MATHEMATICS OF COMPUTATION
VOLUME 43, NUMBER 167
JULY 1984, PAGES 157-187

Collocation for Singular Perturbation Problems II: Linear First Order Systems Without Turning Points

By U. Ascher* and R. Weiss**

Abstract. We consider singularly perturbed linear boundary value problems for ODE's with variable coefficients, but without turning points. Convergence results are obtained for collocation schemes based on Gauss and Lobatto points, showing that highly accurate numerical solutions for these problems can be obtained at a very reasonable cost using such schemes, provided that appropriate meshes are used. The implementation of the numerical schemes and the practical construction of corresponding meshes are discussed.

These results extend those of a previous paper which deals with systems with constant coefficients.

1. Introduction. In part I of this work [2] (hereinafter referred to as "Part I"), we have considered the numerical solution of singularly perturbed boundary value ordinary differential equations with constant coefficients. Our attention was focused on symmetric collocation schemes, which include the midpoint (or box) and the trapezoidal difference schemes as special cases. We have shown that such schemes can be used to compute highly accurate numerical solutions at a very reasonable cost, provided that appropriate meshes are used. Such a mesh consists, in general, of three parts: Two fine grids near the boundaries, to cover the possible two-layer regions, and a coarser grid in between.

Similar results for the variable coefficient case are obtained in Weiss [9] for the trapezoidal and midpoint schemes. The eigenvalue of the "fast component" part of the differential equations are assumed to stay away from the imaginary axis for all values of the independent variable. Thus, in particular, turning points are excluded from the discussion. In the passage from constant to variable coefficients, the analysis had to be significantly extended.

In this paper we extend the results of the two papers mentioned above to include convergence results for the collocation schemes based on Gauss and Lobatto points for linear two-point boundary value problems which have a uniformly bounded inverse and which are restricted as in [9]. Our convergence results are summarized in Theorem 3.3. In addition, we describe an implementation of these schemes, discuss practical mesh construction and demonstrate our results numerically. The ideas presented here have been implemented in Spudich and Ascher [13].

Received June 21, 1982.
1980 *Mathematics Subject Classification*. Primary 65L10.
*The research of this author was supported in part under NSERC grant A4306.
**Supported by Oesterreichischer Forschungsfoerderungsfonds.

0025-5718/84 $1.00 + $.25 per page

The general problem considered in this paper is of order $n + m$, with n equations singularly perturbed,

$$\varepsilon \mathbf{y}' = A_{11}(t, \varepsilon)\mathbf{y} + A_{12}(t, \varepsilon)\mathbf{z} + \mathbf{f}_1(t, \varepsilon), \tag{1.1}$$

$$0 \leqslant t \leqslant 1,$$

$$\mathbf{z}' = A_{21}(t, \varepsilon)\mathbf{y} + A_{22}(t, \varepsilon)\mathbf{z} + \mathbf{f}_2(t, \varepsilon), \tag{1.2}$$

plus the boundary conditions for $\mathbf{x}(t) = \binom{\mathbf{y}(t)}{\mathbf{z}(t)}$

$$B_0(\varepsilon)\mathbf{x}(0) + B_1(\varepsilon)\mathbf{x}(1) = \boldsymbol{\beta}. \tag{1.3}$$

The assumption (2.3) below on the eigenvalues of A_{11} plus the other regularity assumptions lead to the conclusion that the solution of (1.1)–(1.3) consists of a smooth curve away from the boundaries, possibly connected at each end to the boundary by a thin transition layer. As was pointed out in Part I, with Gauss or Lobatto schemes these boundary layer solutions must be approximated accurately, because otherwise layer errors would propagate throughout the entire interval of integration. The meshes used for collocation thus consist of three parts: Two fine grids near each boundary, with maximum mesh spacing $h_L \leqslant K\varepsilon$ for a suitable constant K, connected to a much sparser mesh away from the boundaries with minimum mesh spacing $\underline{h} \gg \varepsilon$. The determination of the sparse mesh is based on the accuracy needed in the approximation of the reduced solution. A similar mesh structure with a symmetric difference scheme for certain second order systems has been considered in Kreiss [12]. The total number of mesh points N required to meet a given error tolerance can be made to be independent of ε.

The essential features of the convergence results, summarized in Theorem 3.3, are as follows. Assume, for simplicity of presentation, that the coarse mesh segment away from boundaries is uniform, with mesh spacing $h \gg \varepsilon$. The error at mesh points in the fast solution components $\mathbf{y}$ is uniformly estimated by

$$O(e + \delta).$$

Here δ is an error tolerance, which controls both the absolute error in layer regions and the locations where matching between the fine mesh segments and the coarse mesh in between takes place; e is the error of approximation away from layers in the idealized situation where no error propagates from the boundary layers. For a given δ, the same layer meshes are constructed for a k-stage Gauss scheme and for a $(k + 1)$-stage Lobatto scheme. However, while for the Lobatto scheme the usual superconvergence order $e = h^{2k}$ is retained as $\varepsilon \to 0$, for the Gauss scheme we only get $e = h^{k+q}$, $q = 0$ if k is even and $q = 1$ if k is odd. In addition, improved estimates for the slow solution components $\mathbf{z}$ are obtained when, up to $O(\varepsilon)$, the boundary conditions (1.3) contain a subset of m linearly independent conditions involving $\mathbf{z}$ alone. In this case, the superconvergence order $O(h^{2k})$ is retained both with the k-stage Gauss scheme and with the $(k + 1)$-stage Lobatto scheme, while the contribution of the layer error to the global error bound is $O(h\delta)$ for Lobatto schemes, and a smaller $O(\varepsilon h^{-1}\delta)$ for Gauss schemes.

Of course, the mesh described above becomes highly nonuniform for very small ε. However, higher order collocation methods can handle such nonuniformity, see Part

I and Ascher, Pruess and Russell [1]. Thus they are preferable to convergence acceleration methods in this context.

Following a short section where some results on the analytic solution of (1.1)–(1.3) are gathered for later use is Section 3, where the numerical schemes, their implementation and properties and the convergence results are presented. In Section 3.1 we describe a careful implementation of the collocation schemes which uses local unknowns elimination (or condensation of parameters), resulting in a well-conditioned system of linear equations (3.14), (3.16) of a familiar sparse structure, independent of the order of the scheme. This implementation is used both for the analysis and for the numerical calculations in following sections. The condition number of the matrix is a modest $O(N)$ and in particular is independent of ε (cf. Theorem 6.2 of Part I).

Indeed, it is a good practice in actual computation to roughly estimate the condition number of the above matrix for two values of ε, say. If that condition number seems to get large as ε decreases, then something is "wrong": The mesh may be inadequate, or (3.54) does not hold or, perhaps most commonly, the differential problem is not well posed uniformly in ε. How to deal with the latter two cases will be discussed in a subsequent paper.

In Section 3.2 we consider a transformation of the dependent variables, needed for the analysis. Whereas in Part I this transformation commutes with the collocation operator, here it does not, and the resulting residue is shown to be sufficiently small in norm so that it can be considered as a small perturbation in regions where the mesh is dense, i.e. in boundary layer regions.

In Section 3.3 the mesh is described, together with the general collocation solution decomposition on each of its three parts. Then, in Section 3.4, our convergence results are stated. Theorem 3.1 summarizes the results for the layer regions near the boundaries while Theorem 3.2 describes our results in the region away from the boundaries. Theorem 3.3 then states the combined results of the previous two theorems on the entire interval.

Sections 4 and 5 are devoted almost entirely to the proofs of Theorems 3.1 and 3.2, respectively. In Section 4 we also discuss the layer mesh construction and show that the number of mesh points needed to achieve overall accuracy δ for any ε, $0 < \varepsilon \leqslant \varepsilon_1$, is $O(\delta^{-1/p})$, where p is the order of superconvergence of the method, defined in (3.42). This, provided that the mesh defined in (3.46), (3.47) is used. If a uniform layer mesh is used instead, then the number of mesh points needed is $O(-\delta^{-1/p} \ln \delta)$. But the actual advantage of (3.46), (3.47) over a uniform layer mesh is more significant than these bounds would indicate; see Table 4.2 of Part I.

It is interesting to note that, perhaps contrary to one's first intuition, the analysis for the "long" interval away from the boundaries, where the solution varies slowly, is much more gruelling than the analysis for the layer intervals, where the solution varies very rapidly. In fact, the solution in the layer is dominated by a rapidly decreasing exponential and so its form is very smooth and simple to approximate, provided that we have a layer mesh with step sizes proportional to ε, affecting a stretching transformation. Indeed, it is the simple, exponential form of the layer solution which enables us to come up with the a priori error equidistributing mesh (3.46), (3.47), whereas in general such meshes can be constructed only adaptively. Markowich and Ringhofer [6] had a similar success with problems on infinite intervals.

In Section 6, we seal this paper with a numerical example demonstrating our theoretical results.

The extension of the analysis presented here to nonlinear problems of a similar form is considered in [10], where nonlinear numerical examples are presented as well.

2. Analytic Preliminaries. In this section we mention some analytic results needed in the sequel and develop some notation. Since this section covers the same ground as Section 2 of Weiss [9], we allow ourselves to omit some details here.

Consider the linear problem (1.1), (1.2) where $A_{ij} = A_{ij}(t, \varepsilon)$ and $\mathbf{f}_i = \mathbf{f}_i(t, \varepsilon)$ are assumed, for simplicity, to be in $C^\infty([0,1] \times [0, \varepsilon_0])$ for some $\varepsilon_0 > 0$, $1 \leqslant i, j \leqslant 2$. Further, assume that

$$A_{11}(t,0) = E(t)\Lambda(t)E^{-1}(t), \tag{2.1}$$

with $E \in C^\infty[0,1]$,

$$\Lambda(t) = \operatorname{diag}\{\lambda_1(t),\ldots,\lambda_n(t)\} \tag{2.2}$$

and

$$\operatorname{re}(\lambda_i(t))\begin{cases} < 0, & i = 1,\ldots,n_-, \\ > 0, & i = n_- + 1,\ldots,n, \end{cases} \qquad t \in [0,1]. \tag{2.3}$$

Let $n_+ := n - n_-$, and denote

$$\Lambda_-(t) = \operatorname{diag}\{\lambda_1(t),\ldots,\lambda_{n_-}(t)\}, \qquad \Lambda_+(t) = \operatorname{diag}\{\lambda_{n_-+1}(t),\ldots,\lambda_n(t)\}.$$

We wish to decouple the slow components $\mathbf{z}$ from the fast ones and to (almost) diagonalize the remaining system for $\mathbf{y}$. With $L(t)$ a smooth solution to

$$\varepsilon L' = -LA_{11} + \varepsilon(A_{22}L - LA_{12}L) + A_{21} \tag{2.4}$$

define the transformation

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} E^{-1} & 0 \\ -\varepsilon L & I \end{pmatrix}\begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}. \tag{2.5}$$

(See [9] for justification.) The system (1.1)–(1.2) is then transformed into

$$\varepsilon\mathbf{u}' = (\Lambda + \varepsilon B_{11})\mathbf{u} + B_{12}\mathbf{v} + \mathbf{g}_1, \tag{2.6}$$

$$\mathbf{v}' = B_{22}\mathbf{v} + \mathbf{g}_2, \tag{2.7}$$

where B_{11}, B_{12}, B_{22}, $\mathbf{g}_1, \mathbf{g}_2$ are smooth functions of t and ε.

For the transformed system (2.6)–(2.7), a desirable representation of the solution is obtained [5]: Writing it compactly as

$$H\mathbf{w} = \mathbf{g} \tag{2.8}$$

with

$$\mathbf{w}(t) = \begin{pmatrix} \mathbf{u}(t) \\ \mathbf{v}(t) \end{pmatrix}, \qquad \mathbf{g}(t) = \begin{pmatrix} \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \end{pmatrix},$$

and introducing the maps $P_- \in \mathbf{R}^{n_- \times n}$ and $P_+ \in \mathbf{R}^{n_+ \times n}$ defined by

$$P_-\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n_-} \end{pmatrix}, \quad P_+\mathbf{x} = \begin{pmatrix} x_{n_-+1} \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \tag{2.9}$$

we have

THEOREM 2.1. *The system* (2.8) *subject to boundary conditions*

$$P_-\mathbf{u}(0) = \eta_- \in \mathbf{R}^{n_-}, \quad P_+\mathbf{u}(1) = \eta_+ \in \mathbf{R}^{n_+}, \quad \mathbf{v}(0) = \eta_0 \in \mathbf{R}^m \tag{2.10}$$

has a unique solution which satisfies

$$\|\mathbf{w}\| \leqslant \text{const}\big(\|\mathbf{g}\| + \|\eta_-\| + \|\eta_+\| + \|\eta_0\|\big), \tag{2.11}$$

provided that ε *is sufficiently small,* $0 < \varepsilon \leqslant \varepsilon_1$. *Also, for any* $q \geqslant 0$ *there is a particular solution* $\mathbf{w}_p(t) = \begin{pmatrix} \mathbf{u}_p(t) \\ \mathbf{v}_p(t) \end{pmatrix}$ *of* (2.8) *which satisfies*

$$\sum_{j=0}^{q} \left\| \frac{d^j \mathbf{w}_p}{dt^j} \right\| \leqslant \text{const}, \qquad 0 < \varepsilon \leqslant \varepsilon_1. \quad \square \tag{2.12}$$

Now, define matrix solutions W_-, W_+ and W_0 to the homogeneous problem (2.8) with $\mathbf{g} = \mathbf{0}$ as follows:

(i) $W_- = \begin{pmatrix} U_- \\ 0 \end{pmatrix}$, $U_- \in \mathbf{R}^{n \times n_-}$, where U_- satisfies

$$\varepsilon U_-' = (\Lambda + \varepsilon B_{11})U_-, \quad P_-U_-(0) = I, \quad P_+U_-(1) = 0,$$

(ii) $W_+ = \begin{pmatrix} U_+ \\ 0 \end{pmatrix}$, $U_+ \in \mathbf{R}^{n \times n_+}$, where U_+ satisfies

$$\varepsilon U_+' = (\Lambda + \varepsilon B_{11})U_+, \quad P_-U_+(0) = 0, \quad P_+U_+(1) = I,$$

(iii) $W_0 = \begin{pmatrix} U_0 \\ V_0 \end{pmatrix}$, $U_0 \in \mathbf{R}^{n \times m}$, $V_0 \in \mathbf{R}^{m \times m}$, where

$$HW_0 = 0; \quad V_0(0) = I, \quad P_-U_0(0) = S_-(\varepsilon), \quad P_+U_0(1) = S_+(\varepsilon),$$

and $S_- \in \mathbf{R}^{n_- \times m}$, $S_+ \in \mathbf{R}^{n_+ \times m}$ can be chosen by Theorem 2.1 such that

$$\sum_{j=0}^{q} \left\| \frac{d^j W_0}{dt^j} \right\| \leqslant \text{const}. \tag{2.13}$$

Then we obtain the desired representation of the general solution to (2.6)–(2.7):

THEOREM 2.3. *Any solution* $\mathbf{w}$ *of* (2.8) *can be written as*

$$\mathbf{w} = W_-\zeta_- + W_+\zeta_+ + W_0\zeta_0 + \mathbf{w}_p, \tag{2.14}$$

with $\zeta_- \in \mathbf{R}^{n_-}$, $\zeta_+ \in \mathbf{R}^{n_+}$ *and* $\zeta_0 \in \mathbf{R}^m$. *The (smooth) particular solution satisfies* (2.12), *and the matrices* W_-, W_+ *and* W_0, *defined above, have the asymptotic expansions*

$$\begin{aligned} W_0(t) &= \sum_{j=0}^{q} W_{0j}(t)\varepsilon^j + O(\varepsilon^{q+1}), \\ U_-(t) &= \sum_{j=0}^{q} U_{-j}(t/\varepsilon)\varepsilon^j + O(\varepsilon^{q+1}), \\ U_+(t) &= \sum_{j=0}^{q} U_{+j}((t-1)/\varepsilon)\varepsilon^j + O(\varepsilon^{q+1}). \end{aligned} \tag{2.15}$$

From the expansions (2.15) *it is clear that*

$$
\begin{aligned}
&P_- U_{-0}(t/\varepsilon) = \exp(\Lambda_-(0)t/\varepsilon), \qquad P_+ U_{-0}(t/\varepsilon) = 0.\\
&P_+ U_{+0}((t-1)/\varepsilon) = \exp(\Lambda_+(1)(t-1)/\varepsilon), \qquad P_- U_{+0}(t/\varepsilon) = 0. \quad \square
\end{aligned} \tag{2.16}
$$

Consider now the linear boundary value problem (1.1)–(1.3). The boundary conditions are transformed by (2.5) into a similar form for $\mathbf{w}(0)$ and $\mathbf{w}(1)$, and substituting the representation (2.14) into these boundary conditions, we get

$$M(\varepsilon)\zeta = \hat{\beta}(\varepsilon), \tag{2.17}$$

where $\zeta = (\zeta_-, \zeta_+, \zeta_0)^T$ and the matrix M has the expansion

$$M(\varepsilon) = \sum_{j=0}^{q} M_j \varepsilon^j + O(\varepsilon^{q+1}). \tag{2.18}$$

We assume that M_0 is nonsingular. This is equivalent to assuming that the boundary value problem (1.1)–(1.3) is well posed, i.e. for ε small enough

$$\left\| \begin{matrix} \mathbf{y} \\ \mathbf{z} \end{matrix} \right\| \leqslant \text{const} \left(\left\| \begin{matrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{matrix} \right\| + \|\beta\| \right) \tag{2.19}$$

with the constant independent of ε.

