

A GLOBALLY UNIFORMLY CONVERGENT FINITE ELEMENT METHOD FOR A SINGULARLY PERTURBED ELLIPTIC PROBLEM IN TWO DIMENSIONS

EUGENE O'RIORDAN AND MARTIN STYNES

ABSTRACT. We analyze a new Galerkin finite element method for numerically solving a linear convection-dominated convection-diffusion problem in two dimensions. The method is shown to be convergent, uniformly in the perturbation parameter, of order $h^{1/2}$ in a global energy norm which is stronger than the L^2 norm. This order is optimal in this norm for our choice of trial functions.

1. INTRODUCTION

In this paper, we will examine a finite element method for the numerical solution of the singularly perturbed linear elliptic boundary value problem

$$(1.1a) \quad Lu \equiv -\varepsilon \Delta u + \vec{a} \cdot \nabla u + a_0 u = f \quad \text{on } \Omega \equiv (0, 1) \times (0, 1),$$

$$(1.1b) \quad u = g \quad \text{on } \partial\Omega,$$

$$(1.1c) \quad \vec{a} = (a_1, a_2) > (0, 0), \quad a_0 > 0 \quad \text{on } \bar{\Omega},$$

where ε is a small positive parameter. This problem is often viewed as a basic model of a steady-state convection-diffusion process. For small values of ε , the solution u will in general vary rapidly in a layer region of width $O(\varepsilon \ln(1/\varepsilon))$ at the outflow boundary $\{(x, y) \in \bar{\Omega} \mid x = 1 \text{ or } y = 1\}$.

When solving (1.1) numerically, the dual nature of the solution causes serious difficulties. When ε is small, this elliptic problem is essentially hyperbolic in its behavior outside the layer region; however, in the numerical approximation of elliptic and hyperbolic problems, different approaches are normally adopted.

For small values of ε , it is well known that classical numerical methods for (1.1) produce wild oscillations throughout the whole domain. Various upwind methods have been proposed to eliminate these oscillations and produce a stable numerical solution. The literature on numerical methods for problem (1.1) is extensive, and we will not attempt to give a comprehensive survey here.

Received March 23, 1990.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65N30; Secondary 35B25, 35B45.

The second author's research was partly supported by the Arts Faculty Research Fund, University College, Cork.

In the context of finite elements, the best-known approach is the streamline-diffusion method, which essentially upwinds along the streamlines. Mathematical analyses of different versions of the streamline-diffusion method have been performed by Johnson et al. [8, 9] and Nijjima [11] for various classes of convection-diffusion problems. These results confirm the accuracy of the streamline-diffusion method in *smooth* regions (i.e., away from the layers). However, they do not yield good error estimates for the behavior of the method on the entire domain Ω ; the global bounds obtained depend on Sobolev norms of the solution u , which are large when ε is small.

We are interested in globally uniformly convergent (GUC) numerical methods. These are methods which converge, uniformly in ε , *throughout all of* Ω . More precisely, we mean that if u is the solution of (1.1) and u^h is an approximation obtained using a GUC method, then we have an inequality of the form

$$(1.2) \quad \|u - u^h\|' \leq Ch^p,$$

where $C > 0$ and $p > 0$ are independent of ε and of the mesh width h , and $\|\cdot\|'$ is some appropriate norm, which measures the behavior on *all of* Ω . Examples of such norms are the discrete and continuous $L^p(\Omega)$ norms and the energy norm defined in (3.3) below. The estimates mentioned above for the streamline-diffusion method do not prove it to be a GUC method, as they are uniform in ε only outside layer regions. It is desirable to have GUC methods, because for a given mesh, their accuracy is retained throughout Ω irrespective of the value of ε .

In evaluating the performance of GUC methods, the choice of norm to use is not universally agreed. From an examination of one-dimensional difference schemes, one concludes that the L^1 norm is too weak for problems such as (1.1), as it does not adequately measure behavior in layers (see Hegarty et al. [6] for details). Thus, a stronger norm is desirable, and we will obtain our error estimates in an energy norm (defined in (3.3) below) and in the discrete $L^2(\Omega)$ norm. Roos [14] has given necessary conditions for a numerical method to be GUC, measured in the discrete L^∞ norm, when applied to (1.1). The streamline-diffusion schemes do not satisfy these conditions, but the method examined in this paper does.

Most GUC methods have been obtained for singularly perturbed ordinary differential equations (see, e.g., Doolan, Miller, and Schilders [1], O'Riordan and Stynes [12], and Gartland [3]). The error analyses of these one-dimensional methods required detailed information about the local behavior of the exact solution. In two dimensions, this behavior can be considerably more complicated (see, e.g., Shih and Kellogg [15]).

Finite difference schemes for problem (1.1) which are GUC have been examined by Emel'janov [2], Hegarty [5], and others. These schemes are, in general, variations on Il'in's scheme [7], which is essentially $O(h^{1/2})$ in the discrete $L^\infty(\Omega)$ norm. However, theoretical finite difference convergence results have

been obtained only for schemes which satisfy a discrete maximum principle. In the present paper, a difference scheme is analyzed by finite element techniques and is not required to satisfy a discrete maximum principle. Numerically, our method appears to be only a slight improvement on Il'in's scheme (see Hegarty et al. [6]). However, by employing the same finite element framework as described in this paper, a family of difference schemes have been derived, of which some are significantly more accurate in numerical tests (see Hegarty et al. [6]) than Il'in's scheme.

A conventional Lax-Milgram approach requires the bilinear form to be both coercive and continuous. Although it is easy to obtain coercivity uniformly in ε (see Theorem 3.1), it does not seem possible to organize the relevant norms in such a way as to obtain a satisfactory upper bound for the bilinear form. Consequently, our analysis is considerably different from that of the standard finite element approach. A preliminary version of this analysis, which proved a weaker result, was given in [16].

Our finite element method is shown to be globally uniformly convergent, in an energy norm, when the coefficients \vec{a} in (1.1) are assumed to be bounded away from zero. To get sharp bounds on the solution u of (1.1) and on its derivatives, we also assume that the data is sufficiently smooth and satisfies certain compatibility conditions (see §2). This effectively eliminates the possibility of interior layers, but boundary layers may still appear at the outflow boundary. We believe that the insight gained in the analysis of (1.1) will be helpful in the study of elliptic problems whose solutions exhibit more complex behavior. We know of no other method for problem (1.1) which has been proven to be globally uniformly convergent in an energy norm.

