JOURNAL OF COMPUTATIONAL PHYSICS 105, 24–32 (1993)

A Comparison of Uniformly Convergent Difference Schemes for Two-Dimensional Convection–Diffusion Problems

ALAN F. HEGARTY

Mathematics Department, University of Limerick, Limerick, Ireland

EUGENE O'RIORDAN

Mathematics Department, Regional Technical College, Dundalk, Ireland

AND

MARTIN STYNES*

Mathematics Department, University College, Cork, Ireland

Received March 4, 1991; revised May 12, 1992

Galerkin and Petrov Galerkin finite element methods are used to obtain new finite difference schemes for the solution of linear two-dimensional convection diffusion problems. Numerical estimates are made of the rates of convergence of these schemes, uniformly with respect to the perturbation parameter, and these uniform rates are shown to compare favourably with those of established methods. © 1993 Academic Press, Inc.

1. INTRODUCTION

In this paper, we will examine methods for the numerical solution of the following linear elliptic singular perturbation problem:

$$Lu \equiv \varepsilon\, \Delta u + \mathbf{a} \cdot \nabla u - bu = f \quad \text{on} \quad \Omega = (0, 1) \times (0, 1), \qquad (1.1)$$
$$u = 0 \quad \text{on } \partial\Omega,$$

where $0 < \varepsilon \leqslant 1$, the coefficients $\mathbf{a}$, b, and f are sufficiently smooth—see O'Riordan and Stynes [10]—and

$$\mathbf{a} = (a_1(x), a_2(y)) \geqslant (\alpha_1, \alpha_2) > 0 \quad \text{on } \bar{\Omega} \qquad (1.2\text{a})$$
$$b(x, y) \geqslant \beta > 0 \quad \text{on } \bar{\Omega} \qquad (1.2\text{b})$$
$$b + \tfrac{1}{2}\nabla \cdot \mathbf{a} > 0 \quad \text{on } \bar{\Omega}. \qquad (1.2\text{c})$$

We impose the following compatibility conditions at the four corners of Ω:

$$f(0, 0) = f(0, 1) = f(1, 0) = f(1, 1) = 0, \qquad (1.3)$$

so that the solution is sufficiently smooth in $\bar{\Omega}$—see [10] for details. These conditions will not always be satisfied in the numerical examples in Section 4, but are required for the theoretical convergence results. For small values of ε, boundary layers will, in general, appear along the two sides $\{(x, 0) : 0 \leqslant x \leqslant 1\}$ and $\{(0, y) : 0 \leqslant y \leqslant 1\}$ of $\bar{\Omega}$.

This problem may be viewed as modelling a steady-state convection–diffusion process, in which u represents some transported quantity—such as heat or momentum; the first-order terms u_x and u_y represent convection; and the second-order term Δu, diffusion. While accurate numerical solution of (1.1) is desirable in itself, its real importance lies in the fact that it can also be regarded as a simple linear model of the Navier–Stokes flow equations, or of the electron and hole continuity equations in semiconductor device modelling. Numerical solution of these equations—or, indeed, of other similarly structured equations—will often be equivalent to the numerical solution of a sequence of linear equations of the form (1.1). In these instances the case of small ε is known as convection dominated flow and, depending on the transported quantity u, small values of the parameter ε are equivalent to large values of the Reynolds, Peclet, or Hartmann number.

Classical numerical methods, such as centred finite difference methods or standard Galerkin finite element

* Research partly supported by Arts Faculty Research Fund, University College, Cork.

0021-9991/93 $5.00

methods, are, in general, unsuitable for such problems, as they become unstable for small values of ε; this is because the reduced problem ($\varepsilon = 0$) is hyperbolic and thus when ε is small the problem is essentially hyperbolic outside the layer region. It is well known that the use of an upwind method will at least produce a stable numerical solution to (1.1); currently, the most popular of these methods is the streamline diffusion method, which upwinds along the streamlines (the characteristics of the reduced hyperbolic problem). Johnson *et al.* [5, 6] and Niijima [9] have proved that variations of the streamline diffusion method are accurate in regions where the solution u of (1.1) does not vary rapidly (i.e., away from the layers). In fact, global error bounds are obtained, but these bounds depend on Sobolev norms of the solution u which are large when ε is small.

We are interested in methods which converge irrespective of the value of ε throughout the entire domain Ω. We shall call such methods "globally uniformly convergent" (GUC); thus, if u is the solution of (1.1) and u^h is a numerical solution obtained using a GUC method, the inequality

$$\|u - u^h\| \leqslant Ch^p,$$

holds for some C and p independent of ε and the mesh-width h and for some appropriate norm $\|\cdot\|$, which measures the behaviour on *all* of Ω. The analyses in [5, 6, 9] mentioned above only show the streamline diffusion schemes to be uniformly convergent outside layer regions, and, measured in the discrete L^∞ norm, they fail to satisfy the necessary conditions for global uniform convergence of Roos [12].

Globally uniformly convergent finite difference schemes for problems like (1.1) have been examined by many authors (see Emel'janov [1], Hegarty [3], Shishkin [13], and the references therein). Many of these schemes are variations of Il'in's scheme [4] and satisfy a discrete maximum principle; however, in O'Riordan and Stynes [11] a uniform convergence result is obtained for one of the schemes examined below, without any recourse to a discrete maximum principle.

Other authors (e.g., Liseikin [8]) have obtained GUC difference schemes by using special meshes which condense in the regions where the solution changes rapidly. In this paper, we will always use a uniform mesh, which does not explicitly attempt to follow the characteristics of the reduced problem.

In Section 2, we discuss various norms and whether they are appropriate for the measurement of errors in the numerical solution of this singularly perturbed problem. In Section 3, the difference schemes are derived by means of Galerkin and Petrov–Galerkin finite element methods. Finally, in Section 4, the numerical performance of our methods is compared with Il'in's finite difference scheme. From these results it appears that some of the methods proposed here are GUC with respect to the discrete L^2 and L^∞ norms, and at least first-order accurate, uniformly in ε, with respect to the discrete L^2 norm.

2. THE CHOICE OF NORM

In this section we assess the suitability of various norms for measuring the error in computed solutions for singular perturbation problems such as (1.1). Much of the discussion will centre on (2.1) below, which is the one-dimensional analogue of (1.1); this is both for simplicity and because several numerical schemes for (2.1) have been thoroughly analysed in the literature.

