BIT

Error estimation based on locally weighted defect for boundary value problems in second order ordinary differential equations

**Winfried Auzinger · Othmar Koch ·
Amir Saboor Bagherzadeh**

Abstract We investigate efficient asymptotically correct a posteriori error estimates for the numerical approximation of two-point boundary value problems for second order ordinary differential equations by piecewise polynomial collocation methods. Our error indicators are based on the defect of the collocation solution with respect to an associated exact difference scheme and backsolving using a cheap, low order finite-difference scheme. We prove high asymptotical correctness of this error indicator and illustrate the theoretical analysis by numerical examples.

W. Auzinger (✉) · O. Koch
Institute for Analysis and Scientific Computing, Vienna University of Technology,
Wiedner Hauptstrasse 8–10/E101, 1040 Vienna, Austria
e-mail: w.auzinger@tuwien.ac.at

O. Koch
e-mail: othmar@othmar-koch.org

A. Saboor Bagherzadeh
Institute of Structural Mechanics, Bauhaus University, Marienstrasse 15,
99423 Weimar, Germany
e-mail: saboorbagherzadeh.a@gmail.com

1 Introduction

This paper is devoted to the efficient and accurate a posteriori error estimation for numerical approximations to boundary value problems (BVPs) for second order ordinary differential equations (ODEs). We consider the class of BVPs

$$L\,u(t) = f(t, u(t), u'(t)) =: F(u)(t), \quad a < t < b, \quad L\,u = u'', \qquad (1.1a)$$

with general linear boundary conditions

$$B_a \cdot \begin{bmatrix} u(a) \\ u'(a) \end{bmatrix} + B_b \cdot \begin{bmatrix} u(b) \\ u'(b) \end{bmatrix} = \begin{bmatrix} s_0 \\ s_1 \end{bmatrix}, \quad B_a, B_b \in \mathbb{R}^{2 \times 2}. \qquad (1.1b)$$

We assume that the problem is well-posed, with Lipschitz continuous right-hand side and a sufficiently smooth solution $u(t)$.

Problem (1.1) can be cast into the standard form of a system of first-order ODEs, and techniques developed for the first-order case may be applied (cf. e.g. [5]). The alternative is direct discretization which we are considering here, since we are explicitly exploiting it in the design and analysis of a method for estimating the global error. As we shall see, this estimate is of a particularly high accuracy on arbitrary meshes. The present paper is of a theoretical nature; more practical questions concerning adaptive mesh selection and comparison which other established adaptive strategies are not addressed. However, in [9] it has been demonstrated that direct treatment of the second-order problem promises computational advantages particularly with respect to adaptive mesh selection, and the conditioning may also be favorable in some situations, see [13].

A preliminary version of the respective analysis has been given in the PhD thesis by the third author [18]. In particular, we consider asymptotically correct global error estimators, designed for the purpose of practical mesh adaptation, for given piecewise polynomial collocation solutions. This continues the work from [5,6,15], where related techniques for first order ODE systems were analyzed. Our particular focus on collocation approximations is motivated by their satisfactory and robust performance for large classes of ODE problems, see for instance [1–3,14].

A collocation solution is a piecewise polynomial function the defect of which (i.e., its pointwise residual with respect to the given ODE) is well-defined and vanishes at the collocation nodes. The overall behavior of the defect is a quality measure which may be used for mesh adaptation, but it yields no direct indication of the behavior of the global error. In [20] it was proposed, in a general setting, to combine defect computation with a backsolving procedure based on an auxiliary, simple finite-difference scheme (SDS) in order to estimate the global error. The heuristic motivation for this approach is based on the well-known principle of defect correction [20]. But already for first order ODEs it turns out that the precise definition of the defect and its interplay with the auxiliary scheme is essential for the successful performance of defect correction algorithms, see for instance [3,7,8,11,15] and references therein, and [4,5], in particular.

The concept of *exact difference schemes* (EDS) for ODEs is very useful in this context. An EDS is a system of difference equations involving integral terms obtained

by locally integrating the given ODE over a grid, which is exactly satisfied by the solution of the ODE. This serves as the conceptual basis for the design of high order schemes. Detailed material on EDS and their direct or iterative computational realization can for example be found in [8,10,12,16,19]. Here, in contrast, we are using an appropriately designed EDS for defect computation only, i.e., the defect of a given numerical solution will be defined as its residual with respect to the EDS. In practice this is realized via sufficiently accurate quadrature.

1.1 Overview

The analysis of a defect-based global error estimate in the context of first-order nonlinear regular problems was given in [5]. It could be shown that for a collocation method with an $\mathcal{O}(h)$ mesh and stage order $\mathcal{O}(h^m)$, the error of the estimate (the difference between the global error and its estimate) is of order $\mathcal{O}(h^{m+1})$.

In the case of second order ODEs, the analysis of the error estimator is significantly more involved. Therefore, for the details of this analysis, we restrict ourselves to a standard class of scalar second order semilinear ODEs with Dirichlet boundary conditions, see Sect. 3. The focus of our analysis is on the robustness with respect to nonequidistant grids and on the particular high asymptotic quality of the estimator. Indeed, it turns out that the deviation of the error estimator for a collocation method of order m is at least of order $m + 2$ due to a supraconvergence effect[1] and may be even higher in special situations. The scope of applicability of our approach is much wider than the presentation in this paper, however, awaiting closer investigation of more general BVPs; see the discussion in Sect. 7.3.

Remark 1.1 In our analysis we are assuming that (1.1) is well-posed and that the methods considered (a collocation method and a simple auxiliary finite-difference scheme) are stable. We do not explicitly recapitulate standard stability arguments for these schemes which are available in the literature. The stability and convergence arguments as given in [2], for instance, are natural in the sense that the stability of the discrete schemes is directly related to the conditioning of the given BVP; see [2, section 5].

However, in order to obtain sharp estimates for the deviation of the advocated global error estimate, we will make use of refined finite-difference stability estimates based on the structure of Green's function of the (linearized) BVP, see Sect. 3.3.

2 Error estimation based on EDS defect: general setting

2.1 Design of an a posteriori error estimate

In an abstract, informal setting, the defect-based error estimation procedure can be described as follows:[2]

---

[1] See Sect. 3.3 for an explanation of the notion of supraconvergence.

[2] In this introductory section, we refrain from denoting continuous and discrete objects in different styles.

– Let

$$L\,u = F(u) \tag{2.1a}$$

represent the given ODE problem, with exact solution u. Here, L is to be identified with the leading, explicit linear part of the given differential operator.

– An EDS is a discrete version of (2.1a) in the form of a set of finite-difference equations

$$L_\Delta u = \mathcal{I}(F(u)), \tag{2.1b}$$

which is exactly satisfied by the solution u of (2.1a). Here, L_Δ is a discretization of L over a given grid Δ, and $\mathcal{I}(\cdot)$ is a certain linear functional typically defined via weighted local integral means of $F(u)$. For computational realization, $\mathcal{I}(\cdot)$ is replaced by a higher-order quadrature approximation $\mathcal{Q}(\cdot)$.

– Furthermore, let $\widehat{u}$ be defined as the solution of an auxiliary (discrete) problem

$$L_\Delta \widehat{u} = \widehat{F}(\widehat{u}) \tag{2.2}$$

which will be chosen as a simple, stable, low-order finite-difference approximation (SDS) to (2.1a) over a grid Δ. Note that L_Δ in (2.1b) and (2.2) are assumed to be identical.

Now, for a given approximation $\widetilde{u} \approx u$ (in our case computed by a high-order collocation method), we consider its (discrete) defect

$$d := L_\Delta \widetilde{u} - \mathcal{Q}(F(\widetilde{u})) \approx L_\Delta \widetilde{u} - \mathcal{I}(F(\widetilde{u})), \tag{2.3a}$$

such that $\widetilde{u}$ is an exact solution of the 'neighboring problem'

$$L_\Delta \widetilde{u} = \mathcal{Q}(F(\widetilde{u})) + d. \tag{2.3b}$$

We approximate (2.3b) by its SDS analog,

$$L_\Delta \widehat{\widetilde{u}} = \widehat{F}(\widehat{\widetilde{u}}) + d. \tag{2.4}$$

For sufficiently small d, we expect

$$\widehat{\widetilde{u}} - \widetilde{u} \approx \widehat{u} - u, \tag{2.5a}$$

i.e., the known error of the SDS approximation $\widehat{\widetilde{u}}$ of the neighboring problem (2.3b) serves as an estimate for the unknown error of the SDS approximation $\widehat{u}$ to the original problem (2.1) over the grid Δ. Equivalently,

$$\varepsilon := \widehat{\widetilde{u}} - \widehat{u} \approx \widetilde{u} - u =: e, \tag{2.5b}$$

i.e., the left-hand side ε of (2.5b), computed at low cost compared to the typical effort for computing $\widetilde{u}$, provides an estimate for the error $e = \widetilde{u} - u$.

Remark 2.1 In the concrete setting which we consider in the sequel, two different grids are involved (see Sects. 3, 4): a coarse grid Δ^N which corresponds to N collocation intervals, and a fine grid Δ comprising Δ^N together with all interior collocation nodes. The auxiliary scheme (2.2) operates on the fine grid Δ.

