A CLASS OF SINGULARLY PERTURBED SEMILINEAR DIFFERENTIAL EQUATIONS WITH INTERIOR LAYERS

P. A. FARRELL, E. O’RIORDAN, AND G. I. SHISHKIN

Abstract. In this paper singularly perturbed semilinear differential equations with a discontinuous source term are examined. A numerical method is constructed for these problems which involves an appropriate piecewise-uniform mesh. The method is shown to be uniformly convergent with respect to the singular perturbation parameter. Numerical results are presented that validate the theoretical results.

1. Introduction

In this paper a class of singularly perturbed semilinear ordinary differential equations is considered on the unit interval \(\Omega = (0, 1)\). A single discontinuity in the source term is assumed to occur at a point \(d \in \Omega\). It is convenient to introduce the notation \(\Omega^- = (0, d)\) and \(\Omega^+ = (d, 1)\) and to denote the jump at \(d\) in any function with \([\omega](d) = \omega(d^+) - \omega(d^-)\). The problem follows.

Find \(u_\varepsilon \in C^1(\bar{\Omega}) \cap C^2(\Omega^- \cup \Omega^+)\) such that

\[
\begin{align*}
-\varepsilon u''_\varepsilon + b(u)u_\varepsilon &= f \quad \text{for all } x \in \Omega^- \cup \Omega^+, \\
u_\varepsilon(0) &= A, \quad u_\varepsilon(1) = B, \\
f(d^-) &\neq f(d^+), \quad b(0) > 0, \\
b &\in C^4(-\infty, \infty), \quad f \in C^4(\bar{\Omega} \setminus \{d\}).
\end{align*}
\]

Below we impose further restrictions on the magnitudes of \(\|f\|_{\bar{\Omega}}\), the boundary values \(|u_\varepsilon(0)|, |u_\varepsilon(1)|\), and the class of nonlinear functions \(b(\cdot)\) that will be examined. These restrictions are introduced at appropriate locations in the paper. Because \(f\) is discontinuous at \(d\), the solution \(u_\varepsilon\) of (1.1) does not necessarily have a continuous second order derivative at the point \(d\). Thus \(u_\varepsilon \notin C^2(\Omega)\), but the first derivative of the solution exists and is continuous. If \(f \in C^1(\Omega)\), then under certain restrictions on the nonlinearity \(b(u)u_\varepsilon\), only boundary layers would appear in the solution of (1.1). The asymptotic structure of the solutions of singularly perturbed semilinear differential equations with both boundary and interior layers is given in [1].
D’Annunzio [2] examined semilinear problems, whose solutions displayed both boundary and interior layer phenomena, but D’Annunzio placed restrictions on the mesh size so the number of mesh intervals employed depended adversely on the small parameter. In this paper our goal is to design numerical methods which are parameter uniform. That is, if \( u_\varepsilon \) is a solution of (1.1) and \( U_\varepsilon \) is a numerical approximation, then

\[
\|u_\varepsilon - U_\varepsilon\|_\infty \leq Cg(N), \quad g(N) \to 0 \text{ as } N \to \infty,
\]

where the number of mesh intervals \( N \) is independent of \( \varepsilon \), and \( C \) is a constant independent of \( \varepsilon \) and \( N \). Shishkin [12] established parameter-uniform convergence for a class of quasi-linear parabolic equations with smooth data using finite difference schemes based on piecewise-uniform meshes. The numerical method presented in this paper is also based on piecewise-uniform meshes. Singularly perturbed linear problems with discontinuous data were treated in [13]. A linear version of (1.1) was studied in [7], where a parameter-uniform numerical method based on a suitably designed piecewise-uniform mesh adapted to the interior layer was shown to converge with \( g(N) = N^{-1} \ln N \). The methodology in [7] is extended in this paper to the nonlinear problem (1.1). In [5] it was shown that numerical methods based on uniform meshes cannot be parameter uniform for semilinear singularly perturbed problems. Sun and Stynes [14] constructed finite difference schemes based on piecewise-uniform meshes for semilinear problems whose solutions exhibit only boundary layer structure. In this paper we are primarily interested in the interior layer behaviour introduced by the discontinuity of \( f \).

2. The continuous problem

We introduce the concepts of upper and lower solutions, which are useful in establishing existence and in determining the character of the solution.

**Definition 1.** A function \( \alpha \in C^0(\bar{\Omega}) \cap C^2(\Omega^- \cup \Omega^+) \) is a lower solution of (1.1) if

\[
\begin{align*}
-\varepsilon \alpha'' + b(\alpha) \alpha &\leq f, \quad x \neq d \\
\alpha'(d+) &\geq \alpha'(d-) \\
\alpha(0) &\leq u_\varepsilon(0), \quad \alpha(1) \leq u_\varepsilon(1)
\end{align*}
\]

An upper solution \( \beta \) is defined in an analogous fashion, with all inequalities reversed.

**Theorem 2.1** ([10]). If \( \alpha, \beta \in C^0(\bar{\Omega}) \cap C^2(\Omega^- \cup \Omega^+) \) are, respectively, lower and upper solutions for the problem (1.1) and \( \alpha(x) \leq \beta(x), \forall x \in \bar{\Omega}, \) then there exists a solution to (1.1) and

\[
\alpha(x) \leq u_\varepsilon(x) \leq \beta(x) \quad \forall x \in \bar{\Omega}.
\]

Hence, to establish existence we are only required to construct a lower and upper solution. First we place a restriction on the magnitude of the boundary conditions and \( \|f\| \).

**Assumption 1.** Assume that there exists a \( \theta > 0 \) such that

\[
\begin{align*}
b(y) \geq \theta &> 0 \quad \forall y \in D_{\theta,K} = [-\frac{K}{\theta}, \frac{K}{\theta}] \\
K &\equiv \max\{\|f\|, \theta|u_\varepsilon(0)|, \theta|u_\varepsilon(1)|\}.
\end{align*}
\]

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Note that since \( b(0) > 0 \) and \( b \) is smooth, then there exists a neighbourhood \([-\delta, \delta]\) such that \( b(y) \geq \theta > 0 \) \( \forall y \in [-\delta, \delta] \). Requiring that \( b(z) \geq \theta > 0 \) \( \forall z \in (-\infty, \infty) \) is considerably more stringent on the extent of the class of nonlinear problems under consideration.

**Theorem 2.2.** Problem (1.1), (2.2) has a solution \( u_\varepsilon \in C^1([0, 1], D_{\theta, K}) \) and
\[
\|u_\varepsilon\| \leq \frac{K}{\theta}.
\]

**Proof.** Let \( \alpha(x) = -\frac{1}{\theta}K = -\beta(x) \). Then \( \alpha(0) \leq u_\varepsilon(0) \leq \beta(0) \) and \( \alpha(1) \leq u_\varepsilon(1) \leq \beta(1) \), with \( \varepsilon \alpha'' = \varepsilon \beta'' = 0 \). Note also that, by virtue of (2.2),
\[
-\varepsilon \alpha''(x) + b(\alpha \alpha = b(-\frac{1}{\theta}K)(-\frac{1}{\theta}K) \leq f
\]
and
\[
-\varepsilon \beta''(x) + b(\beta) \beta \geq f.
\]
Hence, \( \alpha \) and \( \beta \) are lower and upper solutions with \( \alpha(x) \leq \beta(x) \) \( \forall x \in [0, 1] \). By the previous theorem there exists a solution to problem (1.1), (2.2) and\( \alpha(x) \leq u_\varepsilon(x) \leq \beta(x) \) \( \forall x \in [0, 1] \).

