INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 7, Number 3, Pages 477–490

© 2010 Institute for Scientific
Computing and Information

AN ENRICHED SUBSPACE FINITE ELEMENT METHOD FOR CONVECTION-DIFFUSION PROBLEMS

R. BRUCE KELLOGG AND CHRISTOS XENOPHONTOS

This paper is dedicated to G.I. Shishkin on the occasion of his 70th birthday

Abstract. We consider a one-dimensional convection-diffusion boundary value problem, whose solution contains a boundary layer at the outflow boundary, and construct a finite element method for its approximation. The finite element space consists of piecewise polynomials on a uniform mesh but is enriched by a finite number of functions that represent the boundary layer behavior. We show that this method converges at the optimal rate, independently of the singular perturbation parameter, when the error is measured in the energy norm associated with the problem. Numerical results confirming the theory are also presented, which also suggest that in the case of variable coefficients, the number of enrichment functions need not be as high as the theory suggests.

Key Words. finite element method, boundary layers, enriched subspace.

1. Introduction

Let $p > 0$, $q > 0$ be smooth functions, let $\varepsilon \in (0,1]$, and consider the problem

$$Lu := -\varepsilon u'' + p(x)u' + q(x)u = f(x) \text{ in } (0,1),\ u(0) = u(1) = 0. \tag{1}$$

It is well-known that the solution to this problem has a boundary layer at $x = 1$, and that an accurate, robust numerical solution can be obtained by putting a highly refined mesh, often called the "Shishkin mesh", near this boundary point [7, 9] ; see also [8] for other mesh choices used in conjunction with the high order p and hp versions of the finite element method (FEM). In this paper we suggest an alternate way to obtain an accurate and robust numerical method. We use a FEM with a uniform mesh. The finite element subspace consists of the usual piecewise polynomials subspace, enriched by a finite number of functions that represent the boundary layer behavior. It is shown that this results in a numerical solution with an ε-uniform error bound in the energy norm associated with the problem. Numerical results are given to illustrate the method.

Perhaps the first use of boundary layer enrichment was given in the paper of Han and Kellogg, [2]. Subsequent work related to this paper is found in Cheng-Temam, [1], which also considers a singularly perturbed ordinary differential equation. The paper [1] is restricted to an equation with constant coefficients, and uses only piecewise linear functions plus an enrichment function. The results are analogous to those of the present paper. Jung and Temam [3, 4, 5] have applied enriched finite

Received by the editors April 5, 2009 and, in revised form, October 12, 2009.
2000 *Mathematics Subject Classification.* 65N30, 65N15, 65L60.

elements to a model singularly perturbed convection diffusion problem whose solution involves both ordinary and parabolic boundary layers. It would be interesting to apply the enriched technique to problems with interior layers.

Section 2 gives some properties of the solution to (1) that are needed for our error analysis. Section 3 formulates the enriched FEM and gives the error analysis. Section 4 presents some numerical results.

We require that the functions p, q, f are sufficiently smooth. Also we assume

$$0 < p_{\min} \le p(x) \le p_{\max} < \infty, \tag{2a}$$

$$q(x) > 0 \text{ in } [0,1], \tag{2b}$$

$$q(x) - \frac{1}{2}p'(x) > 0 \text{ in } [0,1]. \tag{2c}$$

We let $\|w\|_k$ denote the norm in the Sobolev space $H^k(0,1)$, and we use the notation

$$\|w\|_{k,\infty} = \sup\{|w^{(j)}(x)| : x \in [0,1],\ j = 0, \cdots, k\}.$$

We also use $D_x^j w$ as well as $w^{(j)}(x)$ to denote the j^{th} derivative of w with respect to x. When there is no confusion, we will omit the subscript/variable and simply write $D^j w$ or $w^{(j)}$. The letter C denotes a positive number that may be different in different instances, but is always independent of ε and the mesh spacing h.

2. Solution properties

The solution properties for the problem (1) are well-known and may be found, for example, in [7]. These properties are stated here in a form that is useful for our analysis.

From (2b), solutions of (1) satisfy the maximum principle. Therefore the problem (1) has a solution u, and $\|u\|_{0,\infty} \le C\|f\|_{0,\infty}$. For derivative bounds we cite [7, Lemma 1.8]:

$$|u^{(k)}(x)| \le C(f)\big(1 + \varepsilon^{-k} e^{-p_{\min}(1-x)/\varepsilon}\big).$$

Examining the proof one obtains

$$|u^{(k)}(x)| \le C\|f\|_{k,\infty}\big(1 + \varepsilon^{-k} e^{-p_{\min}(1-x)/\varepsilon}\big). \tag{3}$$

We now give a formal asymptotic expansion of the solution. This expansion is also given in [7, p.22], but we derive it in greater detail in order to obtain the information contained in Lemma 1.

Let $V_{n-1}(x) = \sum_{j=0}^{n-1} \varepsilon^j v_j(x)$. Then

$$\begin{aligned} LV_{n-1} &= \sum_{j=0}^{n-1} \big[-\varepsilon^{j+1} v_j'' + p\varepsilon^j v_j' + q\varepsilon^j v_j \big] \\ &= pv_0' + qv_0 + \sum_{j=1}^{n-1} \varepsilon^j \big[pv_j' + qv_j - v_{j-1}''\big] - \varepsilon^n v_{n-1}''. \end{aligned} \tag{4}$$

Define the functions v_j by

$$\begin{aligned} &pv_0' + qv_0 = f,\ v_0(0) = 0, \\ &pv_j' + qv_j = v_{j-1}'',\ v_j(0) = 0 \text{ for } j = 1, \cdots, n-1. \end{aligned}$$

Equation (4) then gives

$$LV_{n-1} = f - \varepsilon^n v''_{n-1},\ V_{n-1}(0) = 0. \tag{5}$$

Also

$$\|v_k\|_{j,\infty} \le C. \tag{6}$$

Next we introduce the stretched variable $\xi = (1-x)/\varepsilon$, so $x = 1 - \varepsilon\xi$. We use a tilde to denote a function of ξ; thus, $\tilde{p}(\xi) = p(1-\varepsilon\xi)$. Note the formula $L\tilde{w} = -\varepsilon^{-1}\tilde{w}'' - \varepsilon^{-1}\tilde{p}\tilde{w}' + \tilde{q}\tilde{w}$. To construct a formal expansion in functions of ξ we use the Taylor series formulas

$$p(x) = \sum_{k=0}^{\infty} p_k(1-x)^k,\ q(x) = \sum_{k=0}^{\infty} q_k(1-x)^k,\ \text{so}$$
$$\tilde{p}(\xi) = \sum_{k=0}^{\infty} p_k\varepsilon^k\xi^k,\ \tilde{q}(\xi) = \sum_{k=0}^{\infty} q_k\varepsilon^k\xi^k.$$

