INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 4, 1043–1063 (1984)

THE SOLUTION OF NON-LINEAR HYPERBOLIC EQUATION SYSTEMS BY THE FINITE ELEMENT METHOD

R. LÖHNER, K. MORGAN AND O. C. ZIENKIEWICZ

Department of Civil Engineering, University of Wales, Swansea SA28PP, U.K.

SUMMARY

The difficulties experienced in the treatment of hyperbolic systems of equations by the finite element method (or other) spatial discretization procedures are well known. In this paper a temporal discretization precedes the spatial one which in principle is considered along the characteristics to achieve a self adjoint form. By a suitable expansion, the original co-ordinates are preserved and combined with the use of a standard Galerkin process to achieve an accurate discretization. It is shown that the process is equivalent to the Taylor–Galerkin methods of Donea.[17]

Several examples illustrate the accuracy and efficiency attainable in such problems as transport, shallow water equations, transonic flow etc.

1. INTRODUCTION

The use of upwind differencing or more generally of Petrov–Galerkin methods in steady-state convection–diffusion problems is now commonplace and generally accepted, not only to avoid spurious oscillations but in general to improve the accuracy of computation.[1–3] Special forms of such Petrov–Galerkin processes using so-called streamline or balancing diffusion[4,5] are particularly effective, and recent mathematical studies have shown that standard Galerkin processes are not optimal in minimizing the error for such non-self-adjoint problems.[6,7]

For transient phenomena of convection–diffusion or for mixed representation of wave problems it would appear that similar procedures should be used. Here a 'direct transplant' of concepts led to many questions not yet answered satisfactorily. When statically optimal balancing diffusion was used the results were invariably overdamped and semi-empirical factors (such as $1/\sqrt{15}$) have been used by some[8,9] to counterbalance this effect. Further much debate ensued on such questions as the use of Petrov–Galerkin weighting in the context of mass matrices.[10,11]

The problems raised above can be side-stepped by using alternative routes. In the first the differential equations are recast in terms of moving co-ordinates (or along the characteristics) thus making the problem self-adjoint and hence best treated by standard Galerkin spatial discretization. Such procedures have been used with an interpolation after each time-step on the original (fixed) co-ordinates,[12–15] but appear to be very uneconomical and very difficult to extend to systems of equations. More recently a more direct way was introduced, not requiring updating and interpolation.[16] The second path has been proposed by Donea[17] in which a Taylor expansion in

0271–2091/84/11 1043–21$02.10

Received 10 September 1983
Revised 19 December 1983

time precedes the Galerkin space discretization. This procedure leads to excellent results but the Galerkin discretization is not justified by self-adjointness.

In the present paper the authors show that:

(a) the use of moving co-ordinates and re-interpolation on to the original mesh (by a Taylor expansion) is equivalent to the Taylor–Galerkin process in the case of convection–diffusion problems
(b) for *systems* of differential equations both procedures are again equivalent
(c) in multidimensional problems the methods essentially produce a streamline balancing diffusion of the correct kind without recourse to semi-empirical coefficients.

Several examples are included showing that physically correct, not overdamped, solutions are obtained for transport problems, shallow water equations, Euler equations, etc. although in the presence of shock waves additional (gradient proportional) diffusion needs to be added to avoid erroneous results.

2. THE TAYLOR–GALERKIN METHOD

A Taylor–Galerkin method for the solution of hyperbolic problems was recently proposed by Donea.[17] The method may be described by considering a simple scalar conservation law in one dimension expressed as

$$\frac{\partial u}{\partial t}+\frac{\partial F}{\partial x}=0 \tag{1}$$

or, alternatively, in the form

$$\frac{\partial u}{\partial t}+A\frac{\partial u}{\partial x}=0 \tag{2}$$

where

$$A=A(u)=\frac{\partial F}{\partial u} \tag{3}$$

Making a Taylor series expansion in time about $t=t^n$ it is possible to write:

$$u^{n+1}=u^n+\Delta t\frac{\partial u}{\partial t}\bigg|^n+\frac{\Delta t^2}{2}\frac{\partial^2 u}{\partial t^2}\bigg|^n \tag{4}$$

correct to second order, where $t^{n+1}=t^n+\Delta t$. The time derivatives appearing in equation (4) can be expressed in terms of space derivatives by using equations (1) and (2).

