Supplement 2

Stabilized methods and SUPG

14.1 Introduction

We have employed several weighted residual methods to formulate our finite element solutions. Most of the time we have used the standard Galerkin method where we multiply a residual error by a single special spatial weighting function. Within an element we assumed a spatial interpolation for the approximate solution as

\[ u(x) = H(x) \sum_j H_j(x) U_j^e \]  

which in turn defines the residual error

\[ R(H(x)) \neq 0. \]  

The standard Galerkin method (sometimes called the Bubnov-Galerkin) is said to be a process that ‘makes the spatial approximation orthogonal to the residual error’ by requiring the weighed system to be

\[ \int_{\Omega} W_j R(x) d\Omega = 0_j \]  

where the weights are defined to be

\[ W_j(x) \equiv H_j(x), \]  

the element spatial interpolation associated with node \( j \).

The Galerkin method works well for elliptical differential equations. However, when it was applied to other classes of differential equations it was often found to yield "unstable" solutions, i.e., solutions that exhibit non-physical spatial oscillations. Generally the standard Galerkin approach is seen to break down in problems with strong boundary layer effects. The analytical approach to such problems is usually called singular perturbation theory. Typically such problems have a data dependent coefficient multiplying the highest derivative. In common special cases that coefficient approaches zero and the nature of the equation changes because of the loss of the highest derivative term. Alternatively, one can view it as a reduction in the number of boundary conditions which causes a very rapid change in the solution (i.e., a boundary layer) near the region of the "lost" boundary condition data. Some analysts divide the differential equation by the coefficient multiplying the highest derivative term and thereby create an increased
emphasis on the lower order derivative terms in the equation. One of the first studies to successfully apply a new finite element theory "for second order problems with significant first derivatives" was the use of the Petrov-Galerkin method by Christie, et al, in 1976 [4]. Since then the Petrov-Galerkin methods [16, 27] have generally come to be known as ‘stabilized’ formulations because they prevent the spatial oscillations and sometimes yield nodally exact solutions where the classical Galerkin method would fail badly. They are also very important because they allow equal order element interpolation for mixed nodal variables, such as pressure and velocity.

14.2 Petrov-Galerkin method

The Petrov-Galerkin method is assumed to be more general because it does not restrict the weights to just the special case of the spatial distribution of the approximating solution, but adds some additional terms to it:

\[
\int_{\Omega} (W_j + \tau P_j) R(x) \, d\Omega = 0_j
\]

where the \( P_j \) denotes Pertov or ‘stabilization’ terms [10, 24]. In Eq. 14.5 the \( \tau \) multiplier was introduced to recognize that one would often need to account for the difference in units between \( W_j \) and \( P_j \) and to scale their relative importance in the solution. Here we will refer to each such \( \tau \) term as a ‘stabilization parameter’ [24, 25]. If we are going to allow the weighting of the residual error to be more general than the classic Galerkin approach we are faced with selecting a rational for the additional weights. Some methods have been tried and shown to work well for some classes of equations, such as the advective-diffusion equation [16]. Since advection means "to carry along" it often occurs in modeling various transport phenomena. One of the common applications is heat transfer with mass flow. That is usually referred to as a convection-diffusion problem. For such problem classes the Petrov-Galerkin method is often used to create ‘upwind elements’ as one technique to stabilize the solution [3, 5, 6, 8, 11, 12, 17]. One can find many articles on those subjects, but most employ linear elements and zero source terms and may hide more general concepts.

Consider a typical application such as convective-diffusive heat transfer governed by

\[
\rho c \left( \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi \right) = \nabla (K \nabla \phi) + Q
\]

where \( \mathbf{v} \) denotes a given velocity vector field, \( Q \) is the volumetric source and the material properties \( \rho, c, K \) are the mass density, specific heat and thermal conductivities, respectively. We select a generalized weight function

\[
w = (\phi + p)
\]

where \( p(x) \) is the new Petrov or stabilization term(s). Then we invoke the method of weighted residuals:

\[
\int_{\Omega} (\phi + p)[\rho c \left( \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi \right) - \nabla (K \nabla \phi) - Q] \, d\Omega = 0.
\]

Usually \( \phi(x) \) is taken as continuous across element boundaries and thus allows one to employ integration by parts to yield the terms given earlier in the classical Galerkin
Finite Element Analysis with Error Estimators

The Petrov term, \( p(x) \), may or may not be continuous across element boundaries. Usually it is not continuous and we cannot reduce the order of the derivatives in its integrals. In either case, we can view this expanded integral form as

\[
[\text{Classical Galerkin}] + [\text{Stabilization Terms}] = 0
\]  

(14.9)

where the stabilization term for continuous \( p(x) \) is

\[
I_s = \int_{\Omega} p(x) [\rho c \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi] - \nabla (K \nabla \phi) - Q \, d\Omega
\]  

(14.10)

and for discontinuous \( p(x) \) the above domain integral is replaced by the sum of the element domain integrals [12]. In either case this contribution is clearly zero for the exact solution. This is a typical example of a Petrov-Galerkin approach. Note that unless integration by parts can be employed this term retains the highest derivative found in the original equation. That would either increase the interpolation inter-element continuity requirement, or restrict the integral evaluation to each of the element domains rather than the full domain, \( \Omega \). The latter usually occurs. For example, when one includes a least squares weight (partial derivative of the residual error with respect to the unknowns) as a Petrov term [14]. That leads to a Galerkin/Least Squares (GLS) stabilization process [15].

It should be noted that in most low order elements the second derivatives are zero and thus the diffusion term is often omitted in the stabilization calculation. However, the second derivatives can always be estimated using patch methods or other techniques when using an iterative solution.