Note that $p_0 = p(1)$. It is convenient to have a notation for the remainder in these Taylor series. We write

$$R_n(p)(x) = p(x) - \sum_{k=0}^{n} p_k(1-x)^k,\ R_n(q)(x) = q(x) - \sum_{k=0}^{n} q_k(1-x)^k.$$

Writing $\tilde{W}(\xi) = \sum_{j=0}^{\infty} \varepsilon^j \tilde{w}_j(\xi)$ with $\tilde{w}_j$ to be defined shortly, a formal calculation gives

$$L\tilde{W} = -\sum_{j=0}^{\infty} \varepsilon^{j-1}[\tilde{w}''_j + p_0\tilde{w}'_j] - \sum_{j=0}^{\infty} \varepsilon^{j-1}(\tilde{p} - p_0)\tilde{w}'_j + \sum_{j=0}^{\infty} \varepsilon^j \tilde{q}\tilde{w}_j.$$

One has

$$\begin{aligned} \sum_{j=0}^{\infty} \varepsilon^{j-1}(\tilde{p}-p_0)\tilde{w}'_j &= \sum_{j=0}^{\infty}\sum_{k=1}^{\infty} p_k\varepsilon^{j+k-1}\xi^k\tilde{w}'_j = \sum_{\mu=0}^{\infty} \varepsilon^{\mu} \sum_{j=0}^{\mu} p_{\mu-j+1}\xi^{\mu-j+1}\tilde{w}'_j, \\ \sum_{j=0}^{\infty} \varepsilon^j\tilde{q}\tilde{w}_j &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} q_k\varepsilon^{j+k}\xi^k\tilde{w}_j = \sum_{\mu=0}^{\infty} \varepsilon^{\mu}\sum_{j=0}^{\mu} q_{\mu-j}\xi^{\mu-j}\tilde{w}_j. \end{aligned} \tag{7}$$

Using these expansions we obtain

$$\begin{aligned} L\tilde{W} = -\varepsilon^{-1}[\tilde{w}''_0 + p_0\tilde{w}'_0] &- \sum_{\mu=0}^{\infty} \varepsilon^{\mu}[\tilde{w}''_{\mu+1} + p_0\tilde{w}'_{\mu+1}] \\ &- \sum_{\mu=0}^{\infty}\varepsilon^{\mu}\sum_{j=0}^{\mu} p_{\mu-j+1}\xi^{\mu-j+1}\tilde{w}'_j + \sum_{\mu=0}^{\infty}\varepsilon^{\mu}\sum_{j=0}^{\mu} q_{\mu-j}\xi^{\mu-j}\tilde{w}_j. \end{aligned}$$

Motivated by this formal expansion we define the functions $\tilde{w}_\nu$, $\nu = 0, 1, \cdots$, by

$$\begin{aligned} &\tilde{w}''_0 + p_0\tilde{w}'_0 = 0, \\ &\quad \tilde{w}_0(0) = -v_0(1),\ \tilde{w}_0(\xi) \to 0 \text{ as } \xi \to \infty, \\ &\tilde{w}''_\nu + p_0\tilde{w}'_\nu = -\sum_{j=0}^{\nu-1} p_{\nu-j}\xi^{\nu-j}\tilde{w}'_j + \sum_{j=0}^{\nu-1} q_{\nu-j-1}\xi^{\nu-j-1}\tilde{w}_j,\ \nu = 1, 2, \cdots, \\ &\quad \tilde{w}_\nu(0) = -v_\nu(1),\ \tilde{w}_\nu(\xi) \to 0 \text{ as } \xi \to \infty. \end{aligned} \tag{8}$$

With these functions, define $\tilde{W}_n = \sum_{\nu=0}^{n} \varepsilon^\nu \tilde{w}_\nu$ and $W_n(x) = \tilde{W}_n(\xi)$.

Regarding the $\tilde{w}_k$ we have

Lemma 1. *$\tilde{w}_k(\xi) = P_{2k}(\xi)e^{-p_0\xi}$ where P_{2k} is a polynomial of degree $2k$. If $p = p_0$ and $q = q_0$ are constants, then $\tilde{w}_k(\xi) = P_k(\xi)e^{-p_0\xi}$.*

Proof. One has $\tilde{w}_0(\xi) = Ce^{-p_0\xi}$, which proves the result for $k = 0$. From (8), $\tilde{w}_1$ satsfies the differential equation

$$\tilde{w}_1'' + p_0\tilde{w}_1' = -p_1\xi\tilde{w}_0' + q_0\tilde{w}_0.$$

Note that the right hand side is a polynomial of degree 1 times $e^{-p_0\xi}$; solving the differential equation one obtains $\tilde{w}_1(\xi) = P_2(\xi)e^{-p_0\xi}$, which gives the result for $k = 1$. The proof for $k > 1$ proceeds in the same way using induction and the fact that $\tilde{w}_\nu$ satisfies

$$\tilde{w}_\nu'' + p_0\tilde{w}_\nu' = \pi_{2\nu-1}(\xi)e^{-p_0\xi},$$

for some polynomial $\pi_{2\nu-1}(\xi)$ of degree $2\nu - 1$. The proof in the case of constant coefficients is similar but simpler. ∎

From Lemma 1, we see that w_k is a linear combination of the functions $\phi_0, \cdots, \phi_{2k}$, so W_n is a linear combination of the functions $\phi_0, \cdots, \phi_{2n}$, where

$$\phi_j(x) = (1-x)^j e^{-p_0(1-x)/\varepsilon}.$$

This suggests that the functions $\phi_0, \cdots, \phi_{2n}$ be used as enrichment functions in a finite element approximation. If it happens that $p = p_0$ and $q = q_0$ are constants, then Lemma 1 suggests the use of $\phi_0, \cdots, \phi_n$ as enrichment functions. Lemma 2 ahead gives a solution decomposition based on these enrichment functions.

