij}) = D(t_{ij})^- q(t_{ij}) + Q_0(t_{ij})\omega(q(t_{ij}), t_{ij}), \tag{40}$$

for all collocation points t_{ij}. The system (39) together with the boundary conditions (33c) form a classical collocation scheme for $q \in \mathcal{B}_k$. According to Theorem 3.1 in [25], there exists a unique collocation solution $q \in \mathcal{B}_k$ of the scheme (39) subject to (33c), under the assumptions that the underlying analytical problem is well-posed with sufficiently smooth data, and that the grid is sufficiently fine. Finally, $p \in \mathcal{B}_k$ is uniquely specified by the values of p at all collocation points, see (40), and the consistency conditions given by (34) or (36).

The convergence analysis for the collocation scheme (39) subject to appropriately posed boundary conditions has been given in [25]. Especially, the structure of the global error of the collocation solution at the collocation points has been described,

$$q(t_{ij}) - u_\star(t_{ij}) = \epsilon_u(t_{ij})h^k + r(t_{ij}) + O(h^{k+1}), \tag{41}$$

where ϵ_u is a smooth function with $\epsilon_u(0) \in \ker M(0)$, and $r = O(h^k)$ is in $\mathcal{B}_k$. Thus, we can conclude that $q(t_{ij}) - u_\star(t_{ij}) = O(h^k)$ holds. It follows from the solution representation at the collocation points (19) and (40) that

$$\begin{aligned} p(t_{ij}) &= D(t_{ij})^- q(t_{ij}) + Q_0(t_{ij})\omega(q(t_{ij}), t_{ij}),\\ x_\star(t_{ij}) &= D(t_{ij})^- u_\star(t_{ij}) + Q_0(t_{ij})\omega(u_\star(t_{ij}), t_{ij}). \end{aligned}$$

Consequently,

$$p(t_{ij}) - x_\star(t_{ij}) = D(t_{ij})^-(q(t_{ij}) - u_\star(t_{ij})) + \int_0^1 Q_0(t_{ij})\omega_u(sq(t_{ij}) + (1-s)u_\star(t_{ij}), t_{ij})ds\,(q(t_{ij}) - u_\star(t_{ij})).$$

Taking into account the form of ω_u, we finally obtain the error representation

$$p(t_{ij}) - x_\star(t_{ij}) = \int_0^1 \big[\underbrace{I - Q_0(t_{ij})G_1^{-1}(\Omega_{ij}(s))f_x(\Omega_{ij}(s))}_{P_{can}(\Omega_{ij}(s))}\big] ds\, D(t_{ij})^-(q(t_{ij}) - u_\star(t_{ij})), \tag{42}$$

where

$$\begin{aligned} \big(\Omega_{i,j}(s)\big) = \big(&D(t_{ij})\omega(sq(t_{ij}) + (1-s)u_\star(t_{ij}), t_{ij}),\\ &D(t_{ij})^-(sq(t_{ij}) + (1-s)u_\star(t_{ij})) + Q_0(t_{ij})\omega(sq(t_{ij}) + (1-s)u_\star(t_{ij}), t_{ij}),\ t_{ij}\big). \end{aligned}$$

Evidently, in case that the canonical projector function remains bounded, the error $p(t_{ij}) - x_\star(t_{ij})$ inherits the convergence order of $q(t_{ij}) - u_\star(t_{ij})$ which means that

$$p(t_{ij}) - x_\star(t_{ij}) = O(h^k). \tag{43}$$

This makes clear that the commutativity of the decoupling procedure and the collocation discretization, which is due to the numerically qualified DAE formulation (with $\mathcal{R}(D(t)) = \mathbb{R}^n$, $t \in [0,1]$, cf. [14], [15]), ensures the expected convergence order at the collocation points. We summarize the above result in the following theorem.

Theorem 3.1 *Let the nonlinear BVP* (3) *be well-posed and* $x_\star \in C^{k+1}[0,1]$ *be a solution of* (3)*, where* (3a) *has a properly stated leading term and tractability index 1 on* $\mathbb{R}^n \times \mathcal{D} \times (0,1]$*. The matrix* $G_1(y,x,t)$ *may undergo a rank drop for* $t \to 0$*, causing a singularity of the first kind in the associated inherent ODE* (20)*, see also* (21)*. To this end, let*

$$tG_1(y,x,t)^{-1}$$

have a continuous extension for $t \to 0$*. Let the projector function* W *have the continuous extension* W^{ext} *for* $t \to 0$*. The eigenvalues of the matrix* $M(0)$ *in* (21) *are assumed to have either negative real parts or to be*

zero. Moreover, the Jordan boxes associated with the zero eigenvalue are assumed to be diagonal. Then, for sufficiently fine grids, the collocation scheme (33), (34) *provides a uniquely determined pair of collocation polynomials* p *and* q *such that at* k *interior collocation points*

$$p(t_{ij}) - x_\star(t_{ij}) = O(h^k), \quad q(t_{ij}) - u_\star(t_{ij}) = O(h^k), \quad i = 0, \ldots, N-1, \;\; j = 1, \ldots, k,$$

holds.

In case that $q(t_{ij}) - u_\star(t_{ij}) = O(h^{k+1})$, cf. Section 4.1, the order of convergence $k+1$ carries over to all components due to (42).
The same result applies also to the BVP (33), (36) and in case of BVPs where the consistency conditions (34) or (36) have been replaced by the corresponding conditions at $t = 1$.
If the above simple structure associated with the zero eigenvalue of $M(0)$ is not valid, the convergence order is polluted by additional logarithmic terms, $q(t_{ij}) - u_\star(t_{ij}) = O(h^k |\ln h|^{n_0 - 1})$, $n_0 \in \mathbb{N}$, cf. [16].

If the transversality condition is violated for $t \to 0$ and P_{can} becomes unbounded, formula (42) indicates possible order reductions. Such order reductions have already been observed in [26]. It should be emphasized that in such a case the necessary initial condition (26) and the consistency condition (34) or (36) may overlap. This means that consistency conditions have to be accordingly posed at $t = 1$.

4 Collocation Methods – Uniform Convergence

4.1 Differential Components

Recall that the system (39) subject to the boundary conditions (33c) forms a classical collocation scheme for $q \in \mathcal{B}_k$. Therefore, we can use convergence results for explicit singular ODEs developed in [17]. For k equidistant interior collocation points the following convergence orders hold:

$$\|q - u_\star\| = O(h^{k+1}), \;\; \text{for } k \text{ odd}, \quad \|q - u_\star\| = O(h^k), \;\; \text{for } k \text{ even}. \tag{44}$$

Moreover, for Gaussian collocation points the superconvergence behavior $O(h^{2k})$ in the mesh points does not hold in general, a well known fact in the context of singular ODEs [17]. Rather, the order $k+1$ is observed uniformly in t (uniform superconvergence).