Since the Galerkin process works well for elliptical differential equations, but often not for an advective-diffusion equation, we review the standard Galerkin process and seek a change in it that may better capture advective-diffusion type solutions. The most common form of the Petrov-Galerkin method is to pick the weight for node \( j \) as

\[
W_j = h_j + \alpha F_j
\]  

(14.11)

where the sum of the integrals of the \( F_j \) is zero. Some authors, such as Kondo et al [17], like to include additional terms in the summation in an effort to improve the numerical accuracy, but others include different residuals to provide physical insight to the stabilization terms [22].

**Continuous Petrov forms**

Huebner and Thornton [13] present an example formulation where the \( F_j(x) \), and thus \( p(x) \), are picked to be continuous across the element boundaries and therefore they are able to apply integration by parts, over the entire domain, to the new terms and retain the use of \( C^0 \) interpolation. Others have used similar approaches, such as element bubble functions that vanish on the element boundary, but most applications involve functions that are discontinuous across the inter-element boundaries, as discussed next.

**Discontinuous Petrov forms**

In most advection applications the behavior in the streamline direction is usually more important than in the perpendicular "cross-wind" direction. It is possible to bias the Petrov weight by defining it to be the scalar result of the dot product of a unit vector in
the upwind streamline direction (obtained from the velocity vector \( \mathbf{v} \)) and some other assumed vector function, say \( \mathbf{G}(x) \) so:

\[
p(x) = n_x(x) \cdot \mathbf{G}(x) .
\]  

(14.12)

This common special case is known as the Streamline Upwind Petrov-Galerkin (SUPG) method \([2, 11, 12, 14]\). The vector function is usually taken as the gradient of the solution, \( \mathbf{G}(x) = L \nabla \phi(x) \), where \( L \) is constant, with length dimension, introduced to keep the units consistent with those of \( \phi \). Since the gradient is usually discontinuous between elements, we can not use integration by parts on \( p(x) \) over the solution domain. The goal of the Streamline Upwind Petrov-Galerkin is to stabilize the solution by adding information to include a bias on gradients in the flow direction. For a given velocity \( \mathbf{v} \) this is done by defining the SUPG weight function to be

\[
W_j(x) = H_j(x) + \alpha h \nabla H_j(x) \cdot \frac{\mathbf{v}(x)}{||\mathbf{v}||} \]  

(14.13)

where \( h \) is a measure of the element size, and from one-dimensional studies \( \alpha \) can be related to the local element Peclet number so as to obtain optimal accuracy \([2]\). In terms of the notation of Eq. 14.5, we would have \( r_j = \alpha h / ||\mathbf{v}|| \), and \( P_j = \mathbf{v} \cdot \nabla H_j \). Tezduyar and Osama \([25]\) have given mathematical norm definitions for establishing both element and nodal \( r \) values. Other definitions of \( r \) will be considered later.

Lagrangian interpolations have the property that at any point in space

\[
\sum_j H_j(x) = 1
\]  

(14.14)

Likewise, any gradient in the \( x_\gamma \) spatial direction of the above sum is the null vector

\[
\sum_j \frac{\partial H_j}{\partial x_\gamma}(x) = 0_\gamma
\]  

(14.15)

for spatial direction \( 1 \leq \gamma \leq 3 \). Note that the units have changed by the introduction of a length in the denominator due to taking a spatial derivative. For future reference consider a scalar zero term created by dotting Fig. 14.15 with a unit vector \( n_\gamma \):

\[
\sum_\gamma \left( \sum_j \frac{\partial N_j}{\partial x_\gamma} \right) n_\gamma(x) = 0
\]  

(14.16)

Typically, we want an upwind bias and use the velocity vector to define the unit vector as

\[
n_\gamma = \frac{v_\gamma}{||\mathbf{v}||}
\]  

(14.17)

and then the reference length \( h \) may be taken as some appropriate element distance.

Finally, we multiply the result by a constant \( 0 \leq \alpha \leq 1 \) to indicate the relative amount of ‘upwind’ emphasis. From this we see that if \( \mathbf{v} \) is constant

\[
\sum_j \alpha h \nabla H_j \frac{\mathbf{v}}{||\mathbf{v}||} = 0
\]  

(14.18)

so that from Eq. 14.13

\[
\sum_j W_j(x) = \sum_j H_j(x) = 1
\]  

(14.19)

as in the standard Galerkin form. Typically, \( \alpha \) is picked to give the optimal (nodally
exact) result for a known 1-D exact solution. That optimal value is usually defined in a collocation sense in that it exactly satisfies the PDE at a point in a uniform grid (for special choices of $Q$). The appearance of $h$ in the stabilization term, of Eq. 14.13, has lead several authors to propose ways to evaluate the relevant element length to be employed. We will review some of the methods in the next section and later relate them quantitatively to other length measures related to turbulence modeling.

### 14.3 Geometric measures of element and modal lengths for $\tau$

The stabilization parameters, $\tau$, often involve definitions that require a local length in the streamline direction related to the element size. Many researchers assume an average value over the element [2] based solely on the element size and shape, as suggested in Fig. 14.1. Some analysts allow for different lengths (and $\tau$ values) to be associated with each node of the element. However, purely geometric choices, such as those of Fig. 14.1, have been found to be less accurate and consistent that length scales developed on physical arguments, referred to here as $h_{UGN}$ and $h_{RGN}$ developed by Tezduyar for SUPG and discontinuity capturing problems [24, 26]. Tezduyar and Osawa developed a more general process for computing $\tau$ based on the ratio of local element norms [25]. The matrix norm based process of Tezduyar and Osawa was shown to also be efficient and accurate for higher order elements by Akin et al [1].