2. THE CONTINUOUS PROBLEM

We begin with a definition of the set $C^{k,\alpha}(\Omega)$.

Definition. For each integer $k \geq 0$, a function $w(x, y)$ is in the set $C^{k,\alpha}(U)$, where $U \subset \mathfrak{R}^2$, if $w(x, y) \in C^k(U)$ and if on U all the derivatives of w (up to and including order k) satisfy a Hölder condition of order $\alpha \in (0, 1)$, viz.,

$$\max_{|\beta| \leq k} \sup_{(x,y), (x',y') \in U} \frac{|D^\beta w(x, y) - D^\beta w(x', y')|}{((x - x')^2 + (y - y')^2)^{\alpha/2}} \leq M$$

for some constant M , where D^β is the usual multi-index notation for derivatives.

In order to guarantee that the solution of problem (1.1) is sufficiently smooth for our purposes, we impose the following conditions on the data:

$$(2.1) \quad a_1, a_2, a_0, f \in C^{0,\alpha}(\overline{\Omega}) \cap C^{2,\alpha}(\Omega),$$

$$(2.2a) \quad g = 0,$$

$$(2.2b) \quad f(0, 0) = f(1, 0) = f(0, 1) = f(1, 1) = 0.$$

Ladyzhenskaya and Ural'tseva [10, p. 111] show that the problem

$$Lu = f \quad \text{on } D, \quad u = g \quad \text{on } \partial D$$

(where D is a region possessing a smooth boundary ∂D), has a solution in $C^{k,\alpha}(\overline{D})$ if the Poisson equation

$$\Delta u = f \quad \text{on } D, \quad u = g \quad \text{on } \partial D,$$

is solvable in $C^{k,\alpha}(\overline{D})$. Volkov [17] gives necessary and sufficient conditions for the solution of

$$\Delta u = f \quad \text{on } \Omega, \quad u = g \quad \text{on } \partial\Omega,$$

to lie in $C^{k,\alpha}(\overline{\Omega})$. Combining Ladyzhenskaya's continuation argument with Volkov's result, we obtain

Theorem 2.1. *If the data in (1.1) satisfies (2.1), (2.2a), and (2.2b), then there exists a unique solution $u(x, y)$ to (1.1), and*

$$u \in C^{2,\alpha}(\overline{\Omega}) \cap C^{4,\alpha}(\Omega).$$

Remark 2.2. There is no loss of generality in imposing homogeneous boundary conditions, since the general inhomogeneous case is easily reduced to this (see, e.g., Ladyzhenskaya and Ural'tseva [10, p. 111]). However, to do this, we need g to be $C^{2,\alpha}$ on the four sides of Ω ; furthermore, at the four corners, the compatibility conditions (2.2b) will involve g and some of its derivatives (see Volkov [17] for details).

The problem satisfies a weak maximum principle:

Lemma 2.3 (Gilbarg and Trudinger [4, Theorem 3.1]). *For all functions $w(x, y) \in C^2(\Omega) \cap C^0(\overline{\Omega})$, if $w \geq 0$ on $\partial\Omega$ and $Lw \geq 0$ on Ω , then $w \geq 0$ on $\overline{\Omega}$.*

For the sake of clarity, we will restrict our attention in this paper to the case where

$$(2.3) \quad \bar{a} \text{ is constant on } \overline{\Omega}.$$

The case of variable \bar{a} has been outlined in [13].

Note. Throughout this paper, we shall use C (sometimes subscripted) to denote a generic positive constant independent of ε and of the mesh. We also assume that conditions (2.1), (2.2), and (2.3) are satisfied. We shall use $u(\cdot, \cdot)$ to denote the solution of (1.1).

Lemma 2.4. *The following estimates hold:*

- (a) $|u(x, y)| \leq C(1 - e^{-2a_1(1-x)/\varepsilon})$ on $\overline{\Omega}$,
- (b) $|u(x, y)| \leq Cx$ on $\overline{\Omega}$.

Proof. (a) Use the barrier function $\phi(x, y) = C(1 - e^{-2a_1(1-x)/\varepsilon})$:

$$\begin{aligned} L(\phi \pm u)(x, y) &= 2a_1^2 C \varepsilon^{-1} e^{-2a_1(1-x)/\varepsilon} + C a_0 (1 - e^{-2a_1(1-x)/\varepsilon}) \pm f \\ &\geq 0 \quad \text{for } C \text{ sufficiently large.} \end{aligned}$$

Thus, $|u| \leq \phi$ by Lemma 2.3.

- (b) Use the barrier function $\phi(x, y) = Cx$. \square

Note that this lemma implies that $\|u\|_\infty \leq C$.

Corollary 2.5. *We have*

(a) $|u_x(1, y)| \leq C\varepsilon^{-1}$ for $0 < y < 1$,

(b) $|u_x(0, y)| \leq C$ for $0 < y < 1$.

Proof. (a) Use

$$\begin{aligned} |u_x(1, y)| &= \left| \lim_{x \rightarrow 1^-} \frac{u(x, y) - u(1, y)}{x - 1} \right| \\ &\leq \lim_{x \rightarrow 1^-} \frac{C(1 - e^{-2a_1(1-x)/\varepsilon})}{1 - x} = C\varepsilon^{-1}. \end{aligned}$$

(b) Similarly,

$$|u_x(0, y)| = \left| \lim_{x \rightarrow 0} \frac{u(x, y) - u(0, y)}{x} \right| \leq C. \quad \square$$

Lemma 2.6. *The following estimates hold:*

(a) $|u_x(x, y)| \leq C(1 + \varepsilon^{-1}e^{-a_1(1-x)/\varepsilon})$ on Ω ,

(b) $|u_y(x, y)| \leq C(1 + \varepsilon^{-1}e^{-a_2(1-y)/\varepsilon})$ on Ω .