It is clear that the preceding representation of the general solution of (2.8) can be made analogously on any interval $[\underline{t}, \bar{t}] \subset [0,1]$ with the solution matrices appropriately defined. In particular, in (2.15), (2.16), $\underline{t}$ would replace 0 and $\bar{t}$ would replace 1. Denoting by $(U_-)_l$ and $(U_+)_l$ the lth columns of U_- and U_+ we get

$$\left\| \frac{d^j (U_-)_l}{dt^j} \right\| \leqslant \text{const}\, \varepsilon^{-j} \left[\exp\{\text{re}(\lambda_l(\underline{t}))(t-\underline{t})/\varepsilon\} + O(\varepsilon) \right], \tag{2.20}$$

$$\underline{t} \leqslant t \leqslant \bar{t}, \; l = 1, \ldots, n_-, \; j = 0, 1, \ldots, q,$$

$$\left\| \frac{d^j (U_+)_l}{dt^j} \right\| \leqslant \text{const}\, \varepsilon^{-j} \left[\exp\{\text{re}(\lambda_l(\bar{t}))(t-\bar{t})/\varepsilon\} + O(\varepsilon) \right], \tag{2.21}$$

$$\underline{t} \leqslant t \leqslant \bar{t}, \; l = n_- + 1, \ldots, n, \; j = 0, 1, \ldots, q.$$

3. Numerical Solutions and Their Convergence.

3.1. *The Numerical Schemes and Their Implementation.* In Section 3 of Part I we have presented some classes of collocation methods and discussed their equivalent Runge-Kutta formulation and some of their properties. Here we mention only some of these details again and rely on familiarity with Part I for the rest.

A collocation procedure under consideration is completely determined in terms of k points ($k \geqslant 1$),

$$0 \leqslant \rho_1 < \cdots < \rho_k \leqslant 1, \tag{3.1}$$

which we take to be either Gauss or Lobatto points, and a mesh

$$
\begin{aligned}
&\Delta\colon 0 = t_1 < t_2 < \cdots < t_N < t_{N+1} = 1,\\
&h_i := t_{i+1} - t_i, \qquad 1 \leqslant i \leqslant N, \qquad h := \max\{h_i, 1 \leqslant i \leqslant N\}.
\end{aligned} \tag{3.2}
$$

On a given mesh Δ, the collocation solution

$$\mathbf{x}_\Delta(t) = \begin{pmatrix} \mathbf{y}_\Delta(t) \\ \mathbf{z}_\Delta(t) \end{pmatrix}$$

to (1.1)–(1.3) is a continuous piecewise polynomial vector function of degree at most k satisfying the boundary conditions (1.3) and the differential equations (1.1), (1.2) at the collocation points

$$t_{ij} := t_i + h_i\rho_j, \qquad i = 1,\ldots,N, j = 1,\ldots,k. \tag{3.3}$$

Inside each subinterval $[t_i, t_{i+1}]$, the polynomials $\mathbf{y}_\Delta(t)$ and $\mathbf{z}_\Delta(t)$ can be represented in terms of the values

$$\begin{aligned} &\mathbf{y}_i := \mathbf{y}_\Delta(t_i), \quad \mathbf{z}_i := \mathbf{z}_\Delta(t_i), \quad 1 \leqslant i \leqslant N+1, \\ &\mathbf{y}_{ij} := \mathbf{y}_\Delta(t_{ij}), \quad \mathbf{z}_{ij} := \mathbf{z}_\Delta(t_{ij}), \quad 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant k \end{aligned} \tag{3.4}$$

(strictly speaking, for Lobatto points some additional derivative values are required as well), which satisfy the difference equations

$$\frac{\varepsilon}{h_i}(\mathbf{y}_{ij} - \mathbf{y}_i) = \sum_{l=l}^{k} \hat{a}_{jl}\left(A_{11}(t_{il}, \varepsilon)\mathbf{y}_{il} + A_{12}(t_{il}, \varepsilon)\mathbf{z}_{il} + \mathbf{f}_1(t_{il})\right), \tag{3.5}$$

$$1 \leqslant j \leqslant k,$$

$$\frac{1}{h_i}(\mathbf{z}_{ij} - \mathbf{z}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\left(A_{21}(t_{il}, \varepsilon)\mathbf{y}_{il} + A_{22}(t_{il}, \varepsilon)\mathbf{z}_{il} + \mathbf{f}_2(t_{il})\right). \tag{3.6}$$

For Lobatto points, $\rho_k = 1$ and $\rho_1 = 0$. Thus $\mathbf{y}_{i+1} = \mathbf{y}_{ik}$, $\mathbf{z}_{i+1} = \mathbf{z}_{ik}$ and Eqs. (3.5), (3.6) are trivial for $j = 1$. For Gauss points, $\rho_k < 1$, $\rho_1 > 0$, and we extend the range of j in (3.5), (3.6) to include $j = k + 1$ as well, with $\mathbf{y}_{i+1} = \mathbf{y}_{i,k+1}$, $\mathbf{z}_{i+1} = \mathbf{z}_{i,k+1}$ and $\hat{a}_{k+1,l} = \hat{b}_l$, $l = 1,\ldots,k$; see Section 3 of Part I for the definitions of the constants $\hat{a}_{jl}$, $\hat{b}_l$, as well as the matrices $\hat{A}$ and $\bar{A}$ used later.

In the sequel, we shall adhere to the following notational convention, used already above. The collocation approximation to a function $\psi(t)$ is denoted by $\psi_\Delta(t)$. Its values at mesh points are ψ_i, $1 \leqslant i \leqslant N+1$, and those at collocation points are ψ_{ij}, $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant k$. Also, ψ^c will denote the vector formed by the restriction of $\psi(t)$ to $\Delta \cup \{t_{ij};\ 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant k\}$. Furthermore, c, K and c_j, $j = 0,1,2,\ldots$ will denote constants independent of ε and Δ.

Next, we describe a particular, careful implementation of the collocation schemes which is used both for the numerical calculations reported in Section 6 and for the analysis in Section 5. The differential equations (1.1), (1.2) are written as one system

$$\mathbf{x}' = A(t)\mathbf{x} + \mathbf{f}(t), \tag{3.7}$$

for which the numerical method is written in Runge-Kutta form

$$\mathbf{x}_{i+1} = \mathbf{x}_i + h_i \sum_{j=1}^{k} \hat{b}_j \mathbf{F}_{ij}, \qquad 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant k, \tag{3.8}$$

$$\mathbf{F}_{ij} = \mathbf{x}'_\Delta(t_{ij}) = A(t_{ij})\mathbf{x}_{ij} + \mathbf{f}(t_{ij}) = A(t_{ij})\left(\mathbf{x}_i + h_i \sum_{l=1}^{k} \hat{a}_{jl}\mathbf{F}_{il}\right) + \mathbf{f}(t_{ij}). \tag{3.9}$$

The unknowns $\mathbf{F}_{ij}$ (or $\mathbf{x}_{ij}$) for each interval $[t_i, t_{i+1}]$ are local and can be eliminated locally. (This is sometimes referred to as "condensation of parameters"—see Ascher, Pruess and Russell [1].) We choose to locally eliminate the $\mathbf{F}_{ij}$ in case that $\rho_k < 1$, and the $\mathbf{x}_{ij}$ in case that $\rho_k = 1$. These choices avoid unnecessary loss of digits due to cancellation error, as can be readily verified for the example $y' = y/\varepsilon + 1/\varepsilon$ with $0 < \varepsilon \ll 1$.

Consider *Gauss points* first. Equations (3.9) can be written as $(n + m)k$ linear equations

$$\mathbf{J}_i\mathbf{F}_i = \mathbf{R}_i, \tag{3.10}$$

where

$$\mathbf{F}_i = (\mathbf{F}_{i1}, \ldots, \mathbf{F}_{ik})^T, \qquad \mathbf{R}_i = C_A\mathbf{x}_i + \mathbf{f}_i,$$

$$C_A = \begin{bmatrix} A(t_{i1}) \\ \vdots \\ A(t_{ik}) \end{bmatrix}, \qquad \mathbf{f}_i = \begin{pmatrix} \mathbf{f}(t_{i1}) \\ \vdots \\ \mathbf{f}(t_{ik}) \end{pmatrix}, \tag{3.11}$$

$$J_i = I - h_i \begin{bmatrix} \hat{a}_{11}A(t_{i1}) & \hat{a}_{12}A(t_{i1}) & \cdots & \hat{a}_{1k}A(t_{i1}) \\ \hat{a}_{21}A(t_{i2}) & \hat{a}_{22}A(t_{i2}) & \cdots & \hat{a}_{2k}(t_{i2}) \\ \vdots & & & \\ \hat{a}_{k1}A(t_{ik}) & \cdots\cdots\cdots\cdots\cdots & & \hat{a}_{kk}A(t_{ik}) \end{bmatrix} \tag{3.12}$$

$$= I - h_i D_A(\hat{A} \otimes I),$$

in which I stands for an identity matrix of the appropriate dimension $(n + m$ or $(n + m)k)$ and $D_A = \operatorname{diag}\{A(t_{i1}), \ldots, A(t_{ik})\}$. (The dependence on i is suppressed in C_A and D_A.) Introducing for notational purposes the $(n + m) \times (n + m)k$ matrix

$$\hat{B} = [\hat{b}_1 I, \ldots, \hat{b}_k I], \tag{3.13}$$

we can write (3.8) as

$$\mathbf{x}_{i+1} = \Gamma_i\mathbf{x}_i + \mathbf{g}_i, \qquad 1 \leqslant i \leqslant N, \tag{3.14}$$

where

$$\Gamma_i = I + h_i\hat{B}J_i^{-1}C_A, \qquad \mathbf{g}_i = h_i\hat{B}J_i^{-1}\mathbf{f}_i. \tag{3.15}$$

The difference equations (3.14) together with the boundary equations corresponding to (1.3)

$$B_0\mathbf{x}_1 + B_1\mathbf{x}_{N+1} = \boldsymbol{\beta} \tag{3.16}$$

form a set of $(N + 1)(n + m)$ linear equations for the solution values at the mesh points, whose size and structure are independent of k.

For *Lobatto points* we perform a similar elimination of local parameters, but now our parameters are $\mathbf{x}_i = \mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{i,k-1}, \mathbf{x}_{ik} = \mathbf{x}_{i+1}$. Instead of (3.8), (3.9) we write, as in (3.5), (3.6),

$$1/h_i(\mathbf{x}_{ij} - \mathbf{x}_i) = \sum_{l=1}^{k} \hat{a}_{jl}(A(t_{il})\mathbf{x}_{il} + \mathbf{f}(t_{il})), \qquad 2 \leqslant j \leqslant k, \tag{3.17}$$

and this can be written as $(n + m)(k - 1)$ linear equations

$$\bar{J}_i \bar{\mathbf{x}}_i = \bar{\mathbf{R}}_i, \tag{3.18}$$

where

$$\begin{aligned} &\bar{\mathbf{x}}_i = (\mathbf{x}_{i2}, \ldots, \mathbf{x}_{ik})^T, \qquad \bar{\mathbf{R}}_i = (\bar{\mathbf{R}}_{i2}, \ldots, \bar{\mathbf{R}}_{ik})^T, \\ &\bar{\mathbf{R}}_{ij} = \big[I + h_i \hat{a}_{j1} A(t_i)\big]\mathbf{x}_i + h_i \sum_{l=1}^{k} \hat{a}_{jl} \mathbf{f}(t_{il}), \end{aligned} \tag{3.19}$$

$$\bar{J}_i = I - h_i \begin{bmatrix} \hat{a}_{22} A(t_{i2}) \cdots \hat{a}_{2k} A(t_{ik}) \\ \vdots \\ \hat{a}_{k2} A(t_{i2}) \cdots \hat{a}_{kk} A(t_{ik}) \end{bmatrix} = I - h_i (\bar{A} \otimes I) \bar{D}_A, \tag{3.20}$$

where $\bar{D}_A = \operatorname{diag}\{A(t_{i2}), \ldots, A(t_{ik})\}$ and where $\bar{A}$ is a nonsingular matrix, as in (3.14) of Part I.

Since $\rho_k = 1$, $\mathbf{x}_{i+1}$ is obtained as the last $n + m$ rows of $\bar{J}_i^{-1}\mathbf{R}_i$. Partitioning $\bar{J}_i^{-1}$ into blocks of size $(n + m) \times (n + m)$, $\bar{J}_i^{-1} = ((\bar{J}_i^{-1})_{jl})_{j,l=2}^{k}$, we get difference equations of the form (3.14) where now, instead of (3.15),

$$\Gamma_i = \sum_{l=2}^{k} \big(\bar{J}_i^{-1}\big)_{kl} \big[I + h_i \hat{a}_{l1} A(t_i)\big], \qquad \mathbf{g}_i = h_i \sum_{l=2}^{k} \big(\bar{J}_i^{-1}\big)_{kl} \sum_{s=1}^{k} \hat{a}_{ls} \mathbf{f}(t_{is}). \tag{3.21}$$

An advantage of the difference equations (3.14), (3.16), obtained both for Gauss and for Lobatto points, is that, even when some rows of $A(t)$ of (3.7) depend on $1/\varepsilon$ and $\varepsilon \ll h_i$, the components of Γ_i and $\mathbf{g}_i$ remain bounded and are constructed accurately.

3.2. *Transformation of Variables.* Consider the linear problem (1.1)–(1.3) and the transformed system (2.6), (2.7). Since the latter has a structure more amenable to analysis, we will rely on it in parts of our treatment. However, we stress that the actual numerical procedure is applied to (1.1), (1.2) and not to (2.6), (2.7).