![Figure 14.1 Purely geometric element downwind distance options](image-url)
14.4 Review of SUPG concepts

There are numerous publications on the mathematics and application of the SUPG of Brooks and Hughes [2]. Several arguments have been given to describe why the stabilization method drastically improves the results of finite solutions of non-elliptical problems. Here we will review some of the concepts but the main point of this chapter is how we implement these methods when needed. We begin our review of some of the interpretations of how these stabilization methods work with the usual approach of the one-dimensional SUPG which has been proven to exactly satisfy the homogeneous form \((Q = 0)\) of Eq. 14.19 at all nodes in a uniform mesh for all Peclet numbers. Consider the one-dimensional model equation

\[
\nu \frac{\partial \phi}{\partial x} - k \frac{\partial^2 \phi}{\partial x^2} + Q = 0, \quad x \in ]0, L[ \tag{14.20}
\]

satisfying the two boundary conditions of

\[
\phi(0) = \phi_0, \quad \phi(L) = \phi_L \tag{14.21}
\]

where \(u\) is the given flow velocity, \(k\) is the diffusivity coefficient, and \(Q\) is the internal source per unit length. The solution for \(\phi(x)\) is governed by the global Peclet number \(Pe = vL/k\), and the grid Peclet number, \(pe = v h/(2k)\), where \(h\) is the element size. For \(Q = 0\) and \(v\) and \(k\) both constant, the exact solution for \(\phi_0 = 0\) and \(\phi_L = 1\) is given by \(\phi(x) = (1 - e^{-Pe x / L})/(1 - e^{-Pe})\). The classic Galerkin method solutions of this problem appear under-diffuse while most upwind methods appear over-diffuse.

Continuous Petrov Form

Only if \(F\) in Eq. 14.11 is continuous, then the element matrix forms can be written, after integration by parts as the conduction or diffusion parts:

\[
S^e_k = \int_{L_e} \frac{\partial W^T}{\partial x} k \frac{\partial H}{\partial x} dx \tag{14.22}
\]

\[
= \int_{L_e} \frac{\partial H^T}{\partial x} k \frac{\partial H}{\partial x} dx + \alpha \int_{L_e} \frac{\partial F^T}{\partial x} k \frac{\partial H}{\partial x} dx \tag{14.23}
\]

which matches the classical form only if \(\alpha = 0\) or if \(F\) is picked to force the last integral to vanish. [13] Otherwise there will be an additional new diffusion contribution and \(S^e_k\) will usually become unsymmetric. The source resultant is

\[
C^e_Q = \int_{L_e} W^T Q dx = \int_{L_e} H^T Q dx + \alpha \int_{L_e} F^T Q dx \tag{14.24}
\]

which for \(Q \neq 0\), modifies the nodal distribution of the source resultant. For any source distribution, \(Q(x)\), the sum of all the terms in the first integral (of \(H^T Q\)) accounts for the total source effects. This means that the sum of all the terms from the second integral, involving \(F\), must vanish. For a constant \(Q\) that means that the sum of the \(F\) terms must vanish. If we had known fluxes on the boundary they would be coupled to \(W\) (and thus to \(\alpha F\) if it is continuous) like the volumetric source matrices were.

The new moving, or advection, contribution is the matrix
\[
S_v^e = \int_{L^e} W^T_v \frac{\partial H}{\partial x} \, dx
\]  
(14.25)

which splits into

\[
S_v^e = \int_{L^e} H^T_v \frac{\partial H}{\partial x} \, dx + \alpha \int_{L^e} F^T_v \frac{\partial H}{\partial x} \, dx
\]  
(14.26)

which is the standard non-symmetric form plus a new array depending on \( F \) which in general would also be non-symmetric. It is much more common to employ Petrov forms that are discontinuous at the inter-element boundaries.

**Discontinuous Petrov forms**

If, and only if, we consider a special case where \( F \) is proportional to the gradient of the shape functions \( H \) (say \( F = c\partial H / \partial x \)) will the new Petrov-Galerkin advection contribution due to the velocity, \( v \), be a symmetrical matrix and be almost identical to the standard diffusion matrix;

\[
\alpha \int_{L^e} F^T_v \frac{\partial H}{\partial x} \, dx = \alpha \int_{L^e} \frac{\partial H^T}{\partial x} c v \frac{\partial H}{\partial x} \, dx.
\]  
(14.27)

Thus most people like to think of this common case as an element designed to increase the numerical diffusion in a controlled fashion.

When the SUPG is applied to this problem for linear finite elements, it gives nodally exact results by picking the optimal diffusion to add to the system. The SUPG is usually demonstrated with the finite difference pattern it produces when elements are assembled at a typical interior node of a uniform mesh. Here we will take the different approach and look directly at the element matrices that result if the residual vanishes on each element. For the case of \( Q = 0 \);

\[
\begin{bmatrix}
\frac{v}{2} [-1 & 1] + k \frac{h}{h} [-1 & 1] + \frac{\alpha v h}{2h} [-1 & 1] \\
-1 & 1
\end{bmatrix} \phi = 0
\]  
(14.28)

where

\[
\alpha = \text{Coth} (\text{Pe}) - 1/\text{Pe}
\]  
(14.29)

is the optimal upwind coefficient.

The first two matrices are the classical Galerkin advection and diffusion matrices, and the third square matrix is viewed as the added SUPG diffusion necessary for nodally exact solutions. While the third term is intended to emphasize the added diffusion, its units suggest that it could be added to the first matrix for the classical advection term. If we do that and apply the Petrov-Galerkin to a constant source term, \( Q \), then we see an alternate view of the element matrices (when node 2 is downwind) is

\[
\begin{bmatrix}
\frac{v}{2} [-1 + \alpha & 1 - \alpha] + k \frac{h}{h} [-1 & 1] + \frac{\alpha v h}{2h} [-1 & 1] \\
-1 & 1
\end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \frac{hQ}{2} \begin{bmatrix} 1 - \alpha \\ 1 + \alpha \end{bmatrix}
\]  
(14.30)

which serves as a clear reminder that the Petrov-Galerkin method also significantly influences the resultant-source vector. A system assembled from these element matrices gives the nodally exact result for any \( \alpha \). Two common special cases are easily observed;
\[\alpha \rightarrow 1 \text{ (Pe} \rightarrow \infty) \text{ for the maximum upwind correction}
\]
\[
\begin{pmatrix}
\frac{2v}{2} & 0 & 0 \\
\frac{k}{h} & 1 & -1 \\
\end{pmatrix}
\begin{pmatrix}
\phi_1 \\
\phi_2 \\
\end{pmatrix}
= 
\frac{2hQ}{2}
\begin{pmatrix}
0 \\
1 \\
\end{pmatrix}
\]
(14.31)

This lets us think of the upwinding effects as a strong weighing the upwind rows of the element source and the advection matrices rather than modifying the diffusion. This gives some insight into why, for large Peclet numbers, the upwind method of Rice and Schnipke [19] (and its degenerate form by Shemirani and Jambunathan [21]), which deletes the upwind rows of the convection matrix, works so well for low-order multi-dimensional elements for \( Q = 0 \) (and \( \text{Pe} \rightarrow \infty \)).

