

**A NEW FINITE ELEMENT FORMULATION FOR COMPUTATIONAL
FLUID DYNAMICS: IV.
A DISCONTINUITY-CAPTURING OPERATOR FOR MULTIDIMENSIONAL
ADVECTIVE-DIFFUSIVE SYSTEMS***

Thomas J.R. HUGHES and Michel MALLET

Division of Applied Mechanics, Durand Building, Stanford University, Stanford, CA 94305, U.S.A.

Received 5 December 1985

A discontinuity-capturing operator is developed for the 'streamline' formulation of advective-diffusive systems extending previous work on the *scalar* advection-diffusion equation. The operator provides a mechanism for exerting control over strong gradients in the discrete solution which appear, for example, in boundary and interior layers.

1. Introduction

This is the fourth in a series of papers which present a new finite element formulation for computational fluid dynamics. Throughout this paper we refer to Hughes, Franca and Mallet [2] as Part I; Hughes, Mallet and Mizukami [5] as Part II; and Hughes and Mallet [3] as Part III. In Part I we demonstrated the advantages of employing physical 'entropy variables' as a basis for discretization of the compressible Navier–Stokes equations. In addition to providing a correct global statement of entropy production in terms of the discrete solution, entropy variables symmetrize all coefficient matrices appearing in the equations. Thus the entropy-variables form of the compressible Navier–Stokes equations falls within the class of *symmetric incompletely parabolic systems*. A semidiscrete generalization of the streamline concept ('SUPG') to multidimensional advective-diffusive systems of symmetric type was presented in Part III and precise mathematical error estimates were shown to hold for the linear case covering the complete spectrum of advective-diffusive behavior.

As discussed in Part II, SUPG is an excellent method for problems with smooth solutions, but typically introduces localized oscillations about sharp internal and boundary layers. To improve upon the situation, in Part II we added a discontinuity-capturing term to the formulation. This term provides additional control over gradients in the discrete solution and considerably increases the robustness of the methodology. Because the discontinuity-capturing term is a function of the discrete-solution gradient, the numerical method is nonlinear even when the original equation is linear. (SUPG by itself is a linear method.) In Part II the developments were restricted to the *scalar* advection-diffusion equation.

In this paper we generalize the discontinuity-capturing operator to multidimensional

*This research was sponsored by the NASA Langley Research Center under Grant NASA-NAG-1-361.

systems of the type considered in Parts I and III. The discontinuity-capturing operator is defined in terms of certain other operators which are described in detail in Section 2. The exposition makes heavy use of notations and ideas presented in Part III. This has enabled us to keep this work very concise, but requires that the reader be familiar with Part III before embarking upon a serious study of the material presented herein. It may be noted that, when applied to the compressible Euler equations, the present formulation has produced very good results [4]. Conclusions are drawn in Section 3.

2. Analysis

The basic idea is to employ the weak formulation for systems developed in Part III and add a discontinuity-capturing term to it, which generalizes the scalar case presented in Part II. The desired weak form is

$$\begin{aligned} & \int_{\Omega} (\mathbf{W}^h \cdot \tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h + \nabla \mathbf{W}^h \cdot (\tilde{\mathbf{K}} \nabla \mathbf{V}^h)) \, d\Omega \\ & + \sum_{e=1}^{n_{el}} \int_{\Omega^e} (\tilde{\tau}_1 \tilde{\mathbf{A}} \cdot \nabla \mathbf{W}^h + \boxed{\tilde{\tau}_2 \tilde{\mathbf{A}}_{\parallel} \cdot \nabla \mathbf{W}^h}) \cdot (\tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h - \nabla \cdot (\tilde{\mathbf{K}} \nabla \mathbf{V}^h) - \mathcal{F} + \mathbf{A}_0 \mathbf{V}_{,t}^h) \, d\Omega \\ & = \int_{\Omega} \mathbf{W}^h \cdot (\mathcal{F} - \mathbf{A}_0 \mathbf{V}_{,t}^h) \, d\Omega + \text{natural boundary condition terms} . \end{aligned} \quad (1)$$

The boxed term represents the only additional term when compared with the semidiscrete streamline formulation of multidimensional advective-diffusive systems written in symmetric form with Riemannian metric (cf. Part III, equation (128)). The matrix $\tilde{\mathbf{A}}_{\parallel}$ is the *projection of $\tilde{\mathbf{A}}$ onto the direction $\nabla \mathbf{V}^h$ with respect to the Riemannian metric \mathbf{A}_0* , and the matrix $\tilde{\tau}_2$ represents the *intrinsic element time scales associated with gradient information*. Precise mathematical definitions of $\tilde{\mathbf{A}}_{\parallel}$ and $\tilde{\tau}_2$ will be developed in the remainder of this paper. For all remaining notations, the reader will need to consult Part III.