2.2 How to analyze the deviation of the error estimate

Let

$$\theta := e - \varepsilon = (\widetilde{u} - u) - (\widehat{\widetilde{u}} - \widehat{u}) \tag{2.6}$$

denote the *deviation* of the error estimate. For the analysis of the asymptotic behavior of θ we will proceed as follows. From (2.1b) and (2.3b),

$$\begin{aligned} L_\Delta e &= \mathcal{Q}(F(\widetilde{u})) + d - \mathcal{I}(F(u)) \\ &= \mathcal{Q}(F(\widetilde{u}) - F(u)) + (\mathcal{Q} - \mathcal{I})(F(u)) + d. \end{aligned} \tag{2.7}$$

On the other hand, from (2.2) and (2.4),

$$L_\Delta \varepsilon = L_\Delta(\widehat{\widetilde{u}} - \widehat{u}) = \widehat{F}(\widehat{\widetilde{u}}) - \widehat{F}(\widehat{u}) + d. \tag{2.8}$$

Subtracting (2.8) from (2.7) the defect cancels out, giving

$$\begin{aligned} L_\Delta \theta &= (\widehat{F}(\widehat{u}) - \widehat{F}(\widehat{\widetilde{u}})) + \mathcal{Q}(F(\widetilde{u}) - F(u)) + (\mathcal{Q} - \mathcal{I})(F(u)) \\ &= (\widehat{F}(\widehat{u}) - \widehat{F}(u)) - (\widehat{F}(\widehat{\widetilde{u}}) - \widehat{F}(\widetilde{u})) \\ &\quad + \mathcal{Q}(F(\widetilde{u}) - F(u)) - (\widehat{F}(\widetilde{u}) - \widehat{F}(u)) + (\mathcal{Q} - \mathcal{I})(F(u)). \end{aligned} \tag{2.9}$$

Relation (2.9) is sort of a nonlinear difference equation of SDS type for θ. In order to show that ε is *asymptotically correct*, i.e., that θ is of higher order than e itself, one uses quasilinearization combined with stability arguments and, in addition, estimates the asymptotic behavior of both contributions to the inhomogeneity in (2.9),

(i) the difference $(\mathcal{Q}F - \widehat{F})(\widetilde{u}) - (\mathcal{Q}F - \widehat{F})(u)$,
(ii) the quadrature error $(\mathcal{Q} - \mathcal{I})(F(u)) = (\mathcal{Q} - \mathcal{I})(L u)$.

Assuming sufficient smoothness of the problem data and u, (ii) indicates the required order for the quadrature formula $\mathcal{Q}_\Delta(\cdot)$. On the other hand, (i) is typically determined by the asymptotic behavior of the error e and some of its lower derivatives. This has to be analyzed in detail with the goal to ensure that (i) is of higher order than the error e itself.

These relations become particularly simple for linear problems, $F(u) = Hu + q$, where H denotes the linear part of the affine operator F. In this case, ε can be directly obtained by solving the SDS

$$L_\Delta \varepsilon = \widehat{H}\varepsilon + d. \tag{2.10}$$

Using an analogous notation for $\widehat{F}_\Delta$ we obtain from (2.9):

$$L_\Delta \theta = \widehat{H}\theta + (\mathcal{Q}(He) - \widehat{H}e) + (\mathcal{Q} - \mathcal{I})(F(u)). \tag{2.11}$$

The analysis in Sect. 6 will be based on (2.11) and (2.9), respectively.

Remark 2.2 Recapitulating these considerations, with the 'classical', pointwise defect $L\widetilde{u} - F(\widetilde{u})$ instead of the modified EDS-defect (2.3a), one observes that an additional term would arise which influences the quality of ε, namely $(L_\Delta - L)(e)$. Typically, this term depends on a higher derivative of the error e, and it will not show the desired asymptotic behavior in typical applications. We also note that for a collocation solution $\widetilde{u}$ the pointwise defect evaluated at the collocation nodes vanishes and thus provides no information.

Another feature of the EDS formulation is that it is robust with respect to nonequidistant grids, e.g., for the case of Gaussian collocation points. See also the discussion in [5,7] for the case of first order ODE systems. In this case the quadrature realization of the EDS is closely related to a higher-order implicit Runge-Kutta scheme. This way of estimating the global error for first-order problems has also been analyzed for the case of singular BVPs, see [6,15], and it is implemented in the MATLAB[3] package sbvp, an adaptive solver for first order boundary value problems, see [3].

3 Second order boundary value problems, SDS and EDS

3.1 Problem class

Considering the problem class (1.1), a complete analysis will be given for the case of Dirichlet boundary conditions

$$u(a) = u_a, \quad u(b) = u_b. \tag{3.1}$$

We will first consider well-posed linear Dirichlet problems

$$u''(t) = c_1(t)u'(t) + c_0(t)u(t) + q(t) =: (Hu)(t) + q(t), \tag{3.2a}$$
$$u(a) = u_a, \quad u(b) = u_b, \tag{3.2b}$$

with bounded, smooth data functions. It will then be shown how this analysis extends to nonlinear problems (1.1a). The treatment of more general boundary conditions (1.1b) will be briefly discussed from the algorithmic point of view in Sect. 7.

For $v \in C[a, b]$, as usual we denote

$$\|v\|_\infty := \max_{a \le t \le b} |v(t)|.$$

---

[3] MATLAB is a trademark of The MathWorks, Inc. sbvp is available at www.mathworks.de/matlabcentral/fileexchange/1464-sbvp-1-0-package.

In the following subsections we introduce some further notation and briefly recall relevant facts about finite-difference approximations to be used later on.

3.2 Simple finite-difference scheme (SDS)

Consider a grid

$$\Delta := \left(a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\right) \subseteq [a, b]. \tag{3.3}$$

For an interior grid point t_ℓ, we denote

$$\delta_{\ell-\frac{1}{2}} := t_\ell - t_{\ell-1}, \quad \delta_{\ell+\frac{1}{2}} := t_{\ell+1} - t_\ell, \quad \overline{\delta}_\ell := \frac{1}{2}\left(\delta_{\ell-\frac{1}{2}} + \delta_{\ell+\frac{1}{2}}\right). \tag{3.4a}$$

Then,

$$\delta_{\ell-\frac{1}{2}} = \alpha_\ell\,\overline{\delta}_\ell, \quad \delta_{\ell+\frac{1}{2}} = \beta_\ell\,\overline{\delta}_\ell, \quad \text{with} \quad \alpha_\ell + \beta_\ell = 2. \tag{3.4b}$$

Furthermore we denote

$$t_{\ell\pm\frac{1}{2}} = t_\ell \pm \frac{1}{2}\delta_{\ell\pm\frac{1}{2}}, \quad \ell = 1 \ldots n - 1. \tag{3.4c}$$

We assume that Δ is quasiuniform, i.e.,

$$h := \max_{1\leq\ell\leq n-1} \overline{\delta}_\ell \leq K \cdot \min_{1\leq\ell\leq n-1} \overline{\delta}_\ell \tag{3.5}$$

with a moderate-sized constant K, see [2, Assunption 5.92]. On Δ, we define corresponding grid functions

$$v_\Delta = \left(v_0, \ldots, v_n\right).$$

The max-norm on the space of grid functions is denoted by

$$\|v_\Delta\|_\Delta := \max_{0\leq\ell\leq n} |v_\ell|.$$

For a continuous function $v \in C[a, b]$, we denote by R_Δ its pointwise projection onto the space of grid functions,

$$R_\Delta(v) := \left(v(t_0), \ldots, v(t_n)\right).$$

We also denote $v(t_\ell)$ simply by v_ℓ, and $f(t_\ell, \cdots)$ by $f_\ell(\cdots)$. The second difference quotient

$$(L_\Delta v)_\ell := \frac{\frac{v_{\ell+1}-v_\ell}{\delta_{\ell+\frac{1}{2}}} - \frac{v_\ell-v_{\ell-1}}{\delta_{\ell-\frac{1}{2}}}}{\overline{\delta}_\ell} = \frac{\beta_\ell\,v_{\ell-1} - 2\,v_\ell + \alpha_\ell\,v_{\ell+1}}{\alpha_\ell\beta_\ell\,\overline{\delta}_\ell^{\,2}} \tag{3.6a}$$

is an approximation to $(L\,v)(t_\ell)$. Furthermore, let

$$(\partial_\Delta v)_\ell := \frac{v_{\ell+1} - v_{\ell-1}}{2\,\overline{\delta}_\ell} \tag{3.6b}$$

denote the centered difference approximation to $v'(t_\ell)$. Both (3.6a) and (3.6b) are first-order accurate, in the equidistant case they are second-order accurate.

The standard compact, centered finite-difference approximation (SDS) to (1.1a) reads

$$(L_\Delta \widehat{u}_\Delta)_\ell = f_\ell\big(\widehat{u}_\ell, (\partial_\Delta \widehat{u}_\Delta)_\ell\big) =: \widehat{F}_\Delta(\widehat{u}_\Delta)_\ell, \quad \ell = 1 \ldots n-1. \tag{3.7}$$

Solution of (3.7) subject to boundary conditions (3.2b) provides a grid function

$$\widehat{u}_\Delta = \big(\widehat{u}_0, \widehat{u}_1, \ldots, \widehat{u}_{n-1}, \widehat{u}_n\big) \tag{3.8}$$

with $\widehat{u}_0 = u_a$, $\widehat{u}_n = u_b$, and $\widehat{u}_\ell \approx u(t_\ell)$.

For the linear BVP (3.2), the SDS (3.7) takes the form

$$(L_\Delta \widehat{u}_\Delta)_\ell = (\widehat{H}_\Delta \widehat{u}_\Delta)_\ell + q_\ell, \quad \ell = 1 \ldots n-1, \tag{3.9a}$$

where $\widehat{H}_\Delta v_\Delta$ denotes the centered discretization of $H(v)(t_\ell)$, i.e.,

$$\big(\widehat{H}_\Delta v_\Delta\big)_\ell = c_1(t_\ell)(\partial_\Delta v_\Delta)_\ell + c_0(t_\ell)v_\ell. \tag{3.9b}$$

We assume that the SDS discretization is stable with respect to $\|\cdot\|_\Delta$ for sufficiently small h. Typically, stability depends on Δ in an uncritical way under the assumption of a quasiuniform grid Δ.[4]

3.3 Discrete Green's function and supraconvergence of the SDS

Since (3.6a) and (3.6b) are only first-order accurate in the nonequidistant case, the local truncation error of the SDS (3.7) is only $\mathcal{O}(h)$. However, it is well-known that the global order of the resulting approximation is again $\mathcal{O}(h^2)$, a phenomenon called supraconvergence. This has been analyzed in [17], see also references therein.

Here we do not repeat a proof of supraconvergence in detail, but we list the essential technical tools. Namely, in Sect. 6 we will rely on related techniques; in particular, we will show that the global error estimate to be defined in Sect. 5 is also 'supra-accurate'.

We will make use of the following facts. Consider the simple Dirichlet problem

$$u''(t) = q(t), \quad a < t < b, \qquad u(a) = u(b) = 0, \tag{3.10a}$$

---

[4] This remark applies to several related estimates, as for instance (3.13c), (3.16) below, and to all error estimates obtained.

with solution

$$u(t) = \int_a^b G(t, \tau) \, q(\tau) \, d\tau, \quad a \le t \le b, \tag{3.10b}$$

where $G(t, \tau)$ is Green's function

$$G(t, \tau) = \begin{cases} \dfrac{(b - t)(a - \tau)}{b - a}, & a \le \tau \le t \le b, \\ \dfrac{(b - \tau)(a - t)}{b - a}, & a \le t \le \tau \le b. \end{cases} \tag{3.10c}$$

The proof of the following lemma is straightforward.