Notation. Throughout this paper, C denotes a generic constant independent of ε and of any mesh. Given a mesh of diameter h, by $g = O(h^\gamma)$ we mean that, for h sufficiently small, $|g| \leqslant Ch^\gamma$, where γ may be any nonnegative constant independent of ε and h.

Consider thus the two-point boundary value problem

$$\varepsilon z''(x) + r(x)\, z'(x) - s(x)\, z(x) = w(x), \qquad 0 < x < 1, \quad (2.1a)$$

$$z(0) = z(1) = 0, \quad (2.1b)$$

where ε is a parameter in $(0, 1]$, $r(x) > 0$ on $[0, 1]$, and $s(x) \geqslant 0$ on $[0, 1]$. We assume that r, s, and w are smooth. In general the solution $z(x)$ has a boundary layer near $x = 0$.

Fix a positive integer N. Set $h = 1/N$ and $x_i = ih$ for $i = 0, 1, ..., N$. Define the discrete $L^p[0, 1]$ norms by

$$\|g\|_{p,d,1} = \begin{cases} \left\{ h \sum_{i=1}^{N-1} |g(x_i)|^p \right\}^{1/p}, & \text{if } 1 \leqslant p < \infty, \\ \max_i |g(x_i)|, & \text{if } p = \infty, \end{cases}$$

for all functions $g(\cdot)$ defined on the mesh $\{x_i\}$ and satisfying $g(0) = g(1) = 0$.

For all such $g(\cdot)$, one can easily show that

$$\|g\|_{p,d,1} \leqslant \|g\|_{q,d,1}, \qquad \text{if } 1 \leqslant p \leqslant q \leqslant \infty,$$

using Hölder's inequality.

Kellogg and Tsan [7] have analysed the nodal errors of certain difference schemes when applied to (2.1). Letting z_i^1 and z_i^2 denote the solution computed at the node x_i by simple upwind differencing and Il'in's scheme, respectively, it is shown in [7] that, for each i,

$$|z(x_i) - z_i^1| \leqslant \begin{cases} Ch[1 + \varepsilon^{-1} \exp(-r\varepsilon^{-1} x_i)], \\ \quad \text{if } h \leqslant \varepsilon, \\ C[h + \exp(-rx_i/(rh + \varepsilon))], \\ \quad \text{if } h \geqslant \varepsilon, \end{cases} \quad (2.2)$$

and

$$|z(x_i) - z_i^2| \leqslant Ch^2/(h+\varepsilon) + Ch^2\varepsilon^{-1}\exp(-r\varepsilon^{-1}x_i), \qquad (2.3)$$

where $0 < r < \min_{[0,1]} r(x)$.

These inequalities essentially say that upwinding is $O(h)$ accurate outside the boundary layer and only $O(1)$ accurate near $x = 0$, but the Il'in scheme is $O(h)$ accurate at all nodes. Nevertheless, the discrete L^1 norm is unable to discriminate between the two schemes; measured in $\|\cdot\|_{1,d,1}$, both are $O(h)$ accurate, as an easy calculation using (2.2) and (2.3) shows. We consequently reject $\|\cdot\|_{1,d,1}$ as unsuitable for comparing the errors of competing difference schemes used to solve (2.1) or (1.1).

Remark 2.1. If the position of the layer is not known a priori, then the discrete L^1 norm is a valid measure of how accurately this layer has been located; however, if the layer is to be resolved, then a stronger norm is required in order to measure how well this has been done—this is illustrated by Table III and Fig. 4.1.

Apart from $p = 1$, the most common choices of $\|\cdot\|_{p,d,1}$ are $p = 2$ and $p = \infty$. While $\|\cdot\|_{\infty,d,1}$ is, of course, the strongest L^p norm, we shall work with the two-dimensional analogue of $\|\cdot\|_{2,d,1}$. There are several reasons for this: (i) the solution to (1.1) on almost all of Ω is essentially the solution to a first-order hyperbolic problem (obtained by setting $\varepsilon = 0$ in (1.1) and using the boundary conditions on the inflow boundary of Ω), and for such problems it is usual to analyse difference schemes using an L^2-type norm; (ii) in finite element arguments, the L^2 norm is more natural than the L^∞ norm; (iii) unlike the L^1 norm, the L^2 norm is able to distinguish between schemes which are first-order accurate at all nodes and schemes which are first-order accurate only outside the boundary layers, and then degenerate to $O(1)$ accuracy in the layers.

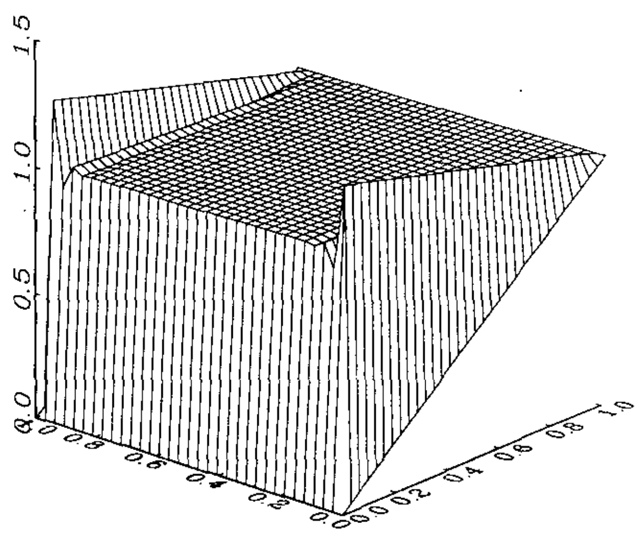

FIG. 4.1. $\varepsilon = 0.00000001$; Scheme PG2, $h = \frac{1}{32}$.

DEFINITION. Place a square grid of diameter $h = 1/N$ on $\bar{\Omega}$, where Ω is as in Section 1, so that the grid points are $\{(ih, jh): i, j = 0, 1, ..., N\}$. Then the discrete L^2 norm of a function defined on this grid is defined as

$$\|v\|_{2,d} = \left\{h^2 \sum_{i,j=0}^{N} v^2(x_i, y_j)\right\}^{1/2},$$

where $(x_i, y_j) \equiv (ih, jh)$ for each pair i, j.