In the constant coefficient case, the operators of collocation and the transformation (2.5) commute. Here they do not, in general. Thus, if we define vector functions $\mathbf{u}_\Delta(t)$, $\mathbf{v}_\Delta(t)$ by

$$\begin{pmatrix} \mathbf{u}_\Delta \\ \mathbf{v}_\Delta \end{pmatrix} = \begin{pmatrix} E^{-1} & 0 \\ -\varepsilon L & I \end{pmatrix} \begin{pmatrix} \mathbf{y}_\Delta \\ \mathbf{z}_\Delta \end{pmatrix}, \tag{3.22}$$

then $\mathbf{u}_\Delta, \mathbf{v}_\Delta$ collocate the transformed equations, but are not necessarily piecewise polynomials of degree at most k. Correspondingly, applying the transformation (3.22) to the difference equations (5.5), (3.6), we obtain

$$\begin{aligned} \frac{\varepsilon}{h_i}(\mathbf{u}_{ij} - \mathbf{u}_i) = \sum_{l=1}^{k} \hat{a}_{jl} \big\{ \big[\Lambda(t_{il}) + \varepsilon B_{11}(t_{il})\big]\mathbf{u}_{il} + B_{12}(t_{il})\mathbf{v}_{il} + \mathbf{g}_1(t_{il}) \big\} \\ + \frac{\varepsilon}{h_i}\mathbf{R}_{ij}, \qquad 1 \leqslant i \leqslant N, \end{aligned} \tag{3.23}$$

$$\frac{1}{h_i}(\mathbf{v}_{ij} - \mathbf{v}_i) = \sum_{l=1}^{k} \hat{a}_{jl} \big\{ B_{22}(t_{il})\mathbf{v}_{il} + \mathbf{g}_2(t_{il}) \big\} + \frac{1}{h_i}\mathbf{S}_{ij}, \tag{3.24}$$

where $\varepsilon/h_i\mathbf{R}_{ij}$ and $1/h_i\mathbf{S}_{ij}$ consist of linear operators acting on $\mathbf{u}_{il}, \mathbf{v}_{il}$, $l = 1,\ldots,k$ for Lobatto points and $l = 1,\ldots,k+1$ for Gauss points, and inhomogeneities. We now show that their norms are $O(h_i)$, and so they can be dealt with as small perturbations when h_i is small.

LEMMA 3.1. *For each* i, $1 \leqslant i \leqslant N$,

$$\begin{pmatrix} \varepsilon/h_i\mathbf{R}_{ij} \\ 1/h_i\mathbf{S}_{ij} \end{pmatrix} = h_i\{\boldsymbol{\phi}_{ij}[\mathbf{u}_{i1},\ldots,\mathbf{u}_{iq};\mathbf{v}_{i1},\ldots,\mathbf{v}_{iq}] + h\boldsymbol{\psi}_{ij}[\mathbf{g}_1(t_{i1}),\ldots,\mathbf{g}_1(t_{iq}),\mathbf{g}_2(t_{i1}),\ldots,\mathbf{g}_2(t_{iq})]\}, \tag{3.25}$$

where $\boldsymbol{\phi}_{ij}$, $\boldsymbol{\psi}_{ij}$ *are bounded linear operators,*

$$\|\boldsymbol{\phi}_{ij}\|, \|\boldsymbol{\psi}_{ij}\| \leqslant c, \qquad 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant q,$$

with $q = k$ *for Lobatto points,* $q = k+1$ *for Gauss points.* □

Those readers who wish to skip the proof of this lemma can do so without loss of continuity.

Proof. Writing $\mathbf{u}'_\Delta(t)$ and $\mathbf{v}'_\Delta(t)$ in terms of their polynomial interpolants of order k on $[t_i, t_{i+1}]$, we get

$$\mathbf{u}'_\Delta(t) = \sum_{j=1}^{k} \mathbf{u}'_\Delta(t_{ij}) L_j\left(\frac{t-t_i}{h_i}\right) + \frac{1}{k!}\mathbf{u}_\Delta^{(k+1)}(\zeta_t)\prod_{j=1}^{k}(t-t_{ij}), \qquad t_i \leqslant \zeta_t \leqslant t_{i+1},$$

where L_j are the Lagrange polynomials. Integrating,

$$\mathbf{u}_{ij} - \mathbf{u}_i = h_i \sum_{l=1}^{k} \mathbf{u}'_\Delta(t_{il})\hat{a}_{jl} + \frac{1}{k!}\int_{t_i}^{t_{ij}} \mathbf{u}_\Delta^{(k+1)}(\zeta_t)\prod_{l=1}^{k}(t-t_{il})\,dt$$

and so, by (3.23),

$$\mathbf{R}_{ij} = \frac{1}{k!}\int_{t_i}^{t_{ij}} \mathbf{u}_\Delta^{(k+1)}(\zeta_t)\prod_{l=1}^{k}(t-t_{il})\,dt, \tag{3.26}$$

with a similar expression for $\mathbf{S}_{ij}$, $\mathbf{v}_\Delta$ replacing $\mathbf{u}_\Delta$.

Next, since $\mathbf{y}_\Delta$ and $\mathbf{z}_\Delta$ are polynomials of degree at most k on $[t_i, t_{i+1}]$, by the transformation (3.22),

$$\mathbf{u}_\Delta^{(k+1)}(\tau) = \sum_{\nu=1}^{k+1}\binom{k+1}{\nu}(E^{-1}(\tau))^{(\nu)}\mathbf{y}_\Delta^{(k+1-\nu)}(\tau), \qquad t_i \leqslant \tau \leqslant t_{i+1}, \tag{3.27}$$

$$\mathbf{v}_\Delta^{(k+1)}(\tau) = -\varepsilon\sum_{\nu=1}^{k+1}\binom{k+1}{\nu}L^{(\nu)}(\tau)\mathbf{y}_\Delta^{(k+1-\nu)}(\tau) \tag{3.28}$$

and

$$\varepsilon\mathbf{y}_\Delta^{(l)}(\tau) = \varepsilon\frac{d^{l-1}}{d\tau^{l-1}}(\mathbf{y}'_\Delta(\tau)) = \sum_{j=1}^{k}\varepsilon\mathbf{y}'_\Delta(t_{ij})\frac{d^{l-1}}{d\tau^{l-1}}L_j\left(\frac{\tau-t_i}{h_i}\right). \tag{3.29}$$

Replacing the vectors $\varepsilon y'_\Delta(t_{ij})$ through the collocation equations (3.9) and the transformation (3.22), and substituting into (3.27), (3.28), we obtain

$$\|\mathbf{u}_\Delta^{(k+1)}\| \leqslant h_i\phi/\varepsilon, \qquad \|\mathbf{v}_\Delta^{(k+1)}\| \leqslant h_i\phi, \tag{3.30}$$

$$\phi = ch_i^{-k}\max\left\{\|\mathbf{u}_{ij}\|, \|\mathbf{v}_{ij}\|, \|\mathbf{g}_1(t_{ij})\|, \|\mathbf{g}_2(t_{ij})\|, 1 \leqslant j \leqslant q\right\}. \tag{3.31}$$

Finally, substituting (3.30), (3.31) into (3.26) and the corresponding expression for $\mathbf{S}_{ij}$, the desired result (3.25) is obtained. Q.E.D.

3.3. *The Mesh and the Decomposition of Numerical Solutions.* The meshes considered in this paper have the following structure. Near the boundaries, the step sizes h_i are comparable to ε. Specifically, there are given numbers $0 < N_0$, $N_1 < N$ and constants K_0, K_1, such that

$$\begin{aligned} h_i &\leqslant K_0\varepsilon, \quad i = 1,\ldots,N_0, \\ h_i &\leqslant K_1\varepsilon, \quad i = N - N_1 + 1,\ldots,N. \end{aligned} \tag{3.32}$$

In between, much larger step sizes may be used, i.e. $h_i \gg \varepsilon$, $i = N_0 + 1,\ldots,N - N_1$. We will assume for convenience of notation that h, the largest step size, occurs away from the boundaries. Such a mesh is depicted in Figure 1 below.

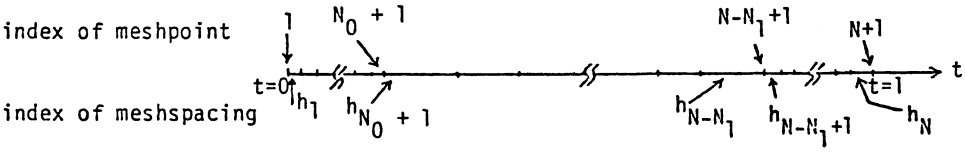

FIGURE 1: *The mesh*

For brevity of notation we set $\underline{i} = N_0 + 1$, $\bar{i} = N - N_1 + 1$ and write

$$t_{\underline{i}} = T_0\varepsilon, \qquad t_{\bar{i}} = 1 - T_1\varepsilon. \tag{3.33}$$

Our next step is to write down a decomposition representation to the discrete solution of (3.23), (3.24), similar to the representation (2.14) for the analytic solution. Moreover, we write down such a representation for each of the three parts of the mesh.

We write the system (3.23), (3.24), in analogy to (2.8) as

$$(H\mathbf{w}_\Delta)(t_{ij}) = \mathbf{g}(t_{ij}), \qquad j = 1,\ldots,q,\ i = 1,\ldots,N, \tag{3.34a}$$

or in shorthand as

$$H_\Delta\mathbf{w}_\Delta = \mathbf{g}_\Delta, \tag{3.34b}$$

where $\mathbf{w}_\Delta(t) = \binom{\mathbf{u}_\Delta(t)}{\mathbf{v}_\Delta(t)}$ is in the class S_Δ of functions defined by (3.22) with $\mathbf{y}_\Delta(t)$, $\mathbf{z}_\Delta(t)$ continuous piecewise polynomial vector functions of degree at most k. Let

$$W_{\Delta s}^M(t) = \begin{pmatrix} U_{\Delta s}^M(t) \\ V_{\Delta s}^M(t) \end{pmatrix} \in \mathbf{R}^{(n+m)\times n_s}, \qquad t \in \left[t_{M_0}, t_{M_1}\right] \tag{3.35}$$

be matrix-valued functions with columns in class S_Δ. Here s stands for $-$, $+$ or 0, $n_0 := m$, M stands for I, II or III, to denote the three mesh regions considered, and so $\mathrm{I}_0 = 1$, $\mathrm{I}_1 = \underline{i} = \mathrm{II}_0$, $\mathrm{II}_1 = \bar{i} = \mathrm{III}_0$, $\mathrm{III}_1 = N + 1$.

On the interval $[0, T_0\varepsilon]$: Define $W_{\Delta s}^{\mathrm{I}}$, $\mathbf{w}_{p\Delta}^{\mathrm{I}} \in S_\Delta$ as follows (omitting the superscript I):

$$H_\Delta W_{\Delta -} = 0, \quad P_- U_{\Delta -}(0) = I, \quad P_+ U_{\Delta -}(T_0\varepsilon) = 0, \quad V_{\Delta -}(0) = 0; \tag{3.36a}$$

(3.36b) $H_\Delta W_{\Delta+} = 0, \quad P_- U_{\Delta+}(0) = 0, \quad P_+ U_{\Delta+}(T_0\varepsilon) = I, \quad V_{\Delta+}(0) = 0;$

(3.36c) $H_\Delta W_{\Delta 0} = 0, \quad P_- U_{\Delta 0}(0) = S_-(\varepsilon), \quad P_+ U_{\Delta 0}(T_0\varepsilon) = P_+ U_0(T_0\varepsilon),$
$V_{\Delta 0}(0) = I;$

(3.36d) $H_\Delta \mathbf{w}_{P\Delta} = \mathbf{g}_\Delta, \quad P_-\mathbf{u}_{P\Delta}(0) = P_-\mathbf{u}_P(0), \quad P_+\mathbf{u}_{P\Delta}(T_0\varepsilon) = P_+\mathbf{u}_P(T_0\varepsilon),$
$\mathbf{v}_{P\Delta}(0) = \mathbf{v}_P(0).$

The general solution of (3.34) on the left layer mesh is written formally as

(3.37) $$\mathbf{w}_\Delta^{\mathrm{I}} = W_{\Delta-}^{\mathrm{I}} \zeta_-^{\mathrm{I}} + W_{\Delta+}^{\mathrm{I}} \zeta_+^{\mathrm{I}} + W_0^{\mathrm{I}} \zeta_0^{\mathrm{I}} + \mathbf{w}_{P\Delta}^{\mathrm{I}},$$

with $\zeta_-^{\mathrm{I}} \in \mathbf{R}^{n_-}$, $\zeta_+^{\mathrm{I}} \in \mathbf{R}^{n_+}$, $\zeta_0^{\mathrm{I}} \in \mathbf{R}^m$.

On the interval $[1 - T_1\varepsilon, 1]$: In precise analogy to the above, define $W_{\Delta s}^{\mathrm{III}}$ and $\mathbf{w}_{P\Delta}^{\mathrm{III}}$, again omitting superscripts:

(3.38a) $H_\Delta W_{\Delta-} = 0, \quad P_- U_{\Delta-}(1 - T_1\varepsilon) = I, \quad P_+ U_{\Delta-}(1) = 0, \quad V_{\Delta-}(1 - T_1\varepsilon) = 0;$

(3.38b) $H_\Delta W_{\Delta+} = 0, \quad P_- U_{\Delta+}(1 - T_1\varepsilon) = 0, \quad P_+ U_{\Delta+}(1) = I, \quad V_{\Delta+}(1 - T_1\varepsilon) = 0;$

(3.38c) $H_\Delta W_{\Delta 0} = 0, \quad P_- U_{\Delta 0}(1 - T_1\varepsilon) = P_- U_0(1 - T_1\varepsilon),$
$P_+ U_{\Delta 0}(1) = S_+(\varepsilon), \quad V_{\Delta 0}(1 - T_1\varepsilon) = V_0(1 - T_1\varepsilon),$

(3.38d) $H_\Delta \mathbf{w}_{P\Delta} = \mathbf{g}_\Delta, \quad P_-\mathbf{u}_{P\Delta}(1 - T_1\varepsilon) = P_-\mathbf{u}_P(1 - T_1\varepsilon),$
$P_+\mathbf{u}_{P\Delta}(1) = P_+\mathbf{u}_P(1), \quad \mathbf{v}_{P\Delta}(1 - T_1\varepsilon) = \mathbf{v}_P(1 - T_1\varepsilon).$

The general solution of (3.34) on the right layer mesh is written formally as

(3.39) $$\mathbf{w}_\Delta^{\mathrm{III}} = W_{\Delta-}^{\mathrm{III}} \zeta_-^{\mathrm{III}} + W_{\Delta+}^{\mathrm{III}} \zeta_+^{\mathrm{III}} + W_0^{\mathrm{III}} \zeta_0^{\mathrm{III}} + \mathbf{w}_{P\Delta}^{\mathrm{III}},$$

with $\zeta_-^{\mathrm{III}} \in \mathbf{R}^{n_-}$, $\zeta_+^{\mathrm{III}} \in \mathbf{R}^{n_+}$, $\zeta_0^{\mathrm{III}} \in \mathbf{R}^m$.

On the interval $[T_0\varepsilon, 1 - T_1\varepsilon]$: Define W_Δ^{II}, $W_{\Delta 0}^{\mathrm{II}}$ and $\mathbf{w}_{P\Delta}^{\mathrm{II}}$ as follows: Let $Y_0(t)$, $Z_0(t)$ be obtained from $U_0(t)$, $V_0(t)$, via the (inverse) transformation (2.5). Then W_Δ^{II} and $W_{\Delta 0}^{\mathrm{II}}$ are obtained via this transformation (3.22)

$$\text{from } X_\Delta = \begin{pmatrix} Y_\Delta \\ Z_\Delta \end{pmatrix} \quad \text{and} \quad X_{\Delta 0} = \begin{pmatrix} Y_{\Delta 0} \\ Z_{\Delta 0} \end{pmatrix},$$

respectively, which are defined as follows:

(3.40a) X_Δ and $X_{\Delta 0}$ satisfy the homogeneous equations (3.8), (3.9) with $\mathbf{f} = \mathbf{0}$.