Lemma 3.1 *Consider an arbitrary interior grid point t_ℓ and let $g_{\Delta;\ell} := G(\cdot, t_\ell)$, i.e., $(g_{\Delta;\ell})_k = G(t_k, t_\ell)$. Then,*

$$(L_\Delta \, g_{\Delta;\ell})_k = \begin{cases} 0, & k \ne \ell, \\ \frac{1}{\bar\delta_\ell}, & k = \ell. \end{cases}$$

Discretization of (3.10a) according to (3.6a), i.e.,

$$(L_\Delta \widehat{u}_\Delta)_\ell = q_\ell, \quad \ell = 1 \ldots n - 1, \tag{3.11}$$

leads to a system of $n - 1$ difference equations for $\widehat{u}_\Delta$, with $\widehat{u}_0 = \widehat{u}_n = 0$. As a consequence of Lemma 3.1, the inverse of this discrete system, i.e., its discrete Green's function, is the exact pointwise restriction of $G(t, \tau)$ onto $\Delta \times \Delta$. The following discrete solution representation is a direct consequence of this fact.

Lemma 3.2 *The unique solution $\widehat{u}_\Delta$ of (3.11) is given by*

$$\widehat{u}_\ell = L_\Delta^{-1} R_\Delta(q) = \sum_{k=1}^{n-1} G(t_\ell, t_k) \, \bar\delta_k \, q_k, \quad \ell = 1 \ldots n - 1. \tag{3.12}$$

In the remainder of this section we discuss the solution structure of an SDS with a special inhomogeneity, which provides the technical basis for supraconvergence proofs and for the analysis to be given in Sect. 6. In particular, we consider

$$(L_\Delta \check{v}_\Delta)_\ell = \frac{1}{\bar\delta_\ell} \left(z_{\ell+\frac{1}{2}} - z_{\ell-\frac{1}{2}} \right) + r_\ell, \quad \ell = 1 \ldots n - 1, \tag{3.13a}$$

with given z_k and r_ℓ. Using Lemma 3.2 and applying partial summation, the solution $\check{v}_\Delta$ of (3.13a) with $\check{v}_0 = \check{v}_n = 0$, can be written as

$$\check{v}_\ell = \sum_{k=0}^{n-1} \left(G(t_\ell, t_k) - G(t_\ell, t_{k+1}) \right) z_{k+\frac{1}{2}} + \sum_{k=0}^{n-1} G(t_\ell, t_k) \, \bar\delta_k \, r_k. \tag{3.13b}$$

Since $G(t, \tau)$ is bounded and smooth away from the diagonal $t = \tau$, this shows that an estimate

$$\|\check{v}_\Delta\|_\Delta \leq C \left(\max_{0 \leq k \leq n-1} |z_{k+\frac{1}{2}}| + \max_{1 \leq k \leq n-1} |r_k| \right) \tag{3.13c}$$

is valid. Furthermore, from (3.13b) we obtain

$$(\partial_\Delta \check{v}_\Delta)_\ell = \sum_{k=0}^{n-1} \left(\partial_\Delta^{[2]} G\right)_{\ell,k} z_{k+\frac{1}{2}} + \left(\partial_\Delta^{[1]} G\right)_{\ell,k} r_k, \tag{3.14}$$

with

$$\left(\partial_\Delta^{[1]} G\right)_{\ell,k} = \frac{1}{2\,\bar{\delta}_\ell} \left(G(t_{\ell+1}, t_k) - G(t_{\ell-1}, t_k) \right) = \begin{cases} \mathcal{O}(h), & k \neq \ell, \\ \mathcal{O}(1), & k = \ell, \end{cases} \tag{3.15a}$$

and

$$\begin{aligned} \left(\partial_\Delta^{[2]} G\right)_{\ell,k} &= \frac{1}{2\,\bar{\delta}_\ell} \left(G(t_{\ell+1}, t_k) - G(t_{\ell-1}, t_k) \right) \\ &- \frac{1}{2\,\bar{\delta}_\ell} \left(G(t_{\ell+1}, t_{k+1}) - G(t_{\ell-1}, t_{k+1}) \right) = \begin{cases} \mathcal{O}(h), & k \notin \{\ell-1, \ell\}, \\ \mathcal{O}(1), & k \in \{\ell-1, \ell\}. \end{cases} \end{aligned} \tag{3.15b}$$

This shows that there exists a constant C such that we also have

$$\|\partial_\Delta \check{v}_\Delta\|_\Delta \leq C \left(\max_{0 \leq k \leq n-1} \left| z_{k+\frac{1}{2}} \right| + \max_{1 \leq k \leq n-1} |r_k| \right). \tag{3.16}$$

For a proof of supraconvergence of the SDS one makes use of (3.13), representing the local truncation error in the form of the right-hand side in (3.13a), i.e., as a difference quotient which is $\mathcal{O}(h)$ plus an $\mathcal{O}(h^2)$ term. This is combined with a stability argument for the full scheme.

In a way analogous to (3.13), the solution $\check{v}_\Delta$ of an SDS of the form

$$(L_\Delta \check{v}_\Delta)_\ell = (\partial_\Delta z_\Delta)_\ell + r_\ell = \frac{1}{2\,\bar{\delta}_\ell} \left(z_{\ell+1} - z_{\ell-1} \right) + r_\ell, \quad \ell = 1 \ldots n-1, \tag{3.17a}$$

with $\check{v}_0 = \check{v}_n = 0$, can be written as

$$\check{v}_\ell = \sum_{k=0}^{n-1} \left(G(t_\ell, t_{k-1}) - G(t_\ell, t_{k+1}) \right) z_k + \sum_{k=0}^{n-1} G(t_\ell, t_k) \, \bar{\delta}_k \, r_k, \tag{3.17b}$$

resulting in estimates analogous to (3.13c) and (3.16).

As a result of these considerations, a perturbation argument combined with stability of the full SDS, involving some minor technicalities, implies supraconvergence of $\widehat{u}_\Delta$ and $(\partial_\Delta \widehat{u}_\Delta)_\Delta$, that is,

$$\|\widehat{u}_\Delta - R_\Delta(u)\|_\Delta = \mathcal{O}(h^2) \quad \text{and} \quad \|\partial_\Delta \widehat{u}_\Delta - R_\Delta(u')\|_\Delta = \mathcal{O}(h^2). \qquad (3.18)$$

3.4 Exact finite-difference scheme (EDS)

Definition 3.1 By

$$\widehat{C}_2[t_{\ell-1}, t_\ell, t_{\ell+1}] := \left\{ v \in C^1[t_{\ell-1}, t_{\ell+1}] : v'' \text{ continuous on } [t_{\ell-1}, t_\ell) \cup (t_\ell, t_{\ell+1}], \right.$$

$$\left. \lim_{t \uparrow t_\ell} v''(t) \in \mathbb{R}, \ \lim_{t \downarrow t_\ell} v''(t) \in \mathbb{R} \text{ exist} \right\}$$

we denote the class of functions from $C^1[t_{\ell-1}, t_{\ell+1}]$ for which v'' is allowed to have a jump discontinuity at $t = t_\ell$.

Lemma 3.3 *For $v \in \widehat{C}_2[t_{\ell-1}, t_\ell, t_{\ell+1}]$ and $L_\Delta v$ defined in* (3.6a)*, we have*

$$(L_\Delta v)_\ell = \int_{-\alpha_\ell}^{\beta_\ell} K_\ell(\xi) \, v''(t_\ell + \bar{\delta}_\ell \, \xi) \, d\xi, \qquad (3.19a)$$

with kernel

$$K_\ell(\xi) = \begin{cases} 1 + \frac{\xi}{\alpha_\ell}, & \xi \in [-\alpha_\ell, 0], \\ 1 - \frac{\xi}{\beta_\ell}, & \xi \in [0, \beta_\ell]. \end{cases} \qquad (3.19b)$$

Proof A routine argument based on integration by parts [12, 19]. □

Lemma 3.3 represents a special instance of an identity sometimes called Marchuk or Chawla identity, relating a general second order differential operator to its finite-difference analog via local weighted integral means.

For functions $w = w(t)$ we adopt the denotation

$$\mathcal{I}_\ell(w) := \int_{-\alpha_\ell}^{\beta_\ell} K_\ell(\xi) \, w\big(t_\ell + \bar{\delta}_\ell \, \xi\big) \, d\xi \qquad (3.20)$$

with K_ℓ from (3.19b). By locally integrating the ODE (1.1a), we see that, at all interior grid points t_ℓ, its solution $u(t)$ exactly satisfies the EDS equations

$$(L_\Delta u)_\ell = \mathcal{I}_\ell(f(\cdot, u(\cdot), u'(\cdot))). \qquad (3.21)$$

Later on, (3.21) together with appropriate quadrature formulas $\mathcal{Q}_\ell \approx \mathcal{I}_\ell$ will be used for defining a higher-order EDS-type defect of a given numerical solution, see Sect. 5.

4 Review of collocation methods

We recapitulate classical results for piecewise polynomial collocation methods for second order BVPs (1.1), see [2, Section 5]. For the numerical approximation we define a mesh

$$\Delta^N := \left(a = \tau_0 < \tau_1 < \cdots < \tau_N = b\right) \subseteq [a, b],$$

and set $l_i := \tau_i - \tau_{i-1}$, $i = 1 \ldots N$. Here $l_i = \mathcal{O}(h)$, with h defined in (3.5).

