INT. J. NUM. METH. HEAT FLUID FLOW, VOL. 3, 205–221 (1993)

PETROV–GALERKIN METHODS ON ISOPARAMETRIC BILINEAR AND BIQUADRATIC ELEMENTS TESTED FOR A SCALAR CONVECTION–DIFFUSION PROBLEM

J. PETERA†, V. NASSEHI† AND J. F. T. PITTMAN

Computer Aided Process Engineering Group, Department of Chemical Engineering, University of Wales, Singleton Park, Swansea SA2 8PP, UK

ABSTRACT

A number of finite element formulations involving discontinuous weighting functions have been tested against analytic solutions for a steady scalar convection–diffusion problem at intermediate Peclet number, with a 'hard' downstream boundary condition. The emphasis is on extending these methods to isoparametric bilinear and biquadratic elements. In order to do this a procedure is given for the exact calculation of shape function Laplacians. Having confirmed the success of the Brooks–Hughes streamline upwind Petrov–Galerkin (SUPG) method for isoparametric bilinear elements, formulations for biquadratic elements are examined. Galerkin least squares offers little advantage over SUPG in the test problem. The generalized Galerkin method of Donea *et al.* gave excellent results, but because of concern over the possibility of cross-streamline artificial diffusion in some cases, a strictly streamline formulation incorporating the optimal parameters of Donea *et al.* is proposed. This gave excellent results on a sufficiently refined mesh.

KEY WORDS Petrov–Galerkin methods Convection–diffusion problem Isoparametric elements

INTRODUCTION

The convection–diffusion problem

It is well known from finite element theory and practice that Galerkin methods applied to non-symmetric operator problems, such as arise when both convection and diffusion terms are included in the energy equation, lack the 'best approximation' property. The solutions of many convection–diffusion problems, where convection is significant, are perturbed by spurious oscillations, or 'wiggles', which make the solution worthless.

Many efforts have been made to overcome this difficulty. Several of them consist in deformation of shape functions so as to upwind the difference stencil, that is to shift it in the upstream direction, to give a better approximation of convection terms at high Peclet numbers. Upwinding techniques for finite elements have been developed intensively since the mid-1970s (1 and references therein). The method has been formalized and mathematically justified in the works of Hughes and co-workers[2–6], and in parallel by Johnson and his group[7–9]. Generally, a new weighting function space is introduced consisting of the sum of the usual Galerkin shape function, plus an 'upwind' term formed from the scalar product of the velocity vector and the gradient of the same shape function, multiplied by a coefficient which may be dependent on the mesh

† Present address: Department of Chemical Engineering, Loughborough University of Technology, Loughborough, UK.

0961–5539/93/030205–17$2.00

Received August 1992
Revised November 1992

Peclet number. Using this space of modified, discontinuous weighting functions, the streamline upwind Petrov–Galerkin (SUPG) method[2] was formulated as the proper weighted residuals variational method, capable of high accuracy and not subject to the deficiencies of the former streamline upwind (SU) method. This latter method, also known as selective upwinding, consists in applying the modified weighting to the convection term only, and can be interpreted[2] as the addition of a portion of artificial diffusion to the physical diffusion in the differential equation, followed by conventional Galerkin treatment. This improves stability, but at the expense of accuracy, and though smooth solutions can be obtained they are often overdiffuse.

Recently, Hughes and co-workers[10,11] have indicated that the SUPG method is a special case of a more general methodology—the so-called mixed variational formulation. In order to improve the stability of the Galerkin method, a Lagrange multiplier is used as an additional variable. In this way, the primary variables, such as temperature in the energy equation, belong to a larger space of functions, giving more flexibility and permitting better approximation. The problem arises that not every mixed formulation is eligible for combination with the Galerkin method, because of the requirement of the Babuska–Brezzi stability conditions[10]. However, Hughes and co-workers[10] have proved that addition to the Galerkin bilinear form of various least-squares-like terms, containing integrals over element interiors, allows one to satisfy or at least circumvent the BB condition. In the case of the convection–diffusion problem, this additional perturbation term is based on residual forms of the Euler–Lagrange equations resulting from momentum or energy conservation. The convergence of this Galerkin least squares (GLS) formulation was proved in Reference 11.

Aims of the present work

The aim of the present work is to test a number of versions of SUPG and GLS formulations, described in the following section, together with the classical Galerkin method, implemented using both bilinear and biquadratic isoparametric elements, in an abstract steady scalar convection–diffusion problem at moderately high Peclet numbers. The test problem has a 'hard' (Dirichlet) downstream boundary condition, and exhibits steep exit boundary layers, providing a quite severe test for the methods. Numerical results are compared with analytic solutions for the cases with and without a source term. It should be noted that most previous tests of these methods have been on purely hyperbolic problems. To our knowledge no tests against exact solutions are available for a two-dimensional mixed hyperbolic–elliptic problem such as is considered here.

The importance of the diffusion term in the exit boundary layers means that the Laplacian, $\nabla^2 T$, cannot generally be neglected in the discretized problem. In bilinear elements, which have previously been used to implement the schemes, the Laplacian is identically zero, but for isoparametric bilinear and biquadratic elements this is not the case. We wish to extend the methods to isoparametric elements, because of their geometrical flexibility, and to biquadratics in particular, because of their generally greater efficiency in providing accurate solutions. We therefore describe a convenient algorithm for evaluation of the second derivatives of shape functions with respect to global co-ordinates in isoparametric elements. Later the test problem is presented, its solution given and extended to the case including a source term. Finite element results are compared with exact solutions in the subsequent section, and finally conclusions are drawn.

GALERKIN LEAST SQUARES AND UPWIND FORMULATIONS

Let us consider the simplest, steady-state convection–diffusion equation (7):

$$\mathscr{L}(T, \kappa) = q \tag{1}$$

where

$$\mathscr{L}(T,\kappa)=\mathbf{v}\cdot\nabla\mathbf{T}-\kappa\,\nabla\cdot\nabla\mathbf{T} \tag{2}$$

$T=T(\mathbf{x})$ is the dependent variable (e.g. temperature), $x\in\Omega\subset\mathbb{R}^d$ is the spatial domain of the problem with d the space dimensionality, $\mathbf{v}$ is the velocity, κ the diffusivity and q a volumetric source term. ∇ denotes the spatial gradient operator. We assume that velocity is divergence free:

$$\nabla\cdot\mathbf{v}=0 \tag{3}$$

and κ is constant.

As usual, Dirichlet and/or Neumann boundary conditions are imposed,

$$T=g \qquad \text{on}\,\Gamma_g \tag{4a}$$

$$\kappa\,\nabla\mathbf{T}\cdot\mathbf{n}=h \qquad \text{on}\,\Gamma_h \tag{4b}$$

where Γ is the boundary of Ω and

$$\Gamma=\mathrm{cl}(\Gamma_g\cup\Gamma_h) \qquad \Gamma_g\cap\Gamma_h=\phi$$

$\mathbf{n}$ is the outward normal vector to Γ, and g and h are given functions of position.

The convection dominated problem corresponds to small κ, or, more exactly, to large element Peclet number,

$$Pe^d=\frac{|\mathbf{v}|\,d}{2\kappa}\gg 1 \tag{5}$$

where d is a characteristic dimension of the mesh element.