Proof. (a) We have $L(u_x) = f_x - (a_0)_x u$, since \vec{a} is constant. Consider the barrier function $\phi(x, y) = C(1 + \varepsilon^{-1}e^{-a_1(1-x)/\varepsilon})$; then

$$L(\phi \pm u_x)(x, y) = Ca_0(1 + \varepsilon^{-1}e^{-a_1(1-x)/\varepsilon}) \pm (f_x - (a_0)_x u) \geq 0$$

for C sufficiently large. Since $u \equiv 0$ on $\partial\Omega$, $u_x(x, 0) = u_x(x, 1) = 0$ for $0 \leq x \leq 1$. Using Corollary 2.5, we obtain $(\phi \pm u_x)(x, y) \geq 0$ on $\partial\Omega$. Now apply Lemma 2.3 to obtain $|u_x| \leq \phi$ on Ω .

(b) follows similarly. \square

Lemma 2.7. *There holds*

(a) $|-\varepsilon u_{xx} + a_1 u_x| \leq C$ on Ω ,

(b) $|-\varepsilon u_{yy} + a_2 u_y| \leq C$ on Ω .

Proof. (a) Let $w \equiv -\varepsilon u_{xx} + a_1 u_x$. On the two sides $y = 0$ and $y = 1$ of Ω , $w \equiv 0$. On the other sides $x = 0$ and $x = 1$ of Ω , $u_{yy} = u_y = u = 0$. Thus from $Lu = f$ and Theorem 2.1, we get $w = f$ on this pair of sides. Hence $|w| \leq C$ on $\partial\Omega$. From Theorem 2.1, $w \in C^{0,\alpha}(\bar{\Omega}) \cap C^{2,\alpha}(\Omega)$, so we can apply Lemma 2.3 to w . We have

$$\begin{aligned} Lw &= -\varepsilon \Delta w + a_1 w_x + a_2 w_y + a_0 w \\ &= -\varepsilon (Lu)_{xx} + a_1 (Lu)_x + \varepsilon (a_0)_{xx} u + 2\varepsilon (a_0)_x u_x - a_1 (a_0)_x u \\ &= -\varepsilon f_{xx} + a_1 f_x + \varepsilon [(a_0)_{xx} u + 2(a_0)_x u_x] - a_1 (a_0)_x u. \end{aligned}$$

Hence, $|Lw| \leq C$, using Lemma 2.6 and $|u| \leq C$. Use the barrier function $\phi = C$ to finish (note that $a_0 > 0$).

(b) We have $-\varepsilon u_{yy} + a_2 u_y = f + \varepsilon u_{xx} - a_1 u_x$. Using (a), we are done. \square

Remark 2.8. For our main convergence result, Theorem 3.2, the only bounds we need from this section are

$$\begin{aligned} \|u\|_\infty &\leq C, \\ \int_0^1 |u_x(x, y)| dx &\leq C \quad \text{for all } y, \\ \int_0^1 |u_y(x, y)| dy &\leq C \quad \text{for all } x, \\ |-\varepsilon u_{xx} + a_1 u_x| &\leq C \quad \text{in } \Omega, \\ |-\varepsilon u_{yy} + a_2 u_y| &\leq C \quad \text{in } \Omega. \end{aligned}$$

Thus, our arguments above may assume more differentiability of $u(x, y)$ on $\overline{\Omega}$ than is necessary in practice. Further evidence of this has been supplied by numerical experiments; see Remark 3.3.

3. DISCRETIZATION OF THE CONTINUOUS PROBLEM

A weak formulation of problem (1.1), with homogeneous boundary conditions, is:

Find $u \in H_0^1(\Omega)$ such that

$$(3.1) \quad \begin{aligned} B(u, v) &\equiv (\varepsilon u_x, v_x) + (\varepsilon u_y, v_y) + (\vec{a} \cdot \nabla u + a_0 u, v) \\ &= (f, v) \quad \text{for all } v \in H_0^1(\Omega), \end{aligned}$$

where (\cdot, \cdot) denotes the usual $L^2(\Omega)$ inner product.

We will discretize this weak form by means of a finite element method. Let N and M be two positive integers. Let $h = 1/(N+1)$, and set $x_i = ih$ for $i = 0, 1, \dots, N+1$. Let $k = 1/(M+1)$, and set $y_j = jk$ for $j = 0, 1, \dots, M+1$. We define a set of trial functions $\{\phi^{i,j}(x, y) : i = 1, \dots, N; j = 1, \dots, M\}$. The trial (test) space is the linear span of these trial (test) functions and will be denoted by S (T). Let

$$u^h(x, y) = \sum_{i,j} u_{i,j}^h \phi^{i,j}(x, y)$$

be our finite element approximation to $u(x, y)$.

The nodal values $\{u_{i,j}^h\}$ are determined from

$$(3.2) \quad \overline{B}(u^h, \psi) = (\overline{f}, \psi) \quad \text{for all } \psi \in T,$$

where

$$\overline{B}(v, w) \equiv \varepsilon(\nabla v \cdot \nabla w, 1) + (\overline{a}_1 v_x + \overline{a}_2 v_y, w) + hk \sum_{i=1}^N \sum_{j=1}^M (a_0 v w)_{i,j}$$

and \overline{a}_1 , \overline{a}_2 , and \overline{f} are piecewise constant approximations to a_1 , a_2 , and f , respectively. Since we are interested in the case of constant \vec{a} , we simply have

$\bar{a}_1 \equiv a_1$ and $\bar{a}_2 \equiv a_2$. We will use the notation $z_{i,j} \equiv z(x_i, y_j)$, for various functions $z(\cdot, \cdot)$.

We define the energy norm $||| \cdot |||$ by

$$(3.3) \quad |||v|||^2 \equiv \varepsilon(v_x, v_x) + \varepsilon(v_y, v_y) + hk \sum_{i,j} v_{i,j}^2 \quad \forall v \in H_0^1(\Omega) \cap C_0(\Omega).$$

Theorem 3.1. *For all $v \in H_0^1(\Omega) \cap C_0(\Omega)$, we have $\bar{B}(v, v) \geq C_1 |||v|||^2$.*

Proof. Since \bar{a} is constant, $(a_1 v_x, v) = (a_1, (v^2/2)_x) = 0$. Thus, $\bar{B}(v, v) = \varepsilon(v_x, v_x) + \varepsilon(v_y, v_y) + hk \sum_{i,j} (a_0 v^2)_{i,j}$, and the result follows from $a_0 > 0$. \square