**Theorem 2.3.** Let \( \alpha, \beta \) be lower and upper solutions. Assume that
\[ y < z \quad \text{implies that } b(y)y < b(z)z \quad \forall y, z \in [-\max\{\|\alpha\|, \|\beta\|\}, \max\{\|\alpha\|, \|\beta\|\}] . \]
With this assumption,
\[ \alpha(x) \leq \beta(x) \quad \forall x \in [0, 1] . \]

**Proof.** Let \( p \) be any point at which \( \omega = \alpha - \beta \) attains its maximum value in \( \Omega \). Assume that \( \omega(p) > 0 \). If \( p \neq d \) and \( p \in \Omega^- \cup \Omega^+ \), then \( \omega''(p) \leq 0 \), and at this point \( x = p \),
\[
\varepsilon \alpha''(p) \geq b(\alpha \alpha - f > b(\beta) \beta - f \geq \varepsilon \beta''(p),
\]
which implies that \( \omega''(p) > 0 \), which is a contradiction. If \( p = d \), then the argument depends on whether or not \( \omega \) is differentiable at \( d \). If \( \omega'(d) \) does not exist, then \( |\omega'(d)| \neq 0 \) and because \( \omega'(d-) \geq 0 \), \( \omega'(d+) \leq 0 \), it is clear that \( |\omega'(d)| < 0 \). However, because \( \alpha \) and \( \beta \) are lower and upper solutions, we also have that \( |\tilde{\alpha}'(y)| \geq 0 \) and \( |\tilde{\beta}'(y)| \leq 0 \), which contradicts \( |\omega'(d)| < 0 \). If \( \omega'(d) \) does exist, then \( \omega'(d) = 0 \) and one can follow the argument as in the linear problem [7] to arrive again at a contradiction. Hence, the assumption that \( \omega(p) > 0 \) always leads to a contradiction.

This result and assumption (2.3) guarantee uniqueness of the solution of (1.1), (2.2). Let \( u_1, u_2 \) be two solutions of problem (1.1), (2.2). Then, by Theorem 2, we have that
\[
\|u_i\| \leq \frac{K}{\theta}, \quad i = 1, 2.
\]
Assuming (2.3), \( u_1, u_2 \) can be viewed as lower and upper solutions and so \( u_1 \leq u_2 \). Reversing the roles of \( u_1, u_2 \) provides uniqueness.

Follow the arguments in [9] to get
\[ u_\varepsilon|_{k, \Omega^- \cup \Omega^+} \leq C \frac{K}{\theta}(1 + \varepsilon^{-k/2}), \quad 0 \leq k \leq 4, \]
where the seminorms \( |\cdot|_{k, D} \) are defined by
\[
|y|_{k, \Omega^- \cup \Omega^+} = \left| \frac{d^k y}{dx^k} \right|_{\Omega^- \cup \Omega^+}.
\]
The reduced problem (i.e., set \( \varepsilon = 0 \) in (1.1)) is

\[
(2.5) \quad b(y_0)y_0 = f(x), \quad x \neq d.
\]

Note that if \( f \equiv 0 \) on the subinterval \( x < d \) (or \( x > d \)), then by (1.1), \( y_0 \equiv 0 \) is a solution to the reduced problem on the subinterval \( x < d \) (or \( x > d \)). By (2.2) we know that if \( \|f\| \neq 0 \), then

\[
(2.6) \quad b(-\frac{\|f\|}{\theta})(-\frac{\|f\|}{\theta}) - f \leq \frac{\|f\|^2}{\theta^2} - f.
\]

Hence in the interval \( D_{\theta,\|f\|} = [-\frac{\|f\|}{\theta}, \frac{\|f\|}{\theta}] \), for each \( x \neq d \), there exists an associated \( y \) such that

\[
b(y)y - f(x) = 0, \quad x \neq d.
\]

Thus, it is sufficient for existence of a reduced solution that

\[
(2.7) \quad b(y) \geq \theta > 0 \quad \forall y \in [-\frac{\|f\|}{\theta}, \frac{\|f\|}{\theta}].
\]

This is implied by (2.2), and hence a reduced solution \( y_0 \) exists, by (2.3) is unique within the interval \( D_{\theta,\|f\|} \subset D_{\theta,K} \), and

\[
(2.8) \quad b(y_0) \geq \theta > 0.
\]

We now impose a further condition on the strength of the nonlinearity. Assume that given \( \theta \) in (2.2) there exists a \( \gamma > 0 \) such that

\[
(2.9) \quad \frac{d}{dy}(b(y)y) \geq \gamma > 0 \quad \forall y \in D_{\theta,\|f\|} = [-\frac{\|f\|}{\theta}, \frac{\|f\|}{\theta}].
\]

Assuming (2.9) guarantees (via the implicit function theorem) that if \( f \in C^k(\Omega^- \cup \Omega^+) \), then a reduced solution \( y_0 \in C^k(\Omega^- \cup \Omega^+) \) exists and is unique, with \( \|y_0\| \leq \frac{\|f\|}{\theta} \). For \( u_\varepsilon \) to be unique, we can assume that

\[
(2.10) \quad \frac{d}{dy}(b(y)y) \geq \gamma > 0 \quad \forall y \in D_{\theta,K} = [-\frac{K}{\theta}, \frac{K}{\theta}].
\]

Note that (2.10) implies (2.3), which yields uniqueness of the solution \( u_\varepsilon \).

To establish the parameter-robust properties of the numerical methods involved in this paper, the following decomposition of \( u_\varepsilon \) into regular \( v_\varepsilon \) and singular \( w_\varepsilon \) components will be used. The regular component \( v_\varepsilon \) is defined as the solution of

\[
-\varepsilon v''_\varepsilon + b(v_\varepsilon)v_\varepsilon = f, \quad x \neq d,
\]

\[
b(v_\varepsilon)v_\varepsilon = f, \quad x \neq d,
\]

\[
v_\varepsilon(0) = v_\varepsilon(0), \quad v_\varepsilon(d-) = v_\varepsilon(d-),
\]

\[
v_\varepsilon(d+) = v_\varepsilon(d+), \quad v_\varepsilon(1) = v_\varepsilon(1).
\]

Note that \( \|v_\varepsilon\| \leq \frac{\|f\|}{\theta} \), which implies that \( |v_\varepsilon(0)|, |v_\varepsilon(d-)| \leq \|f\|/\theta \). Hence, using the arguments in Theorem (2.2) on \( \Omega^- \), \( \Omega^+ \) separately, we deduce that \( v_\varepsilon \) exists and

\[
\|v_\varepsilon\| \leq \frac{\|f\|}{\theta}.
\]

The singular component \( w_\varepsilon \) is given implicitly by \( u_\varepsilon = v_\varepsilon + w_\varepsilon \), where \( u_\varepsilon \) is the solution of the problem (1.1). Since the solution of (1.1), (2.2), (2.10) is unique

\[
\|w_\varepsilon\| = \|u_\varepsilon - v_\varepsilon\| \leq \frac{\|f\| + K}{\theta}.
\]
In order to derive sharp pointwise bounds on the singular component \(w_\varepsilon\), we are required to strengthen the restriction given in (2.10) to the following assumption.

