On the Numerical Solution of Singular Boundary Value Problems of Second Order by a Difference Method

By Ewa Weinmüller

Abstract. The standard three-point discretization applied to the numerical solution of nonlinear boundary value problems for second-order systems with a singularity of the first kind is investigated. The results are extended to the boundary value problems arising in practical problems from mechanics and chemistry. A number of numerical examples illustrating the theoretical results is presented.

1. Introduction. The main purpose of this paper is to investigate the application of a finite difference scheme (with three-point discretization) to the following nonlinear boundary value problems:

\begin{align}
(1.1a) \quad y''(t) - \frac{A_1}{t} y'(t) - \frac{A_0}{t^2} y(t) &= f\left(t, y(t), y'(t)\right), \quad 0 < t \leq 1, \\
(1.1b) \quad B\left(y(0); y(1), y'(1)\right) &= 0,
\end{align}

and

\begin{align}
(1.2a) \quad y''(t) - \frac{A_1}{t} y'(t) - \frac{A_0}{t^2} y(t) &= f\left(t, y(t), \frac{y(t)}{t}\right), \quad 0 < t \leq 1, \\
(1.2b) \quad B\left(y(1), y'(1)\right) &= 0, \quad y(0) = 0.
\end{align}

Here $y$ and $f$ are vector-valued functions of dimension $n$, $A_0$ and $A_1$ are constant $n \times n$ matrices and $B$ in (1.1) is an $m$-dimensional vector-valued function, with $m \leq 2n$, while $B$ in (1.2) is $n$-dimensional. We also assume $y$ to be real-valued and continuously differentiable on $[0,1]$ and its second derivative to be continuous on $(0,1]$. The problems (1.1) and (1.2) occur often in applications from mechanics and chemistry, see for example Keller and Wolfe [5], Parter, Stein and Stein [10] and Rentrop [12], typically when transforming partial differential equations to ordinary differential equations.

The numerical solution of scalar equations of this type has been investigated by Jamet [3], Natterer [9], Russell and Shampine [13]. Brabston and Keller [1] and de Hoog and Weiss [2] have considered first-order systems with a singularity of the first kind.
The basic analytical properties of nonlinear, second-order systems have been studied by the author in [14]. Questions of existence, smoothness and uniqueness of the analytical solutions have been discussed. In [15] the finite difference method has been applied to linear systems with variable coefficient matrices $A_0(t)$ and $A_1(t)$, the stability of the scheme has been shown, and the convergence order has been derived. The analysis in [14] and [15] is based on the transformation of the system of second order to a system of first order, cf. (3.2a), where the $2n \times 2n$ matrix $M(t)$ occurs. It has been shown in [15] that if the analytical problem is well-posed, then for every equidistant gridspacing $h$ on $[0,1]$ there exists a $t_0(h) > 0$ such that the difference system associated with the linear boundary value problem has a unique solution on $[t_0(h),1]$. Furthermore, the order of convergence depends on the eigenvalues of $M \equiv M(0)$ and the smoothness properties of $y$. It has been shown for $f \in C^2$ that the order of convergence is $h^{q} \ln h |p|$, $p > 0$, where $q = 2$ if all eigenvalues of $M$ have nonpositive real parts, and $q = \min(\sigma_+, 2)$, where $\sigma_+$ is the smallest positive real part, otherwise. In a fairly large number of computer examples (also nonlinear) it has been observed that the choice $t_0(h) = 0$ worked well, and this suggests that the restriction $t_0(h) > 0$ is rather technical (contraction arguments were used in [14]) and can be omitted. This will be shown here.

The outline of the paper follows. In Section 2 we collect the notations and preliminary results used in the subsequent analysis. In Section 3 we present the basic analytical properties of the solution of the linear problem, which we require for the discussion of the numerical method. Furthermore, we formulate assumptions for the treatment of nonlinear problems. Because in many practical nonlinear problems the solution is not unique, cf. [5], we only assume that a solution exists and is isolated. In Section 4, we formulate the numerical schemes to be considered and show existence and uniqueness of the solution of the difference systems for the linear case on the whole interval $[0,1]$. To avoid repetitions in the analysis, we state the convergence results without proofs, which can be found in [15]. The results are extended to the nonlinear case in Section 5. If the problems (1.1) and (1.2) have an isolated solution, we show that the associated discrete problems have also a unique solution in a neighborhood of the continuous solution, when the gridspacing $h$ is sufficiently small. In the analysis of the nonlinear problems we use the results of Keller [6]. The order of convergence depends on the smoothness of the solution $y$ as well as on the spectral properties of $A_0$ and $A_1$. Typically, one has convergence order $h^{2} \ln h |p|$, $p > 0$. Finally, Section 6 contains numerical examples illustrating the theory, and Section 7 (appendix) technical details.

2. Notations and Preliminary Results. We denote by $X^n$ the space of complex-valued vectors of dimension $n$ and use $| \cdot |$ to denote the maximum norm on $X^n$,

$$|x| := \max_{1 \leq i \leq n} |x_i|.$$ 

Let $\Delta$ be an equidistant partition of the interval $[0,1]$

$$\Delta = \{ t_i, i = 0(1) N | t_i = ih, t_N = 1 \},$$

where $i = 0,1,\ldots,N$ is denoted by $i = 0(1)N$. With each partition $\Delta$ we associate the linear space $X_\Delta$ with elements

$$x_\Delta = (x_0, x_1, \ldots, x_N).$$
where \( x_m = (x_{m1}, x_{m2}, \ldots, x_{mn})^T \in X^n, m = 0(1)N, \) and the norm on \( X_\Delta \) is
\[
\|x_\Delta\| := \max_{0 \leq m \leq N} |x_m|.
\]
\( C^p[0, 1] \) is the space of vector-valued functions as well as the space of complex-valued matrices, whose elements are \( p \) times continuously differentiable on \([0, 1], \) and \( C^p(0, 1] \) is defined analogously. We use \( C = C[0, 1] = C^0[0, 1] \) and \( C(0, 1] = C^0(0, 1]. \) For each vector \( y \in C \) we define the norm
\[
\|y\| := \max_{0 \leq t \leq 1} |y(t)|,
\]
and for each matrix \( A \in C, \|A\| \) is the induced norm. For any continuous function \( y \) the modulus of continuity is defined by
\[
\omega(y; \delta) := \max_{0 \leq t \leq 1-\delta} |y(t + \delta) - y(t)|.
\]
Finally, \( R_\Delta: C \to X_\Delta \) is the bounded linear map such that
\[
R_\Delta y = (y(t_0), y(t_1), \ldots, y(t_N)), \quad R_\Delta y' = (y'(t_0), y'(t_1), \ldots, y'(t_N)).
\]
We now formulate the results which are basic in our analysis.

**Lemma 2.1.** Given a complex number \( \lambda = \sigma + i\kappa, \sigma > 0 \) and \( \Omega(\lambda) = \{ \mu | |\lambda - \mu| \leq \sigma/2\}; \) for \( \mu \in \Omega(\lambda) \) define

\[
(2.1) \quad z_{kj}(\mu) := \begin{cases} 1, & k = j, \\ \frac{1}{l-k} \prod_{l=k}^{j-1} (1 - \frac{\mu}{l}), & 1 \leq k < j, j = 2(1)N + 1. \end{cases}
\]

Then there exists an \( \eta > 0 \) such that
\[
|z_{kj}(\mu)| \leq \text{const} (\frac{\mu}{l})^\eta, \quad k \leq j, j = 1(1)N,
\]
for all \( \mu \in \Omega(\lambda). \)

**Proof.** Let \( \mu \in \Omega(\lambda) \) and \( l \geq l_0 = \frac{3}{2} \sigma. \) Then \( |l - \mu| \leq |\lambda - \nu|, \) where \( \nu = \sigma/2 - i(\kappa + \sigma/2). \) Let \( k \geq l_0. \) Then \( |z_{kj}(\mu)| \leq |v_{kj}|, \) where
\[
v_{kj} = \prod_{l=k}^{j-1} (l - \nu)/l.
\]
Using
\[
\Gamma(z + n) = \left( \prod_{k=0}^{n-1} (z + k) \right) \Gamma(z)
\]
and the asymptotic expansion
\[
\Gamma(s + a)/\Gamma(s + b) = s^{a-b}(1 + O(1/s)), \quad \text{Re}(s) > 0,
\]
cf. [7], we can rewrite \( v_{kj} \) and obtain
\[
v_{kj} = \prod_{l=k}^{j-1} (l - \nu)/l = \Gamma(j - \nu)\Gamma(k)/\Gamma(k - \nu)\Gamma(j) = (k/j)^\nu(1 + O(1/k)).
\]
which yields the desired result for \( k \geq l_0. \) The result for \( k < l_0 \) now follows, since \( z_{kj}(\mu) = z_{k,l_0}(\mu)z_{l_0,j}(\mu) \) and \( z_{k,l_0} \) consists of at most \( l_0 \) terms. \( \square \)
Lemma 2.2. Given a complex number $\lambda = \sigma + i\kappa$, $\sigma > 0$, $\Omega_1(\lambda) = \{\mu|\lambda - \mu| \leq \sigma/4\}$ and $\Omega_2(\lambda) = \{\mu|\lambda - \mu| \leq \sigma/2\}$; for any $\mu \in \Omega_1(\lambda)$ and complex number $\delta$ define

$$z_{kj}(\mu, \delta) := \begin{cases} 1, & k = j, \\ \prod_{l=k}^{j-1} \left(1 - \frac{\mu + \delta}{l}\right), & 1 \leq k < j, j = 2(1)N + 1. \end{cases}$$

Then there exists an $\eta > 0$ such that

$$|z_{kj}(\mu, \delta)| \leq \text{const}(t_k/t_j)^{\eta}, \quad k \leq j, j = 1(1)N,$$

for all $\mu \in \Omega_1(\lambda)$.