2.1. The matrix $\tilde{\mathbf{A}}_{\parallel}$

The situation is somewhat complicated by the presence of a Riemannian metric. To simplify the derivation, let us temporarily assume that \mathbf{A}_0 and $\tilde{\mathbf{A}}$ are constant and employ the transformation introduced in Part III to convert to Euclidian metric, namely

$$\mathbf{X}^h = \mathbf{L}^t \mathbf{V}^h , \quad (2)$$

$$\hat{\mathbf{A}}_i = \mathbf{L}^{-1} \tilde{\mathbf{A}}_i \mathbf{L}^{-t} , \quad (3)$$

where \mathbf{L} is a constant matrix which factorizes \mathbf{A}_0 , that is

$$\mathbf{A}_0 = \mathbf{L} \mathbf{L}^t , \quad (4)$$

but is otherwise arbitrary. Note that:

$$\tilde{\hat{A}} = \begin{pmatrix} \tilde{\hat{A}}_1 \\ \vdots \\ \tilde{\hat{A}}_d \end{pmatrix} \quad (m \cdot d \times m), \quad \hat{A} = \begin{pmatrix} \hat{A}_1 \\ \vdots \\ \hat{A}_d \end{pmatrix} \quad (m \cdot d \times m), \quad (5)$$

in which the submatrices are symmetric. Consider the definition:

$$\hat{A}_{\parallel}^t = \hat{A}^t \hat{\Pi}, \quad (6)$$

where

$$\hat{\Pi} = \frac{\nabla X^h (\nabla X^h)^t}{(\nabla X^h)^t \nabla X^h}. \quad (7)$$

The following properties are easy to establish:

- (i) Range $(\hat{\Pi}) = \text{span}(\nabla X^h)$;
- (ii) $\hat{\Pi}^2 = \hat{\Pi}$;
- (iii) $\hat{\Pi} = \hat{\Pi}^t$;
- (iv) for all $(m \cdot d)$ -dimensional vectors u which are orthogonal to ∇X^h (i.e., $u \cdot \nabla X^h = u^t \nabla X^h = 0$),

$$\hat{A}_{\parallel} \cdot u = \hat{A}_{\parallel}^t u = 0;$$

$$(v) \hat{A}_{\parallel} \cdot \nabla X^h = \hat{A} \cdot \nabla X^h.$$

By properties (i)–(iii) it may be seen that $\hat{\Pi} : \mathbb{R}^{m \cdot d} \rightarrow \mathbb{R}^{m \cdot d}$ is the *orthogonal projection* onto $\text{span}(\nabla X^h)$, where $\nabla X^h \in \mathbb{R}^{m \cdot d}$. By properties (iv) and (v) we see that the null space of \hat{A}_{\parallel} is $(\text{span}(\nabla X^h))^{\perp}$ and that the action of \hat{A}_{\parallel} on $\text{span}(\nabla X^h)$ is identical to that of \hat{A} .

The definition of $\tilde{\hat{A}}_{\parallel}$ may now be obtained by transformation, viz.

$$\tilde{\hat{A}}_{\parallel}^t = \tilde{\hat{A}}^t \tilde{\Pi}, \quad (8)$$

where

$$\tilde{\Pi} = \frac{\nabla V^h (\nabla V^h)^t \text{diag}(A_0, \dots, A_0)}{(\nabla V^h)^t \text{diag}(A_0, \dots, A_0) \nabla V^h}, \quad (9)$$

and $\text{diag}(A_0, \dots, A_0)$ stands for d copies of A_0 along the diagonal of an $(m \cdot d) \times (m \cdot d)$ matrix which is otherwise zero. The present definition satisfies the appropriate analogs of conditions (i)–(v) with respect to the inner product induced by $\text{diag}(A_0, \dots, A_0)$ on $\mathbb{R}^{m \cdot d}$, in particular

$$\tilde{\hat{A}}_{\parallel} \cdot \nabla V^h = \tilde{\hat{A}} \cdot \nabla V^h \quad (10)$$