4.2 All Components

We now consider the convergence behavior of the whole numerical solution including algebraic components. We perform this error analysis relying on techniques presented in [26]. Here, we consider BVPs which can be reduced to well-posed IVPs using shooting arguments. This means a restriction on the spectrum of $M(0)$ specified before. As a prerequisite for the following investigations we quote additional uniform convergence results for the differential components of x, see Theorem 3.1 in [25],

$$\|q - u_\star\| = O(h^k), \quad \|q' - u_\star'\| = O(h^k). \tag{45}$$

Moreover, we have

$$q(t) - u_\star(t) = tO(h^k) + RO(h^k). \tag{46}$$

From the analysis of the linear case in [26] and formula (42), we know that an additional order reduction can be attributed to the fact that the canonical projector is unbounded for $t \to 0$. Therefore, we focus on the case where this projector remains bounded.

Let the exact solution and the related collocation polynomial be given as

$$\widehat{x}_\star(t) := \begin{pmatrix} x_\star(t) \\ u_\star(t) \end{pmatrix}, \quad \widehat{p}(t) := \begin{pmatrix} p(t) \\ q(t) \end{pmatrix},$$

respectively. Let us introduce the error function $\widehat{e}(t) := (e(t), e_u(t))^T$, $\widehat{e} \in \mathcal{B}_k$, by

$$\widehat{e}(0) = \begin{pmatrix} x_\star(0) - p(0) \\ u_\star(0) - q(0) \end{pmatrix}, \quad \widehat{e}'(t_{i,j}) = \widehat{x}_\star'(t_{ij}) - \widehat{p}'(t_{ij}), \;\; j = 1, \ldots, k, \;\; i = 0, \ldots N-1,$$

which yields

$$\widehat{e}'(t) = \widehat{x}_\star'(t) - \widehat{p}'(t) + O(h^k), \quad \widehat{e}(t) = \widehat{x}_\star(t) - \widehat{p}(t) + t \begin{pmatrix} r(t) \\ s(t) \end{pmatrix},$$

where $r(t) = O(h^k)$ and $s(t) = O(h^k)$ are interpolation errors. Note that for special choices of the nodes t_{ij}, such that the related collocation scheme shows superconvergence, the interpolation error also shows the corresponding higher order. For instance, for an odd number of equidistant abscissae, order $k+1$ results. This reasoning can also be applied to improve (45) above and (54) below.

From (31) and (33a), (33b) we know that

$$\begin{aligned} f(u_\star'(t_{ij}), x_\star(t_{ij}), t_{ij}) = 0, \quad D(t_{ij}) x_\star(t_{ij}) - u_\star(t_{ij}) = 0, \\ f(q'(t_{ij}), p(t_{ij}), t_{ij}) = 0, \quad D(t_{ij}) p(t_{ij}) - q(t_{ij}) = 0. \end{aligned}$$

This yields

$$\mathcal{A}(t_{ij})(q'(t_{ij}) - u_\star'(t_{ij})) + \mathcal{B}(t_{ij})(p(t_{ij}) - x_\star(t_{ij})) = 0, \tag{47a}$$
$$D(t_{ij})(p(t_{ij}) - x_\star(t_{ij})) - (q(t_{ij}) - u_\star(t_{ij})) = 0, \tag{47b}$$

with coefficients

$$\mathcal{A}(t) = \int_0^1 \tilde{\mathcal{A}}(t,\tau) d\tau, \quad \mathcal{B}(t) = \int_0^1 \tilde{\mathcal{B}}(t,\tau) d\tau,$$

$$\begin{aligned} \tilde{\mathcal{A}}(t,\tau) &= f_y(u_\star'(t) + \tau(q'(t) - u_\star'(t)), x_\star(t) + \tau(p(t) - x_\star(t)),\, t), \quad t, \tau \in [0,1], \\ \tilde{\mathcal{B}}(t,\tau) &= f_x(u_\star'(t) + \tau(q'(t) - u_\star'(t)), x_\star(t) + \tau(p(t) - x_\star(t)),\, t), \quad t, \tau \in [0,1]. \end{aligned}$$

Let us also introduce

$$\begin{aligned} \mathcal{G}_1(t) &:= \mathcal{A}(t) D(t) + \mathcal{B}(t) Q_0(t) = \int_0^1 \big(\tilde{\mathcal{A}}(t,\tau) D(t) + \tilde{\mathcal{B}}(t,\tau) Q_0(t)\big) d\tau \\ &= \int_0^1 G_1(u_\star'(t) + \tau(q'(t) - u_\star'(t)), x_\star(t) + \tau(p(t) - x_\star(t)),\, t) d\tau. \end{aligned}$$

The matrix functions $\mathcal{A}$ and $\mathcal{B}$ are continuous on $[0,1]$ because $x_\star, p, u_\star', q'$ are. Thus, also $\mathcal{G}_1$ is continuous on $[0,1]$.
From (47) and from the representations

$$e_u'(t_{ij}) = u_\star'(t_{ij}) - q'(t_{ij}), \tag{48a}$$
$$e_u(t_{ij}) = u_\star(t_{ij}) - q(t_{ij}) + t_{ij} s(t_{ij}), \tag{48b}$$
$$e(t_{ij}) = x_\star(t_{ij}) - p(t_{ij}) + t_{ij} r(t_{ij}), \tag{48c}$$

it follows that

$$\mathcal{A}(t_{ij})e'_u(t_{ij}) + \mathcal{B}(t_{ij})e(t_{ij}) = t_{ij}\mathcal{B}(t_{ij})r(t_{ij}), \tag{49a}$$
$$D(t_{ij})e(t_{ij}) - e_u(t_{ij}) = t_{ij}D(t_{ij})r(t_{ij}) - t_{ij}s(t_{ij}), \tag{49b}$$

subject to

$$e(0) = x_\star(0) - p(0), \quad e_u(0) = u_\star(0) - q(0) = RO(h^k). \tag{50}$$

Let us define the function $\varphi(u,t) := Q_0(t)\omega(u,t)$, $u \in \mathcal{D}_\omega$, $t > 0$. The function φ is continuous with continuous partial derivative $\varphi_u = -Q_{can}D^-$. We know that $Q_0(t)x_\star(t) = Q_0(t)\omega(u_\star(t),t) = \varphi(u_\star(t),t)$, $t \in (0,1]$, and there exists the limit

$$Q_0(0)x_\star(0) = \lim_{t\to 0}\varphi(u_\star(t),t) =: \varphi(u_\star(0),0).$$

