

A PETROV-GALERKIN FINITE ELEMENT FORMULATION FOR
SYSTEMS OF CONSERVATION LAWS WITH
SPECIAL REFERENCE TO THE COMPRESSIBLE EULER EQUATIONS

T.J.R. Hughes, T.E. Tezduyar and A.N. Brooks

(Division of Applied Mechanics, Stanford University, USA)

1. INTRODUCTION

In this paper we are concerned with certain classes of Petrov-Galerkin finite element methods for convection dominated flows. Petrov-Galerkin methods are distinguished from the more widely known (Bubnov-) Galerkin method in that weighting functions and trial solutions are selected from different classes of functions. This computational framework opens the way to the development of numerical methods which are optimal, or nearly so, according to various measures of accuracy.

This may be contrasted with the behaviour of Galerkin methods for this class of problems, which suffer from both accuracy and stability deficiencies. Dupont (1973) was the first to point out that the Galerkin method did not generally achieve optimal asymptotic error estimates for even smooth solutions of the simplest one-dimensional, first-order hyperbolic model problem. When solutions are rough, Galerkin methods are known to produce highly oscillatory error components. Thus the "best approximation property", and its stability and optimal-accuracy by-products, which the Galerkin method achieves in typical elliptic operator problems, is lost in the context of highly nonsymmetrical convection-type operators.

Due to the potential of Petrov-Galerkin methods for flow problems, a large literature is already accumulating. The first reference that we are aware of on this topic was the paper by Dendy (1974). Works which analyzed, further developed, and generalized Dendy's approach were those of Wahlbin (1974), Raymond and Garder (1976) and Baker (1979). Dendy's formulation applied to first-order hyperbolic problems. The introduction of second-order terms into the equations posed technical problems unless interpolations with C^1 -continuity were introduced. These are, of course, very inconvenient in practice and this, perhaps, was the main reason why this approach did not have a more significant impact initially.

A separate school of thought was later originated by the Dundee and Swansea research groups (see, e.g., Christie et al., 1976, Heinrich et al., 1977, Griffiths and Mitchell, 1979, and Heinrich and Zienkiewicz, 1979). In these works various perturbations of the normal Galerkin weighting functions were introduced ostensibly to account for upwind influence.* Because these perturbations could be continuous, there were no impediments to applying them to equations including second-order spatial derivatives (e.g. advection-diffusion equations). Morton and colleagues (see, e.g., Barrett and Morton, 1980, 1981, Morton, 1981, Morton and Barrett, 1980, Morton and Parrott, 1980) extended the development of these ideas to time-dependent cases and defined methods achieving optimality conditions in various norms.

Our work in this area began with the development of finite element methods for the advection-diffusion and incompressible Navier-Stokes equations which we called the "streamline-upwind method" (Hughes and Brooks, 1979, Brooks and Hughes, 1980). The method was similar in intent to the skew-differencing procedure of Raithby (1976). In attempting to generalize the procedure to a Petrov-Galerkin method we encountered the same difficulty that faced Dendy's method. This was finally overcome in (Hughes and Brooks, 1982) for the advection-diffusion equation and in (Brooks and Hughes, 1982a,b) for the incompressible Navier-Stokes equations.

We have solved a wide variety of steady and time-dependent linear and nonlinear problems to demonstrate the accuracy and stability of this formulation.

Claes Johnson (1981) has analyzed the streamline method and established optimal convergence rates and a strong discontinuity capturing property, even when the discontinuity is skew to the mesh.

In this paper we present a general Petrov-Galerkin formulation of streamline type which accommodates systems of conservation laws. The formulation can specialize to advection-diffusion, compressible Euler, compressible Navier-Stokes equations,

*Frequently, Petrov-Galerkin finite element methods are described as "upwind element methods". Fully upwind finite difference methods are so inaccurate that the terminology "upwind" has taken on a perjorative connotation. It is to be emphasized in no uncertain terms that "upwind Petrov-Galerkin finite element methods" are very different from such upwind difference methods and do not suffer the pathological, excess numerical diffusion of the latter.

etc. In this work we emphasize the advection-diffusion and, in particular, the compressible Euler equations. The former are simple and delineate the methodology, and the latter are of considerable contemporary interest in computational aerodynamics. An outline of the remainder of the paper follows.

After describing notational preliminaries in Section 2 we introduce general systems of conservation laws in Section 3. In Section 4 we propose a weighted residual formulation of Petrov-Galerkin type. Particular attention is paid to formulating a variety of physically important boundary conditions. In our opinion this is a subject which is often not adequately discussed in computational fluid dynamics. Some possible simplifications of the basic formulation are described. In Section 5 we present further details concerning the selection of optimization parameters and describe a family of new methods prompted by Jean Donea's recently developed Taylor-Galerkin schemes (Donea, 1982). Some sample calculations are presented in Section 6 and conclusions are drawn in Section 7.

To keep the paper to a reasonable size we have not discussed solution algorithms herein. The interested reader may consult our other works which contain detailed treatments of this subject (see, e.g., Brooks and Hughes, 1982a,b and Tezduyar and Hughes, 1982).

2. PRELIMINARIES

Let n_{sd} denote the number of space dimensions. Let Ω be an open region in $\mathbb{R}^{n_{sd}}$ with piecewise smooth boundary Γ . Let $\underline{x} = \{x_i\}$, $i = 1, 2, \dots, n_{sd}$, denote a general point in Ω and let $\underline{n} = \{n_i\}$ be the unit outward normal vector to Γ . We assume Γ admits the following decomposition

$$\Gamma = \overline{\Gamma_G \cup \Gamma_H} \quad (2.1)$$

$$\emptyset = \Gamma_G \cap \Gamma_H \quad (2.2)$$

where Γ_G and Γ_H are subsets of Γ . The superposed bar in (2.1) represents set closure and $\emptyset$, in (2.2), denotes the empty set. The significance of Γ_G and Γ_H will be made apparent in the sequel.

The summation convention on repeated indices is assumed in force and the subscript n is used to denote the normal component of a vector (e.g., if $n_{sd} = 3$, then $\underline{F}_n = \underline{F}_j n_j = \underline{F}_1 n_1 + \underline{F}_2 n_2 + \underline{F}_3 n_3$). A comma is used to denote partial differentiation

(e.g., $F_{j,j} = \partial F_j / \partial x_j$) and t denotes time. The Kronecker delta is denoted by δ_{ij} ; if $i = j$, then $\delta_{ij} = 1$, otherwise $\delta_{ij} = 0$.

Consider a discretization of Ω into element subdomains Ω^e , $e = 1, 2, \dots, n_e$, where n_e is the number of elements. Each Ω^e is taken to be an open set and its boundary is denoted Γ^e . We assume

$$\bar{\Omega} = \bigcup_e \bar{\Omega}^e \quad (2.3)$$

$$\Gamma \subset \bigcup_e \Gamma^e \quad (2.4)$$

The set $\bigcup_e \Omega^e$ will be referred to as the element interiors. The element boundaries, modulo Γ , play an important role in what follows. We call this set the interior boundary, viz.

$$\Gamma_{\text{int}} = \bigcup_e \Gamma^e - \Gamma. \quad (2.5)$$

Two classes of functions are important in the developments which follow. The classes are distinguished by their continuity properties across Γ_{int} . It suffices to assume that all functions considered herein are smooth on the element interior. Functions of the first class are continuous across Γ_{int} . These functions are denoted by C^0 . Functions of the second class are allowed to be discontinuous across Γ_{int} and are denoted by C^{-1} . The C^0 functions may be recognized as containing the standard finite element interpolations.