However, we are interested in a general process for higher order (e.g., p-adaptive) elements, so we will consider the change from Eq. 14.27 when a quadratic line element is employed with nodes 2 and 3 being downwind. We make the common assumption that \( \alpha \) is a single scalar term. Again we view the diffusion matrix as unchanged from the standard Galerkin form (with zero row and column sums), where node 2 is the interior node:

\[
S_k = \frac{k}{3h}
\begin{pmatrix}
7 & -8 & 1 \\
-8 & 16 & -8 \\
1 & -8 & 7 \\
\end{pmatrix}
\]
(14.32)

but the source vector due to \( Q \) has an additional term

\[
C_Q = \frac{Qh}{6}
\begin{pmatrix}
1 \\
4 \\
1 \\
\end{pmatrix}
\begin{pmatrix}
-6 \\
0 \\
6 \\
\end{pmatrix}
+ \frac{\alpha QH}{6}
\begin{pmatrix}
1 - 6\alpha \\
4 \\
1 + 6\alpha \\
\end{pmatrix}
\]
(14.33)

From this we see that the upwinding has a strong effect on the "corner" sources but no effect on the interior (and, thus, another downwind) node. Of course, the total source contributed \((Qh)\) is still accounted for at all values of \( \alpha \) (and \( \text{Pe} \)) since the sum of the coefficients multiplying it is unity.

The standard advection matrix for this quadratic element is the Galerkin form (with zero row sums)

\[
S_a = \frac{v}{6}
\begin{pmatrix}
-3 & 4 & -1 \\
-4 & 0 & 4 \\
1 & -4 & 3 \\
\end{pmatrix}
\]
(14.34)

and the SUPG correction is obtained by using \( k = \alpha v h / 2 \) in \( S_k \) (from Eq. 14.25). The combined upwind element matrices for the SUPG method are

\[
\frac{v}{6}
\begin{pmatrix}
(7\alpha - 3) & (4 - 8\alpha) & (\alpha - 1) \\
(-4 - 8\alpha) & 16\alpha & (4 - 8\alpha) \\
(1 + \alpha) & (-4 - 8\alpha) & (3 + 7\alpha) \\
\end{pmatrix}
+ \frac{k}{3h}
\begin{pmatrix}
7 & -8 & 1 \\
-8 & 16 & -8 \\
1 & -8 & 7 \\
\end{pmatrix}
\phi
= \frac{Qh}{6}
\begin{pmatrix}
1 - 6\alpha \\
4 \\
1 + 6\alpha \\
\end{pmatrix}
\]
(14.35)

which is the classic Galerkin form when \( \alpha = 0 \). From Eq. 14.35 we see that for higher order elements, the upwinding effects on advection and source terms are not as simple as the effect of \( \alpha \) in Eq. 14.23 may have implied. For maximum upwinding \((\alpha = 1)\) this
The above single element based $\alpha$ term for the quadratic element no longer gives exact results at the nodes, even for $Q = 0$. To accomplish that Codina, et al. [5] have shown that a nodal based approached may be needed with one upwind constant, $\alpha$, for the two corner nodes and a second, $\beta$, for the interior node. They show the constants to be

$$\beta = (\coth(\text{Pe}/2) - 2/\text{Pe})/2,$$

and

$$\alpha = \frac{(3 + 3\text{ Pe} \beta) \tanh(\text{Pe}) - (3\text{Pe} + \text{Pe}^2 \beta)}{(2 - 3\beta \tanh(\text{Pe}))\text{Pe}^2}$$

which gives a $\beta$ that is about half of $\alpha$ for most Pe, and $\beta \to \alpha/2$ for $\text{Pe} \to \infty$. Their form from an element matrix viewpoint is

$$\begin{pmatrix} v/6 & (7\alpha - 3) & (4 - 8\alpha) & (\alpha - 1) \\ -4 - 8\beta & 16\beta & (4 - 8\alpha) \\ 1 + \alpha & (-4 - 8\alpha) & (3 + 7\alpha) \end{pmatrix} \phi = C_0 + S_k \begin{pmatrix} -5 \\ 4 \\ 7 \end{pmatrix}$$

which, compared to Eq. 14.35, reduces the upwind effect on the interior node. If we use the gross approximation that $\beta = \alpha/2$ (which is exact for $\text{Pe} \to \infty$), we get

$$S_a \approx \begin{pmatrix} (7\alpha - 3) & (4 - 8\alpha) & (\alpha - 1) \\ 4(-1 - \alpha) & 8\alpha & 4(1 - \alpha) \\ (1 + \alpha) & (-4 - 8\alpha) & (3 + 7\alpha) \end{pmatrix}$$

which has zero row sums, but not zero column sums. In the limit of $\text{Pe} \to \infty$, this gives

$$S_a = \begin{pmatrix} 4 & -4 & 0 \\ -8 & 8 & 0 \\ 2 & -12 & 10 \end{pmatrix}$$

which again significantly differs from the Galerkin form of Eq. 14.27 and the single element based $\alpha$ form given in Eq. 14.29. This short review of SUPG concepts suggests that the extension to higher order elements may, in general, benefit from different upwind coefficients for corner nodes, edge nodes, and internal nodes.