From the representation

$$\varphi(u,t) = \varphi(u_\star(t),t) + \int_0^1 \varphi_u(su + (1-s)u_\star(t),t)ds(u - u_\star(t)),$$

together with the boundedness of Q_{can} for $t \to 0$, we conclude that $\varphi(u,t)$ is bounded for $t \to 0$ and hence the bounded extension of φ for $t \to 0$ exists. The consistency condition (34) now yields

$$P_0(0)p(0) = D(0)^-q(0), \quad Q_0(0)p(0) = \varphi(q(0),0),$$

such that

$$\begin{aligned}
x_\star(0) &= D(0)^-u_\star(0) + \varphi(u_\star(0),0),\\
p(0) &= D(0)^-q(0) + \varphi(q(0),0),\\
x_\star(0) - p(0) &= D(0)^-(u_\star(0) - q(0)) + \int_0^1 \varphi_u(su_\star(0) + (1-s)q(0),0)ds(u_\star(0) - q(0))\\
&= \int_0^1 P_{can}(\ldots)ds(u_\star(0) - q(0)).
\end{aligned}$$

Since P_{can} is bounded, $x_\star(0) - p(0)$, and hence $e(0)$ inherits the convergence order from $u_\star(0) - q(0)$.

Next, we take a closer look at the coefficients $\mathcal{A}$ and $\mathcal{B}$ and show that $\mathcal{A}(t_{ij}) \in \mathbb{R}^{m\times n}$ has full column rank n and $\mathcal{G}_1(t_{ij}) = \mathcal{A}(t_{ij})D(t_{ij}) + \mathcal{B}(t_{ij})Q_0(t_{ij})$ is nonsingular provided that the stepsize h is sufficiently small.

Rewrite

$$\begin{aligned}
\tilde{\mathcal{A}}(t,\tau) &= f_y(u_\star'(t), x_\star(t),\, t) \\
&+ \int_0^1 \{ f_{yy}(u_\star'(t) + (1-\xi)\tau(q'(t) - u_\star'(t)), x_\star(t) + (1-\xi)\tau(p(t) - x_\star(t)),\, t)\tau(q'(t) - u_\star'(t)) \\
&+ f_{yx}(u_\star'(t) + (1-\xi)\tau(q'(t) - u_\star'(t)), x_\star(t) + (1-\xi)\tau(p(t) - x_\star(t)),\, t)\tau(p(t) - x_\star(t))\}\, d\xi \\
&= A_\star(t) + \int_0^1 \{\cdots\}\, d\xi, \\
\tilde{\mathcal{B}}(t,\tau) &= f_x(u_\star'(t), x_\star(t),\, t) \\
&+ \int_0^1 \{ f_{xy}(u_\star'(t) + (1-\xi)\tau(q'(t) - u_\star'(t)), x_\star(t) + (1-\xi)\tau(p(t) - x_\star(t)),\, t)\tau(q'(t) - u_\star'(t)) \\
&+ f_{xx}(u_\star'(t) + (1-\xi)\tau(q'(t) - u_\star'(t)), x_\star(t) + (1-\xi)\tau(p(t) - x_\star(t)),\, t)\tau(p(t) - x_\star(t))\}\, d\xi \\
&= B_\star(t) + \int_0^1 \{\cdots\}\, d\xi.
\end{aligned}$$

Taking into account the convergence results formulated in (43) and (45) we obtain

$$\begin{aligned}
\mathcal{A}(t_{ij}) &= A_\star(t_{ij}) + \int_0^1 \int_0^1 \{ f_{yy}(\cdots, t_{ij})\tau(\underbrace{q'(t_{ij}) - u_\star'(t_{ij})}_{O(h^k)}) + f_{yx}(\cdots, t_{ij})\tau(\underbrace{p(t_{ij}) - x_\star(t_{ij})}_{O(h^k)})\}\, d\xi d\tau \\
&= A_\star(t_{ij}) + O(h^k),
\end{aligned}$$

analogously

$$\mathcal{B}(t_{ij}) = B_\star(t_{ij}) + O(h^k),$$

and hence

$$\mathcal{G}_1(t_{ij}) = G_{\star 1}(t_{ij}) + O(h^k).$$

We now assume the grid to be sufficiently fine so that the terms $O(h^k)$ are small and therefore $\mathcal{A}(t_{ij})$ inherits from $A_\star(t_{ij})$ the full rank n, while $\mathcal{G}_1(t_{ij})$ inherits the invertibility of $G_{\star 1}(t_{ij})$.

Let us now introduce

$$\mathcal{M}(t_{ij}) := -t_{ij} D(t_{ij}) \mathcal{G}_1(t_{ij})^{-1} \mathcal{B}(t_{ij}) D(t_{ij})^-.$$

We multiply (49a) by $D(t_{ij})\mathcal{G}_1(t_{ij})^{-1}$ and by $Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}$ and also (49b) by $D(t_{ij})^-$. Taking into account the relations

$$D\mathcal{G}_1^{-1}\mathcal{A} = D\mathcal{G}_1^{-1}\mathcal{A}DD^- = D\mathcal{G}_1^{-1}\mathcal{G}_1 D^- = I,\;\; Q_0\mathcal{G}_1^{-1}\mathcal{A} = Q_0\mathcal{G}_1^{-1}\mathcal{A}DD^- = Q_0\mathcal{G}_1^{-1}\mathcal{G}_1 D^- = 0,$$

$$D\mathcal{G}_1^{-1}\mathcal{B}Q_0 = D\mathcal{G}_1^{-1}\mathcal{G}_1 Q_0 = 0,\;\; D\mathcal{G}_1^{-1}\mathcal{B} = D\mathcal{G}_1^{-1}\mathcal{B}D^- D,$$

we find

$$\begin{aligned}
&e_u'(t_{ij}) = \frac{1}{t_{ij}}\mathcal{M}(t_{ij})e_u(t_{ij}) - \mathcal{M}(t_{ij})s(t_{ij}), \\
&Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij})P_0(t_{ij})e(t_{ij}) + Q_0(t_{ij})e(t_{ij}) = t_{ij}Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij})r(t_{ij}), \\
&P_0(t_{ij})e(t_{ij}) - D(t_{ij})^- e_u(t_{ij}) = t_{ij}P_0(t_{ij})r(t_{ij}) - t_{ij}D(t_{ij})^- s(t_{ij}).
\end{aligned}$$