Proof. Let $\mu_j = \mu + \delta/j$. Then we can rewrite (2.2) and obtain

$$z_{kj}(\mu, \delta) = \prod_{l=k}^{j-1} \left(1 - \frac{\mu}{l}\right).$$

If $\mu \in \Omega_1(\lambda)$ and $l \geq l_1 = 4|\delta|/\sigma$, then $\mu_j \in \Omega_2(\lambda)$. Let $l \geq \max\{l_0, l_1\}$. Then $|l - \mu_j| \leq |l - \mu|$, and hence the assertion follows as in Lemma 2.1. \(\Box\)

Lemma 2.3. Let $\lambda, \Omega_1(\lambda)$ and $\Omega_2(\lambda)$ be as in Lemma 2.2. Let $\delta(l, \mu)$ be a bounded sequence of complex numbers such that $|\delta(l, \mu)| \leq \delta$ for each $l \geq \rho \geq 1$ and all $\mu \in \Omega_1(\lambda)$. For $\mu \in \Omega_1(\lambda)$ and $\delta(l, \mu)$ define

$$z_{kj}(\mu, \delta(\cdot)) := \begin{cases} 1, & k = j, \\ \prod_{l=k}^{j-1} \left(1 - \frac{\mu + \delta(l, \mu)}{l}\right), & 1 \leq k < j, j = 2(1)N + 1. \end{cases}$$

Then there exists an $\eta > 0$ such that

$$|z_{kj}(\mu, \delta(\cdot))| \leq \text{const}(t_k/t_j)^{\eta}, \quad k \leq j, j = 1(1)N,$$

for all $\mu \in \Omega_1(\lambda)$.

Proof. On setting $\mu_{l,8} = \mu + \delta(l, \mu)/l$, we have

$$z_{kj}(\mu, \delta(\cdot)) = \prod_{l=k}^{j-1} \left(1 - \frac{\mu_{l,8}}{l}\right).$$

Let $\mu \in \Omega_1(\lambda)$. Then $\mu_{l,8} \in \Omega_2(\lambda)$ for each $l \geq l_1 = \max(\rho, 4\delta/\sigma)$, and the same arguments as in the proof of Lemma 2.2 apply. \(\Box\)

Lemma 2.4. For every $k > j \geq 1$ and $\gamma \in \mathbb{R}$,

$$\sum_{l=j}^{k-1} h t_{l, \gamma}^{-1} \leq \begin{cases} \text{const}\left|t_k^\gamma - t_j^\gamma\right|, & \gamma \neq 0, \\ \text{const}\ln(t_k/t_j), & \gamma = 0. \end{cases}$$

Proof. See [15, Lemma 2.2]. \(\Box\)

Finally, for the matrix $A$ and an analytic function $\xi(\lambda)$, we write the matrix function $\xi(A)$ as

$$\xi(A) = \frac{1}{2\pi i} \oint_{\Gamma} \xi(\lambda)(\lambda I - A)^{-1} d\lambda,$$

where $\Gamma$ is a closed curve containing all eigenvalues of $A$. 
3. Analytical Problems. In this section we recall the formulation of the linear problem given in [14], some analytical properties of the solutions of this problem and the assumptions made for the nonlinear case. For technical details, proofs, etc., see [14] and [15].

Consider the linear boundary value problem

\begin{align}
(3.1a) & \quad y''(t) - \frac{A_1(t)}{t} y'(t) - \frac{A_0(t)}{t^2} y(t) = f(t), \quad 0 < t \leq 1, \\
(3.1b) & \quad B_0 Y(0) + B_1 Y(1) = \beta, \quad Y(t) = (y(t), y'(t))^T,
\end{align}

where \( y, f \) are \( n \)-vectors, \( A_0, A_1 \) are \( n \times n \) matrices, \( B_0, B_1 \) are \( m \times 2n \) constant matrices and \( \beta \) is an \( m \)-vector, with \( m \leq 2n \). By the linear transformation \( z(t) = (z_1(t), z_2(t))^T = (y(t), ty'(t))^T \), the system (3.1a) can be reduced to a first-order system of the form

\begin{align}
(3.2a) & \quad z'(t) = \frac{1}{t} \begin{bmatrix} 0 & I \\ A_0(t) & I + A_1(t) \end{bmatrix} z(t) + t \begin{bmatrix} 0 \\ f(t) \end{bmatrix} = \frac{1}{t} M(t) z(t) + t f(t), \\
(3.2b) & \quad B_0 Y(0) + B_1 Y(1) = \beta.
\end{align}

Let \( f \in C \) and assume that \( A_0(t) \) and \( A_1(t) \) are

\begin{align}
(3.3) & \quad A_0(t) = A_0 + t^r C_0(t), \quad A_1(t) = A_1 + t^r C_1(t), \quad r \geq 0,
\end{align}

where \( A_0, A_1 \) are constant matrices and \( C_0, C_1 \in C \). Then we can rewrite (3.2) and obtain

\begin{align}
(3.4a) & \quad z'(t) = \frac{1}{t} \begin{bmatrix} 0 & I \\ A_0 & I + A_1 \end{bmatrix} z(t) + t^{r-1} \begin{bmatrix} 0 \\ C_0(t) \\ C_1(t) \end{bmatrix} z(t) + t \begin{bmatrix} 0 \\ f(t) \end{bmatrix} = \frac{1}{t} M z(t) + t^{r-1} \hat{C}(t) z(t) + t f(t), \quad 0 < t \leq 1, \\
(3.4b) & \quad B_0 Y(0) + B_1 Y(1) = \beta.
\end{align}

In order to formulate the existence result for the solution of (3.4a), we have to introduce the following notations. We denote by \( R \) the spectral projection onto the eigenspace of \( M = M(0) \) corresponding to the eigenvalue \( \lambda = 0 \), and by \( S \) the spectral projection onto the invariant subspace of \( M \) corresponding to the eigenvalues with positive real parts. We also set

\begin{align}
(3.5) & \quad P = R + S, \quad Q = I - P,
\end{align}

and have the following result.

**Theorem 3.1.** Let \( Q z(0) = 0 \); then for every \( f \in C \) and constant \( 2n \)-vector \( \gamma \) there exists a unique, continuous solution \( z(t) \) of (3.4a) satisfying \( P z(1) = P \gamma \). Since \( y(t) = z_1(t) \), we obtain a solution \( y(t) \) of (3.1a), and \( y \in C \cap C^2(0,1) \).

The question as to whether this solution satisfies the boundary value problem (3.1) can be answered if we substitute the solution \( y \) and its derivative \( y' \) into (3.1b). On noting that \( P z(1) = P Y(1) = P \gamma \) we see that we need \( m = \text{rank}[P] \) conditions to make the solution \( y \) unique and these conditions have to be given by (3.1b). It has been shown in [14, Theorem 4.2] that the \( m \) constants can be uniquely determined from (3.1b) if and only if the inverse of a certain \( m \times m \) matrix obtained by the above substitution exists.

For the nonlinear problem (1.1) we make the following assumptions.
N.1.1. Problem (1.1) has a solution \( y \in C^1 \). We define the sphere \( S_\rho \) associated with the solution \( y \) of (1.1) by
\[
S_\rho(y(t)) = \{ v \in X^n \mid |v - y(t)| \leq \rho, \rho > 0 \}
\]
and the sphere \( S_\delta \) associated with its derivative \( y' \) by
\[
S_\delta(y'(t)) = \{ w \in X^n \mid |w - y'(t)| \leq \delta, \delta > 0 \}.
\]
Additionally, we set
\[
T_{\rho,\delta} = \left\{ (t, v, w) \mid 0 \leq t \leq 1, v \in S_\rho(y(t)), w \in S_\delta(y'(t)) \right\}.
\]

N.1.2. \( B : D_B \to X^n \) is a nonlinear map and \( f : D_f \to X^n \) is a nonlinear map, which is continuous on \([0, 1] \times X^n \times X^n \), \( D_B \) and \( D_f \) are open sets and \( m = \text{rank}[P] \).

N.1.3. \( f \) is continuously differentiable with respect to \( v \) and \( w \), \( f_v(t, v, w) \), \( f_w(t, v, w) \) are continuous on \( T_{\rho,\delta} \); \( B \) is continuously differentiable with respect to all variables on
\[
S_\rho(y(0)) \times S_\rho(y(1)) \times S_\delta(y'(1)).
\]

N.1.4. If \( M \) has eigenvalues \( \lambda \) with positive real parts \( \sigma \), then \( \sigma > 1 \).

N.1.5. The solution \( y \) of (1.1) is isolated. This is equivalent to the condition that the following (linear) problem
\[
(3.5a) \quad L(y)v(t) = v''(t) - \frac{A_1}{t} v'(t) - C_1(t) v'(t) - \frac{A_0}{t^2} v(t) - C_0(t) v(t) = 0,
\]
(3.5b) \[
B_{00}v(0) + B_{10}v(1) + B_{11}v'(1) = 0,
\]
where \( f \equiv f(\cdot, v(\cdot), w(\cdot)) \), \( B \equiv B(u_1; u_2, u_3) \) and
\[
B_{00} = \frac{\partial B}{\partial u_1}(y(0); y(1), y'(1)),
\]
(3.6a) \[
B_{10} = \frac{\partial B}{\partial u_2}(y(0); y(1), y'(1)), \quad B_{11} = \frac{\partial B}{\partial u_3}(y(0); y(1), y'(1)),
\]
has only the trivial solution \( v(t) = 0 \).

Let \( \hat{P} \) be the \( 2n \times m \) matrix consisting of the linearly independent columns of \( P \) and \( \hat{P} \hat{P} = P \). Let \( I_1 \) be the \( n \times 2n \) matrix consisting of the first rows of the identity matrix \( I \) and \( I_2 \) the \( n \times 2n \) matrix consisting of the last rows of \( I \). Then any continuous solution of (1.1) satisfies
\[
(3.7a) \quad y(t) = I_1 \{(Hf(\cdot, y, y'))(t) + \phi(t) \hat{P} \alpha \},
\]
(3.7b) \[
y'(t) = I_2 \{(Hf(\cdot, y, y'))(t) + \phi(t) \hat{P} \alpha \}/t,
\]
(3.7c) \[
\alpha = \alpha - B(y(0); y(1), y'(1)),
\]
where \( \alpha = \hat{P} Y(1), \alpha \in X^m \) and
\[
(3.8a) \quad (Hf(\cdot, y, y'))(t) = t^2 \int_0^1 Qs^{-M}s^f(ts, y(ts), y'(ts)) \, ds + t^M \int_1^t P s^{-M}s^f(s, y(s), y'(s)) \, ds,
\]
(3.8b) \[
\phi(t) = t^M \hat{P}.
\]
We can write (3.7) as

\[ x = N(x), \]

where \( x = (y, y', a) \) and \( N: U_{p, \delta} \times X^m \to C^m \) is a compact nonlinear operator, \( U_{p, \delta} = \{ u \in C^1 | u(t) \in S_p(y(t)), \ u'(t) \in S_p(y'(t)), \ 0 \leq t \leq 1 \} \) and \( C^m = C^1 \times X^m \). A simple modification of contraction arguments given in [14, Section 5] yields the stability of the solution of (1.1), i.e., the continuous dependence of the solution \( y \) on small perturbations in the right-hand side of the differential equations and boundary conditions.