**Assumption 2.** Given \(\theta\) in (2.2), assume that there exists a \(\gamma > 0\) such that

\[
(2.11a) \quad \frac{d}{dy}(b(y)y) \geq \gamma > 0 \quad \forall y \in D_{\theta,2\|f\|+K} = \left[ -\frac{2\|f\| + K}{\theta}, \frac{2\|f\| + K}{\theta} \right],
\]

where

\[
(2.11b) \quad K = \max\{\|f\|, \theta|u_\varepsilon(0)|, \theta|u_\varepsilon(1)|\}.
\]

**Theorem 2.4.** Let \(u_\varepsilon\) be the solution of the problem (2.10), (2.2), (2.11). Then \(u_\varepsilon = v_\varepsilon + w_\varepsilon\), and for each integer \(j\) satisfying \(0 \leq j \leq 4\), the components \(v_\varepsilon\) and \(w_\varepsilon\) satisfy the following bounds for \(\varepsilon\) sufficiently small:

\[
|v_\varepsilon(x)|_j \leq \begin{cases} 
  C(1 + \varepsilon^{1-j}), & x \in \Omega^-, \\
  C(1 + \varepsilon^{1-j}), & x \in \Omega^+,
\end{cases}
\]

\[
|w_\varepsilon(x)|_j \leq \begin{cases} 
  C(e^{-\frac{\varepsilon}{2}(e^{-x}\sqrt{\gamma} + e^{-(d-x)}\sqrt{\gamma})}), & x \in \Omega^-, \\
  C(e^{-\frac{\varepsilon}{2}(e^{-d}\sqrt{\gamma} + e^{-(1-x)}\sqrt{\gamma})}), & x \in \Omega^+,
\end{cases}
\]

where \(C\) is a constant independent of \(\varepsilon\) and \(|\cdot|_j\) denotes the maximum pointwise norm of the \(j\)th derivative.

**Proof.** Note that \(|v_0|_{j,\Omega-\cup\Omega^+} \leq C\). Introduce the notation \(g(y) = (b(y)y)y\) and, by assumption (2.11), \(g(y) \geq \gamma > 0 \quad \forall y \in D_{\theta,2\|f\|}\). We have that, for some \(t \in [0,1]\),

\[-\varepsilon(v_\varepsilon - v_0)'' + (b(v_\varepsilon)v_\varepsilon - b(v_0)v_0) = -\varepsilon(v_\varepsilon - v_0)'' + g(v_0 + t(v_\varepsilon - v_0))(v_\varepsilon - v_0) = \varepsilon v_0'',
\]

\[(v_\varepsilon - v_0)(0) = (v_\varepsilon - v_0)(d-) = (v_\varepsilon - v_0)(d+) = (v_\varepsilon - v_0)(1) = 0.
\]

Note that both \(v_\varepsilon, v_0 \in D_{\theta,\|f\|}\) and hence

\[g(v_0 + t(v_\varepsilon - v_0)) \geq \gamma > 0.
\]

Consider the nonlinear problem

\[-\varepsilon y'' + g(v_0 + ty)y = \varepsilon y_0'', \quad x \in \Omega^-,
\]

\[y(0) = A_1, \quad y(d) = B_1.
\]

Use \(\alpha = -\varepsilon\|v_0''\|/\gamma = -\beta\) as lower and upper solutions. For \(\varepsilon\) sufficiently small,

\[v_0 \pm \varepsilon t\|v_0''\|/\gamma \in D_{\theta,2\|f\|}, \quad 0 < t < 1,
\]

and, hence, \(g(v_0 + t\alpha) \geq \gamma\) and \(g(v_0 + t\beta) \geq \gamma\). Thus, we have that

\[\|v_\varepsilon - v_0\|_{\Omega^-\cup\Omega^+} \leq C\varepsilon.
\]

Then, follow the argument in [9] to get the bounds

\[|v_\varepsilon - v_0|_{j,\Omega^-\cup\Omega^+} \leq C\varepsilon(1 + \varepsilon^{-j/2}), \quad 1 \leq j \leq 4.
\]

The singular component is the solution of

\[-\varepsilon w_\varepsilon'' + b(u_\varepsilon)(v_\varepsilon + w_\varepsilon) = b(v_\varepsilon)v_\varepsilon, \quad x \neq d,
\]

\[w_\varepsilon = u_\varepsilon - v_\varepsilon, \quad x = 0,1,
\]

\[|w_\varepsilon|(d) = -|v_\varepsilon|(d), \quad |w_\varepsilon''|(d) = -|v_\varepsilon''|(d).
\]
Note that
\[-\varepsilon w'' + (b(u_\varepsilon)v_\varepsilon - b(v_\varepsilon)u_\varepsilon) = -\varepsilon w'' + g(u_\varepsilon + t(v_\varepsilon - u_\varepsilon))w_\varepsilon = 0.\]
We have that \(u_\varepsilon \in D_{\theta,K}, v_\varepsilon \in D_{\theta,\|f\|}\) and hence
\[g(u_\varepsilon + t(v_\varepsilon - u_\varepsilon)) = g(v_\varepsilon + (1-t)w_\varepsilon) \geq \gamma > 0\]
from (2.11).

On the interval \(\Omega^-\), consider the nonlinear problem
\[-\varepsilon y'' + g(v_\varepsilon + (1-t)y)y = 0, \quad x \in \Omega^{-},\]
\[y(0) = A_1, \quad y(d) = B_1.\]

Consider the barrier function
\[\phi(x) = \max\{|w_\varepsilon(0)|, |w_\varepsilon(d^-)|\} e^{-\gamma x/\varepsilon} + e^{-(d-x)\gamma/\varepsilon} \quad \frac{1}{1 + e^{-d\gamma/\varepsilon}}.\]

Note that \(\varepsilon \phi'' = 0\), and \(\|\phi\| \leq \frac{\|f\|+K}{\gamma} \). Hence, \(v_\varepsilon \pm (1-t)\phi \in D_{\theta,2\|f\|+K}\) and, by (2.11), this gives
\[g(v_\varepsilon \pm (1-t)\phi) \geq \gamma.\]

Let \(\alpha = -\phi = \beta\). Then
\[-\varepsilon \alpha'' + g(v_\varepsilon + (1-t)\alpha)\alpha \leq -\varepsilon \beta'' + g(v_\varepsilon + (1-t)\beta)\beta.\]

Then \(-\phi \leq w_\varepsilon \leq \phi\) in the interval \(\Omega^-.\) Follow the arguments in [9] to get bounds on the derivatives of \(w_\varepsilon\) in the interval \(\Omega^-\). The result for the interval \(\Omega^+\) can be obtained in a similar manner.