and

$$\tilde{\hat{A}}_{\parallel} \cdot w = 0, \quad (11)$$

for all w orthogonal to ∇V^h with respect to $\text{diag}(A_0, \dots, A_0)$. Clearly, $\tilde{\hat{A}}_{\parallel}$ has rank 1. The role

played by the \tilde{A}_{\parallel} -term in (1) may be seen by examining the terms

$$(\tilde{\tau}_1 \tilde{A} \cdot \nabla W^h + \tilde{\tau}_2 \tilde{A}_{\parallel} \cdot \nabla W^h) \cdot \tilde{A} \cdot \nabla V^h = \nabla W^h \cdot \underbrace{\tilde{A} \tilde{\tau}_1 \tilde{A}^t}_{\text{generalized streamline operator}} \nabla V^h + \nabla W^h \cdot \underbrace{\tilde{A}_{\parallel} \tilde{\tau}_2 \tilde{A}_{\parallel}^t}_{\text{discontinuity-capturing operator}} \nabla V^h.$$

Note that we have used (10) in deriving the last term of (12). The discontinuity-capturing operator engenders extra control over the gradient in the discrete solution. By virtue of the fact that \tilde{A}_{\parallel} is a function of ∇V^h , (1) is a *nonlinear* method even when the original equation system is linear. It is clear from (12) that the maximal rank of the discontinuity-capturing operator is 1.

REMARK 2.1. The intent of the discontinuity-capturing term is to bring about an effect similar to that of Davis' 'rotated difference' scheme in which the difference stencil aligns itself perpendicular to the gradients in the discrete solution [1].

2.2. The matrix $\tilde{\tau}_{\parallel}$

As a prelude to the definition of $\tilde{\tau}_2$, we first consider a related matrix, $\tilde{\tau}_{\parallel}$. Roughly speaking, $\tilde{\tau}_{\parallel}$ is computed from \tilde{A}_{\parallel} in the same way that $\tilde{\tau}$ was computed from \tilde{A} in Part III. Let

$$(\tilde{B}_{\parallel})_i = (\partial \xi_i / \partial x_j) (\tilde{A}_{\parallel})_j, \quad (13)$$

where the $(\tilde{A}_{\parallel})_j$ are the submatrices of \tilde{A}_{\parallel}^t , viz.

$$\tilde{A}_{\parallel}^t = [(\tilde{A}_{\parallel})_1, \dots, (\tilde{A}_{\parallel})_d]. \quad (14)$$

Note that the $(\tilde{A}_{\parallel})_j$ are generally not symmetric and that

$$\tilde{B}_{\parallel}^t = [(\tilde{B}_{\parallel})_1, \dots, (\tilde{B}_{\parallel})_d], \quad (15)$$

like \tilde{A}_{\parallel}^t , has rank 1. Let p be an *even positive integer*. The eigenproblem

$$\left(\sum_{i=1}^d ((\tilde{B}_i A_0^{-1} \tilde{B}_i^t A_0^{-1})^{p/2-1} (\tilde{B}_i A_0^{-1} \tilde{B}_i^t)) - \mu_{\parallel}^p A_0 \right) \varphi_{\parallel} = 0 \quad (16)$$

provides essential information for defining $\tilde{\tau}_{\parallel}$. We shall confine our attention to the case $p = 2$, for which

$$(\tilde{B}_{\parallel}^t \text{diag}(A_0^{-1}, \dots, A_0^{-1}) \tilde{B}_{\parallel} - \mu_{\parallel}^2 A_0) \varphi_{\parallel} = 0. \quad (17)$$

This eigenproblem can be solved in closed form. The following result is useful:

$$(\tilde{B}_{\parallel})_i = \frac{\partial \xi_i}{\partial x_j} (\tilde{A}_{\parallel})_j = \frac{\partial \xi_i}{\partial x_j} \frac{\tilde{A} \cdot \nabla V^h (V_{,j}^h)^t A_0}{(\nabla V^h)^t \text{diag}(A_0, \dots, A_0) \nabla V^h}. \quad (18)$$

Introducing

$$\mathbf{DV}^h = \begin{pmatrix} \mathbf{V}^h_{\cdot j} \frac{\partial \xi_1}{\partial x_j} \\ \vdots \\ \mathbf{V}^h_{\cdot j} \frac{\partial \xi_d}{\partial x_j} \end{pmatrix} \quad (19)$$

enables us to construct the index-free version of (18), namely

$$\tilde{\mathbf{B}}_{\parallel}^t = \frac{\tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h (\mathbf{DV}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0)}{(\nabla \mathbf{V}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0) \nabla \mathbf{V}^h}. \quad (20)$$