3. SYSTEMS OF CONSERVATION LAWS

We consider the following system of m partial differential equations:

$$U_t + F_{j,j} + G = 0 \text{ on } \Omega, \quad (3.1)$$

where

$$U = U(x, t) \quad (3.2)$$

$$F_j = F_j(U, \nabla U, x, t) \quad (3.3)$$

$$G = G(U, x, t). \quad (3.4)$$

The vector F_j is referred to as a flux vector and G is a source term. We assume that for each $k = \{k_1\} \in \mathbb{R}^{n_{sd}}$ there exists a

matrix $\tilde{S}$ such that

$$\tilde{S}^{-1}(\tilde{k}_j \tilde{A}_j) \tilde{S} = \tilde{\Lambda}, \quad (3.5)$$

where

$$\tilde{A}_j = \tilde{D}_1 \tilde{F}_j = \partial \tilde{F}_j / \partial \tilde{U} \quad (3.6)$$

and $\tilde{\Lambda}$ is a real, diagonal matrix. If, in addition

$$\tilde{D}_2 \tilde{F}_j = \partial \tilde{F}_j / \partial (\nabla \tilde{U}) = 0, \quad (3.7)$$

then (3.1) is said to be a first-order hyperbolic system.

Let the total flux be decomposed into partial fluxes, $\tilde{F}_j^{(1)}$ and $\tilde{F}_j^{(2)}$, as follows:

$$\tilde{F}_j(\tilde{U}, \nabla \tilde{U}, \tilde{x}, t) = \tilde{F}_j^{(1)}(\tilde{U}, \tilde{x}, t) + \tilde{F}_j^{(2)}(\tilde{U}, \nabla \tilde{U}, \tilde{x}, t). \quad (3.8)$$

Thus, if dissipative mechanisms are present, as evidenced by the appearance of the argument $\nabla \tilde{U}$, we assume they are confined to the partial flux $\tilde{F}_j^{(2)}$.

The initial/boundary-value problem for (3.1) consists of finding a function $\tilde{U}$ which satisfies (3.1), the initial condition

$$\tilde{U}(\tilde{x}, 0) = \tilde{U}_0(\tilde{x}), \quad (3.9)$$

where $\tilde{U}_0$ is a given function of $\tilde{x} \in \tilde{\Omega}$, and appropriately specified boundary conditions. In this paper, for sake of simplicity, we shall limit attention to boundary conditions of the following form:

Dirichlet type

In this case we assume

$$\partial \tilde{U} = \tilde{G} \text{ on } \Gamma_{\tilde{G}} \quad (3.10)$$

where ∂ is a boundary operator and $\tilde{G}$ is a prescribed function.

Neumann type

In this case we assume

$$-\tilde{F}_n^{(2)} = \tilde{H} \text{ on } \Gamma_{\tilde{H}} \quad (3.11)$$

where $\tilde{H}$ is a prescribed function on $\Gamma_{\tilde{H}}$. This can be interpreted as a partial flux boundary condition or, if $F_{\tilde{j}}^{(1)} = 0$, a total flux boundary condition.

Finally, we allow for no boundary condition on $\Gamma_{\tilde{H}}$. This is viewed as a special case of (3.11) in which both $F_{\tilde{j}}^{(2)}$ and $\tilde{H}$ are assumed to be identically zero. Clearly, various combinations of the above boundary conditions may also be specified on portions of Γ , although this is not explicitly spelled out in the sequel.

4. WEIGHTED RESIDUAL FORMULATION

Consider a point x in Γ_{int} . Designate (arbitrarily) one side of Γ_{int} to be the "plus side" and the other to be the "minus side". Let $\tilde{n}^+$ and $\tilde{n}^-$ be unit normal vectors to Γ_{int} at x which point in the plus and minus directions, respectively. Clearly, $\tilde{n}^- = -\tilde{n}^+$. Let $F_{\tilde{j}}^+$ and $F_{\tilde{j}}^-$ denote the values of $F_{\tilde{j}}$ obtained by approaching x from the positive and negative sides, respectively. The "jump" in $F_{\tilde{n}}$ at x is defined to be

$$\begin{aligned} [F_{\tilde{n}}] &= (F_{\tilde{j}}^+ - F_{\tilde{j}}^-) \tilde{n}_i^+ \\ &= F_{\tilde{j}}^+ \tilde{n}_i^+ + F_{\tilde{j}}^- \tilde{n}_i^-. \end{aligned} \quad (4.1)$$

As may be readily verified from (4.1), the jump is invariant with respect to reversing the plus and minus designations.

Throughout, we shall assume that trial solutions, U , satisfy $\partial U = G$ on Γ_G and weighting functions, W , satisfy $\partial W = 0$ on Γ_G . Thus all Dirichlet type boundary conditions are treated as essential boundary conditions in the present formulation.

The variational equation is assumed to take the form

$$\begin{aligned} \int_{\Omega} \{ W_{\tilde{t}}^T (U_{\tilde{t}} + F_{\tilde{j},j}^{(1)} + G) - W_{\tilde{j}}^T F_{\tilde{j}}^{(2)} \} d\Omega \\ + \sum_{e=1}^{n_{el}} \int_{\Omega} \tilde{P}^T (U_{\tilde{t}} + F_{\tilde{j},j} + G) d\Omega = \int_{\Gamma_{\tilde{H}}} W_{\tilde{H}}^T H d\Gamma. \end{aligned} \quad (4.2)$$

In (4.2), U and W are assumed to be taken from the same class of typical C^0 finite element interpolations; $\tilde{P}$ is a C^{-1} perturbation to the weighting function.

The Euler-Lagrange conditions emanating from (4.2) may be deduced by way of integration by parts:

$$0 = \sum_{e=1}^{n_e} \int_{\Omega_e} \tilde{W}^T (U_{\sim,t} + F_{\sim,j,j} + G_{\sim}) d\Omega - \int_{\Gamma_H} \tilde{W}^T (F_{\sim,n}^{(2)} + H_{\sim}) d\Gamma - \int_{\Gamma_{int}} \tilde{W}^T [F_{\sim,n}^{(2)}] d\Gamma. \quad (4.3)$$

From (4.3) we see that the Euler-Lagrange equations are (3.1) restricted to the element interiors, (3.11), and the (partial) flux continuity condition across inter-element boundaries, namely

$$[F_{\sim,n}^{(2)}] = 0. \quad (4.4)$$

Note that (3.11) is a natural boundary condition.

Remarks

- 1) Note that if $\tilde{P} = 0$ we have a Galerkin weighted residual formulation; if $\tilde{P} \neq 0$ we have a Petrov-Galerkin formulation.

The modified weighting function, that is

$$\tilde{W} = W + \tilde{P}, \quad (4.5)$$

is confined to the element interiors and thus does not affect boundary or continuity conditions.

- 2) The preceding formulation generalizes (Hughes and Brooks, 1982) (which was restricted to the linear advection-diffusion equation) to systems of conservation laws. A related but somewhat different formulation for the incompressible Navier-Stokes equations is described in (Brooks and Hughes, 1982a,b). The case in which $F_{\sim,j}^{(2)}$ and $H_{\sim}$ are 0 has been employed in (Tezduyar and Hughes, 1982) for the compressible Euler equations. See also (Griffiths, 1981) for a related formulation.

The following examples will illustrate potential applications of the preceding formulation.