(3.40b) $$Y_\Delta(T_0\varepsilon) = I, \qquad Z_\Delta(T_0\varepsilon) = 0;$$

(3.40c) $$Y_{\Delta 0}(T_0\varepsilon) = Y_0(T_0\varepsilon), \qquad Z_{\Delta 0}(T_0\varepsilon) = Z_0(T_0\varepsilon).$$

The particular solution $\mathbf{w}_{P\Delta}^{\mathrm{II}}$ is defined, e.g., by

(3.40d) $$H_\Delta \mathbf{w}_{P\Delta} = \mathbf{g}_\Delta, \qquad \mathbf{w}_{P\Delta}(T_0\varepsilon) = \mathbf{w}_P(T_0\varepsilon).$$

The general solution of (3.34) on the long interval away from the layers is written formally as

(3.41) $$\mathbf{w}_\Delta^{\mathrm{II}} = W_\Delta^{\mathrm{II}} \zeta^{\mathrm{II}} + W_{\Delta 0}^{\mathrm{II}} \zeta_0^{\mathrm{II}} + \mathbf{w}_{P\Delta}^{\mathrm{II}},$$

with $\zeta^{\mathrm{II}} \in \mathbf{R}^n$, $\zeta_0^{\mathrm{II}} \in \mathbf{R}^m$.

3.4. *Convergence Results.* Below we state the various results regarding the convergence of the numerical methods, culminating in Theorem 3.3. The proofs for those results which have not been given elsewhere are contained in the next two sections.

Denote by p the "regular" superconvergence order of the schemes under consideration, i.e.

$$p = \begin{cases} 2k & \text{for a } k\text{-stage Gauss scheme,} \\ 2(k-1) & \text{for a } k\text{-stage Lobatto scheme.} \end{cases} \tag{3.42}$$

Also, define the seminorms on collocation solutions,

$$\|\psi_\Delta\|_\Delta := \max\{\|\psi_i\|;\ 1 \leqslant i \leqslant N+1\}, \tag{3.43a}$$

$$\|\psi_\Delta\|_c := \max\{\|\psi_{ij}\|;\ 1 \leqslant i \leqslant N,\ 1 \leqslant j \leqslant k\}, \tag{3.43b}$$

where the vector norms used are maximum norms. Thus $\|\psi_\Delta^c\| = \max\{\|\psi_\Delta\|_\Delta, \|\psi_\Delta\|_c\}$. Also $\|\psi_\Delta\|_\Delta^M$, $\|\psi_\Delta\|_c^M$ will denote the seminorm where the range of i in (3.43) is restricted to $M_0 \leqslant i \leqslant M_1$, $M = \mathrm{I}, \mathrm{II}$ or III. For a matrix whose columns are collocation solutions, a maximum on the column norms is taken.

For the "short" intervals $[0, t_{\underline{i}}]$ and $[t_{\bar{i}}, 1]$ we have

THEOREM 3.1. (a) *The solution representations* (3.37) *and* (3.39) *are valid* (*i.e. their components can be computed in a stable way*).

(b) *With* h_L *the maximum step size in the layers* ($h_L \leqslant \varepsilon \max\{K_0, K_1\}$ *by* (3.32)), *the "smooth" components satisfy*

$$\|W_{\Delta 0} - W_0\|_\Delta^{\mathrm{I}}, \|W_{\Delta 0} - W_0\|_\Delta^{\mathrm{III}}, \|\mathbf{w}_{P\Delta} - \mathbf{w}_P\|_\Delta^{\mathrm{I}}, \|\mathbf{w}_{P\Delta} - \mathbf{w}_P\|_\Delta^{\mathrm{III}} \leqslant c h_L^P. \tag{3.44}$$

(c) *The auxiliary solution components in the layers* (*for which there are no counterparts in the exact solution decomposition*) *satisfy*

$$W_{\Delta+}^{\mathrm{I}}(T_0\varepsilon) = \begin{pmatrix} 0_{n_- \times n_+} \\ I_{n_+ \times n_+} \\ 0 \end{pmatrix} + O(\varepsilon), \qquad W_{\Delta-}^{\mathrm{III}}(1 - T_1\varepsilon) = \begin{pmatrix} I_{n_- \times n_-} \\ 0 \end{pmatrix} + O(\varepsilon). \tag{3.45}$$

(d) *For a given accuracy tolerance* δ, $\delta \geqslant c\varepsilon$, *the layer meshes can be constructed as follows*: *With*

$$\mu := \max\{|\lambda_j(0)|, j = 1, \ldots, n_-\}, \qquad \nu := \min\{-\mathrm{re}(\lambda_j(0)), j = 1, \ldots, n_-\} > 0,$$

define

$$h_1 := \frac{\varepsilon}{\mu}\left[\frac{\nu}{\mu |c_\gamma|}\right]^{1/p} \delta^{1/p}, \tag{3.46}$$

$$h_i := h_{i-1} \exp\left\{\frac{1}{p}\,\frac{\nu}{\varepsilon} h_{i-1}\right\} \quad \textit{until } t_{i+1} \geqslant T_0\varepsilon, \tag{3.47}$$

where c_γ *is a known constant* (*cf. Part* I) *and*

$$T_0 := \nu^{-1} |\ln \delta|. \tag{3.48}$$

The right end layer is constructed in an analogous way, depending on $\lambda_j(1)$, $j = n_- + 1, \ldots, n$. *Then*

$$U_{\Delta-}^{\mathrm{I}}(t_i) = \begin{pmatrix} \exp(\Lambda_-(0) t_i/\varepsilon) \\ 0 \end{pmatrix} + O(\delta), \tag{3.49a}$$

$$V_{\Delta-}^{\mathrm{I}}(t_i) = \varepsilon O(\delta), \qquad 1 \leqslant i \leqslant \underline{i},$$

$$U^{\mathrm{III}}_{\Delta+}(t_i) = \begin{pmatrix} 0 \\ \exp(\Lambda_+(1)(t_i - 1)/\varepsilon) \end{pmatrix} + O(\delta), \tag{3.49b}$$

$$V^{\mathrm{III}}_{\Delta+}(t_i) = \varepsilon O(\delta), \qquad \bar{i} \leqslant i \leqslant N+1. \quad \square$$

The proof of this theorem is given in Section 4. We note that the assumption $\delta \geqslant c\varepsilon$ is not essential, see Section 6, it just leads to a simpler presentation of the results.

For the "long" interval $[t_{\underline{i}}, t_{\bar{i}}]$ we have

THEOREM 3.2. *Let*

$$\begin{aligned} &\bar{\kappa} := \varepsilon \underline{h}^{-1}, \qquad \underline{h} := \min\{h_i; \underline{i} \leqslant i \leqslant \bar{i}\}, \\ &\kappa := (\bar{i} - \underline{i})\bar{\kappa} \geqslant \varepsilon \sum_{i=\underline{i}}^{\bar{i}-1} h_i^{-1}. \end{aligned} \tag{3.50}$$

(a) *The solution representation* (3.41) *is valid for* ε *sufficiently small such that* $\kappa < c$, *where* c *is a constant of order* 1.

(b) *The first* n *fundamental solution components satisfy*

$$W^{\mathrm{II}}_{\Delta}(t_i) = (-1)^{k(i-\underline{i})} \begin{pmatrix} E^{-1}(t_i) \\ 0 \end{pmatrix} + \begin{pmatrix} O(\kappa) \\ O(\bar{\kappa}) \end{pmatrix}, \qquad \underline{i} \leqslant i \leqslant \bar{i}, \tag{3.51a}$$

for Gauss points and

$$W^{\mathrm{II}}_{\Delta}(t_i) = (-1)^{(k+1)(i-\underline{i})} \begin{pmatrix} \Lambda^{-1}(t_i) E^{-1}(t_i) A_{11}(t_{\underline{i}}) \\ 0 \end{pmatrix} + \begin{pmatrix} O(\kappa) \\ O(h) \end{pmatrix}, \qquad \underline{i} \leqslant i \leqslant \bar{i}, \tag{3.51b}$$

for Lobatto points.

(c) *Define the error* e *as follows: For a* k*-stage Gauss scheme,* $e := h^k$ *and, if* k *is odd and the mesh is locally almost uniform, i.e.,*

$$h_{i+1} = h_i(1 + O(h_i)) \quad \text{for all } i \text{ odd or all } i \text{ even}, \tag{3.52}$$

then $e := h^{k+1}$. *For a* k*-stage Lobatto scheme* $e := Kh^P + \varepsilon h^{k-1}$, *and, if* k *is even and the mesh is locally almost uniform, then* $e = Kh^P + \varepsilon h^k$. *Also, if the slow components* $\mathbf{z}$ *are absent from* (1.1), (1.3), *then* $K = 0$.

Then we have

$$\left\|\mathbf{w}^{\mathrm{II}}_{P\Delta} - \mathbf{w}_P\right\|^{\mathrm{II}}_{\Delta}, \left\|W^{\mathrm{II}}_{\Delta 0} - W_0\right\|^{\mathrm{II}}_{\Delta} \leqslant ce \tag{3.53a}$$

and, for the slow components, also

$$\left\|\mathbf{v}^{\mathrm{II}}_{P\Delta} - \mathbf{v}_P\right\|^{\mathrm{II}}_{\Delta}, \left\|V^{\mathrm{II}}_{\Delta 0} - V_0\right\|^{\mathrm{II}}_{\Delta} \leqslant ch^P, \tag{3.53b}$$

$$\left\|\mathbf{z}^{\mathrm{II}}_{P\Delta} - \mathbf{z}_P\right\|^{\mathrm{II}}_{\Delta}, \left\|Z^{\mathrm{II}}_{\Delta 0} - Z_0\right\|^{\mathrm{II}}_{\Delta} \leqslant c(h^P + \varepsilon e). \quad \square \tag{3.53c}$$

The proof of this theorem is given in Section 5.

The central theorem, summarizing our effors for linear problems with variable coefficients, follows:

THEOREM 3.3. *Assume that the boundary value problem* (1.1)–(1.3) *is well posed, uniformly for* $0 < \varepsilon \leqslant \varepsilon_0$, *and denote the solution by* $\mathbf{x}(t) = \binom{\mathbf{y}(t)}{\mathbf{z}(t)}$. *Assume further that*

$$\det\begin{pmatrix} P_- E^{-1}(0) \\ P_+ E^{-1}(1) \end{pmatrix} \neq 0. \tag{3.54}$$

Then for $0 < \varepsilon \leqslant \varepsilon_1$ *there are positive constants* c_0, δ_0, h_0 *and* κ_0 *such that for any* δ, $0 < c_0\varepsilon \leqslant \delta \leqslant \delta_0$, *and any mesh* $\Delta = \Delta(\varepsilon)$ *satisfying* (3.46), (3.47), *and similar conditions at the right end layer,* $\kappa \leqslant \kappa_0$ *and* $h \leqslant h_0$, *the k-stage collocation scheme based on Gauss or Lobatto points has a unique solution* $\mathbf{x}_\Delta(t)$ *which satisfies*

$$\|\mathbf{x}_\Delta - \mathbf{x}\|_\Delta \leqslant c(e + \delta), \tag{3.55a}$$

where e is defined in part (c) *of Theorem* 3.2.

Further, the following improved estimates hold for the slow components when, up to $O(\varepsilon)$, *the boundary conditions* (1.3) *contain a subset of m linearly independent conditions involving* $\mathbf{z}$ *alone. For Gauss schemes,*

$$\|\mathbf{z}_\Delta - \mathbf{z}\|_\Delta \leqslant c\big(h^P + \bar{\kappa}(e + \delta)\big), \tag{3.55b}$$

while for Lobatto schemes

$$\|\mathbf{z}_\Delta - \mathbf{z}\|_\Delta \leqslant c\big(h^P + h(e + \delta)\big). \tag{3.55c}$$

The proof of Theorem 3.3 is similar in essence to that of Theorem 5.3 in Weiss [9], so we only give a detailed outline here.

Proof Outline for Theorem 3.3. The basic task is to patch together the solution representations (3.37), (3.41) and (3.39) on the three segments of the mesh. The problem in transformed variables (3.23), (3.24) (or (3.34) for short) is considered, first under the boundary conditions

$$P_-\mathbf{u}_1 = \boldsymbol{\eta}_- \in \mathbf{R}^{n_-}, \quad P_+\mathbf{u}_{N+1} = \boldsymbol{\eta}_+ \in \mathbf{R}^{n_+}, \quad \mathbf{v}_1 = \boldsymbol{\eta}_0 \in \mathbf{R}^m. \tag{3.56}$$

This corresponds to the differential problem (2.8), (2.10) which Theorem 2.1 guarantees to be well-behaved for any given parameter vectors $\boldsymbol{\eta}_-$, $\boldsymbol{\eta}_+$ and $\boldsymbol{\eta}_0$.

Thus, the $3(n + m)$ components of the parametric representations (3.37), (3.41) and (3.39), i.e. of $\boldsymbol{\zeta}^\Delta := (\boldsymbol{\zeta}_-^{\mathrm{I}}, \boldsymbol{\zeta}_+^{\mathrm{I}}, \boldsymbol{\zeta}_0^{\mathrm{I}}, \boldsymbol{\zeta}^{\mathrm{II}}, \boldsymbol{\zeta}_0^{\mathrm{II}}, \boldsymbol{\zeta}_-^{\mathrm{III}}, \boldsymbol{\zeta}_+^{\mathrm{III}}, \boldsymbol{\zeta}_0^{\mathrm{III}})$, are fixed by the $3(n + m)$ linear equations consisting of (3.56) plus the matching conditions

$$\mathbf{w}_\Delta^{\mathrm{I}}(t_{\underline{i}}) = \mathbf{w}_\Delta^{\mathrm{II}}(t_{\underline{i}}), \qquad \mathbf{w}_\Delta^{\mathrm{II}}(t_i^-) = \mathbf{w}_\Delta^{\mathrm{III}}(t_i^-). \tag{3.57}$$

In analogy to (2.11), the resulting $3(n + m) \times 3(n + m)$ constraint matrix should have a uniformly bounded inverse for δ, h and κ sufficiently small. Theorems 3.1 and 3.2 furnish us with information on the structure of key blocks of this matrix in (3.45), (3.49) and (3.51). Examining the resulting structure, it becomes apparent that the principal part of the constraint matrix is nonsingular if and only if the matrix

$$\begin{pmatrix} P_- E^{-1}(t_{\underline{i}}) \\ P_+ E^{-1}(t_i^-) \end{pmatrix}$$

is nonsingular. The condition for the latter to hold uniformly in ε is (3.54).

Now Theorem 2.3 guarantees that the exact solution $\mathbf{w}(t)$ has a decomposition similar to that of the discrete problem, with a parameter vector $\boldsymbol{\zeta}$ corresponding to

ζ^{Δ}. Up to exponentially small terms,

$$\zeta = (\zeta_-, \mathbf{0}, \zeta_0, \mathbf{0}, \zeta_0, \mathbf{0}, \zeta_+, \zeta_0),$$

with $\zeta_-, \zeta_+, \zeta_0$ determined by (2.17). The stability of the constraint matrix plus the convergence results of Theorems 3.1 and 3.2 imply that

$$\|\zeta - \zeta^{\Delta}\| \leqslant c(e + \delta). \tag{3.58}$$

Further, considering only those blocks of the constraint matrix which pertain to the smooth solution components and using (3.49), (3.51), (3.53b) and (3.58), we obtain for Gauss schemes

$$\|\zeta_0 - \zeta_0^M\| \leqslant c\big(h^P + \bar{\kappa}(e + \delta)\big), \tag{3.59}$$

and for Lobatto schemes

$$\|\zeta_0 - \zeta_0^M\| \leqslant c\big(h^P + h(e + \delta)\big), \tag{3.60}$$

$M = \mathrm{I}, \mathrm{II}, \mathrm{III}$. Combining (3.58)–(3.60) with (3.44), (3.45), (3.49), (3.51) and (3.53) yields the estimates (3.55a) for $\mathbf{w}$ instead of $\mathbf{x}$ and (3.55b, c) for $\mathbf{v}$ instead of $\mathbf{z}$, for the special set of boundary conditions (3.56). Representing $\mathbf{w}$ in terms of η_-, η_+, η_0 and returning to the original variables

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{y}(t) \\ \mathbf{z}(t) \end{pmatrix} = \begin{pmatrix} E & 0 \\ \varepsilon LE & I \end{pmatrix} \begin{pmatrix} \mathbf{u}(t) \\ \mathbf{v}(z) \end{pmatrix},$$

the parameters η_-, η_+, η_0 are determined such that the original boundary conditions are satisfied, leading to (3.55a, b, c) and completing the proof of Theorem 3.3.