We shall also work with an energy norm which is stronger than the discrete L^2 norm. For $v \in H_0^1(\Omega)$, this norm is defined to be

$$|||v||| = \left\{\varepsilon(|\nabla v|^2, 1) + h^2 \sum_{i,j=1}^{N-1} v^2(x_i, y_j)\right\}^{1/2},$$

where for $v, w \in L^2(\Omega)$, (v, w) denotes the usual $L^2(\Omega)$ inner product.

We can prove that one of the finite element methods presented in Section 3 is $O(h^{1/2})$ accurate in $|||\cdot|||$. We now show (in the one-dimensional case) that such an estimate excludes the possibility that the nodal accuracy falls to $O(1)$ near the outflow boundary. That is, near the layer, upwind schemes which are not of exponential type perform significantly worse than our schemes.

In the one-dimensional case, by analogy with the schemes presented in Section 3, a basis for our trial space S_1 would be $\{\phi_i : i = 1, 2, ..., N-1\}$, where each ϕ_i satisfies

$$\varepsilon\phi_i''(x) + r(x)\,\phi_i'(x) = 0 \quad \text{on} \quad [0, 1]\backslash\{x_0, x_1, ..., x_N\}$$

$$\phi_i(x_j) = \delta_{i,j}, \qquad \text{for} \quad j = 0, 1, ..., N;$$

here r is a piecewise constant approximation to $r(x)$ defined by $r|_{(x_{i-1}, x_i]} \equiv r_i \equiv r(x_i)$ for $i = 1, 2, ..., N$. On $[0, 1]$, the energy norm is

$$|||v|||_1 = \left\{\int_0^1 \varepsilon(v'(x))^2\,dx + \|v\|_{2,d,1}^2\right\}^{1/2}.$$

PROPOSITION 2.2. *Let $z^h \in S_1$ satisfy $|||z - z^h|||_1 \leqslant Ch^{1/2}$. Let M be any fixed positive integer. Then*

$$|z(x_i) - z^h(x_i)| \leqslant C'h^{1/2} \qquad \textit{for} \quad 1 \leqslant i \leqslant M,$$

where C' depends on M.

Proof. Let $z_I \in S_1$ interpolate to $z(x)$ at all the nodes. Then $|||z - z_I|||_1 \leqslant Ch^{1/2}$ [14]. By the triangle inequality,

$$|||z_I - z^h|||_1 \leqslant Ch^{1/2}. \qquad (2.4)$$

On the other hand, a direct computation yields

$$|||z_I - z^h|||_1 = \left\{ \sum_{i=1}^{N} \frac{r_i(1+e^{-r_i h/\varepsilon})}{2(1-e^{-r_i h/\varepsilon})} (\zeta(x_i) - \zeta(x_{i-1}))^2 + h \sum_{i=1}^{N-1} \zeta^2(x_i) \right\}^{1/2},$$

where $\zeta \equiv z - z^h$. Combining this with (2.4), we have in particular

$$\sum_{i=1}^{N} (\zeta(x_i) - \zeta(x_{i-1}))^2 \leqslant Ch.$$

Since $\zeta(x_0) = 0$, this implies the desired result. ∎

Remark 2.3. The result of Proposition 2.2 shows that a scheme which is convergent of order $h^{1/2}$ with respect to the energy norm will yield an accurate approximation even in the boundary layer.

Remark 2.4. In the energy norm $|||\cdot|||_1$, the optimal order of accuracy attainable using the trial functions described above is, in fact, $O(h^{1/2})$, for consider the case where $r(x)$ and $w(x)$ are positive constants and $s(x) \equiv 0$. Let $v \in S_1$ be arbitrary. Then, on computing $|||z - v|||$ explicitly in terms of $\{(z-v)(x_i) : i = 1, ..., N-1\}$, it is easy to see that $|||z - v|||$ is minimised when we take $v = z_I$. One then has

$$|||z - z_I|||_1 = (w/r)\, \varepsilon^{1/2}\, (rh/(2\varepsilon) \coth(rh/(2\varepsilon)) - 1)^{1/2}.$$

This implies that, uniformly in ε, $\inf_{v \in S_1} |||z - v||| = |||z - z_I|||_1$ is $O(h^{1/2})$ but not $O(h^\gamma)$ for any $\gamma > \frac{1}{2}$.

3. FINITE ELEMENT METHODS

A weak form of problem (1.1) is: find $u \in H_0^1(\Omega)$, such that

$$B(u, v) \equiv (-\varepsilon \nabla u \cdot \nabla v, 1) + (\mathbf{a} \cdot \nabla u - bu, v) = (f, v) \quad \forall v \in H_0^1(\Omega). \tag{3.1}$$

In this section, we will generate finite difference schemes for problem (1.1) by discretizing the above weak form using Galerkin and Petrov–Galerkin finite element methods. Let $\Omega^h = \{(x_i, y_j) : i, j = 0, 1, ..., N\}$ be the set of mesh points defined in Section 2. We will define a set of trial functions $\{\phi^{i,j}(x, y) : i, j = 1, ..., N-1\}$ and a set of test functions $\{\psi^{i,j}(x, y) : i, j = 1, ..., N-1\}$. The trial (test) space is the linear span of these trial (test) functions and will be denoted by S (T). Let

$$u^h(x, y) = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} u^h_{i,j} \phi^{i,j}(x, y) \tag{3.2}$$

be our finite element approximation to $u(x, y)$. Here the nodal values are determined from

$$\bar{B}(u^h, \psi^{i,j}) = (\bar{f}, \psi^{i,j}) \quad \forall \psi^{i,j} \in T, \tag{3.3}$$

where

$$\bar{B}(v, w) \equiv (-\varepsilon \nabla v \cdot \nabla w, 1) + (\bar{a}_1 v_x + \bar{a}_2 v_y, w) - h^2 \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} (bvw)(x_i, y_j),$$

for all $v \in S$ and $w \in T$. We define $\bar{a}_1$, $\bar{a}_2$, and $\bar{f}$ to be

$$\bar{a}_1(x) = (a_1(x_{i-1}) + a_1(x_i))/2, \quad \text{if } x \in [x_{i-1}, x_i), \tag{3.4a}$$

$$\bar{a}_2(y) = (a_2(y_{j-1}) + a_2(y_j))/2, \quad \text{if } y \in [y_{j-1}, y_j), \tag{3.4b}$$

$$(\bar{f}, \psi^{i,j}) = h^2 f(x_i, y_j) \quad \text{for each test function } \psi^{i,j}. \tag{3.4c}$$