For collocation, m points are inserted in each subinterval $[\tau_i, \tau_{i+1}]$. This yields the (fine) grid[5]

$$\Delta := \left\{ t_{i,j} = \tau_i + \rho_j l_i, \ i = 0 \ldots N - 1, \ j = 0 \ldots m + 1 \right\}, \qquad (4.1a)$$

with given nodes

$$0 =: \rho_0 < \rho_1 < \cdots < \rho_m < \rho_{m+1} := 1. \qquad (4.1b)$$

This grid Δ contains $n + 1 = (m + 1)N + 1$ points and is to be identified with Δ from (3.3), with flattened indexing

$$t_\ell \equiv t_{i,j}, \quad \ell = (m+1)i + j, \quad i = 0 \ldots N - 1, \ j = 0 \ldots m. \qquad (4.2)$$

Definition 4.1 Consider a quasiuniform grid Δ as specified in (4.1a). A continuously differentiable collocation solution $\widetilde{u}(t) := \widetilde{u}_i(t)$, $t \in [\tau_i, \tau_{i+1}]$, $i = 0 \ldots N - 1$, of (1.1) is a piecewise polynomial function of degree $m + 1$, corresponding to m collocation points in each subinterval, which satisfies (1.1a) at the m collocation points in each of the N subintervals, i.e.

$$\widetilde{u}''(t_{i,j}) = f(t_{i,j}, \widetilde{u}(t_{i,j}), \widetilde{u}'(t_{i,j})), \quad i = 0 \ldots N - 1, \ j = 1 \ldots m,$$

and, in addition, the boundary conditions (1.1b). Here, $\widetilde{u}(t)$ and $\widetilde{u}'(t)$ are required to be continuous at the endpoints of the subintervals $[\tau_i, \tau_{i+1}]$, but $\widetilde{u}''(t)$ will have jump discontinuities at the τ_i. Thus, in the neighborhood of τ_i, $\widetilde{u}(t)$ is of the type $\widehat{C}_2[\tau_{i-1}, \tau_i, \tau_{i+1}]$ in the sense of Definition 3.1.

Let

$$e(t) := \widetilde{u}(t) - u(t)$$

---

[5] For convenience, we denote τ_i by $t_{i,0} \equiv t_{i-1,m+1}$, $i = 1 \ldots N - 1$.

denote the collocation error. Under standard assumptions on the given BVP, a collocation solution is stable and convergent (see [2, Chapter 5]). The following theorem states a convergence result for the problem (1.1a). The proof of the following theorem follows from general convergence results, see [2, Sect. 5.6].

Theorem 4.1 *Assume that the given BVP is well-posed, i.e., it has a locally unique and sufficiently smooth solution $u(t)$. For a collocation solution $\widetilde{u}(t)$ according to Definition 4.1, the following uniform estimates hold:*

$$\left\|e^{(\nu)}\right\|_\infty = \left\|\widetilde{u}^{(\nu)} - u^{(\nu)}\right\|_\infty = \mathcal{O}(h^m), \quad \nu = 0, 1, 2. \tag{4.3a}$$

The higher derivatives satisfy

$$\left\|e^{(\nu)}\right\|_\infty = \left\|\widetilde{u}^{(\nu)} - u^{(\nu)}\right\|_\infty = \mathcal{O}(h^{m+2-\nu}), \quad \nu = 3 \ldots m+1. \tag{4.3b}$$

In the special case where m is odd and the nodes ρ_j are symmetrically distributed, i.e., $\rho_j = \rho_{m+1-j}$, the following improved uniform estimates hold:

$$\left\|e^{(\nu)}\right\|_\infty = \left\|\widetilde{u}^{(\nu)} - u^{(\nu)}\right\|_\infty = \mathcal{O}(h^{m+1}), \quad \nu = 0, 1. \tag{4.4}$$

5 Construction of the error estimate

We design an a posteriori estimate for the global error $e(t)$ of a given collocation solution $\widetilde{u}(t)$. The construction principle has been introduced in Sect. 2.1 and it is now applied in our concrete setting. We mainly consider the case of Dirichlet conditions (3.1) where no defect occurs at the boundary. Remarks on handling more general boundary conditions are given in Sect. 7.

Let us describe the procedure and its algorithmic components in detail:

- The given approximation is the collocation solution $\widetilde{u}(t)$, for which the error $e(t) = \widetilde{u}(t) - u(t)$ is to be estimated.
- The (discrete) defect d_Δ of $\widetilde{u}$ is computed by inserting $\widetilde{u}(t)$ into the EDS reformulation of the original problem [see (2.1b)], which has been introduced in Sect. 3.4. In the computational realization, the integral operators $\mathcal{I}_\ell = \mathcal{I}_{i,j}$ are replaced by a set of higher-order quadrature formulas $\mathcal{Q}_\ell = \mathcal{Q}_{i,j}$, see (2.3a) and Sect. 5.1 below.
- The auxiliary scheme (2.2) is the SDS (3.7) operating on the fine grid Δ given by (4.1a). Its solution is denoted by $\widehat{u}_\Delta = (\widehat{u}_0, \widehat{u}_1, \ldots, \widehat{u}_{n-1}, \widehat{u}_n)$, see (3.8).
- Solving the auxiliary scheme with additional inhomogeneity d_Δ [see (2.4)] gives another grid function $\widehat{\widetilde{u}}_\Delta = \left(\widehat{\widetilde{u}}_0, \widehat{\widetilde{u}}_1, \ldots, \widehat{\widetilde{u}}_{n-1}, \widehat{\widetilde{u}}_n\right)$.

With these settings, the (discrete) error estimator ε_Δ is defined according to (2.5b) as

$$\varepsilon_\Delta := \widehat{\widetilde{u}}_\Delta - \widehat{u}_\Delta \approx R_\Delta(e). \tag{5.1}$$

5.1 Quadrature formulas for defect computation

At each interior grid point $t_\ell = t_{i,j} \in \Delta$, the exact EDS defect of $\widetilde{u}$ reads [see (3.21)]

$$(L_\Delta \widetilde{u})_{i,j} - \mathcal{I}_{i,j}(f(\cdot, \widetilde{u}(\cdot), \widetilde{u}'(\cdot)). \tag{5.2}$$

Here,

$$\mathcal{I}_{i,j}(w) = \int_{-\alpha_{i,j}}^{\beta_{i,j}} K_{i,j}(\xi)\, w(t_{i,j} + \overline{\delta}_{i,j}\, \xi)\, d\xi \tag{5.3a}$$

is the integral operator defined in (3.20), with kernel

$$K_{i,j}(\xi) = \begin{cases} 1 + \frac{\xi}{\alpha_{i,j}}, & \xi \in [-\alpha_{i,j}, 0], \\ 1 - \frac{\xi}{\beta_{i,j}}, & \xi \in [0, \beta_{i,j}], \end{cases} \tag{5.3b}$$

from (3.19b). Practical approximation of integrals of the form (5.3) by interpolatory quadrature can be performed in a standard way. We consider two different cases:

– For the case where $t_{i,j} \in (\tau_i, \tau_{i+1})$ is an interior collocation node, we construct interpolatory quadrature formulas $\mathcal{Q}_{i,j} \approx \mathcal{I}_{i,j}$ in the usual way via

$$\mathcal{I}_{i,j}(w) \approx \mathcal{Q}_{i,j}(w) := \int_{-\alpha_{i,j}}^{\beta_{i,j}} K_{i,j}(\xi)\, \pi_i\big(t_{i,j} + \overline{\delta}_{i,j}\, \xi\big)\, d\xi = \sum_{k=0}^{m+1} \omega_{j,k}\, w(t_{i,k}),$$

where π_i is the local Lagrange interpolant of degree $m+1$ to w at the nodes $t_{i,0}, \ldots, t_{i,m+1}$.

– For the case where $t_{i,0} = \tau_i \in (a, b)$ is an endpoint between two neighboring collocation subintervals ($i = 1 \ldots N - 1$), we split the integral into two parts,

$$\mathcal{I}_{i,j}(w) = \mathcal{I}_{i,0}^-(w) + \mathcal{I}_{i,0}^+(w) :=$$

$$= \int_{-\alpha_{i,0}}^{0} \left(1 + \frac{\xi}{\alpha_{i,0}}\right) w(\tau_i + \overline{\delta}_{i,0}\, \xi)\, d\xi + \int_{0}^{\beta_{i,0}} \left(1 - \frac{\xi}{\beta_{i,0}}\right) w\big(\tau_i + \overline{\delta}_{i,0}\, \xi\big)\, d\xi,$$

and we separately approximate the two terms using data from the left and right subintervals, respectively, in the same way as before. This gives

$$\mathcal{I}_{i,0}(w) \approx \mathcal{Q}_{i,0}(w) = \mathcal{Q}_{i,0}^-(w) + \mathcal{Q}_{i,0}^+(w)$$

$$= \sum_{k=0}^{m+1} \omega_{0,k}^-\, w(t_{i-1,k}) + \sum_{k=0}^{m+1} \omega_{0,k}^+\, w(t_{i,k})$$

for $i = 1 \ldots N - 1$.

Since the distribution of the collocation nodes is fixed [defined via $\rho_j \in [0, 1]$, see (4.1)], the coefficients $\omega_{j,k}^{[\pm]}$ of these quadrature formulas do not depend on i and can be computed in a preprocessing phase. In this way, the defect of $\widetilde{u}(t)$ is defined by the computable object

$$d_{i,j} := (L_\Delta \widetilde{u})_{i,j} - \mathcal{Q}_{i,j}(f(\cdot, \widetilde{u}(\cdot), \widetilde{u}'(\cdot)), \tag{5.4}$$

at all interior grid points $t_{i,j} \in \Delta$.

Standard arguments show

$$\mathcal{Q}_{i,j}(w) - \mathcal{I}_{i,j}(w) = \mathcal{O}(h^{m+2}), \tag{5.5a}$$

for sufficiently smooth $w = w(t)$. In the special case where m is odd and the nodes ρ_j are symmetrically distributed, i.e., $\rho_j = \rho_{m+1-j}$, we have

$$\mathcal{Q}_{i,j}(w) - \mathcal{I}_{i,j}(w) = \mathcal{O}(h^{m+3}). \tag{5.5b}$$

In particular, we shall make use of (5.5) for $w = u''$.

6 Analysis of the error estimate

For simplicity of notation we write e, ε, θ, etc. instead of $R_\Delta(e)$, ε_Δ, θ_Δ, etc., and we use the flattened index notation (4.2).

For the deviation $\theta = e - \varepsilon$, an estimate of the form $\|\theta\|_\Delta = \|e - \varepsilon\|_\Delta = \mathcal{O}(h^{m+1})$ can be derived in a similar manner as in [5] for the case of first-order systems. Here, the asymptotic behavior of the collocation error described in Theorem 4.1 is essential. In the sequel we prove that the error estimate is even supra-accurate, i.e., $\|\theta\|_\Delta = \|e - \varepsilon\|_\Delta = \mathcal{O}(h^{m+2})$ is valid. In particular, this implies that the error estimate is asymptotically correct also in the case where already $e = \mathcal{O}(h^{m+1})$ holds, see Theorem 4.1.