The extension of this result to the problem (1.2) is rather straightforward, if we change assumptions N.1.1–N.1.5 properly.

N.2.1. Problem (1.2) is well-posed and has a solution \( y \in C^1, y(t)/t \in C \). We define the following sphere for the solution \( y \)

\[ S_p(y(t)/t) = \left\{ v \in X^n \left| \frac{v}{t} - \frac{y(t)}{t} \right| \leq \varepsilon, \varepsilon > 0 \right\} \]

and note that if \( v \in S_p(y(t)/t) \) then \( v \in S_p(y(t)) \). We also set

\[ T_{p,t} = \left\{ (t,v,w) \mid 0 \leq t \leq 1, \ v \in S_p(y(t)), \ w \in S_p(y(t)/t) \right\}. \]

N.2.2. \( B: D_B \to X^n \) is a nonlinear map and \( D_B \) is an open set.

N.2.3. \( B \) is continuously differentiable on \( S_p(y(1)) \times S_p(y'(1)) \).

N.2.4. If the smallest real part of the eigenvalues of \( M \) is 1, then \( \lambda = 1 \) and the Jordan box associated with \( \lambda = 1 \) is diagonal.

N.2.5. The solution \( y \) of (1.2) is isolated, i.e., the problem

\[ \begin{align*}
&v''(t) - \frac{A_1}{t} v'(t) - \frac{A_0}{t^2} v(t) - C_0(t) v(t) = 0, \\
&B_{10} v(1) + B_{11} v'(1) = 0,
\end{align*} \]

where

\[ C_0(t) = \frac{\partial f}{\partial v} \left( t, y(t), \frac{y(t)}{t} \right) + \frac{\partial f}{\partial w} \left( t, y(t), \frac{y(t)}{t} \right)/t, \]

has only the trivial solution.

All other assumptions remain valid with respect to the above changes and new definitions.

4. **Numerical schemes.** We consider the partition \( \Delta \) of Section 2 and the nonlinear problem (1.1). Then the three-point discretization yields

\[ \begin{align*}
&\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} - \frac{A_1}{t_i} \left( \frac{y_{i+1} - y_i}{2h} \right) - \frac{A_0}{t_i^2} y_i \\
&= f \left( t_i, y_i, \frac{y_{i+1} - y_i}{2h} \right), \quad i = 1(1) N,
\end{align*} \]

\[ B \left( y_0; y_N, \frac{y_{N+1} - y_N}{2h} \right) = 0. \]

Without loss of generality we assume that the boundary conditions, which are to be posed at \( t = 0 \) for the continuity of the solution, i.e., \( Qz(0) = 0 \), are given by

\[ \begin{align*}
&\tilde{Q} Y_0 = \tilde{Q} \left[ \left( y_1 - y_0 \right)/h - hA^{-1} f(0, y_0, 0)/2 \right] = 0,
\end{align*} \]
where \( \hat{Q} \) is a constant \( q \times 2n \) matrix and \( q = \text{rank}[Q] \). To see that the lower expression in (4.1c) is an approximation for \( y'(0) \) we assume \( y \in C^2 \) and apply Taylor’s theorem to (1.1a). Then we have

\[
Ay''(0) = (I - A_1 - A_0/2)y''(0) = f(0, y(0), y'(0))
\]

if and only if \( (A_1 + A_0)y'(0) = 0 \) and \( A_0y(0) = 0 \). By N.1.4 and N.1.2 we have immediately \( y'(0) = 0 \) (cf. [14, Lemmas 3.1, 3.2 and 3.3]), and provided that \( A^{-1} \) exists, the result follows from

\[
y''(0) := \left( y_{-1} - 2y_0 + y_1 \right)/h^2, \quad y'(0) := \left( y_1 - y_{-1} \right)/2h.
\]

The difference scheme associated with (1.2) is

\[
\begin{align*}
y_{i+1} - \frac{2y_i + y_{i-1}}{h^2} - \frac{A_1}{i} \left( \frac{y_{i+1} - y_{i-1}}{2h} \right) &= \frac{A_0}{i^2} y_i, \\
&= f\left(t_i, y_i, \frac{y_i}{t_i}\right), \quad i = 1(1)N, \\
B \left( y_N, \frac{y_{N+1} - y_{N-1}}{2h} \right) &= 0,
\end{align*}
\]

\[
y_0 = 0.
\]

In the following subsection we consider the linear case and show how the stability results from [15] can be extended to the whole interval \( 0 \leq t \leq 1 \). The convergence results are repeated, because their proofs do not change.

**Numerical Results for Linear Problems.** We first formulate the difference scheme for the case when the coefficient matrices \( A_0(t) \) and \( A_1(t) \) are constant, i.e., \( C_0(t) = C_1(t) = 0 \), and then transform the second-order system to the first-order one. For (3.1) we have

\[
\begin{align*}
(I - A_1 + A_0)
\left( y_{i+1} - \left( 2I + \frac{A_0}{i^2} \right) y_i + \left( I + \frac{A_1}{2i} \right) y_{i-1} = h^2 f(t_i), \quad i = 1(1)N, \\
B_0 \left[ \left( y_1 - y_0 \right)/h - hA^{-1} f(0)/2 \right] + B_1 \left[ \left( y_{N+1} - y_{N-1} \right)/2h \right] = \beta.
\end{align*}
\]

We define

\[
u_{1,i} = y_i, \quad u_{2,i} = (i + 1)(y_{i+1} - y_i), \quad i = 0(1)N,
\]

and obtain immediately

\[
u_{1,i} = u_{1,i-1} + \frac{1}{i} u_{2,i-1}, \quad i = 1(1)N.
\]

The application of (4.4) to (4.3a) yields

\[
u_{2,i} = \left( \frac{A_0}{i} + \frac{1}{i^2} \Theta_1(i) \right) u_{1,i-1} + \left( I + \frac{1}{i} (I + A_1) + \frac{1}{i^2} \Theta_2(i) \right) u_{2,i-1} + \Theta_3(i) t_{i+1} hf(t_i), \quad i = 1(1)N,
\]
where
\[ \Theta_1(i) = \left( I - \frac{A_1}{2i} \right)^{-1} \left( I + \frac{A_1}{2i} \right) A_0, \]
\[ \Theta_2(i) = \left( I - \frac{A_1}{2i} \right)^{-1} \left[ \left( 1 + \frac{1}{i} \right) A_0 + \left( I + \frac{A_1}{2i} \right) A_1 \right], \]
\[ \Theta_3(i) = \left( I - \frac{A_1}{2i} \right)^{-1}. \]

Clearly, we have to assume that for \( 1 < i < |A_1|/2 \), \( (I - A_1/2i)^{-1} \) exist. From (4.5) and (4.6) we have for \( u_i = (u_{1,i}, u_{2,i})^T \)
\[ u_i = u_{i-1} + \frac{1}{i} Mu_{i-1} + \frac{1}{i^2} \Theta_4(i) u_{i-1} + t_{i+1} h \Theta_5(i) \hat{f}(t_i), \quad i = 1(1)N, \]
where \( \Theta_4 \) and \( \Theta_5 \) are appropriate \( 2n \times 2n \) matrices. Since \( J = E^{-1}ME \), \( v_i = E^{-1}u_i \) solves
\[ v_i = v_{i-1} + \frac{1}{i} Jv_{i-1} + \frac{1}{i^2} \Theta_4(i) v_{i-1} + t_{i+1} h \psi(i) g_i, \]
where
\[ \Theta_4(i) = E^{-1}Q_4(i)E, \quad \psi(i) = E^{-1}Q_5(i)E, \quad \phi_i = E^{-1}\hat{f}(t_i), \quad i = 1(1)N. \]

For (4.8) we write
\[ v_i - v_{i-1} - \frac{1}{i} Jv_{i-1} - \frac{1}{i^2} \Theta_4(i) v_{i-1} = t_{i+1} h \psi(i) g_i, \quad i = 1(1)N, \]
and this is formally
\[ GV = R, \]
where \( G: X_\Delta \rightarrow X_\Delta \). In [15] we proceeded similarly, but there the expression on the left-hand side of (4.9) was \( v_i - v_{i-1} - (1/i)Jv_{i-1} \). In order to show that \( G \) has an inverse bounded independently of \( h \), we assume that \( J \) consists of one Jordan box
\[ J = \begin{bmatrix} \lambda & 1 \\ & \ddots & 1 \\ & & \lambda \end{bmatrix}, \quad \lambda = \sigma + i\kappa, \]
and consider the following three cases \( \sigma < 0, \lambda = 0 \) and \( \sigma > 0 \) separately. Let us write (4.8) as
\[ v_i = v_{i-1} + \frac{1}{i} Jv_{i-1} + \frac{1}{i^2} \Theta_4(i) v_{i-1} + r_i, \quad i = 1(1)N, \]
and consider

**Case 1.** \( \sigma < 0 \). Let \( v_0 = \gamma \in R^{2n} \). Then
\[ v_i = \prod_{l=0}^{i-1} \left( I + \frac{1}{l} J + \frac{1}{l^2} \Theta(l) \right) \gamma + \sum_{k=1}^{i-1} \left( \prod_{l=k}^{i-1} \left( I + \frac{1}{l} J + \frac{1}{l^2} \Theta(l) \right) \right) r_k + r_i, \]
\[ = \tilde{Z}_{i,1} \gamma + \sum_{k=1}^{i-1} \tilde{Z}_{i,k+1} r_k + r_i, \quad i = 1(1)N. \]
Let
\[ Z_{j,k+1} = \prod_{i=j}^{k+1} \left( I + \frac{1}{l} J + \frac{1}{l^2} \Delta(l, J) \right) \]
\[ = \frac{1}{2\pi i} \int_{\Gamma} z_{k+1,j+1}(-\lambda, -\delta(\cdot))(\lambda I - J)^{-1} d\lambda, \]
where \( \Gamma = \{ \mu | |\lambda - \mu| = \sigma/4 \} \), \( \Delta(l, J) \) is a diagonal matrix, \( \Delta(l, J) = \text{diag}(\delta(l, \lambda)) \) and \( \delta(l, \lambda) \) is a complex number such that
\[ \text{Re}(\delta(l, \lambda)) = \text{sign}(\text{Re}(\lambda))|\lambda|, \quad \text{Im}(\delta(l, \lambda)) = \text{sign}(\text{Im}(\lambda))|\lambda| \]
for each \( 1 \leq l \leq N \). We notice that if \( |A_1| = a_1 \) and \( |A_0| = a_0 \), then for each \( l \geq \rho > a_1/2 \)
\[ |\Theta(l)| \leq |\Theta(\rho)| \leq \text{const} \left[ (1 + a_1/2)(a_0 + a_1) + 2a_0 / (1 - a_1/2\rho) \right] = \Theta \]
and \( |\Theta(l)| \leq \text{const} \) otherwise. Using results shown in the appendix and Lemma 2.3, we now have
\[ |\tilde{Z}_{j,k+1}| \leq |Z_{j,k+1}| \leq \text{const}(t_{k+1}/t_{j+1})^{\eta}, \quad k < j, \]
and
\[ |V_i| \leq \text{const} \left( |\gamma| + \sum_{k=1}^{j} (t_{k+1}/t_{i+1})^{\eta}|r_k| \right), \quad i = 1(1)N. \]