Thus

$$\begin{aligned} \tilde{\mathbf{B}}_{\parallel}^t \text{diag}(\mathbf{A}_0^{-1}, \dots, \mathbf{A}_0^{-1}) \tilde{\mathbf{B}}_{\parallel} \\ = \frac{(\mathbf{DV}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0) \mathbf{DV}^h \tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h (\tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h)^t}{((\nabla \mathbf{V}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0) \nabla \mathbf{V}^h)^2}. \end{aligned} \quad (21)$$

From this expression it may be concluded that the eigenvector and eigenvalue of (17) are, respectively,

$$\boldsymbol{\varphi}_{\parallel} = \frac{\mathbf{A}_0^{-1} \tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h}{((\tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h)^t \mathbf{A}_0^{-1} \tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h)^{1/2}}, \quad (22)$$

$$\mu_{\parallel}^2 = \frac{((\tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h)^t \mathbf{A}_0^{-1} \tilde{\mathbf{A}} \cdot \nabla \mathbf{V}^h) ((\mathbf{DV}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0) \mathbf{DV}^h)}{((\nabla \mathbf{V}^h)^t \text{diag}(\mathbf{A}_0, \dots, \mathbf{A}_0) \nabla \mathbf{V}^h)^2}. \quad (23)$$

The definition of $\tilde{\tau}_{\parallel}$ may now be given:

$$\tilde{\tau}_{\parallel} = \boldsymbol{\varphi}_{\parallel} \tau_{\parallel} \boldsymbol{\varphi}_{\parallel}^t, \quad (24)$$

where

$$\tau_{\parallel} = \tilde{\xi}(\alpha_{\parallel}) / \mu_{\parallel}, \quad (25)$$

$$\alpha_{\parallel} = \mu_{\parallel} / \sigma_{\parallel}, \quad (26)$$

$$\sigma_{\parallel} = \frac{1}{d} \begin{pmatrix} \boldsymbol{\varphi}_{\parallel} \\ \vdots \\ \boldsymbol{\varphi}_{\parallel} \end{pmatrix}^t \tilde{\mathbf{K}}_{\xi} \begin{pmatrix} \boldsymbol{\varphi}_{\parallel} \\ \vdots \\ \boldsymbol{\varphi}_{\parallel} \end{pmatrix}, \quad (27)$$

in which

$$\tilde{\xi}(\alpha_{\parallel}) = \coth(\alpha_{\parallel}) - \alpha_{\parallel}^{-1} \text{ (cf. Part III, (50), and discussion thereafter)}, \quad (28)$$

and \tilde{K}_ξ is defined by Part III, equations (135), (136).

REMARK 2.2. The eigenvalue may also be written in the alternative form:

$$\mu_\parallel = |\lambda_\parallel| / \frac{1}{2} h_\parallel, \quad (29)$$

where

$$|\lambda_\parallel| = \sqrt{\frac{(\tilde{A} \cdot \nabla V^h)^t A_0^{-1} \tilde{A} \cdot \nabla V^h}{(\nabla V^h)^t \text{diag}(A_0, \dots, A_0) \nabla V^h}}, \quad (30)$$

$$\frac{1}{2} h_\parallel = \sqrt{\frac{(\nabla V^h)^t \text{diag}(A_0, \dots, A_0) \nabla V^h}{(D V^h)^t \text{diag}(A_0, \dots, A_0) D V^h}}. \quad (31)$$

With these expressions we can write

$$\tau_\parallel = \frac{1}{2} h_\parallel \tilde{\xi}(\alpha_\parallel) / |\lambda_\parallel|, \quad (32)$$

$$\alpha_\parallel = \frac{1}{2} h_\parallel |\lambda_\parallel| / \kappa_\parallel, \quad (33)$$

$$\kappa_\parallel = (\frac{1}{2} h_\parallel)^2 \sigma_\parallel, \quad (34)$$

in which σ_\parallel is defined by (27).