3. **Discrete Problem**

On \(\Omega^+ \cup \Omega^-\) a piecewise-uniform mesh of \(N\) mesh intervals, where \(N\) is a multiple of 8, is constructed as follows. The interval \(\Omega^-\) is subdivided into the three subintervals
\[\{0, \sigma_1\}, \quad \{\sigma_1, d - \sigma_1\}, \quad \{d - \sigma_1, d\}\]
for some \(\sigma_1\) that satisfies \(0 < \sigma_1 \leq \frac{d}{2}\). On \(\{0, \sigma_1\}\) and \(\{d - \sigma_1, d\}\) a uniform mesh with \(\frac{N}{3}\) mesh intervals is placed, while \(\{\sigma_1, d - \sigma_1\}\) has a uniform mesh with \(\frac{N}{4}\) mesh intervals. The subintervals \([d, d + \sigma_2], [d + \sigma_2, 1 - \sigma_2], [1 - \sigma_2, 1]\) are treated analogously for some \(\sigma_2\) satisfying \(0 < \sigma_2 \leq \frac{1-d}{4}\). The interior points of the mesh are denoted by
\[\Omega_e^N = \{x_i : 1 \leq i \leq \frac{N}{2} - 1\} \cup \{x_i : \frac{N}{2} + 1 \leq i \leq N - 1\}.\]

Clearly \(x_{\frac{N}{2}} = d\) and \(\Omega_e^N = \{x_i\}_{0}^{N}\). Note that, for the case \(d = 1/2\), this piecewise-uniform mesh is a uniform mesh when \(\sigma_1 = \frac{d}{4}\) and \(\sigma_2 = \frac{1-d}{4}\). It is fitted to the singular perturbation problem (1.11) by choosing \(\sigma_1\) and \(\sigma_2\) to be the functions
\[(3.1) \quad \sigma_1 = \min \left\{ \frac{d}{4}, M \sqrt{\varepsilon} \ln N \right\}, \quad \sigma_2 = \min \left\{ \frac{1-d}{4}, M \sqrt{\varepsilon} \ln N \right\}, \quad M \geq \frac{1}{\sqrt{\gamma}},\]
of \(N\) and \(\varepsilon\), where \(\gamma\) is specified in (2.11). On the piecewise-uniform mesh \(\Omega_e^N\), a standard centred finite difference operator is used. Then the fitted mesh method for problem (1.1) follows.
Find a mesh function $U_\varepsilon$ such that
\begin{align}
\tag{3.2a} -\varepsilon \delta^2 U_\varepsilon(x_i) + b(U_\varepsilon(x_i)) U_\varepsilon(x_i) &= f(x_i) \quad \text{for all } x_i \in \Omega^N,
\tag{3.2b} U_\varepsilon(0) &= u_\varepsilon(0), \quad U_\varepsilon(1) = u_\varepsilon(1),
\tag{3.2c} D^- U_\varepsilon(x_{\frac{\varepsilon}{2}}) = D^+ U_\varepsilon(x_{\frac{\varepsilon}{2}}),
\end{align}
where
\begin{equation}
\delta^2 Z_i = \left( \frac{Z_{i+1} - Z_i}{x_{i+1} - x_i} - \frac{Z_i - Z_{i-1}}{x_i - x_{i-1}} \right) \frac{1}{x_{i+1} - x_{i-1}}.
\end{equation}

Let $G : R^{N+1} \to R^{N+1}$ be the mapping associated with this finite difference scheme. For mesh function $Y$ we have an associated vector $Y \in R^{N+1}$, where $Y_i = Y(x_i)$. Let
\begin{equation}
(GY)_i = \begin{cases} Y(0), & i = 0, \\
-\varepsilon \delta^2 Y_i + b(Y_i) Y_i, & i \neq N/2, 1 \leq i \leq N, \\
-\varepsilon \delta^2 Y_i, & i = N/2, \\
Y(1), & i = N + 1.
\end{cases}
\end{equation}

We also define a vector $F$ by
\begin{equation}
F_i = \begin{cases} A, 0, B, & i = 0, N/2, N + 1, \\
f(x_i), & \text{otherwise}.
\end{cases}
\end{equation}

The finite difference scheme (3.2a) can then be written in the form $GU_\varepsilon = F$.

**Definition 2.** Given any vector $H \in R^{N+1}$, a lower mesh solution $V$ for the problem $GW = H$ is a mesh function which satisfies $GV \leq H$.

There is an analogous definition for an upper mesh solution to $GW = H$.

**Theorem 3.1.** If $\Phi, \Psi$ are lower and upper mesh solutions, respectively, for the problem $GW = H$ with $\Phi(x_i) \leq \Psi(x_i)$ $\forall x_i \in \Omega^N$, then there exists a solution to $GW = H$ such that
\begin{equation}
\Phi(x_i) \leq W(x_i) \leq \Psi(x_i) \quad \forall x_i \in \Omega^N.
\end{equation}

**Proof.** We follow the argument from Lorentz. Let $\Phi_1, \Phi_2$ be two lower mesh functions. Define the mesh function $\Phi_3$ by $\Phi_3(x_i) = \max\{\Phi_1(x_i), \Phi_2(x_i)\}$. At some point $x_j$, we assume without loss of generality that $\Phi_3(x_j) = \Phi_1(x_j)$. Note that $-\Phi_3(x_i) \leq -\Phi_1(x_i) \forall x_i$, and
\begin{align}
-\varepsilon \delta^2 \Phi_3(x_j) + b(\Phi_3) \Phi_1(x_j) &\leq -\varepsilon \delta^2 \Phi_1(x_j) + b(\Phi_1) \Phi_1(x_j) \leq H(x_j), \quad x_j \neq d, \\
-\varepsilon \delta^2 \Phi_3(x_j) &\leq -\varepsilon \delta^2 \Phi_1(x_j) \leq H(d), \quad x_j = d, \\
\Phi_3(0) &\leq H(0), \quad \Phi_3(1) \leq H(1).
\end{align}

Then $\Phi_3$ is also a lower mesh solution. Let $L = \{\phi : G\phi \leq H, \Phi \leq \phi \leq \Psi\}$. Define $U(x_i) = \sup_{\phi \in L} \phi(x_i)$. First note that $U \in L$ exists and $GU \leq H$. Assume that $\Phi_3$ satisfy the lower mesh solution. Let $\Phi_3(x_i) = \sup_{\phi \in L} \phi(x_i)$. Then $\Phi_3$ is a lower mesh solution. Let $V(x_i) = \sup_{\phi \in L} \phi(x_i)$. Then $V$ is a lower mesh solution. Then $U \leq V$.
Theorem 3.2. Let $\Phi, \Psi$ be lower and upper mesh solutions of $GW = H$ and $M_2 = \max\{\|\Phi\|, \|\Psi\|\}$. Assume that (2.3) holds over the interval $[-M_2, M_2]$, then

$$\Phi(x_i) \leq \Psi(x_i) \quad \forall x_i \in \Omega^N.$$ 

Proof. Let $x_j$ be that mesh point at which $\Phi - \Psi$ attains its maximum value in $\Omega^N$. Assume that $\Phi(x_j) > \Psi(x_j)$. If $x_j \neq d$, then since $\Phi, \Psi$ are lower and upper mesh solutions,

$$\varepsilon \delta^2 \Phi(x_j) \geq b(\Phi) - H > b(\Psi)\Psi - H \geq \varepsilon \delta^2 \Psi(x_j),$$

which contradicts $\delta^2 (\Phi - \Psi)(x_j) \leq 0$, which would occur if $\Phi - \Psi$ had its maximum at $x_j$. If $x_j = d$, then since $\Phi, \Psi$ are lower and upper mesh solutions, $\delta^2 (\Phi - \Psi)(d) \geq 0$. To avoid a contradiction,

$$(\Phi - \Psi)(d) = (\Phi - \Psi)(x_{N/2-1}) = (\Phi - \Psi)(x_{N/2+1}),$$

and apply the first part of the argument to $()$.