Examples

1. The scalar linear advection-diffusion equation

In this case we have

$$\phi_{,t} + \sigma_{j,j} - \delta = 0 \quad (4.6)$$

where

$$\sigma_j = \sigma_j^a + \sigma_j^d \quad (\text{total flux}) \quad (4.7)$$

$$\sigma_j^a = \sigma_j^a(\phi) = u_j \phi \quad (\text{advective flux}) \quad (4.8)$$

$$\sigma_j^d = \sigma_j^d(\phi) = -k_{jk} \phi_{,k} \quad (\text{diffusive flux}). \quad (4.9)$$

In the above, δ is a source term, u_j is the flow velocity, and k_{jk} is diffusivity. Each of δ , u_j and k_{jk} is assumed to be a given function of x and t .

If Γ_g and Γ_h decompose the boundary such that

$$\Gamma = \Gamma_g \cup \Gamma_h \quad (4.10)$$

$$\emptyset = \Gamma_g \cap \Gamma_h \quad (4.11)$$

then we shall assume Dirichlet conditions on Γ_g , that is

$$\phi = g \text{ on } \Gamma_g \quad (4.12)$$

where g is a given function defined on Γ_g . The three possibilities on Γ_h are:

total flux boundary condition

$$-\sigma_n = h \text{ on } \Gamma_h \quad (4.13)$$

where h is a given function defined on Γ_h ;

diffusive flux boundary condition

$$-\sigma_n^d = h \text{ on } \Gamma_h; \quad (4.14)$$

and, finally,

no boundary condition on Γ_h .

This last condition occurs in cases of pure advection in which Γ_h is defined to be that part of Γ on which $u_n \geq 0$.

The initial condition is

$$\phi(\tilde{x}, 0) = \phi_0(\tilde{x}) \quad (4.15)$$

where ϕ_0 is a given function of $\tilde{x} \in \Omega$.

The linear advection-diffusion equation may be brought within the general framework by the following conditions:

$$m = 1 \quad (4.16)$$

$$\tilde{U} = \phi \quad (4.17)$$

$$\tilde{W} = w \quad (4.18)$$

$$\tilde{P} = \tilde{p} \quad (4.19)$$

$$\tilde{F}_j = \sigma_j \quad (4.20)$$

$$\tilde{G} = -\phi \quad (4.21)$$

$$\tilde{\partial} = 1 \quad (4.22)$$

$$\tilde{G} = g \quad (4.23)$$

$$\tilde{H} = h \quad (4.24)$$

$$\tilde{U}_0 = \phi_0 \quad (4.25)$$

To attain the various conditions, $\tilde{F}_j^{(1)}$ and $\tilde{F}_j^{(2)}$ need to be set as follows:

total flux boundary condition

$$\tilde{F}_j^{(1)} = 0; \quad \tilde{F}_j^{(2)} = \tilde{F}_j = \sigma_j \quad (4.26)$$

diffusive flux boundary condition

$$\tilde{F}_j^{(1)} = \sigma_j^a; \quad \tilde{F}_j^{(2)} = \sigma_j^d \quad (4.27)$$

no boundary condition on Γ_h

$$\tilde{F}_j^{(1)} = F_j = \sigma_j = \sigma_j^a; \quad \tilde{F}_j^{(2)} = \sigma_j^d = 0; \quad \tilde{H} = h = 0. \quad (4.28)$$

We have found the diffusive flux boundary condition to be very effective in numerically simulating outflow conditions (see, e.g., Hughes and Brooks, 1982, Brooks and Hughes, 1982a,b). Cohen (1982) has employed the total flux condition successfully in certain problems of oil reservoir simulation.

The perturbation to the weighting function, $\tilde{p}$, may be selected in various ways. In the streamline-upwind Petrov-Galerkin formulation, originated in (Hughes and Brooks, 1979, 1982) and (Brooks and Hughes, 1980), it takes the form

$$\tilde{p} = \tau_{i,i} \sigma_{i,i}^a(w) = \tau(u_i w)_{,i} \quad (4.29)$$

where τ is an optimization parameter. Selection of τ is discussed in Section 5. Kelly et al. (1980) later independently proposed a similar idea (see also (Nakazawa 1982) for further developments).

2. The compressible Euler equations (inviscid gas dynamics)

The equations are

$$\rho_{,t} + (\rho u_j)_{,j} = 0 \quad (\text{continuity}) \quad (4.30)$$

$$(\rho u_i)_{,t} + (\rho u_j u_i)_{,j} = \sigma_{ij,j} + f_i \quad (\text{momentum}) \quad (4.31)$$

$$(\rho e)_{,t} + (\rho u_j e)_{,j} = r + f_i u_i + (\sigma_{ij} u_i)_{,j} \quad (\text{energy}) \quad (4.32)$$

where

$$\sigma_{ij} = -p \delta_{ij}, \quad p = \hat{p}(\rho, \epsilon) \quad (4.33)$$

$$e = \epsilon + u^2/2, \quad u = \|u\| = (u_i u_i)^{1/2} \quad (4.34)$$

in which ρ is the density, u_j is the velocity, σ_{ij} is the Cauchy stress tensor, f_i is the prescribed body force (per unit volume), e is the total energy density, r is the heat supply (per unit volume), p is the pressure, and ϵ is the internal energy density.

These equations may be put into the format of a hyperbolic system of conservation laws by employing the following definitions:

$$m = n_{sd} + 2 \quad (4.35)$$

$$U_1 = \rho \quad (4.36)$$

$$U_{j+1} = \rho u_j, \quad 1 \leq j \leq n_{sd} \quad (4.37)$$

$$U_m = \rho e \quad (4.38)$$

$$F_j = (U_{j+1}/U_1) U + \begin{Bmatrix} 0 \\ p \delta_j \\ p U_{j+1}/U_1 \end{Bmatrix} \quad (4.39)$$

$$p = \hat{p} \left(U_1, U_m/U_1 - \frac{1}{2} \sum_{i=1}^{n_{sd}} U_{i+1}^2 / U_1^2 \right) \quad (4.40)$$

$$\delta_j = \begin{Bmatrix} \delta_{j1} \\ \cdot \\ \cdot \\ \delta_{jn_{sd}} \end{Bmatrix} \quad (4.41)$$

$$G = - \begin{Bmatrix} 0 \\ f \\ r + f_i u_i \end{Bmatrix} \quad (4.42)$$

Consideration of the characteristics and particular physical situation lead to appropriate boundary condition specifications. A variety of Dirichlet type conditions (3.10) may be envisaged. Three potentially useful boundary conditions on Γ_H may be set up as follows:

total flux boundary condition

$$\tilde{F}_j^{(1)} = 0 ; \tilde{F}_j^{(2)} = \tilde{F}_j \quad (4.43)$$

pressure ("traction") boundary condition

Pressure may be specified as a natural boundary condition by selecting

$$\tilde{F}_j^{(1)} = (U_{j+1}/U_1) \tilde{U} , \quad (4.44)$$

$$\tilde{F}_j^{(2)} = \begin{Bmatrix} 0 \\ p \delta_j \\ p U_{j+1}/U_1 \end{Bmatrix} , \quad (4.45)$$

and

$$\tilde{H} = - \begin{Bmatrix} 0 \\ h n \\ h u_n \end{Bmatrix} , \quad (4.46)$$

where h is the prescribed value of pressure defined on Γ_h .

no boundary condition on Γ_H

$$\tilde{F}_j^{(1)} = \tilde{F}_j ; \tilde{F}_j^{(2)} = 0 ; \tilde{H} = 0. \quad (4.47)$$

In (Tezduyar and Hughes, 1982) we have employed the formulation defined by (4.47). In modelling outflow conditions, (4.47) appears to be effective although pressure specification may be more appropriate in some circumstances.