Using Lemma 2.4, we obtain for the solution of (4.8)
\[ (4.13) \quad |V_i| \leq \text{const} \left[ |\gamma| + t_i^2\|f_\delta\| \right], \quad i = 1(1)N. \]

Case 2. \( \lambda = 0 \). Before studying (4.8), we investigate the growth of solutions of the following system
\[ (4.14) \quad w_i = \left( I + \frac{1}{i} J + \frac{1}{i^2} \tilde{\Theta}(i) \right)w_{i-1} + r_i, \quad i = 1(1)N, \]
where \( \tilde{\Theta}(i) \) is a diagonal matrix and each diagonal element is equal to \( |\Theta(i)| = n(i) \).
Let \( w_{0r} = \delta_{0r}, \quad r = 2(1)2n \) and \( w_{N1} = \delta_{N1} \), and let us consider the system (4.14) componentwise. For the \( k \)th component we have
\[ w_{ik} = w_{i-1,k} + \frac{1}{i} w_{i-1,k+1} + \frac{1}{i^2} n(i)w_{i-1,k} + r_{ik}, \quad k = 1(1)2n - 1, \]
\[ w_{i,2n} = w_{i-1,2n} + \frac{1}{i^2} n(i)w_{i-1,2n} + r_{i,2n}. \]
From the last equation we obtain
\[ w_{i,2n} = \prod_{m=1}^{i} \left( 1 + \frac{1}{m^2} n(m) \right) \delta_{0,2n} + \sum_{j=1}^{i-1} \left( \prod_{l=j+1}^{i} \left( 1 + \frac{1}{l^2} n(l) \right) \right) r_{j,2n} + r_{i,2n}. \]
and it follows immediately from Lemma 2.4 that
\[ (4.15) \quad |w_{i,2n}| \leq \text{const} \left[ |\delta_{0,2n}| + t_i^2\|f_\delta\| \right], \quad i = 1(1)N. \]
For \( k = 2n - 1 \) we have

\[
\begin{aligned}
w_{i,2n-1} &= \prod_{m=1}^{i} \left( 1 + \frac{1}{m^2} n(m) \right) \delta_{0,2n-1} \\
&\quad + \sum_{j=1}^{i-1} \left( \prod_{l=j+1}^{i} \left( 1 + \frac{1}{l^2} n(l) \right) \right) \left( \frac{1}{j} w_{j-1,2n} + r_{j,2n-1} \right) + \frac{1}{i} w_{i-1,2n} + r_{i,2n-1}
\end{aligned}
\]

and the following estimate can be obtained using (4.15) and Lemma 2.4,

\[
|w_{i,2n-1}| \leq \text{const}\left( |\delta_{0,2n-1}| + |\delta_{0,2n}| |\ln h| + t_i^2 \|f_\delta\| \right).
\]

Clearly,

\[
(4.16) \quad |w_{i,r}| \leq \text{const}\left( \sum_{k=r}^{2n} |\delta_{0,k}| |\ln h|^{k-r} + t_i^2 \|f_\delta\| \right), \quad i = 1(1)N, \quad r = 2(1)2n.
\]

Finally, the first component is

\[
w_{11} = \prod_{m=i+1}^{N} \left( 1 + \frac{1}{m^2} n(m) \right) \delta_{N1} \\
- \sum_{j=i+1}^{N} \left( \prod_{l=j+1}^{i} \left( 1 + \frac{1}{l^2} n(l) \right) \right) \left( \frac{1}{j} w_{j-1,2} + r_{j1} \right), \quad i = 1(1)N,
\]

and the estimate for \( w_{11} \) follows in a very similar way,

\[
(4.17) \quad |w_{i1}| \leq \text{const}\left( |\delta_{N1}| + \sum_{k=2}^{2n} |\delta_{0,k}| |\ln h|^{k-1} + \|f_\delta\| \right), \quad i = 1(1)N.
\]

Let \( v_\lambda \) be a solution of (4.8) and \( v_{N1} = \delta_1, v_{0r} = \delta_r, r = 2(1)2n \). Since

\[
\left| I + \frac{1}{i} J + \frac{1}{i^2} \Theta(i) \right| \leq \left| I + \frac{1}{i} J + \frac{1}{i^2} \Theta(i) \right|, \quad i = 1(1)N,
\]

it follows from (4.16) and (4.17) that

\[
(4.18) \quad |v_i| \leq \text{const}\left( |\delta_1| + \max_{2 \leq r \leq 2n} |\delta_r| |\ln h|^{2n-1} + \|f_\delta\| \right).
\]

**Remark 4.1.** Consider the case when the Jordan box associated with \( \lambda = 0 \) is diagonal; this is equivalent to \( J = 0 \). In addition, let \( v_0 = \delta_0 \) be the initial condition; then it follows immediately from (4.15) that in this case

\[
(4.19) \quad |v_i| \leq \text{const}\left( |\delta_0| + t_i^2 \|f_\delta\| \right), \quad i = 1(1)N.
\]

**Remark 4.2.** It follows also from [14, Lemma 3.2] that the case when the initial condition is of the form \( v_0 = \delta_0 \) is of interest. Analytically, this condition is equivalent to \( Qz(0) = 0 \) and \( Rz(0) = R_\gamma \). From (4.16) it is easy to see that the estimate for \( v_i \) is now

\[
(4.20) \quad |v_i| \leq \text{const}\left( |\delta_{01}| + \max_{2 \leq r \leq 2n} |\delta_{0r}| |\ln h|^{2n-1} + t_i^2 \|f_\delta\| \right).
\]
Case 3. \( \sigma > 0 \). Let \( v_N = \gamma \in \mathbb{R}^{2n} \). Then,

\[
 u_j = \prod_{k=j+1}^{N} \left( I + \frac{1}{k} J + \frac{1}{k^2} \Theta(k) \right)^{-1} \gamma \tag{4.21}
\]

\[
 = \sum_{k=j+1}^{N} \left( \prod_{l=j+1}^{k} \left( I + \frac{1}{l} J + \frac{1}{l^2} \Theta(l) \right)^{-1} \right) r_k 
\]

\[
 = Z_{i+1,N} \gamma - \sum_{k=i+1}^{N} Z_{i+1,k} r_k, \quad i = 0(1)N - 1.
\]

Before considering Case 3 in detail, let us briefly discuss the question of the existence of the matrix

\[
 M_{i+1,k} = \prod_{l=i+1}^{k} \left( I + \frac{1}{l} M + \frac{1}{l^2} \Theta_4(l) \right)^{-1}, \quad k \geq i + 1, i = 0(1)N - 1. \tag{4.22}
\]

Clearly, there exists an index \( i_0 \) such that for each \( i \geq i_0 \)

\[
 \left| \frac{1}{i^2} \Theta_4(i) \right| < \left| \left( I + \frac{1}{i} M \right)^{-1} \right| \leq \text{const},
\]

and hence \( M_{i+1,k} \) exists for each \( i \geq i_0 \). It is not so simple to solve this problem for \( 1 \leq i < i_0 \) in the most general case, without any additional assumptions on the matrices \( A_0 \) and \( A_1 \). Therefore, we assume that \( M_{i+1,k} \) exists for each \( i = 0(1)N - 1 \) and show that this assumption holds in some special cases that are important for applications, cf. [5], [10], [12].

Scalar case. The eigenvalues \( \lambda_1, \lambda_2 \) of the matrix \( M \) are solutions of the following equation,

\[
 \lambda^2 - \lambda (1 + a_1) - a_0 = 0; \quad a_0, a_1 \in \mathbb{R},
\]

and hence if \( \text{Re}(\lambda_1) > 0, i = 1, 2 \), then it follows immediately that

\[
 a_1 > -1. \tag{4.23}
\]

Assume that an index \( i \geq 1 \) exists such that the matrix

\[
 I + \frac{1}{i} M + \frac{1}{i^2} \Theta_4(i) \tag{4.24}
\]

is singular, or equivalently,

\[
 1 + \frac{1}{i} (1 + a_1) + \frac{1}{i^2} \Theta_2(i) = \frac{1}{i^2} a_0 + \frac{1}{i^3} \Theta_1(i).
\]

Then we have from the last equation

\[
 2i^2 + (2 + a_1)i + a_1 = 0
\]

and \( a_1 = -2i \), which contradicts (4.23).

\( A_0, A_1 \) are real diagonal matrices. As in the scalar case, if \( \lambda_i, i = 1(1)n \), are the eigenvalues of the matrix \( M \) and \( \text{Re}(\lambda_i) > 0 \), then

\[
 a_{1,k} > -1, \quad k = 1(1)n,
\]

where \( a_{1,k} \) denotes the \( k \)th element of the matrix \( A_1 \). If the matrix (4.24) is singular, then

\[
 \det \left[ -\left( \frac{1}{i^2} A_0 + \frac{1}{i^3} \Theta_1(i) \right) + I + \frac{1}{i} (I + A_1) + \frac{1}{i^2} \Theta_2(i) \right] = 0,
\]
or equivalently,
\[
\det\left[2i^2I + (2I + A_1)i + A_1\right] = 0,
\]
and the contradiction is obvious.

\(A_1\) has real eigenvalues, \(A_0 = a_0I\), \(a_0 \in \mathbb{R}\). The result can be shown as in the previous case, via the eigenvalues of \(A_1\).

Both \(A_1\) and \(A_0\) have real eigenvalues, \(A_0\) and \(A_1\) are semisimple and \(A_0A_1 = A_1A_0\). The result follows on noting that in this case \(A_0\), \(A_1\) have a complete set of eigenvectors in common.