Since the solution of the continuous problem has boundary layers, it seems natural to incorporate boundary layer behaviour into the trial space; hence, the trial functions are chosen to be the tensor product of one-dimensional exponential trial functions. That is,

$$\phi^{i,j}(x, y) = \phi^i(x)\, \phi_j(y) \qquad \forall i, j, \tag{3.5a}$$

where each $\phi^i(x)$ satisfies

$$\varepsilon(\phi^i)'' + \bar{a}_1(\phi^i)' = 0, \qquad x \in (0, 1) \backslash \{x_1, ..., x_{N-1}\}, \qquad \phi^i(x_j) = \delta_{i,j}, \tag{3.5b}$$

and each $\phi_j(y)$ satisfies

$$\varepsilon(\phi_j)'' + \bar{a}_2(\phi_j)' = 0, \qquad y \in (0, 1) \backslash \{y_1, ..., y_{N-1}\}, \qquad \phi_j(y_i) = \delta_{i,j}. \tag{3.5c}$$

We shall refer to such functions $\phi^{i,j}$ as $\bar{L}$ splines. The test functions are chosen to be of the form

$$\psi^{i,j}(x, y) = \psi^i(x)\, \psi_j(y)$$

and satisfy

$$\psi^i(x_j) = \delta_{i,j}, \qquad \psi_j(y_i) = \delta_{i,j},$$
$$\operatorname{supp}(\psi^{i,j}) = [x_{i-1}, x_{i+1}] \times [y_{j-1}, y_{j+1}].$$

These test functions are not yet fully specified. The finite difference scheme (3.3) may be written in the form

$$\frac{\varepsilon}{h^2}\sum_{r=i-1}^{i+1}\sum_{q=j-1}^{j+1}\alpha_{r,q}u^h_{r,q}-b(x_i,y_j)\,u^h_{i,j}=f(x_i,y_j) \quad (3.6a)$$

for $i=1, ..., N-1$, and $j=1, ..., N-1$; also

$$\begin{aligned} u^h_{0,j}=u^h_{N,j}=0 \quad &\text{for } j=0,1,...,N,\\ u^h_{i,0}=u^h_{i,N}=0 \quad &\text{for } i=0,1,...,N. \end{aligned} \quad (3.6b)$$

Here,

$$\begin{pmatrix}\alpha_{i-1,j+1} & \alpha_{i,j+1} & \alpha_{i+1,j+1}\\ \alpha_{i-1,j} & \alpha_{i,j} & \alpha_{i+1,j}\\ \alpha_{i-1,j-1} & \alpha_{i,j-1} & \alpha_{i+1,j-1}\end{pmatrix} = \begin{pmatrix} R_x^-S_y^+ + S_x^-R_y^+ & R_x^cS_y^+ + S_x^cR_y^+ & R_x^+S_y^+ + S_x^+R_y^+\\ R_x^-S_y^c + S_x^-R_y^c & R_x^cS_y^c + S_x^cR_y^c & R_x^+S_y^c + S_x^+R_y^c\\ R_x^-S_y^- + S_x^-R_y^- & R_x^cS_y^- + S_x^cR_y^- & R_x^+S_y^- + S_x^+R_y^-\end{pmatrix}$$

$$R_x^+=\sigma(\rho_x^{i+1}),\quad R_x^-=\sigma(-\rho_x^i),\quad R_x^c=-(R_x^-+R_x^+);$$

$$R_x^+=\sigma(\rho_y^{j+1}),\quad R_y^-=\sigma(-\rho_y^j),\quad R_y^c=-(R_y^-+R_y^+);$$

$$\rho_x^i=\frac{\bar a_1(x_{i-1})h}{\varepsilon},\quad \rho_y^j=\frac{\bar a_2(y_{j-1})h}{\varepsilon},\quad \sigma(x)=\frac{x}{1-e^{-x}}; \quad (3.6c)$$

and

$$S_x^+=\frac{1}{h}(\phi^{i+1},\psi^i),\quad S_x^c=\frac{1}{h}(\phi^i,\psi^i),\quad S_x^-=\frac{1}{h}(\phi^{i-1},\psi^i);$$
$$S_y^+=\frac{1}{h}(\phi_{j+1},\psi_j),\quad S_y^c=\frac{1}{h}(\phi_j,\psi_j),\quad S_y^-=\frac{1}{h}(\phi_{j-1},\psi_j). \quad (3.6d)$$

Note that $S^+ + S^c + S^- = 1$.

Remark 3.1. Irrespective of the choice of test functions, the above difference scheme satisfies the necessary conditions C1, C2, and C3, given by Roos [12] for uniform convergence in the discrete L^∞ norm. It can also be shown [15] that, in order to obtain uniform convergence of order greater than $\frac{1}{2}$ in the discrete L^2 norm, it is necessary that the difference scheme should satisfy conditions C1 and C2 of [12].

Remark 3.2. If we lump—i.e., set $S^+=S^-=0$ and $S^c=1$—then the difference scheme (3.6) is Il'in's exponentially fitted finite difference scheme (irrespective of the choice of test functions). This lumped difference scheme satisfies a discrete maximum principle which guarantees the stability of the scheme. However, one of the non-lumped difference schemes which we propose below—(3.8c)—appears numerically superior to this lumped scheme.

For the test functions, we examine three possible choices:

(a) $\bar L$ splines, i.e.,

$$\psi^{i,j}(x,y)=\phi^{i,j}(x,y); \quad (3.7a)$$

(b) $\bar L^*$ splines, i.e.,

$$\begin{aligned} \varepsilon(\psi^i)''-\bar a_1(\psi^i)'=0,\quad &\psi^i(x_j)=\delta_{i,j},\\ \varepsilon(\psi_j)''-\bar a_2(\psi_j)'=0,\quad &\psi_j(y_i)=\delta_{i,j}, \end{aligned} \quad (3.7b)$$

(c) bilinear functions, i.e.,

$$\psi^i=\begin{cases}\dfrac{x-x_{i-1}}{h}, & x_{i-1}\leqslant x\leqslant x_i\\ \dfrac{x_{i+1}-x}{h}, & x_i\leqslant x\leqslant x_{i+1},\end{cases}$$
$$\psi^j(y)=\begin{cases}\dfrac{y-y_{j-1}}{h}, & y_{j-1}\leqslant y\leqslant y_j\\ \dfrac{y_{j+1}-y}{h}, & y_j\leqslant y\leqslant y_{j+1}.\end{cases} \quad (3.7c)$$

Choice (a) yields a Galerkin method, while (b) and (c) are Petrov–Galerkin methods.