We calculate

$$\begin{aligned} L\tilde{W}_n &= \sum_{\nu=0}^{n} \varepsilon^\nu L\tilde{w}_\nu = -\sum_{\nu=0}^{n} \varepsilon^{\nu-1}[\tilde{w}_\nu'' + \tilde{p}\tilde{w}_\nu'] + \sum_{\nu=0}^{n} \varepsilon^\nu \tilde{q}\tilde{w}_\nu \\ &= -\sum_{\nu=0}^{n} \varepsilon^{\nu-1}[\tilde{w}_\nu'' + p_0\tilde{w}_\nu'] - \sum_{\nu=0}^{n} \varepsilon^{\nu-1}(\tilde{p} - p_0)\tilde{w}_\nu' + \sum_{\nu=0}^{n} \varepsilon^\nu \tilde{q}\tilde{w}_\nu. \end{aligned}$$

Using (8) we have

$$\begin{aligned} L\tilde{W}_n &= \sum_{\nu=1}^{n} \sum_{j=0}^{\nu-1} p_{\nu-j}\varepsilon^{\nu-1}\xi^{\nu-j}\tilde{w}_j' - \sum_{\nu=0}^{n} \varepsilon^{\nu-1}(\tilde{p} - p_0)\tilde{w}_\nu' \\ &\quad - \sum_{\nu=1}^{n} \sum_{j=0}^{\nu-1} q_{\nu-j-1}\varepsilon^{\nu-1}\xi^{\nu-j-1}\tilde{w}_j + \sum_{\nu=0}^{n} \varepsilon^\nu \tilde{q}\tilde{w}_\nu \\ &:= A + B. \end{aligned}$$

To obtain a formula for A we use the identity $\sum_{\nu=1}^{n} \sum_{j=0}^{\nu-1} = \sum_{j=0}^{n-1} \sum_{\nu=j+1}^{n}$. In the inner sum, set $k = \nu - j$, so k ranges from 1 to $n - j$. We then get, also changing the summation index in the second sum from ν to j,

$$\begin{aligned} A &= \sum_{j=0}^{n-1}\sum_{k=1}^{n-j} p_k \varepsilon^{j+k-1}\xi^k \tilde{w}_j' - \sum_{j=0}^{n} \varepsilon^{j-1}(\tilde{p}-p_0)\tilde{w}_j' \\ &= -\sum_{j=0}^{n-1}\varepsilon^{j-1}\left[\tilde{p} - \sum_{k=0}^{n-j}\varepsilon^k\xi^k p_k\right]\tilde{w}_j' = -\sum_{j=0}^{n-1}\varepsilon^{j-1}R_{n-j}(p)\tilde{w}_j'. \end{aligned}$$

A similar result holds for B: One sees that $\sum_{\nu=1}^{n-1}\sum_{j=0}^{\nu-1} = \sum_{j=0}^{n-2}\sum_{\nu=j+1}^{n-1}$. In the inner sum, set $k=\nu-j-1$, so k ranges from 0 to $n-j-2$. We then get, also changing the summation index in the second sum from ν to j,

$$\begin{aligned} B &= \sum_{j=0}^{n}\varepsilon^j\tilde{q}\tilde{w}_j - \sum_{j=0}^{n-2}\sum_{\nu=j+1}^{n-1} q_{\nu-j-1}\varepsilon^{\nu-1}\xi^{\nu-j-1}\tilde{w}_j \\ &= \sum_{j=0}^{n-2}\varepsilon^j\left[\tilde{q} - \sum_{k=0}^{n-j-2}\varepsilon^k\xi^k q_k\right]\tilde{w}_j + \tilde{q}(\varepsilon^{n-1}\tilde{w}_{n-1}+\varepsilon^n\tilde{w}_n) \\ &= \sum_{j=0}^{n-2}\varepsilon^j R_{n-j-2}(q)\tilde{w}_j + \tilde{q}(\varepsilon^{n-1}\tilde{w}_{n-1}+\varepsilon^n\tilde{w}_n). \end{aligned}$$

Combining the formulas for A and B gives

$$L\tilde{W}_n = -\sum_{j=0}^{n-1}\varepsilon^{j-1}R_{n-j}(p)\tilde{w}_j' + \sum_{j=0}^{n-2}\varepsilon^j R_{n-j-2}(q)\tilde{w}_j + \tilde{q}(\varepsilon^{n-1}\tilde{w}_{n-1}+\varepsilon^n\tilde{w}_n). \tag{9}$$

If $n=1$ the second sum in (9) is not present.

We now estimate the derivatives of $L\tilde{W}_n$ and LW_n. One has

$$|D_x^k R_n(p)| \le C(1-x)^{n+1-k},\ |D_x^k R_n(q)| \le C(1-x)^{n+1-k}.$$

Using Lemma 1 and the inequality $\xi^k e^{-p_0\xi} \le C(a)e^{-a\xi}$ for $a\in(0,p_0)$, a computation gives

$$|D_\xi^k L\tilde{W}_n(\xi)| \le C(a)\varepsilon^{n-1}e^{-a\xi} \text{ for } a\in(0,p_0).$$

In terms of the unstretched variables, $W_n(x)$ satisfies

$$|D_x^k LW_n(x)| \le C(a)\varepsilon^{n-1-k}e^{-a(1-x)/\varepsilon} \text{ for } a\in(0,p_0), \tag{10a}$$

$$|W_n(0)| \le C(a)e^{-a/\varepsilon} \text{ for } a\in(0,p_0), \tag{10b}$$

$$W_n(1)=0. \tag{10c}$$

Combining (5) and (10) we have

$$u = V_{n-1} + W_n + \mathcal{R}_n, \tag{11}$$

where the remainder $\mathcal{R}_n$ satisfies

$$\begin{aligned} &|D_x^k L\mathcal{R}_n| \le C\varepsilon^n + C(a)\varepsilon^{n-1-k}e^{-a(1-x)/\varepsilon} \text{ for } a\in(0,p_0), \\ &|\mathcal{R}_n(0)| \le C(a)e^{-a/\varepsilon} \text{ for } a\in(0,p_0), \\ &\mathcal{R}_n(1)=0. \end{aligned} \tag{12}$$