Non-self-adjoint singularly perturbed problems such as (1.1) are usually solved using Petrov-Galerkin methods (i.e., one chooses $S \neq T$). However, in this paper, we will examine a Galerkin method (i.e., $S = T$), where the trial (and test) functions are the tensor product of one-dimensional exponential trial functions. That is, for $i = 1, \dots, N$ and $j = 1, \dots, M$,

$$(3.4a) \quad \phi^{i,j}(x, y) = \phi^i(x)\phi_j(y),$$

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

$$(3.4b) \quad -\varepsilon(\phi^i)_{xx} + a_1(\phi^i)_x = 0 \quad \text{on } (0, 1) \setminus \{x_1, \dots, x_N\}, \quad \phi^i(x_j) = \delta_{i,j},$$

and each $\phi_j(y)$ satisfies

$$(3.4c) \quad -\varepsilon(\phi_j)_{yy} + a_2(\phi_j)_y = 0 \quad \text{on } (0, 1) \setminus \{y_1, \dots, y_M\}, \quad \phi_j(y_i) = \delta_{i,j}.$$

Note that $\iint_{\Omega} \phi^{i,j} dx dy = hk$ for each (i, j) .

Our Galerkin finite element approximation u^h satisfies

$$(3.5) \quad \bar{B}(u^h, \phi^{i,j}) = (\bar{f}, \phi^{i,j}) \quad \text{for all } \phi^{i,j} \in S,$$

where \bar{f} is chosen so that $|f - \bar{f}| \leq C(h+k)$ in Ω . From Theorem 3.1, it follows that u^h is defined uniquely by (3.5).

Remark. A possible choice for \bar{f} is obtained by taking $(\bar{f}, \phi^{i,j}) \equiv hk f_{i,j}$. The finite difference scheme (3.5) for this choice of \bar{f} is written out explicitly in Hegarty et al. [6]. Various choices for the piecewise constant functions \bar{a}_1, \bar{a}_2 (in the case of variable \bar{a}), and \bar{f} are also examined in [6]. When these piecewise constant functions are specified, all the integrals in (3.5) can be evaluated exactly, so effectively we are specifying a quadrature rule.

Let $u_I = \sum_{i,j} u(x_i, y_j) \phi^{i,j}(x, y)$ be the function in S which interpolates to u at the nodes. From Theorem 3.1,

$$(3.6) \quad C_1 |||u - u^h|||^2 \leq \bar{B}(u - u^h, u - u^h) = \bar{B}(u - u^h, u - u_I) + \bar{B}(u - u^h, u_I - u^h).$$

In the next two sections we will show (Propositions 4.2 and 5.6) that

$$(3.7) \quad \bar{B}(u - u^h, u_I - u^h) \leq (C_1 hk/2) \sum_{i,j} (u - u^h)_{i,j}^2 + C(h+k)$$

and

$$(3.8) \quad \overline{B}(u - u^h, u - u_I) \leq Ch + (C_1/4) \| \| u - u^h \| \|^2$$

provided that $C_2 \leq k/h \leq C_3$.

Combining (3.6), (3.7), and (3.8) shows that the Galerkin method described above is uniformly convergent in ε in the energy norm $\| \| \cdot \| \|$. More precisely, we have

Theorem 3.2. *Let u^h be as in (3.5). If $C_2 \leq k/h \leq C_3$, then $\| \| u - u^h \| \| \leq Ch^{1/2}$.*

Remark 3.3. Combining Lemma 5.4 below with Theorem 3.2 shows that $\| \| u - u^h \| \|_{L^2(\Omega)} \leq Ch^{1/2}$. We have observed this rate of convergence in several numerical experiments. See Hegarty et al. [6] for details. This indicates that the above error estimate is sharp. The same rate of convergence has also been observed in many problems which do not satisfy condition (2.2b).

Remark 3.4. Consider the one-dimensional problem $- \varepsilon u'' + au' = f$ on $[0, 1]$, where the functions a and f are constant and $u(0) = u(1) = 0$. Using the exponential elements (3.4b), it can be shown that, for the analogue of (3.5) applied to this problem, $\| \| u - u_I \| \|' = (f/a)\varepsilon^{1/2}(\rho \coth \rho - 1)^{1/2}$, where $\| \| \cdot \| \|'$ is the one-dimensional analogue of $\| \| \cdot \| \|$ and $\rho \equiv ah/(2\varepsilon)$. This implies that, uniformly in ε , $\| \| u - u_I \| \|$ is $O(h^{1/2})$ but not $O(h^\alpha)$ for any $\alpha > 1/2$. Hence, the optimal order of accuracy attainable using the trial functions (3.4a) is achieved in Theorem 3.2.

4. QUADRATURE ERROR

In this section we will derive (3.7). We have

$$\begin{aligned} \overline{B}(u - u^h, u_I - u^h) &= \overline{B}(u, u_I - u^h) - \overline{B}(u^h, u_I - u^h) \\ &= \overline{B}(u, u_I - u^h) - B(u, u_I - u^h) + (f - \overline{f}, u_I - u^h) \\ &\hspace{15em} \text{(from } Lu = f \text{ and (3.5))} \\ &= hk \sum_{i,j} (a_0 u(u_I - u^h))_{i,j} - (a_0 u, u_I - u^h) \\ &\hspace{15em} + (f - \overline{f}, u_I - u^h). \end{aligned}$$

Since $(u)_{i,j} = (u_I)_{i,j}$, we get

$$(4.1) \quad \begin{aligned} &\overline{B}(u - u^h, u_I - u^h) \\ &= \sum_{i=1}^N \sum_{j=1}^M (u - u^h)_{i,j} \{ hk(a_0 u)_{i,j} - (a_0 u, \phi^{i,j}) + (f - \overline{f}, \phi^{i,j}) \}. \end{aligned}$$

The bounding of (4.1) is organized into the following two results.

Lemma 4.1. For each i and j ,

$$\begin{aligned} & |hk(a_0u)_{i,j} - (a_0u, \phi^{i,j}) + (f - \bar{f}, \phi^{i,j})| \\ & \leq Ch \iint_{\Omega_{i,j}} |u_x| + Ck \iint_{\Omega_{i,j}} |u_y| + Chk(h+k), \end{aligned}$$

where $\Omega_{i,j} \equiv (x_{i-1}, x_{i+1}) \times (y_{j-1}, y_{j+1})$.