The terms in (3.6d) can now be evaluated, where the test functions are:

(a) $\bar L$ splines:

$$\begin{aligned} S^-=\eta(\rho),\quad &S^+=\eta(\rho),\\ \eta(x)\equiv\frac{e^{-x}}{(1-e^{-x})^2}&\left[\frac{\sinh x}{x}-1\right], \end{aligned} \quad (3.8a)$$

i.e., $S_x^-=\eta(\rho_x^i)$, $S_y^-=\eta(\rho_y^j)$, $S_x^+=\eta(\rho_x^{i+1})$, $S_y^+=\eta(\rho_y^{j+1})$.

(b) $\bar L^*$ splines:

$$\begin{aligned} S^-=R^-\xi(\rho),\quad &S^+=R^+\xi(\rho);\\ \xi(x)\equiv\left(\frac{2}{x^2}\right)&\left(\frac{x}{2}\coth\frac{x}{2}-1\right); \end{aligned} \quad (3.8b)$$

(c) bilinear:

$$\begin{aligned} S^-=R^-\tau(\rho),\quad &S^+=R^+\tau(-\rho);\\ \tau(x)\equiv(e^x-(1+x&+x^2/2))/x^3. \end{aligned} \quad (3.8c)$$

Remark 3.3. If we were to employ $\bar L^*$ splines in both the trial and test spaces, the resulting difference scheme would coincide with that obtained using $\bar L$ splines in both the trial and test spaces. Similarly, if we take bilinear trial functions and $\bar L^*$ spline test functions, the resulting difference scheme coincides with that generated by $\bar L$ spline trial functions and bilinear test functions.

For the choice (a) of test functions, we have the following energy norm estimate of the error:

THEOREM 3.4. (O'Riordan and Stynes [10]). *Assume* **a** *is constant in* $\bar{\Omega}$. *Suppose the trial functions are chosen as in* (3.5) *and the test functions as in* (3.7a). *If the conditions* (1.2) *and* (1.3) *are satisfied, then*

$$|||u - u^h||| \leqslant Ch^{1/2}.$$

Remark 3.5. Recall that, by Remark 2.3, $O(h^{1/2})$ is the optimal order of accuracy obtainable in one dimension. The result of Theorem 3.4 may be extended to the case of variable **a**, using a modified discretization of the term (bu, u); see O'Riordan and Stynes [11] for details. However, theoretical considerations indicate that a "good" choice of test functions should approximately satisfy the homogeneous adjoint equation

$$L^* v \equiv \varepsilon \Delta v - \nabla \cdot (\mathbf{a} v) - bv = 0.$$

This motivates the second choice of test functions (3.7b).

THEOREM 3.6. *Assume* **a** *is constant in* $\bar{\Omega}$. *Suppose the trial functions are chosen as in* (3.5) *and the test functions as in* (3.7b); *then, given any* $v \in S$, *there exists* $\hat{v} \in T$ *such that*

$$|\bar{B}(v, \hat{v})| \geqslant C_1 \, |||v|||^2.$$

Proof. Note that for our choice of trial and test functions,

$$\psi^i(x) = e^{a_1(x - x_i)/\varepsilon} \phi^i(x).$$

Hence, for any $v \in S$,

$$(-\varepsilon v_x, \psi^i_x) + (a_1 v_x, \psi^i) = (-\varepsilon v_x, e^{a_1(x - x_i)/\varepsilon} \phi^i_x).$$

Given $v \in S$, let $\hat{v} \in T$ be such that

$$\hat{v}_{i,j} = e^{a_1 x_i/\varepsilon} e^{a_2 y_j/\varepsilon} v_{i,j}.$$

Then,

$$\begin{aligned} -\bar{B}(v, \hat{v}) &= \sum_{i,j} \hat{v}_{i,j} \bar{B}(v, \psi^{i,j}) \\ &= \sum_{i,j} \hat{v}_{i,j} \,[(\varepsilon v_x, e^{a_1(x - x_i)/\varepsilon} e^{a_2(y - y_j)/\varepsilon} \phi^i_x \phi_j) \\ &\quad + (\varepsilon v_y, e^{a_1(x - x_i)/\varepsilon} e^{a_2(y - y_j)/\varepsilon} \phi^i (\phi_j)_y)] \\ &\quad + h^2 \sum_{i,j} b_{i,j} v_{i,j} \hat{v}_{i,j} \\ &= (\varepsilon v_x, w v_x) + (\varepsilon v_y, w v_y) + h^2 \sum_{i,j} b_{i,j} v^2_{i,j} w_{i,j}, \end{aligned}$$

where $w = e^{a_1 x/\varepsilon} e^{a_2 y/\varepsilon} \geqslant 1$. Consequently,

$$|\bar{B}(v, \hat{v})| \geqslant (\varepsilon v_x, v_x) + (\varepsilon v_y, v_y) + \beta h^2 \sum_{i,j} v^2_{i,j} \geqslant C_1 \, |||v|||^2. \quad \blacksquare$$

Remark 3.7. This guarantees the invertibility of the matrix associated with the test functions (3.7b). However, it does not guarantee stability, since $\hat{v}$ is, in general, exponentially large; in fact, the reduced matrix, i.e., that obtained by setting $\varepsilon = 0$ in the difference scheme matrix, is singular when $b \equiv 0$. The instability of (3.6) and (3.8b) has been observed numerically.