Remark 6.1 We believe that for the case where $e = \mathcal{O}(h^{m+1})$, i.e., for m odd and symmetrically distributed ρ_j, the estimates (4.4) and (5.5b) can be exploited in order to show that even $\|\theta\|_\Delta = \mathcal{O}(h^{m+3})$ is valid, which is also observed in numerical experiments, see Sect. 7. This would involve further technicalities, in particular an extension of Lemma 6.1 below, and coping with the reduced order $e'' = \mathcal{O}(h^m)$. This would however exceed the scope of the present work.

For the sake of clarity we consider the linear case first. The proof for the nonlinear case in Sect. 6.2 is an extension of the linear version, based on quasilinearization and supraconvergence arguments.

6.1 The linear case

We first derive a representation for the leading term in the inhomogeneity of the linear difference equation (2.11) for θ in the form of a difference quotient plus a higher-order term.

Lemma 6.1 *For the linear BVP* (3.2), *at all grid points t_ℓ we have*

$$
\begin{aligned}
&\mathcal{I}_\ell(He) - (\widehat{H}_\Delta e)_\ell \\
&= \frac{1}{\overline{\delta}_\ell} \left(\delta_{\ell+\frac{1}{2}}\, z^{[1]}_{\ell+\frac{1}{2}} - \delta_{\ell-\frac{1}{2}}\, z^{[1]}_{\ell-\frac{1}{2}} \right) + \frac{1}{\overline{\delta}_\ell} \left(\delta_{\ell+\frac{1}{2}}\, z^{[0]}_{\ell+\frac{1}{2}} - \delta_{\ell-\frac{1}{2}}\, z^{[0]}_{\ell-\frac{1}{2}} \right) + \mathcal{O}(h^{m+2}).
\end{aligned}
\tag{6.1}
$$

Here,

$$
z^{[0]}_{\ell\pm\frac{1}{2}} = \delta_{\ell\pm\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\tfrac{\zeta^2}{2} + \tfrac{1}{8} \right) (c_0\, e)' \left(t_{\ell\pm\frac{1}{2}} + \zeta\, \delta_{\ell\pm\frac{1}{2}} \right) d\zeta = \mathcal{O}(h^{m+1}),
\tag{6.2a}
$$

$$
\begin{aligned}
z^{[1]}_{\ell\pm\frac{1}{2}} = {}&-c_1 \left(t_{\ell\pm\frac{1}{2}} \right) \int_{-\frac{1}{2}}^{\frac{1}{2}} \zeta\, e' \left(t_{\ell\pm\frac{1}{2}} + \zeta\, \delta_{\ell\pm\frac{1}{2}} \right) d\zeta \\
&-\delta_{\ell\pm\frac{1}{2}}\, c_1' \left(t_{\ell\pm\frac{1}{2}} \right) \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\zeta^2 - \tfrac{1}{4} \right) e' \left(t_{\ell\pm\frac{1}{2}} + \zeta\, \delta_{\ell\pm\frac{1}{2}} \right) d\zeta = \mathcal{O}(h^{m+1}).
\end{aligned}
\tag{6.2b}
$$

Proof By definition of H and $\widehat{H}_\Delta$ [see (3.2), (3.9b)],

$$
\mathcal{I}_\ell(He) - (\widehat{H}_\Delta e)_\ell = \underbrace{\left(\mathcal{I}_\ell(c_1\, e') - c_1(t_\ell)(\partial_\Delta e)_\ell \right)}_{\mathbf{D}_1} + \underbrace{\left(\mathcal{I}_\ell(c_0\, e) - (c_0\, e)(t_\ell) \right)}_{\mathbf{D}_0}.
\tag{6.3}
$$

Representation and estimation of $\mathbf{D}_0$. By definition of $\mathcal{I}_\ell$ [see (3.20)] and with $\alpha_\ell + \beta_\ell = 2$,

$$
\begin{aligned}
\mathbf{D}_0 &= \int_{-\alpha_\ell}^{\beta_\ell} K_\ell(\xi)\, (c_0\, e)\big(t_\ell + \xi\, \overline{\delta}_\ell\big)\, d\xi - (c_0\, e)(t_\ell) \\
&= \underbrace{\int_{-\alpha_\ell}^{0} \left(1 + \frac{\xi}{\alpha_\ell} \right) (c_0\, e)\big(t_\ell + \xi\, \overline{\delta}_\ell\big)\, d\xi - \frac{1}{2}\, \alpha_\ell\, (c_0\, e)(t_\ell)}_{\mathbf{I}_0^-}
\end{aligned}
$$

$$+ \int_0^{\beta_\ell} \left(1 - \frac{\xi}{\beta_\ell}\right) (c_0\, e)\left(t_\ell + \xi\, \overline{\delta}_\ell\right) d\xi - \frac{1}{2}\, \beta_\ell\, (c_0\, e)(t_\ell).$$

$$\underbrace{\phantom{\int_0^{\beta_\ell} \left(1 - \frac{\xi}{\beta_\ell}\right) (c_0\, e)\left(t_\ell + \xi\, \overline{\delta}_\ell\right) d\xi - \frac{1}{2}\, \beta_\ell\, (c_0\, e)(t_\ell).}}_{\mathbf{I}_0^+}$$

In $\mathbf{I}_0^-$ we substitute $\frac{1}{2} + \frac{\xi}{\alpha_\ell} = \zeta$ and integrate by parts to obtain

$$\mathbf{I}_0^- = \alpha_\ell \left(\frac{3}{8}\, (c_0\, e)(t_\ell) + \frac{1}{8}\, (c_0\, e)(t_{\ell-1}) \right.$$

$$\left. - \delta_{\ell-\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\zeta^2}{2} + \frac{\zeta}{2}\right) (c_0\, e)' \left(t_{\ell-\frac{1}{2}} + \zeta\, \delta_{\ell-\frac{1}{2}}\right) d\zeta \right)$$

$$= \alpha_\ell \left(\frac{1}{2}\, (c_0\, e)(t_{\ell-1}) - \delta_{\ell-\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\zeta^2}{2} + \frac{\zeta}{2} - \frac{3}{8}\right) (c_0\, e)' \left(t_{\ell-\frac{1}{2}} + \zeta\, \delta_{\ell-\frac{1}{2}}\right) d\zeta \right),$$

hence

$$\mathbf{I}_0^- - \frac{1}{2}\, \alpha_\ell\, (c_0\, e)(t_\ell) = -\alpha_\ell\, \delta_{\ell-\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\zeta^2}{2} + \frac{\zeta}{2} + \frac{1}{8}\right) (c_0\, e)' \left(t_{\ell-\frac{1}{2}} + \zeta\, \delta_{\ell-\frac{1}{2}}\right) d\zeta$$

$$= -\alpha_\ell\, z_{\ell-\frac{1}{2}}^{[0]} - \alpha_\ell\, \delta_{\ell-\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\zeta}{2}\, (c_0\, e)' \left(t_{\ell-\frac{1}{2}} + \zeta\, \delta_{\ell-\frac{1}{2}}\right) d\zeta,$$

with $z_{\ell-\frac{1}{2}}^{[0]}$ defined in (6.2a). Here, $z_{\ell-\frac{1}{2}}^{[0]} = \mathcal{O}(h^{m+1})$ follows from $e' = \mathcal{O}(h^m)$, and after integration by parts,

$$\delta_{\ell-\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\zeta}{2}\, (c_0\, e)' \left(t_{\ell-\frac{1}{2}} + \zeta\, \delta_{\ell-\frac{1}{2}}\right) d\zeta = \mathcal{O}(h^{m+2})$$

follows from $e'' = \mathcal{O}(h^m)$ [see (4.3a)].

For $\mathbf{I}_0^+$, substituting $-\frac{1}{2} + \frac{\xi}{\beta_\ell} = \zeta$ we obtain in an analogous way

$$\mathbf{I}_0^+ - \frac{1}{2}\, \beta_\ell\, (c_0\, e)(t_\ell) = \beta_\ell\, z_{\ell+\frac{1}{2}}^{[0]} - \beta_\ell\, \delta_{\ell+\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\zeta}{2}\, (c_0\, e)' \left(t_{\ell+\frac{1}{2}} + \zeta\, \delta_{\ell+\frac{1}{2}}\right) d\zeta,$$

with analogous estimates. Altogether, this shows

$$\mathbf{D}_0 = \beta_\ell \, z_{\ell+\frac{1}{2}}^{[0]} - \alpha_\ell \, z_{\ell-\frac{1}{2}}^{[0]} + \mathcal{O}(h^{m+2})$$

$$= \frac{1}{\delta_\ell} \left(\delta_{\ell+\frac{1}{2}} \, z_{\ell+\frac{1}{2}}^{[0]} - \delta_{\ell-\frac{1}{2}} \, z_{\ell-\frac{1}{2}}^{[0]} \right) + \mathcal{O}(h^{m+2}), \qquad z_{\ell\pm\frac{1}{2}}^{[0]} = \mathcal{O}(h^{m+1}). \quad (6.4a)$$

Representation and estimation of $\mathbf{D}_1$. By definition of $\mathcal{I}_\ell$ (see (5.3)) and with $\alpha_\ell + \beta_\ell = 2$,

$$\mathbf{D}_1 = \mathcal{I}_\ell(c_1 \, e') - c_1(t_\ell)(\partial_\Delta e)_\ell$$

$$= \int_{-\alpha_\ell}^{\beta_\ell} K_\ell(\xi) \, (c_1 \, e')(t_\ell + \xi \, \overline{\delta}_\ell) \, d\xi - \frac{1}{2} \, c_1(t_\ell) \int_{-\alpha_\ell}^{\beta_\ell} e'(t_\ell + \xi \, \overline{\delta}_\ell) \, d\xi$$

$$= \int_{-\alpha_\ell}^{\beta_\ell} \left(K_\ell(\xi) \, c_1(t_\ell + \xi \, \overline{\delta}_\ell) - \frac{1}{2} \, c_1(t_\ell) \right) e'(t_\ell + \xi \, \overline{\delta}_\ell) \, d\xi$$

$$= \underbrace{\int_{-\alpha_\ell}^{0} \left((1 + \frac{\xi}{\alpha_\ell}) \, c_1(t_\ell + \xi \, \overline{\delta}_\ell) - \frac{1}{2} \, c_1(t_\ell) \right) e'(t_\ell + \xi \, \overline{\delta}_\ell) \, d\xi}_{\mathbf{I}_1^-}$$

$$+ \underbrace{\int_{0}^{\beta_\ell} \left((1 - \frac{\xi}{\beta_\ell}) \, c_1(t_\ell + \xi \, \overline{\delta}_\ell) - \frac{1}{2} \, c_1(t_\ell) \right) e'(t_\ell + \xi \, \overline{\delta}_\ell) \, d\xi}_{\mathbf{I}_1^+}.$$