Proof. Since $|f - \bar{f}| \leq C(h+k)$, we have

$$|(f - \bar{f}, \phi^{i,j})| \leq C(h+k)(1, \phi^{i,j}) = Chk(h+k).$$

To bound the remaining terms, we will simply use integration by parts appropriately. Let $w(x, y) \equiv (a_0u)_{i,j} - (a_0u)(x, y)$. Note that $w_{i,j} = 0$. Now,

$$\begin{aligned} hk(a_0u)_{i,j} - (a_0u, \phi^{i,j}) &= (w, \phi^{i,j}) \quad (\text{since } (1, \phi^{i,j}) = hk) \\ &= \int_{y_{j-1}}^{y_{j+1}} \left\{ \int_{x_{i-1}}^{x_i} w \phi^{i,j} dx + \int_{x_i}^{x_{i+1}} w \phi^{i,j} dx \right\} dy \\ &= \int_{y_{j-1}}^{y_{j+1}} \left\{ w(x_i, y) \int_{t=x_{i-1}}^{x_{i+1}} \phi^{i,j}(t, y) dt \right. \\ &\quad \left. - \int_{x=x_{i-1}}^{x_i} w_x(x, y) \int_{t=x_{i-1}}^x \phi^{i,j}(t, y) dt dx \right. \\ &\quad \left. + \int_{x=x_i}^{x_{i+1}} w_x(x, y) \int_{t=x}^{x_{i+1}} \phi^{i,j}(t, y) dt dx \right\} dy, \end{aligned}$$

using integration by parts. The first term in the above right-hand side is

$$\begin{aligned} & \int_{t=x_{i-1}}^{x_{i+1}} \left\{ \int_{y=y_{j-1}}^{y_i} w(x_i, y) \phi^{i,j}(t, y) dy + \int_{y=y_j}^{y_{j+1}} w(x_i, y) \phi^{i,j}(t, y) dy \right\} dt \\ &= \int_{t=x_{i-1}}^{x_{i+1}} \left\{ - \int_{y=y_{j-1}}^{y_j} w_y(x_i, y) \int_{s=y_{j-1}}^y \phi^{i,j}(t, s) ds dy \right. \\ &\quad \left. + \int_{y=y_j}^{y_{j+1}} w_y(x_i, y) \int_{s=y}^{y_{j+1}} \phi^{i,j}(t, s) ds dy \right\} dt, \end{aligned}$$

using $w_{i,j} = 0$. Now $|\phi^{i,j}| \leq 1$. Hence,

$$\begin{aligned} |hk(a_0u)_{i,j} - (a_0u, \phi^{i,j})| &\leq Ch \iint_{\Omega_{i,j}} |w_x| + Ck \iint_{\Omega_{i,j}} |w_y| \\ &= Ch \iint_{\Omega_{i,j}} |(a_0u)_x| + Ck \iint_{\Omega_{i,j}} |(a_0u)_y| \\ &\leq Chk(h+k) + Ch \iint_{\Omega_{i,j}} |u_x| + Ck \iint_{\Omega_{i,j}} |u_y|. \quad \square \end{aligned}$$

Proposition 4.2. *We have*

$$|\overline{B}(u - u^h, u_I - u^h)| \leq (C_1 hk/2) \sum_{i=1}^N \sum_{j=1}^M |u - u^h|_{i,j}^2 + C(h + k),$$

where C_1 is the constant appearing in Theorem 3.1.

Proof. For each fixed y , let $\theta_i^x(y) \equiv \int_{x_{i-1}}^{x_{i+1}} |u_x(x, y)| dx$. Then

$$\sum_{i=1}^N \theta_i^x(y) \leq 2 \int_{x=0}^1 |u_x(x, y)| dx \leq C,$$

from Lemma 2.6. Analogously define $\theta_j^y(x)$. From (4.1) and Lemma 4.1, we have

$$\begin{aligned} & |\overline{B}(u - u^h, u_I - u^h)| \\ & \leq C \sum_{i,j} |u - u^h|_{i,j} \left\{ h \int_{y=y_{j-1}}^{y_{j+1}} \theta_i^x(y) dy + k \int_{x=x_{i-1}}^{x_{i+1}} \theta_j^y(x) dx \right. \\ & \qquad \qquad \qquad \left. + Chk(h + k) \right\}. \end{aligned}$$

Using the arithmetic-geometric mean inequality,

$$\begin{aligned} & \sum_{i,j} |u - u^h|_{i,j} \int_{y=y_{j-1}}^{y_{j+1}} \theta_i^x(y) dy \\ & \leq (C_1 k/6) \sum_{i,j} |u - u^h|_{i,j}^2 + Ck^{-1} \sum_{i,j} \left(\int_{y_{j-1}}^{y_{j+1}} \theta_i^x(y) dy \right)^2 \\ & \leq (C_1 k/6) \sum_{i,j} |u - u^h|_{i,j}^2 \\ & \quad + Ck^{-1} \sum_{i,j} \left(\int_{y=y_{j-1}}^{y_{j+1}} 1^2 dy \right) \left(\int_{y=y_{j-1}}^{y_{j+1}} (\theta_i^x(y))^2 dy \right) \\ & \leq (C_1 k/6) \sum_{i,j} |u - u^h|_{i,j}^2 + C \sum_{j=1}^M \int_{y_{j-1}}^{y_{j+1}} \sum_{i=1}^N (\theta_i^x(y))^2 dy. \end{aligned}$$

Since $\theta_i^x(y) \geq 0$, we have $\sum_{i=1}^N (\theta_i^x(y))^2 \leq (\sum_{i=1}^N \theta_i^x(y))^2 \leq C$. Thus,

$$h \sum_{i,j} |u - u^h|_{i,j} \left(\int_{y=y_{j-1}}^{y_{j+1}} \theta_i^x(y) dy \right) \leq (C_1 hk/6) \sum_{i,j} |u - u^h|_{i,j}^2 + Ch.$$

Dealing similarly with the other terms yields

$$|\overline{B}(u - u^h, u_I - u^h)| \leq (C_1 hk/2) \sum_{i,j} |u - u^h|_{i,j}^2 + C(h + k). \quad \square$$

Remark 4.3. The only properties of the trial functions needed to attain the bounds in this section are

$$\begin{aligned}\phi^{i,j}(x, y) &= \phi^i(x)\phi_j(y), \\ |\phi^i(x)| &\leq 1, \quad |\phi_j(y)| \leq 1, \\ \int_{x=x_{i-1}}^{x_{i+1}} \phi^i dx &= h \quad \text{and} \quad \int_{y=y_{j-1}}^{y_{j+1}} \phi_j dy = k.\end{aligned}$$

5. INTERPOLATION ERROR

In this section we will derive (3.8). We initially bound the error $u - u_I$.