The finite element Galerkin least squares (GLS) variational formulation may be stated as flows. Let $\mathscr{S}^d$ be the shape functions space, and $\mathscr{V}^d$ the weighting functions space, both composed of the same class of functions up to inhomogeneous Dirichlet boundary conditions. The problem consists in finding $T^d\in\mathscr{S}^d$ such that for all $W^d\in\mathscr{V}^d$:

$$B_{\mathrm{GLS}}(W^d,T^d)=L_{\mathrm{GLS}}$$

where

$$B_{GLS}(W^d,T^d)=\int_\Omega(W^d\mathbf{v}\cdot\nabla T^d+\kappa\nabla W^d\cdot\nabla T^d)\,\mathrm{d}\Omega+\sum_{e=1}^{n_1}\int_{\Omega^e}\tau P_1(W^d,\kappa)D(T_1^d\kappa)\,\mathrm{d}\Omega \tag{6a}$$

$$L_{GLS}(W^d)=\int_\Omega W^dq\,\mathrm{d}\Omega+\int_{\Gamma_h}W^dh\,\mathrm{d}\Gamma+\sum_{e=1}^{n_1}\int_{\Omega^e}\tau P_2(W^d,\kappa)q\,\mathrm{d}\Omega \tag{6b}$$

$$P_1(W^d,\kappa)=P_2(W^d,\kappa)=\mathbf{v}\cdot\nabla W^d-\kappa\nabla^2W^d \tag{6c}$$

$$D(T^d,\kappa)=\mathbf{v}\cdot\nabla T^d-\kappa\nabla^2T^d \tag{6d}$$

and the Lagrange multiplier is τ. The SUPG method in its simplest form may be obtained from the above formulation by putting:

$$P_1(W^d,\kappa)=P_2(W^d,\kappa)=\mathbf{v}\cdot\nabla W^d \tag{7}$$

Both SUPG and GLS are weighted residuals methods, and can be called consistent upwinding methods. The corresponding Euler–Lagrange equations may be written as follows:

$$\sum_{e=1}^{n_1}\int_\Omega(W^d+P_1(W^d,\kappa))(\mathscr{L}(T^d,\kappa)-q)\,\mathrm{d}\Omega+\int_{\Gamma_h}W^d(\kappa\nabla T\cdot\mathbf{n}-h)\,\mathrm{d}\Gamma$$
$$+\int_{\Gamma_{\mathrm{int}}}W^d\kappa(\nabla T^+-\nabla T^-)\cdot\mathbf{n}^+\,\mathrm{d}\Gamma=0 \tag{8}$$

where $\Gamma_{\text{int}} = \bigcup_e \Gamma^e - \Gamma$, $\mathbf{n}^+$ is the unit vector normal to the 'positive' side of Γ_{int} (chosen arbitrarily), and $\nabla T^+ - \nabla T^-$ is the difference between the limits of gradients of T as Γ_{int} is approached from positive and negative sides.

In contrast to the formulation presented above, the streamline or selective upwinding method (SU) is not of the weighted residuals type. It can, however, be obtained from (6) by setting P_2 equal zero and deleting the Laplacian in P_1 and D:

$$P_1(W^d, \kappa) = \mathbf{v}\cdot\nabla W^d \tag{9a}$$

$$P_2(W^d, \kappa) = 0 \tag{9b}$$

$$D(T^d, \kappa) = \mathbf{v}\cdot\nabla T^d \tag{9c}$$

A question that arises immediately in the use of these methods is: what should be the value of the multiplier τ? For the one-dimensional constant coefficient homogeneous convection–diffusion problem with linear interpolation on regular elements of length d, analysis gives the well-known result (2), yielding nodally exact solutions:

$$\tau = \frac{1}{2}\frac{\alpha d}{v} \tag{10a}$$

$$\alpha_{\text{opt}} = \coth(Pe^d) - 1/Pe^d \tag{10b}$$

$$Pe^d = \frac{vd}{2\kappa} \quad \text{(mesh Peclet number)} \tag{10c}$$

No corresponding result is available for multi-dimensional problems, and *ad hoc* generalizations have been used. The structure used by Brooks and Hughes with bilinear elements is equivalent to taking the multiplier as:

$$\tau = \frac{\alpha_\xi d_\xi v_\xi + \alpha_\eta d_\eta v_\eta}{|\mathbf{v}|}\cdot\frac{1}{2|\mathbf{v}|} \tag{11a}$$

This is evaluated at integration points, using quantities expressed with respect to local co-ordinates:

$$\alpha_\xi = \coth(Pe_\xi^d) - \frac{1}{Pe_\xi^d} \tag{11b}$$

$$Pe_\xi^d = \frac{v_\xi d_\xi}{2\kappa} \tag{11c}$$

and correspondingly for the η-direction. The characteristic element dimensions d_ξ and d_η were defined for each element as, respectively, the distance between images of the (ξ, η) points $(-1, 0)$, $(+1, 0)$ and $(0, -1)$, $(0, +1)$ on the computational domain. Velocity components were obtained from:

$$v_\xi = \mathbf{e}_\xi\cdot\mathbf{v} \qquad v_\eta = \mathbf{e}_\eta\cdot\mathbf{v} \tag{12}$$

where $\mathbf{e}_\xi$ and $\mathbf{e}_\eta$ are unit vectors directed positively along the mapped local co-ordinate axes.

In isoparametric elements the local coordinate directions may vary from point to point, so in place of the above element-wise definition of unit vectors, we prefer to calculate integration point values as follows:

$$\mathbf{e}_\xi(\xi_G, \eta_G) = (x_I \partial_\xi W_I, y_I \partial_\xi W_I)/D_\xi^{1/2} \tag{13a}$$

$$D_\xi = (x_I \partial_\xi W_I)^2 + (y_I \partial_\xi W_I)^2 \tag{13b}$$

and correspondingly for $\mathbf{e}_\eta$. ∂_ξ denotes differentiation with respect to ξ, etc. Repeated subscripts

I indicate summation over nodes, and W_I is the weighting (shape) function at node I. (x_I, y_I) are the Cartesian nodal coordinates.

It is also convenient to use point-wise definitions of element characteristic dimensions in the x and y directions:

$$d_x(\xi_G, \eta_G) = l|x_I \partial_\xi W_I + x_I \partial_\eta W_I| \tag{14a}$$

$$d_y(\xi_G, \eta_G) = l|y_I \partial_\xi W_I + y_I \partial_\eta W_I| \tag{14b}$$

where $l = 1$ for biquadratic elements, and $l = 2$ for bilinear. This gives similar dimensions for a biquadratic element and for the 4 bilinear elements formed on the same 9 nodes. Local coordinate forms are then obtained as:

$$d_\xi = \mathbf{e}_\xi \cdot \mathbf{d} \qquad d_\eta = \mathbf{e}_\eta \cdot \mathbf{d} \tag{15}$$

The question of optimal upwinding parameters for use with biquadratic elements has been considered by Donea *et al.*[12]. From analysis of the simplified one-dimensional problem, they found that α, as given by (10b), is appropriate for centre (mid-side) nodes, but that for inter-element (corner) nodes, $\frac{1}{2}\alpha$ is replaced by β:

$$\beta = \frac{2 - \cosh(2Pe^d) - 2(Pe^d)^{-1}\tanh(Pe^d) + (2Pe^d)^{-1}\sinh(2Pe^d)}{4\tanh(Pe^d) - \sinh(2Pe^d) - 3(Pe^d)^{-1}\sinh(2Pe^d)\tanh(Pe^d)} \tag{16}$$

These results were generalized to multi-dimensions using a vectorial multiplier, λ, leading to a formulation in which the discontinuous part of the weighting $\tau \mathbf{v} \cdot \nabla W$ in SUPG (6) is replaced by $\lambda \cdot \nabla W$. The components of λ are defined in terms of α or β, depending upon whether mid-side or corner character is appropriate. Donea *et al.* write the expansion of $\lambda \cdot \nabla W$ in Cartesian components; however, it seems to us that mid-side or corner character of the components of λ can only be defined by reference to local coordinates. We thus write the discontinuous part of the weighting for nodes of the element illustrated in *Figure 1* as:

Node 1 $\quad \lambda_\xi^c W_{1,\xi} + \lambda_\eta^c W_{1,\eta}$

Node 2 $\quad \lambda_\xi^M W_{2,\xi} + \lambda_\eta^c W_{2,\eta}$

Node 3 $\quad \lambda_\xi^c W_{3,\xi} + \lambda_\eta^c W_{3,\eta}$

Node 4 $\quad \lambda_\xi^c W_{4,\xi} + \lambda_\eta^M W_{4,\eta}$

Node 5 $\quad \lambda_\xi^c W_{5,\xi} + \lambda_\eta^c W_{5,\eta}$

Node 6 $\quad \lambda_\xi^M W_{6,\xi} + \lambda_\eta^c W_{6,\eta}$

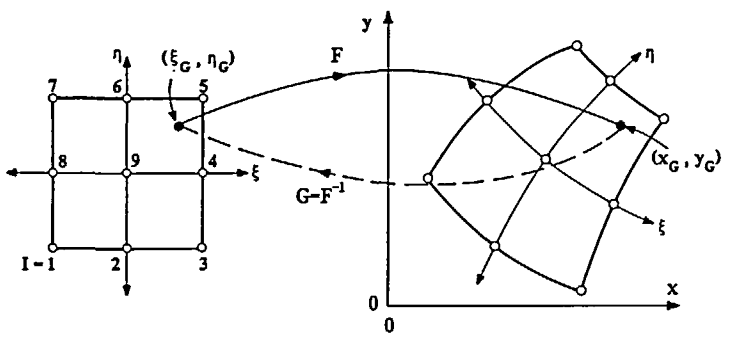

Figure 1 Transformation between local and global co-ordinate systems. A local node numbering system is shown on the reference element