REMARK 2.3. The situation is somewhat simplified if we assume that the element geometry is sufficiently regular to permit replacement of $\partial \xi_i / \partial x_j$ by $(2/h) \delta_{ij}$ (see Part III, discussion after (83) and Fig. 1). In this case

$$D V^h = (2/h) \nabla V^h, \quad (35)$$

$$\mu_\parallel = |\lambda_\parallel| / \frac{1}{2} h \quad (36)$$

$$\kappa_\parallel = (\frac{1}{2} h)^2 \sigma_\parallel = \frac{1}{d} \begin{pmatrix} \varphi_\parallel \\ \vdots \\ \varphi_\parallel \end{pmatrix}^t \tilde{K} \begin{pmatrix} \varphi_\parallel \\ \vdots \\ \varphi_\parallel \end{pmatrix}. \quad (37)$$

2.3. The matrices $\tilde{\tau}_1$ and $\tilde{\tau}_2$

Corresponding to the scalar case presented in Part II we may consider the following two possibilities:

$$(i) \quad \tilde{\tau}_1 = \tilde{\tau} \quad (\tilde{\tau} \text{ is defined in Part III}), \quad \tilde{\tau}_2 = \tilde{\tau}_\parallel. \quad (38)$$

As described in Part II, this choice causes a doubling effect under certain circumstances. Numerical results also appear to be a bit too diffusive (see Part II). These deficiencies are mitigated by the following alternative.

(ii) In this case we endeavor to remove the component of the generalized streamline

operator which acts in the “direction” of the discontinuity-capturing operator. Specifically, we define a scalar τ by way of the expression

$$(\nabla V^h)^t \tilde{A}_{\parallel} \varphi_{\parallel} \tau \varphi_{\parallel}^t \tilde{A}_{\parallel}^t \nabla V^h = (\nabla V^h)^t \tilde{A} \tilde{\tau} \tilde{A}^t \nabla V^h, \quad (39)$$

from which it follows that

$$\tau = \frac{(\tilde{A} \cdot \nabla V^h)^t \tilde{\tau} \tilde{A} \cdot \nabla V^h}{(\varphi_{\parallel} \cdot \tilde{A}_{\parallel} \cdot \nabla V^h)^2} = \frac{(A_0 \varphi_{\parallel})^t \tilde{\tau} A_0 \varphi_{\parallel}}{(\varphi_{\parallel} \cdot A_0 \varphi_{\parallel})^2}. \quad (40)$$

Then we set

$$\tilde{\tau}_1 = \tilde{\tau}, \quad \tilde{\tau}_2 = \varphi_{\parallel} \max(0, \tau_{\parallel} - \tau) \varphi_{\parallel}^t. \quad (41)$$

These expressions have proven successful in practical calculations with the compressible Euler equations which will be presented in future works in this series (see also [4] for preliminary results).

3. Conclusions

In this paper we have presented a discontinuity-capturing operator applicable to the semidiscrete SUPG formulation of multidimensional advective-diffusive *systems* developed in Part III of this series of papers. The present work generalizes Part II, which was restricted to *scalar* advective-diffusive equations. The discontinuity-capturing operator provides a means within the SUPG framework of exercising control over gradients in the discrete solution. Application of the methodology described herein (and described in Parts I and III) to the compressible Euler equations has proven quite successful (see [4] for preliminary results, and subsequent papers in this series).

Acknowledgements

We wish to thank Leo Franca, Claes Johnson, Akira Mizukami and Farzin Shakib for helpful comments. Michel Mallet was supported by an IBM Graduate Fellowship.

References

- [1] S.F. Davis, A rotationally biased upwind difference scheme for the Euler equations, *J. Comput. Phys.* 56 (1984) 65–92.
- [2] T.J.R. Hughes, L.P. Franca and M. Mallet, A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier–Stokes equations and the second law of thermodynamics, *Comput. Meth. Appl. Mech. Engrg.* 54 (1986) 223–234.

- [3] T.J.R. Hughes and M. Mallet, A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems, *Comput. Meths. Appl. Mech. Engrg.* 58 (1986) 305–328 (this issue).
- [4] T.J.R. Hughes, M. Mallet and L.P. Franca, Entropy-stable finite element methods for compressible fluids: application to high Mach number flows with shocks, in: P. Bergan, K.J. Bathe and W. Wunderlich, eds., *Finite Element Methods for Nonlinear Problems* (Springer, Berlin, 1986) 761–773.
- [5] T.J.R. Hughes, M. Mallet and A. Mizukami, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, *Comput. Meths. Appl. Mech. Engrg.* 54 (1986) 341–355.