Lemma 5.1. *If the trial functions are chosen as in (3.4), then*

- (a) $|(u - u_I)(x, y_j)| \leq C(x - x_{i-1})(1 - e^{-a_1(x_i - x)/\varepsilon})$ on each line segment $[x_{i-1}, x_i] \times \{y_j\}$,
- (b) $|(u - u_I)(x_i, y)| \leq C(y - y_{j-1})(1 - e^{-a_2(y_j - y)/\varepsilon})$ on each line segment $\{x_i\} \times [y_{j-1}, y_j]$,
- (c) $\max_{\bar{Q}} |(u - u_I)(x, y)| \leq Ch(1 - e^{-a_1 h/\varepsilon}) + Ck(1 - e^{-a_2 k/\varepsilon})$.

Proof. Fix i and j .

(a) We will confine ourselves to the line segment $[x_{i-1}, x_i] \times \{y_j\}$. Set $M_1 z \equiv -\varepsilon z_{xx} + a_1 z_x$. Then M_1 satisfies a maximum principle on $[x_{i-1}, x_i] \times \{y_j\}$. On $(x_{i-1}, x_i) \times \{y_j\}$,

$$M_1(u - u_I) = -\varepsilon u_{xx} + a_1 u_x \leq C \quad \text{by Lemma 2.7.}$$

Use the barrier function $\phi(x) = C(x - x_{i-1})(1 - e^{-a_1(x_i - x)/\varepsilon})$. Then

$$M_1\phi = Ca_1(1 + e^{-a_1(x_i - x)/\varepsilon}) > |M_1(u - u_I)|$$

for C sufficiently large. Hence, $|u - u_I| \leq \phi$ on $[x_{i-1}, x_i] \times \{y_j\}$.

(b) is proven similarly.

(c) We work on the square $\bar{Q}_{i,j} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$. On the open square $Q_{i,j}$,

$$|L(u - u_I)| = |f - a_0 u_I| \leq C.$$

We employ the barrier function

$$\phi(x, y) = C_1(x - x_{i-1})(1 - e^{-a_1(x_i - x)/\varepsilon}) + C_2(y - y_{j-1})(1 - e^{-a_2(y_j - y)/\varepsilon}),$$

where C_1 and C_2 are chosen sufficiently large. By (a) and (b) above, $\phi \geq |u - u_I|$ on $\partial Q_{i,j}$. We also have

$$\begin{aligned}L\phi &= a_0\phi + C_1 a_1(1 + e^{-a_1(x_i - x)/\varepsilon}) + C_2 a_2(1 + e^{-a_2(y_j - y)/\varepsilon}) \\ &\geq |L(u - u_I)|.\end{aligned}$$

Since L satisfies a maximum principle, $|u - u_I| \leq \phi$ on $\bar{Q}_{i,j}$. \square

Theorem 5.2. *If the trial functions are chosen as in (3.4), and $C_2 \leq k/h \leq C_3$, then $\|u - u_I\| \leq Ch^{1/2}$.*

Proof. From Theorem 3.1,

$$\begin{aligned} C_1 \|u - u_I\|^2 &\leq \overline{B}(u - u_I, u - u_I) \\ &= \varepsilon(\nabla(u - u_I) \cdot \nabla(u - u_I), 1) + (\vec{d} \cdot \nabla(u - u_I), u - u_I). \end{aligned}$$

Integrating by parts,

$$\begin{aligned} C_1 \|u - u_I\|^2 &\leq (-\varepsilon\Delta(u - u_I) + \vec{d} \cdot \nabla(u - u_I), u - u_I)^\wedge \\ &\quad + \varepsilon \int_{y=0}^1 \sum_{i=1}^N (u - u_I)(x_i, y) \mathcal{F}_{x_i}((u_I)_x)(y) dy \\ &\quad + \varepsilon \int_{x=0}^1 \sum_{j=1}^M (u - u_I)(x, y_j) \mathcal{F}_{y_j}((u_I)_y)(x) dx, \end{aligned}$$

since $u \in C^2(\Omega)$, where $\widehat{}$ denotes that integration is over $\bigcup_{i,j} Q_{i,j}$ (defined in the proof of Lemma 5.1) and

$$\begin{aligned} \mathcal{F}_{x_i}(w)(y) &\equiv w(x_i^+, y) - w(x_i^-, y), \\ \mathcal{F}_{y_j}(w)(x) &\equiv w(x, y_j^+) - w(x, y_j^-). \end{aligned}$$

By our choice of trial functions,

$$\begin{aligned} &|(-\varepsilon\Delta(u - u_I) + \vec{d} \cdot \nabla(u - u_I), u - u_I)^\wedge| \\ &= |(f - a_0 u, u - u_I)^\wedge| \leq C \|u - u_I\|_\infty \\ &\leq Ch \quad (\text{by Lemma 5.1, using } C_2 \leq k/h \leq C_3). \end{aligned}$$

For $(x, y) \in (x_{i-1}, x_i) \times (0, 1)$ and all $v \in S$,

$$v_x = (v(x_i, y) - v(x_{i-1}, y)) \phi_x^i(x).$$

Now, setting $\rho_1 = a_1 h / \varepsilon$,

$$\begin{aligned} \varepsilon |\mathcal{F}_{x_i}((u_I)_x)(y)| &= \varepsilon |(u_I(x_{i+1}, y) - u_I(x_i, y)) \phi_x^{i+1}(x_i^+) \\ &\quad - (u_I(x_i, y) - u_I(x_{i-1}, y)) \phi_x^i(x_i^-)| \\ &= a_1 (1 - e^{-\rho_1})^{-1} |(u_I(x_{i+1}, y) - u_I(x_i, y)) e^{-\rho_1} \\ &\quad - (u_I(x_i, y) - u_I(x_{i-1}, y))|, \end{aligned}$$

since $\varepsilon \phi_x^{i+1}(x_i^+) = a_1 e^{-\rho_1} / (1 - e^{-\rho_1})$ and $\varepsilon \phi_x^i(x_i^-) = a_1 / (1 - e^{-\rho_1})$, as can be computed from the definition of $\phi^i(x)$. Hence,

$$\begin{aligned} \varepsilon |\mathcal{F}_{x_i}((u_I)_x)(y)| &\leq a_1 (1 - e^{-\rho_1})^{-1} (|u_I(x_{i+1}, y) - u_I(x_i, y)| \\ &\quad + |u_I(x_i, y) - u_I(x_{i-1}, y)|). \end{aligned}$$