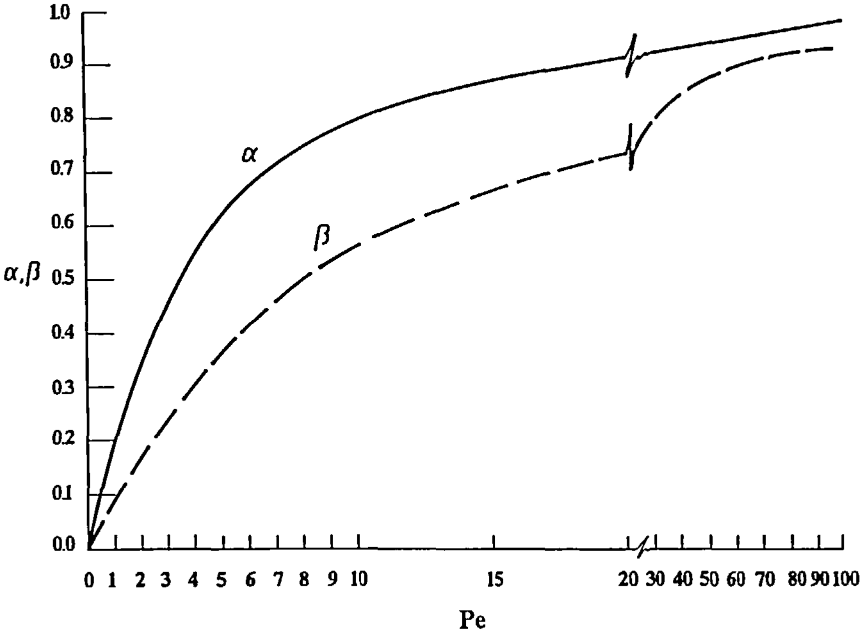

Figure 2 The 'optimal' upwind parameters α and β. (Both are odd functions of Peclet number, *Pe*)

$$
\begin{aligned}
&\textit{Node 7} \quad \lambda_\xi^c W_{7,\xi} + \lambda_\eta^c W_{7,\eta} \\
&\textit{Node 8} \quad \lambda_\xi^c W_{8,\xi} + \lambda_\eta^M W_{8,\eta} \\
&\textit{Node 9} \quad \lambda_\xi^M W_{9,\xi} + \lambda_\eta^M W_{9,\eta}
\end{aligned} \tag{17a}
$$

where

$$\lambda_\xi^M = \tfrac{1}{2}\alpha_\xi d_\xi, \qquad \lambda_\xi^c = \beta_\xi d_\xi \tag{17b}$$

and correspondingly for the η components. The functions α and β are shown in *Figure 2.*

It should be noted that the Brooks–Hughes SUPG involves a tensorial diffusivity-like term $k\mathbf{v}\,\mathbf{v}/|\mathbf{v}|^2$, with a corresponding flux $-k\mathbf{v}(\mathbf{v}\cdot\nabla T)$ directed strictly along the streamline. The Donea form, however, introduces a diffusivity $\lambda\mathbf{v}$ giving a flux $-\lambda(\mathbf{v}\cdot\nabla T)$. λ is not in general parallel to $\mathbf{v}$, so the possibility of artificial diffusion across streamlines exists. This occurrence was considered a serious disadvantage of early upwinding methods. However, as can be seen from *Figure 2,* α and β are approximately linear functions for low Peclet number, and in this case, at nodes 1, 3, 5, 7 and 9 in *Figure 1,* λ will be approximately parallel to $\mathbf{v}$, provided $d_\xi \approx d_\eta$. At nodes 2, 4, 6 and 8, where we have the mixed mid-side and corner forms, λ will be to one side or the other of $\mathbf{v}$, with, presumably, some possibility for cancellation of cross-streamline diffusion over the element. The situation at high Peclet, when $\alpha, \beta \rightarrow 1$, seems less favourable, with the direction of λ being determined by the relative magnitudes of d_ξ and d_η.

In view of these concerns we have experimented with a version of the Brooks–Hughes form, (11a) for biquadratic elements, that incorporates the Donea optimal parameters. Best results, which will be described later, were obtained by blending the λ components in a second order fashion, as follows (the components are taken as the corner or mid-side forms as appropriate):

$$\tau = \frac{|\lambda_\xi| v_\xi^2 + |\lambda_\eta| v_\eta^2}{|\mathbf{v}|^2} \cdot \frac{1}{|\mathbf{v}|} \tag{18}$$

This, of course, reduces to the expected one-dimensional result, but otherwise gives slightly lower values of τ than the first order blending corresponding to (11a).

Finally, for comparison with these various 'optimal' forms, we have carried out solutions with the multiplier defined as:

$$\tau = c\frac{d_x v_x + d_y v_y}{|\mathbf{v}|}\cdot\frac{1}{|\mathbf{v}|} \tag{19}$$

where c is constant over all elements and is chosen by numerical experiment.

THE LAPLACIAN FOR ISOPARAMETRIC BILINEAR AND BIQUADRATIC ELEMENTS

As can be seen from (6) it is necessary to calculate integrals over element interiors of the Laplacian of the shape functions, and as a consequence of the use of numerical integration, values of the Laplacian are required at the integration points. The way to obtain these, in isoparametric elements, is not immediately obvious.

In order to derive proper formulae it is convenient first to depict the transformations involved, in *Figure 1*.

Let us denote by F the one-to-one transformation from the local to the global co-ordinate system:

$$F:\quad x = \tilde{x}(\xi, \eta), \qquad y = \tilde{y}(\xi, \eta) \tag{20}$$

and the corresponding inverse transformation as follows:

$$G = F^{-1}:\quad \xi = \tilde{\xi}(x, y), \qquad \eta = \tilde{\eta}(x, \eta) \tag{21}$$

For isoparametric elements we have:

$$\tilde{x}(\xi, \eta) = W_J(\xi, \eta)x_J, \qquad \tilde{y}(\xi, \eta) = W_J(\xi, \eta)y_J \tag{22}$$

where summation on J is implied. Let us note that the functions $\tilde{x}$ and $\tilde{y}$ for bilinear and biquadratic elements are contained in the forms:

$$x = \tilde{x}(\xi, \eta) = \hat{x}_0 + \hat{x}_1\xi + \hat{x}_2\eta + \hat{x}_3\xi^2 + \hat{x}_4\xi\eta + \hat{x}_5\eta^2 + \hat{x}_6\xi^2\eta + \hat{x}_7\xi\eta^2 + \hat{x}_8\xi^2\eta^2 \tag{23a}$$

$$y = \tilde{y}(\xi, \eta) = \hat{y}_0 + \hat{y}_1\xi + \hat{y}_2\eta + \hat{y}_3\xi^2 + \hat{y}_4\xi\eta + \hat{y}_5\eta^2 + \hat{y}_6\xi^2\eta + \hat{y}_7\xi\eta^2 + \hat{y}_8\xi^2\eta^2 \tag{23b}$$

where

$$\hat{x}_0 = \tilde{x}(0, 0),\ \hat{x}_1 = \partial_\xi\tilde{x}(0, 0),\ \hat{x}_2 = \partial_\eta\tilde{x}(0, 0), \ldots, \hat{x}_7 = \frac{1}{3!}3\partial^3_{\xi\eta\eta}\tilde{x}(0, 0),\ \hat{x}_8 = \frac{1}{4!}6\partial^4_{\xi\xi\eta\eta}\tilde{x}(0, 0)$$