The reduced difference scheme for (3.6) and (3.8a) is the same as Il'in's reduced scheme, which in turn coincides with the reduced standard upwinded scheme (no exponential fitting). For $\varepsilon = 0$, (3.6) and (3.8a) reduce to

$$\begin{aligned} &\frac{\bar{a}_1(x_i)}{h} (u^h_{i+1,j} - u^h_{i,j}) + \frac{\bar{a}_2(y_j)}{h} (u^h_{i,j+1} - u^h_{i,j}) \\ &\quad - b(x_i, y_j)\, u^h_{i,j} = f(x_i, y_j). \end{aligned} \tag{3.9a}$$

which is first-order accurate. However, the reduced difference scheme for (3.6) and (3.8c) is significantly different. That is, for $\varepsilon = 0$, (3.6) and (3.8c) reduce to (Keller's box scheme):

$$\begin{aligned} &\frac{\bar{a}_1(x_i)}{2h} [(u^h_{i+1,j} - u^h_{i,j}) + (u^h_{i+1,j+1} - u^h_{i,j+1})] \\ &\quad + \frac{\bar{a}_2(y_j)}{2h} [(u^h_{i,j+1} - u^h_{i,j}) \\ &\quad + (u^h_{i+1,j+1} - u^h_{i+1,j})] \\ &\quad - b(x_i, y_j)\, u^h_{i,j} = f(x_i, y_j). \end{aligned} \tag{3.9b}$$

In the case of constant a_1 and a_2, this becomes Wendroff's implicit scheme for first-order hyperbolic problems [16].

Experience of one-dimensional singularly perturbed problems suggests that higher uniform accuracy can be achieved by choosing appropriate quadrature rules for the inhomogeneous term $(\bar{f}, \psi^{i,j})$ in (3.3). Hence, we propose the following alternative to (3.4c):

$$\begin{aligned} \bar{f}(x, y) = \tfrac{1}{4}(&f(x_i, y_j) + f(x_{i-1}, y_j) \\ &+ f(x_i, y_{j-1}) + f(x_{i-1}, y_{j-1})), \\ &\text{if} \quad (x, y) \in [x_{i-1}, x_i) \times [y_{j-1}, y_j). \end{aligned} \tag{3.10}$$

For this choice of quadrature rule, the right-hand side of the difference scheme (3.3) now depends on the choice of test space. Thus, we will fix our choice of test functions as $\bar{L}^*$ splines (i.e., (3.7b)). By virtue of Remark 3.7, taking $\bar{L}$ splines in the trial space is then of no value. Hence, we will examine the schemes generated by taking $\bar{L}^*$ splines in the test space, coupled with either $\bar{L}^*$ splines or bilinear functions in the trial space. This does not conflict with our earlier work; recall Remark 3.3.

Incorporating (3.10) into (3.3), the difference scheme is of the form

$$\frac{\varepsilon}{h^2}\sum_{r=i-1}^{i+1}\sum_{j=j-1}^{j+1}\alpha_{r,q}u^h_{r,q}-b(x_i,y_j)\,u^h_{i,j}=\sum_{r=i-1}^{i+1}\sum_{q=j-1}^{j+1}\gamma_{r,q}f(x_r,y_q) \qquad (3.11a)$$

for $i=1,\ldots,N-1$ and $j=1,\ldots,N-1$, where

$$\begin{pmatrix}\gamma_{i-1,j+1} & \gamma_{i,j+1} & \gamma_{i+1,j+1}\\ \gamma_{i-1,j} & \gamma_{i,j} & \gamma_{i+1,j}\\ \gamma_{i-1,j-1} & \gamma_{i,j-1} & \gamma_{i+1,j-1}\end{pmatrix}=\tfrac{1}{4}\begin{pmatrix}Q_x^-Q_y^+ & Q_x^cQ_y^+ & Q_x^+Q_y^+\\ Q_x^-Q_y^c & Q_x^cQ_y^c & Q_x^+Q_y^c\\ Q_x^-Q_y^- & Q_x^cQ_y^- & Q_x^+Q_y^-\end{pmatrix}$$

$$Q_x^-=\frac{1}{h}\int_{x_{i-1}}^{x_i}\psi^i(x)\,dx=\frac{1-\sigma(-\rho_x^i)}{\rho_x^i};$$

$$Q_y^+=\frac{1}{h}\int_{y_j}^{y_{j+1}}\psi_j(y)\,dy=1-\frac{1-\sigma(-\rho_y^{j+1})}{\rho_y^{j+1}}; \qquad (3.11b)$$

$$Q^c=Q^-+Q^+;$$

and because of the choice of $\bar{L}^*$ splines as test functions, the terms $\{\alpha_{r,q}\}$ are as given in (3.6), with the terms in (3.6d) now determined by the choice of trial space. If the trial functions are

(a) $\bar{L}^*$ splines: the terms in (3.6d) are as in (3.8a);

(b) bilinear: the terms in (3.6d) are as in (3.8c).

We summarise this section by listing the four difference schemes whose numerical performance will be examined in the next section:

Method	Trial space	Test space	Quadrature	Difference scheme LHS	Difference scheme RHS
G1	$\bar{L}^*$	$\bar{L}^*$	(3.4)	(3.6) (3.8a)	(3.6)
G2	$\bar{L}^*$	$\bar{L}^*$	(3.4a), (3.4b) (3.10)	(3.6) (3.8a)	(3.11)
PG1	bilinear	$\bar{L}^*$	(3.4)	(3.6) (3.8c)	(3.6)
PG2	bilinear	$\bar{L}^*$	(3.4a), (3.4b) (3.10)	(3.6) (3.8c)	(3.11)

4. NUMERICAL RESULTS

In this section, we examine the performance of the above methods in the solution of some test problems and compare it with that of full upwinding and of Il'in's method. Convergence rates are estimated as in Farrell [2]. The problems are solved on a sequence of meshes of mesh-widths $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$, and $\frac{1}{128}$. The errors $|u^h_{i,j}-u(ih,jh)|$ are approximated for successive values of ε on the four coarsest meshes by

$$e^h_\varepsilon(ih,jh)=|u^h_{i,j}-u^{h/2}_{2i,2j}|,$$

where the superscript indicates the mesh diameter. The discrete L^2 norm of the error on each mesh, $\|u^h-u\|_{2,d}$ is thus approximated by

$$E_{2,\varepsilon,h}=h\left[\sum_{i,j}e^h_\varepsilon(ih,jh)^2\right]^{1/2}$$

for each value of ε; similarly, for each ε the discrete L^∞ norm $\|u^h-u\|_{\infty,d}$ of the error is approximated by

$$E_{\infty,\varepsilon,h}=\max_{i,j}e^h_\varepsilon(ih,jh).$$

Convergence rates with respect to each of these norms for each ε and for $h=\frac{1}{8}$, $\frac{1}{16}$, and $\frac{1}{32}$, are estimated by the quantities $p_{2,\varepsilon,h}$, and $p_{\infty,\varepsilon,h}$, where

$$p_{2,\varepsilon,h}=\log_2\left(\frac{E_{2,\varepsilon,h}}{E_{2,\varepsilon,h/2}}\right),$$

$$p_{\infty,\varepsilon,h}=\log_2\left(\frac{E_{\infty,\varepsilon,h}}{E_{\infty,\varepsilon,h/2}}\right).$$

The average of these estimates over the three values of h is taken as an estimate for the order of convergence for each value of ε.