Note also that for $y \in (y_{j-1}, y_j)$,

$$\begin{aligned} u_I(x_i, y) &= u_I(x_i, y_j)\phi_j(y) + u_I(x_i, y_{j-1})\phi_{j-1}(y) \\ &= u_{i,j}\phi_j(y) + u_{i,j-1}\phi_{j-1}(y). \end{aligned}$$

Thus, using $|\phi_j| \leq 1$ for all j ,

$$\begin{aligned} & \left| \varepsilon \int_{y=0}^1 \sum_{i=1}^N (u - u_I)(x_i, y) \mathcal{F}_{x_i}((u_I)_x) dy \right| \\ & \leq C \|u - u_I\|_\infty (1 - e^{-\rho_1})^{-1} \max_j \sum_{i=1}^{N+1} |u_{i,j} - u_{i-1,j}| \\ & = C(1 - e^{-\rho_1})^{-1} \|u - u_I\|_\infty \max_j \sum_i \left| \int_{x_{i-1}}^{x_i} u_x(x, y_j) dx \right| \\ & \leq C(1 - e^{-\rho_1})^{-1} \|u - u_I\|_\infty \max_j \int_0^1 |u_x(x, y_j)| dx \\ & \leq C(1 - e^{-\rho_1})^{-1} \|u - u_I\|_\infty \quad (\text{by Lemma 2.6}) \\ & \leq Ch \quad (\text{by Lemma 5.1, since } C_2 \leq k/h \leq C_3). \end{aligned}$$

It follows that $C_1 \|u - u_I\|^2 \leq Ch$. \square

Remark 5.3. In the classical case $\varepsilon = 1$, an inspection of the above proof shows that by writing

$$\begin{aligned} & (u_I(x_{i+1}, y) - u_I(x_i, y))e^{-\rho_1} - (u_I(x_i, y) - u_I(x_{i-1}, y)) \\ & = -(u_I(x_{i+1}, y) - u_I(x_i, y))(1 - e^{-\rho_1}) + O(h^2) \end{aligned}$$

we can improve the above result to

$$\|u - u_I\|^2 \leq C \|u - u_I\|_\infty \leq Ch^2.$$

We also require suitable bounds for $u^h(x, y)$ and its first derivatives. First we define the continuous and discrete L^2 norms

$$(5.1a) \quad \|v\|^2 = (v, v) \quad \text{for all } v \in L^2(\Omega),$$

$$(5.1b) \quad \|v\|_d^2 = hk \sum_{i=1}^N \sum_{j=1}^M v_{i,j}^2 \quad \text{for all } v \in S.$$

A useful relationship between these norms is given by

Lemma 5.4. *We have $\|v\| \leq C \|v\|_d$ for all $v \in S$.*

Proof. Let $v \in S$. Then $v = \sum_{i,j} v_{i,j} \phi^{i,j}$, so

$$\begin{aligned} \|v\|^2 &= \sum_{i=1}^{N+1} \sum_{j=1}^{M+1} \int_{x_{i-1}}^{x_i} \int_{y_{j-1}}^{y_j} [v_{i-1,j-1} \phi^{i-1,j-1} + v_{i-1,j} \phi^{i-1,j} \\ & \quad + v_{i,j-1} \phi^{i,j-1} + v_{i,j} \phi^{i,j}]^2 dx dy. \end{aligned}$$

Note that $0 \leq \phi^{i,j} \leq 1$ for all i and j and $(a+b)^2 \leq 2(a^2+b^2)$. Consequently,

$$\|v\|^2 \leq 16hk \sum_{i,j} v_{i,j}^2 = 16\|v\|_d^2. \quad \square$$

The next lemma relates the L^1 and L^2 norms of the derivatives of our trial functions.

Lemma 5.5. *If the trial functions are chosen as in (3.4), then*

$$\int_0^1 \int_0^1 |v_x| dx dy \leq Ch^{-1/2} (1 - e^{-\rho_1})^{1/2} \varepsilon^{1/2} \|v_x\| \quad \text{for all } v \in S,$$

where $\rho_1 = a_1 h / \varepsilon$.

Proof. For $(x, y) \in (x_{i-1}, x_i) \times (0, 1)$ and $v \in S$,

$$(5.2) \quad v_x(x, y) = (v(x_i, y) - v(x_{i-1}, y)) \phi_x^i(x).$$

We then have

$$\begin{aligned} \int_0^1 \int_0^1 |v_x| dx dy &= \sum_{i=1}^{N+1} \left(\int_{x_{i-1}}^{x_i} |\phi_x^i| dx \right) \left(\int_{y=0}^1 |v(x_i, y) - v(x_{i-1}, y)| dy \right) \\ &= \sum_{i=1}^{N+1} \left(\int_{y=0}^1 |v(x_i, y) - v(x_{i-1}, y)| dy \right) \\ &\leq \left(\sum_{i=1}^{N+1} 1 \right)^{1/2} \left(\sum_{i=1}^{N+1} \left(\int_{y=0}^1 |v(x_i, y) - v(x_{i-1}, y)| dy \right)^2 \right)^{1/2} \\ &\quad \text{(using the discrete Cauchy-Schwarz inequality)} \\ &\leq h^{-1/2} \left(\sum_{i=1}^{N+1} \int_{y=0}^1 |v(x_i, y) - v(x_{i-1}, y)|^2 dy \right)^{1/2} \\ &= h^{-1/2} (2\varepsilon/a_1)^{1/2} (1 - e^{-\rho_1})^{1/2} (1 + e^{-\rho_1})^{-1/2} \\ &\quad \times \left(\sum_{i=1}^{N+1} \int_{x_{i-1}}^{x_i} (\phi_x^i)^2 dx \int_{y=0}^1 |v(x_i, y) - v(x_{i-1}, y)|^2 dy \right)^{1/2}, \end{aligned}$$

since for each i ,

$$\varepsilon \int_{x_{i-1}}^{x_i} (\phi_x^i)^2 dx = (a_1/2)(1 + e^{-\rho_1})(1 - e^{-\rho_1})^{-1}.$$

The result follows, using (5.2). \square

We are now ready to derive (3.8).