For $\mathbf{I}_1^-$ we substitute $\frac{1}{2} + \frac{\xi}{\alpha_\ell} = \zeta$ to obtain

$$\mathbf{I}_1^- = \alpha_\ell \int_{-\frac{1}{2}}^{\frac{1}{2}} \left((\zeta + \frac{1}{2}) \, c_1 \left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}} \right) \right.$$

$$\left. - \frac{1}{2} \, c_1 \left(t_{\ell-\frac{1}{2}} + \frac{1}{2} \, \delta_{\ell-\frac{1}{2}} \right) \right) e' \left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}} \right) \, d\zeta.$$

By Taylor expansion of $c_1(t)$ about $t = t_{\ell-\frac{1}{2}}$,

$$\left(\zeta + \frac{1}{2} \right) c_1 \left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}} \right) - \frac{1}{2} c_1 \left(t_{\ell-\frac{1}{2}} + \frac{1}{2} \, \delta_{\ell-\frac{1}{2}} \right)$$

$$= \zeta \, c_1 \left(t_{\ell-\frac{1}{2}} \right) + \delta_{\ell-\frac{1}{2}} (\zeta^2 + \frac{\zeta}{2} - \frac{1}{4}) \, c_1' \left(t_{\ell-\frac{1}{2}} \right) + \mathcal{O}(h^2),$$

and using analogous arguments as for $\mathbf{D}_0$ above we obtain

$$
\mathbf{I}_1^- = \alpha_\ell \, c_1(t_{\ell-\frac{1}{2}}) \underbrace{\int_{-\frac{1}{2}}^{\frac{1}{2}} \zeta \, e'\left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}}\right) d\zeta}_{=\mathcal{O}(h^{m+1})}
$$

$$
+\alpha_\ell \, \delta_{\ell-\frac{1}{2}} \, c_1'\left(t_{\ell-\frac{1}{2}}\right) \underbrace{\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\zeta^2 - \frac{1}{4}\right) e'\left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}}\right) d\zeta}_{=\mathcal{O}(h^m)}
$$

$$
+\alpha_\ell \, \delta_{\ell-\frac{1}{2}} \, c_1'\left(t_{\ell-\frac{1}{2}}\right) \underbrace{\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\zeta}{2} e'\left(t_{\ell-\frac{1}{2}} + \zeta \, \delta_{\ell-\frac{1}{2}}\right) d\zeta}_{=\mathcal{O}(h^{m+1})} + \mathcal{O}(h^{m+2}),
$$

i.e.,

$$
\mathbf{I}_1^- = -\alpha_\ell \, z_{\ell-\frac{1}{2}}^{[1]} + \mathcal{O}(h^{m+2}),
$$

with $z_{\ell-\frac{1}{2}}^{[1]}$ defined in (6.2b), which is indeed seen to be $\mathcal{O}(h^{m+1})$.

For $\mathbf{I}_1^+$, substituting $-\frac{1}{2} + \frac{\xi}{\beta_\ell} = \zeta$ we obtain in an analogous way

$$
\mathbf{I}_1^+ = \beta_\ell \, z_{\ell+\frac{1}{2}}^{[1]} + \mathcal{O}(h^{m+2}),
$$

with $z_{\ell+\frac{1}{2}}^{[1]} = \mathcal{O}(h^{m+1})$ defined in (6.2b). Altogether, this shows

$$
\mathbf{D}_1 = \beta_\ell \, z_{\ell+\frac{1}{2}}^{[1]} - \alpha_\ell \, z_{\ell-\frac{1}{2}}^{[1]} + \mathcal{O}(h^{m+2})
$$
$$
= \frac{1}{\delta_\ell} \left(\delta_{\ell+\frac{1}{2}} \, z_{\ell+\frac{1}{2}}^{[1]} - \delta_{\ell-\frac{1}{2}} \, z_{\ell-\frac{1}{2}}^{[1]} \right) + \mathcal{O}(h^{m+2}), \quad z_{\ell\pm\frac{1}{2}}^{[1]} = \mathcal{O}(h^{m+1}). \tag{6.4b}
$$

Combining (6.3) with (6.4a) and (6.4b) completes the proof. □

Theorem 6.1 *For the linear BVP* (3.2)*, the deviation of the error estimate satisfies*

$$
\|\theta\|_\Delta = \|e - \varepsilon\|_\Delta = \mathcal{O}(h^{m+2}).
$$

Proof According to (2.11), the deviation $\theta = e - \varepsilon = (\widetilde{u} - u) - (\widehat{\widetilde{u}} - \widehat{u})$ satisfies the SDS-type difference equation

$$(L_\Delta \theta)_\ell = (\widehat{H}_\Delta \theta)_\ell + \big(\mathcal{Q}_\ell(He) - (\widehat{H}_\Delta e)_\ell\big) + (\mathcal{Q}_\ell - \mathcal{I}_\ell)(F(u))$$

$$= (\widehat{H}_\Delta \theta)_\ell + \underbrace{\big(\mathcal{I}_\ell(He) - (\widehat{H}_\Delta e)_\ell\big)}_{\mathbf{J}_1} + \underbrace{(\mathcal{Q}_\ell - \mathcal{I}_\ell)(He)}_{\mathbf{J}_2} + \mathcal{O}(h^{m+2}), \quad (6.5)$$

for $\ell = 1 \ldots n-1$, with homogeneous boundary conditions. Here $(\mathcal{Q}_\ell - \mathcal{I}_\ell)(F(u)) = (\mathcal{Q}_\ell - \mathcal{I}_\ell)(L u)$ is a higher-order quadrature error, see Sect. 5.1.

Representation and estimation of $\mathbf{J}_1$ *and* $\mathbf{J}_2$. According to Lemma 6.1,

$$\mathbf{J}_1 = \frac{1}{\delta_\ell}\left(z_{\ell+\frac{1}{2}} - z_{\ell-\frac{1}{2}}\right) + \mathcal{O}(h^{m+2}), \quad (6.6a)$$

with

$$z_{\ell\pm\frac{1}{2}} = \delta_{\ell\pm\frac{1}{2}}\left(z_{\ell\pm\frac{1}{2}}^{[1]} + z_{\ell\pm\frac{1}{2}}^{[0]}\right) = \mathcal{O}(h^{m+2}), \quad (6.6b)$$

see (6.2). The quadrature error $\mathbf{J}_2$ depends on the $(m + 2)$-nd derivative of $He = c_1 e' + c_0 e$,

$$|\mathbf{J}_2| \le C h^{m+2}, \quad C = C\left((c_1 e' + c_0 e)^{(m+2)}\right).$$

In order to guarantee $\mathbf{J}_2 = \mathcal{O}(h^{m+2})$ we thus have to show that the higher derivatives $e^{(\nu)} = \tilde{u}^{(\nu)} - u^{(\nu)}$ are bounded independently of h up to $\nu = m+3$. Since the degree of $\tilde{u}$ is $m+1$, we have $e^{(\nu)} = -u^{(\nu)}$ for $\nu = m+2, m+3$. For $\nu = 0 \ldots m+1$, the uniform boundedness of $e^{(\nu)}$ immediately follows from (4.3) (Theorem 4.1). Altogether this shows

$$\mathbf{J}_2 = \mathcal{O}(h^{m+2}). \quad (6.7)$$

Estimation of θ. From (6.5)–(6.7),

$$(L_\Delta \theta)_\ell = \big(\widehat{H}_\Delta \theta\big)_\ell + \frac{1}{\delta_\ell}\left(z_{\ell+\frac{1}{2}} - z_{\ell-\frac{1}{2}}\right) + \mathcal{O}(h^{m+2}), \quad z_{\ell\pm\frac{1}{2}} = \mathcal{O}(h^{m+2}), \quad (6.8a)$$

with homogeneous boundary conditions. Let us first consider the simplified scheme

$$(L_\Delta \check{\theta})_\ell = \frac{1}{\delta_\ell}\left(z_{\ell+\frac{1}{2}} - z_{\ell-\frac{1}{2}}\right) + \mathcal{O}(h^{m+2}) \quad (6.8b)$$

of type (3.13a) with solution according to (3.13b),

$$\check{\theta}_\ell = \sum_{k=0}^{n-1}\big(G(t_\ell, t_k) - G(t_\ell, t_{k+1})\big)z_{k+\frac{1}{2}} + \sum_{k=0}^{n-1}G(t_\ell, t_k) \cdot \mathcal{O}(h^{m+2}) = \mathcal{O}(h^{m+2}),$$

$$(6.9)$$

see (3.13c). It remains to be shown that $\|\theta - \check{\theta}\|_\Delta = \mathcal{O}(h^{m+2})$ holds. Subtracting (6.8b) from (6.8a) yields

$$\left(L_\Delta(\theta - \check{\theta})\right)_\ell = \left(\widehat{H}_\Delta(\theta - \check{\theta})\right)_\ell + \left(\widehat{H}_\Delta \check{\theta}\right)_\ell,$$

and, due to stability of the SDS, it remains to show that

$$\left(\widehat{H}_\Delta \check{\theta}\right)_\ell = c_1(t_\ell)\left(\partial_\Delta \check{\theta}\right)_\ell + c_0(t_\ell)\check{\theta}_\ell = \mathcal{O}\left(h^{m+2}\right) \tag{6.10}$$

is valid. Here, $c_0(t_\ell)\check{\theta}_\ell = \mathcal{O}(h^{m+2})$, and from (6.9) we obtain

$$\left(\partial_\Delta \check{\theta}\right)_\ell = \sum_{k=0}^{n-1} \left(\partial_\Delta^{[2]} G\right)_{\ell,k} z_{k+\frac{1}{2}} + \left(\partial_\Delta^{[1]} G\right)_{\ell,k} \cdot \mathcal{O}(h^{m+2}),$$

with $z_{k+\frac{1}{2}} = \mathcal{O}(h^{m+2})$ and $(\partial_\Delta^{[v]} G)_{\ell,k}$ from (3.15), which implies $c_1(t_\ell)(\partial_\Delta \check{\theta})_\ell = \mathcal{O}(h^{m+2})$. Altogether this shows that (6.10) indeed holds, which completes the proof. □

6.2 The nonlinear case

Supra-accuracy of the error estimate extends to the nonlinear case. We present the proof in a concise style, using quasilinearization arguments. The focus is on how to organize this in a way such that Lemma 6.1 and (supra)convergence of the SDS can be exploited.