Let us develop the functions $\tilde{x}$ and $\tilde{y}$ as Taylor series in the neighbourhood of a point (ξ_G, η_G).

$$\Delta x = \Delta\tilde{x}(\Delta\xi, \Delta\eta) = a_1\Delta\xi + a_2\Delta\eta + a_3(\Delta\xi)^2 + a_4(\Delta\xi)(\Delta\eta) + a_5(\Delta\eta)^2 + \cdots + a_8(\Delta\xi)^2(\Delta\eta)^2 \tag{24a}$$

$$\Delta y = \Delta\tilde{y}(\Delta\xi, \Delta\eta) = b_1\Delta\xi + b_2\Delta\eta + b_3(\Delta\xi)^2 + b_4(\Delta\xi)(\Delta\eta) + b_5(\Delta\eta)^2 + \cdots + b_8(\Delta\xi)^2(\Delta\eta)^2 \tag{24b}$$

where $\Delta\xi = \xi - \xi_G$, $\Delta\eta = \eta - \eta_G$

$$\Delta x = \tilde{x}(\xi, \eta) - \tilde{x}(\xi_G, \eta_G) \qquad \Delta y = \tilde{y}(\xi, \eta) - \tilde{y}(\xi_G, \eta_G)$$

and

$$\begin{aligned} a_1 &= \partial_\xi\tilde{x}(\xi_G, \eta_G), a_2 = \partial_\eta\tilde{x}(\xi_G, \eta_G), \ldots, a_8 = \tfrac{1}{4}\partial^4_{\xi\xi\eta\eta}\tilde{x}(\xi_G, \eta_G) \\ b_1 &= \partial_\xi\tilde{y}(\xi_G, \eta_G), b_2 = \partial_\eta\tilde{y}(\xi_G, \eta_G), \ldots, b_8 = \tfrac{1}{4}\partial^4_{\xi\xi\eta\eta}\tilde{y}(\xi_G, \eta_G) \end{aligned} \tag{24d}$$

Other derivatives vanish identically on element interiors as a consequence of (23). Of course, we can easily find relationships between the coefficients $\hat{x}_i$ and a_i, and $\hat{y}_i$ and b_i, by noting that $\partial_\xi x = \partial_\xi \Delta x$ etc. and differentiating (23). For example,

$$a_4 = \hat{x}_4 + 2\hat{x}_6\xi_G + 2\hat{x}_7\eta_G + 4\hat{x}_8\xi_G\eta_G \tag{25}$$

However, it is not necessary to pursue this, as the derivatives can be obtained via shape functions in the usual finite element way.

Because the transformation F has an inverse G, i.e.

$$J = \det\begin{bmatrix} \partial_\xi\tilde{x}(\xi_G, \eta_G) & \partial_\eta\tilde{x}(\xi_G, \eta_G) \\ \partial_\xi\tilde{y}(\xi_G, \eta_G) & \partial_\eta\tilde{y}(\xi_G, \eta_G) \end{bmatrix} \neq 0 \tag{26a}$$

and if

$$F(\xi_G, \eta_G) = (\tilde{x}(\xi_G, \eta_G), \tilde{y}(\xi_G, \eta_G)) = (x_G, y_G) \tag{26b}$$

then

$$F(G(x_G, \eta_G)) = (x_G, y_G), \qquad (FF^{-1} = id) \tag{26c}$$

It follows that we can also develop the functions $\tilde{\xi}$ and $\tilde{\eta}$ as Taylor series in the sufficiently small neighbourhood of the point (x_G, y_G) giving:

$$\Delta\xi = \Delta\tilde{\xi}(\Delta x, \Delta y) = \alpha_1\Delta x + \alpha_2\Delta y + \alpha_3(\Delta x)^2 + \alpha_4(\Delta x)(\Delta y) + \alpha_5(\Delta y)^2 + \alpha_6(\Delta x)^3 + \alpha_7(\Delta x)^2(\Delta y) + \cdots \tag{27a}$$

$$\Delta\eta = \Delta\tilde{\eta}(\Delta x, \Delta y) = \beta_1\Delta x + \beta_2\Delta y + \beta_3(\Delta x)^2 + \beta_4(\Delta x)(\Delta y) + \beta_5(\Delta y)^2 + \beta_6(\Delta x)^3 + \beta_7(\Delta x)^2(\Delta y) + \cdots \tag{27b}$$

where

$$\Delta\xi = \tilde{\xi}(x, y) - \tilde{\xi}(x_G, y_G), \qquad \Delta\eta = \tilde{\eta}(x, y) - \tilde{\eta}(x_G, y_G)$$
$$\Delta x = x - x_G, \qquad \Delta y = y - y_G$$
$$\alpha_1 = \partial_x\tilde{\xi}(x_G, y_G), \alpha_2\tilde{\xi}(x_G, y_G), \alpha_3 = \tfrac{1}{2}\partial^2_{xx}\tilde{\xi}(x_G, y_G), \ldots, \alpha_5 = \tfrac{1}{2}\partial^2_{yy}\tilde{\xi}(x_G, y_G)$$
$$\beta_1 = \partial_x\tilde{\eta}(x_G, y_G), \beta_2 = \partial_y\tilde{\eta}(x_G, y_G), \beta_3 = \tfrac{1}{2}\partial^2_{xx}\tilde{\eta}(x_G, y_G), \ldots, \beta_5 = \tfrac{1}{2}\partial^2_{y}\tilde{\eta}(x_G, y_G) \tag{27c}$$

We now write expressions for the Cartesian derivatives of the shape function, (recalling that this is a compound function):

$$\partial_x W = (\partial_\xi W)\alpha_1 + (\partial_\eta W)\beta_1, \qquad \partial_y W = (\partial_y W)\alpha_2 + (\partial_y W)\beta_2 \tag{28a}$$

$$\partial^2_{xx} W = (\partial^2_{\xi\xi} W)\alpha_1^2 + 2(\partial^2_{\xi\eta} W)\alpha_1\beta_1 + (\partial^2_{\eta\eta} W)\beta_1^2 + 2(\partial_\xi W)\alpha_3 + 2(\partial_\eta W)\beta_3 \tag{28b}$$

$$\partial^2_{yy} W = (\partial^2_{\xi\xi} W)\alpha_2^2 + 2(\partial^2_{\xi\eta})W\alpha_2\beta_2 + (\partial^2_{\eta\eta} W)\beta_2^2 + 2(\partial_\xi W)\alpha_5 + 2(\partial_\eta W)\beta_5 \tag{28c}$$

To calculate the coefficients $\alpha_1, \ldots, \alpha_5$ and $\beta_1, \ldots, \beta_5$ we substitute (27) into (24), and as a consequence of (26) obtain an identity.

Then, comparing coefficients of corresponding powers of Δx and Δy, we obtain for the first order terms:

$$\begin{aligned} 1 &= a_1\alpha_1 + a_2\beta_2, \qquad 0 = a_1\alpha_2 + a_2\beta_2 \\ 0 &= b_1\alpha_1 + b_2\beta_1, \qquad 1 = b_1\alpha_2 + b_2\beta_2 \end{aligned} \tag{29a}$$

giving

$$\alpha_1 = \frac{b_2}{J}, \qquad \alpha_2 = -\frac{a_2}{J}, \qquad \beta_1 = -\frac{b_1}{J}, \qquad \beta_2 = \frac{a_1}{J} \tag{29b}$$

Similarly for the second order terms:

$$\alpha_3 = \frac{A_1 b_2 - B_1 a_2}{J}, \qquad \alpha_5 = \frac{A_2 b_2 - B_2 a_2}{J}$$
$$\beta_3 = \frac{B_1 a_1 - A_1 b_1}{J}, \qquad \beta_5 = \frac{B_2 a_1 - A_2 b_1}{J} \tag{30a}$$

where

$$\begin{aligned} A_1 &= -(a_3\alpha_1^2 + a_4\alpha_1\beta_1 + a_5\beta_1^2) \\ A_2 &= -(a_3\alpha_2^2 + a_4\alpha_2\beta_2 + a_5\beta_2^2) \\ B_1 &= -(b_3\alpha_1^2 + b_4\alpha_1\beta_1 + b_5\beta_1^2) \\ B_2 &= -(b_3\alpha_2^2 + b_4\alpha_2\beta_2 + b_5\beta_2^2) \end{aligned} \tag{30b}$$

J is given in (26a), and the coefficients a_I and b_I are calculated in the usual way, e.g. $a_1 = W_{J,\xi}(\xi_G, \eta_G)x_J$ etc. Since the bilinear interpolation is contained within the biquadratic, the above relationships are valid for bilinear isoparametric elements, when

$$\begin{aligned} a_3 = a_5 = a_6 = a_7 = a_8 = 0 \\ b_3 = b_5 = b_6 = b_7 = b_8 = 0 \end{aligned} \tag{31}$$

The computational costs involved in the calculation are not high, because values of the shape function Laplacians at integration points need to be calculated only once, and then stored on disc. It should also be noted that the method used here gives exact values, and can be extended for higher order interpolation as well.