Thus, the values of $e(t_{ij})$ for the interpolation process from which we can recover the error function $e(t)$ read $e(0) = O(h^k)$ and

$$\begin{aligned} e(t_{ij}) &= (P_0(t_{ij}) + Q_0(t_{ij}))e(t_{ij}) \\ &= (I - Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij}))(D(t_{ij})^- e_u(t_{ij}) + t_{ij}P_0(t_{ij})r(t_{ij}) - t_{ij}D(t_{ij})^- s(t_{ij})) \\ &\quad + t_{ij}Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij})r(t_{ij}). \end{aligned}$$

It can be seen that $e_u(t_{ij}) = t_{ij}O(h^k) + RO(h^k)$. This follows from $e_u(t) = u_\star(t) - q(t) + ts(t)$ and (46). Consequently, the error function e is determined by the values $e(0)$ and

$$e(t_{ij}) = (I - Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij}))D(t_{ij})^-(u_\star(t_{ij}) - q(t_{ij}) + t_{ij}D(t_{ij})r(t_{ij})) + t_{ij}Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij})r(t_{ij}),$$

that is

$$e(t_{ij}) = (I - Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij}))D(t_{ij})^-(u_\star(t_{ij}) - q(t_{ij})) + t_{ij}r(t_{ij}). \tag{51}$$

The crucial terms in the above expression are $\mathcal{G}_1(t_{ij})^{-1}$ for t_{ij} close to zero. First observe that $\mathcal{G}_1(t_{ij})^{-1}$ appears only within the expression

$$Q_0(t_{ij})\mathcal{G}_1(t_{ij})^{-1}\mathcal{B}(t_{ij}) =: \mathcal{Q}(t_{ij}),$$

and therefore, we shall investigate this expression in more detail. It holds that $\mathcal{Q}(t_{ij})Q_0(t_{ij}) = Q_0(t_{ij})$, $\mathcal{Q}(t_{ij})^2 = \mathcal{Q}(t_{ij})$, and $\mathcal{Q}(t_{ij})$ is a projector onto $N_0(t_{ij}) = \mathcal{R}(Q_0(t_{ij}))$ along

$$\ker \mathcal{Q}(t_{ij}) = \{z \in \mathbb{R}^m : \mathcal{B}(t_{ij})z \in \mathcal{R}(\mathcal{A}(t_{ij})D(t_{ij})) = \mathcal{R}(\mathcal{A}(t_{ij}))\} =: \mathcal{S}(t_{ij}).$$

This implies $\dim \mathcal{S}(t_{ij}) = n$ and

$$N_0(t_{ij}) \oplus \mathcal{S}(t_{ij}) = \mathbb{R}^m. \tag{52}$$

Now recall that for $t > 0$, $Q_{\star can}(t) = Q_0(t)G_{\star 1}(t)^{-1}B_\star(t)$ is the canonical projector onto $N_0(t)$ along

$$S_\star(t) = \{z \in \mathbb{R}^m : B_\star(t)z \in \mathcal{R}(A_\star(t))\} = S((D(t)x_\star(t))', x_\star(t), t),$$

and there exists the continuous extension $Q^{ext}_{\star can}$ of $Q_{\star can}$, which is defined on $[0,1]$. Owing to the condition (27) the decomposition

$$N_0(t) \oplus S^{ext}_\star(t) = \mathbb{R}^m, \quad t \in [0,1], \tag{53}$$

is valid with

$$S^{ext}_\star(t) = S^{ext}((D(t)x_\star(t))', x_\star(t), t) = \ker W^{ext}_\star(t)B_\star(t), \quad W^{ext}_\star(t) := W^{ext}((D(t)x_\star(t))', x_\star(t), t).$$

We know already that $\mathcal{A}(t_{ij})$ and $\mathcal{B}(t_{ij})$ approximate $A_\star(t_{ij})$ and $B_\star(t_{ij})$ respectively, and so the subspace $\mathcal{S}(t_{ij})$ appears to be a perturbed version of $S_\star(t_{ij})$. Roughly speaking, if the $O(h^k)$ terms are sufficiently small, then the projectors $\mathcal{Q}(t_{ij})$ realizing (52) are uniformly bounded as a consequence of the uniform boundedness of $Q^{ext}_\star(t_{ij})$ associated with (53). A precise justification for the case that $\mathcal{R}(f_y(y,x,t))$ does not depend on (y,x) follows from Lemma 8.1 in the Appendix and continuity arguments. Namely, due to Lemma 8.1, for any sufficiently small matrix function $\Delta B(t)$, the subspaces $N_0(t)$ and

$$\mathcal{S}(t) := \ker W_\star(t)(B_\star(t) + \Delta B(t))$$

are related in such a way that

$$N_0(t) \oplus \mathcal{S}(t) = \mathbb{R}^m, \quad t \in [0,1], \quad \mathcal{S}(t) = (I + (W^{ext}_\star(t)B_\star(t))^- W^{ext}(t)\Delta B(t))^{-1} S^{ext}_\star(t), \quad t \in [0,1],$$

and the projection function $\mathcal{Q}$ onto N_0 along $\mathcal{S}$ is continuous on the compact interval $[0,1]$. Therefore, the values of $\mathcal{Q}(t_{ij})$ are uniformly bounded. Consequently, this yields the uniform convergence result for the global error of the collocation solution to the DAE system,

$$x_\star(t) - p(t) = O(h^k). \tag{54}$$

If the consistency condition is posed at $t = 1$, we construct the error function $\widehat{e}$ in a similar way, but replace the above condition $e(0) = x_\star(0) - p(0)$ by $e(1) = x_\star(1) - p(1)$ such that

$$\begin{aligned} \widehat{e}(t) &= \widehat{x}_\star(t) - \widehat{p}(t) + \begin{pmatrix} (1-t)\tilde{r}(t) \\ ts(t) \end{pmatrix}, \\ e(1) &= \int_0^1 P_{can}(\ldots) d\tau D(1)^-(u_\star(1) - q(1)), \end{aligned}$$

and the terms $t_{ij} r(t_{ij})$ are replaced by $(1 - t_{ij})\tilde{r}(t_{ij})$. With these modifications formulae (51) and (53) hold.