Theorem 6.2 *For the nonlinear BVP* (1.1a)*,* (3.1)*, the deviation of the error estimate satisfies*

$$\|\theta\|_\Delta = \|e - \varepsilon\|_\Delta = \mathcal{O}(h^{m+2}).$$

Proof We use vector notation for continuous and discrete functions,

$$V(t) := \begin{bmatrix} v(t) \\ v'(t) \end{bmatrix}, \qquad V_\ell := \begin{bmatrix} v_\ell \\ (\partial_\Delta v)_\ell \end{bmatrix},$$

i.e., capital or boldface capital letters typographically represent the named object together with its derivative or centered difference quotient, respectively. Furthermore, emplyoing the the mean value theorem we will denote[6]

$$H(t, V, W) := \int_0^1 \nabla f(t, V + \sigma W) \, d\sigma, \quad \widehat{H}_\ell(V, W) := H(t_\ell, V, W).$$

---

[6] Here, $\nabla f(t, V)$ is the gradient with respect to V.

Difference equation for θ via quasilinearization. According to (2.9), the deviation $\theta = e - \varepsilon = (\widetilde{u} - u) - (\widehat{\widetilde{u}} - \widehat{u})$ satisfies

$$
\begin{aligned}
(L_\Delta \, \theta)_\ell &= \big(\widehat{F}(\widetilde{u})_\ell - \widehat{F}(u)_\ell\big) - \big(\widehat{F}(\widehat{\widetilde{u}})_\ell - \widehat{F}(\widehat{u})_\ell\big) \\
&\quad + \mathcal{Q}_\ell(F(\widetilde{u}) - F(u)) - \big(\widehat{F}(\widetilde{u})_\ell - \widehat{F}(u)_\ell\big) + (\mathcal{Q}_\ell - \mathcal{I}_\ell)(F(u)) \\
&= \mathcal{I}_\ell(F(\widetilde{u}) - F(u)) - \big(\widehat{F}(\widehat{\widetilde{u}})_\ell - \widehat{F}(\widehat{u})_\ell\big) \\
&\quad + (\mathcal{Q}_\ell - \mathcal{I}_\ell)(F(\widetilde{u}) - F(u)) + \underbrace{(\mathcal{Q}_\ell - \mathcal{I}_\ell)(L\,u)}_{=\,\mathcal{O}(h^{m+2})},
\end{aligned}
\tag{6.11}
$$

for $\ell = 1 \ldots n - 1$, with homogeneous boundary conditions. Here,

$$
\begin{aligned}
(\widehat{F}(\widetilde{u})_\ell - \widehat{F}(u)_\ell) &= f_\ell(\widetilde{U}_\ell) - f_\ell(U_\ell) = \widehat{H}_\ell(U_\ell, E_\ell) \cdot E_\ell, \\
(\widehat{F}(\widehat{\widetilde{u}})_\ell - \widehat{F}(\widehat{u})_\ell) &= f_\ell(\widehat{\widetilde{U}}_\ell) - f_\ell(\widehat{U}_\ell) = \widehat{H}_\ell(\widehat{U}_\ell, \mathcal{E}_\ell) \cdot \mathcal{E}_\ell,
\end{aligned}
$$

and

$$
(F(\widetilde{u}) - F(u))(t) = f(t, \widetilde{U}(t)) - f(t, U(t)) = H(t, U(t), E(t)) \cdot E(t).
$$

After rearranging terms, (6.11) now takes the form of a linearized scheme of SDS type for $\boldsymbol{\Theta}_\ell = \boldsymbol{E}_\ell - \boldsymbol{\mathcal{E}}_\ell$,

$$
\begin{aligned}
(L_\Delta \, \theta)_\ell &= \widehat{H}_\ell(\widehat{U}_\ell, \mathcal{E}_\ell) \cdot \boldsymbol{\Theta}_\ell \\
&\quad + \underbrace{\big(\widehat{H}_\ell(U(t_\ell), E(t_\ell)) - \widehat{H}_\ell(\widehat{U}_\ell, \mathcal{E}_\ell)\big) \cdot \boldsymbol{E}_\ell}_{\mathbf{J}_0} \\
&\quad + \underbrace{\big(\mathcal{I}_\ell\big(H(\cdot, U(\cdot), E(\cdot)) \cdot E(\cdot)\big) - \widehat{H}_\ell(U(t_\ell), E(t_\ell)) \big) \cdot \boldsymbol{E}_\ell}_{\mathbf{J}_1} \\
&\quad + \underbrace{(\mathcal{Q}_\ell - \mathcal{I}_\ell)\big(H(\cdot, U(\cdot), E(\cdot)) \cdot E(\cdot)\big)}_{\mathbf{J}_2} + \mathcal{O}(h^{m+2}).
\end{aligned}
$$

Estimation of $\mathbf{J}_0$. In the linear case this term vanishes. For the nonlinear case we exploit supraconvergence (3.18) of the SDS, i.e., $U(t_\ell) - \widehat{U}_\ell = \mathcal{O}(h^2)$, and $E(t) = \mathcal{O}(h^m)$ [Theorem 4.1, (4.3)] to conclude $\mathbf{J}_0 = \mathcal{O}(h^{m+2})$.

Estimation of $\mathbf{J}_1$. For all grid points t_ℓ,

$$
H(t, U(t), E(t)) \cdot E(t) = c_1(t)e'(t) + c_0(t)e(t)
$$

with smooth coefficient functions

$$
c_0(t) = c_0(t, U(t), E(t)) = \int_0^1 \partial_v f(t, U(t) + \sigma\, E(t))\, d\sigma,
$$

$$c_1(t) = c_1(t, U(t), E(t)) = \int_0^1 \partial_{v'} f(t, U(t) + \sigma E(t)) \, d\sigma,$$

additionally depending on $U = (u, u')$ and $E = (e, e')$. Furthermore,

$$\widehat{H}_\ell(U(t_\ell), E(t_\ell)) \cdot \mathbf{E}_\ell = c_1(t_\ell)(\partial_\Delta e)_\ell + c_0(t_\ell)e(t_\ell)$$

with the same functions c_0 and c_1. Thus, for $\mathbf{J}_1$ we can resort to Lemma 6.1 as in the linear case.

Estimation of $\mathbf{J}_2$. The estimate

$$\mathbf{J}_2 = (\mathcal{Q}_\ell - \mathcal{I}_\ell)(c_1 \, e' + c_0 \, e) = \mathcal{O}(h^{m+2})$$

is obtained in the same way as in the linear case.

The assertion of the theorem is now also obtained in an analogous way as in the proof of Theorem 6.1. □

7 Implementation, extensions, and numerical illustration

7.1 Implementation

The collocation method and the a posteriori error estimator were implemented in a test code on the basis of MATLAB. For an arbitrary given collocation mesh Δ^N and given collocation nodes $\{\rho_j\}$ the algebraic equations defining the piecewise polynomial solution $\widetilde{u}(t)$ are generated via differentiation matrices representing the evaluation of $\widetilde{u}'(t_{i,j})$ and $\widetilde{u}''(t_{i,j})$ in terms of the $\widetilde{u}(t_{i,j})$. The respective coefficients as well as the quadrature coefficients required for evaluation of the defect are static data which can be generated in a preprocessing step.

For linear problems, the respective systems of linear algebraic equations are solved by direct elimination. For the nonlinear case, Newton iteration with residual-based control was implemented.

7.2 Treatment of general linear boundary conditions

For boundary conditions (1.1b), the idea again is to convert them into exact difference equations, employing the given differential equation. For illustration, let us consider the case where Neumann data are given at $t_n = b$, i.e., $u'(b) = s_1$. In the SDS (2.2) for computing $\widehat{u}_\Delta$, the difference equation at the Neumann boundary is discretized by the one-sided difference quotient

$$(\partial_1^- \widehat{u})_n := \frac{\widehat{u}_n - \widehat{u}_{n-1}}{\delta_{n-\frac{1}{2}}} = s_1. \tag{7.1a}$$

Due to the Peano representation

$$(\partial_1^- v)_n - v'(t_n) = -\delta_{n-\frac{1}{2}} \mathcal{I}_n(v''), \quad \mathcal{I}_n(w) := \int_0^1 \xi \, w(t_{n-1} + \xi \, \delta_{n-\frac{1}{2}}) \, d\xi,$$

the solution $u(t)$ of the BVP exactly satisfies

$$(\partial_1^- u)_n = s_1 - \delta_{n-\frac{1}{2}} \mathcal{I}_n(f(\cdot, u(\cdot), u'(\cdot))).$$

The corresponding defect of the collocation solution $\widetilde{u}(t)$ is

$$d_n = (\partial_1^- \widetilde{u})_n + \delta_{n-\frac{1}{2}} \mathcal{Q}_n f(\cdot, \widetilde{u}(\cdot), \widetilde{u}'(\cdot)) - s_1,$$

where $\mathcal{Q}_n$ is the quadrature approximation to $\mathcal{I}_n$ obtained in the same way as in Sect. 5.1, with quadrature nodes from the rightmost collocation interval $[\tau_{N-1}, \tau_N]$. In contrast to Dirichlet boundary conditions, which are exactly satisfied by $\widetilde{u}$, we now have to account for the defect with respect to the boundary condition. This means that the boundary equation

$$(\partial_1^- \widehat{\widetilde{u}})_n = s_1 + d_n$$

is incorporated into the SDS (2.4) defining the auxiliary grid function $\widehat{\widetilde{u}}_\Delta$.