From (1):

$$\frac{\partial u}{\partial t}\bigg|^n=-\frac{\partial F}{\partial x}\bigg|^n=-A\frac{\partial u}{\partial x}\bigg|^n \tag{5}$$

and differentiating (1) gives

$$\frac{\partial^2 u}{\partial t^2}\bigg|^n=-\frac{\partial}{\partial t}\left(\frac{\partial F}{\partial x}\right)\bigg|^n=-\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial t}\right)\bigg|^n \tag{6}$$

This expression may be further reduced by using (3), so that

$$\frac{\partial^2 u}{\partial t^2}\bigg|^n=-\frac{\partial}{\partial x}\left(A\frac{\partial u}{\partial t}\right)\bigg|^n=\frac{\partial}{\partial x}\left(A\frac{\partial F}{\partial x}\right)=\frac{\partial}{\partial x}\left(A^2\frac{\partial u}{\partial x}\right)\bigg|^n \tag{7}$$

Introducing (5) and (7) in (4) leads to the generalized second-order (Euler) time-stepping scheme:

$$u^{n+1} = u^n - \Delta t \frac{\partial F}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\frac{\partial}{\partial x}\left(A\frac{\partial F}{\partial x}\right)\bigg|^n$$

$$= u^n - \Delta t A\frac{\partial u}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\frac{\partial}{\partial x}\left(A^2\frac{\partial u}{\partial x}\right)\bigg|^n \tag{8}$$

By using an expansion correct to third order in place of equation (4), Donea[17] has been able to produce a third order equivalent of equation (8) and he has also examined the behaviour of other time-stepping schemes, obtained by simply taking the Taylor expansion at other values of t. The second order Euler–Taylor–Galerkin method, which is of interest here, results from applying the Galerkin finite element method to equation (8). The resulting algorithm, which may be regarded as a finite element implementation of the one-step Lax–Wendroff method,[18] has been shown to possess excellent phase-accuracy and minimum numerical damping when applied to the scalar advection equation.

3. A CHARACTERISTICS BASED METHOD LEADING TO THE EULER–TAYLOR GALERKIN SCHEME

The solution of (2) may be obtained by the method of characteristics. The characteristics are the lines satisfying

$$\frac{\mathrm{d}x}{\mathrm{d}t} = A \tag{9}$$

and along these lines equation (2) is reduced to the characteristics relation:

$$\frac{\partial u}{\partial t} = 0 \tag{10}$$

or

$$u = \text{const.} \tag{11}$$

Thus starting from a given initial distribution and a given spatial discretization, equation (11) may be used to construct the solution at later times with a continuous mesh updating via equation (9). This computationally undesirable feature of mesh updating can be avoided as follows. We introduce here a characteristic co-ordinate ξ, such that ξ is a constant along a characteristic, and consider the situation over a time interval $t^n \leqslant t \leqslant t^{n+1}$. Figure 1 shows the characteristic ξ which passes through the point P with co-ordinates (x_ρ, t^n) and the point Q with coordinates (x, t^{n+1}).

From equation (11)

$$u(x, t^{n+1}) = u(x_\rho, t^n) \tag{12}$$

and, using (9)

$$x_\rho = x - \Delta t \bar{A} \tag{13}$$

$\bar{A}$ being a suitable average value of A over the characteristic curve between the points P and Q. Using (13) in (12) and expanding as a Taylor series results in

$$u^{n+1} = u^n - \Delta t \bar{A}\frac{\partial u}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\bar{A}^2\frac{\partial^2 u}{\partial x^2}\bigg|^n \tag{14}$$