THE TEST PROBLEM

Let us consider the model steady, scalar convection–diffusion problem, illustrated in *Figure 3*, with differential equation:

$$\mathbf{K}\cdot\nabla T - \nabla^2 T = Q \tag{32}$$

If this is interpreted as the energy conservation equation, then $\mathbf{K}$ is identified as:

$$K_x = \frac{v_x}{\kappa}, \qquad K_y = \frac{v_y}{\kappa} \tag{33}$$

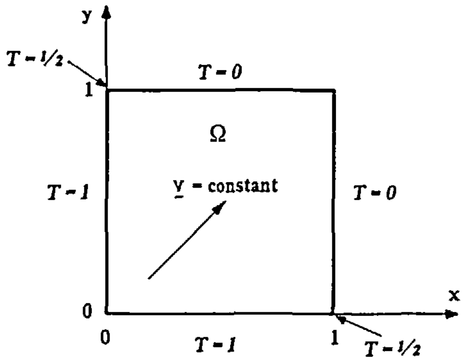

Figure 3 Test problem definition, with boundary conditions

and Q is a heat source term, corresponding for example to viscous heat generation. The following Dirichlet boundary conditions are used:

$$\begin{aligned} &x=0, \quad 0\leqslant y<1: && T(x,y)=1\\ &y=0, \quad 0\leqslant x<1: && T(x,y)=1\\ &x=1, \quad 0<y\leqslant 1: && T(x,y)=0\\ &y=1, \quad 0<x\leqslant 1: && T(x,y)=0\\ &x=0, \quad y=1: && T(x,y)=1/2\\ &x=1, \quad y=0: && T(x,y)=1/2 \end{aligned} \tag{34}$$

For significant values of the Peclet number we can expect steep exit boundary layers, and when discretization is such that mesh Peclet numbers

$$Pe_x^d=\frac{K_x d_x}{2} \qquad Pe_y^d=\frac{K_y d_y}{2} \tag{35}$$

(where d_x, d_y are characteristic element dimensions in x and y directions) are also large, but not necessarily very large compared to unity, we expect the classical Galerkin finite element method to yield oscillatory results. The problem poses a severe test for methods seeking to deal correctly with 2-dimensional convection–diffusion involving sharp flow direction gradients, and there also exists opportunity for spurious cross-streamline diffusion.

For this problem, series solutions are available, which may be evaluated to high accuracy, for the cases of zero and non-zero source term.

(*a*) $Q=0$. The solution is obtainable using separation of variables[13]:

$$T(x,y)=8\pi\exp\left(\frac{K_x x+K_y y}{2}\right)\sum_{n=1}^{\infty}\frac{n}{1-e^{\alpha_n}}\left[\frac{1-(-1)^n e^{-K_x/2}}{K_x^2+4\pi^2n^2}\sin(\pi n x)(e^{-\alpha_n y/2}-e^{-\alpha_n(1-y/2)})+\frac{1-(-1)^n e^{-K_y/2}}{K_y^2+4\pi^2n^2}\sin(\pi n y)(e^{-\alpha_n x/2}-e^{-\alpha_n(1-x/2)})\right]$$

where

$$\alpha_n=\sqrt{K_x^2+K_y^2+4\pi^2n^2} \tag{36}$$

(*b*) $Q\neq 0$. We extend the series solution to the case of a non-zero source term in the following way. Let ϕ satisfy (32) with $Q=0$ on the domain Ω, with boundary conditions on Γ as stated in (34). Let

$$g\in C^2, \qquad g(x,y)=0 \quad \text{for } (x,y)\in\Gamma \tag{37}$$

Define

$$T=\phi+g \tag{38}$$

Then T satisfies the boundary conditions, and it can easily be shown that T is a solution of (32) if:

$$Q=K_x\partial_x g+K_x\partial_y g-\partial_{xx}^2 g-\partial_{yy}^2 g \tag{39}$$

For instance, if we choose

$$g(x,y)=Hx(1-x)y(1-y) \tag{40}$$

Then

$$Q(x,y)=H[y(1-y)(K_x(1-2x)+2)+x(1-x)(K_y(1-2y)+2)] \tag{41}$$

This is the form used in tests with a non-zero source term.

The series solutions were evaluated, for comparison with numerical results, by summing until:

$$n > 3000 \qquad \text{and} \qquad \left| \frac{(\text{term})_{n+1}}{\sum_{1}^{n}} \right| < 10^{-5} \tag{42}$$

Selected tests showed that

$$\left| \sum_{1}^{3,000} - \sum_{1}^{10,000} \right| < 10^{-10} \tag{43}$$

Note that (36) is written avoiding positive exponents, which would cause computer overflow.

RESULTS AND DISCUSSION

Cases studied and finite element implementation

Meshes with the same 121 nodes, symmetrically placed about the diagonal $x = y$ of the square domain, were used for both bilinear and biquadratic elements. As seen in *Figure 4*, elements are chosen not all to be rectangular, to provide some test of the algorithms on non-orthogonal isoparametric elements. The Figure also shows the lines I and II along which nodal values were selected for comparison with series solution results.

Some additional computations were carried out with biquadratic elements on the refined mesh shown in *Figure 5*, obtained by a 2 × 2 subdivision of the previous mesh.

Solutions to the test problem were carried out, with parameter values $K_x = K_y = 80$ (33) and $H = 10$ (40), giving flow at 45° to the axes, and an appreciable effect from the source term. Analytic results are shown in *Figure 6*, along the mesh lines I and II indicated in *Figures 4* and *5*.

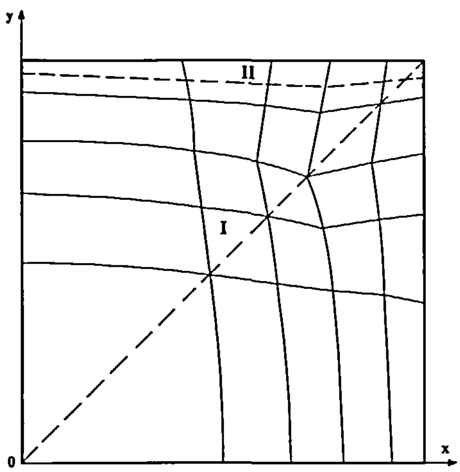

Figure 4 Mesh of biquadratic elements, coarser version, used for solution of the test problem, showing lines I and II on which numerical results are examined. A bilinear mesh is formed using the same nodes