In order to further enhance the asymptotic quality of the error estimate obtained in this way, one can modify the SDS using a one-sided second-order approximation $\partial_2^- v$ to v' involving three consecutive grid points (as in BDF2 schemes), and replaces (7.1a) by

$$(\partial_2^- \widehat{u})_n = s_1. \tag{7.1b}$$

The respective coefficients, the corresponding 'EDS integral operator' $\mathcal{I}_n$ and its quadrature approximation $\mathcal{Q}_n$ are obtained by a routine calculation. This redefines d_n, and the rest of the procedure is the same as before. For a numerical illustration, see Sect. 7.3.

It is evident how the procedure extends to the case of general linear boundary conditions (1.1b). For the treatment of nonlinear boundary conditions, see Sect. 7.3.

7.3 Examples

We present two selected numerical examples. These serve to illustrate the theoretical results but, in particular, Example 2 also indicates extensions of the convergence results obtained, for the case of m odd and symmetric collocation nodes, and for general linear boundary condition. Further numerical results are reported in [18].

Example 1 A nonlinear BVP with Dirichlet boundary conditions. For

$$u''(t) = 1 - (u'(t))^2,$$
$$u(0) = \frac{1}{2}, \quad u(1) = \frac{1}{4},$$

the analytic solution is known. We choose $m = 4$ (with collocation degree $m + 1 = 5$), $\rho = (0, \rho_1, \rho_2, \rho_3, \rho_4, 1)$ including Chebyshev nodes $\rho_1, \rho_2, \rho_3, \rho_4$, and N collocation intervals of length $1/N$. The results are shown in Table 1.

As the practical purpose of an error estimate is its use for mesh adaptation, it is essential that collocation intervals of variable lengths are admitted. We note that our theoretical results from Sect. 6 do not rely on the assumption of an equidistant collocation mesh Δ^N. We illustrate this point by repeating the experiment documented in Table 1, but starting with a non-uniformly distributed mesh Δ^N, followed by coherent refinement to observe orders. The results are shown in Table 2.

Example 2 A linear BVP with general linear boundary conditions. For

$$u''(t) = \frac{1}{2} u'(t) + \frac{1}{2} u(t) - \left(\frac{1}{2} + 3t\right) e^t$$

with boundary conditions

$$\begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u(a) \\ u'(a) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} u(b) \\ u'(b) \end{bmatrix} = \begin{bmatrix} -2 \\ -e \end{bmatrix}$$

and known analytic solution,

Table 1 Results for Example 1

N	$\|R_\Delta(e)\|_\Delta$	p_e	$\|\theta_\Delta\|_\Delta$	p_θ
4	7.171e−08		6.588e−10	
8	4.524e−09	3.99	9.187e−12	6.16
16	2.809e−10	4.01	1.212e−13	6.24

Observed orders: $p_e \approx m = 4$ for the collocation error, and $p_\theta \approx m + 2 = 6$ for the deviation of the error estimate. These results are in accordance with the assertion of Theorem 6.2

Table 2 Results for Example 1, with initial mesh $\Delta^N = (0, 0.01, 0.1, 0.3, 1)$ and coherent refinement

N	$\|R_\Delta(e)\|_\Delta$	p_e	$\|\theta_\Delta\|_\Delta$	p_θ
4	4.481e−06		5.234e−07	
8	3.267e−07	3.77	4.254e−09	6.94
16	1.936e−08	4.08	6.054e−11	6.13

All other data are the same as for Table 1. Here, the maximal stepsize is significantly larger, resulting in a larger error. However, the behavior of the observed orders has not changed

Table 3 Results for Example 2

N	$\|R_\Delta(e)\|_\Delta$	p_e	$\left\|\theta_\Delta^{[1]}\right\|_\Delta$	$p_\theta^{[1]}$	$\left\|\theta_\Delta^{[2]}\right\|_\Delta$	$p_\theta^{[2]}$
4	1.609 e−06		1.205 e−07		2.683 e−09	
8	1.050 e−07	4.01	3.210 e−09	5.23	4.414 e−11	5.93
16	6.544 e−09	4.00	1.003 e−10	5.00	8.687 e−13	5.67

Here, $\theta_\Delta^{[1]}$ and $\theta_\Delta^{[2]}$ correspond to a treatment of the derivatives in the boundary conditions in the sense of (7.1a) and (7.1b) respectively. The observed orders are of the expected quality, namely $p_\theta^{[1]} \approx m+1 = 5$ and $p_\theta^{[2]} \approx m+2 = 6$

we choose $m = 3$ (with collocation degree $m+1 = 4$), $\rho = (0, \rho_1, \rho_2, \rho_3, 1)$ including Chebyshev nodes ρ_1, ρ_2, ρ_3, and N collocation intervals of length $1/N$. The results are shown in Table 3.

7.4 Further extensions

General nonlinear boundary conditions

Linear boundary conditions frequently occur in applications, but the case of nonlinear boundary conditions

$$b_i(u(a), u'(a), u(b), u'(b)) = 0, \quad i = 1, 2, \tag{7.2}$$

is of course also relevant. For the collocation solver, incorporation of (7.2) is straightforward; for the finite-difference schemes, $u'(a)$ and $u'(b)$ are again approximated by difference quotients. Since our approach relies on an underlying EDS for the computation of the defect, the question is how to translate (7.2) to the EDS context. This can be accomplished on the basis of the considerations from Sect. 7.2. Consider for instance a right boundary condition of the form

$$b_2(u(b), u'(b)) = 0. \tag{7.3}$$

In the notation from Sect. 7.2, the solution $u(t)$ of the BVP exactly satisfies

$$u'(b) = u'(t_n) = \left(\partial_1^- u\right)_n + \delta_{n-\frac{1}{2}} \, \mathcal{I}_n\big(f(\cdot, u(\cdot), u'(\cdot))\big).$$

On the basis of this identity, the defect of an approximate solution $\widetilde{u}$ with respect to (7.3) reads

$$d_n = b_2\big(\widetilde{u}_n, (\partial_1^- \widetilde{u})_n + \delta_{n-\frac{1}{2}} \, \mathcal{Q}_n\big(f(\cdot, u(\cdot), \widetilde{u}'(\cdot))\big)\big).$$

In the SDS defining the auxiliary grid function $\widehat{\widetilde{u}}_\Delta$, the corresponding boundary equation becomes

$$b_2\big(\widetilde{\widehat{u}}_n, (\partial_1^-\widetilde{\widehat{u}})_n\big) = d_n.$$

Variable coefficients, ODE systems, and higher-order problems

As the idea of EDS is rather general and well-developed, see [12,19], it can be expected that a similar approach can be successfully applied to more general problem classes, e.g., for the case where Lu is a differential operator with variable coefficients or of higher order. Also, the extension to systems or mixed systems seems to be within the scope. Again we stress that in our approach the EDS formulation is only used for evaluation of the defect, which is significantly simpler to realize than solving the EDS equations up to high order.

References

1. Ascher, U., Christiansen, J., Russell, R.: A collocation solver for mixed order systems of boundary values problems. Math. Comp. **33**, 659–679 (1978)
2. Ascher, U., Mattheij, R., Russell, R.: Numerical Solution of Boundary Value Problems for Ordinary Differential Equations. Prentice-Hall, Englewood Cliffs (1988)
3. Auzinger, W., Kneisl, G., Koch, O., Weinmüller, E.: A collocation code for boundary value problems in ordinary differential equations. Numer. Algorithms **33**, 27–39 (2003)
4. Auzinger, W., Koch, O., Praetorius, D., Weinmüller, E.: New a posteriori error estimates for singular boundary value problems. Numer. Algorithms **40**, 79–100 (2005)
5. Auzinger, W., Koch, O., Weinmüller, E.: Efficient collocation schemes for singular boundary value problems. Numer. Algorithms **31**, 5–25 (2002)
6. Auzinger, W., Koch, O., Weinmüller, E.: Analysis of a new error estimate for collocation methods applied to singular boundary value problems. SIAM J. Numer. Anal. **42**(6), 2366–2386 (2005)
7. Auzinger, W., Kreuzer, W., Hofstätter, H., Weinmüller, E.: Modified defect correction algorithms for ODEs. Part I: general theory. Numer. Algorithms **36**, 135–155 (2004)
8. Butcher, J.C., Cash, J., Moore, G.: Defect correction for 2-point boundary value-problems on non-equdistant meshes. Math. Comp. **64**, 629–648 (1995)
9. Cash, J., Kitzhofer, G., Koch, O., Moore, G., Weinmüller, E.: Numerical solution of singular two point BVPs, JNAIAM. J. Numer. Anal. Indust. Appl. Math. **4**, 129–149 (2009)
10. Doedel, E.J.: The construction of finite difference approximations to ordinary differential equations. SIAM J. Numer. Anal. **15**, 451–465 (1978)
11. Dutt, A., Greengard, L., Rokhlin, V.: Spectral deferred correction methods for ordinary differential equations. BIT **40**, 241–266 (2000)
12. Gavriluk, I.P., Hermann, M., Makarov, V.L., Kutniv, M.V.: Exact and Truncated Difference Schemes for Boundary Value ODEs. Birkhäuser, Basel (2010)
13. Kitzhofer, G., Koch, O., Lima, P., Weinmüller, E.: Efficient numerical solution of the density profile equation in hydrodynamics. J. Sci. Comput. **32**, 411–424 (2007)
14. Kitzhofer, G., Koch, O., Pulverer, G., Simon, C., Weinmüller, E.: The new MATLAB code BVPSUITE for the solution of singular implicit boundary value problems. JNAIAM J. Numer. Anal. Indust. Appl. Math. **5**, 113–134 (2010)
15. Koch, O.: Asymptotically correct error estimation for collocation methods applied to singular boundary value problems. Numer. Math. **101**, 143–164 (2005)
16. Lynch, R.E., Rice, J.R.: A high-order difference method for differential equations. Math. Comp. **34**, 333–372 (1980)
17. Manteuffel, T.A., White Jr, A.B.: The numerical solution of second-order boundary value problems on nonuniform meshes. Math. Comp. **47**, 511–535 (1986)

18. Saboor Bagherzadeh, A.: Defect-based error estimation for higher order differential equations. PhD thesis, Vienna University of Technology (2011)
19. Samarskii, A.A.: The Theory of Difference Schemes. Marcel Dekker Inc., New York, Basel (2001)
20. Stetter, H.J.: The defect correction principle and discretization methods. Numer. Math. **29**, 425–443 (1978)