4. Boundary Layer Regions. In this section we consider the linear problem (1.1), (1.2) on the subinterval $[0, T_0\varepsilon]$, where a fine mesh satisfying (3.32) is assumed. Analogous results hold for the subinterval $[1 - T_1\varepsilon, 1]$. Let

$$h_L := \max\{h_i, 1 \leqslant i \leqslant N_0\}. \tag{4.1}$$

Following Weiss [9] we consider the transformed system (2.6)–(2.7) for an easier analysis.

4.1. *Stability and Convergence Results for Smooth and Auxiliary Solution Components.* First, consider one equation

$$\varepsilon y' = \lambda y + f, \qquad 0 \leqslant t \leqslant T_0\varepsilon, \tag{4.2}$$

with $\mathrm{re}(\lambda(t)) \leqslant -\bar{\lambda} < 0$ and $\lambda(t)$ is piecewise constant: $\lambda(t) = \lambda(t_i)$, $t_i \leqslant t < t_{i+1}$. From (3.5) we get for Gauss points

$$y_{ij} = y_i + \frac{h_i}{\varepsilon} \sum_{l=1}^{k} \hat{a}_{jl}\big[\lambda(t_i) y_{il} + f(t_{il})\big], \qquad 1 \leqslant j \leqslant k. \tag{4.3}$$

So, eliminating the local unknowns y_{ij} and substituting into the corresponding expression for y_{i+1}, we get (see Part 1, Section 4)

$$y_{i+1} = \gamma\left(\frac{\lambda(t_i)h_i}{\varepsilon}\right) y_i + \frac{h_i}{\varepsilon} \hat{\mathbf{b}}^T \left[\left(\frac{\varepsilon}{\lambda(t_i)h_i} - \hat{A}\right)^{-1} \hat{A} + I\right] \mathbf{f}_i, \tag{4.4}$$

where $\mathbf{f}_i = (f(t_{i1}), \ldots, f(t_{ik}))^T$, $\mathbf{1} = (1, \ldots, 1)^T \in \mathbf{R}^k$, and

$$\gamma(\zeta) = 1 + \hat{\mathbf{b}}^T(\zeta^{-1}I - \hat{A})^{-1}\mathbf{1} \tag{4.5}$$

is the growth factor. The matrix $\zeta^{-1}I - \hat{A}$ is nonsingular for all ζ satisfying $\mathrm{re}(\zeta) < 0$.

Solving the recurrence relation (4.4), we get

$$(4.6)\quad y_{i+1} = \left[\prod_{l=1}^{i}\gamma\left(\frac{\lambda(t_l)h_l}{\varepsilon}\right)\right]y_1 + \frac{1}{\varepsilon}\sum_{j=0}^{i-1}\left[\prod_{l=i-j+1}^{i}\gamma\left(\frac{\lambda(t_l)h_l}{\varepsilon}\right)\right]h_{i-j}\xi_{i-j}^{T}\mathbf{f}_{i-j},$$

where

$$\xi_{i-j}^{T} := \hat{\mathbf{b}}^{T}\left[\left(\frac{\varepsilon}{\lambda(t_i)h_i} - \hat{A}\right)^{-1}\hat{A} + I\right]$$

is a bounded vector by (3.32). Now, since the method is A-stable, we have

$$(4.7a)\qquad |\gamma(\zeta)| \leqslant 1, \qquad \mathrm{re}(\zeta) < 0.$$

Furthermore, since $\gamma'(0) = 1$, it follows that for any set S of the form

$$S \equiv S(\alpha_1, \alpha_2, \beta) = \left\{\zeta \mid 0 < |\zeta| \leqslant \beta, \frac{\pi}{2} + \alpha_1 \leqslant \arg\zeta \leqslant \frac{3\pi}{2} - \alpha_2\right\}$$

with $\alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 \leqslant \pi$, $\beta < \infty$, there is a positive constant $\mu = \mu(\alpha_1, \alpha_2, \beta)$ such that

$$(4.7b)\qquad |\gamma(\zeta)| \leqslant e^{\mu\,\mathrm{re}(\zeta)}, \qquad \zeta \in S.$$

By (3.32).

$$\left|\frac{\lambda(t_l)h_l}{\varepsilon}\right| \leqslant |\lambda(t_l)|K_0 \leqslant \beta$$

for some well-defined constant β of moderate size. Using (4.7b), we get

$$\left|\prod_{l=i-j+1}^{i}\gamma\left(\frac{\lambda(t_l)h_l}{\varepsilon}\right)\right| \leqslant \prod_{l=i-j+1}^{i} e^{-\mu\bar{\lambda}h_l/\varepsilon} = e^{-\mu\bar{\lambda}(t_{i+1}-t_{i+1-j})/\varepsilon}$$

and

$$\left|\frac{1}{\varepsilon}\sum_{j=0}^{i-1}\left[\prod_{l=i-j+1}^{i}\gamma\left(\frac{\lambda(t_l)h_l}{\varepsilon}\right)\right]h_{i-j}\right| \leqslant \frac{1}{\varepsilon}\sum_{j=0}^{i-1} e^{-\mu\bar{\lambda}(t_{i+1}-t_{i+1-j})/\varepsilon}h_{i-j}$$

$$\leqslant \frac{1}{\varepsilon}\int_{t_1}^{t_{i+1}} e^{-\mu\bar{\lambda}(t_{i+1}-s)/\varepsilon}\,ds \leqslant \frac{1}{\mu\bar{\lambda}}.$$

So, substituting in (4.6), we get

LEMMA 4.1. *When applying a Gauss or a Lobatto difference scheme to* (4.2), *with* $h_j \leqslant K_0\varepsilon$, $1 \leqslant j \leqslant i$, *the following stability result holds*:

$$(4.8)\qquad |y_{i+1}| \leqslant |y_1| + c\|f^c\|,$$

where

$$(4.9)\qquad c = (\bar{\lambda}\mu)^{-1}\max_{1\leqslant j\leqslant i}\|\xi_j\|. \quad \square$$

Note that we have proved the lemma above only for Gauss points. However, it is straightforward to show that a similar result is obtained also for the Lobatto points.

This is the desired stability result for one equation. Next, consider the differential system (2.6)–(2.7) and its corresponding collocation discretization (3.23)–(3.24).

THEOREM 4.1. *The difference equations* (3.23), (3.24), *subject to the boundary conditions*

$$P_-\mathbf{u}_1 = \boldsymbol{\eta}_- \in \mathbf{R}^{n_-}, \quad P_+\mathbf{u}_{\underline{l}} = \boldsymbol{\eta}_+ \in \mathbf{R}^{n_+}, \quad \mathbf{v}_{1'} = \boldsymbol{\eta}_0 \in \mathbf{R}^m, \tag{4.10}$$

have a unique solution $\mathbf{u}_\Delta^c, \mathbf{v}_\Delta^c$ *which satisfies*

$$\|\mathbf{u}_\Delta^c\|, \|\mathbf{v}_\Delta^c\| \leqslant K\{\|\boldsymbol{\eta}_-\| + \|\boldsymbol{\eta}_+\| + \|\boldsymbol{\eta}_0\| + \|\mathbf{g}^c\|\}, \tag{4.11}$$

provided that ε *is small enough*, $0 < \varepsilon \leqslant \varepsilon_0$.

Note: ε_0 is sufficiently small to enable a contraction argument below and depends on the bounds in Lemma 3.1 and on $\max_{0\leqslant t\leqslant T_0\varepsilon}\|\Lambda'(t)\|$. (To recall, by $\mathbf{u}^c$ we mean the restriction of $\mathbf{u}_\Delta(t)$ to the mesh points plus the collocation points.)

Proof. We consider the case for Gauss points; the case for Lobatto points is treated similarly. Our strategy is to consider first the simplified difference equations

$$\frac{\varepsilon}{h_i}(\mathbf{u}_{ij} - \mathbf{u}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\{\Lambda(t_{il})\mathbf{u}_{il} + \mathbf{f}_1(t_{il})\}, \qquad 1 \leqslant j \leqslant k+1, \tag{4.12}$$

$$\frac{1}{h_i}(\mathbf{v}_{ij} - \mathbf{v}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\{B_{22}(t_{il})\mathbf{v}_{il} + \mathbf{g}_2(t_{il})\}, \tag{4.13}$$

where $\mathbf{f}_1(t_{il}) := \mathbf{g}_1(t_{il}) + B_{12}(t_{il})\mathbf{v}_{il}$, and to treat the difference between (4.12)–(4.13) and (3.23)–(3.24) as a perturbation term of order h_L.

The components $\{\mathbf{v}_{ij}\}$ in (4.13), (4.10) are now completely separated from the components $\{\mathbf{u}_{ij}\}$. For $\mathbf{v}_\Delta(t)$ the usual theory applies. This is a Runge-Kutta scheme for a nonstiff initial value problem, and certainly for ε small and h_L satisfying (3.32), $\mathbf{v}_\Delta(t)$ exists and satisfies

$$\|\mathbf{v}^c\| \leqslant c\{\|\boldsymbol{\eta}_0\| + \|\mathbf{g}_2\|\}. \tag{4.14}$$

Now, for (4.12) note that since $\Lambda(t)$ is diagonal, the vector system decouples into n scalar components, so Lemma 4.1 can be applied to each equation separately. For each of the first n_- components we can apply the estimate (4.8) directly, since $\mathrm{re}(\lambda_j(t_i)) < 0$, $1 \leqslant j \leqslant n_-$. For the last n_+ components, $\mathrm{re}(\lambda_j(t_i)) > 0$, and we have to reverse the direction of integration, from right to left. Thus, for such a component, (4.8) is changed to read

$$|y_i| \leqslant |y_{\underline{i}}| + c\|f^c\|, \tag{4.15}$$

which is compatible with the end conditions (4.10). We obtain that the difference equations (4.12) subject to (4.10) possess a solution $\mathbf{u}_\Delta$ satisfying

$$\begin{aligned}\|\mathbf{u}_\Delta\|_\Delta &\leqslant \|\boldsymbol{\eta}_-\| + \|\boldsymbol{\eta}_+\| + c\|\mathbf{f}_1\| \\ &\leqslant \|\boldsymbol{\eta}_-\| + \|\boldsymbol{\eta}_+\| + c_1\{\|g_\Delta\| + \|\boldsymbol{\eta}_0\|\}.\end{aligned} \tag{4.16}$$

It is now easy to show a similar result for $\mathbf{u}_{ij}$ by expressing them in terms of $\mathbf{u}_i$ using (4.12).

This completes the treatment of the major part of the difference operator. Now, the equations (4.12)–(4.13) differ from (3.23)–(3.24) by terms of order h_i (or ε) only, and a standard perturbation argument completes the proof. Q.E.D.

Now, with the stability result (4.11) and the linearity of the problem, part (a) of Theorem 3.1 easily follows. Next, consider the "smooth" components $W_{\Delta 0}(t)$ and $\mathbf{w}_{P\Delta}(t)$. These correspond to the components in the exact solution decomposition which vary slowly across the boundary layer region. Substitution of $W_{\Delta 0} - W_0$ and $\mathbf{w}_{P\Delta} - \mathbf{w}_P$ into (4.11) immediately yields that

$$\|W^c_{\Delta 0} - W^c_0\|, \|\mathbf{w}^c_{P\Delta} - \mathbf{w}^c_P\| \leqslant ch_L = O(\varepsilon), \tag{4.17}$$

and this is really all we need. However, more can be obtained by applying the standard collocation analysis (Russell [7], Weiss [8]) to the original variables (i.e., analyzing the error in (3.5), (3.6)). After transforming back to $\mathbf{w}$, part (b) of Theorem 3.1 is obtained.

Consider part (c) of Theorem 3.1. We write

$$W^{\mathrm{I}}_{\Delta +} = F + G, \quad F = \begin{pmatrix} 0 \\ F_+ \\ 0 \end{pmatrix}, \quad G = \begin{pmatrix} U^{\mathrm{I}}_+ \\ G_+ \\ V^{\mathrm{I}}_+ \end{pmatrix}, \tag{4.18}$$

with F satisfying the homogeneous equations (4.12), (4.13), subject to

$$F(t_i) = \begin{pmatrix} 0 \\ I \\ 0 \end{pmatrix} \tag{4.19}$$

and G is the rest. The difference equations for F are again decoupled, and so A-stability immediately implies that F is bounded. We now have to show that G is small. But comparing (3.36b) with what F satisfies, it is apparent that G satisfies the difference equations (3.23), (3.24), with inhomogeneous terms of size $O(\varepsilon + h_L) = O(\varepsilon)$ and under homogeneous boundary conditions as in (3.36b). Using stability, part (c) of Theorem 3.1 is proven.

4.2. *Mesh Selection in the Layer Regions.* In Section 4.1 we have shown parts (a)–(c) of Theorem 3.1. Here we treat the dominant components of the solution decomposition, $W^{\mathrm{I}}_{\Delta -}(t)$. Analogous results for $W^{\mathrm{III}}_{\Delta +}(t)$ will be omitted.

First, consider one homogeneous equation with constant coefficients,

$$\varepsilon y' = \lambda y, \qquad y(0) = 1, \tag{4.20}$$

with the solution $y(t) = \exp(\lambda t/\varepsilon)$, and denote $\hat{\lambda} := -\mathrm{re}(\lambda) > 0$. In Part I it was shown that, given a tolerance $\delta < 1$, the following mesh generates an approximation accurate to within this tolerance,

$$h_1 := \frac{\varepsilon}{|\lambda|} c_p \delta^{1/p}, \qquad c_p := \left[\frac{\hat{\lambda}}{|\lambda|\,|c_\gamma|} \right]^{1/p}, \tag{4.21}$$

$$h_i := h_{i-1} \exp\left\{ \frac{1}{p} \frac{\hat{\lambda}}{\varepsilon} h_{i-1} \right\} = h_1 \exp\left\{ \frac{1}{p} \frac{\hat{\lambda}}{\varepsilon} t_i \right\}, \qquad i = 2, \ldots, N_0. \tag{4.22}$$

Here c_γ is a known constant depending only on p, and N_0 is determined so that $t_{N_0+1} \geqslant T_0 \varepsilon > t_{N_0}$. Since we would like the contribution of the fast decaying solution

to be below δ on the "long" interval $[t_{\underline{i}}, t_{\bar{i}}]$, it is natural to set T_0 so that

$$\exp\left(\frac{-\hat{\lambda}}{\varepsilon} T_0 \varepsilon\right) = \delta,$$

i.e.

$$T_0 = \frac{1}{\hat{\lambda}}(-\ln \delta). \tag{4.23}$$

Note that the mesh defined by (4.21), (4.22) satisfies (3.32) because its steps are monotonically increasing and $h_{\underline{i}} = \varepsilon c_p/|\lambda|$. Also, beyond $t_{\underline{i}}$ the mesh becomes much sparser, depending only upon the accuracy needs for the reduced solution. Thus, the magnitude of $|y(t_{\underline{i}})|$ is propagated essentially undamped by the numerical scheme outside the layer region.