The uniform convergence rate is estimated as in Farrell [2]: for each value of h the quantities

$$E_{2,h}=\max_\varepsilon E_{2,\varepsilon,h}, \qquad E_{\infty,h}=\max_\varepsilon E_{\infty,\varepsilon,h}$$

are computed, and having calculated

$$p_{2,h}=\log_2\left(\frac{E_{2,h}}{E_{2,h/2}}\right),$$

$$p_{\infty,h}=\log_2\left(\frac{E_{\infty,h}}{E_{\infty,h/2}}\right)$$

for each h, the average values $p_2=\frac{1}{3}\sum_h p_{2,h}$ and $p_\infty=\frac{1}{3}\sum_h p_{\infty,h}$ are used as estimates of the uniform convergence rate with respect to the discrete L^2 and L^∞ norms, respectively.

Four singular perturbation problems were examined; all have zero boundary conditions. All calculations were carried out in double-precision Fortran 77 on a micro VAXII.

Problem 1. $\varepsilon \Delta u + 2u_x + u_y = -100xy(1-x)(1-y)$.

Table I shows the averages of the computed rates of convergence over $h = \frac{1}{8}$, $\frac{1}{16}$, and $\frac{1}{32}$, for each value of ε with respect to the discrete L^2 and L^∞ norms and also the computed rates of uniform convergence with respect to each of the norms. We also consider the following three problems.

Problem 2. $\varepsilon \Delta u + (1+x^2)\, u_x + (1+y^2)\, u_y = -100xy(1-x)(1-y)$.

Problem 3. $\varepsilon \Delta u + (1+x+x^2)\, u_x + (1+y+y^2)\, u_y = -100xy(1-x)(1-y)$.

Problem 4. $\varepsilon \Delta u + (1+x^2)\, u_x + (1+y^2)\, u_y = -1$.

Note that Problems 1, 2, and 3 all satisfy the compatibility conditions (1.3), but that problem 4 does not. None of the problems satisfies all of the assumptions (1.2).

From [11] we see that, for Problems 1, 2, and 3, we have

$$|u_x(x, y)| \leqslant C(1 + \varepsilon^{-1} e^{-x/\varepsilon})$$

and

$$|u_y(x, y)| \leqslant C(1 + \varepsilon^{-1} e^{-y/\varepsilon}).$$

That is, these problems exhibit boundary layers at the sides $x = 0$ and $y = 0$ of $\bar{\Omega}$. Problem 4, which does not satisfy (1.3), exhibits similar boundary layers and, furthermore, has corner layers due to the incompatibility of the data at the corners of $\bar{\Omega}$.

Table II lists the computed rates of uniform convergence with respect to each of the norms for all four problems; it also lists (in italics) the computed classical convergence rates—i.e., the average over the three values of h of the computed rates of convergence for $\varepsilon = 1$:

$$\frac{1}{3}\sum_h p_{2,1,h} = \log_2\left(\frac{E_{2,1,h}}{E_{2,1,h/2}}\right),$$

$$\frac{1}{3}\sum_h p_{\infty,1,h} = \log_2\left(\frac{E_{\infty,1,h}}{E_{\infty,1,h/2}}\right).$$

Remark 4.3. Method PG2 exhibits, in general, a significantly higher rate of convergence than all the other methods. The L^∞ estimates given for G2 and PG2 are similar; however, the error approximations $E_{\infty,\varepsilon,h}$ are significantly smaller in the case of PG2.

Remark 4.4. Both in the classical and in the reduced case, PG2 appears to be of second order. This is the only scheme, among those considered, which possesses this desirable property.

Remark 4.5. Our results show that the quadrature rule for the right-hand side should be chosen with great care; witness the considerable improvement from G1 to G2, and from PG1 to PG2.

TABLE I

	L^2 rates						L^∞ rates					
ε	G1	G2	PG1	PG2	Il'in	Upw'd	G1	G2	PG1	PG2	Il'in	Upw'd
1.0000000000	2.02	1.98	2.02	1.98	2.01	0.82	1.90	1.89	1.91	1.88	1.99	0.84
0.5000000000	2.01	1.99	2.01	1.99	1.98	0.83	1.69	1.75	1.72	1.71	1.99	0.83
0.2500000000	1.99	2.02	1.99	1.99	1.96	0.81	1.35	1.58	1.47	1.45	1.96	0.81
0.1250000000	1.94	2.01	1.95	1.99	1.89	0.77	1.07	1.52	1.24	1.26	1.90	0.71
0.0625000000	1.81	1.87	1.86	1.94	1.72	0.69	1.07	1.51	1.13	1.10	1.84	0.56
0.0312500000	1.48	1.62	1.61	1.87	1.43	0.57	1.33	1.78	1.32	1.10	1.58	0.38
0.0156250000	1.22	1.34	1.43	1.78	1.14	0.45	1.28	1.55	1.47	1.69	1.24	0.11
0.0078125000	0.96	1.10	1.20	1.71	0.92	0.40	0.94	1.20	1.25	1.21	0.91	−0.06
0.0039062500	0.81	0.96	1.04	1.74	0.81	0.43	0.75	1.12	1.11	1.38	0.72	−0.19
0.0019531250	0.75	0.91	0.96	1.83	0.76	0.52	0.67	1.06	1.03	1.76	0.66	−0.13
0.0009765625	0.75	0.91	0.96	1.89	0.75	0.62	0.66	1.06	1.02	1.80	0.65	0.05
0.0004882813	0.75	0.91	0.96	1.92	0.75	0.70	0.65	1.05	1.02	1.79	0.65	0.28
0.0002441406	0.75	0.91	0.96	1.93	0.75	0.73	0.65	1.05	1.02	1.78	0.65	0.56
0.0001220703	0.75	0.91	0.96	1.94	0.75	0.74	0.65	1.05	1.02	1.78	0.65	0.56
0.0000610352	0.75	0.91	0.96	1.94	0.75	0.75	0.65	1.05	1.02	1.77	0.65	0.60
Uniform	0.75	0.91	0.96	1.83	0.75	0.53	0.65	1.05	1.02	1.10	0.65	0.05