The decomposition (11) can be put into a more convenient form by removing the non-zero boundary condition. Let $\mathcal{R}_{n,1} = (1-x)\mathcal{R}_n(0)$, $\mathcal{R}_{n,2} = \mathcal{R}_n - \mathcal{R}_{n,1}$. Thus

$$|\mathcal{R}'_{n,1}| \le C|\mathcal{R}_{n,1}(0)| \le C(a)e^{-a/\varepsilon} \text{ for } a \in (0, p_0),\ D^k\mathcal{R}_{n,1} = 0 \text{ for } k > 1. \tag{13}$$

Also $\mathcal{R}_{n,2}(0) = \mathcal{R}_{n,2}(1) = 0$, and $L\mathcal{R}_{n,2} = L\mathcal{R}_n + \mathcal{R}_n(0)\big(-p(x) + (1-x)q(x)\big)$. It follows that $\mathcal{R}_{n,2}$ satisfies, for $a = p_0/2$,

$$\begin{aligned} |D_x^k L\mathcal{R}_{n,2}| &\le C\varepsilon^n + C\varepsilon^{n-1-k}e^{-a(1-x)/\varepsilon}, \\ \mathcal{R}_{n,2}(0) &= 0, \\ \mathcal{R}_{n,2}(1) &= 0. \end{aligned} \tag{14}$$

Applying [6, Lemma 2.3] to the problem derived from (14) and satisfied by $\varepsilon^{-n}\mathcal{R}_{n,2}$ we obtain the bound

$$|D_x^k\mathcal{R}_{n,2}(x)| \le C[\varepsilon^n + \varepsilon^{n-k}e^{-a(1-x)/\varepsilon}]. \tag{15}$$

Setting $r_n = V_{n-1} + \mathcal{R}_n$ and using (6), (11), (13), (15), we have

$$u = r_n + W_n \text{ where } |D_x^k r_n(x)| \le C[1 + \varepsilon^n + \varepsilon^{n-k}e^{-a(1-x)/\varepsilon}]. \tag{16}$$

This leads to a solution decomposition that is used in the finite element error analysis.

Lemma 2. *There are numbers $\gamma_0, \cdots, \gamma_{2n}$ such that*

$$u = r_n + \sum_{j=0}^{2n} \gamma_j\phi_j, \tag{17a}$$

$$|D_x^k r_n(x)| \le C[1 + \varepsilon^n + \varepsilon^{n-k}e^{-a(1-x)/\varepsilon}], \tag{17b}$$

$$\|r_n^{(k)}\|_2 \le C[1 + \varepsilon^{n-k+1/2}], \tag{17c}$$

$$|\gamma_k| \le C\varepsilon^{-\frac{1}{2}k} \textit{ for } k \textit{ even, } |\gamma_k| \le C\varepsilon^{-\frac{1}{2}(k-1)} \textit{ for } k \textit{ odd.} \tag{17d}$$

In the case of constant coefficients, (17a) *is replaced by $u = r_n + \sum_{j=0}^n \gamma_j\phi_j$ and $|\gamma_k| \le C$.*

Proof. We have $\tilde{W}_n(\xi) = \sum_{j=0}^n \varepsilon^j P_{2j}(\xi)e^{-p_0\xi}$, where $P_{2j}(\xi) = \sum_{k=0}^{2j} c_{2j,k}\xi^k$ with $|c_{2j,k}| \le C$. Therefore, using (16) and the fact that $\xi = \varepsilon^{-1}(1-x)$, we immediately get (17a) and (17b), with (17c) obtained by integration. The γ_j's in (17a) are given by

$$\begin{aligned} \gamma_0 &= c_{0,0} + \varepsilon c_{2,0} + \varepsilon^2 c_{4,0} + ... + \varepsilon^n c_{2n,0}, \\ \gamma_1 &= c_{2,1} + \varepsilon c_{4,1} + \varepsilon^2 c_{6,1} + ... + \varepsilon^{n-1} c_{2n,1}, \\ \gamma_2 &= \varepsilon^{-1}c_{2,2} + c_{4,2} + \varepsilon c_{6,2} + ... + \varepsilon^{n-2}c_{2n,2}, \\ \gamma_3 &= \varepsilon^{-1}c_{4,3} + c_{6,3} + \varepsilon c_{8,3} + ... + \varepsilon^{n-3}c_{2n,3}, \\ \gamma_4 &= \varepsilon^{-2}c_{4,4} + \varepsilon^{-1}c_{6,4} + c_{8,4} + ... + \varepsilon^{n-4}c_{2n,4}, \\ \gamma_5 &= \varepsilon^{-2}c_{6,5} + \varepsilon^{-1}c_{8,5} + c_{10,5} + ... + \varepsilon^{n-5}c_{2n,5}, \\ &\vdots \end{aligned}$$

from which (17d) follows. In the case of constant coefficients the situation is similar (but simpler). ∎

3. The enriched finite element method

The solution of (1) satisfies

$$B(u,v)=\int_0^1 fv\,dx \text{ for } v\in H_0^1(0,1), \tag{18}$$

where the bilinear form B is defined by

$$B(v,w)=\int_0^1 \big(\varepsilon v'w'+pv'w+qvw\big)dx.$$

For the FEM we pick an N-dimensional subspace $\mathcal{S}\subset H_0^1(0,1)$. The finite element solution $u_N\in\mathcal{S}$ is defined to be the function which satisfies

$$B(u_N,v)=\int_0^1 fv\,dx \text{ for } v\in\mathcal{S}. \tag{19}$$

Equation (19) gives a linear system of N equations in N unknowns.

The inequality (2c) implies

$$B(v,v)\geq c\int_0^1 \big(\varepsilon v'^2+v^2\big)dx, \tag{20}$$

and (20) in turn implies the nonsingularity of the system (19). We let u_N denote the solution to (19).