Corollary 3.3. Assuming (2.2) and (2.11), there exists a unique solution $U_\varepsilon$ to the problem (2.2) and

$$\|U_\varepsilon\| \leq \frac{K}{\theta}.$$ 

Proof. Follow an analogous argument to that used in the proof of Theorem 2.2.

4. Error analysis

We begin by looking at the truncation error. By classical estimates, for all $x_i \in \Omega^N \cap \Omega^-$,

$$| - \varepsilon (\frac{d^2}{dx^2} - \delta^2)w_\varepsilon(x_i)| \leq \frac{\varepsilon}{3}(x_{i+1} - x_{i-1})|w_\varepsilon| \leq C\sqrt{\varepsilon}N^{-1},$$

and from [8] we have

$$| - \varepsilon (\frac{d^2}{dx^2} - \delta^2)w_\varepsilon(x_i)| \leq \begin{cases} \varepsilon(x_{i+1} - x_{i-1})|w_\varepsilon|, & \text{(a)} \\ 2\varepsilon \max_{x \in [x_{i-1}, x_{i+1}]} |w_\varepsilon''(x)|, & \text{(b)} \end{cases}$$

Using (b) outside the layers and at $x = \sigma_1$, $x = d - \sigma_1$ gives

$$| - \varepsilon (\frac{d^2}{dx^2} - \delta^2)w_\varepsilon(x_i)| \leq \varepsilon C\varepsilon^{-1} \max_{x \in [x_{i-1}, x_{i+1}]} (e^{-\sqrt{\gamma/\varepsilon}} + e^{-(d-x)\sqrt{\gamma/\varepsilon}}) \leq CN^{-1}.$$ 

Using (a) inside the layers gives, as above for $v_\varepsilon$,

$$| - \varepsilon (\frac{d^2}{dx^2} - \delta^2)w_\varepsilon(x_i)| \leq C\varepsilon \frac{\sigma_1}{N} \varepsilon^{-1} \leq CN^{-1} \ln N.$$ 

Similar bounds on the truncation error are valid for all $x_i \in \Omega^N \cap \Omega^+$. Hence,

$$\tag{4.1a} | - \varepsilon (\frac{d^2}{dx^2} - \delta^2)u_\varepsilon(x_i)| \leq CN^{-1} \ln N, \quad x_i \neq d.$$ 

At the point $x_i = d$,

$$(D^+ - D^-)(U_\varepsilon - u_\varepsilon)(d) = -(D^+ - D^-)u_\varepsilon(d).$$
Let \( h^\pm \) be the mesh interval sizes on either side of the point \( x = d \) and \( h = \max\{h^-, h^+\} \). Then
\[
|(D^+ - D^-)u_\varepsilon(x)| \leq |(D^+ - \frac{d}{dx})u_\varepsilon(x)| + |(D^- - \frac{d}{dx})u_\varepsilon(x)|
\leq \frac{1}{2} h^+|u_\varepsilon|_2 + \frac{1}{2} h^-|u_\varepsilon|_2.
\]
Thus,
\[ (4.1b) \quad |(D^+ - D^-)(U_\varepsilon - u_\varepsilon)(d)| \leq \frac{Ch}{\varepsilon}. \]
We are now ready to bound the nodal error \( |(u_\varepsilon - U_\varepsilon)(x_i)| \).

**Theorem 4.1.** Let \( u_\varepsilon \) be the solution of problem (1.1), (2.2), (2.11) and \( U_\varepsilon \) the solution of (4.2a). Then, for \( \varepsilon \) sufficiently small,
\[
\max_{x_i \in \Omega^\varepsilon} |U_\varepsilon(x_i) - u_\varepsilon(x_i)| \leq CN^{-1} \ln N,
\]
where \( C \) is a constant independent of \( \varepsilon \) and \( N \).

**Proof.** At the internal mesh points,
\[
-\varepsilon \delta^2 U_\varepsilon(x_i) + b(U_\varepsilon(x_i))U_\varepsilon(x_i) = (-\varepsilon u_\varepsilon'' + b(u_\varepsilon)u_\varepsilon)(x_i),
\]
\[
-b(U_\varepsilon - u_\varepsilon)(x_i) = b(u_\varepsilon)u_\varepsilon - (\varepsilon u_\varepsilon'' + \varepsilon \delta^2 u_\varepsilon)(x_i).
\]
Note that by (2.11),
\[
b(U_\varepsilon(x_i))U_\varepsilon(x_i) - b(u_\varepsilon)u_\varepsilon(x_i) = g(Z)(U_\varepsilon(x_i) - u_\varepsilon(x_i)),
\]
where, since \( u_\varepsilon, U_\varepsilon \in D_{\theta,K}, \)
\[
Z = u_\varepsilon + t_i(U_\varepsilon - u_\varepsilon) \in D_{\theta,K},
\]
and
\[
g(Z) \geq \gamma > 0 \quad 0 < t_i < 1.
\]
Define the linear operator \( L^N_{U_\varepsilon} \) as follows.

For any mesh function \( V \)
\[
L^N_{U_\varepsilon} V(x_i) = -\varepsilon \delta^2 V(x_i) + g(u_\varepsilon(x_i)) + t_i(U - u)V(x_i), \quad x_i \neq d,
\]
\[
L^N_{U_\varepsilon} V(d) = D^- V(d) - D^+ V(d).
\]
From [4] we have the following discrete comparison principle. *If \( V \) is a mesh function such that \( V(0) \geq 0, V(1) \geq 0, L^N_{U_\varepsilon} V \geq 0, x_i \in \Omega^N, \) and \( D^+ V(d) - D^- V(d) \leq 0, \) then \( V(x_i) \geq 0 \) for all \( x_i \in \Omega^N. \) By the truncation error estimates (4.1),
\[
|L^N_{U_\varepsilon}(U - u)| \leq CN^{-1} \ln N, \quad x_i \neq d,
\]
and at the mesh point \( x_i = d \)
\[
|L^N_{U_\varepsilon}(u - U)(d)| \leq C \frac{h}{\varepsilon}.
\]
Consider the mesh function
\[
\Xi(x_i) = C_1 N^{-1} \ln N + C_2 \frac{h}{\varepsilon} \phi_d(x_i) \pm U - u,
\]
where \( C_1 \) and \( C_2 \) are suitably large constants and \( \Phi_d \) is defined as the solution of

\[
-\varepsilon^2 \Phi_d(x_i) + \gamma \Phi_d(x_i) = 0 \quad \text{for all } x_i \in \Omega^N,
\]

\[
\Phi_d(0) = 0, \quad \Phi_d(d) = 1, \quad \Phi_d(1) = 0.
\]