5 Numerical Experiments

In this section we consider the following nonlinear DAE system:

$$A(t)(D(t)x(t))' + B(t)x(t) + h(x(t), t) = 0, \quad t \in (0, 1],$$

where $n = 2$, $m = 4$, and

$$A(t) = \begin{pmatrix} tI \\ 0 \end{pmatrix}, \quad D(t) = (I \quad 0), \quad B(t) = \begin{pmatrix} -11 & -18 & 3 & -1 \\ 12 & 19 & -2 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 3 & 0 & \frac{1}{5} \end{pmatrix}, \quad D(t)^- = \begin{pmatrix} I \\ 0 \end{pmatrix},$$

$$h(x, t) = t \begin{pmatrix} x_1 \sin(x_2) + x_3 e^{-x_1} \\ x_2 \cos(x_4) + x_4 \sin(x_1 + x_3) \\ x_1 x_2^3 + x_3 x_1 \\ x_1 x_2^2 + x_4 x_2^2 \end{pmatrix} + \beta(t), \quad x \in \mathcal{D}, \quad t \in [0, 1],$$

where $\mathcal{D} := \{x \in \mathbb{R}^4 : x_1 > -0.5\}$. The function $\beta(t)$ is chosen in such a way that

$$x_\star(t) = \begin{pmatrix} t^2 \sin(t) \\ te^t \\ t\cos(t) \\ \sin(t) \end{pmatrix}$$

is the exact solution and $\beta(0) = 0$ holds. Moreover,

$$Q_0(t) = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix},$$

$$G_1((x_1, x_2)^T, (x_3, x_4)^T, t) = \begin{pmatrix} t & 0 & 3 - te^{-x_1} & -1 \\ 0 & t & -2 + tx_4 \cos(x_1 + x_3) & 1 - tx_2 \sin(x_4) + t \sin(x_1 + x_3) \\ 0 & 0 & 1 + tx_1 & 0 \\ 0 & 0 & 0 & 0.2 + tx_2^2 \end{pmatrix}.$$

One can see that $G_1(y, x, t)$ is nonsingular for $t > 0$ and $tG_1(y, x, t)^{-1}$ remains bounded for $t \to 0$. The canonical projector $Q_{can}(y, x, t)$ has a continuous extension on $\mathbb{R}^2 \times \mathcal{D} \times [0, 1]$ due to the fact that the following matrix has full column rank, see (28):

$$\begin{pmatrix} W^{ext}(y, x, 0)B(0) \\ D(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 2 & 3 & 0 & 0.2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

The nonlinear inherent ODE is singular with a singularity of the first kind,

$$D(t)\omega(u,t) = \frac{1}{t}Mu + g(u,t),$$

where

$$M = \begin{pmatrix} 4 & 6 \\ -4 & -6 \end{pmatrix},$$

and $g(u,t)$ is smooth. We now augment the DAE system by the necessary number of boundary conditions given by

$$2x_1(0) + 3x_2(0) = 0, \quad x_1(1) + x_2(1) = \sin(1) + e, \tag{55a}$$

$$x_1(0) + x_2(0) + x_3(0) = 0, \quad 2x_1(0) + 3x_2(0) + 0.2x_4(0) = 0. \tag{55b}$$

In the following tables, we have collected results of our numerical experiments. All calculations have been carried out with MATLAB. In Table 1 we report on the global error of the solution x and its differential $x_d = (x_1, x_2)^T$ and algebraic $x_a = (x_3, x_4)^T$ components. In the upper part of the table we illustrate the asymptotical properties of the differential components x_d with error at points τ_i defined as $\max_{\tau_i}\{\max\{|x_1(\tau_i)|, |x_2(\tau_i)|\}\}$. Similarly, the global error at the collocation points is given by $\max_{t_{ij}}\{\max\{|x_1(t_{ij})|, |x_2(t_{ij})|\}\}$. For the algebraic components the above quantities are specified in an analogous way. The order and the error constant are computed from two consecutive steps in the usual fashion. We can see that the observed order of convergence is $k+1$ (since we are using an odd number $k = 3$ of equidistant collocation points).
In the lower part of the table we report on the asymptotical behavior of the whole solution vector. The left column serves as an illustration for the estimate given in (43). In order to illustrate (54), in the right column the maximal global error of x is calculated by considering its values at 1000 uniformly spaced points in the interval $[0, 1]$. In both cases, we also observe the convergence of order $4 = k + 1$.

In Table 2 the analogous results are given for $k = 4$ and equidistant points. We observe the order of convergence $k = 4$, as predicted by the analysis.

uniform mesh		differential components x_d at points τ_i			differential components x_d at points t_{ij}		
N	h	error	order	const.	error	order	const.
10	1.00e-01	1.719e-06			1.109e-06		
20	5.00e-02	1.081e-07	4.0	1.684e-02	7.901e-08	3.8	7.187e-03
40	2.50e-02	6.786e-09	4.0	1.698e-02	5.256e-09	3.9	9.653e-03
80	1.25e-02	4.254e-10	4.0	1.711e-02	3.390e-10	4.0	1.138e-02
160	6.25e-03	2.664e-11	4.0	1.722e-02	2.153e-11	4.0	1.254e-02
320	3.13e-03	1.667e-12	4.0	1.732e-02	1.357e-12	4.0	1.328e-02
uniform mesh		algebraic components x_a at points τ_i			algebraic components x_a at points t_{ij}		
N	h	error	order	const.	error	order	const.
10	1.00e-01	1.223e-05			1.839e-06		
20	5.00e-02	8.253e-07	3.9	9.495e-02	1.091e-07	4.1	2.187e-02
40	2.50e-02	5.341e-08	3.9	1.135e-01	7.128e-09	3.9	1.439e-02
80	1.25e-02	3.392e-09	4.0	1.257e-01	4.595e-10	4.0	1.548e-02
160	6.25e-03	2.146e-10	4.0	1.284e-01	2.921e-11	4.0	1.693e-02
320	3.13e-03	1.350e-11	4.0	1.340e-01	1.842e-12	4.0	1.795e-02
uniform mesh		solution x at points t_{ij}			solution x at 1000 uniform points		
N	h	error	order	const.	error	order	const.
10	1.00e-01	1.839e-06			1.223e-05		
20	5.00e-02	1.091e-07	4.1	2.187e-02	8.253e-07	3.9	9.495e-02
40	2.50e-02	7.128e-09	3.9	1.439e-02	5.341e-08	3.9	1.135e-01
80	1.25e-02	4.595e-10	4.0	1.548e-02	2.774e-09	4.3	3.661e-01
160	6.25e-03	2.921e-11	4.0	1.693e-02	1.734e-10	4.0	1.136e-01
320	3.13e-03	1.842e-12	4.0	1.795e-02	1.085e-11	4.0	1.126e-01

Table 1: Numerical experiment for collocation with three equidistant collocation points.