We shall use the norms $\|v\|_{1,\varepsilon}$ and $\|v\|_{1,\varepsilon,\varepsilon^{-1}}$ defined by

$$\|v\|_{1,\varepsilon}^2=\int_0^1[\varepsilon v'^2+v^2]dx,$$
$$\|v\|_{1,\varepsilon,\varepsilon^{-1}}^2=\int_0^1[\varepsilon v'^2+\varepsilon^{-1}v^2]dx.$$

The inequality (20) gives $B(v,v)\geq c\|v\|_{1,\varepsilon}^2$. Also one has

$$\begin{aligned}|B(v,w)|&\leq C\|v\|_1\|w\|_{1,\varepsilon},\\ |B(v,w)|&\leq C\|v\|_{1,\varepsilon,\varepsilon^{-1}}\|w\|_{1,\varepsilon}.\end{aligned} \tag{21}$$

Subtracting (19) from (18) one obtains $B(u-u_N,v)=0$ for $v\in\mathcal{S}$. This implies $B(v-u_N,v-u_N)=B(v-u,v-u_N)$. So

$$c\|v-u_N\|_{1,\varepsilon}^2\leq B(v-u,v-u_N) \text{ for } v\in\mathcal{S}.$$

From this inequality and the first inequality of (21) we obtain

$$\|v-u_N\|_{1,\varepsilon}\leq C\|v-u\|_1^{1/2}\|v-u_N\|_{1,\varepsilon}^{1/2},$$

so

$$\|v-u_N\|_{1,\varepsilon}\leq C\|v-u\|_1.$$

Hence

$$\|u-u_N\|_{1,\varepsilon}\leq\|u-v\|_{1,\varepsilon}+\|v-u_N\|_{1,\varepsilon}\leq C\|u-v\|_1.$$

A similar argument gives $\|u-u_N\|_{1,\varepsilon}\leq C\|u-v\|_{1,\varepsilon,\varepsilon^{-1}}$. Collecting these inequalities and taking the supremum over $v\in\mathcal{S}$, we obtain

$$\|u-u_N\|_{1,\varepsilon}\leq C\inf_{v\in\mathcal{S}}\|u-v\|_1, \tag{22a}$$

$$\|u-u_N\|_{1,\varepsilon}\leq C\inf_{v\in\mathcal{S}}\|u-v\|_{1,\varepsilon,\varepsilon^{-1}}. \tag{22b}$$

Let $M > 0$ be an integer, let $h = M^{-1}$, and let $\mathcal{V}_{M,\ell}$ be the space of piecewise polynomials of degree ℓ on a uniform mesh of width h, which vanish at $x = 0, 1$. Thus, $\mathcal{V}_{M,\ell}$ has dimension $\ell M - 1$ and has the approximation property

$$\inf_{v\in\mathcal{V}_{M,\ell}} \left\{\|w - v\|_0 + h\|w - v\|_1\right\} \leq Ch^{\ell+1}\|w\|_{\ell+1} \text{ for } w \in H^{\ell+1}(0, 1). \tag{23}$$

Let n be a positive integer and let

$$\mathcal{S}_{M,\ell,n} = \mathcal{V}_{M,\ell} + \{\phi_0, \cdots, \phi_{2n}\},\ n = 0, 1, \cdots .$$

Thus, $\mathcal{S}_{M,\ell,n}$ has dimension $N = \ell M + 2n$, $n = 0, 1, \cdots$. The space $\mathcal{S}_{M,\ell,n}$ is our "enriched" finite element space. The following theorem gives our error estimate. The proof uses the solution decomposition (17a) but with n replaced by $n+1$; that is, $u = r_{n+1} + \sum_{j=0}^{2n+2} \gamma_j\phi_j$. To use (23) one needs $r_{n+1} \in H^{\ell+1}(0, 1)$, and from (17c), this requires $n \geq \ell$.

Theorem 1. *Let $n \geq \ell$. Let u be the solution of* (18) *and $u_N \in \mathcal{S}_{M,\ell,n}$ the solution of* (19). *We have*

$$\|u - u_N\|_{1,\varepsilon} \leq Ch^{\ell}, \tag{24}$$

with $C \in \mathbb{R}$ a positive constant independent of $M = h^{-1}$ and ε.

Proof. Note that $\phi_j(x) = \phi_j(1 - \varepsilon\xi) = \varepsilon^j\xi^j e^{-p_0\xi}$, so $D_x^k\phi_j(x) = \varepsilon^{j-k}D_\xi^k(\xi^j e^{-p_0\xi})$ and

$$\|D_x^k\phi_j\|_0 \leq C\varepsilon^{j-k}\left[\int_0^\infty \xi^{2j}e^{-2p_0\xi}\varepsilon d\xi\right]^{1/2} \leq C\varepsilon^{j-k+\frac{1}{2}}. \tag{25}$$

Let $v \in \mathcal{V}_{M,\ell}$ be a good approximation to r_{n+1} in the sense of (23), and let $s = \sum_0^{2n} \gamma_i\phi_i + v \in \mathcal{S}_{M,\ell,n}$. Using (17c), (17d), and (25) we have

$$\begin{aligned}
\|u - s\|_0 &\leq \|\gamma_{2n+1}\phi_{2n+1}\|_0 + \|\gamma_{2n+2}\phi_{2n+2}\|_0 + \|r_{n+1} - v\|_0 \\
&\leq C[\varepsilon^{n+3/2} + h^{\ell+1}], \\
\|u' - s'\|_0 &\leq C[\varepsilon^{n+1/2} + h^{\ell}].
\end{aligned}$$

Therefore $\|u - s\|_1 \leq C[\varepsilon^{n+1/2} + h^{\ell}]$. From (22a) we obtain

$$\|u - u_N\|_{1,\varepsilon} \leq C[\varepsilon^{n+1/2} + h^{\ell}].$$

Since $n \geq \ell$, this gives (24) in the case $\varepsilon \leq h$.