As in the previous example, there are many possibilities for defining the weighting function perturbation, $\tilde{P}$. In (Tezduyar and Hughes, 1982) we tested the following definition

$$\tilde{p} = \tilde{z}_i \tilde{w}_{,i} \quad (4.48)$$

where

$$\tilde{z}_i = \tau_i A_i \quad (\text{no sum}) \quad (4.49)$$

or

$$\tilde{z}_i = \tau_i A_i^T \quad (\text{no sum}). \quad (4.50)$$

(Recall that $A_i = \partial F_i / \partial U_i$.) Each of (4.49) and (4.50) leads to interesting formulations. We have experimented with both, and also with the definition of τ_i . At present we favour (4.50) over (4.49) due to its superior behaviour on nonlinear problems. This is discussed more fully in (Tezduyar and Hughes, 1982). A sketch of the procedures used to define τ_i is contained in Section 5.

It is interesting to observe that if $\hat{p}(\rho, \epsilon) = \rho f(\epsilon)$, where f is an arbitrary function,* then $F_j(U)$ is a homogeneous function of degree 1, that is $F_j(\alpha U) = \alpha F_j(U)$ for all $\alpha \in \mathbb{R}$. In this case it follows that $F_j = A_j U$. The individual partial fluxes defined by (4.44) and (4.45) also are homogeneous functions of degree 1.

Remarks

1. The full, compressible Navier-Stokes equations, including thermal effects, can also be subsumed by the general weighted residual format of (4.2). Likewise, various natural boundary conditions can be built into the formulation. However, this lies outside the scope of the present paper.
2. Discretization of (4.2) is carried out by expanding U and W in terms of a set of finite element basis, or shape, functions. For example, the expression for U might take the form[†]

$$U(x, t) = \sum_B N_B(x) U_B(t) \quad (4.51)$$

*For example, this is the case for a perfect gas defined by $\hat{p}(\rho, \epsilon) = (\gamma - 1)\rho\epsilon$, $\gamma \in \mathbb{R}$

† One could also employ different shape functions for the individual components of U . This idea becomes important in "constrained" cases, such as incompressibility (see, e.g., Thomasset, 1981). Similar considerations need to be made if compressible flow algorithms are to be exercised at very low Mach numbers.

where B is a nodal index, N_B is the shape function associated with node B , and $\underline{U}_B$ is the value of $\underline{U}$ at node B . This leads to a weighted residual formulation in the strict sense. Certain modifications of the basic formulation--such as reduced integration or use of lower-order interpolants--are interesting from the standpoints of efficiency and, occasionally, lead to improved accuracy.* We have been fond of the use of reduced integration techniques in our finite element work (see e.g. Hughes et al., 1979) and have experimented with them in the present context in (Tezduyar and Hughes, 1982).

The use of interpolants of $\underline{F}_j$ and $\underline{G}$ has an interesting consequence. It produces schemes which are reminiscent of, and at the same time generalize, classical "conservative differencing schemes". The basic idea is to approximate $\underline{F}_j$ and $\underline{G}$ by expansions in terms of shape functions and the nodal values of $\underline{F}_j$ and $\underline{G}$, respectively. For example, the expansion for $\underline{F}_j$ might take the following form:

$$\underline{F}_j(\underline{x}) = \sum_B N_B(\underline{x}) \underline{F}_{jB} \quad (4.52)$$

where

$$\underline{F}_{jB}(t) = \underline{F}_j(\underline{U}_B(t), (\nabla \underline{U})_B(t), \underline{x}_B, t), \quad (4.53)$$

$$(\nabla \underline{U})_B(t) = \sum_C \nabla N_C(\underline{x}_B) \underline{U}_C(t) \quad (4.54)$$

and likewise for $\underline{G}$. Christie et al. (1981) have proposed and examined schemes of this kind. They term them "product approximations". Spradley et al. (1980) have also adopted this idea in their "general interpolants method". Various other interesting finite element and finite difference concepts are synthesized in their approach.

The use of reduced integration and/or the use of interpolants simplifies element calculations. Their relative merits from the standpoints of accuracy, stability, shock structure, etc., do not seem certain at this point and thus further investigations are warranted.

* These violations of the weighted residual recipe are often referred to as "variational crimes", a terminology coined by Strang (see Strang and Fix, 1973).

3. The development of so-called finite volume techniques also emanates from integral forms of the conservation equations (see, e.g. Caughey and Jameson, 1979). These techniques, which are often thought of as finite difference methods, have essential features in common with the integral difference methods originated at Lawrence Livermore Laboratory (see e.g. Wilkins, 1969). More recently it has been shown how to derive the latter class of methods via finite element/reduced integration concepts (see Goudreau and Hallquist, 1982 for a survey of these ideas).

We thus anticipate that the finite volume method will also find a place within the finite element hierarchy.

5. PETROV-GALERKIN METHODS OF CHARACTERISTIC TYPE

In this Section we discuss further details of the class of Petrov-Galerkin methods described in Section 4 and propose some further generalizations.

Selection of τ_i

We begin with a presentation of the formulae we have used for the optimization parameters. As yet no universal scheme has been formulated. The following criteria have been investigated so far.

Temporal criterion

In conjunction with certain time-stepping algorithms, a criterion based upon the time step, Δt , proves effective:

$$\tau_i = \tau = F \alpha \Delta t, \quad 1 \leq i \leq n_{sd}. \quad (5.1)$$

In (5.1), α is an algorithmic parameter and F is a parameter which enables us to adjust the magnitude of τ_i for various purposes. The value of α is usually set to $\frac{1}{2}$ in cases in which we are interested in time accuracy, whereas it is set to 1 in cases in which we wish to rapidly obtain a steady flow. An interesting explicit algorithm which attains second-order time accuracy can be constructed by setting $\alpha = \frac{1}{2}$ and $F = 1$ (see Dukowitz and Ramshaw, 1979 and Tezduyar and Hughes, 1982 for further details).

Since (5.1) pertains to all elements in a mesh it is a global criterion.

Spatial criteria

Two spatial criteria have been employed:

$$(a) \quad \tau_i = \tau = F \alpha \tilde{\xi} h / \lambda, \quad 1 \leq i \leq n_{sd} \quad (5.2)$$

$$(b) \quad \tau_i = F \alpha \tilde{\xi} h_i / \lambda_i, \quad (\text{no sum}), \quad 1 \leq i \leq n_{sd} \quad (5.3)$$

where λ_i is the spectral radius of A_i ,

$$\lambda = \|\tilde{\lambda}\| = (\lambda_i \lambda_i)^{1/2} \quad (5.4)$$

$$h_i = 2 \|\tilde{\nabla} x_i\|^* \quad (5.5)$$

$$h = h_i \lambda_i / \lambda \quad (= 0 \text{ if } \lambda = 0) \quad (5.6)$$

$$\tilde{\xi} = (\coth \alpha) - 1/\alpha \quad (5.7)$$

$$\alpha = \lambda h / (2d). \quad (5.8)$$

In (5.5), $\tilde{\nabla}$ denotes the gradient operator with respect to the canonical isoparametric coordinates. The factor of 2 appears since the isoparametric parent domain is usually scaled to have a length of 2 (Zienkiewicz, 1977). The following examples should make the definition clear:

$$h = 2 |\partial x / \partial \xi| \quad (\text{one dimension}) \quad (5.9)$$

$$h_i = 2 [(\partial x_i / \partial \xi)^2 + (\partial x_i / \partial \eta)^2]^{1/2} \quad (\text{two dimensions}). \quad (5.10)$$

In (5.8), d is a diffusivity coefficient which takes on the following definitions:

$$d = k_{ij} \lambda_i \lambda_j / \lambda^2 \quad (\text{advection-diffusion eq.}) \quad (5.11)$$

$$d = 0 \quad (\text{Euler eqs.}). \quad (5.12)$$

Note that $d = 0$ implies $\tilde{\xi} = 1$.