TABLE II

	Problem 1		Problem 2		Problem 3		Problem 4	
Method	L^2	L^∞	L^2	L^∞	L^2	L^∞	L^2	L^∞
G1	0.75	0.65	0.64	0.53	0.61	0.47	0.65	0.45
	2.02	*1.90*	*2.03*	*1.93*	*2.11*	*1.89*	*2.05*	*1.30*
G2	0.91	1.05	0.87	0.74	0.84	0.67	0.64	0.45
	1.98	*1.89*	*1.98*	*1.95*	*1.99*	*1.96*	*2.06*	*1.28*
PG1	0.96	1.02	0.93	0.93	0.92	0.88	0.72	0.02
	2.02	*1.91*	*2.03*	*1.95*	*2.02*	*1.91*	*2.05*	*1.28*
PG2	1.83	1.10	1.72	0.70	1.26	0.66	1.13	0.39
	1.98	*1.88*	*1.98*	*1.91*	*1.98*	*1.86*	*2.04*	*1.31*
Il'in	0.75	0.65	0.60	0.48	0.61	0.41	0.53	0.34
	1.99	*1.99*	*2.00*	*1.99*	*1.99*	*1.99*	*1.98*	*1.97*
Upwinding	0.53	0.05	0.48	−0.04	0.55	−0.07	0.45	−0.12
	0.82	*0.83*	*0.77*	*0.78*	*0.77*	*0.79*	*1.03*	*1.01*

TABLE III

ε	L^1	L^2	L^∞
1.0000000000	2.34	2.02	1.03
0.2500000000	2.14	2.05	1.07
0.0625000000	1.99	1.93	0.93
0.0156250000	1.96	1.95	0.83
0.0039062500	2.05	2.03	1.30
0.0009765625	2.13	1.93	1.55
0.0002441406	1.93	1.57	1.33
0.0000610352	1.56	1.08	0.73
0.0000152588	1.16	0.69	0.25
0.0000038147	1.01	0.53	0.04
0.0000009537	0.96	0.48	−0.03
0.0000002384	0.95	0.46	−0.05
0.0000000596	0.95	0.46	−0.05
Uniform	0.95	0.46	−0.05

Remark 4.6. Problem 4 does not satisfy the compatibility conditions (1.3); nevertheless, PG2 yields satisfactory results.

Finally, we examine briefly the validity of measuring uniform convergence with respect to the various error norms; in this case we also include the discrete L^1 norm

$$E_{1,\varepsilon,h} = h^2 \sum_{i,j} |e_\varepsilon^h(ih, jh)|$$

and estimate the convergence rates for the different values of ε and uniformly, as with the other two norms. PG2 is now used to solve the following parabolic layer problem. We note that, in practice, it is difficult to obtain accurate numerical results in the neighbourhood of parabolic layers on a uniform mesh.

Problem 5. $\varepsilon \Delta u + u_x = 0$, with boundary conditions

$$u(x, 0) = u(x, 1) = x; \qquad u(0, y) = 0; \qquad u(1, y) = 1.$$

The orders of convergence with respect to the discrete L^1, L^2, and L^∞ norms are listed in Table III. The numerical solution of Problem 5, using PG2, is depicted in Fig. 4.1. Spurious oscillations are evident in the numerical solution in the region of the parabolic boundary layers. This bears out the lack of uniform convergence in L^∞ and the low uniform convergence rate in L^2. The measured uniform convergence rate of near 1 in L^1 would seem to indicate the inappropriateness of this norm for measuring uniform convergence. However, it does indicate that the scheme has correctly located the parabolic boundary layers along the sides $y = 0$ and $y = 1$ of $\bar{\Omega}$, even though the scheme is unable to solve the problem accurately near these layers (as evidenced by the L^2 and L^∞ convergence rates, which are low in comparison with the corresponding PG2 rates in Table II).

5. CONCLUSIONS

In this paper, we have considered the use of exponentially fitted Galerkin and Petrov–Galerkin methods to solve convection-dominated elliptic problems in two dimensions. We have discussed the suitability of various norms for error measurement in such problems and have obtained numerical results for several new schemes generated using our approach. One of these schemes, which exhibits exceptional accuracy, seems very promising.

REFERENCES

1. K. V. Emel'janov, *Numer. Methods Mech. Cont. Media* **1**, 20 (1970). [Russian]
2. P. A. Farrell, *IMA J. Numer. Anal.* **7**, 459 (1987).
3. A. F. Hegarty, in *BAIL II*, edited by J. J. H. Miller (Boole Press, Dublin, 1982), p. 263.
4. A. M. Il'in, *Mat. Zametki* **6**, 237 (1969); English trans., *Math. Notes* **6**, 596 (1969).
5. C. Johnson, V. Nävert, and I. Pitkäranta, *Comput. Methods Appl. Mech. Eng.* **45**, 285 (1984).
6. C. Johnson, A. H. Schatz, and L. B. Wahlbin, *Math. Comput.* **49**, 25 (1987).
7. R. B. Kellogg and A. Tsan, *Math. Comput* **32**, 1025 (1978).
8. V. D. Liseikin, *Numer. Methods Mech. Cont. Media* **14**, 110 (1983) [Russian].
9. K. Niijima, *Numer. Math.* **56**, 707 (1990).
10. E. O'Riordan and M. Stynes, *Math. Comput.* **57**, 47 (1991).
11. E. O'Riordan and M. Stynes, in *Proceedings, Conf. "Discretization Methods in Singular Perturbations and Flow Problems," Technical University "Otto von Guericke" Magdeburg, Germany, May 1989*, pp. 48–55.
12. H.-G. Roos, *Int. J. Numer. Methods Eng.* **21**, 1459 (1985).
13. G. I. Shishkin, *USSR Comput. Math. Math. Phys.* **26**, 38 (1986).
14. M. Stynes and E. O'Riordan, *Math. Comput.* **56**, 663 (1991).
15. M. Stynes and L. Tobiska, Necessary L^2-uniform convergence conditions for difference schemes for two-dimensional convection-diffusion problems, Preprint 10/92, Mathematics Department, Technical University "Otto von Guericke" Magdeburg (1992).
16. B. Wendroff, *J. SIAM* **8**, 549 (1960).