To treat the case $\varepsilon \geq h$, let $v \in \mathcal{V}_{M,\ell}$ be a good approximation to $\gamma_{2n+1}\phi_{2n+1} + \gamma_{2n+2}\phi_{2n+2} + r_{n+1}$ in the sense of (23), and let $s = \sum_0^{2n} \gamma_i\phi_i + v \in \mathcal{S}_{M,\ell,n}$. Using (17c), (17d), and (25) we have

$$\begin{aligned}
\|u - s\|_0 &\leq Ch^{\ell+1}\left[\|\gamma_{2n+1}\phi_{2n+1}\|_{\ell+1} + \|\gamma_{2n+2}\phi_{2n+2}\|_{\ell+1} + \|r_{n+1}\|_{\ell+1}\right] \\
&\leq C[1 + \varepsilon^{n-\ell+1/2}]h^{\ell+1}, \\
\|u' - s'\|_0 &\leq C[1 + \varepsilon^{n-\ell-1/2}]h^{\ell}.
\end{aligned}$$

Therefore

$$\|u - s\|_{1,\varepsilon,\varepsilon^{-1}} \leq C\left[\varepsilon^{n-\ell}h^{\ell+1} + \varepsilon^{-1/2}h^{\ell+1} + \varepsilon^{n-\ell}h^{\ell} + \varepsilon^{1/2}h_{\ell}\right] \tag{26}$$

From (22b) we obtain

$$\|u - u_N\|_{1,\varepsilon} \leq C\left[1 + \varepsilon^{n-\ell}\right]h^{\ell}, \tag{27}$$

Since $n \geq \ell$ this gives (24) in the case $\varepsilon \geq h$. ∎

Remark 1. *The following remarks, related to Theorem* 1*, are in order:*

- *For the variable coefficient case, the condition* $n \geq \ell$ *means that if piecewise linear polynomials are used then at least three enrichment functions should be added. However, in the numerical experiments, given in the following section, we find that optimal convergence rates and satisfactory errors are attained with only one enrichment function.*
- *From Lemma* 2*, if the coefficients* p, q *are constant, one can replace* $\mathcal{S}_{M,\ell,n}$ *by the space* $\mathcal{V}_{M,\ell}+\{\phi_0, \cdots, \phi_n\}$*. This means that the number of enrichment functions that should be added is equal to the degree of the approximating piecewise polynomials.*

The above remarks will be illustrated in the following section.

4. Numerical experiments

In this section we present the results of numerical computations for two model problems: one with constant coefficients and another with variable coefficients. In both cases we will show convergence plots of the percentage relative error in the energy norm,

$$Error = 100 \times \frac{\|u - u_N\|_{1,\varepsilon}}{\|u\|_{1,\varepsilon}},$$

versus the number of degrees of freedom N, in a log-log scale.

4.1. The constant coefficient case. We consider the problem (1), with $p(x) = q(x) = f(x) = 1$ which allows us to obtain its exact solution and thus, the results we report are reliable. It is well known that the standard h version FEM with piecewise linear basis functions on a uniform mesh yields an approximation that contains oscillations. If, however, we enrich the finite element space with the function $\phi_0(x) = e^{-(1-x)/\varepsilon}$, then the situation changes dramatically. Figure 1 shows the plot of the true and approximate/enriched solutions, as well as the error between them at the nodes. We see that there are no oscillations and the absolute error at the nodes is of the order $O(h^2)$. (Other values for ε gave the same results.)

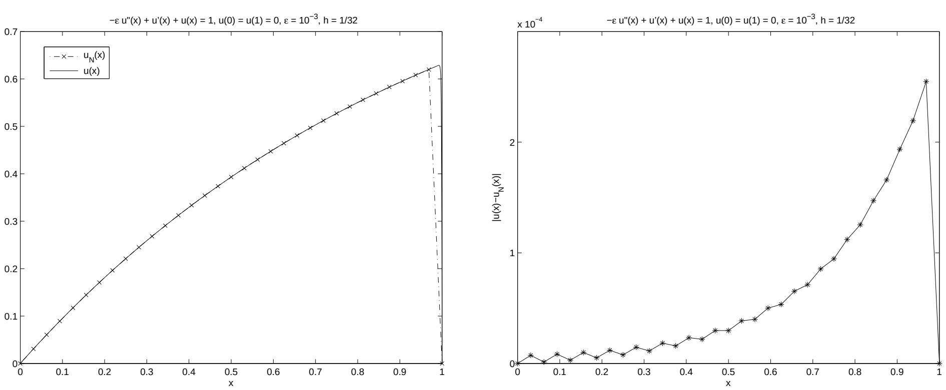

FIGURE 1. Left: Plot of $u(x)$ and $u_N(x)$ for $\varepsilon = 10^{-3}$ and $h = 1/32$, with u_N the enriched FE solution. Right: Error at the nodes.

Figure 2 shows the robustness and optimal convergence rate $O(h)$ of the enriched method, for various values of ε.

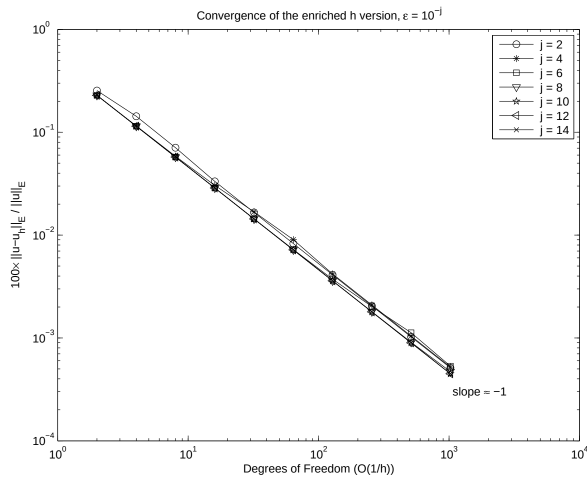

FIGURE 2. Convergence of the enriched finite element solution obtained with piecewise linear functions plus one enrichment function on a uniform mesh.

When the finite element space consists of piecewise quadratics, Lemma 2 suggests that we should add the two enrichment functions

$$\phi_0(x) = e^{-(1-x)/\varepsilon}\ ,\ \phi_1(x) = (1-x)e^{-(1-x)/\varepsilon}.$$

Figure 3 shows the convergence of this method, and again we see that it is robust with optimal rate $O(h^2)$.

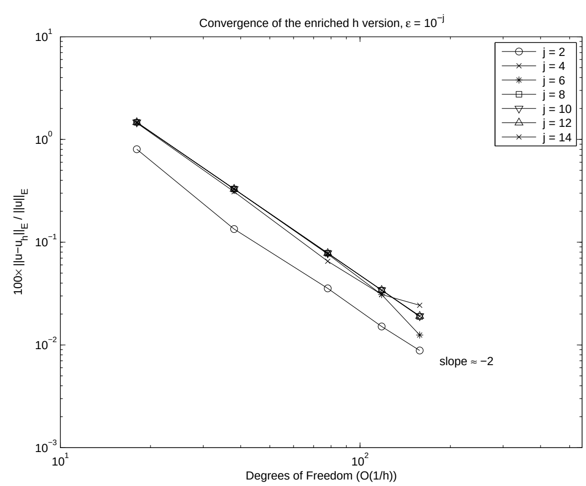

FIGURE 3. Convergence of the enriched finite element solution obtained with piecewise quadratics plus two enrichment functions on a uniform mesh.