*This definition was prompted by the work of Nakazawa (1982).

The values of τ_i given by (5.2) and (5.3) depend upon the geometry and state of an individual element. In this sense (5.2) and (5.3) are local criteria.

Various model equations have been employed to select $F\alpha$ in (5.2) and (5.3). The value $F\alpha = \frac{1}{2}$ has been determined by considering the one-dimensional, steady advection-diffusion equation (Hughes and Brooks, 1979). A boundary-layer investigation also points to this value (Hughes and Brooks, 1980). An analysis of the semi-discrete, time-dependent, advection equation in one dimension yields a value $F\alpha = 1/\sqrt{15}$ (Raymond and Garder, 1976). Fourth-order phase accuracy is thereby attained for linear elements. We have employed this value in transient advection-diffusion and Navier-Stokes calculations such as the ones described in Section 6. Values of $F\alpha > \frac{1}{2}$ have been experimented with to achieve effective resolution of shocks governed by first-order hyperbolic systems.

Accuracy, stability and shock structure

In the context of the advection-diffusion equation, the streamline-upwind/Petrov-Galerkin method described in Section 4, and above, leads to difference stencils which are skewed with respect to streamlines and are centred about points upwind of the nodal point to which they are associated. In this way the accuracy of centred, skew schemes is attained while, at the same time, the stability of upwind procedures is achieved. Claes Johnson (1981) has proved that this formulation leads to optimal error estimates* and furthermore that effects are propagated in the discrete problem approximately as in the continuous problem. In particular, Johnson (1981) has established that discontinuities of exact solutions across characteristics (i.e., streamlines in the present case) will be numerically captured in a very narrow region around the characteristic, even when the characteristics are arbitrarily skew to the mesh.

The methods proposed for hyperbolic systems possess similar interesting structure and our experience in non-linear cases with shocks is very good so far. However, it should be mentioned that, for this class of problems, a strong case can be made for monotone methods (see Harten et al., 1976 and the works of Osher and collaborators, e.g., Osher, 1980, 1981a,b, 1982, Engquist and Osher, 1980, Goorjian and Van Buskirk, 1981). We are still hopeful that within the framework that we have delineated, by appropriate local selection of the optimization parameter, we

* In the Galerkin method, optimal asymptotic error estimates fail to hold in general, even when the exact solution is smooth (Dupont, 1973). Rough exact solutions create particular havoc, causing spurious oscillations which pollute the solution globally.

will achieve the attributes of monotone methods, concomitantly with the desirable features delineated above. Admittedly, this is an area requiring further investigation.

We wish to now consider a definition of $\tilde{P}$ which generalizes the forms that we have previously considered. This definition is motivated by Jean Donea's development of the class of Taylor-Galerkin methods (Donea, 1982) which possess excellent phase-accuracy properties.

$$\text{Let} \quad \tilde{P} = L W \quad (5.13)$$

where L is a linear operator defined by the expansion

$$L = L(n, L) = L + \frac{1}{2} L^2 + \dots + \frac{1}{n!} L^n \quad (5.14)$$

In (5.14), L is also a linear operator. Clearly, if $n = 1$, $L = L$. The definitions of the perturbation that we have used in our previous work are special cases of (5.13), (5.14) with $n = 1$. This can be seen as follows:

advection-diffusion equation; cf. (4.29)

$$\tilde{p} = L W = L W = \tau(u_i w)_{,i} \quad (5.15)$$

compressible Euler equations; cf. (4.48)

$$\tilde{P} = L W = L W = T_i W_i \quad (5.16)$$

Formally generalizing (5.16) to the case of $n \geq 1$ leads to the expression

$$\tilde{P} = L W = (T_i \frac{\partial}{\partial x_i} + \dots + (T_i \frac{\partial}{\partial x_i})^n) W \quad (5.17)$$

If in (5.17)

$$T_i = \frac{\Delta t}{2} A_i^T \quad (5.18)$$

we arrive at a family of methods similar to Donea's. By using, say

$$T_i = \tau_i A_i^T \quad (\text{no sum}) \quad (5.19)$$

with τ_i selected according to spatial criteria, we generate other interesting possibilities. The spatial criteria have the advantage that steady flows are independent of the size of the time step employed, and seem better suited to meshes involving

significant variation in element size. Clearly these ideas are in a speculative phase and require further investigation.

6. NUMERICAL EXAMPLES

The rotating cone (advection-diffusion equation)

An elevation plot for the rotating-cone problem is shown in Fig. 1. The cone is shown after having rotated 360° . The trapezoidal-rule algorithm was used for temporal discretization of the streamline-upwind Petrov-Galerkin formulation. A 30×30 mesh of bilinear elements was employed.

For additional numerical examples and comparisons with different approaches, see (Hughes and Brooks, 1982).

Karman vortex street (incompressible Navier-Stokes equations)

Our procedures for solving the incompressible Navier-Stokes equations are described in (Brooks and Hughes, 1982a,b) which may be consulted for further details.

Stationary streamlines are shown in Fig. 2 for a computation of the Karmen vortex street at a Reynolds number of 100. The elements employed assumed bilinear velocity and constant pressure. The mesh consisted of 1436 elements. The time-stepping algorithm treats velocity degrees-of-freedom explicitly and pressures implicitly.

Transonic nozzle flow (compressible Euler equations)

Shock profiles are shown in Fig. 3. The exact solution (shown solid) has been provided by Lomax et al. (1981). This problem is one of steady flow, consequently a value of $\alpha = 1$ (backward differences) was employed. Various values of F were tested and their effects on the shock profile may be seen in Fig. 3. At $F = 1$, the shock front is captured across one element (40 linear elements were employed). As F is increased the shock tends to be smeared somewhat. In the calculation shown

$\tilde{T} = \tilde{A}^T$ and the local criterion, (5.2), was employed.

Shocks structure/entropy condition test problem

This problem was suggested to us by S. Osher. The equation is

$$u_t + \left(\frac{u^2}{2}\right)_x = 0 \quad (6.1)$$

and the initial condition is shown in frame 0 of Fig. 4. Note that the initial condition involves two stable shocks and one

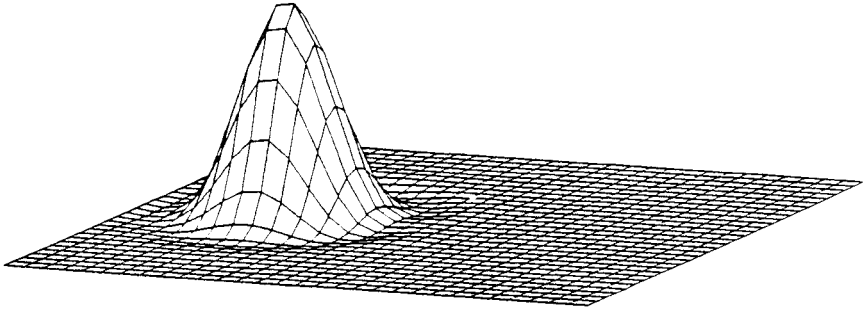
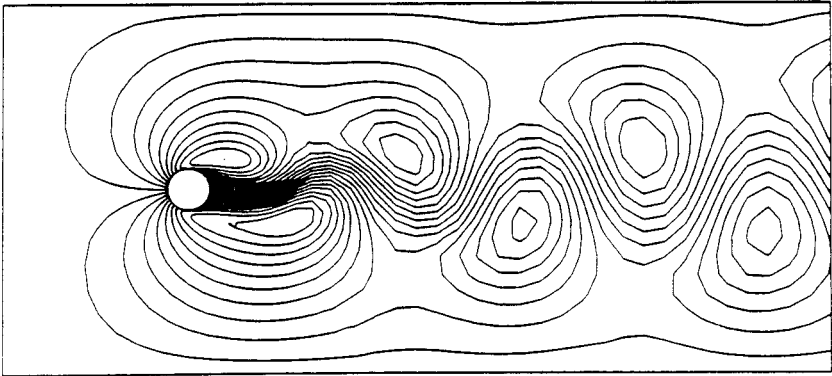


Fig. 1 Advection of a cone.