### 14.5 One-dimensional example

The application of the SUPG for linear line elements is, as expected, the most common way to illustrate the process. There are aspects of the computation that are obtained by inspection that require little extra programming in general. It is clear that the reference length to be used in calculating the Peclet number is simply the length of the element. Likewise, the gradient of the solution, and $\mathbf{H}$, in the $x$-direction is also the gradient in the direction tangent to the streamline. Finally, since the second derivatives of such approximations are zero (unless iterative results are used) one avoids having to bring
into the analysis the information on second derivatives that occur in the SUPG theory.

For very high Peclet numbers the governing equation needs to be modeled analytically with what is known as ‘singular perturbation theory’. Usually such problems involve a parameter (here 1/Pe) associated with the highest order derivative in the differential equation. As that parameter approaches zero one has essentially a lower order differential equation with more boundary conditions than the reduced equation requires. In other words a thin "boundary layer" develops in a part of the solution domain near the redundant boundary condition and the solution must change very rapidly in that small region as it tries to satisfy the original higher order derivative terms.

When such a problem is approximated by a classic Galerkin finite element solution a least squares spatial response develops to try to capture the very sharp gradients in thin boundary layer. While it may do that, it over shoots the spatial solution in the region adjacent to the boundary layer and gives huge errors, or physically impossible results, as it oscillates about the true solution in the main domain that is reasonably modeled by the lower order differential equation. This type of behavior should not come as a surprise since we have changed to a new class of differential equation that is unlike the elliptical one used in most of our examples. Here the system is parabolic in nature and a polynomial approximation may not be the best choice for our finite element model. Since the response is basically exponential in nature near the boundary layer we should consider using exponential interpolation functions or adding new terms to our solutions to accurately dampen out the incorrect responses. The SUPG approach does the latter. The non-polynomial interpolations would also work, but tend to be expensive to compute and sensitive to the word length of the computer employed.

We begin with a classic Galerkin solution in one-dimension with no source term and where the diffusion term, $k$, is decreased to make $Pe = 100$. A ten element model based on linear interpolation is shown in Fig. 14.2 along with the exact solution (dashed line). It is easily seen that the effect of the sharp boundary layer has propagated back into the full domain and gives a very inaccurate result. By way of comparison, the addition of the SUPG terms give the result in Fig. 14.3. The improvement is drastic, with the SUPG giving numerical results that are essentially exact at the nodes. If one reviews the error estimate for the Galerkin result (see Fig. 14.4) it is tempting to try smaller elements near the boundary layer. However, that does not give much improvement (as seen in Fig. 14.5) and a much finer mesh would be required to attempt to get reasonable accuracy and thus the SUPG modifications are much more cost effective.

The one-dimensional source code, in Fig. 14.6, illustrates typical considerations of SUPG methods. This version is designed to let the student switch from the standard Galerkin to SUPG by supplying the keyword $supg$ in the input file. In the one-dimensional case we know that the fluid velocity is constant and acts over the full length of the element. Thus, one can select to compute the upwind parameters outside the element numerical integration loop. They are illustrated in lines 46 to 51. Other changes that relate to the SUPG selection occur between lines 70 to 91. The consistent SUPG method introduces second derivatives of the interpolation functions. They may or may not be zero. Lines 72 through 78 address their inclusion, as do lines 84 and 85, even though most programmers choose to omit them. If we neglect the second derivative question we always have to append new matrices to the element source vector and square matrix.
Figure 14.2 Galerkin solution of $u \phi_x - k \phi_{xx} = 0$ for $Pe = 100$

Figure 14.3 SUPG solution of $u \phi_x - k \phi_{xx} = 0$ for $Pe = 100$
Figure 14.4 Galerkin energy error norm estimate

Figure 14.5 Revised mesh Galerkin solution for Pe = 100
The main difference in the element matrices, lines 89 through 91, compared to previous examples is that we have both the interpolations for $W$ and $H$ appearing in the matrix products instead of just $H$. Note that the physical derivatives of $H$ in the $x$-direction, $DGH(1,:)$ is actually a derivative taken tangent to the fluid streamline so the term $v \ast DGH(1,:)$ in line 91 is related to the speed of the flow times the gradient of the unknown along the streamline. Also remember that line 91 makes the square matrix nonsymmetric. The data file for this example is in Fig. 14.7. The exact solution to be compared with (case 18) is identified in line 3 while its use is invoked in lines 19 and 20. It requires the global Peclet number and that is supplied as miscellaneous data as the last line (52) and is not used anywhere in the application source code of Fig. 14.6.

14.6 Generalizing to higher dimensions

Here we will outline the generalization of the previous process to a single implementation that can handle 1-D, 2-D, 3-D, or axisymmetric domains for any element in the MODEL library. Generalizing the SUPG method requires much more data to describe the velocity field and required items along the streamline directions. The current version allows the choice of four different definitions of $\tau$, but the element norm method is now the default. Thus the example program is quite a bit longer but is easily broken into four conceptual tasks.

For higher dimensional problems we can form a single element based upwind length measure by evaluating the velocity at the element center to form nodal distances, such as in Figs. 14.1,b and 1,b, and then convert them to a single length as illustrated in Fig. 14.1,c. Denote that length as $h_{geom}$. Along the streamline we could use the 1-D optimal length scaling, $\alpha_{opt}$, based on the local element Peclet number (at the center) to compute a corresponding stabilization term:

$$
\tau_{geom} = \alpha_{opt} h_{geom} / 2\|v\|.
$$

Tezduyar and Park [23] defined an alternate geometric length that usually gives better results. It is known as $h_{ugn}$ since it comes from the dot product of $u$ and the gradient of the generalized element interpolation functions, $N$. For our scalar variable examples it is:

$$
h_{ugn} = 2\|v\| \left( \sum_j |v \cdot \nabla H_j | \right)^{-1}.
\tag{14.41}
$$

They define the corresponding advection dominated flow stabilization parameter to be

$$
\tau_{ugn} = h_{ugn} / 2\|v\| \tag{14.42}
$$

and a similar form for stabilizing the least squares incompressibility constraint in Navier-Stokes flows.