4.2. The variable coefficient case. We now turn our attention to a variable coefficient problem (1), in which $p(x) = (x + 1), q(x) = 3/2$ and $f(x)$ chosen so that the exact solution is known. Theorem 1 states that if the finite element space consists of piecewise linear basis functions on a uniform mesh, then we should add at least 3 enrichment functions ϕ_j. Nevertheless, we wish to examine the effect of adding just one enrichment function (as was done for the constant coefficient case). Based on the data of the problem, we enrich the space with the function $\phi_0(x) = e^{-2(1-x)/\varepsilon}$ and in Figure 4 we show the plot of the exact and approximate solution obtained with this method, as well as the error between them at the nodes, for $\varepsilon = 10^{-3}, h = 1/32$; other values of ε gave similar results. We do not observe any oscillations in the solution, but when we look at the error at the nodes we see that some oscillations are present, although of very small scale and appearing throughout the interval $(0, 1)$. This suggests that their presence is not due to the boundary layer, and if one wishes, one can "post-process" the approximate solution (by, for example, averaging) to smooth them out, something that falls outside the scope of the present article. Figure 5 shows the performance of this method for various values of ε. It is clear that the method is robust with optimal rate $O(h)$, even though only one enrichment function was added to the finite element space.

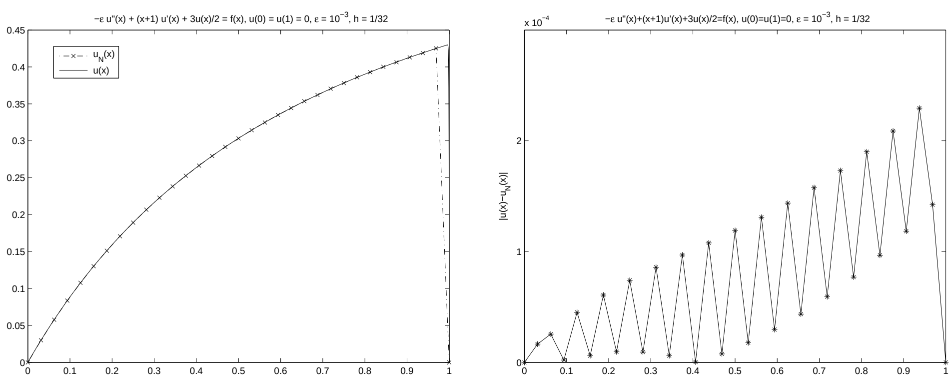

FIGURE 4. Left: Plot of $u(x)$ and $u_N(x)$ for $\varepsilon = 10^{-3}$ and $h = 1/32$, with u_N the enriched FE solution (by one function). Right: Error at the nodes.

Next, we would like to see what effect (if any) the addition of a second enrichment function would have. To this end, we enrich the finite element space with the functions $\phi_0(x) = e^{-2(1-x)/\varepsilon}$ and $\phi_1(x) = (1-x)e^{-2(1-x)/\varepsilon}$, and repeat the previous computations. Figure 6 shows the plot of the exact and approximate solution obtained with this method, as well as the error between them at the nodes, for $\varepsilon = 10^{-3}, h = 1/32$. There is no significant difference between Figures 4 and 6, hence it seems that adding only one enrichment function is sufficient, even for the variable coefficient case. Figure 7 shows the performance of the method for various values of ε, and again the robustness and optimal convergence rate are clearly visible.

As a final experiment, we would like to see what happens when a third enrichment function is added to the subspace, as Theorem 1 requires. Since two enrichment

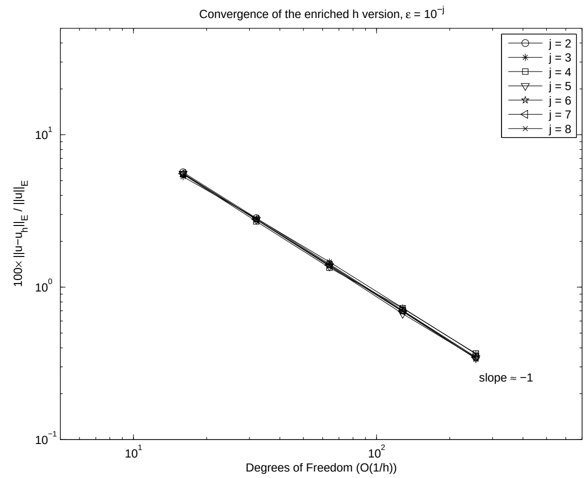

FIGURE 5. Convergence of the enriched finite element solution obtained with piecewise linear functions plus one enrichment function on a uniform mesh.

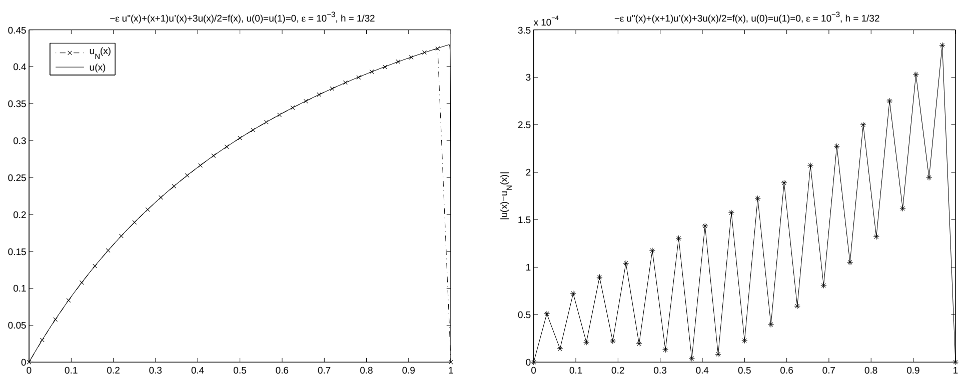

FIGURE 6. Left: Plot of $u(x)$ and $u_N(x)$ for $\varepsilon = 10^{-3}$ and $h = 1/32$, with u_N the enriched FE solution (by two functions). Right: Error at the nodes.

functions do not yield better results than using just one, this last computation would, in some sense, test the sharpness of the result stated in Theorem 1. The third enrichment function is given by $\phi_2(x) = (1-x)^2 e^{-2(1-x)/\varepsilon}$, and Figures 8–9 show the results of this computation. It is clear, from these figures, that a third enrichment function is not required. Moreover, by comparing Figures 4, 6 and 8, we see that the error at the nodes does not decrease as the number of enrichment functions increases – in fact it seems to be (slightly) increasing. This, we believe, is due to the fact that the coefficient matrix in the linear system becomes ill-conditioned (for small values of ε) as the number of enrichement functions increases. Concluding, we find that the "best" results are obtained with only one enrichment function.