Fig. 2 Karman vortex street with $Re = 100$.

unstable shock. Each successive frame corresponds to a time step ($\Delta t = 2.174$). As may be seen, the unstable shock immediately breaks down and eventually the stable shocks coalesce to form a steady shock (frame 23). The finite scheme nodally interpolates the exact steady state.

Equation (6.1) is a special case of the hyperbolic system formulation described in Section 3. In the calculation shown in Fig. 4, 40 linear elements were employed, the spatial criterion (5.2) was employed, $F = 1$ and $\alpha = \frac{1}{2}$.

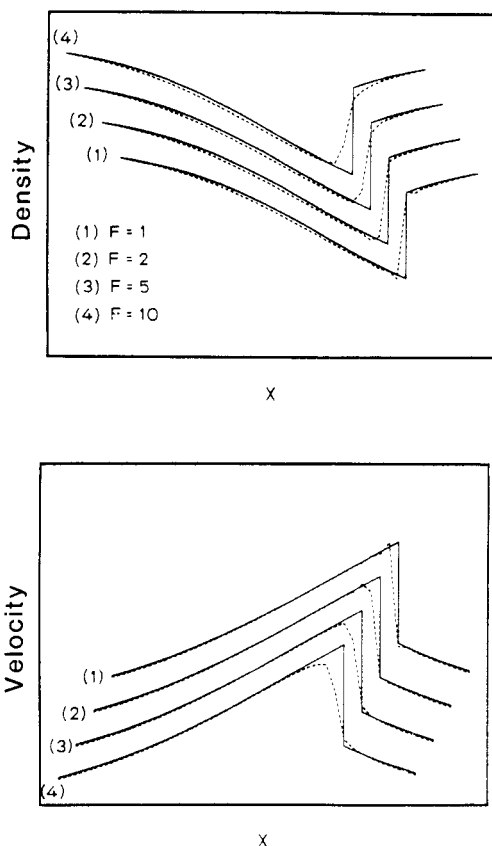


Fig. 3 Transonic nozzle flow problem.

Thin biconvex airfoil (compressible Euler equations)

Sample results are presented for a transonic flow about a thin biconvex airfoil. Details about the problem and additional results may be found in (Tezduyar and Hughes, 1982). "Medium" and "fine" meshes are illustrated in Fig. 5. In the medium mesh 16 elements are used to discretize the airfoil whereas 32 elements are used in the fine mesh. (The fine mesh represents a bisection of the medium mesh.) For this problem the weighting function perturbation was defined by (4.50), the temporal criterion (5.1) was employed, $F = 1$ and $\alpha = 1$. Pressure coefficient plots are presented in Fig. 6 for channel and free-stream boundary conditions along the upper edge. The results show a shift in the shock location of about 6% for the two boundary condition cases which is to be expected. Mach number contours are illustrated in Fig. 7 for the channel boundary condition case.

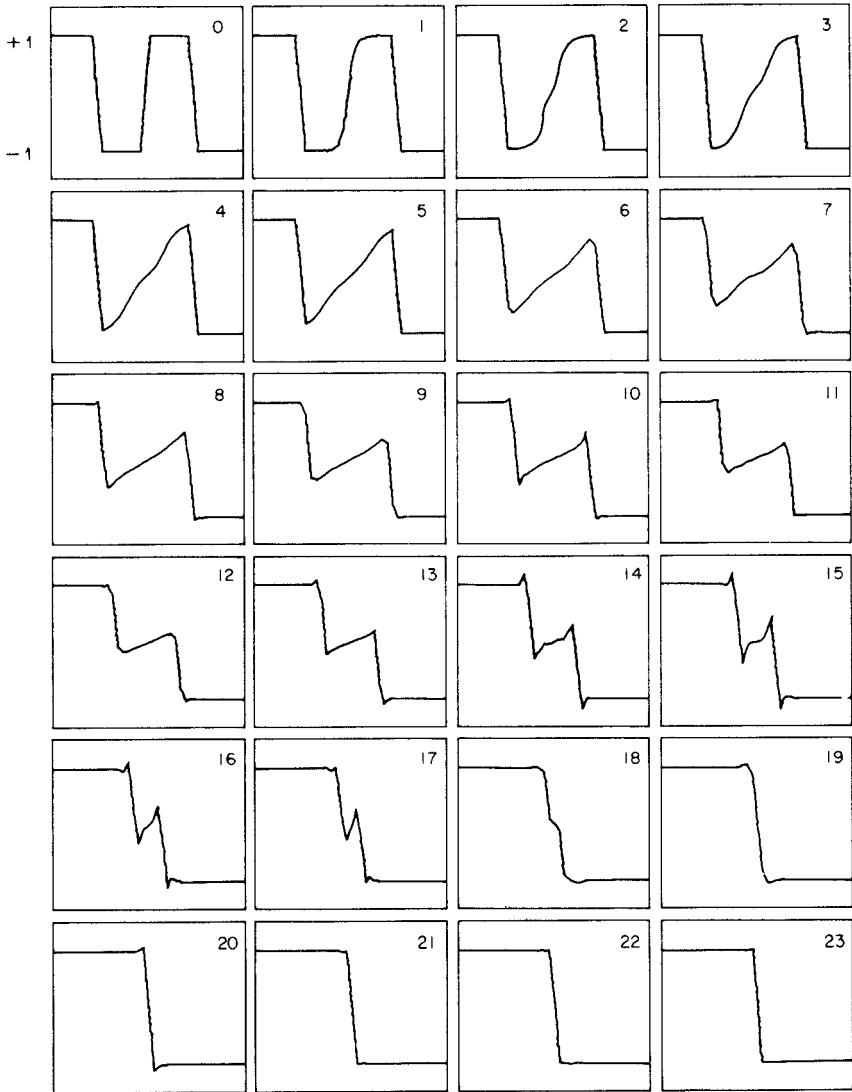
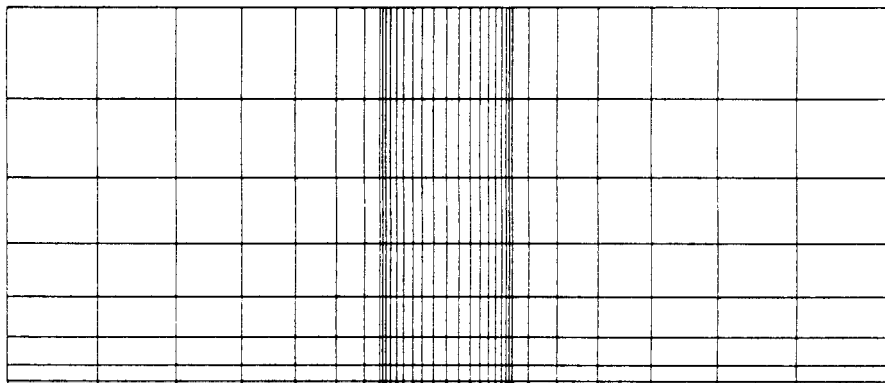


Fig. 4 Burger's equation test problem.