More recently Tezduyar and Osama [25] suggested a $\tau$ parameter based on scaling the Galerkin and stabilization matrices to be of the same order in each element. Thus they use the ratios of two matrix norms to establish the stabilization parameters for advection, diffusion, and transient dominated regions. Since the resulting $\tau$ involves the ratio of the norms of two matrices it has been found to be relatively insensitive to the method chosen to evaluate a matrix norm. Here we will use the norm to be the square
Supplement 2: Stablized methods and SUPG

---

! ******** ELEM_SQ_MATRIX PROBLEM DEPENDENT STATEMENTS FOLLOW ****
!! Define any new array or variable types, then statements
!! (Library example number 112)
!! Galerkin or SUPG 1-d Advection-Diffusion Problem
!! u * dp/dx - d(k * dp/dx)/dx = Q, assume u, k, Q constant
!!
!! u = GET_REAL_LP (1) ! velocity
!! k = GET_REAL_LP (2) ! conductivity
!! Q = GET_REAL_LP (3) ! source per unit length
!! SUPG = global logical flag ! F=Galerkin (default), T=SUPG
!!
!! LT_N = number of nodes for this element type
!! MISC_FX = number of integer miscellaneous properties
!! N_LP_FLO = number of real element properties
!! W =Petrov weight, DGW its global derivative
!
REAL(DP) :: W (LT_N), DGW (1, LT_N) ! SUPG & deriv
REAL(DP) :: D2GH (1, LT_N) ! SUPG, zero ?
REAL(DP) :: DL, DX_DR, DL_A ! Length, Jacobian
REAL(DP) :: u, k, Q ! input data
REAL(DP) :: Pe, ALPHA, COTH ! Peclet data, L2
INTEGER :: IQ ! Loops
REAL(DP), SAVE :: Pe_max ! debugging
!
DL = COORD (LT_N, 1) - COORD (1, 1) ! Element length
DX_DR = DL / 2. ! constant Jacobian
DL_A = DL / (LT_N - 1) ! SUPG length
!
DATA READS AND SAVES
u = GET_REAL_LP (1) ! velocity
k = GET_REAL_LP (2) ; E = k ! conductivity
Q = 0.d0 ; IF ( N_LP_FLO > 2 ) Q = GET_REAL_LP (3) ! source
E = k ! constitutive
!
IF ( IE == 1 ) THEN ! FIRST ELEMENT, ONE TIME ACTIONS
Pe_max = 0.d0 ! initialize
!
IF ( .NOT. SUPG ) THEN ! echo choice
PRINT *, 'NOTE: Galerkin method' ! default
ELSE ; PRINT *, 'NOTE: SUPG method' ; END IF ! supg
END IF ! FIRST ELEMENT
!
SUPG TERMS (ASSUMING L2 ELEMENT), ? Bias for L3, L4?
Pe = ABS(u) * DL / k ! Grid Peclet
IF ( Pe > Pe_max ) Pe_max = Pe ! for debug
COTH = COSH (Pe/2) / SINH (Pe/2) ! Optimal SUPG
ALPHA = ABS (COTH) - 1.d0 / ABS (Pe/2) ! Optimal SUPG L2
ALPHA = SIGN (ALPHA, u) ! abs(ALPHA)*sign of u

Figure 14.6a One-dimensional SUPG interface
Figure 14.6b One-dimensional SUPG element matrices
Supplement 2: Stabilized methods and SUPG

525

Figure 14.7 Data for 1-D SUPG test
root of the sum of the squares of all the terms in the matrix.

Note from Eq. 14.8 that the Galerkin contribution will produce three square matrices. They come from transient, advection, and diffusion dominated terms. A discontinuous Petrov stabilization term would also contribute similar terms but usually linear elements are employed so the second derivative contribution is zero in that case. The setting of $\tau$ by a matrix norm method is an attempt to assure that the Galerkin and Petrov terms are of the same order of magnitude, relative to the effect that is dominating the flow. For advective dominated flow the definition of the $\tau_{\text{norm}}$ given by the ratio norms of the two matrices arising from the $v \cdot \nabla \phi$ terms:

$$\tau_{\text{norm}} = \frac{\| \int_{L^e} H^T v \cdot \nabla H \, dx \|}{\| \int_{L^e} v \cdot \nabla H^T v \cdot \nabla H \, dx \|}. \quad (14.43)$$

14.7 Two-dimensional examples

The extension of SUPG methods two higher dimensions is relatively clear but there are choices to be made on the most cost effective way to get the effective Peclet number. This may involve elements where the velocity is clearly constant in an element, or it may have significant changes over an element (as in a high degree p-method formulation). Some logical options will be considered after some typical numerical results are presented. We will begin using only the classic linear triangular element (T3) and consider higher degree elements in later examples. A common test case is where a fluid enters the lower left edge of a rectangle and exits at the lower right edge as shown in Fig. 14.10. The boundary condition on the unknown is that it varies rapidly from zero to two along the inflow boundary and is zero at the impervious sides. Along the inflow edge (of the negative x-axis) the given value is $T = 1 + \tanh((2 * x + 1) * 10)$ for $x = [-1, 0]$. The outflow and interior values are to be determined. For an infinite Peclet number (no diffusion) the outflow values should be the mirror images of the input curve. The velocity components are $u = 2y(1 - x^2)$ and $v = -2x(1 - y^2)$ which means that for relatively large elements both the magnitude and direction of the velocity may change significantly within the element. This velocity field is maximum at the origin, zero on three sides, and is clockwise about the origin.