with

$$u^{n+1} = u(x, t^{n+1})$$

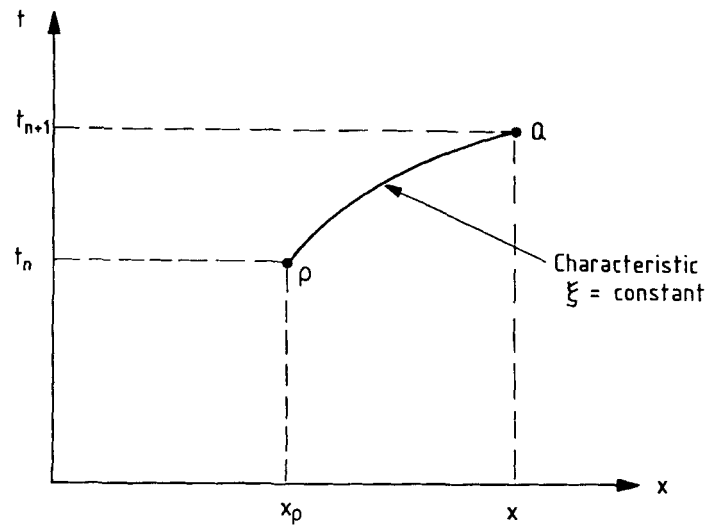


Figure 1. Path of a characteristic over the time interval (t_n, t_{n+1})

and

$$u^n = u(x, t^n)$$

correct to second order. Various approximations will result from different choices of $\bar{A}$, but here consideration is restricted to the case when $\bar{A}$ is defined to be the value of A at the midpoint of the spatio-temporal interval, i.e.

$$\bar{A} = A^n + \frac{\Delta t}{2}\left(\frac{\partial A}{\partial t}\bigg|^n - A\frac{\partial A}{\partial x}\bigg|^n\right) \tag{15}$$

correct to first order. Noting that

$$\frac{\partial A}{\partial t} = \frac{\partial A}{\partial u}\frac{\partial u}{\partial t} = -A\frac{\partial A}{\partial u}\frac{\partial u}{\partial x} = -A\frac{\partial A}{\partial x} \tag{16}$$

it follows that:

$$\bar{A} = A^n - \Delta t A\frac{\partial A}{\partial x}\bigg|^n \tag{17}$$

Combining (17) and (14) we obtain:

$$u^{n+1} = u^n - \Delta t A\frac{\partial u}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\left\{A^2\frac{\partial^2 u}{\partial x^2}\bigg|^n + 2A\frac{\partial A}{\partial x}\frac{\partial u}{\partial x}\bigg|^n\right\} \tag{18}$$

or, rearranging,

$$\begin{aligned} u^{n+1} &= u^n - \Delta t\frac{\partial F}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\frac{\partial}{\partial x}\left(A\frac{\partial F}{\partial x}\right)\bigg|^n \\ &= u^n - \Delta t A\frac{\partial u}{\partial x}\bigg|^n + \frac{\Delta t^2}{2}\frac{\partial}{\partial x}\left(A^2\frac{\partial u}{\partial x}\right)\bigg|^n \end{aligned} \tag{19}$$

It can be observed that (19) is just the second-order generalized Euler time-stepping scheme (8). The third order schemes of Donea[17] can be obtained similarly, although the analysis is now much more complicated. This section has established the characteristics nature of the Taylor–Galerkin schemes.

Either of the schemes of equations (8), (19) can now be used to apply a discretization procedure in the spatial co-ordinates in some 'weak' integral form,[19,20] essentially a reverse order for that usually practised (in which the spatial discretization precedes the temporal one). The question immediately arises as to the nature of spatial discretization to be followed, i.e. whether Petrov or Bubnov–Galerkin procedures should be applied.

If the nature of (19) and its derivation is examined, we observe that it represents simply an adjoint equation (namely equation (10)), for which the Bubnov–Galerkin process is *optimal*—thus justifying the Taylor–Galerkin algorithm.