We now can estimate \(v_i\) from (4.21). A simple modification of Lemma 2.3 and the existence of a matrix function
\[
Z_{j+1,k} = \prod_{l=j+1}^k \left( I + \frac{1}{\ell}J + \frac{1}{\ell^2}\Delta(l, J) \right)^{-1}
= \frac{1}{2\pi i} \int_{\Gamma} Z_{j+1,k+1}(\lambda, \delta(\cdot))(\lambda I - J)^{-1} d\lambda,
\]
where \(\Gamma = \{ \mu | |\mu - \lambda| = \sigma/4 \}\) and \(|\hat{Z}_{j+1,k}| \leq \text{const}|Z_{j+1,k}|\) for each \(k > j\), yield the existence of a constant \(0 < \eta < 1\) such that
\[
|\hat{Z}_{j+1,k}| \leq \text{const}(t_{j+1}/t_{k+1})^\eta
\]
and finally,
\[
|v_i| \leq \text{const}\left(t_\eta|\gamma| + \sum_{k=i+1}^{N} (t_{i+1}/t_{k+1})^\eta|r_k|\right), \quad i = 1(1)N.
\]
Using (4.25) and Lemma 2.4, we have for the solution of (4.8)
\[
|v_i| \leq \text{const}\left(t_\eta\left(|\gamma| + \|f_\delta\|\right)\right), \quad i = 1(1)N.
\]
Technical details may be found in the appendix.

For \(u_i = Ev_i, 1 \leq i \leq N\), we can now extend the results of Cases 1, 2 and 3 to the general situation when the matrix \(M\) has different eigenvalues. Let \(d_0\) be the dimension of the largest Jordan box associated with \(\lambda = 0\) and \(d_+\) the dimension of the largest Jordan box associated with \(\lambda_+ = \sigma_+ + i\kappa_+\), where \(\sigma_+\) is the smallest positive real part. Then we have the following

**Lemma 4.1.** The system (4.7) has a unique solution for any \(Qu_0 = \delta_0, Pu_N = \delta_N\) and \(f_\Delta\). Furthermore

(i) \(\|u_\Delta\| \leq \text{const}\{|\delta_0|\ln h|^{d_0-1}| + |\delta_N| + \|f_\Delta\|\}\).

(ii) Let \(y_\Delta\) be a solution of (4.3a) and \(f \in C^2\). Then
\[
\|y_\Delta - R_\Delta y\| \leq \begin{cases}
\text{const } h^\sigma |\ln h|^{d_+ - 1}, & 0 < \sigma < 2, \\
\text{const } h^2\left(|\ln h|^{d_+} + |\ln h|^{d_0 - 1}\right), & \sigma = 2, \\
\text{const } h^2|\ln h|^{d_0 - 1}, & \sigma > 2 \text{ or } S = 0.
\end{cases}
\]

The constants appearing in (i) and (ii) are independent of \(h\).

**Proof.** (i) The proof follows from the results of Cases 1–3.

(ii) [15, Lemma 4.4]; we note that if zero is not an eigenvalue of \(M\) or its Jordan box is diagonal, then we have to replace \(\ln h|^{d_0 - 1}\) by 1 in (ii). The convergence results for the case when \(f \in C\) can be found in [15, Lemma 4.4]. □
We now consider the numerical scheme defined in (4.3a), applied to the problem (3.1a). Let $A_0(t)$ and $A_1(t)$ be as in (3.3); then we have

\begin{align}
(4.27a) \quad \Theta_1(i, h) &= \left( I - \frac{A_1 + t_i^j C(t_i)}{2i} \right)^{-1} \left( I + \frac{A_1(t_i)}{2} \right) A_0(t_i), \\
(4.27b) \quad \Theta_2(i, h) &= \left( I - \frac{A_1 + t_i^j C(t_i)}{2i} \right)^{-1} \left[ \left( I + \frac{1}{i} \right) A_0(t_i) \\
&\quad + \left( I + \frac{A_1(t_i)}{2} \right) A_1(t_i) \right], \\
(4.27c) \quad \Theta_3(i, h) &= \left( I - \frac{A_1 + t_i^j C(t_i)}{2i} \right)^{-1},
\end{align}

and assuming that for $1 \leq i \leq \max_{0 \leq t \leq 1} |A_1(t)|/2$, $\Theta_j(i, h)$, $j = 1, 2, 3$ exist, we obtain the following system for $u_i$,

\begin{align}
(4.28a) \quad u_i &= u_{i-1} + \frac{1}{i} M u_{i-1} + \frac{t_i^j}{i} \hat{C}(t_i) u_{i-1} + \frac{1}{i^2} \Theta_4(i, h) u_{i-1} \\
&\quad + t_{i+1} \Theta_5(i, h) f_i, \quad i = 1(1) N_0, \quad N_0 < N,
\end{align}

when $\Theta_4$ and $\Theta_5$ are defined as before with respect to the new definitions of $\Theta_1$, $\Theta_2$, and $\Theta_3$. To show the existence and uniqueness of the solution of (4.28a), subject to the boundary conditions

\begin{align}
(4.28b) \quad Q u_0 &= \delta_0, \quad P u_{N_0} = \delta_{N_0},
\end{align}

we use a contraction argument. For that purpose we consider the following iteration scheme

\begin{align}
(4.29) \quad u^{(k+1)}_i &= u^{(k)}_i + \frac{1}{i} M u^{(k)}_{i-1} - \frac{1}{i^2} \Theta_4(i, h) u^{(k)}_{i-1} \\
&\quad + \frac{t_i^j}{i} \hat{C}(t_i) u^{(k)}_{i-1} + t_{i+1} \Theta_5(i, h) f_i, \quad i = 1(1) N_0,
\end{align}

which can be formally (cf. (4.10)) written as

\begin{align}
(4.29) \quad G U^{(k+1)} &= C U^{(k)} + F,
\end{align}

where $G, C : X_\Delta \rightarrow X_\Delta$ are linear maps. Since the existence of a bounded inverse of $G$ follows from Lemma 4.1, it remains to show that $(I - G^{-1} C)^{-1}$ exists. This holds if $\|G^{-1} C\| < 1$. It has been shown in [15, Section 4] that the latter condition is satisfied if $\tau = h N_0$ is sufficiently small. The standard contraction argument now yields the existence and uniqueness of the solution of (4.28) on $[0, \tau]$. This solution can be uniquely continued to $t = 1$. We formulate this result as

**Lemma 4.2.** Let $Q u_0 = \delta_0$ and $P u_N = \delta_N$. Then for each $f_\Delta$ there exists a unique solution $u_\Delta$ of (4.28a), i.e., the existence of the solution $y_\Delta$ of the associated second-order difference equation, and the following estimate holds,

\begin{align}
\| u_\Delta \| \leq \text{const} \left\{ |\delta_0| \| \ln h \|^{d_0 - 1} + |\delta_N| + \| f_\Delta \| \right\}.
\end{align}

Finally, for the linear boundary value problem consisting of (4.28a) and the boundary conditions (4.3b) we have the following result corresponding to Lemma 4.2, cf. [15, Remark (i) and Lemma 4.5].
Lemma 4.3. Let us assume that the homogeneous boundary value problem (3.1) has only the trivial solution. Then the difference system (4.28a), subject to the boundary conditions (4.3b) and

\[(4.30) \quad \delta = \left[ \frac{y_0}{(y_1 - y_0)/h - hA^{-1}f(0)/2} \right] = \delta, \]

has a unique solution for each \(f_\Delta\) and \(\beta\) if \(h\) is sufficiently small and the following estimate holds,

(i) \(\|y_\Delta\| \leq \text{const} \{ |\delta| |\ln h|^{d_0 - 1} + |\beta| + \|f_\Delta\| \}.\)

Clearly, if \(\delta = 0\), then we have

(ii) \(\|y_\Delta\| \leq \text{const} \{ |\beta| + \|f_\Delta\| \}.\)

(iii) Let \(y_\Delta\) be a solution of the difference system (4.28a) and \(f, C_1 \in C^2, C_0 \in C^3\). Then

\[\|y_\Delta - R_\Delta y\| \leq \begin{cases} \text{const} h^{\sigma_-} |\ln h|^{d_0 - 1}, & 0 < \sigma_- < 2, \\ \text{const} h^2 (|\ln h|^{d_0} + |\ln h|^{d_0 - 1}), & \sigma_- = 2, \\ \text{const} h^2 |\ln h|^{d_0 - 1}, & \sigma_- > 2 \text{ or } S = 0. \end{cases}\]

5. The Nonlinear Problems. We now consider the nonlinear problem (1.1) and write it first as the following nonlinear operator equation

\[
\begin{aligned}
N y(t) &\equiv y''(t) - \frac{A_1}{t} y'(t) - \frac{A_0}{t^2} y(t) - f(t, y(t), y'(t)) = 0, \\
B(y(0); y(1), y'(1)) &= 0, \\
\bar{Q} Y(0) &= 0.
\end{aligned}
\]

The discrete problem related to (5.1) is of the form (cf. (4.1))

\[
\begin{aligned}
N_\delta v_i &\equiv \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} - \frac{A_1}{t_i} \left( \frac{v_{i+1} - v_{i-1}}{2h} \right) - \frac{A_0}{t_i^2} v_i \\
-f(t_i, v_i, \frac{v_{i+1} - v_{i-1}}{2h}) &= 0, \quad i = 1(1)N, \\
B(v_0; v_N, \frac{v_{N+1} - v_{N-1}}{2h}) &= 0, \\
\bar{Q} V_0 &= 0.
\end{aligned}
\]

Before proceeding, the following additional notations are required:

\[
\begin{aligned}
B_\delta^{-1}(y(0); y(1), (y(1 + h) - y(1 - h))/2h), \\
B_\delta^{-1}(y(0); y(1), (y(1 + h) - y(1 - h))/2h), \\
B_\delta^{-1}(y(0); y(1), (y(1 + h) - y(1 - h))/2h), \\
C_\delta(0)(t_i) = \frac{\partial f}{\partial v}(t_i, y(t_i), (y(t_{i+1}) - y(t_{i-1}))/2h), \\
C_\delta(1)(t_i) = \frac{\partial f}{\partial w}(t_i, y(t_i), (y(t_{i+1}) - y(t_{i-1}))/2h). \\
\end{aligned}
\]
The linearization of the nonlinear difference operator defined by (5.2) (at the solution \( y(t) \)) and applied to \( v_\Delta \) is

\[
L_\Delta(y_\Delta)v_i \equiv \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} - \left( \frac{A_1}{t_i} + \frac{\partial f}{\partial y}(t_i, y(t_i), (y(t_{i+1}) - y(t_{i-1}))/2h) \right) \left( \frac{v_{i+1} - v_{i-1}}{2h} \right) - \left( \frac{A_0}{t_i^2} + \frac{\partial f}{\partial y}(t_i, y(t_i), (y(t_{i+1}) - y(t_{i-1}))/2h) \right)v_i, \quad i = 1(1)N,
\]