An initially uniform mesh of linear, T3, triangular elements is selected as shown in Fig. 14.10 along with essential boundary condition flags at the nodes, and a typical low velocity Galerkin solution. That mesh was designed so that the same nodes can be used to form a similar mesh made with quadratic, T6, triangles, or the corresponding Q4 or Q9 quadrilateral elements. When the local Peclet numbers are low a reasonable Galerkin solution can be obtained without stabilization, as illustrated in Fig. 14.10. But if one increases the maximum Peclet number then an unstable Galerkin solution results. Retaining the same data but stabilizing the solution (simply by adding control keyword $supg$ to the input) renders a drastically improved solution whose back view is seen in Figs. 14.12.
Supplement 2: Stabilized methods and SUPG

527

Figure 14.8a Storage for general advection-diffusion
FINITE ELEMENT ANALYSIS WITH ERROR ESTIMATORS

DEFINE ELEMENT PROPERTIES
RATE = 0 ; SOURCE = 0 ; THICK = 1 ; VISCOSITY = 1 ! initialize
IF ( REALS > 0 ) VISCOSITY = GET_REAL_MISC (1) ! diffusity
IF ( REALS > 1 ) SOURCE = GET_REAL_MISC (2) ! constant Q
IF ( REALS > 2 ) RATE = GET_REAL_MISC (3) ! constant r
IF ( REALS > 3 ) THICK = GET_REAL_MISC (4) ! thickness
CENTER = SUM (COORD, DIM=1) / LT_N ! center point
CALL APPLICATION_E_MATRIX (IE, CENTER, E) ! constitutive law
IF ( SUPG ) THEN ! Streamline Upwind Petrov-Galerkin additions
INITIALIZE STABILIZATION ARRAYS
S_UP = 0.d0 ; C_UP = 0.d0 ; H_INTG = 0.d0
S_K_BAR = 0.d0 ; S_R_BAR = 0.d0
GET CENTER VELOCITY AND DIFFUSION
CALL VELOCITY_AT_POINT (CENTER, U, UNIT_V, SPEED)
GET DIFFUSION ALONG STREAMLINE
CALL DIFFUSION_UPWIND (E, UNIT_V, E_UP, E_CROSS)
IF ( TAU_GEOM .OR. TAU_VOL ) THEN
CALL GET_RADIALS_FROM_CENTER (CENTER, RADIAL)
CALL GET_DOWNWIND_LOGIC (RADIAL, UNIT_V, DOWN, IS_DOWNWIND)
CALL GET_MAX_DOWNWIND_DIST (DOWN, DOWN_NODAL)
END IF
IF ( TAU_GEOM ) THEN
GEOM_H = ABS (MINVAL (DOWN)) + MAXVAL (DOWN)
PECLET = 0.5d0 * SPEED * GEOM_H / E_UP
CALL PECLET_OPTIMAL_RULE (PECLET, ALPHA)
GEOM_TAU = 0.5d0 * GEOM_H * ALPHA / SPEED
END IF ! Tau_geom
IF ( TAU_UGN ) THEN
CALL GET_ONE_PT_RULE (ONE_PT, ONE_WT)
CALL SCALAR_DERIVS (ONE_PT, DLH)
CALL INVERT_JACOBIAN (AJ, AJ_INV, DET, N_SPACE)
DGH = MATMUL (AJ_INV, DLH)
U_DGH = MATMUL (U, DGH)
UGN_TAU = 1.d0 / SUM (ABS (U_DGH))
END IF ! Tau_ugn
END IF ! Initialize upwinding
STORE NUMBER OF POINTS FOR FLUX CALCULATIONS
CALL STORE_FLUX_POINT_COUNT ! Save LT_QP

Figure 14.8b Computations at element center
Supplement 2: Stabilized methods and SUPG

Figure 14.8c Numerical integration of advection-diffusion items
Employing the element error estimators suggests that the current element sizes should be scaled as shown in Fig. 14.11. Supplying those guidelines for element densities to an automatic mesh generator yield a second mesh, in the same figure, and whose new solution surface is displayed in Fig. 14.12. This is a typical example of how stabilized solutions are essential for non-elliptical differential equations.

Figure 14.8d Final selection of $\tau$ and SUPG stabilizations

Figure 14.9 New control options for stabilized solutions
\[ u = y, \quad v = -x \]
\[ T (x, 0) = [1 + \tanh (2x + 1)] \]

**Figure 14.10** Initial T3 mesh, and low Pe Galerkin solution
Figure 14.11 Estimated sizes and revised SUPG T3 mesh
**Figure 14.12** Revised SUPG T3 solution with $\tau_{\text{norm}}$

**Figure 14.13** The Cosine Hill initial Q4 mesh
Figure 14.14 Typical Galerkin Q4 solution and mesh refinement suggestion
Figure 14.15 The Q4 result and new mesh estimate from $\tau_{\text{agn}}$
As a second stabilization example consider a common test problem called the Cosine Hill. It assumes almost pure advection \((k = 1e-8)\) using a counterclockwise circular velocity field centered on a square, with \(-1/2 \leq x, y \leq 1/2\). The fluid speed proportional to the radial distance from the center of the square. As shown in Fig. 14.13, the initial mesh has \(\phi = 0\) on the boundary of the square. On the interior line at \(x = 0\) and \(y \leq 0\) it varies as \(\phi = \sin(\pi(1 - 2y))\). In this case the source term, \(Q\), is zero. For pure advection, the solution surface should circle back on itself with no change. That is, in the limit we should see a zero value of \(\partial\phi/\partial x\) as the solution approaches \(x = 0\) from \(x < 0\). Typical Galerkin solutions would have very large oscillations as they approach \(x = 0\).

The initial Q4 element mesh, in Fig. 14.13, has been chosen so that it is uniform and so that without changing the nodal count or locations one can employ a mesh for Q4, Q9, Q16, T3, T6, or T10 elements. This allows one to compare linear, quadratic, and cubic elements. In addition, all of these meshes yield nodal results that can be projected to a common Q4 mesh for visual comparisons. This mesh can be refine uniformly to retain this feature, or if one uses an error estimator new non-uniform meshes can be developed. The results depend on the element degree, the choice of \(\tau\), and the relative element sizes and locations. Thus, it is well suited for parametric studies of those variables. Many such studies have been carried out but space here limits us to a few examples.