It is readily seen that both algorithms (8), (19) introduce a second spatial derivative in the last term—and hence effectively a balancing diffusion of the type associated with upwind procedures. Indeed it will be shown later that this diffusion is streamline oriented in multi-dimensional problems and hence of the form similar to that given in References 4 and 5. However now there is no arbitrariness with regards to its magnitude.

In Appendix I we show how the presence of real diffusion in (1) can be treated with the same approach.

4. APPLICATION TO SYSTEMS OF HYPERBOLIC EQUATIONS IN 1-D

A system of coupled hyperbolic equations may be treated similarly. Writing such a system, with source term, as

$$\frac{\partial \mathbf{u}}{\partial t}+\frac{\partial \mathbf{F}}{\partial x}=\mathbf{S}(\mathbf{u}) \tag{20}$$

or, alternatively, as

$$\frac{\partial \mathbf{u}}{\partial t}+\mathbf{A}\frac{\partial \mathbf{u}}{\partial x}=\mathbf{S}(\mathbf{u}) \tag{21}$$

where

$$\mathbf{A}=\frac{\partial \mathbf{F}}{\partial \mathbf{u}} \tag{22}$$

and proceeding as in Section 2, a direct Taylor-series expansion leads to:

$$\begin{aligned}\mathbf{u}^{n+1} &= \mathbf{u}^n+\Delta t\left(\mathbf{S}-\frac{\partial \mathbf{F}}{\partial x}\right)\Big|^n+\frac{\Delta t^2}{2}\left[\mathbf{C}\left(\mathbf{S}-\frac{\partial \mathbf{F}}{\partial x}\right)-\frac{\partial}{\partial x}\left[\mathbf{A}\left(\mathbf{S}-\frac{\partial \mathbf{F}}{\partial x}\right)\right]\right]\Big|^n \\ &= \mathbf{u}^n+\Delta t\left(\mathbf{S}-\mathbf{A}\frac{\partial \mathbf{u}}{\partial x}\right)\Big|^n+\frac{\Delta t^2}{2}\left[\mathbf{C}\left(\mathbf{S}-\mathbf{A}\frac{\partial \mathbf{u}}{\partial x}\right)-\frac{\partial}{\partial x}\left[\mathbf{A}\left[\mathbf{S}-\mathbf{A}\frac{\partial \mathbf{u}}{\partial x}\right)\right]\right]\Big|^n\end{aligned} \tag{23}$$

where

$$\mathbf{C}=\frac{\partial \mathbf{S}}{\partial \mathbf{u}} \tag{24}$$

A characteristics-based interpretation of this relation is again possible, its derivation proceeding exactly as in Section 3 (see Appendix II).

5. SPATIAL DISCRETIZATION

The spatial discretization of the recurrence relations obtained in (19) and (23) can now proceed to obtain an algorithm for a numerical solution. With the use of standard (or *Bubnov*–) Galerkin processes (justified in Section 3) we can proceed in the usual finite element manner discussed in detail in various texts.[19,20] In the 1-D case two-noded linear elements were used with shape functions $N_j(x)$ associated with node j. The approximate solution, $\mathbf{u}^n$, at time t^n is expressed as:

$$\mathbf{u}^n = \sum_j \hat{\mathbf{u}}_j^n N_j \tag{25}$$

where $\hat{\mathbf{u}}_j^n$ is the approximation to the value of $\mathbf{u}$ at node $\mathbf{j}$ and time t^n. Once the interpolation of $\mathbf{u}$ is chosen, the discretization can proceed directly. The integration of the occurring terms can be done either exactly or numerically. Numerical experiments show that it is convenient (just as in many other problems[19]) to use low order ('reduced') integration rules, for which wave velocities are most correctly given. An alternative which we have used here is simply to employ the identical interpolation rule to the non-linear terms $\mathbf{F}, \mathbf{A}, \mathbf{S}$ and $\mathbf{G}$, i.e.