(5.3a)

\[
B_\Delta(y_\Delta)v_\Delta \equiv B_{00}^\Delta v_0 + B_{10}^\Delta v_N + B_{11}^\Delta \left( \frac{v_{N+1} - v_{N-1}}{2h} \right).
\]

(5.3b)

\[
Q_\Delta v_\Delta \equiv \tilde{Q} \left[ \left( v_1 - v_0 \right)/h - hA^{-1} \frac{\partial f}{\partial y}(0, y(0), 0)v_0/2 \right].
\]

(5.3c)

Let us apply the difference scheme to the linear problem (3.5) augmented by the condition \( QV(0) = 0 \). This yields

\[
\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} - \left( \frac{A_1}{t_i} + C_1^\Delta(t_i) \right) \left( \frac{v_{i+1} - v_{i-1}}{2h} \right) - \left( \frac{A_0}{t_i^2} + C_0^\Delta(t_i) \right)v_i = 0, \quad i = 1(1)N,
\]

(5.4a)

\[
B_{00}^\Delta v_0 + B_{10}^\Delta v_N + B_{11}^\Delta \left( \frac{v_{N+1} - v_{N-1}}{2h} \right) = 0.
\]

(5.4b)

To derive a discretization of \( \tilde{Q}V(0) = 0 \) we assume \( v \in C^2 \), apply Taylor's theorem to (3.5a) and obtain

\[
A v''(0) = (I - A_1 - A_0/2)v''(0) = \frac{\partial f}{\partial v}(0, y(0), 0)v(0)
\]

if \((A_1 + A_0)v'(0) = 0 \) and \( A_0v(0) = 0 \). It follows that

\[
\tilde{Q} \left[ \left( v_1 - v_0 \right)/h - hA^{-1} \frac{\partial f}{\partial y}(0, y(0), 0)v_0/2 \right] = 0,
\]

(5.4c)

and we can see from (3.6) that (5.4) is equivalent to

\[
L_\Delta(y_\Delta)v_i = 0, \quad 1 \leq i \leq N; \quad B_\Delta(y_\Delta)v_\Delta = 0; \quad Q_\Delta v_\Delta = 0.
\]

(5.5)

The main result of this section is formulated in Theorem 5.1. The proof of this theorem follows from the theory developed by Keller [6]. Thus the proof of Theorem 5.1 reduces to verifying the hypotheses of Theorem 4.7 in [6]. For notational convenience we define \( D_\Delta x = (D_0 x_0, D_1 x_1, \ldots, D_N x_N) \), where \( D_0 x_0 = (x_1 - x_0)/h - hA^{-1}f(0, x_0, 0)/2 \) and \( D_i x_i = (x_{i+1} - x_{i-1})/2h, \quad i = 1(1)N \).

**Theorem 5.1.** Let the conditions N.1.1–N.1.5 be satisfied. Then for some \( \varepsilon > 0 \), \( h_0 > 0 \) sufficiently small and all \( h < h_0 \),

(i) the difference system (5.2) has a unique solution \( v_\Delta \) and \( \|v_\Delta - R_\Delta y\| < \varepsilon \);
The difference solution can be computed by Newton's method, which converges quadratically for any initial iterate $y_0$ with $\|y_0 - R\| \leq \varepsilon$ and $\|Dy_0 - R\| \leq \varepsilon$ provided that $\varepsilon (\varepsilon \leq \rho, \delta)$ is sufficiently small.

(iii) \[
\|y_0 - R\| \leq \begin{cases} 
\text{const} \left( h \ln h \right)^{1/2} + \omega \left( f(.), y(.), y'(.) \right), & \text{if } 1 < \sigma < 2, \\
\text{const} \left( h \ln h \right)^{1/2} + \omega \left( f(.), y(.), y'(.) \right), & \text{if } \sigma = 2, \\
\text{const} \left( h + \omega \left( f(.), y(.), y'(.) \right) \right), & \text{if } \sigma > 2 \text{ or } S = 0.
\end{cases}
\]

Proof. (i), (ii). The necessary conditions for (i) and (ii) to hold are (1) stability of (5.5), provided that (3.5) has a unique solution and (2) uniform Lipschitz continuity of the operators $L_\Delta(z_\Delta)$ and $B_\Delta(z_\Delta)$ for all $z \in U_{\rho, \delta}$, i.e., the existence of constants $K_L$ and $K_B$ such that
\[
\|L_\Delta(z_\Delta) - L_\Delta(\xi_\Delta)\| \leq K_L \|z_\Delta - \xi_\Delta\|,
\]
\[
\|B_\Delta(z_\Delta) - B_\Delta(\xi_\Delta)\| \leq K_B \|z_\Delta - \xi_\Delta\|,
\]
for all $z, \xi \in U_{\rho, \delta}$. We point out that the operators $L_\Delta(z_\Delta)$ and $B_\Delta(z_\Delta)$ can be represented by the coefficient matrices from (5.3a, b). Since (1) and (2) hold according to Lemma 4.3, and by the Lipschitz continuity of the derivatives $f_e$ and $f_u$, see N.1.3, respectively, the assertions (i) and (ii) follow from [6, Theorem 4.7].

(iii) We consider the system (4.1). We substitute $R_\Delta y$ into the scheme and obtain
\[
\frac{y(t_{i+1}) - 2y(t_i) + y(t_{i-1})}{h^2} - \frac{A_1}{t_i} \left( \frac{y(t_{i+1}) - y(t_{i-1})}{2h} \right) - \frac{A_0}{t_i^2} y(t_i)
\]
\[
= f(t_i, y(t_i), y'(t_i)) = \delta_i, \quad i = 1(1) N,
\]
\[
B \left( y(0); y(1), \frac{y(1 + h) - y(1 - h)}{2h} \right) = \delta_B,
\]
\[
\tilde{Q} \left( \frac{y(0)}{(y(h) - y(0))/h - hA^{-1}f(0, y(0), 0)/2} \right) = \delta_0.
\]

From the mean-value theorem and $y \in C^2[0,1]$ we have
\[
\delta_i = y''(\xi_i) - y''(t_i) - \frac{A_1}{t_i} (y'(\eta_i) - y'(t_i))
\]
\[
= - \left( f(t_i, y(t_i), y'(\eta_i)) - f(t_i, y(t_i), y'(t_i)) \right)
\]
\[
y''(\xi_i) - y''(t_i) - \left( \frac{A_1}{t_i} + F_1(t_i) \right) (y'(\eta_i) - y'(t_i)), \quad i = 1(1) N,
\]
where $t_{i-1} < \xi_i, \eta_i < t_{i+1},$
\[
F_1(t_i) = \int_0^1 \frac{\partial f}{\partial y'}(t_i, y(\eta_i), y'(\eta_i) + \tau (y'(t_i) - y'(\eta_i))) d\tau,
\]
\[
\delta_B = B \left( y(0); y(1), y'(\eta_{N+1}) \right) - B \left( y(0); y(1), y'(1) \right)
\]
\[
= \int_0^1 \frac{\partial B}{\partial u_3} \left( y(0); y(1), y'(\eta_{N+1}) \right.
\]
\[
+ \tau (y'(1) - y'(\eta_{N+1})) \big) d\tau \left( y'(\eta_{N+1}) - y'(1) \big),
\]
License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
where \( 1 - h < \eta_{N+1} < 1 + h \), and

\[
(5.7c) \quad \delta_0 = \tilde{Q} \left[ y'(\eta_0) - y'(0) \right],
\]

where \( 0 < \eta_0 < h \).

On the other hand, subtracting (4.1) from (5.6) and defining \( e_i = y(t_i) - y_i, \) \( i = 0(1)N + 1 \), we obtain the following system for \( e_\Delta \),

\[
\frac{e_{i+1} - 2e_i + e_{i-1}}{h^2} - \left( \frac{A_1}{t_i} + G_1(t_i) \right) \left( \frac{e_{i+1} - e_{i-1}}{2h} \right)
\]

\[
- \left( \frac{A_0}{t_i^2} + G_0(t_i) \right) e_i = \delta_i, \quad i = 1(1)N,
\]

where

\[
G_1(t_i) = \int_0^1 \frac{\partial f}{\partial y} \left( t_i, y, \frac{y(t_{i+1}) - y(t_{i-1})}{2h} - \frac{e_{i+1} - e_{i-1}}{2h} \right) d\tau,
\]

\[
G_0(t_i) = \int_0^1 \frac{\partial f}{\partial y} \left( t_i, y(t_i) - \tau e_i, \frac{y(t_{i+1}) - y(t_{i-1})}{2h} \right) d\tau,
\]

and

\[
\int_0^1 \frac{\partial B}{\partial u_1} \left( y(0) - \tau e_0; y(1), \frac{y(1+h) - y(1-h)}{2h} \right) d\tau e_0
\]

\[
+ \int_0^1 \frac{\partial B}{\partial u_2} \left( y_o; y(1) - \tau e_N, \frac{y(1+h) - y(1-h)}{2h} \right) d\tau e_N
\]

\[
+ \int_0^1 \frac{\partial B}{\partial u_3} \left( y_0; y_N, \frac{y(1+h) - y(1-h)}{2h} \right.
\]

\[
- \frac{e_{N+1} - e_{N-1}}{2h} d\tau \frac{e_{N+1} - e_{N-1}}{2h} = \delta_B.
\]

\[
(5.8c) \quad \tilde{Q} \left[ \frac{y(0) - y_0}{h - hA^{-1}f(0, y(0), 0)/2 - (y_1 - y_0)/h + hA^{-1}f(0, y_0, 0)/2} \right]
\]

\[
= \tilde{Q} \left[ \frac{e_0}{h - hA^{-1} \int_0^1 \frac{\partial f}{\partial y} (0, y(0) - \tau e_0, 0) d\tau e_0/2} \right] = \delta_0.
\]

The result follows by virtue of Lemma 4.3 and by smoothness properties of \( f \) and \( y \), following immediately from (3.8), see also [15, Lemma 4.4].

Let \( f \in C^p[T_{p,\delta}] \) denote that \( f(t, v, w) \) is \( p \) times continuously differentiable on \( T_{p,\delta} \). We see from Theorem 5.1(iii) that if \( \sigma_+ > 2 \) and \( f \in C^1[T_{p,\delta}] \), then the method converges at least as \( O(h) \). This result can be improved, if we assume that \( f \in C^2[T_{p,\delta}] \).