A typical Galerkin solution result is shown in Fig. 14.14 along with the estimated required mesh refinement. One could continually refine the mesh in this fashion and possibly obtain a useful Galerkin solution, but it is more more cost effective to employ a stabilization method.

Figure 14.15 shows an initial solution using Q4 elements and the \(\tau_{\text{uga}}\) choice. Note that the vertical axis of these plots gives the minimum and maximum function value found anywhere in the mesh. That allows some extra comparisons of the various following surface plots which at first glance look very similar when projected onto a common reference surface. This solution yielded error estimates that suggested a new non-uniform mesh with local element sizes also shown there. Note that the new sizes are not symmetrically spaced. The false high degree of refinement at the center is due to the large gradient discontinuity imposed boundary conditions. The corresponding figures for \(\tau_{\text{norm}}\), using the same Q4 elements, are given in Fig. 14.16 and are almost identical to the results of Fig. 14.15.

We will illustrate the other linear through cubic elements, that can use the same nodes, by applying the \(\tau_{\text{norm}}\) to the initial mesh only. We begin with the Q9 and Q16 elements. The Q9 solution results are shown in Fig. 14.17 and those for the Q16 elements are in Fig. 14.18. These are cube meshes even though the polynomial degree is higher. The corresponding first suggested mesh revisions are given in Fig. 14.19, but one probably needs even smaller sizes (which can be done via control keywords). The corresponding linear through cubic T3, T6, and T10 triangular family element results and suggested mesh refinements are given in Fig. 14.20 to 22, respectively.

The higher order element solutions shown here are based on the common practice of omitting the second order derivatives of the interpolation functions. All choices for \(\tau\) (tested here) seem to give similar results for a problem with no source term. One can apply the same grid spacing for meshes with Serendipity quadratic, Q8, and cubic, Q12,
Figure 14.16 The Q4 result and new mesh estimate from $\tau_{\text{norm}}$
Figure 14.17 Quadratic Q9 results for $\tau_{\text{norm}}$, plotted on Q4 mesh

Figure 14.18 Cubic Q16 results for $\tau_{\text{norm}}$, plotted on Q4 mesh
Figure 14.19 Initial suggested refinements for Q9 (top), and Q16 elements with $r_{\text{norm}}$.
Figure 14.20 Linear T3 triangle initial results and sizes for $\tau_{\text{norm}}$. 

Finite Element Analysis with Error Estimators
Figure 14.21 Quadratic T6 triangle results for $\tau_{norm}$, plotted on Q4 mesh

Figure 14.22 Cubic T10 triangle results for $\tau_{norm}$, plotted on Q4 mesh
Finite Element Analysis with Error Estimators

Figure 14.23 Reference Q4 solution for advection of a constant source
quadrilaterals. However, fewer nodes and equations are used since they do not have interior nodes. Considering that effect they seem to perform equally well in this test without a source term.

Since stabilization methods should also be tested on problems with source terms the final example will consider a problem with a constant source term distributed over a square, with the function having essential boundary values of zero on all four edges, and having a constant diagonal flow with a unit velocity. For very high advection rates the solution is pushed into the corner at \( x = y = 1/2 \) and the rapidly drops to zero through a sharp boundary layer along the lines \( x = 1/2 \) and \( y = 1/2 \). Again a uniform initial mesh is selected so that the same set of nodes can be employed in meshes that use different element types taken from the linear through cubic degree elements in list of T3 or T6 or T10 or Q4 or Q9 or Q16 elements. All of these give results that can be projected onto a common Q4 mesh to simplify visual comparisons. The initial mesh was chosen to be relatively crude and consisted of a \( 18 \times 18 \) grid of square Q4 linear elements. A repeatedly refined mesh of Q4 elements was employed to obtain a fine scale reference solution to which other results will be compared. Its maximum solution value is 5.14. In the initial uniform mesh \( \tau \) was a single constant over the full domain. Now since the adaptive element sizes are changing each element has a different \( \tau \) value. One might expect that the stabilization terms will be highest in the boundary layer, but that is not the case. When quadratic and cubic elements are utilized the solution improves but the distribution of \( \tau \) remains about the same in shape. Their peak solution values are closer to the reference solution. The same is not true for the Serendipity Q8 and Q12 elements which yield strange results near the two far corners off the main diagonal of flow. Using mesh refinements guided by the error estimator all solutions approach the same value.

### 14.8 Exercises

1. A model equation with a non-uniform source and a boundary layer near \( x = 1 \) is 
\[-u'' + ku' = Q\] when \( k >> 1 \) and \( u(0) = 0 = u(1) \). Obtain a finite element solution when the source per unit length is defined as:
   
a) \( Q = 3k x^2 \) so that \( u(x) = x^3 + A x^2 + B x + C(e^{kx} - 1)/(e^k - 1) \) where \( A = 3/k, B = 6/k^2, \) and \( C = -(1 + A + B) \).

   b) \( Q = 3k x^2 + 2k \pi \cos(2\pi x) + 4\pi^2 \sin(2\pi x) \) so \( u(x) \) is the above expression plus \( \sin(2\pi x) \).

   Note that in both cases the total applied source is \( k \). Use \( k = 1 \) and \( k = 60 \) to see solutions without and with a boundary layer, respectively. (For \( k = 60 \) these two exact solutions correspond to exact_case 28 and 29 in MODEL, respectively.)

2. A one-dimensional problem with boundary layer at \( x = 1 \) is 
\[-ku''(x) + u'(x) = 1, \quad 0 \leq x \leq 1\]
with the boundary conditions of \( u(0) = 0 = u(1) \). Obtain finite element solutions for \( k = 0.1 \) and \( k = 0.01 \) and compare them to the exact result of
\[u(x) = x - \exp[(x - 1)/k] - \exp[-1/k]/1 - \exp[-1/k].\]
14.9 Bibliography