$$\mathbf{F}^n = \sum_j \hat{\mathbf{F}}_j^n N_j; \quad \mathbf{A}^n = \sum_j \hat{\mathbf{A}}_j^n N_j \tag{26}$$

$$\mathbf{S}^n = \sum_j \hat{\mathbf{S}}_j^n N_j; \quad \mathbf{C}^n = \sum_j \hat{\mathbf{C}}_j^n N_j$$

Application of the Galerkin method produces the equation system:

$$\mathbf{M}\Delta\hat{\mathbf{U}}^n = \hat{\mathbf{f}}^n \tag{27}$$

where $\mathbf{M}$ is the standard (symmetric) mass matrix, $\Delta\hat{\mathbf{U}}^n = \hat{\mathbf{U}}^{n+1} - \hat{\mathbf{U}}^n$ and $\hat{\mathbf{U}}, \hat{\mathbf{f}}$ are defined by

$$\hat{\mathbf{U}}^T = [\hat{\mathbf{U}}_1, \hat{\mathbf{U}}_2, \hat{\mathbf{U}}_3, \ldots]$$
$$\hat{\mathbf{f}}^T = [\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \hat{\mathbf{f}}_3, \ldots] \tag{28}$$

$$\hat{\mathbf{f}}_i = \Delta t \int \left(\mathbf{S} - \frac{\partial \mathbf{F}}{\partial x}\right) N_i \mathrm{d}x + \frac{\Delta t^2}{2} \int \mathbf{C}\left(\mathbf{S} - \frac{\partial \mathbf{F}}{\partial x}\right) N_i \mathrm{d}x + \frac{\Delta t^2}{2} \int \mathbf{A}\left(\mathbf{S} - \frac{\partial \mathbf{F}}{\partial x}\right) \frac{\partial N_i}{\partial x} \mathrm{d}x \tag{29}$$

In general, $\mathbf{M}$ is a non-diagonal matrix, but for computational expediency it is convenient to use a lumped representation $\mathbf{M}_l$. Computational results show however that superior results are obtained by the use of the full matrix $\mathbf{M}$. A convenient scheme uses an iterative procedure[17] of the form:

$$\mathbf{M}_l(\Delta\hat{\mathbf{U}}_r^n - \Delta\hat{\mathbf{U}}_{r-1}^n) = \hat{\mathbf{f}}^n - \mathbf{M}\Delta\mathbf{U}_{r-1}^n, \qquad 1 \leqslant r \leqslant \text{niter} \tag{30}$$

where $\Delta\hat{\mathbf{U}}_r^n$ denotes the rth iterate. Unless otherwise stated, all the results presented in this paper were obtained for niter = 3 and $\Delta\hat{\mathbf{U}}_0 = \mathbf{0}$.

5.1 *Time-step limits*

If linear elements are employed on a regular mesh, and if furthermore A = const. in equation (2), a standard stability analysis shows that the local Courant number

$$C = \Delta t A/h \tag{31}$$

(h = mesh size, Δt = time-step size) must satisfy the inequalities

$$C \leqslant 1 \tag{32}$$

when niter = 1 in equation (30), so that a lumped mass matrix is used, whereas

$$C \leqslant 1/\sqrt{3} = 0{\cdot}5773 \tag{33}$$

for equation (27) when the complete mass matrix is employed.

6. NUMERICAL RESULTS IN 1-D

6.1 *Shallow water equations*

The non-linear shallow water equations in 1-D can be written in the form of equation (20) with

$$\mathbf{u} = \begin{bmatrix} H+\eta \\ (H+\eta)u \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} (H+\eta)u \\ (H+\eta)u^2 + g(H+\eta)^2/2 \end{bmatrix}; \quad \mathbf{S} = \begin{bmatrix} 0 \\ g\dfrac{dH}{dx}(H+\eta) \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ g(H+\eta) - u^2 & 2u \end{bmatrix}; \quad \mathbf{C} = \begin{bmatrix} 0 & 0 \\ g\dfrac{dH}{dx} & 0 \end{bmatrix} \tag{34}$$

Here H denotes the depth, η the surface elevation, u the velocity and g the acceleration due to gravity.