Next we let λ in (4.20) vary as in (4.2), i.e. $\lambda(t) = \lambda(t_i)$, $t_i \leqslant t < t_{i+1}$. Then (cf. Part I, Section 4.2)

$$y_{i+1} = \prod_{j=1}^{i} \gamma\left(\frac{\lambda(t_j)h_j}{\varepsilon}\right) = \prod_{j=1}^{i} \gamma\left(\frac{\lambda(0)h_j}{\varepsilon}\right) + R_i, \tag{4.24}$$

where

$$R_i = \prod_{j=1}^{i} \gamma\left(\frac{\lambda(0)h_j}{\varepsilon}\right)\left[\sum_{j=1}^{i}\left(1 + O\left(\frac{t_j h_j}{\varepsilon}\right)\right) - 1\right]. \tag{4.25}$$

It is easily verified that

$$|R_i| \leqslant c\varepsilon, \tag{4.26}$$

provided that $\varepsilon(\ln \delta)^2$ is bounded by a constant, see [9, Lemma 4.1]. Thus the mesh (4.21), (4.22) with $\lambda(0)$ replacing λ yields an approximation of $\exp(\lambda(0)t/\varepsilon)$ to within $O(\delta) + O(\varepsilon)$, establishing (3.49a) for the case of one equation.

Turning to the differential system (2.6), (2.7), we once again consider the difference equations (3.23), (3.24) as an $O(\varepsilon)$ perturbation of (4.12), (4.13). The homogeneity and boundary conditions of (3.36a) plus the decoupling of (4.13) from (4.12) clearly imply that $V_{\Delta_-}(t) \equiv 0$ and $P_+ U_{\Delta_-}(t) \equiv 0$. Also, for each of the first n_- components, the previous result for one equation applies, provided that the mesh in (4.21), (4.22) is chosen accordingly. Taking the most stringent of these choices will produce $O(\delta)$ accuracy for all fast components. This is clearly achieved by the choice (3.46), (3.47). The result (3.49) for the slow components is easily obtained by applying standard collocation theory. Part (d) of Theorem 3.1 is then proven and hence, the proof of Theorem 3.1 is complete.

The practical importance of using the mesh (3.46), (3.47) instead of, say, a uniform mesh has been demonstrated in Table 4.2 of Part I. We now supplement this by some a priori estimates of N_0, the number of mesh points needed in the layer.

THEOREM 4.2. *Asymptotically, for ε and δ small,*

$$N_0 = O(\delta^{-1/p}). \quad \square \tag{4.27}$$

Note that in (4.27) N_0 is independent of ε. Also, a uniform mesh with T_0 given by (4.23) would yield $N_0 = O(\delta^{-1/p}(-\ln \delta))$.

Proof. It is sufficient to consider (4.21), (4.22). Then

$$N_0 = \sum_{i=1}^{N_0} h_i/h_i = \frac{1}{h_1}\sum_{i=1}^{N_0} h_i \exp\left\{-\frac{1}{p}\frac{\hat{\lambda}}{\varepsilon}t_i\right\} \approx \frac{1}{h_1}\int_0^{T_0\varepsilon} \exp\left\{-\frac{1}{p}\frac{\hat{\lambda}}{\varepsilon}t\right\} dt$$

$$= -\frac{p\varepsilon}{\hat{\lambda}h_1}\left[1 - \exp\left\{-\frac{\hat{\lambda}T_0}{p}\right\}\right].$$

Substituting (4.23) for T_0 and (4.21) for h_1,

$$N_0 \approx \frac{p}{c_p}\frac{|\lambda|}{\hat{\lambda}}(\delta^{-1/p} - 1) \approx \frac{|\lambda|}{\hat{\lambda}} p c_p^{-1}\delta^{-1/p}. \tag{4.28}$$

This proves our claim. Q.E.D.

Further, the constants c_p of (4.21) can be shown to increase as p is increased (see Part I for $|c_\gamma|$). Thus, the estimate (4.28) also indicates that for a given accuracy δ, N_0 decreases as p (or k) is increased. Note that c_p also reflects a relative efficiency of higher-order methods for problems where the eigenvalues have significant imaginary parts.

The mesh (3.46), (3.47) may be more demanding than necessary in case that eigenvalues of different magnitude are present in $\Lambda_-(t)$. At a given t, $0 \leqslant t \leqslant T_0\varepsilon$, the eigenvalue which imposes the smallest step size is the one for which the magnitude of the pth derivative of the solution, $(|\lambda|/\varepsilon)^p \exp\{-\hat{\lambda}t/\varepsilon\}$, is largest. Thus, if for instance, $|\lambda_1| = \max\{|\lambda_j|,\ j = 1,\ldots,n_-\}$, then we can use (4.21) with $\lambda := \lambda_1$ in place of (3.46) and then construct the mesh using (4.22) (with $\lambda := \lambda_1$) until $t_{i+1} \geqslant \hat{t}_1$, where

$$\hat{t}_1 := \min\{\hat{t}_{1j};\ \hat{t}_{1j} > 0\}, \qquad \hat{t}_{ij} := \varepsilon p \frac{\mathrm{re}(\lambda_j) - \mathrm{re}(\lambda_1)}{\ln|\lambda_1/\lambda_j|}. \tag{4.29}$$

Then, in case that $\hat{t}_1 < T_0\varepsilon$, switch to $\lambda := \lambda_l$ where l gives the minimum in (4.29) and continue with (4.22), etc. However, the overhead involved in constructing such a mesh is worthwhile only in special cases, as described above.

5. The Long Interval. On the "long" interval $[t_i, t_i^-]$ we use the original problem variables and do not apply the transformation (2.5), because we can deal with the system (1.1), (1.2) directly in a simpler fashion. Thus our difference equations are (3.5), (3.6). For ease of presentation we treat Gauss and Lobatto points separately. Throughout this section we suppress the dependence of the problem coefficients on ε.

5.1. *Gauss Points.* To examine the stability of the scheme we consider

$$\frac{\varepsilon}{h_i}(\mathbf{y}_{ij} - \mathbf{y}_i) = \sum_{l=1}^{k} \hat{a}_{jl}(A_{11}(t_{il})\mathbf{y}_{il} + A_{12}(t_{il})\mathbf{z}_{il}) + \mathbf{r}_{ij}, \qquad 1 \leqslant j \leqslant k+1, \tag{5.1}$$

$$\frac{1}{h_i}(\mathbf{z}_{ij} - \mathbf{z}_i) = \sum_{l=1}^{k} \hat{a}_{jl}(A_{21}(t_{il})\mathbf{y}_{il} + A_{22}(t_{il})\mathbf{z}_{il}) + \mathbf{s}_{ij}, \tag{5.2}$$

where $\mathbf{r}_{ij} \in \mathbf{R}^n$, $\mathbf{s}_{ij} \in \mathbf{R}^m$. With

$$\hat{\mathbf{r}}_i^* = \begin{bmatrix} \mathbf{r}_{i1}^* \\ \vdots \\ \mathbf{r}_{ik}^* \end{bmatrix} := (\hat{A} \otimes I)^{-1} \begin{bmatrix} \mathbf{r}_{i1} \\ \vdots \\ \mathbf{r}_{ik} \end{bmatrix} \tag{5.3}$$

and

$$\mathbf{r}_i^* := \varepsilon^{-1} h_i (\mathbf{r}_{i,k+1} - \hat{B}\hat{\mathbf{r}}_i^*), \tag{5.4}$$

we rewrite (5.1) as

$$\frac{\varepsilon}{h_i}(\mathbf{y}_{ij} - \mathbf{y}_i) = \sum_{l=1}^{k} \hat{a}_{jl}(A_{11}(t_{il})\mathbf{y}_{il} + A_{12}(t_{il})\mathbf{z}_{il} + \mathbf{r}_{il}^*), \qquad i \leqslant j \leqslant k, \tag{5.5}$$

$$\frac{\varepsilon}{h_i}(\mathbf{y}_{i+1} - \mathbf{y}_i) = \sum_{l=1}^{k} \hat{b}_l(A_{11}(t_{il})\mathbf{y}_{il} + A_{12}(t_{il})\mathbf{z}_{il} + \mathbf{r}_{il}^*) + \frac{\varepsilon}{h_i}\mathbf{r}_i^*. \tag{5.6}$$

With

$$\hat{\mathbf{f}}_{ij} := A_{12}(t_{ij})\mathbf{z}_{ij} + \mathbf{r}_{ij}^*, \qquad \hat{\mathbf{f}}_i = (\hat{\mathbf{f}}_{i1}, \ldots, \hat{\mathbf{f}}_{ik})^T, \tag{5.7}$$

we can consider (5.5), (5.6) separately from (5.2). As in (3.10)–(3.15) we obtain

$$\mathbf{y}_{i+1} = \Gamma_i \mathbf{y}_i + \mathbf{g}_i + \mathbf{r}_i^*, \tag{5.8}$$

where Γ_i and $\mathbf{g}_i$ are given by (3.15), (3.11)–(3.13) with n replacing $n + m$, $\varepsilon^{-1}A_{11}$ replacing A and $\varepsilon^{-1}\hat{\mathbf{f}}_i$ replacing $\mathbf{f}_i$. We have

$$J_i^{-1} = \left(I - h_i\varepsilon^{-1}D_{A_{11}}(\hat{A} \otimes I)\right)^{-1} = \varepsilon h_i^{-1}\left(\varepsilon h_i^{-1} I - \hat{G}_i\right)^{-1}, \tag{5.9}$$

$$\hat{G}_i = D_{A_{11}}(\hat{A} \otimes I).$$

As in Eq. (4.15) of Part I we write for the singular matrix $\hat{G}_i$

$$\left(\varepsilon h_i^{-1} I - \hat{G}_i\right)^{-1} = \left(\varepsilon h_i^{-1}\hat{C}_i - I\right)\hat{G}_i^{-1}, \tag{5.10}$$

provided that $\varepsilon < h_i \|\hat{G}_i^{-1}\|^{-1}$. Now,

$$\hat{B}\hat{G}_i^{-1}C_{A_{11}} = \hat{B}(\hat{A} \otimes I)^{-1}D_{A_{11}}^{-1}C_{A_{11}} = (\hat{\mathbf{b}}^T\hat{A}^{-1}\mathbf{1})I,$$

where $\mathbf{1} = (1, \ldots, 1) \in \mathbf{R}^k$. From (4.17) of Part I,

$$I - \hat{B}\hat{G}_i^{-1}C_{A_{11}} = (1 - \hat{\mathbf{b}}^T\hat{A}^{-1}\mathbf{1})I = (-1)^k I, \tag{5.11}$$

and this is the leading term of Γ_i. We get in (5.8)

$$\mathbf{y}_{i+1} = (-1)^k \mathbf{y}_i + \varepsilon h_i^{-1}\hat{H}_i\mathbf{y}_i + \hat{B}\left(\varepsilon h_i^{-1}\hat{C}_i - I\right)\hat{G}_i^{-1}\hat{\mathbf{f}}_i + \mathbf{r}_i^*, \tag{5.12}$$

where the matrix $\hat{H}_i$ is bounded and depends on $\hat{\mathbf{b}}$, $\hat{A}$ and $A_{11}(t_{ij})$, $j = 1, \ldots, k$. Clearly, both $\hat{H}_i$ and the matrix multiplying $\hat{\mathbf{f}}_i$ in (5.12) vary smoothly with i. Further, since

$$\mathbf{y}_{ij} = A_{11}^{-1}(t_{ij})\left(-\hat{\mathbf{f}}(t_{ij}) + \varepsilon\mathbf{F}_{ij}\right) \tag{5.13}$$

(cf. (3.9)), we get for

$$\hat{\mathbf{y}}_i := (\mathbf{y}_{i1}, \ldots, \mathbf{y}_{ik})^T \tag{5.14}$$

the equations

$$\hat{\mathbf{y}}_i = -D_{A_{11}}^{-1}\hat{\mathbf{f}}_i - \varepsilon h_i^{-1}D_{A_{11}}^{-1}\big(\varepsilon h_i^{-1}\hat{C}_i - I\big)\hat{G}_i^{-1}\big(\hat{\mathbf{f}}_i + C_{A_{11}}\mathbf{y}_i\big). \tag{5.15}$$

Equations (5.12), (5.15) are equivalent to (5.1). We next state and prove the stability result, recalling (3.50).

LEMMA 5.1. *There are constants κ_0 and h_0 such that, provided $\kappa \leqslant \kappa_0$ and $h \leqslant h_0$, the initial value scheme* (5.12), (5.15), (5.2) *for $i = \underline{i},\ldots,\bar{i}-1$ has a unique solution which satisfies*

$$\|\mathbf{y}_\Delta\|_c \leqslant d, \qquad \|\mathbf{z}_\Delta^c\| \leqslant c\big\{\|\mathbf{z}_{\underline{i}}\| + \bar{\kappa}\|\mathbf{y}_{\underline{i}}\| + \|\mathbf{r}_\Delta^c\| + \|\mathbf{s}_\Delta^c\|\big\}, \tag{5.16}$$

$$\|\mathbf{y}_\Delta\|_\Delta \leqslant d + c \max_{\underline{i}\leqslant i<\bar{i}} \left\| \sum_{j=\underline{i}}^{i-1} (-1)^{k(i-j)}\tilde{\mathbf{r}}_j \right\|, \tag{5.17}$$

where c is a constant,

$$d := c\big\{\|\mathbf{y}_{\underline{i}}\| + \|\mathbf{z}_{\underline{i}}\| + \|\mathbf{r}_\Delta\|_c + \|\mathbf{s}_\Delta^c\|\big\}, \tag{5.18}$$

$$\tilde{\mathbf{r}}_j := \hat{B}G_j^{-1}\hat{\mathbf{r}}_j^* - \mathbf{r}_j^*, \qquad i \leqslant j \leqslant \bar{i}, \tag{5.19}$$

(*Thus $\tilde{\mathbf{r}}_j$ relates to the original inhomogeneities through* (5.3), (5.4).) □

Proof. Consider (5.12), (5.15) with $\varepsilon = 0$ first. Together with (5.2), we refer to this as the "reduced system". Then (5.15) yields

$$\mathbf{y}_{ij} = -A_{11}^{-1}(t_{ij})\big[A_{12}(t_{ij})\mathbf{z}_{ij} + \mathbf{r}_{ij}^*\big], \qquad 1 \leqslant j \leqslant k, \tag{5.20}$$

and substituting into (5.2), we see that

$$h_i^{-1}(\mathbf{z}_{ij} - \mathbf{z}_i) - \sum_{l=1}^{k} \hat{a}_{jl}\big[A_{22}(t_{il}) - A_{21}(t_{il})A_{11}^{-1}(t_{il})A_{12}(t_{il})\big]\mathbf{z}_{il}$$
$$= -\sum_{l=1}^{k} \hat{a}_{jl}A_{21}(t_{il})\mathbf{r}_{il}^* + \mathbf{s}_{ij}, \qquad i \leqslant j \leqslant k+1. \tag{5.21}$$

The difference operator on the left-hand side of this equality is the collocation scheme for the differential operator

$$\mathbf{z}' - \big[A_{22} - A_{21}A_{11}^{-1}A_{12}\big]\mathbf{z}.$$

Hence standard collocation theory not only implies (5.16) (using (5.20)), but also yields the explicit dependence of $\mathbf{z}_{ij}$ on $\mathbf{z}_{\underline{i}}$, $\mathbf{r}_{\mu\nu}^*$, $\mathbf{s}_{\mu\nu}$, in terms of a discrete analogue of the variation of constant formula, as discussed for the box scheme in Weiss [9]. To derive (5.17), note first that by (5.12),

$$\mathbf{y}_i = (-1)^{k(i-\underline{i})}\mathbf{y}_{\underline{i}} - \sum_{j=\underline{i}}^{i-1} (-1)^{k(i-1-j)}\big(\hat{B}\hat{G}_j^{-1}\hat{\mathbf{f}}_j - \mathbf{r}_j^*\big). \tag{5.22}$$

Hence the contribution of $\tilde{\mathbf{r}}_j$ to (5.17) is correct, in view of (5.7), (5.19).