uniform mesh		differential components x_d at points τ_i			differential components x_d at points t_{ij}		
N	h	error	order	const.	error	order	const.
10	1.00e-01	2.043e-07			1.886e-07		
20	5.00e-02	1.268e-08	4.0	2.087e-03	1.225e-08	3.9	1.662e-03
40	2.50e-02	7.916e-10	4.0	2.043e-03	7.787e-10	4.0	1.818e-03
80	1.25e-02	4.946e-11	4.0	2.030e-03	4.907e-11	4.0	1.909e-03
160	6.25e-03	3.088e-12	4.0	2.040e-03	3.076e-12	4.0	1.973e-03
320	3.13e-03	1.895e-13	4.0	2.308e-03	1.891e-13	4.0	2.269e-03
uniform mesh		algebraic components x_a at points τ_i			algebraic components x_a at points t_{ij}		
N	h	error	order	const.	error	order	const.
10	1.00e-01	1.127e-06			1.457e-07		
20	5.00e-02	7.074e-08	4.0	1.110e-02	9.416e-09	4.0	1.304e-03
40	2.50e-02	4.430e-09	4.0	1.122e-02	5.483e-10	4.1	2.046e-03
80	1.25e-02	2.770e-10	4.0	1.131e-02	3.266e-11	4.1	1.812e-03
160	6.25e-03	1.728e-11	4.0	1.149e-02	1.993e-12	4.0	1.557e-03
320	3.13e-03	1.464e-12	3.6	1.220e-03	1.401e-13	3.8	5.517e-04
uniform mesh		solution x at points t_{ij}			solution x at 1000 uniform points		
N	h	error	order	const.	error	order	const.
10	1.00e-01	1.886e-07			1.127e-06		
20	5.00e-02	1.225e-08	3.9	1.662e-03	7.074e-08	4.0	1.110e-02
40	2.50e-02	7.787e-10	4.0	1.818e-03	4.430e-09	4.0	1.122e-02
80	1.25e-02	4.907e-11	4.0	1.909e-03	2.770e-10	4.0	1.131e-02
160	6.25e-03	3.076e-12	4.0	1.973e-03	1.727e-11	4.0	1.154e-02
320	3.13e-03	1.891e-13	4.0	2.269e-03	1.464e-12	3.6	1.213e-03

Table 2: Numerical experiment for collocation with four equidistant collocation points.

6 A Special Class of Quasi-Linear DAEs

The aim of this section is to derive *explicit sufficient conditions in terms of the original problem data* to ensure the special structure of the inherent ODE (21) for a quasi-linear DAE of the form

$$A(t)(D(t)x(t))' + B(t)x(t) + h(x(t), t) = 0. \tag{56}$$

As for the general case we assume $\mathcal{R}(D(t)) = \mathbb{R}^n$, $t \in [0, 1]$ and $\ker A(t) = \{0\}$, $t \in (0, 1]$. Here, we find that it is very convenient to use the orthogonal projectors $Q_0(t)$ and $P_0(t) = I - Q_0(t)$ onto $\ker D(t)$ and $\ker D(t)^\perp$, respectively. We also make the following assumptions:

(i) Let the matrix $\tilde{G}_1(t) := A(t)D(t) + B(t)Q_0(t)$ be nonsingular for all $t \in (0, 1]$.

(ii) Let $t\tilde{G}_1(t)^{-1}$ have a continuous extension on $[0, 1]$.

(iii) Let $\tilde{G}_1(t)^{-1}h(x, t)$ and $\tilde{G}_1(t)^{-1}h_x(x, t)$ have smooth extensions on $\mathcal{D} \times [0, 1]$.

(iv) Let $Q_0(t)\tilde{G}_1(t)^{-1}B(t)$ be bounded for $t \to 0$.

(v) Let the function h have the structure

$$(I - A(t)A(t)^-)(h(x, t) - h(P_0(t)x, t)) = 0, \quad t > 0.$$

The properties (i)–(iv) have already been useful in the analysis of linear problems, cf. [26]. The structural assumption (v) is satisfied in case that at least $h(x(t), t) = \bar{h}(D(t)x(t), t)$.

In Section 2 we assumed that

A1 $G_1(y, x, t) = A(t)D(t) + (B(t) + h_x(x, t))Q_0(t)$ is nonsingular for $t > 0$,

A2 $tG_1(y, x, t)^{-1}$ has a continuous extension for $t \to 0$,

A3 the decoupling function has the form, see (21),

$$D(t)\omega(u, t) = \frac{1}{t}M(t)u + g(u, t),$$

A4 the canonical projector function has a continuous extension for $t \to 0$.

Note that the example described in the previous section satisfies (i)–(iv), but the structural condition (v) is violated. Nevertheless, all assumptions A1–A4 hold for this example. This means that A1–A4 do not necessarily imply (i)–(v). However, conversely it holds that properties A1–A3 follow from (i)–(v).

We first note that (i), (iii) and (v) imply

$$\begin{aligned} Q_0(t)\tilde{G}_1(t)^{-1}h(x, t) &= Q_0(t)\tilde{G}_1(t)^{-1}[\underbrace{A(t)}_{\tilde{G}_1(t)P_0(t)} A(t)^- + I - A(t)A(t)^-]h(x, t) \\ &= Q_0(t)\tilde{G}_1(t)^{-1}[I - A(t)A(t)^-]h(P_0(t)x, t) = Q_0(t)\tilde{G}_1(t)^{-1}h(P_0(t)x, t), \end{aligned} \tag{57}$$

and additionally

$$Q_0(t)\tilde{G}_1(t)^{-1}h_x(x, t) = Q_0(t)\tilde{G}_1(t)^{-1}h_x(P_0(t)x, t)P_0(t).$$

Thus

$$Q_0(t)\tilde{G}_1(t)^{-1}h_x(x, t)Q_0(t) = 0.$$

Now, we show A1. For $t > 0$ we have

$$\begin{aligned} G_1(y,x,t) &= \tilde{G}_1(t) + h_x(x,t)Q_0(t) = \tilde{G}_1(t)(I + \tilde{G}_1(t)^{-1}h_x(x,t)Q_0(t)) \\ &= \tilde{G}_1(t)(I + P_0(t)\tilde{G}_1(t)^{-1}h_x(x,t)Q_0(t)). \end{aligned}$$

Since $I + P_0(t)\tilde{G}_1(t)^{-1}h_x(x,t)Q_0(t)$ has the inverse $I - P_0(t)\tilde{G}_1(t)^{-1}h_x(x,t)Q_0(t)$, the matrix $G_1(y,x,t)$ is nonsingular for $t > 0$ together with $\tilde{G}_1(t)$.