MEDIUM MESH



FINE MESH

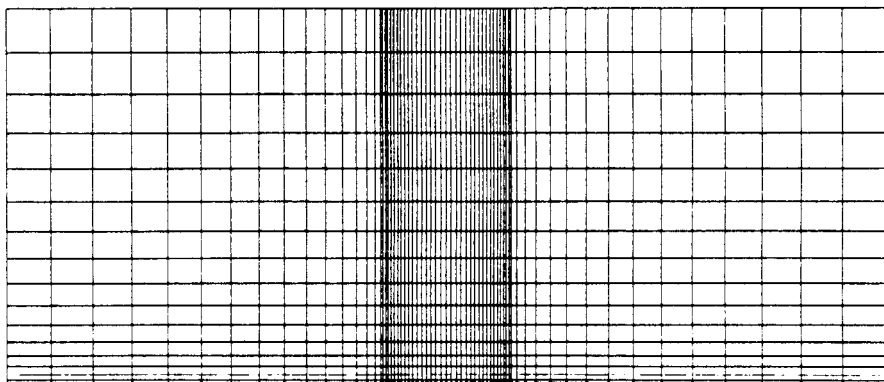


Fig. 5 Finite element meshes. Medium mesh, 256 elements;
fine mesh, 1016 elements.

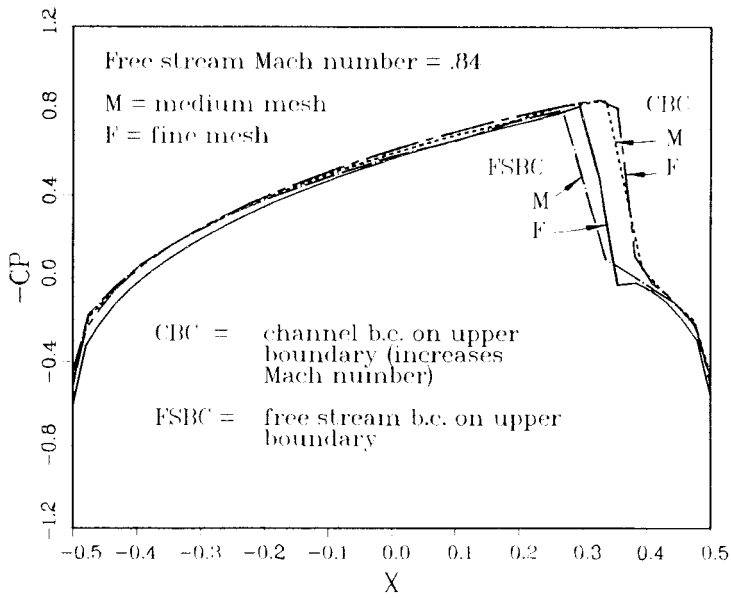


Fig. 6 Pressure coefficient for biconvex airfoil.

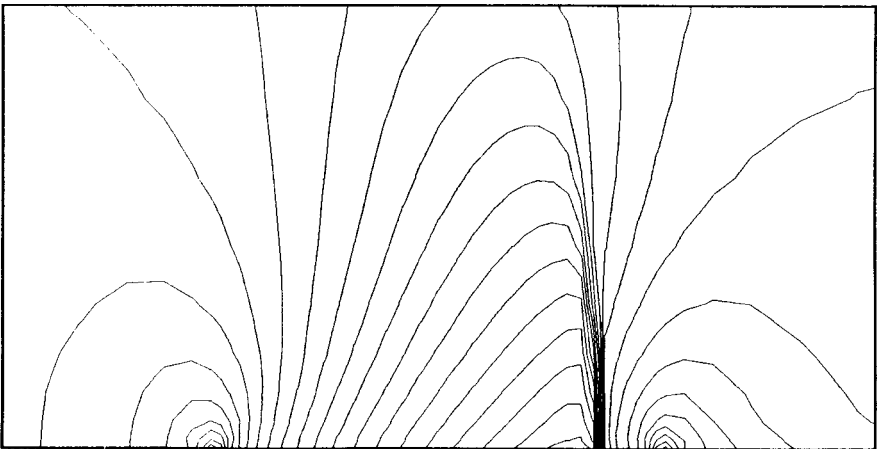


Fig. 7 Transonic case, fine mesh, isomachs
Mach number contours for the airfoil problem.

7. CONCLUSIONS

In this paper we have presented a Petrov-Galerkin formulation which generalizes the streamline upwind concept to systems of conservation laws. Emphasis has been placed on the compressible Euler equations in the present work. Numerical examples indicate the good behaviour of the method on a variety of problems, including ones with shocks.

Future work needs to focus on optimizing the formulation with respect to shock structure and entropy conditions. Furthermore, it is hoped to tackle a number of aerodynamic flows of a complex nature. Finally, the development of element-by-element, and subdomain-by-subdomain, implicit algorithms (see, e.g., Hughes et al., 1982, 1983) for the present class of problems will hopefully open the way to much more efficient solution of large problems, particularly three-dimensional ones.

ACKNOWLEDGEMENT

We would like to thank the Computational Fluid Dynamics Branch of the NASA Ames Research Center for supporting our research, and especially H. Lomax for guidance and encouragement.

REFERENCES

- BAKER, A.J. 1979 Research on numerical algorithms for the three-dimensional Navier-Stokes equations, 1. Accuracy, Convergence and Efficiency, Technical Report AFFDL-TR-79-3141, Wright-Patterson Air Force Base, Ohio, USA.
- BARRETT, J.W. and MORTON, K.W. 1980 Optimal finite element solutions to diffusion convection problems in one dimension, *Int. J. Num. Meth. Eng.*, **15**, pp. 1457-1474.
- BARRETT, J.W. and MORTON, K.W. 1981 Optimal Petrov-Galerkin Methods through approximate symmetrization, *IMA J. Num. Analysis*, **1**, pp. 439-468.
- BROOKS, A. and HUGHES, T.J.R. 1980 Streamline-upwind/Petrov-Galerkin methods for advection dominated flows. In Proceedings of the Third International Conference on Finite Element Methods in Fluid Flow, Banff.
- BROOKS, A. and HUGHES, T.J.R. 1982a Streamline-upwind/Petrov-Galerkin formulation for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. To appear in the proceedings of FENOMECH '81 and Computer Methods in Applied Mechanics and Engineering.