On the interval \([0, d]\) consider the barrier function

\[
\omega(x_j) = \frac{\prod_{i=1}^j (1 + \sqrt[2\varepsilon]{h_i}/\sqrt{2\varepsilon})}{\prod_{i=1}^j (1 + \sqrt[2\varepsilon]{h_i}/\sqrt{2\varepsilon})}, \quad \omega(0) = 0, \quad \omega(1) = 1,
\]

where \( h_i = x_i - x_{i-1} \). Note that

\[
D^+ \omega(x_i) = \frac{\sqrt{\gamma}}{2\varepsilon} \omega(x_i), \quad D^- \omega(x_i) = \frac{\sqrt{\gamma}}{2\varepsilon(1 + \sqrt[2\varepsilon]{h_i}/\sqrt{2\varepsilon})} \omega(x_i),
\]

which implies that

\[-\varepsilon^2 \omega(x_i) + \gamma \omega(x_i) < 0, \quad 0 < x_i < d.\]

Hence, \( \Phi_d(x_i) \leq \omega(x_i) \) and then

\[
D^- \Phi_d(d) = \frac{1 - \Phi_d(d - h)}{h} \geq \frac{1 - \omega(d - h)}{h} = \frac{\sqrt{\gamma}/\sqrt{2\varepsilon}}{1 + \sqrt[2\varepsilon]{h}/\sqrt{2\varepsilon}} \geq \frac{C}{\sqrt[2\varepsilon]}.
\]

From this and using an analogous argument on the interval \([d, 1]\), we have that

\[
\sqrt{\varepsilon}(D^+ \Phi_d(d) - D^- \Phi_d(d)) \leq -C_2.\]

We conclude that

\[
\|u_\varepsilon - u_{\varepsilon}\| \leq C N^{-1} \ln N.
\]

5. Numerical results

In this section we present numerical results, which validate the theoretical results established in the previous section. In order to solve the nonlinear difference scheme we use a variant of the continuation method from [5] §10.3.

\[\begin{align}
(5.1a) \quad & -\varepsilon^2 \Phi_d(x_i) + \gamma \Phi_d(x_i) = 0, \quad x_i \neq d, \quad j = 1, \ldots, K, \\
(5.1b) \quad & D_j^+ U_\varepsilon(x_i, t_j) = D_j^- U_\varepsilon(x_i, t_j), \quad j = 1, \ldots, K, \\
(5.1c) \quad & U_\varepsilon(0, t_j) = u_\varepsilon(0), \quad U_\varepsilon(1, t_j) = u_\varepsilon(1), \quad \text{for all } j, \\
(5.1d) \quad & U_\varepsilon(x, 0) = u_{\text{init}}(x).
\end{align}\]

In all cases in this paper the initial guess for the nonlinear solver is taken to be \( u(0) + (u(1) - u(0)) x \). We can interpret (5.1) as a discretization of the following time-dependent version of the problem

\[\begin{align}
(5.2a) \quad & \text{Find } \quad u \in C^1([0, 1] \times [0, T]) \quad \text{such that} \\
(5.2b) \quad & -\varepsilon u_{xx} + b(u(x, t)) u + u_t = f(x), \quad (x, t) \in (0, 1) \setminus \{d\} \times (0, T), \\
(5.2c) \quad & u(0, t) = u_\varepsilon(0), \quad u(1, t) = u_\varepsilon(1), \quad t \geq 0, \\
(5.2d) \quad & u(x, 0) = u_{\text{init}}(x), \quad 0 < x < 1.
\end{align}\]

The choices of the uniform time-like step \( k = t_j - t_{j-1} \) and the number of iterations \( K \) are determined as follows. Defining

\[\begin{align}
(5.3a) \quad & e(j) = \max_{1 \leq i \leq N} |U_\varepsilon(x_i, t_j) - U_\varepsilon(x_i, t_{j-1})|/k, \quad \text{for } j = 1, 2, \ldots, K, \\
(5.3b) \quad & e(j) \leq e(j - 1), \quad \text{for all } j \text{ satisfying } 1 < j \leq K.
\end{align}\]
Then the number of iterations $K$ is chosen such that
\[ e(K) \leq \text{TOL}, \]
where TOL is a suitably prescribed small tolerance. In the case of this paper the tolerance TOL is chosen to be $10^{-7}$. The numerical solution is computed using the following algorithm. Start from $t_0$ with the initial time step $k = 1.0$. If, at some value of $j$, \((5.3b)\) is not satisfied, then discard the time step from $t_{j-1}$ to $t_j$ and restart from $t_{j-1}$ with half the time step, that is $k_{\text{new}} = k/2$, and continue halving the time step until one finds a $k$ for which \((5.3b)\) is satisfied. Assuming that \((5.3b)\) is satisfied at each time step, continue until either \((5.3c)\) is satisfied, \(t_j = 1000\), or the $t_j - t_{j-1} < 0.000001$. If \((5.3c)\) is not satisfied, we assume that the time stepping process stalled due to a too large choice of the initial time step. In this case we repeat the entire process again from $t_0$, halving the initial time step again. If \((5.3c)\) is satisfied, the resulting values of $U_\varepsilon(x, K)$ are taken as the approximations to the solution of the continuous problem. Numerical results are presented for the problem
\[
\varepsilon u''_{\varepsilon}(x) - (1 - u^2_{\varepsilon})u_{\varepsilon}(x) = f(x),
\]
\[
f(x) = \begin{cases} 
\delta_1 + x(0.5 - x), & x < 0.5, \\
-\delta_2 + (x - 0.5)(x - 1), & x > 0.5,
\end{cases}
\]
\[
u_{\varepsilon}(0) = A, \quad u_{\varepsilon}(1) = B.
\]
Let us examine in more detail the effects of the various constraints, such as those necessary for existence, in this particular case. Note first that for the problem \((5.4)\)
\[
b(y) = 1 - y^2 \geq \theta > 0 \quad \text{for} \quad |y| \leq \sqrt{1 - \theta}.
\]
In this case the restriction \((2.4)\) on $\|f\|$ sufficient for existence of the reduced solution is
\[
\|f\| \leq \theta \sqrt{1 - \theta}.
\]
The range of $f$ allowed by this constraint is maximized when $\theta = 2/3$, in which case it becomes
\[
\|f\| \leq \frac{2}{3\sqrt{3}} \approx 0.3849.
\]
We remark that in order to guarantee that the solution $u_{\varepsilon}$ exists, for all $\varepsilon$, we require in addition that \((2.4)\) be satisfied; that is,
\[
|A|, |B| \leq \sqrt{1 - \theta},
\]
which for this choice of $\theta = 2/3$ gives
\[
|A|, |B| \leq \frac{1}{\sqrt{3}} \approx 0.57735026919.
\]
However, the restriction \((2.11)\) when $K = \|f\|$, required to prove convergence of the numerical method, imposes the additional condition
\[
\|f\| \leq \frac{\theta \sqrt{1 - \gamma}}{3\sqrt{3}}, \quad 0 < \gamma < 1.
\]
Figure 1. Solutions of (5.4) with (a) \( A = B = 0 \) and \( \delta_1 = -0.1, \delta_2 = 0.15 \) and (b) \( -A = B = 0.57735 \) and \( \delta_1 = -0.3849001794597, \delta_2 = 0.15 \) for \( \varepsilon = 10^{-2}, 10^{-4}, 10^{-6}, 10^{-9} \).