A2 is now a simple consequence of the representation

$$tG_1(y,x,t)^{-1} = (I - P_0(t)\tilde{G}_1(t)^{-1}h_x(x,t)Q_0(t))t\tilde{G}_1(t)^{-1}.$$

Moreover, A4 is also valid, since

$$\begin{aligned} Q_{can}(y,x,t) &:= Q_0(t)G_1(y,x,t)^{-1}(B(t) + h_x(x,t)) = Q_0(t)\tilde{G}_1(t)^{-1}(B(t) + h_x(x,t)) \\ &= Q_0(t)\tilde{G}_1(t)^{-1}B(t) + Q_0(t)\tilde{G}_1(t)^{-1}h_x(x,t) \end{aligned}$$

has a continuous extension for $t \to 0$.
Now we turn to A3. The equation defining the decoupling function $\omega(u,t)$ reads:

$$A(t)D(t)w + B(t)(D(t)^{-}u + Q_0(t)w) + h(D(t)^{-}u + Q_0(t)w, t) = 0.$$

For $t > 0$ this is equivalent to

$$w = -\tilde{G}_1(t)^{-1}B(t)D(t)^{-}u - \tilde{G}_1(t)^{-1}h(D(t)^{-}u + Q_0(t)w, t).$$

Taking into account (57) we find

$$Q_0(t)w = -Q_0(t)\tilde{G}_1(t)^{-1}B(t)D(t)^{-}u - Q_0(t)\tilde{G}_1(t)^{-1}h(D(t)^{-}u, t).$$

This means that the function $\omega(u,t)$ satisfies

$$Q_0(t)\omega(u,t) = -Q_0(t)\tilde{G}_1(t)^{-1}B(t)D(t)^{-}u - Q_0(t)\tilde{G}_1(t)^{-1}h(D(t)^{-}u, t).$$

The right-hand side of the above identity has a continuous extension for $t \to 0$. We now insert the last expression into the identity

$$D(t)\omega(u,t) = \underbrace{-D(t)\tilde{G}_1(t)^{-1}B(t)D(t)^{-}}_{=:\frac{1}{t}M(t)}u\ \underbrace{-D(t)\tilde{G}_1(t)^{-1}h(D(t)^{-}u + Q_0(t)\omega(u,t), t)}_{=:g(u,t)}$$

and see that A3 follows, with $M(t)$ and $g(u,t)$ continuous for $t \to 0$.

7 Conclusions

In this work, we discussed the convergence behavior of collocation schemes applied to solve BVPs in nonlinear index 1 DAEs, with type 0-critical point at the left endpoint $t = 0$ of the interval of integration. Here, such a critical point of the DAE results in a singularity of the first kind in the inherent ODE system. Remarkably, we can use standard polynomial collocation at k interior[1] collocation points to solve the DAE system written in its *original form*. We specified conditions under which for a certain class of well-posed boundary value problems in DAEs having sufficiently smooth solutions, the global error at the *collocation points* shows the convergence behavior known from the ODE case. Here, the boundedness of the so-called canonical projector function plays a crucial role. *Uniform convergence* of order k was shown to hold for DAEs which expose the derivative free part. The superconvergence order in the mesh points known for special choices of collocation points (Gaussian) for regular ODEs, does not hold even for singular ODEs, and thus neither for the DAEs we considered in this paper, in general.

For the case that the canonical projector function becomes unbounded for $t \to 0$, order reductions have to be expected.

[1] and therefore avoiding the evaluation of the DAE at the singular point

8 Appendix

Lemma 8.1 *Let W, B, ΔB be $m \times m$ matrices such that $W^2 = W = W^T$, $\operatorname{rank} W = m - n$. Let $N \subset \mathbb{R}^m$ be an $(m-n)$–dimensional subspace with $S := \ker WB$ and $N \oplus S = \mathbb{R}^m$. Let Q_{can} be the projector onto N along S, $P_{can} := I - Q_{can}$, and $\tilde{S} := \ker W(B + \Delta B)$. Then, for sufficiently small ΔB it holds*

$$\tilde{S} = (I + (WB)^- W\Delta B)^{-1} S, \quad N \oplus \tilde{S} = \mathbb{R}^m,$$

where $(WB)^-$ is the reflexive generalize inverse of WB with $(WB)^- WB = Q_{can}$ and $WB(WB)^- = W$.

Proof: By the assumptions, S has dimension n, WB has rank $m - n$ and $\mathcal{R}(WB) = \mathcal{R}(W)$. This justifies the choice of $(WB)^-$. We now represent $\tilde{z} \in \tilde{S}$,

$$\begin{aligned}
&\tilde{z} = P_{can}\tilde{z} + Q_{can}\tilde{z} \Rightarrow \\
&W(B + \Delta B)(P_{can}\tilde{z} + Q_{can}\tilde{z}) = 0 \Rightarrow \\
&(WB + W\Delta B)Q_{can}\tilde{z} = -W(B + \Delta B)P_{can}\tilde{z} \Rightarrow \\
&Q_{can}\tilde{z} + (WB)^- W\Delta B Q_{can}\tilde{z} = -(WB)^- W(B + \Delta B)P_{can}\tilde{z} \Rightarrow \\
&Q_{can}\tilde{z} = -(I + (WB)^- W\Delta B)^{-1}(WB)^- W(B + \Delta B)P_{can}\tilde{z},
\end{aligned}$$

and finally,

$$\tilde{z} = (I + (WB)^- W\Delta B)^{-1} P_{can}\tilde{z},$$

provided that ΔB is sufficiently small for $I + (WB)^- W\Delta B$ to be nonsingular. This yields the relation $\tilde{S} = (I + (WB)^- W\Delta B)^{-1} S$ and $\dim \tilde{S} = \dim S = n$. It remains to show that $\tilde{S} \cap N = \{0\}$. Let $\tilde{z} \in \tilde{S} \cap N$. Then,

$$z := (I + (WB)^- W\Delta B)\tilde{z} \in S = \ker WB$$

and $P_{can}\tilde{z} = 0$. From $P_{can}(WB)^- = 0$ we obtain $P_{can}\tilde{z} = P_{can}z = z$. Thus, $z = 0$ and $\tilde{z} = 0$. □

Lemma 8.2 *Suppose that* $\operatorname{rank} f_y(y, x, t) = n$, *$x \in \mathcal{D}$, $t \in (0, 1]$ and let $\mathcal{R}(f_y(y, x, t))$ be independent of y. Let $W(x, t)$ be the orthogonal projection onto $\mathcal{R}(f_y(y, x, t))^\perp$ for $t > 0$, and W^{ext} be the continuous extension of W for $t \geq 0$. Then, the identity*

$$W^{ext}(x, t)f(y, x, t) \equiv W^{ext}(x, t)f(0, x, t)$$

holds.