The remaining term to be dealt with in (5.22) is

$$\hat{B}\sum_{j=\underline{i}}^{i-1}(-1)^{k(i-1-j)}G_j^{-1}\begin{pmatrix}A_{12}(t_{j1})z_{j1}\\ \vdots \\ A_{12}(t_{jk})z_{jk}\end{pmatrix}. \tag{5.23}$$

We distinguish between two cases. If k is odd, then we essentially have to consider a sum of pairs of terms of the form

$$A_{12}(t_{jl})\mathbf{z}_{jl} - A_{12}(t_{j-1,l})\mathbf{z}_{j-1,l}. \tag{5.24}$$

In view of the discrete variation of constant formula mentioned above, the expression in (5.24) is bounded by $ch_j(\|\mathbf{z}_{\underline{i}}\| + \|\mathbf{r}_\Delta^*\|_c + \|\mathbf{s}_\Delta^c\|)$. The sum in (5.23) is bounded and (5.17) is obtained. If k is even, then there is no sign alternation in (5.23), but by (5.9), (5.11),

$$\hat{B}\hat{G}_j^{-1} = \hat{B}(\hat{A}\otimes I)^{-1}D_{A_{11}}^{-1}, \qquad \hat{\mathbf{b}}^T\hat{A}^{-1}\mathbf{1} = 0. \tag{5.25}$$

Thus the jth term in the sum of (5.23) can be written as

$$\hat{B}(\hat{A}\otimes I)^{-1}\begin{pmatrix}A_{11}^{-1}(t_{j1})A_{12}(t_{j1})\mathbf{z}_{j1}\\ \vdots \\ A_{11}^{-1}(t_{jk})A_{12}(t_{jk})\mathbf{z}_{jk}\end{pmatrix} \tag{5.26}$$

$$= \hat{B}(\hat{A}\otimes I)^{-1}\left[\begin{pmatrix}A_{11}^{-1}(t_{j1})A_{12}(t_{j1})\mathbf{z}_{j1}\\ \vdots \\ A_{11}^{-1}(t_{jk})A_{12}(t_{jk})\mathbf{z}_{jk}\end{pmatrix} - \begin{pmatrix}A_{11}^{-1}(t_{j1})A_{12}(t_{j1})\mathbf{z}_{j1}\\ \vdots \\ A_{11}^{-1}(t_{j1})A_{12}(t_{j1})\mathbf{z}_{j1}\end{pmatrix}\right].$$

This brings us to examine again terms like (5.24) and the remainder of the proof is, therefore, as for the case when k is odd, yielding (5.17).

Next we turn to the "full system" (5.12), (5.15), (5.2), with $\varepsilon > 0$. We write it as the reduced system plus additional terms of size εh_i^{-1}. We also write the solution as

$$\mathbf{y}_\Delta(t) = \bar{\mathbf{y}}_\Delta(t) + \boldsymbol{\eta}_\Delta(t), \qquad \mathbf{z}_\Delta(t) = \bar{\mathbf{z}}_\Delta(t) + \boldsymbol{\zeta}_\Delta(t), \tag{5.27}$$

where $\bar{\mathbf{y}}_\Delta(t)$, $\bar{\mathbf{z}}_\Delta(t)$ solve the reduced system and

$$\boldsymbol{\eta}_\Delta(t_{\underline{i}}) = \boldsymbol{\zeta}_\Delta(t_{\underline{i}}) = 0.$$

Then $\boldsymbol{\eta}_\Delta(t)$, $\boldsymbol{\zeta}_\Delta(t)$ satisfy the "full system" with inhomogeneities $\mathbf{r}_i^*$ and $\mathbf{r}_{ij}^*$ of size εh_i^{-1} and $\mathbf{s}_{ij} = \mathbf{0}$. Applying (5.16), (5.17) for the reduced system, we get the bounds

$$\begin{aligned}\|\boldsymbol{\eta}_\Delta^c\| &\leqslant c\hat{\kappa}\left(\|\bar{\mathbf{y}}_\Delta^c\| + \|\bar{\mathbf{z}}_\Delta^c\| + \|\boldsymbol{\eta}_\Delta^c\| + \|\boldsymbol{\zeta}_\Delta^c\|\right),\\ \|\boldsymbol{\zeta}_\Delta^c\| &\leqslant c\bar{\kappa}\left(\|\bar{\mathbf{y}}_\Delta^c\| + \|\bar{\mathbf{z}}_\Delta^c\| + \|\boldsymbol{\eta}_\Delta^c\| + \|\boldsymbol{\zeta}_\Delta^c\|\right),\end{aligned} \tag{5.28}$$

where

$$\hat{\kappa} := \varepsilon\sum_{i=\underline{i}}^{\bar{i}-1}h_i^{-1} \leqslant \kappa. \tag{5.29}$$

For κ small enough, a contraction argument now yields the bounds (5.16), (5.17) for the full system. This completes the proof of Lemma 5.1. Q.E.D.

Part (a) of Theorem 3.2 now follows for Gauss points.

Next, consider the first n fundamental solution components. The collocation approximations $Y_\Delta(t)$, $Z_\Delta(t)$ are defined by the homogeneous difference schemes (3.5), (3.6) and the initial conditions (3.40b). Thus, for the "reduced" problem with $\varepsilon = 0$ we get

$$\bar{Z}_\Delta(t) \equiv 0, \qquad \bar{Y}_\Delta(t_i) = (-1)^{k(i-\underline{i})} I, \qquad \underline{i} \leqslant i < \bar{i}. \tag{5.30}$$

Furthermore, by repeating the argument of the above lemma it becomes clear that $Y_\Delta(t_i)$, $Z_\Delta(t_i)$ are only $O(\kappa)$, $O(\bar{\kappa})$ away from $\bar{Y}_\Delta(t_i)$, $\bar{Z}_\Delta(t_i)$, respectively. Thus, applying the transformation (3.22), (3.51a) is obtained.

Finally, for a smooth solution $\mathbf{x}(t) = \binom{\mathbf{y}(t)}{\mathbf{z}(t)}$, which may be a transformation of $\mathbf{w}_p(t)$ or of a column of $W_0(t)$, consider the approximation error. We write the difference scheme (5.1), (5.2) for the error, with

$$\mathbf{r}_{ij} = \varepsilon O\big(h_i^k\big), \quad \mathbf{s}_{ij} = O\big(h_i^k\big), \quad \mathbf{r}_{i,k+1} = \varepsilon O\big(h_i^p\big)$$

with $p = 2k$. From Lemma 5.1 it immediately follows that

$$\|\mathbf{y} - \mathbf{y}_\Delta\| = O(h^k), \qquad \|\mathbf{z} - \mathbf{z}_\Delta\| = O(h^k). \tag{5.31}$$

Furthermore, we now show that sharper bounds hold.

Assume for simplicity that the mesh on the long interval is quasiuniform (this assumption can be easily dispensed with, as in de Boor and Swartz [11]). Then, since $\mathbf{y}_\Delta$, $\mathbf{z}_\Delta$ are piecewise polynomials of degree at most k and $\mathbf{y}, \mathbf{z}$ are smooth, we obtain using (5.31)

$$\big\|(\mathbf{y} - \mathbf{y}_\Delta)^{(l)}\big\|, \big\|(\mathbf{z} - \mathbf{z}_\Delta)^{(l)}\big\| \leqslant \text{const}, \qquad l = 0, 1, \ldots, p+1. \tag{5.32}$$

Consider the approximation error in

$$\mathbf{e}_\Delta(t) := \mathbf{v}_\Delta(t) - \mathbf{v}(t) = -\varepsilon L(t)\big(\mathbf{y}_\Delta(t) - \mathbf{y}(t)\big) + \mathbf{z}_\Delta(t) - \mathbf{z}(t) \tag{5.33}$$

(cf. (2.5), (3.22)). Since (2.7) does not contain the fast solution components any more, we can follow [11], even though $\mathbf{v}_\Delta(t)$ is not a polynomial. Thus we write

$$L\mathbf{e}_\Delta(t) := \mathbf{e}_\Delta'(t) - B_{22}(t)\mathbf{e}_\Delta(t) := \boldsymbol{\psi}_\Delta(t) \tag{5.34}$$

and note that, since $\boldsymbol{\psi}_\Delta(t_{ij}) = 0$, $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant k$, $\boldsymbol{\psi}_\Delta(t)$ can be written for $t_i \leqslant t \leqslant t_{i+1}$ as a remainder of polynomial interpolation using the divided difference form,

$$\boldsymbol{\psi}_\Delta(t) = \boldsymbol{\psi}_\Delta[t_{i1}, \ldots, t_{ik}, t] \prod_{l=1}^{k} (t - t_{il}). \tag{5.35}$$

Upon writing $\mathbf{e}_\Delta(t)$ in terms of $\boldsymbol{\psi}_\Delta(t)$, using Green's function, it becomes apparent that for $t_i \leqslant t \leqslant t_{i+1}$, $1 \leqslant i \leqslant N$,

$$\mathbf{e}_\Delta\big(t_{ij}\big) = h_i^{k+1}\big(\boldsymbol{\varphi}_j(t_i) + O(h_i)\big) + O(h^p), \tag{5.36}$$

for some smooth functions $\boldsymbol{\varphi}_j(t)$, and

$$\|\mathbf{e}_\Delta(t_i)\| = O(h^p), \tag{5.37}$$

provided that $\boldsymbol{\psi}_\Delta^{(k)}(t)$ is bounded; see [11]. But, substituting (5.33) into (5.34) and using (5.32), this is evidently true, hence (5.36), (5.37) are established, yielding (3.53b).

Now we replace $\mathbf{z} - \mathbf{z}_\Delta$ by $\mathbf{e}_\Delta$ in (5.1) and, using (5.36), apply the stability result (5.17), obtaining

$$\|\mathbf{y}_\Delta - \mathbf{y}\|_\Delta \leqslant c\left(\sum_{j=\underline{i}}^{\bar{i}-1} (-1)^{k(\bar{i}-1-j)} K_j h_j^{k+1} + h^{k+1}\right)$$

for suitable constants K_j satisfying $K_j - K_{j-1} = O(h)$.

Thus, if k is odd and (3.52) holds,

$$\|\mathbf{y}_\Delta - \mathbf{y}\|_\Delta = O(h^{k+1}). \tag{5.38}$$

Now, for the slow components $\mathbf{z}$, the improved estimate

$$\|\mathbf{z} - \mathbf{z}_\Delta\| = O(h^p) + \varepsilon\|\mathbf{y} - \mathbf{y}_\Delta\| \tag{5.39}$$

(with (5.31) or (5.38) used for $\mathbf{y}$) is obtained by combining (5.33) and (5.37). This completes the proof of Theorem 3.2 for Gauss points.

5.2. *The Case of Lobatto Points.* For Lobatto points we proceed, as before, to establish stability first. Consider

$$\varepsilon h_i^{-1}(\mathbf{y}_{ij} - \mathbf{y}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\{A_{11}(t_{il})\mathbf{y}_{il} + A_{12}(t_{il})\mathbf{z}_{il}\} + \mathbf{r}_{ij}, \tag{5.40}$$

$$\underline{i} \leqslant i \leqslant \bar{i},\ 2 \leqslant j \leqslant k,$$

$$h_i^{-1}(\mathbf{z}_{ij} - \mathbf{z}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\{A_{21}(t_{il})\mathbf{y}_{il} + A_{22}(t_{il})\mathbf{z}_{il}\} + \mathbf{s}_{ij}. \tag{5.41}$$

Rewrite (5.40) as

$$\varepsilon h_i^{-1}(\mathbf{y}_{ij} - \mathbf{y}_i) = \sum_{l=1}^{k} \hat{a}_{jl}\{A_{11}(t_{il})\mathbf{y}_{il} + A_{12}(t_{il})\mathbf{z}_{il} + \mathbf{r}_{il}^*\}, \qquad 2 \leqslant j \leqslant k, \tag{5.42}$$

where

$$\sum_{l=1}^{k} \hat{a}_{jl}\mathbf{r}_{il}^* = \mathbf{r}_{ij}, \qquad \mathbf{r}_{il}^* = \mathbf{r}_{i-1,k}^*, \tag{5.43}$$

so

$$\bar{\mathbf{r}}_i^* := \begin{pmatrix} \mathbf{r}_{i2}^* \\ \vdots \\ \mathbf{r}_{ik}^* \end{pmatrix} = (\bar{A} \otimes I)^{-1}\left\{\begin{pmatrix} \mathbf{r}_{i2} \\ \vdots \\ \mathbf{r}_{ik} \end{pmatrix} - (\bar{a} \otimes I)\mathbf{r}_{i1}^*\right\}. \tag{5.44}$$

(See Lemma 3.1 of Part I for notation.) Denoting the last rows of $(\bar{A} \otimes I)^{-1}$ by $[\hat{a}_{k2}^- I, \ldots, \hat{a}_{kk}^- I]$, we have $\sum_{l=1}^{k} \hat{a}_{kl}^- \hat{a}_{l1} = -\gamma(-\infty) = (-1)^k$ (cf. Eqs. (4.19), (4.20) of Part I). Thus, the last n rows of (5.44) yield the recursion

$$\mathbf{r}_{ik}^* = (-1)^{k+1}\mathbf{r}_{i-1,k}^* + \tilde{\mathbf{r}}_i, \qquad \underline{i} \leqslant i < \bar{i}, \tag{5.45}$$

where

$$\tilde{\mathbf{r}}_i := \sum_{l=2}^{k} \hat{a}_{kl}^- \mathbf{r}_{il}, \qquad \underline{i} \leqslant i < \bar{i}, \tag{5.46}$$

whence

$$\mathbf{r}_{ik}^{*} = \sum_{j=i}^{i} (-1)^{(k+1)(i-j)} \tilde{\mathbf{r}}_j + (-1)^{(k+1)(i-i-1)} \mathbf{r}_{\underline{i}1}^{*}, \tag{5.47}$$

and $\mathbf{r}_{\underline{i}1}^{*}$ is an initial value parameter. Thus, using (5.47), (5.46) and (5.44), we may consider the form (5.42) for arbitrary inhomogeneities $\mathbf{r}_{ij}$ and initial values $\mathbf{y}_{\underline{i}}$.