- BROOKS, A. and HUGHES, T.J.R. 1982b An algorithm for solving the Navier-Stokes equations based upon the streamline-upwind Petrov-Galerkin formulation. To appear in *Numerical Methods for Coupled Problems*, London, J. Wiley & Sons.
- CAUGHEY, D.A. and JAMESON, A. 1979 Numerical calculation of transonic potential flow about wing-body combinations, *AIAA J.* **17**, pp. 175-181.
- CHRISTIE, I., GRIFFITHS, D.F., MITCHELL, A.R. and SANZ-SERNA, J.M. 1981 Product approximation for non-linear problems in the finite element method, *IMA J. Num. Analysis*, **1**, pp. 253-266.
- CHRISTIE, I., GRIFFITHS, D.F., MITCHELL, A.R. And ZIENKIEWICZ, O.C. 1976 Finite element methods for second order differential equations with significant first derivatives, *Int. J. Num. Meth. Eng.*, **10**, pp. 1389-1396.
- COHEN, M. 1982 Private Communication.
- DENDY, J.E. 1974 Two methods of Galerkin type achieving optimum L^2 rates of convergence for first order hyperbolics, *SIAM J. Num. Analysis*, **11**, pp. 637-653.
- DONEA, J. 1982 A Taylor-Galerkin method for convective transport problems, preprint.
- DUKOWICZ, J.K. and RAMSHAW, J.D. 1979 Tensor viscosity method for convection in numerical fluid dynamics, *J. Comput. Phys.*, **32**, pp. 71-79.
- DUPONT, T. 1973 Galerkin methods for first-order hyperbolics: an example, *SIAM J. Num. Analysis*, **10**, pp. 890-899.
- ENGQUIST, B. and OSHER, S. 1980 One-sided difference schemes and transonic flow, *Proc. Natl. Acad. Sci. USA*, **77**, pp. 3071-3074.
- GOORJIAN, P.M. and VAN BUSKIRK, R. 1981 Implicit calculations of transonic flows using monotone methods, AIAA-81-0331, AIAA 19th Aerospace Sciences Meeting, St. Louis, Missouri.
- GOUDREAU, G.L. and HALQUIST, J. O. 1982 Recent developments in large scale finite element Lagrangian hydrocode technology, *Comp. Meth. Appl. Mech. Eng.*, to appear.
- GRIFFITHS, D.F. 1981 A Petrov-Galerkin method for hyperbolic equations. To appear in *Proceedings of MAFELAP '81* (J. Whiteman, Ed.), Brunel University.

- GRIFFITHS, D.F. and MITCHELL, A.R. 1979 On generating upwind finite element methods, *AMD* **34**. In *Finite Element Methods for Convection Dominated Flows* (T.J.R. Hughes, Ed.), New York, ASME.
- HARTEN, A., HYMAN, J.M. and LAX, P.D. 1979 On finite difference approximations and entropy conditions for shocks, *Comm. Pure Appl. Maths.*, **XXIX**, pp. 297-322.
- HEINRICH, J.C., HUYAKORN, P.S., ZIENKIEWICZ, O.C. and MITCHELL, A.R. 1977 An 'upwind' finite element scheme for two-dimensional convective transport equations, *Int. J. Num. Meth. Eng.*, **11**, pp. 134-143.
- HEINRICH, J. and ZIENKIEWICZ, O.C. 1979 The finite element method and 'upwinding' techniques in the numerical solution of convection dominated flow problems, *AMD* **34**. In *Finite Element Methods for Convection Dominated Flows* (T.J.R. Hughes, Ed.), New York, ASME.
- HUGHES, T.J.R. and BROOKS, A. 1979 A multidimensional upwind scheme with no crosswind diffusion, *AMD* **34**. In *Finite Element Methods for Convection Dominated Flows* (T.J.R. Hughes, Ed.), New York, ASME.
- HUGHES, T.J.R. and BROOKS, A. 1980 Galerkin/upwind finite element mesh partitions in fluid mechanics. In *Boundary and Interior Layers - Computational and Asymptotic Methods*, (J.J.H. Miller, Ed.), Dublin, Boole Press.
- HUGHES, T.J.R. and BROOKS, A. 1982 A theoretical framework for Petrov-Galerkin methods with discontinuous weighting functions: application to the streamline upwind procedure. To appear in *Finite Elements in Fluids*, **4** (R.H. Gallagher, Ed.), London, J. Wiley & Sons.
- HUGHES, T.J.R., LEVIT, I. and WINGET, J. 1982 An implicit unconditionally stable element-by-element algorithm for heat conduction analysis. To appear in *J. Eng. Mech. Div.*, ASCE.
- HUGHES, T.J.R., LEVIT, I. and WINGET, J. 1983 An element-by-element solution algorithm for problems of structural and solid mechanics. To appear in *Comput. Meth. Appl. Mech. Eng.*
- HUGHES, T.J.R., LIU, W.K. and BROOKS, A. 1979 Review of finite element analysis of incompressible viscous flows by the penalty function formulation, *J. Comput. Phys.*, **30**, pp. 1-60.
- JOHNSON, C. 1981 Finite element methods for convection-diffusion problems, Fifth International Symposium on Computing Methods

- in Engineering and Applied Sciences, INRIA, Versailles.
- KELLY, D.W., NAKAZAWA, S., ZIENKIEWICZ, O.C. and HEINRICH, J.C. 1980 A note on upwinding and anisotropic balancing dissipation in finite element approximations to convective diffusion problems, *Int. J. Num. Meth. Eng.*, **15**, pp. 1705-1711.
- LOMAX, H. et al. 1981 Private Communication.
- MORTON, K.W. 1981 Finite element methods for non-self-adjoint problems, Numerical Analysis Report 3/81, University of Reading.
- MORTON, K.W. and BARRATT, J.W. 1980 Optimal finite element methods for diffusion-convection problems. In Boundary and Interior Layers- Computational and Asymptotic Methods, (J.J.H. Miller, Ed.), Dublin, Boole Press.
- MORTON, K.W. and PARROTT, A.K. 1980 Generalized Galerkin methods for first-order hyperbolic equations, *J. Comput. Phys.*, **36**, pp. 249-270.
- NAKAZAWA, S. 1982 Finite element analysis applied to polymer processing, PhD Thesis, University of Swansea.
- OSHER, S. 1980 Approximation par éléments finis avec décentrage pour des lois de conservation hyperboliques non linéaires multi-dimensionnelles, *C.R. Acad. Sc. Paris*, t.**290**, Série A., pp. 819-821.
- OSHER, S. 1981a Approximation par éléments finis avec décentrage de problèmes de perturbations singulières quasi linéaires et multi dimensionnels, *C.R. Acad. Sc. Paris*, t.**292**, Série I, pp. 99-101.
- OSHER, S. 1981b Nonlinear singular perturbation problems and one sided difference schemes, *SIAM J. Num. Analysis*, **18**, pp. 129-144.
- OSHER, S. 1982 Numerical solution of singular perturbation problems and hyperbolic systems of conservation laws. North Holland Math. Studies # 47, (Axelsson et al. Eds.).
- RAITHBY, G.D. 1976 Skew upstream differencing schemes for problems involving fluid flow, *Comput. Meth. Appl. Mech. Eng.*, **9**, pp. 153-164.
- RAYMOND, W.H. and GARDER, A. 1976 Selective damping in a Galerkin method for solving wave problems with variable grids, *Mon. Weath. Rev.*, **104**, pp. 1583-1590.

- SPRADLEY, L.W., STALNAKER, J.F. and RATLIFF, A.W. 1980 Computation of three-dimensional viscous flows with the Navier-Stokes equations, AIAA-80-1348, AIAA 13th Fluid and Plasma Dynamics Conference, Snowmass, Colorado, USA.
- STRANG, G. and FIX, G. 1973 An analysis of the finite element method, Englewood Cliffs, New Jersey, Prentice Hall.
- TEZDUYAR, T.E. and HUGHES, T.J.R. 1982 Development of time-accurate finite element techniques for first-order hyperbolic systems with particular emphasis on the compressible Euler equations, Report prepared under NASA-Ames University Consortium Interchange No. NCA2-OR745-104.
- THOMASSET, F. 1981 Implementation of finite element methods for Navier-Stokes equations, New York, Springer-Verlag.
- WAHLBIN, L.B. 1974 A dissipative Galerkin method for the numerical solution of first order hyperbolic equations. In Mathematical Aspects of Finite Elements in Partial Differential Equations, (C. de Boor, Ed.), New York, Academic Press.
- WILKINS, M.L. 1969 Calculation of elastic-plastic flows, Lawrence Livermore Laboratory, Rept. UCRL-7322, Rev. 1, Livermore, California, USA.
- ZIENKIEWICZ, O.C. 1977 The finite element method, London, McGraw-Hill.