Proof: The result follows from

$$W^{ext}(x, t)(f(y, x, t) - f(0, x, t)) = \int_0^1 W^{ext}(x, t) f_y(sy, x, t) ds = 0$$

for all $y \in \mathbb{R}^m$, $x \in \mathcal{D}$, and $t \in [0, 1]$. □
The projection W^{ext} splits the DAE (3a) into two parts,

$$(I - W^{ext}(x(t), t))f((D(t)x(t))', x(t), t) = 0, \quad W^{ext}(x(t), t)f(0, x(t), t) = 0.$$

Since W^{ext} has constant rank $m-n$, these subsystems can be reduced to a system of n and $m-n$ equations, respectively, resulting in the form (35).

References

[1] P. Amodio, T. Levitina, G. Settanni, and E.B. Weinmüller. On the calculation of the finite Hankel transform eigenfunctions. In preparation.

[2] W. Auzinger, O. Koch, and E.B. Weinmüller. Analysis of a new error estimate for collocation methods applied to singular BVPs. *SIAM J. Numer. Anal.*, 42:2366–2386, 2005.

[3] W. Auzinger, E. Karner, O. Koch, and E.B. Weinmüller. Collocation methods for the solution of eigenvalue problems for singular ordinary differential equations. *Opuscula Math.*, 26:229–241, 2006.

[4] P. Bailey, W. Everitt, and A. Zettl. Computing eigenvalues of singular Sturm-Liouville problems. *Results Math.*, 20:391–423, 1991.

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

[6] C.J. Budd and R. Kuske. Localised periodic pattern for the non-symmetric generalized Swift-Hohenberg equations. *Physica D*, 208:73–95, 2005.

[7] C.J. Budd and J.F. Williams. Parabolic Monge-Ampère methods for blow-up problems in several spatial dimensions. *J. Phys. A*, 39:5425–5463, 2006.

[8] J. Cash, G. Kitzhofer, O. Koch, G. Moore, and E.B. Weinmüller. Numerical solution of singular two-point BVPs. *JNAIAM J. Numer. Anal. Indust. Appl. Math.*, 4:129–149, 2009.

[9] M. Drmota, R. Scheidl, H. Troger, and E.B. Weinmüller. On the imperfection sensitivity of complete spherical shells. *Comp. Mech.*, 2:63–74, 1987.

[10] G.B. Froment and K.B. Bischoff. *Chemical reactor analysis and design.* John Wiley & Sons Inc., New York, 1990.

[11] S. Golub. Measures of restrictions in inward foreign direct investment in OECD countries. *OECD Economics Dept. WP*, Nr. 357.

[12] R. Hammerling, O. Koch, C. Simon, E.B. Weinmüller. Numerical solution of singular ODE eigenvalue problems in electronic structure computations. *J. Comput. Phys.*, 181: 1557–1561, 2010.

[13] E. Helpman, M.J. Melitz, and S.R. Yeaple. Export versus FDI with heterogeneous firms. *Amer. Econ. Rev.*, 94(1):300–316, 2004.

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

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

[16] F.R. de Hoog and R. Weiss. Difference methods for boundary value problems with a singularity of the first kind. *SIAM J. Numer. Anal.*, 13:775–813, 1976.

[17] F.R. de Hoog and R. Weiss. Collocation methods for singular boundary value problems. *SIAM J. Numer. Anal.*, 15:198–217, 1978.

[18] F.R. de Hoog and R. Weiss. The numerical solution of boundary value problems with an essential singularity. *SIAM J. Numer. Anal.*, 16:637–669, 1979.

[19] F.R. de Hoog and R. Weiss. On the boundary value problem for systems of ordinary differential equations with a singularity of the second kind. *SIAM J. Math. Anal.*, 11:41–60, 1980.

[20] T. Kapitula. Existence and stability of singular heteroclinic orbits for the Ginzburg-Landau equation. *Nonlinearity*, 9:669–685, 1996.

[21] B. Karabay. Foreign direct investment and host country policies: A rationale for using ownership restrictions. Technical report, University of Virginia, WP, 2005.

[22] H. Keller. Approximation methods for nonlinear problems with application to two-point boundary value problems. *Math. Comp.*, 29:464–474, 1975.

[23] G. Kitzhofer, O. Koch, and E.B. Weinmüller. Pathfollowing for essentially singular boundary value problems with application to the complex Ginzburg-Landau equation. *BIT.*, 49:217–245, (2009).

[24] G. Kitzhofer, O. Koch, G. Pulverer, Ch. Simon, and E.B. Weinmüller. Numerical treatment of singular BVPs: The new Matlab code `bvpsuite`. *JNAIAM J. Numer. Anal. Indust. Appl. Math.*, 5:113–134, 2010.

[25] O. Koch. Asymptotically correct error estimation for collocation methods applied to singular boundary value problems. *Numer. Math.*, 101:143–164, 2005.

[26] O. Koch, R. März, D. Praetorius, and E.B. Weinmüller. Collocation methods for linear index 1 DAEs with a singularity of the first kind. *Math. Comp.*, 79:281-304, 2010.

[27] O. Koch and E.B. Weinmüller. Analytical and numerical treatment of a singular initial value problem in avalanche modeling. *Appl. Math. Comput.*, 148(2):561–570, 2004.

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

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

[30] R. März. Differential algebraic equations anew. *Appl. Numer. Math.*, 42:315–335, 2002.

[31] R. März and R. Riaza. On linear differential-algebraic equations with properly stated leading terms: A-critical points. *Math. Comp. Model. Dyn. Sys.*, 13:291–314, 2004.

[32] D.M. McClung and A.I. Mears. Dry-flowing avalanche run-up and run-out. *J. Glaciol.*, 41(138):359–369, 1995.

[33] G. Moore. Geometric methods for computing invariant manifolds. *Appl. Numer. Math.*, 17:319–331, 1995.

[34] V.V. Ranade. *Computational flow modeling for chemical engineering.* Academic Press, San Diego, 2002.

[35] K. Sundmacher and U. Hoffmann. Multicomponent mass and energy transport on different length scales in a packed reactive distillation column for heterogeneously catalyzed fuel ether production. *Chem. Eng. Sci.*, 49:4443–4464, 1994.

[36] Y. Chin-Yu, A.B. Chen, D.M. Nicholson, and W.H. Butler. Full-potential Korringa-Kohn-Rostoker band theory applied to the Mathieu potential. *Phys. Rev. B*, 42:10976–10982, 1990.