Note that the parameter \( M \) in the definition of the transition points of the mesh is bounded below by

\[
M \geq \frac{1}{\sqrt{\gamma}}.
\]

There is a trade-off between two competing constraints on \( \gamma \). To allow maximum flexibility on choice of the mesh, we would like \( \gamma \) to be as large as possible. However, to maximize the range of \( f \) values, it is better to choose \( \gamma \) smaller. To maximize the acceptable range of \( f \), while keeping \( \gamma \) as large as possible, we choose to make (5.5a) and (5.5c) equal. Requiring this, we obtain

\[
\theta \sqrt{1 - \theta} = \frac{\theta \sqrt{1 - \gamma}}{3\sqrt{3}},
\]

which implies, using the fact that \( \gamma > 0 \), that

\[
\theta > \frac{26}{27} \approx 0.962963.
\]

Assume \( \theta = \frac{26}{27} \), then the restrictions (2.2), (2.11) on the data for the particular problem (5.4), (that is (5.5a), (5.5b), and (5.5c)) become

(5.6a) \( \|f\| < \frac{26}{81\sqrt{3}} \approx 0.1853 \),

(5.6b) \( \max\{|u(0)|, |u(1)|\} \leq \frac{1}{3\sqrt{3}} \approx 0.19245 \).
With these restrictions the parameter in the transition point for the mesh is given by

\[(5.7)\quad M \geq \frac{1}{\sqrt{\gamma}},\quad \text{where} \quad \gamma = 1 - \frac{27^2}{26^2}||f||^2.\]

Figure 1 shows the solution of problem (5.1) using method (5.1) on a mesh with \(M = 2.5\) and \(N = 1024\). The solution is shown for homogeneous boundary conditions \(A = B = 0\) and \(\delta_1 = -0.1, \delta_2 = 0.15\) on the left, and for \(A = -0.57735, B = 0.57735\) and \(\delta_1 = -0.3849001794597, \delta_2 = 0.15\) on the right. The left-most plot satisfies conditions (5.6). The data for the right-most plot are close to the extremes of the condition (5.7a) for the existence of the solution \(u_\varepsilon\).

Tables 1, 2, and 3 present numerical results generated using the scheme (3.2c), where

\[\sigma_1 = \min \left\{ \frac{1}{8}, M \sqrt{\varepsilon} \ln N \right\}, \quad \sigma_2 = \min \left\{ \frac{1}{8}, M \sqrt{\varepsilon} \ln N \right\}, \quad \text{and} \quad M = 1.75,\]

to solve problem (5.1) with \(A = -0.14, B = 0.12, \delta_1 = -0.1, \delta_2 = 0.15\). Table 1 gives the errors \(E^N\) and the uniform errors \(E^N\) with respect to the finest mesh. Table 2 displays the computed rates of convergence \(p^N\) with respect to the finest mesh and the uniform rates of convergence \(p^N\). Table 3 presents the number of iterations taken by the nonlinear solver. See [6] for details on how the quantities \(E^N, \sigma^N, p^N,\) and \(p^N\) are calculated. Note that in this case the condition on \(\gamma (5.7)\)

\[\gamma = 1 - \frac{27^2}{26^2}||f||^2 \approx 0.344871\]

and

\[M \geq \frac{1}{\sqrt{\gamma}} \approx 1.7028.\]

**Table 1.** Computed maximum pointwise errors with respect to the finest mesh for \(A = -0.14, B = 0.12\) and \(\delta_1 = -0.1, \delta_2 = 0.15\) and \(M = 1.75.\)

<table>
<thead>
<tr>
<th>(\varepsilon)</th>
<th>2^{-8}</th>
<th>2^{-7}</th>
<th>2^{-6}</th>
<th>2^{-5}</th>
<th>2^{-4}</th>
<th>2^{-3}</th>
<th>2^{-2}</th>
<th>2^{-1}</th>
<th>2^{0}</th>
<th>2^{1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E^N)</td>
<td>0.00141</td>
<td>0.00079</td>
<td>0.00039</td>
<td>0.00019</td>
<td>0.00009</td>
<td>0.00001</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td>(\varepsilon)</td>
<td>32</td>
<td>64</td>
<td>128</td>
<td>256</td>
<td>512</td>
<td>1024</td>
<td>2048</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Table 2. Computed rates of convergence for \( A = -0.14, B = 0.12 \) and \( \delta_1 = -0.1, \delta_2 = 0.15 \) and \( M = 1.75 \).

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>( 32 )</th>
<th>( 64 )</th>
<th>( 128 )</th>
<th>( 256 )</th>
<th>( 512 )</th>
<th>( 1024 )</th>
<th>( 2048 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^{-1} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-2} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-3} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-4} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-7} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-8} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
<tr>
<td>( 2^{-9} )</td>
<td>1.14</td>
<td>1.03</td>
<td>1.10</td>
<td>1.22</td>
<td>1.15</td>
<td>0.98</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Table 3. Number of iterations for \( A = -0.14, B = 0.12 \) and \( \delta_1 = -0.1, \delta_2 = 0.15 \) and \( M = 1.75 \).

<table>
<thead>
<tr>
<th>( \varepsilon )</th>
<th>( 32 )</th>
<th>( 64 )</th>
<th>( 128 )</th>
<th>( 256 )</th>
<th>( 512 )</th>
<th>( 1024 )</th>
<th>( 2048 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2^{-1} )</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>( 2^{-2} )</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( 2^{-3} )</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( 2^{-4} )</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>( 2^{-7} )</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>( 2^{-8} )</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>( 2^{-9} )</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>( 2^{-10} )</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>( 2^{-11} )</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>( 2^{-12} )</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

To determine the sensitivity of the rates of convergence to the choice of the value \( M \), in Table 4 we present a summary of the maximum errors over all values of \( \varepsilon \) above with respect to the finest mesh, and the double mesh rates of convergence for various values of \( M \) for the same problem. In all cases the number of iterations taken was the same as in Table 3.

It is of some interest to the computational scientist to see if the method still converges in the region where there is no formal theoretical proof. This is, in fact, the case. Although the theory guarantees existence and uniqueness of the solution and convergence of the numerical method if (5.6) holds, in practice, the numerical
method converges for a wider range of choices of $f$, $u(0)$, and $u(1)$. Tables 5 and 6 give examples of problems which just satisfy condition (2.2) but not (5.6a) and/or (5.6b), and hence (2.11), which was required for the proof of the convergence of the numerical solutions. Table 5 is for the problem (5.4) with homogeneous boundary conditions $A = B = 0$ and $\delta_1 = -0.3849$, $\delta_2 = 0.15$, and thus the condition (5.6a) holds.