Next, with $\hat{\mathbf{f}}_{ij}$ as in (5.7), we obtain (5.8) where Γ_i and $\mathbf{g}_i$ are given by (3.21), (3.20), with n replacing $n + m$, $\varepsilon^{-1}A_{11}$ replacing A and $\varepsilon^{-1}\hat{\mathbf{f}}_i$ replacing $\mathbf{f}_i$. Then,

$$\bar{J}_i^{-1} = \varepsilon h_i^{-1}\big(\varepsilon h_i^{-1} I - \bar{G}_i\big)^{-1}, \qquad \bar{G}_i = (\bar{A} \otimes I)\bar{D}_{A_{11}}, \tag{5.48}$$

and we write, as in (5.10), provided that $\varepsilon < h_i \|\bar{G}_i^{-1}\|^{-1}$,

$$\big(\varepsilon h_i^{-1} I - \bar{G}_i\big)^{-1} = \big(\varepsilon h_i^{-1}\bar{C}_i - I\big)\bar{G}_i. \tag{5.49}$$

The last rows of $\bar{G}_i^{-1}$ are

$$A_{11}^{-1}(t_{i+1})\big(\hat{a}_{k2}^{-} I, \ldots, \hat{a}_{kk}^{-} I\big). \tag{5.50}$$

Substituting in (3.21) yields that the difference equations (5.42) are equivalent to

$$\begin{aligned} \mathbf{y}_{i+1} = {} & (-1)^{k+1} A_{11}^{-1}(t_{i+1}) A_{11}(t_i)\mathbf{y}_i + A_{11}^{-1}(t_{i+1})\big((-1)^k \hat{\mathbf{f}}_{i1} - \hat{\mathbf{f}}_{ik}\big) \\ & + \varepsilon h_i^{-1}\big(\bar{H}_i \mathbf{y}_i + \bar{Q}\hat{\mathbf{f}}_i\big), \end{aligned} \tag{5.51}$$

where $\bar{H}_i$ and $\bar{Q}_i$ are bounded matrices independently of ε, which vary smoothly with i. For the other collocation points we obtain in precisely the same way (cf. (3.18))

$$\mathbf{y}_{ij} = c_{ij} A_{11}^{-1}(t_{ij}) A_{11}(t_i)\mathbf{y}_i + \varepsilon h_i^{-1}\bar{H}_{ij}\mathbf{y}_i + \big(\varepsilon h_i^{-1}\bar{C}_i - I\big)\bar{Q}_{ij}\hat{\mathbf{f}}_i, \qquad 2 \leqslant j \leqslant k-1, \tag{5.52}$$

for appropriate constants c_{ij} and bounded matrices $\bar{H}_{ij}$, $\bar{Q}_{ij}$. Now, to obtain the equivalent of Lemma 5.1, set $\varepsilon = 0$ in (5.42) (or (5.51)) first. Then

$$\mathbf{y}_{il} = -A_{11}^{-1}(t_{il})\big(A_{12}(t_{il})\mathbf{z}_{il} + \mathbf{r}_{il}^{*}\big), \qquad 1 \leqslant i \leqslant N, 1 \leqslant l \leqslant k. \tag{5.53}$$

This is substituted in (5.41), obtaining

$$\begin{aligned} h_i^{-1}(\mathbf{z}_{ij} - \mathbf{z}_i) = {} & \sum_{l=1}^{k} \hat{a}_{jl}\big\{\big[A_{22}(t_{il}) - A_{21}(t_{il})A_{11}^{-1}(t_{il})A_{12}(t_{il})\big]\mathbf{z}_{il} \\ & - A_{21}(t_{il})A_{11}^{-1}(t_{il})\mathbf{r}_{il}^{*}\big\} + \mathbf{s}_{ij}, \\ & 1 \leqslant i \leqslant N, 2 \leqslant j \leqslant k. \end{aligned} \tag{5.54}$$

These are collocation equations for the decoupled slow components and the usual stability theory applies. Since, by (5.43), (5.47) and (5.53),

$$\sum_{l=1}^{k} \hat{a}_{jl} A_{21}(t_{il}) A_{11}^{-1}(t_{il})\mathbf{r}_{il}^{*} = A_{21}(t_i)A_{11}^{-1}(t_i)\mathbf{r}_{ij} + O\big(\|\mathbf{r}_{\Delta}^{c}\|\big) + O(h_i)\|\mathbf{y}_{\underline{i}}\|,$$

we obtain

$$\|\mathbf{z}_{\Delta}^{c}\| \leqslant c\big\{\|\mathbf{z}_{\underline{i}}\| + h\|\mathbf{y}_{\underline{i}}\| + \|\mathbf{s}_{\Delta}^{c}\| + \|\mathbf{r}_{\Delta}^{c}\|\big\}. \tag{5.55}$$

Substitution in (5.53) and use of (5.47) then yield

$$\|\mathbf{y}_{\Delta}^{c}\| \leqslant c\big\{\|\mathbf{y}_{\underline{i}}\| + \|\mathbf{z}_{\underline{i}}\| + \|\mathbf{s}_{\Delta}^{c}\| + \|\mathbf{r}_{\Delta}^{*c}\|\big\}, \tag{5.56}$$

$$\|\mathbf{y}_\Delta^c\| \leq c\left\{\|\mathbf{y}_{\underline{i}}\| + \|\mathbf{z}_{\underline{i}}\| + \|\mathbf{s}_\Delta^c\| + \|\mathbf{r}_\Delta^c\| + \max_{\underline{i}\leq i<\bar{i}}\left\|\sum_{j=\underline{i}}^{i}(-1)^{(k+1)(i-j)}\tilde{\mathbf{r}}_j\right\|\right\}. \tag{5.57}$$

A standard contraction argument for $\varepsilon > 0$ small then completes the proof of the following lemma.

LEMMA 5.2. *There are constants κ_0 and h_0 such that for $\kappa \leq \kappa_0$, $h \leq h_0$, the difference scheme* (5.41), (5.42) *with $\mathbf{y}_{\underline{i}}$, $\mathbf{z}_{\underline{i}}$ prescribed has a unique solution which satisfies* (5.55)–(5.57).

Remark. Comparing Lemma 5.2 to Lemma 5.1, note that unlike for Gauss points, the contribution of $\mathbf{y}_{\underline{i}}$ in (5.55) does not vanish as $\varepsilon \to 0$. On the other hand, no bound like (5.56) is available for Gauss schemes.

This establishes part (a) of Theorem 3.2 for Lobatto points. Next, consider the collocation approximation to $Y_\Delta(t)$, $Z_\Delta(t)$. Inserting the trial solution

$$\bar{Z}_\Delta(t) \equiv 0, \qquad \bar{Y}_\Delta(t_i) = (-1)^{(k+1)(i-\underline{i})}A_{11}^{-1}(t_i)A_{11}(t_{\underline{i}}) \tag{5.58}$$

into the (homogeneous) difference equations (5.41), (5.42), we readily find residuals of order h_i for Z_Δ and εh_i^{-1} for Y_Δ. Lemma 5.2 then yields (3.51b).

Finally, consider the approximation error for a smooth solution $\mathbf{x}(t)$. Writing the difference equations for the error, we have

$$\mathbf{r}_{ij} = \varepsilon O(h_i^k), \quad \mathbf{s}_{ij} = O(h_i^k), \quad \mathbf{r}_{ik} = \varepsilon O(h_i^p), \quad \mathbf{s}_{ik} = O(h_i^p),$$

where $p = 2(k-1)$. Proceeding precisely as for Gauss points, we obtain from Lemma 5.2 the estimates (5.32)–(5.37).

Substituting these estimates in (5.51) interpreted as the equations for the error in the fast components, we obtain inhomogeneities of the form $(-1)^k\mathbf{q}_i - \mathbf{q}_{i+1}$ with $\|\mathbf{q}_i\| = O(h^p)$, $\underline{i} \leq i \leq \bar{i}$, and of the form $\varepsilon h_i^k \boldsymbol{\varphi}(t_{ij})$, where $\boldsymbol{\varphi}(t)$ is smooth. As in Theorem 5.2 of Part I (cf. (5.15) there) this readily yields the bound

$$\|\mathbf{y}_\Delta - \mathbf{y}\|_\Delta \leq ce, \tag{5.59}$$

with e defined in part (c) of Theorem 3.2. Now, combine (5.33), (5.37), (5.59) to obtain for the slow solution components,

$$\|\mathbf{z}_\Delta - \mathbf{z}\|_\Delta \leq c(h^p + \varepsilon e). \tag{5.60}$$

This completes the proof of Theorem 3.2.

6. Numerical Examples. The following numerical results were computed on an Amdahl 470-V/8 computer with a 14-hexadecimal-digits mantissa. The notation $a - b \equiv a \times 10^{-b}$ is used throughout.

Example (*Hemker* [3]). Consider

$$\varepsilon u'' + (2 + \cos \pi t)u' - u = f(t), \qquad 0 \leq t \leq 1, \tag{6.1a}$$

where

$$f(t) = -(1 + \varepsilon\pi^2)\cos \pi t - \pi(2 + \cos \pi t)\sin \pi t + \left(1 - \alpha + \frac{3}{2\varepsilon}\pi^2 t^2\right)e^{-3t/\varepsilon} \tag{6.1b}$$

subject to

$$u(0) = \alpha, \qquad u(1) = -1. \tag{6.1c}$$

The solution is

$$u(t) = \cos \pi t + (\alpha - 1)e^{-3t/\varepsilon} + O(\varepsilon^2). \tag{6.2}$$

Thus, when $\alpha \neq 1$ we have a boundary layer at $t = 0$ only.

When converting to a first order system note that if we use the usual variables u, u', then the problem does not have a bounded inverse (since $u'(0) \sim 1/\varepsilon$). Instead we integrate once, as in Kreiss and Kreiss [4], Kreiss and Nichols [5], obtaining with $y := u$ the system

$$\varepsilon y' = -(2 + \cos \pi t)y + z, \tag{6.3a}$$

$$z' = (1 - \pi \sin \pi t)y + f(t), \tag{6.3b}$$

$$y(0) = \alpha, \qquad y(1) = -1. \tag{6.3c}$$

So our matrix A_{11} is a negative scalar function of t here.

First we choose $\alpha = 1$ and use uniform meshes. The results are listed in Table 1, where under "E" we list the maximum error at mesh points and under "rate" the measured convergence rate in h. The results for Gauss and Lobatto points confirm part (c) of Theorem 3.2. In addition, we list for comparison numerical results using collocation at the unsymmetric Radau points (see Part I). The usage of the latter schemes is possible here because all the eigenvalues of A_{11} have the same sign in their real part. For the examples discussed in Weiss [9] or in Section 6 of Part I, for instance, the Radau schemes are unstable unless the transformation (2.5) is explicitly applied (and this time not just for analysis) and the schemes are upwinded. Therefore, we stay with the symmetric schemes.

TABLE 1

Example 1 *with a smooth solution throughout*, $\varepsilon = 10^{-10}$

	Gauss points					Radau points					Lobatto points				
k	N	cond	E	rate	k	N	cond	E	rate	k	N	cond	E	rate
1	10	.23 + 3	.64 − 1		1	10	.62 + 2	.20		2	10	.40 + 3	.65 − 1	
	20	.45 + 3	.16 − 1	2.0		20	.12 + 3	.10	0.9		20	.76 + 3	.17 − 1	2.0
	40	.88 + 3	.40 − 2	2.0		40	.25 + 3	.53 − 1	1.0		40	.15 + 4	.43 − 2	2.0
2	10	.87 + 2	.47 − 2		2	10	.66 + 2	.33 − 3		3	10	.16 + 3	.30 − 4	
	20	.16 + 3	.12 − 2	2.0		20	.13 + 3	.40 − 4	3.0		20	.29 + 3	.19 − 5	4.0
	40	.31 + 3	.29 − 3	2.0		40	.25 + 3	.49 − 5	3.0		40	.55 + 3	.12 − 6	4.0
3	10	.23 + 3	.16 − 3		3	10	.66 + 2	.18 − 5		4	10	.40 + 3	.41 − 6	
	20	.45 + 3	.98 − 5	4.0		20	.13 + 3	.54 − 7	5.0		20	.76 + 3	.68 − 8	5.9
	40	.88 + 3	.61 − 6	4.0		40	.25 + 3	.17 − 8	5.0		40	.15 + 4	.11 − 9	6.0
4	10	.88 + 2	.88 − 5		4	10	.66 + 2	.21 − 8		5	10	.16 + 3	.70 − 10	
	20	.16 + 3	.55 − 6	4.0		20	.13 + 3	.17 − 10	7.0		20	.29 + 3	.28 − 12	8.0
	40	.31 + 3	.34 − 7	4.0		40	.25 + 3	.13 − 12	7.0		40	.55 + 3	.12 − 13	*

*rate polluted by roundoff errors

Next we set $\alpha = 0$, obtaining a steep boundary layer near $t = 0$. Results are listed in Table 2. Here the meshes are constructed by taking the corresponding meshes of Table 1 and adding a layer mesh according to (4.21), (4.22) with $\hat{\lambda} = -\lambda = 3$. The accuracy tolerance δ is chosen to be just below the smooth solution error for the

finest mesh in Table 1, for each scheme. Here we list under "E" the maximum error of all mesh points with "rate" the convergence rate in the maximum mesh width h. Note the relatively small number of mesh points needed to achieve high accuracy with the higher-order schemes, particularly of Lobatto types.

Also listed in Table 2 are some results when $\delta \ll \varepsilon \ll 1$. This is not covered by our analysis (see part (d) of Theorem 3.1), because we are primarily concerned in this paper with what happens when $\varepsilon \to 0$. However, as indicated by the numerical results, the analysis can be extended to cover this case as well. Indeed, when $\delta \ll \varepsilon$, then a denser mesh is constructed in the layer regions and this only makes the situation more regular.

Other examples have been tried as well. In particular, numerical solutions for the example in Weiss [9], which for some particular values violates condition (3.54), have been computed. Their behavior is similar to that reported in [9] for the midpoint and trapezoidal schemes and their discussion is therefore omitted.

TABLE 2

Example 1 *with a boundary layer*, $\alpha = 0$

	Gauss points					Lobatto points				
ε	k	δ	N	E	rate	k	δ	N	E	rate
10^{-10}	1	1.–3	32	.21 – 1		2	1.–3	32	.13 – 1	
			42	.54 – 2	2.0			42	.32 – 2	2.0
			62	.15 – 2	1.8			62	.80 – 3	2.0
	2	1.–4	20	.63 – 2		3	1.–7	57	.22 – 4	
			30	.16 – 2	2.0			67	.13 – 5	4.0
			50	.39 – 3	2.0			87	.82 – 7	4.0
	3	1.–7	26	.10 – 3		4	1.–10	54	.75 – 7	
			36	.62 – 5	4.1			64	.11 – 8	6.0
			56	.39 – 6	4.0			84	.10 – 9	3.5
	4	1.–8	22	.12 – 4		5	1.–10	30	.11 – 9	
			32	.73 – 6	4.0			40	.70 – 10	
			52	.45 – 7	4.0			60	.70 – 10	
10^{-4}	3	1.–7	25	.10 – 3		3	1.–7	56	.20 – 4	
			35	.62 – 5	4.1			66	.11 – 5	4.2
			55	.38 – 6	4.0			86	.86 – 7	3.7
	4	1.–8	21	.12 – 4		4	1.–10	53	.61 – 7	
			31	.66 – 6	4.1			63	.11 – 8	5.7
			51	.26 – 7	4.6			83	.94 – 10	3.6

Department of Computer Science
University of British Columbia
Vancouver, B.C., Canada V6T 1W5

Institut für Angewandte und Numerische Mathematik
Technische Universität Wien
Gusshausstrasse 27-29
1040 Wien, Austria

1. U. ASCHER, S. PRUESS & R. D. RUSSELL, "On spline basis selection for solving differential equations," *SIAM J. Numer. Anal*, v. 20, 1983, pp. 121–142.

2. U. ASCHER & R. WEISS, "Collocation for singular perturbation problems I: First order systems with constant coefficients," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 537–557.

3. P. W. HEMKER, *A Numerical Study of Stiff Two-Point Boundary Value Problems*, Math. Centrum, Amsterdam, 1977.

4. B. KREISS & H. O. KREISS, "Numerical methods for singular perturbation problems," *SIAM J. Numer. Anal.*, v. 18, 1981, pp. 262–276.

5. H. O. KREISS & N. NICHOLS, *Numerical Methods for Singular Perturbation Problems*, Dept. of Computer Science Report #57, Uppsala University, 1975.

6. P. A. MARKOWICH & C. A. RINGHOFER, "Collocation methods for boundary value problems on 'long' intervals," *Math. Comp.*, v. 40, 1983, pp. 123–150.

7. R. D. RUSSELL, "Collocation for systems of boundary value problems," *Numer. Math.* v. 23, 1974, 119–133.

8. R. WEISS, "The application of implicit Runge-Kutta and collocation methods to boundary-value problems," *Math. Comp.*, v. 28, 1974, pp. 449–464.

9. R. WEISS, "An analysis of the box and trapezoidal schemes for linear singularly perturbed boundary value problems," *Math. Comp.*, v. 42, 1984, pp. 41–67.

10. U. ASCHER & R. WEISS, *Collocation for Singular Perturbation Problems* III: *Nonlinear Problems Without Turning Points*, Tech. Report 82–9, Univ. of British Columbia, Vancouver, Canada.

11. C. DE BOOR & B. SWARTZ, "Collocation at Gaussian points," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 582–606.

12. H. O. KREISS, *Centered Difference Approximation to Singular Systems of ODEs*, Symposia Mathematica X (1972), Istituto Nazionale di Alta Matematica.

13. P. SPUDICH & U. ASCHER, "Calculation of complete theoretical seismograms in vertically varying media using collocation methods," *Geoph. J. Roy. Astr. Soc.*, v. 75, 1983, pp. 101–124.