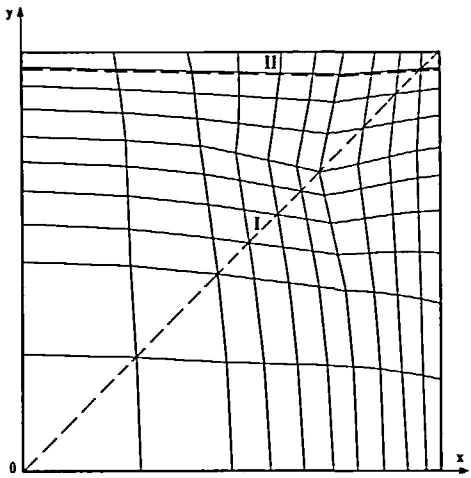

Figure 5 Refined mesh (biquadratic elements) for solution of the test problem

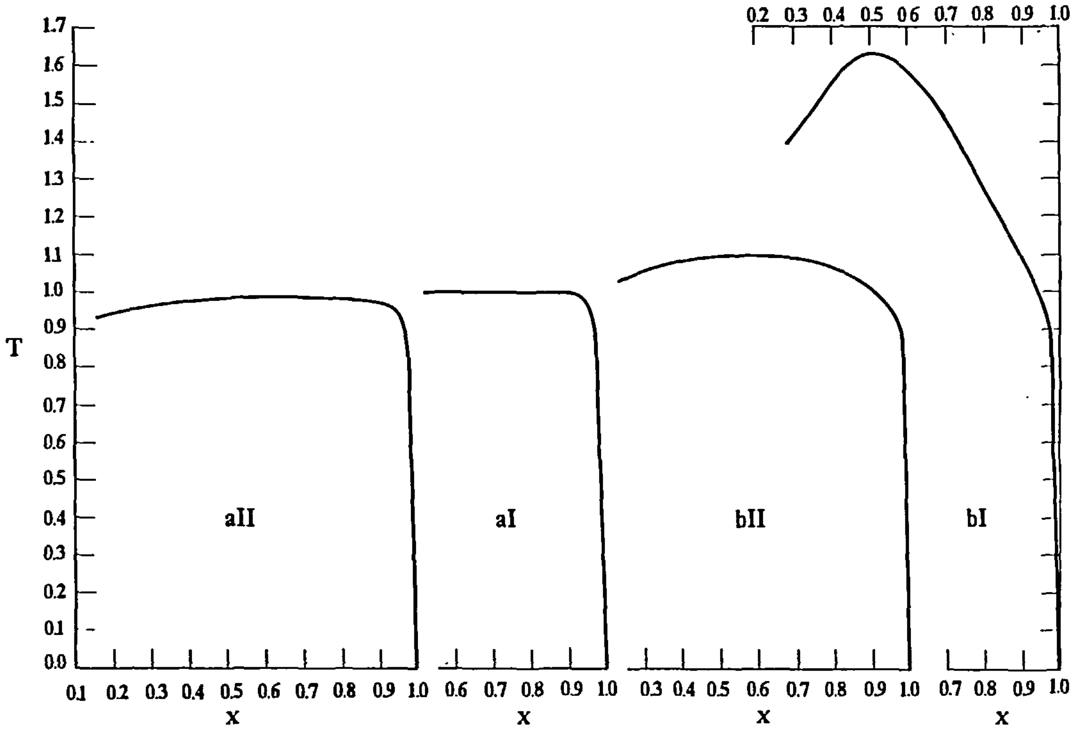

Figure 6 Analytic solution to the test problem on lines I and II (see *Figures 4* and *5*). (a) Source term zero; (b) source term non-zero

The resulting minimum mesh Peclet numbers, occurring in the exit region, are 4 for biquadratic and bilinear elements on the coarse mesh, and 2 on the refined mesh.

Finite element calculations were carried out using 64 bit working.

Computations on the coarser mesh (Figure 4) *constant upwind parameter* (*19*)

We first consider methods implemented using the multiplier τ defined according to (19), and examine the dependence of results on the constant upwinding parameter c.

SU method. Computed results are displayed in terms of plots of departures from the analytic solutions. *Figures 7a* and *b* refer to SU on biquadratic, and *Figures 8a* and *b* to bilinear elements. sharing the same nodes. The biquadratic results are highly oscillatory, and although the amplitude of the oscillations decreases as upwinding constant c is increased, the solution is then overdamped showing significant negative errors. It is apparent that no value of c gives solutions that are both smooth and accurate. This is anticipated, as a consequence of the non-consistent upwind weighting, which is particularly noticeable in the presence of the source term, and which is equivalent to the introduction of a streamline directed artificial diffusivity.

For the zero source term case, SU on the bilinear isoparametric elements is virtually equivalent to the consistent SUPG formulation, on account of the small values of shape function Laplacian; hence the reasonable results obtained for this case on the diagonal, line I (*Figure 8a*). When the source term is non-zero (*Figure 8b*) sharply increased negative errors clearly show the consequences of inconsistent weighting, with overdamping by artificial diffusivity.

Accuracy on line II is overall less good, since all these nodes lie within the steep exit boundary layer; additionally, the proximity of the corner (0, 1), where a discontinuity in the boundary

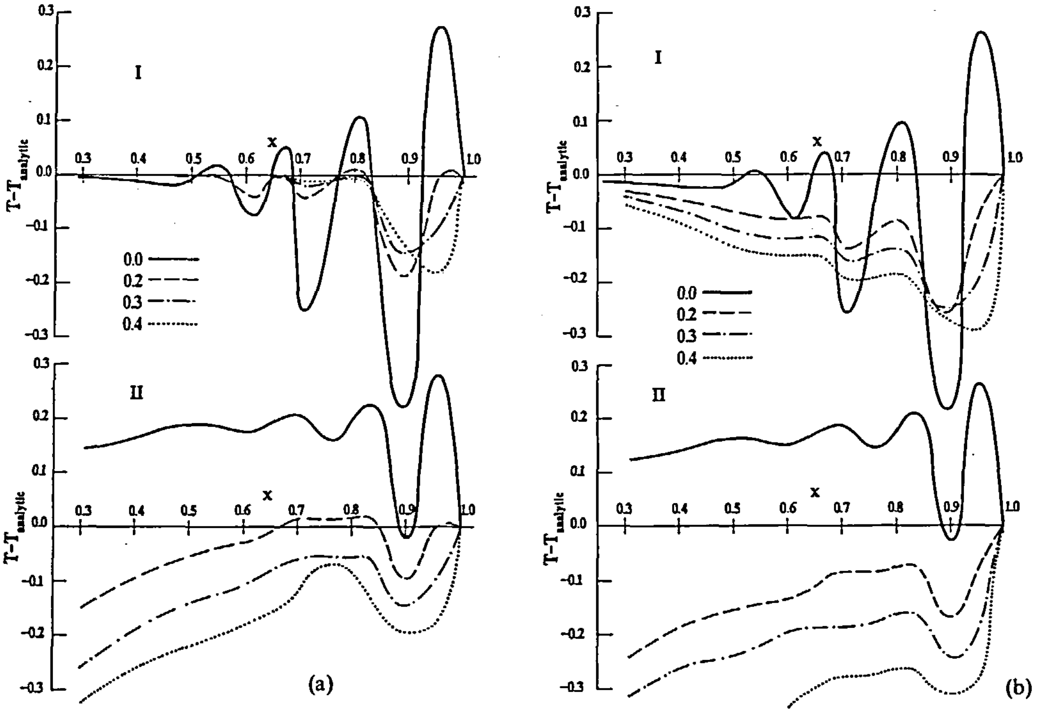

Figure 7 Departure of SU results from analytic solution. Biquadratic elements. Mesh as *Figure 4*. (a) Source term zero; (b) source term non-zero. Parameter: upwinding constant *c*