Table 5. Maximum errors $E_N$ and computed rates of convergence $p_N$ for $A = B = 0$ and $\delta_1 = -0.3849$, $\delta_2 = 0.15$ for several values of $M$.

<table>
<thead>
<tr>
<th>$M$</th>
<th>$32$</th>
<th>$64$</th>
<th>$128$</th>
<th>$256$</th>
<th>$512$</th>
<th>$1024$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_N$</td>
<td>0.180983</td>
<td>0.144543</td>
<td>0.110590</td>
<td>0.089335</td>
<td>0.065783</td>
<td>0.028762</td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.34</td>
<td>0.38</td>
<td>0.40</td>
<td>0.41</td>
<td>0.44</td>
<td>1.43</td>
</tr>
<tr>
<td>$E_N$</td>
<td>0.098538</td>
<td>0.061915</td>
<td>0.041651</td>
<td>0.028841</td>
<td>0.016853</td>
<td>0.007813</td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.47</td>
<td>0.51</td>
<td>0.60</td>
<td>0.78</td>
<td>1.11</td>
<td>1.74</td>
</tr>
<tr>
<td>$E_N$</td>
<td>0.045875</td>
<td>0.020415</td>
<td>0.012320</td>
<td>0.006593</td>
<td>0.002730</td>
<td></td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.55</td>
<td>0.61</td>
<td>0.71</td>
<td>0.90</td>
<td>1.27</td>
<td>1.91</td>
</tr>
<tr>
<td>$E_N$</td>
<td>0.020005</td>
<td>0.010547</td>
<td>0.005849</td>
<td>0.002621</td>
<td>0.001054</td>
<td></td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.69</td>
<td>0.74</td>
<td>0.84</td>
<td>1.04</td>
<td>1.41</td>
<td>2.05</td>
</tr>
<tr>
<td>$E_N$</td>
<td>0.02249</td>
<td>0.01350</td>
<td>0.00731</td>
<td>0.00327</td>
<td>0.00107</td>
<td></td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.79</td>
<td>0.84</td>
<td>0.94</td>
<td>1.14</td>
<td>1.55</td>
<td>1.48</td>
</tr>
<tr>
<td>$E_N$</td>
<td>0.03950</td>
<td>0.02427</td>
<td>0.01355</td>
<td>0.00674</td>
<td>0.00256</td>
<td></td>
</tr>
<tr>
<td>$p_N$</td>
<td>0.84</td>
<td>0.87</td>
<td>0.97</td>
<td>1.08</td>
<td>1.19</td>
<td>1.34</td>
</tr>
</tbody>
</table>
Table 6. Maximum errors $E_N$ and computed rates of convergence $p_N$ for $A = 0.0, B = -0.3849$ and $\delta_1 = -0.1, \delta_2 = 0.15$ for several values of $M$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M = 0.5$</td>
<td>$E_N = 0.04491$</td>
<td>0.03159</td>
<td>0.02295</td>
<td>0.01493</td>
<td>0.00955</td>
<td>0.00535</td>
</tr>
<tr>
<td></td>
<td>$p_N = 0.50$</td>
<td>0.55</td>
<td>0.65</td>
<td>0.83</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>$M = 1.0$</td>
<td>$E_N = 0.00856$</td>
<td>0.00429</td>
<td>0.00215</td>
<td>0.00106</td>
<td>0.00049</td>
<td>0.00020</td>
</tr>
<tr>
<td></td>
<td>$p_N = 0.98$</td>
<td>1.00</td>
<td>1.03</td>
<td>1.10</td>
<td>1.26</td>
<td>1.65</td>
</tr>
<tr>
<td>$M = 2.5$</td>
<td>$E_N = 0.00014$</td>
<td>0.00077</td>
<td>0.00023</td>
<td>0.00009</td>
<td>0.00003</td>
<td>0.00014</td>
</tr>
<tr>
<td></td>
<td>$p_N = 0.04$</td>
<td>0.08</td>
<td>0.14</td>
<td>0.17</td>
<td>0.22</td>
<td>0.56</td>
</tr>
</tbody>
</table>

on $f$ is violated. Table 6 is for the problem (5.4) with $A = 0.0, B = -0.3849$, and $\delta_1 = -0.1, \delta_2 = 0.15$, which satisfy condition (2.2) with $\theta = 2/3$.

We remark that in these cases where the right-hand side and the boundary conditions are such that condition (2.2) is close to being violated, the number of iterations also increases dramatically. Table 7 gives the iteration counts for $M = 2.5$.

Table 7. Number of Iterations for $A = B = 0$ and $\delta_1 = -0.3849, \delta_2 = 0.15$ and $M = 2.5$.

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
<th>2048</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{-2}$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>$2^{-8}$</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>$2^{-10}$</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>$2^{-12}$</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>$2^{-14}$</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>$2^{-16}$</td>
<td>28</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>$2^{-18}$</td>
<td>37</td>
<td>35</td>
<td>34</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>$2^{-20}$</td>
<td>50</td>
<td>47</td>
<td>44</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>$2^{-22}$</td>
<td>67</td>
<td>63</td>
<td>59</td>
<td>55</td>
<td>53</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>$2^{-24}$</td>
<td>89</td>
<td>83</td>
<td>78</td>
<td>73</td>
<td>69</td>
<td>67</td>
<td>66</td>
</tr>
</tbody>
</table>

Table 8. Number of Iterations for $A = -0.57735, B = 0.57735$ and $\delta_1 = -0.3849001794597, \delta_2 = 0.15$ and $M = 2.5$.

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
<th>2048</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{-2}$</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>$2^{-8}$</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>$2^{-10}$</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>$2^{-12}$</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>$2^{-14}$</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>$2^{-16}$</td>
<td>33</td>
<td>32</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>$2^{-18}$</td>
<td>45</td>
<td>42</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>$2^{-20}$</td>
<td>62</td>
<td>57</td>
<td>53</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>$2^{-22}$</td>
<td>84</td>
<td>77</td>
<td>72</td>
<td>68</td>
<td>65</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>$2^{-24}$</td>
<td>113</td>
<td>104</td>
<td>97</td>
<td>91</td>
<td>85</td>
<td>82</td>
<td>82</td>
</tr>
</tbody>
</table>
for the case $A = B = 0$ and $\delta_1 = -0.3849, \delta_2 = 0.15$. Table gives the iteration counts for the case $A = -0.55735, B = 0.55735, \delta_1 = -0.3849001794597, \delta_2 = 0.15$, and $M = 2.5$. The iteration counts for other values of $M$ also increased to similar numbers. In these cases, however, the number of iterations did vary somewhat with the choice of $M$. On the other hand, if the sufficient condition for existence of the reduced solution, that is,

$$|f| \leq \frac{2}{3\sqrt{3}} = 0.384900179459750967,$$

is exceeded, then for sufficiently small $\varepsilon$ the algorithm does not converge. For example, for $A = 0.0, B = 0.0, \delta_1 = -0.39, \delta_2 = 0.15$, and $M = 2.5$, the algorithm fails to converge for $\varepsilon < 2^{-19}$.

REFERENCES