**Theorem 5.2.** Let \( f \in C^2[T_{p,\delta}] \). Then

\[
\| y_\Delta - R_\Delta y \| \leq \begin{cases} 
\text{const } h^{\sigma_+} |\ln h|^{d_+ - 1}, & 1 < \sigma_+ < 2, \\
\text{const } h^2 \left( |\ln h|^{d_+} + |\ln h|^{d_0 - 1} \right), & \sigma_+ = 2,
\end{cases}
\]

\[
\text{const } h^2 |\ln h|^{d_0 - 1}, \quad \sigma_+ > 2 \text{ or } S = 0.
\]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Proof. Since $y \in C^4(0,1]$, we apply Taylor's theorem to (5.6) and obtain

\begin{equation}
\delta_i = h^2 cy^{(4)}(\xi_i) - \left( \frac{A_i}{t_i} + F_i(t_i) \right) h^2 dy^{(\nu)}(\eta_i), \quad i = 1(1)N,
\end{equation}

where $t_{i-1} < \xi_i$, $\eta_i < t_{i+1}$ and

$$F_i(t_i) = \int_0^1 \frac{\partial f}{\partial y}(t_i, y(t_i), y'(t_i) + \tau dy^{(\nu)}(\eta_i)) \, d\tau,$$

\begin{equation}
\delta_B = \int_0^1 \frac{\partial B}{\partial u_3}(y(0); y(1), y'(1) + \tau dy^{(\nu)}(\eta_{N+1})) \, d\tau dy^{(\nu)}(\eta_{N+1}),
\end{equation}

where $1 - h < \eta_{N+1} < 1 + h$, and

\begin{equation}
\delta_0 = \hat{Q} \left[ h^2 dy^{(\nu)}(\eta_0) \right],
\end{equation}

where $0 < \xi_0 < h$. The result follows now in a very similar way, cf. [15, Lemma 4.4].

The extension of the results of Theorems 5.1 and 5.2 to the problem (1.2) is straightforward. In particular (i) and (ii) of Theorem 5.1 hold by [6, Theorem 4.7], so we shall restrict our attention to deriving the associated convergence results, which we formulate in

**Theorem 5.3.** Let the assumptions N.2.1–N.2.5 hold. Then

(i)

$$\|y - R_\Delta y\| \leq \begin{cases} \text{const} h |\ln h|^{d_s-1} + \omega (f(\cdot, y(\cdot), \xi(\cdot)), h), & \lambda = 1 \text{ or } 1 < \sigma_s < 2, \\ \text{const} h^{2} |\ln h|^{d_s} + \omega (f(\cdot, y(\cdot), \xi(\cdot)), h), & \sigma_s = 2, \\ \text{const} (h + \omega (f(\cdot, y(\cdot), \xi(\cdot)), h)), & \sigma_s > 2 \text{ or } S = 0, \end{cases}$$

where $y(t)/t = \xi(t)$.

(ii) If $f \in C^2[T_{\rho,1}]$, then

$$\|y - R_\Delta y\| \leq \begin{cases} \text{const} h^{a_s} |\ln h|^{d_s-1}, & \lambda = 1 \text{ or } 1 < \sigma_s < 2, \\ \text{const} h^{2} |\ln h|^{d_s}, & \sigma_s = 2, \\ \text{const} h^{2}, & \sigma_s > 2 \text{ or } S = 0. \end{cases}$$

**Proof.** (i) In this case we have

\begin{equation}
\delta_i = \left( y''(\xi_i) - y''(t_i) \right) - \frac{A_i}{t_i} \left( y'(\eta_i) - y'(t_i) \right), \quad i = 1(1)N,
\end{equation}

\begin{equation}
\delta_B = \int_0^1 \frac{\partial B}{\partial u_2}(y(1); y'\eta_{N+1}) \\
+ \tau \left( y'(1) - y'\eta_{N+1}) \right) \, d\tau \left( y'\eta_{N+1} - y'(1) \right),
\end{equation}

\begin{equation}
\delta_0 = 0.
\end{equation}

Furthermore, $\delta_i$ can be estimated as in [15, Lemma 4.4] and $\delta_B = O(h)$. This completes the proof of (i).
(ii) From Taylor's theorem we have

$$\delta_i = h^2 cy^{(4)}(\xi_i) - \frac{A_1}{t_i} h^2 dy^{(4)}(\eta_i) = \begin{cases} O(t_i^{-\sigma - 2}h^2|\ln h|^{d_i-1}), & \lambda = 1, 1 < \sigma < 2, \\ O(t_i^{-\sigma}h^2|\ln h|^{d_i}), & \sigma = 2, \\ O(t_i^{-\sigma}h^2), & \sigma > 2, \\ O(h^2), & S = 0, \end{cases}$$

and $\delta_B = O(h^2)$. The proof follows now by [15, Lemma 4.4], see also [15, formulas (4.16), (4.21), (4.23)]. □


Example 1. To illustrate the results of Lemma 4.1(ii) we consider a homogeneous $2 \times 2$ system, where

$$A_1 = \begin{bmatrix} a_1 & 0 \\ 0 & a_1 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & a_0 \\ 0 & 0 \end{bmatrix}.$$ 

If $a_1 = 2$ and $a_0 = 1$, then $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \lambda_4 = 2$, and the boundary conditions $y_1(0) = 3$, $y_2(0) = 0$, $y_2(1) = 4$, $y_1(1) = -2$
yield the solution

$$y(t) = \begin{bmatrix} 3 - 2t^2(1 - \ln t) \\ 4t^2 \end{bmatrix}.$$ 

For $a_1 = 2.5$ and $a_0 = 1$ we have $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = \lambda_4 = 2.5$. We choose $y_1(0) = 3$, $y_2(0) = 0$, $y_2(1) = 5$, $y_1(1) = -3$, and get the following solution

$$y(t) = \begin{bmatrix} 3 - 2t^{2.5}(1 - \ln t) \\ 5t^{2.5} \end{bmatrix}.$$ 

Furthermore,

$$J = \begin{bmatrix} a_1 & 1 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and the error behavior following from Lemma 4.1(ii) is illustrated in Tables (6.1) and (6.2). We note that in this case $\delta_0 = 0$.

| $h$ | $\Delta(h)$ | $\Delta(\Delta(h))/h^2|\ln h|$ | $\Delta(h)/\Delta(h)$ |
|-----|-------------|----------------|------------------|
| 1/10 | 2.133 E - 2 | 0.9264 | * |
| 1/20 | 6.199 E - 3 | 0.8277 | 0.2906 |
| 1/40 | 1.766 E - 3 | 0.7660 | 0.2846 |
| 1/80 | 4.956 E - 4 | 0.7240 | 0.2807 |
| 1/160 | 1.375 E - 4 | 0.6936 | 0.2774 |
| 1/320 | 3.775 E - 5 | 0.6701 | 0.2745 |
| 1/640 | 1.028 E - 5 | 0.6517 | 0.2723 |

$$\lim_{h \to 0} \frac{\Delta(h/2)}{\Delta(h)} = \left(\frac{1}{2}\right)^2 = 0.25.$$ 

Here, $\Delta(h) = \max_{0 \leq i \leq N} |y_1(t_i) - y_{i,1}|$. The second component has been computed without errors.
Example 2. We now investigate the $2 \times 2$ system, where
\[
A_1 = \begin{bmatrix} 2\sqrt{2} - 1 & 0 \\ 2\sqrt{2} & -1 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad f(t) = \begin{bmatrix} 6 - 4\sqrt{2} \\ -10 - 4\sqrt{2} \end{bmatrix}
\]
and
\[
J = \begin{bmatrix} 2 + \sqrt{2} & 0 & 0 \\ 0 & 2 - \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

The boundary conditions are
\[
y_1'(0) - y_2'(0) = 0, \quad y_1(1) = 7, \quad y_2'(1) = \sqrt{2} - 4,
\]
and the solution is
\[
y(t) = \begin{bmatrix} t^{\sqrt{2} + 2} + t^2 + 5 \\ t^{\sqrt{2} + 2} - 3t^2 - 5 \end{bmatrix}.
\]
In this case $y \in C^3[0,1]$ and therefore
\[
(i) \quad \frac{y(h) - y(0)}{h} - y'(0) = O(h),
\]
\[
(ii) \quad \left( \frac{y(h) - y(0)}{h} - hA^{-1}f(0)/2 \right) - y'(0) = O(h^2),
\]
since $(A_1 + A_0)y'(0) = 0$ and $A_0 y(0) = 0$. We can now see that even the order $h$ approximation of the first derivative at $t = 0$ does not influence the error behavior (cf. [14, Lemma 4.2]).
Example 3. Finally, we consider the nonlinear scalar equation
\[ y''(t) + \frac{2}{t}y'(t) = -y^5(t), \quad y'(0) = 0, \quad y(1) = \frac{\sqrt{3}}{4}. \]
The solution of this problem is \( y(t) = \frac{1}{y_1 + \frac{r_2}{3} g C_0}. \) The eigenvalues of \( M \) are \( \lambda_1 = 0, \lambda_2 = -1, \) and it follows from Theorem 5.2 that the order of convergence is 2.

<table>
<thead>
<tr>
<th>Table 6.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
</tr>
<tr>
<td>1/4</td>
</tr>
<tr>
<td>1/8</td>
</tr>
<tr>
<td>1/16</td>
</tr>
<tr>
<td>1/32</td>
</tr>
<tr>
<td>1/64</td>
</tr>
<tr>
<td>1/128</td>
</tr>
</tbody>
</table>

\[ \lim_{h \to 0} \frac{\Delta(h)/h^2}{\Delta(h)} = 0.25. \]

All examples were computed on a CDC Cyber 170/172 in single precision.

7. Appendix.

7.1. Case 1. \( \sigma < 0. \) In order to study the growth of the solution \( v_i \) of the difference system (4.12) we consider the scalar case first and the case when \( n = 2 \) afterwards. In both cases we define a matrix function
\[ Z_{i,k+1} = \prod_{l=i}^{k+1} \left( I + \frac{1}{l} J + \frac{1}{l} \Delta(l, J) \right) \]
in such a way that
\[ |Z_{i,k+1}| \leq |Z_{i,k+1}|, \quad k < i. \]
Furthermore, we define the matrix \( \Delta(l, J) \) as a diagonal matrix to make the system (4.12) decoupled.