Proposition 5.6. *If the trial and test functions are chosen as in (3.4), and $C_2 \leq k/h \leq C_3$, then*

$$\bar{B}(u - u^h, u - u_I) \leq Ch + (C_1/4) \|u - u^h\|^2.$$

Proof. We have

$$\bar{B}(u - u^h, u - u_I) = (\varepsilon \nabla(u - u^h) \cdot \nabla(u - u_I), 1) + (\bar{a} \cdot \nabla(u - u^h), u - u_I).$$

We will bound these two terms separately:

$$\begin{aligned} (\varepsilon \nabla(u - u^h) \cdot \nabla(u - u_I), 1) &\leq (\varepsilon^{1/2} \|\nabla(u - u^h)\|)(\varepsilon^{1/2} \|\nabla(u - u_I)\|) \\ &\leq Ch^{1/2}(\varepsilon^{1/2} \|\nabla(u - u^h)\|) \quad (\text{by Theorem 5.2}) \\ &\leq Ch + (C_1/12)\varepsilon \|\nabla(u - u^h)\|^2. \end{aligned}$$

Also,

$$\begin{aligned} |(a_1(u - u^h)_x, u - u_I)| &= |(a_1(u - u_I)_x, u - u_I) + (a_1(u_I - u^h)_x, u - u_I)| \\ &= |(a_1(u_I - u^h)_x, u - u_I)| \\ &\leq C \|u - u_I\|_\infty \| (u_I - u^h)_x \|_{L^1(\Omega)} \\ &\leq Ch^{1/2} (1 - e^{-\rho_1})^{3/2} \varepsilon^{1/2} \|\nabla(u_I - u^h)\| \\ &\quad (\text{by Lemmas 5.1 and 5.5}) \\ &\leq Ch^{1/2} \varepsilon^{1/2} \|\nabla(u_I - u^h)\| \\ &\leq Ch^{1/2} \varepsilon^{1/2} (\|\nabla(u - u_I)\| + \|\nabla(u - u^h)\|) \\ &\leq Ch + (C_1/12)\varepsilon \|\nabla(u - u^h)\|^2, \end{aligned}$$

using Theorem 5.2 and the arithmetic-geometric mean inequality.

The term $(a_2(u - u^h)_y, u - u_I)$ is bounded in a similar manner. \square

This completes the proof of Theorem 3.2

BIBLIOGRAPHY

1. E. P. Doolan, J. J. H. Miller, and W. H. A. Schilders, *Uniform numerical methods for problems with initial and boundary layers*, Boole Press, Dublin, 1980.
2. K. V. Emel'janov, *A difference scheme for a three dimensional elliptic equation with a small parameter multiplying the highest derivative*, Boundary Value Problems for Equations of Mathematical Physics, Ural Scientific Center, U.S.S.R. Academy of Sciences, 1973, pp. 30–42.
3. E. C. Gartland, Jr., *Uniform high-order difference schemes for a singularly perturbed two-point boundary value problem*, Math. Comp. **48** (1987), 551–564.
4. D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order*, Springer, Berlin, 1977.
5. A. F. Hegarty, *Analysis of finite difference methods for two-dimensional elliptic singular perturbation problems*, Ph.D. thesis, Trinity College, Dublin, 1986.
6. A. F. Hegarty, E. O'Riordan, and M. Stynes, *A comparison of uniformly convergent difference schemes for two-dimensional convection-diffusion problems* (submitted for publication).
7. A. M. Il'in, *Differencing scheme for a differential equation with a small parameter affecting the highest derivative*, Mat. Zametki **6** (1969), 237–248; English transl., Math. Notes **6** (1969), 596–602.
8. C. Johnson, V. Nävert, and I. Pitkäranta, *Finite element methods for linear hyperbolic problems*, Comput. Methods Appl. Mech. Engrg. **45** (1984), 285–312.

9. C. Johnson, A. H. Schatz, and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, *Math. Comp.* **49** (1987), 25–38.
10. O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations*, Academic Press, New York, 1968.
11. K. Nijjima, *Pointwise error estimates for a streamline diffusion finite element scheme*, *Numer. Math.* **56** (1990), 707–719.
12. E. O'Riordan and M. Stynes, *Analysis of difference schemes for singularly perturbed differential equations using a discretized Green's function*, BAIL IV, Proc. Fourth Internat. Conf. on Interior and Boundary Layers (J. J. H. Miller, ed.), Boole Press, Dublin, 1986, pp. 157–168.
13. ———, *A uniformly convergent finite difference scheme for an elliptic singular perturbation problem*, *Discretization Methods of Singular Perturbations and Flow Problems* (L. Tobiska, ed.), Technical University "Otto von Guericke" Magdeburg, 1989, pp. 48–55.
14. H.-G. Roos, *Necessary convergence conditions for upwind schemes in the two-dimensional case*, *Internat. J. Numer. Methods Engrg.* **21** (1985), 1459–1469.
15. S. D. Shih and R. B. Kellogg, *Asymptotic analysis of a singular perturbation problem*, *SIAM J. Math. Anal.* **18** (1987), 1467–1511.
16. M. Stynes and E. O'Riordan, *Finite element methods for elliptic convection-diffusion problems*, BAIL V, Proc. Fifth Internat. Conf. on Interior and Boundary Layers (J. J. H. Miller, ed.), Boole Press, Dublin, 1988, pp. 65–76.
17. E. A. Volkov, *Differentiability properties of solutions of boundary value problems for the Laplace and Poisson equations*, *Proc. Steklov Inst. Math.* **77** (1965), pp. 101–126.

DEPARTMENT OF MATHEMATICS, REGIONAL TECHNICAL COLLEGE, DUNDALK, IRELAND

DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE, CORK, IRELAND

E-mail address: stynes@iruccvax.bitnet