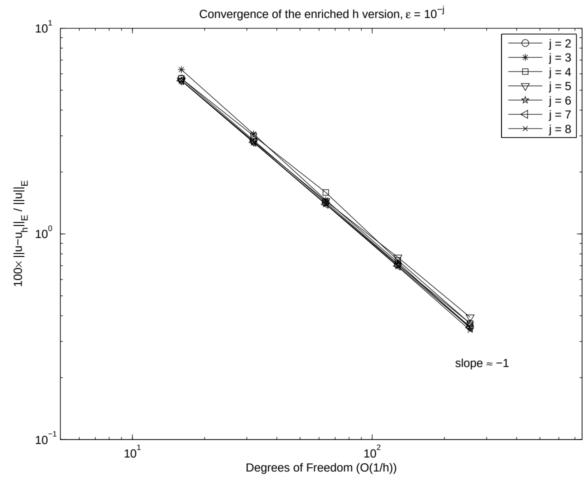

FIGURE 7. Convergence of the enriched finite element solution obtained with piecewise linear functions plus two enrichment functions on a uniform mesh.

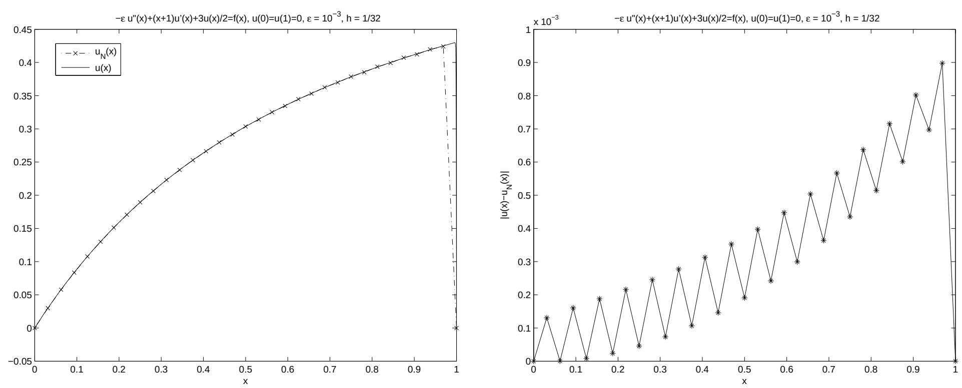

FIGURE 8. Left: Plot of $u(x)$ and $u_N(x)$ for $\varepsilon = 10^{-3}$ and $h = 1/32$, with u_N the enriched FE solution (by three functions). Right: Error at the nodes.

Acknowledgments

The authors thank the anonymous referees whose suggestions helped improve the paper.

References

[1] Wenfang Cheng and Roger Temam, Numerical approximation of one-dimensional stationary diffusion equations with boundary layers, Computers & Fluids, 31(2002), 453–466.

[2] H. Han and R. B. Kellogg, The use of enriched subspaces for singular perturbation problems, Proceedings of the China-France Symposium on finite element methods, Beijing, 1982 (Scinece Press, Beijing, 1983), pp. 293–305.

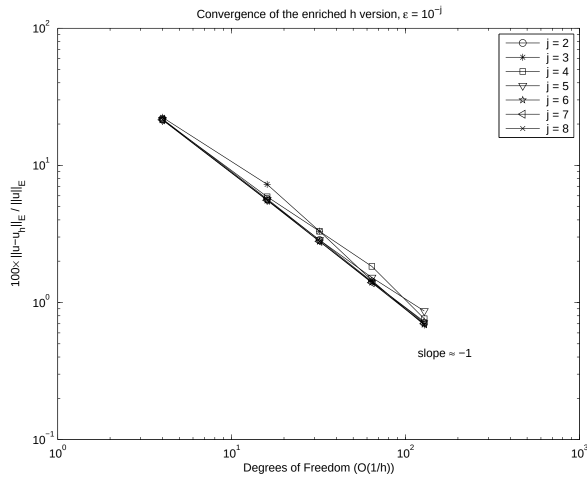

FIGURE 9. Convergence of the enriched finite element solution obtained with piecewise linear functions plus three enrichment functions on a uniform mesh.

[3] Chang-yeol Jung and Roger Temam, Numerical approximation of two-dimensional convection-diffusion equations with multiple boundary layers, International Journal of Numerical Analysis and Modeling, 2(2005), 367–408.

[4] Chang-yeol Jung and Roger Temam, Construction of boundary layer elements for singularly perturbed convection-diffusion equations and L^2-stability analysis, International Journal of Numerical Analysis and Modeling, 5(2008), 729–748.

[5] Chang-yeol Jung and Roger Temam, Interaction of boundary layers and corner singularities, Discrete and Continuous Dynamical Systems, 23(2009), 315–339.

[6] R. B. Kellogg and Alice Tsan, Analysis of some difference approximations for a singularly perturbed problem without turning points, Math. of Comp., 32(1978), 1025–1039.

[7] H.-G. Roos, M. Stynes and L. Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations, Second Edition, Springer Verlag, 2008.

[8] C. Schwab, *p*- and *hp*- Finite Element Methods, Oxford Science Publications, 1998.

[9] G. I. Shishkin, Discrete approximation of singularly perturbed elliptic and parabolic equations (in Russian), Russian Academy of Sciences, Ural Section, Ekaterinburg, 1992.

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA
E-mail: `rbmjk@windstream.net`

Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, Nicosia, 1678, CYPRUS
E-mail: `xenophontos@ucy.ac.cy`
URL: `http://www2.ucy.ac.cy/~xenophon/`