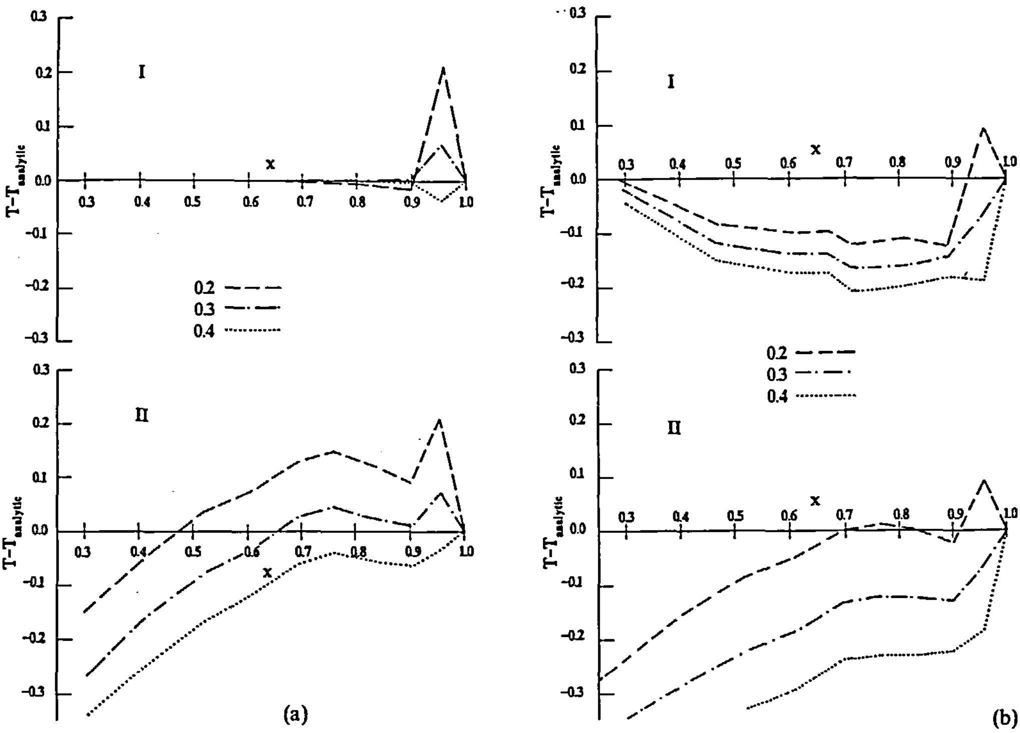

Figure 8 Departure of SU results from analytic solution. Bilinear elements. Mesh nodes as *Figure 4*. (a) Source term zero; (b) source term non-zero. Parameter: upwinding constant *c*

conditions of the test problem exists, produces larger discrepancies at small x. As will be seen, a finer mesh is required to resolve the solution satisfactorily along line II.

SUPG. We first consider results using the upwind structure based on (19) for biquadratic elements. Here, errors for the computations with and without source term are almost identical, and plots are shown only for the case with non-zero source (*Figure 9*). Results are significantly improved, as compared with those from SU (*Figures 7a* and *b*) as anticipated, particularly in the presence of the source term. However, oscillations persist at higher values of c, when solutions show a tendency towards overdamping, though much less pronounced than in the SU calculations.

The pattern of behaviour exhibited by the bilinear results is very similar, (*Figure 10*) though overall less oscillatory as a consequence of the absence of higher order terms in the interpolation.

GLS. Again, GLS copes equally well in the cases with and without source term, and we display only the results for the latter case. As compared with SUPG, GLS on biquadratic elements yields marginally better results (*Figure 11*) with errors being a little more evenly balanced negative and positive. Using bilinear elements, virtually no difference from SUPG was seen, as expected (Laplacian terms negligible), and these results are not displayed.

Coarse mesh, optimal upwind formulations

We now consider the various 'optimal' formulations: the Brooks–Hughes form for bilinear

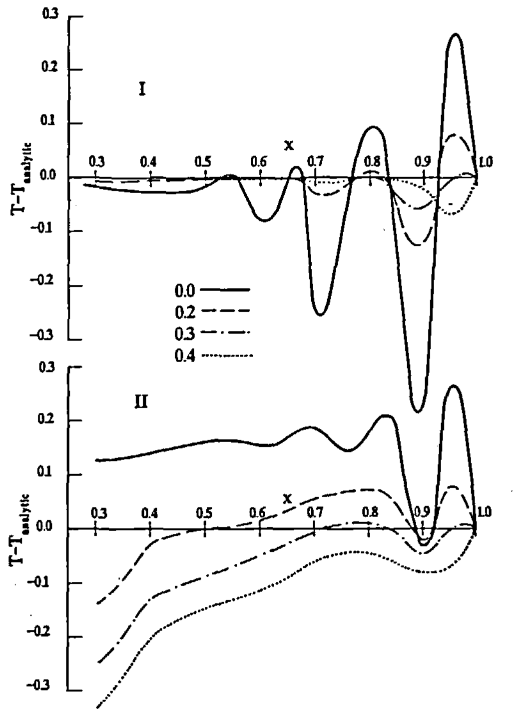

Figure 9 Departure of SUPG results from analytic solution. Biquadratic elements. Mesh as *Figure 4*. Source term non-zero. Parameter: upwinding constant c

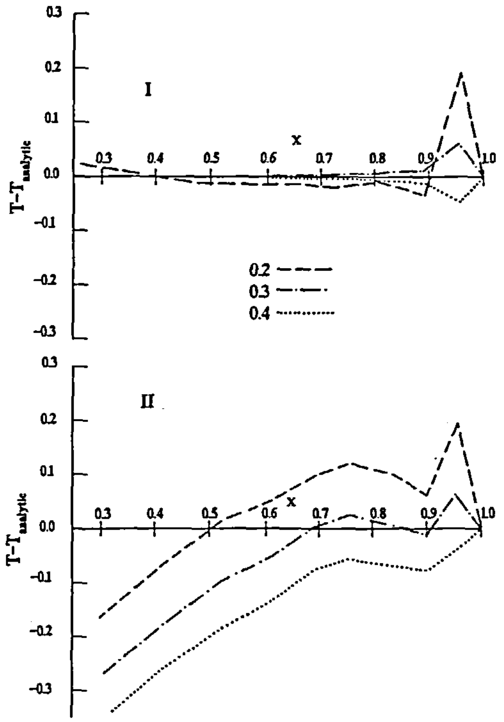

Figure 10 Departure of SUPG results from analytic solution. Bilinear elements. Mesh nodes as *Figure 4*. Source term non-zero. Parameter: upwinding constant c

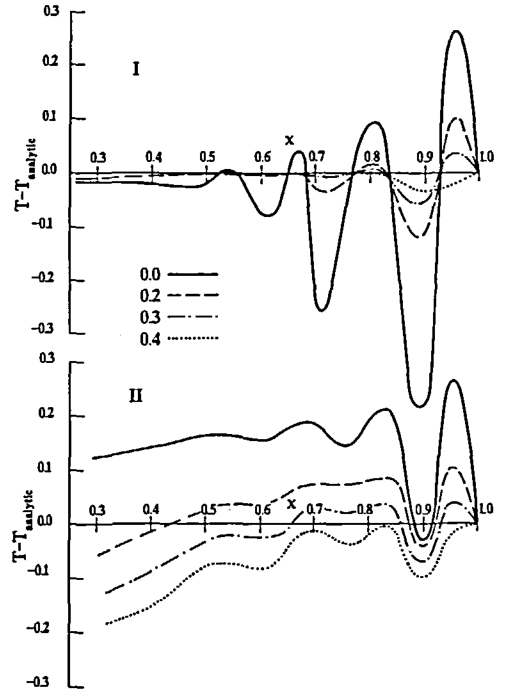

Figure 11 Departure of GLS results from analytic solution. Biquadratic elements. Mesh as *Figure 4*. Source term non-zero. Parameter: upwinding constant c

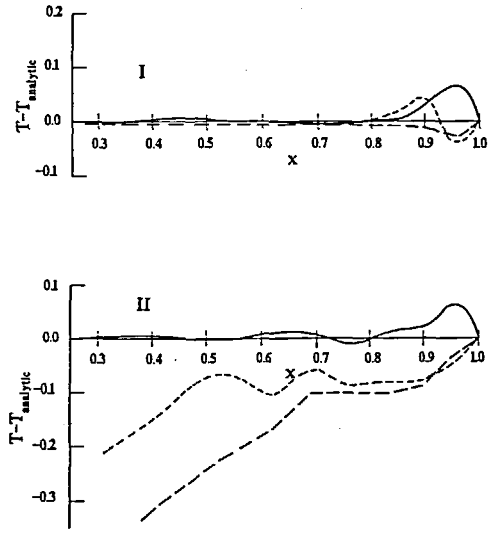

Figure 12 Departure of 'optimal upwind' SUPG results from analytic solution. Source term non-zero. Mesh as *Figure 4*. ---, Optimal SUPG on bilinear elements; ———, Donea *et al.* form on biquadratic elements; -----, streamline form on biquadratic elements

elements, (11)–(15); the Donea form for biquadratic elements, (17); and the strictly streamline form incorporating Donea optimal parameters for biquadratic elements, (18).

Figure 12 compares results from the three formulations. The optimal bilinear form shows little improvement over the best result obtained using a constant upwind parameter ($c = 0.4$, *Figure 10*); however, in the absence of other guidance about how to select the best value of c, it is of obvious value.