Let \( n = 1, a_0, a_1 \in \mathbb{R} \) and let us assume that the eigenvalues of \( M \) are \( \lambda_1 = \lambda \) and \( \lambda_2 = \bar{\lambda}. \) It follows immediately from the form of \( M \) that
\[
E = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}, \quad E^{-1} = \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix}.
\]
We rewrite \( \Theta_1 \) and \( \Theta_2 \) as
\[
\Theta_1(i) = \frac{2i}{2i - a_1} a, \quad \Theta_2(i) = \frac{2i}{2i - a_1} b
\]
and obtain
\[
\Theta(i) = E^{-1} \Theta_4(i) E = \frac{2i}{2i - a_1} \cdot \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} -(a + \lambda_1 b) & -(a + \lambda_2 b) \\ a + \lambda_1 b & a + \lambda_2 b \end{bmatrix}.
\]
Finally, with \( c = 2i/(2i - a_1)(\lambda_2 - \lambda_1) \) we have
\[
I + \frac{1}{i} J + \frac{1}{i^2} \Theta(i) = \begin{bmatrix} 1 + \frac{1}{i} \lambda_1 - \frac{c}{i^2}(a + \lambda_1 b) & -\frac{c}{i^2}(a + \lambda_2 b) \\ \frac{c}{i^2}(a + \lambda_1 b) & 1 + \frac{1}{i} \lambda_2 + \frac{c}{i^2}(a + \lambda_2 b) \end{bmatrix}
\]
\[ \equiv N(i), \]

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
and it can easily be seen that
\[ |N(i)| = \left| 1 + \frac{1}{i} \lambda_1 - \frac{c}{i^2} (a + \lambda_1 b) \right| + \left| \frac{c}{i^2} (a + \lambda_1 b) + \frac{1}{i} \lambda_2 + \frac{c}{i^2} (a + \lambda_2 b) \right|.
\]
Furthermore, since for any \( \alpha \in \mathbb{C} \) and positive real number \( \beta \), \( |\alpha| + \beta \leq |\eta| \), where
\[
\text{Re}(\eta) = \text{Re}(\alpha) + \text{sign}(\text{Re}(\alpha)) \beta, \quad \text{Im}(\eta) = \text{Im}(\alpha) + \text{sign}(\text{Im}(\alpha)) \beta,
\]
we have for each \( i \)
\[
|N(i)| \leq \left| 1 + \frac{1}{i} \lambda \right| + \frac{2}{i} |c(a + \lambda b)| = \left| 1 + \frac{1}{i} \lambda \right| + \frac{1}{i^2} |\Theta(i)|
\]
\[
\leq \left| 1 + \frac{1}{i} \lambda + \frac{1}{i^2} \delta(i, \lambda) \right|,
\]
where
\[
\text{Re}(\delta(i, \lambda)) = \text{sign}(\text{Re}(1 + \frac{1}{i} \lambda)) |\Theta(i)|, \quad \text{Im}(\delta(i, \lambda)) = \text{sign}(\text{Im}(\lambda)) |\Theta(i)|.
\]
Clearly, we define
\[
(7.3) \quad \Delta(I, J) := \begin{bmatrix} \delta(i, \lambda) & 0 \\ 0 & \delta(i, \lambda) \end{bmatrix}.
\]
We notice that although the matrix \( N(i) \) is a full matrix, a certain structure is present; the norm obtained from the first row belonging to \( \lambda \) and the norm obtained from the second row belonging to \( \bar{\lambda} \) are equal. We shall now show that this structure remains unchanged if \( n = 2 \).

Let \( n = 2 \). Let us assume that \( a_i, b_i \in \mathbb{R} \), \( i = 1(1)4 \) and
\[
M = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ a_1 & a_2 & b_1 & b_2 \\ a_3 & a_4 & b_3 & b_4 \end{bmatrix}, \quad J = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \bar{\lambda} & 1 \\ 0 & 0 & 0 & \bar{\lambda} \end{bmatrix};
\]
then we have the following transformation matrix \( E \),
\[
E = \begin{bmatrix} a & b & \bar{a} & \bar{b} \\ 1 & 0 & 1 & 0 \\ a\lambda & a + b\lambda & a\bar{\lambda} & a + b\bar{\lambda} \\ \lambda & 1 & \bar{\lambda} & 1 \end{bmatrix},
\]
where
\[
a = a(\lambda) = (\lambda^2 - \lambda b_4 - a_4)/(a_3 + b_3\lambda), \quad \bar{a} = a(\bar{\lambda}),
\]
\[
b = b(\lambda) = (2\lambda - b_4 - ab_3)/(a_3 + b_3\lambda), \quad \bar{b} = b(\bar{\lambda}).
\]
Furthermore, with \( d = \det(E) = |b(\lambda - \bar{\lambda})^2 - (a - \bar{a})^2| \in \mathbb{R} \),
\[
E^{-1} = \frac{1}{d} \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} b - \bar{b} & b(a + b\lambda) - \bar{b}(a + b\lambda) \\ (a - \bar{a}) + \bar{b}(\lambda - \bar{\lambda}) & (a + a\lambda)(a - \bar{a}) + \bar{b}(a\lambda - a\lambda) \\ -b - \bar{b} & -b(a + b\lambda) - \bar{b}(a + b\lambda) \\ -a - \bar{a} - b(\lambda - \bar{\lambda}) & -(a + b\lambda)(a - \bar{a}) + b(\bar{\lambda} - a\lambda) \end{bmatrix}.
\]
Finally, we calculate $\Theta(i) = (m_{kj})/d$, $k, j = 1(1)4$, and by comparing the first row with the third, and the second row with the fourth, we see that

\begin{align*}
(7.4a) \quad & m_{11} = \overline{m_{33}}, \quad m_{12} = \overline{m_{34}}, \quad m_{13} = \overline{m_{31}}, \quad m_{14} = \overline{m_{32}}, \\
(7.4b) \quad & m_{21} = m_{43}, \quad m_{22} = m_{44}, \quad m_{23} = m_{41}, \quad m_{24} = m_{42}.
\end{align*}

For $N(i) = I + J/i + \Theta(i)/i^2 = (n_{kj})$, $k, j = 1(1)4$, (7.4) holds and this yields

\begin{align*}
(7.5) \quad & \sum_{j=1}^{4} |n_{1j}| = \sum_{j=1}^{4} |n_{3j}|, \quad \sum_{j=1}^{4} |n_{2j}| = \sum_{j=1}^{4} |n_{4j}|.
\end{align*}

Motivated by (7.5), we choose for $\Delta(l, J)$ in (7.1),

\begin{align*}
(7.6) \quad & \Delta(l, J) := \text{diag}(\delta(l, \lambda), \delta(l, \lambda), \delta(l, \bar{\lambda}), \delta(l, \bar{\lambda})),
\end{align*}

where $\delta(\cdot)$ is defined as in the scalar case. Furthermore, (7.5) justifies the "Jordan-boxwise" considerations of the system (4.12).

7.2. Case 3. $\sigma > 0$. We now restrict our attention to the scalar case. The results can be carried over to systems in a similar manner as in Subsection 7.1. Let $n = 1$ and consider $(N(i))^{-1}$,

\begin{align*}
(7.7) \quad & |(N(i))^{-1}| = 1/\left|1 + \frac{1}{i} \lambda - \frac{c}{i^2} (a + \lambda b)\right|.
\end{align*}

Since there exists an index $i_0$ such that for all $i \geq i_0$, $|1 + \lambda/i| \geq |2c(a + \lambda b)/i^2|$, we conclude that

\begin{align*}
(7.8) \quad & \text{Re}(\delta(i, \lambda)) = -|\Theta(i)| \text{cos} \varphi_i = -\text{sign}(\text{Re}(1 + \lambda/i)) |\Theta(i)| \text{cos} \varphi_i, \\
& \text{Im}(\delta(i, \lambda)) = -\text{sign}(\text{Im} \lambda) |\Theta(i)| \text{sin} \varphi_i, \quad \varphi_i = \arctan\left(\frac{\kappa}{i + \sigma}\right).
\end{align*}

Because of (7.7) we define

\begin{align*}
(7.9) \quad & Z_{j+1,k} := \prod_{l=j+1}^{k} \left(I + \frac{1}{l} J + \frac{1}{l^2} \Delta(l, J)\right)^{-1}, \quad k > j,
\end{align*}

where $\Delta(l, J) = \text{diag}(\delta(l, \lambda), \delta(l, \bar{\lambda}))$ and $\delta(l, \lambda)$ is given by (7.8). Finally, we define for $\lambda, \Omega_1(\lambda)$ and $\Omega_2(\lambda)$ as in Lemma 2.2 and for $\delta(l, \mu)$ as in Lemma 2.3,
and show that there exists a constant $\eta > 0$ such that
\[
|z_{kj}(\mu, \delta(\cdot))| \leq \text{const}(t_k/t_j)^{\eta}, \quad k < j, \; j = 1(1)N.
\]
To see this, we choose an index $l_0$ so large that for all $l > l_0$, $|\mu + \delta(l, \mu)/l| < 1$. Then we have
\[
\prod_{l=k}^{j-1} \left(1 + \frac{1}{l}\left(\mu + \frac{1}{l}\delta(l, \mu)\right)\right) = \prod_{l=k}^{j-1} \left(1 - \frac{1}{l}\left(\mu + \frac{1}{l}\delta(l, \mu)\right)\right),
\]
and there exists an index $\hat{\rho} > 1$ such that $|\delta(l, \mu)| \leq \hat{\delta}$ for all $l \geq \hat{\rho}$. The result now follows from Lemma 2.3 for all $k \geq k_0 = \max\{i_0, l_0, \hat{\rho}, 4\hat{\delta}/\sigma\}$. Since for $k \leq k_0$, $z_{k,k_0}(\mu, \delta(\cdot))$ consists of at most $k_0$ terms, the result holds for all $k \geq 1$. Since
\[
Z_{j+1,k} = \frac{1}{2\pi i} \int_{\Gamma} z_{j+1,k+1}(\lambda, \delta(\cdot))(\lambda I - J)^{-1} d\lambda, \quad \Gamma = \{\mu| |\lambda - \mu| \leq \sigma/4\},
\]
and $|\hat{Z}_{j+1,k}| \leq |Z_{j+1,k}|$ for all $k > j$, the estimate (4.25) holds. It should be mentioned that in the case when $n \geq 2$ we have
\[
\left|\left(I + \frac{1}{l}J + \frac{1}{l^2}\Theta(l)\right)^{-1}\right| \leq m(l, n)\left|\left(I + \frac{1}{l}J + \frac{1}{l^2}\Delta(l, J)\right)^{-1}\right|
\]
for $1 \leq l \leq N$, and $m(l, n)$ is uniformly bounded with respect to $l$ and $n$.

Institut für Angewandte und Numerische Mathematik
Technische Universität Wien
Wiedner Hauptstrasse 6-10
A-1040 Wien, Austria