The Donea form for biquadratic elements produces impressive results with, for the first time, relatively good performance in the edge region, on line II. Similar accuracy was obtained in a further run (not shown), where the flow angle was varied from 45° to 26° degrees from x-axis. In view of the concerns expressed earlier about the possibility of cross-streamline diffusion in this formulation, it is probably not safe to generalize from these results, and further testing on problems with less simple flow fields is desirable.

The strictly streamline form for biquadratic elements (18) gives results on the diagonal, (line I) that are comparable in accuracy to the Donea formulation, but performs significantly less well in the edge region, (line II). Noticeably less good results, which are not shown, were obtained with a first order blending of λ components, as used in (11a).

Computations on the refined mesh (Figure 5), *constant upwind parameter* (*equation 19*)

SU calculations on the refined mesh showed decreased oscillations, with errors under half the magnitudes obtained on the coarser mesh. However, qualitative behaviour was identical, with

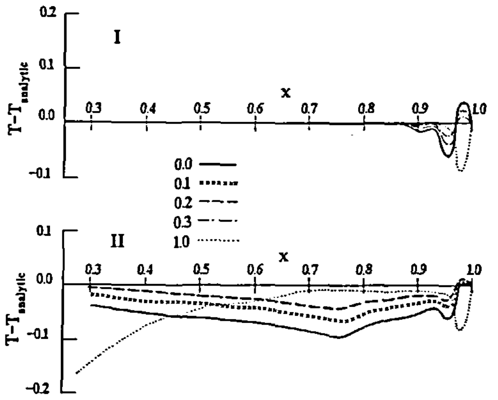

Figure 13 Departure of SUPG results from analytic solution. Biquadratic elements on refined mesh, *Figure 5*. Parameter: upwinding constant c

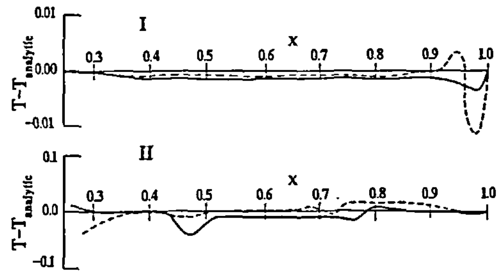

Figure 14 Departure of 'optimal upwind' SUPG results from analytic solution. Biquadratic elements on refined mesh, *Figure 5*. Note: the scale for results on line I is expanded ×10 compared with previous plots. ———, Donea *et al.* form; -----, streamline form

overdamping of the solution occurring before complete elimination of oscillations, and the results are not shown.

SUPG results on biquadratic elements are more significant, and are shown in *Figures 13*. Accuracy, as expected, is substantially improved, and it is interesting to note that in the edge region, on line II, the exact solution is now approached from below as c is increased from 0 to 0.2 (cf. *Figure 9*, where numerical results fall as c rises from 0 to 0.4). The refined mesh solution for $c = 0.1$ shows little oscillation, and in the absence of knowledge of the exact result, could well be taken as satisfactory. It is clear, though, that a better solution is obtained using a somewhat higher value of c. Thus lack of obvious oscillations cannot be taken as a safe indicator that the best possible solution has been calculated. Note that for a much higher value, $c = 1.0$, the solution deteriorates sharply near the exit corners, indicating the existence of an optimum value.

It should be noted that although mesh Peclet numbers in the critical exit region are around 2 on this refined mesh, significant oscillations still occur in the solution given by the classical Galerkin formulation.

Refined mesh, 'optimal' formulations

Figure 14 compares the Donea and strictly streamline formulations for biquadratic elements. On the diagonal, the latter shows somewhat larger errors in the exit boundary layer, but is nevertheless within 1% of the analytic solution (note the expanded scale of the plot). In the edge region, line II, both methods are of comparable, high accuracy.

CONCLUSIONS

A number of formulations involving discontinuous weighting functions have been applied to a quite severe convection–diffusion test problem. The success of SUPG on bilinear elements, particularly when 'optimal' parameters are incorporated, as proposed by Brooks & Hughes[2], has been confirmed. Nevertheless, extension of these methods to isoparametric bilinear, and particularly to biquadratic elements, is of interest, because of the geometrical flexibility of these elements, and the generally greater efficiency of biquadratic interpolation. Implementation requires calculation of integration point values of shape function Laplacians, and a convenient, exact procedure has been given that is applicable to isoparametric bilinear, biquadratic and higher order elements.

The GLS formulation showed little advantage over SUPG on biquadratic elements, and the emphasis has thus been on investigating SUPG, particularly with the incorporation of 'optimal' upwind parameters.

The generalized Galerkin method, proposed by Donea *et al.*[12] for biquadratic elements, gave remarkably good results, even on a relatively coarse mesh. However, because of concerns that this is not a strictly streamline method, a form of SUPG was proposed, having a similar structure to that used by Brooks and Hughes with bilinear elements, and incorporating the optimal parameters of Donea *et al.* On a sufficiently refined mesh (but one where the classical Galerkin formulation still fails) this gives excellent results.

The test problem used in this work was necessarily idealized, in order to obtain analytic solutions; in particular, the velocity field is trivially simple. Generalization from our results should, therefore, be made only with caution; nevertheless, we believe that it has been shown how the SUPG method can be extended successfully to higher order and isoparametric elements.

REFERENCES

1 Henrich, J. C. and Zienkiewicz, O. C. The finite element method and 'upwinding' techniques in the numerical solution of convection dominated flow problems, in *Finite Element Methods for Convection Dominated Flows* (Ed. T. J. R. Hughes), pp. 105–136 (1980)

2 Brooks, A. N. and Hughes, T. J. R. Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, *Comp. Meth. Appl. Mech. Eng.* **32**, 199–259 (1982)

3 Hughes, T. J. R. and Brooks, A. A theoretical framework for Petrov–Galerkin methods with discontinuous weighting functions: application to the streamline-upwind procedure, in *Finite Elements in Fluids*, Vol. 4, Wiley, Chichester, pp. 47–65 (1982)

4 Hughes, T. J. R., Mallet, M. and Mizukanin, A. A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, *Comp. Meth. Appl. Mech. Eng.* **54**, 341–355 (1986)

5 Hughes, T. J. R. and Mallet, M. A new finite element formulation for computational fluid dynamics: III. The generalised streamline operator for multidimensional convection-diffusion systems, *Comp. Meth. Appl. Mech. Eng.* **58**, 305–328 (1986)

6 Hughes, T. J. R. Recent progress in the development and understanding of SUPG methods with special reference to the compressible Euler and Navier–Stokes equations, *Int. J. Num. Meth. Eng.* **7**, 1261–1275 (1987)

7 Johnson, C., Nävert, U. and Pitkärenta, J. Finite element methods for linear hyberbolic problems, *Comp. Meth. Appl. Mech. Eng.* **45**, 285–312 (1984)

8 Johnson, C. Streamline diffusion methods for problems in fluid mechanics, in: *Finite Elements in Fluids* (Eds. R. H. Gallagher, G. F. Carey, J. T. Oden and O. C. Zienkiewicz), Vol. 6, Wiley, Chichester, pp. 251–261 (1986)

9 Johnson, C. and Saranen, J. Streamline diffusion methods for incompressible Euler and Navier–Stokes equations, *Math. Comput.* **4**, 1–18 (1986)

10 France, L. P. and Hughes, T. J. R. Two classes of mixed finite element methods, *Comp. Meth. Appl. Mech. Eng.* **69**, 89–129 (1988)

11 Hughes, T. J. R., Franca, L. P. and Hulbert, G. M. A new finite element formulation for computational fluid dynamics: VII. The Galerkin/least squares method for convective-diffusive equations, *Comp. Meth. Appl. Mech. Eng.* **73**, 173–189 (1989)

12 Donea, J., Belytschko, T. and Smolinski, P. A generalized Galerkin method for steady convection diffusion problems with application to quadratic shape function elements, *Comp. Meth. Appl. Mech. Eng.* **48**, 25–43 (1985)

13 Christie, I. Finite element methods for conduction–convection problems, *PhD Thesis*, University of Dundee (1977)