(a) *Shoaling of a wave.* The method of solution was applied to the classical problem of a shoaling solitary wave. The variation of the depth is shown in Figure 2(a) and the initial conditions are given by:

$$u = -(1 + \frac{a}{2})\eta/(\alpha x + \eta)$$

$$\eta = a\,\mathrm{sech}^2\left[\frac{\sqrt{3a}}{2}(x - 1/\alpha)\right]$$

and

$$a = 0{\cdot}1, \quad g = 1{\cdot}0, \quad \alpha = 1/30$$

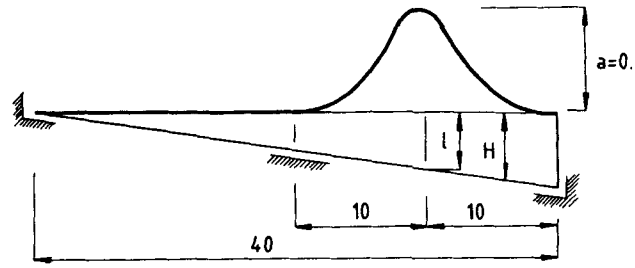
This problem has been studied previously by many authors.[21,22] The results of computations employing 40, 80 and 160 equal elements and a time step of $\Delta t = 0{\cdot}125$ are shown in Figure (2b). For 160 elements the total fluid mass was found to increase by only 0·1 per cent between $t = 0$ and $t = 16$ (see Figure 2(b).

(b) *Breaking of dam.* As a second example we studied the breaking of a dam. The results obtained are shown in Figure 3. The region was covered by 40 equal elements of unit length and initially:

$$u = 0$$
$$\eta = 1, \quad 0 \leqslant x \leqslant 20$$
$$\eta = 0, \quad 20 \leqslant x \leqslant 40$$

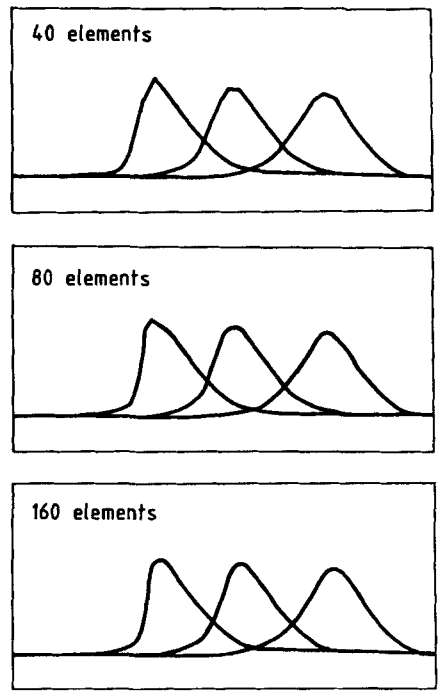
The depth H and g are constant and equal to unity. The results were obtained for a time step of $\Delta t = 0{\cdot}25$.

We note the correct steep front in 'shallow' water and the decreasing gradient for the 'deep' region.



(a) $\eta = a\ \mathrm{sech}^2 \tfrac{1}{2}(3a)^{\frac{1}{2}}\ (x-\alpha^{-1})$

$u = -(1+\tfrac{1}{2}a)\,\eta/(\alpha x+\eta)$ $a=0.1$, $g=1.0$, $\alpha=\tfrac{1}{30}$



(b)

Figure 2. Shoaling of a wave: (a) problem statement: (b) solution for 40, 80 and 160 elements

(c) *Outflow of a river.* As a third example we studied the propagation of a bore originated by tidal motion. The situation is depicted in Figure 4. At $t = 0$ we assumed a uniform flow and depth-field:

$$u = 1; \quad H = 1; \quad \eta = 0$$

Then the height at the right end of the domain was increased according to the function:

$$\eta = 1 - \cos(\pi t/T_0)$$

with $T_0 = 30$. The spatial discretization consisted of 40 elements of unit length. The time step was taken as $\Delta t = 0{\cdot}25$ and $g = 0{\cdot}667$.

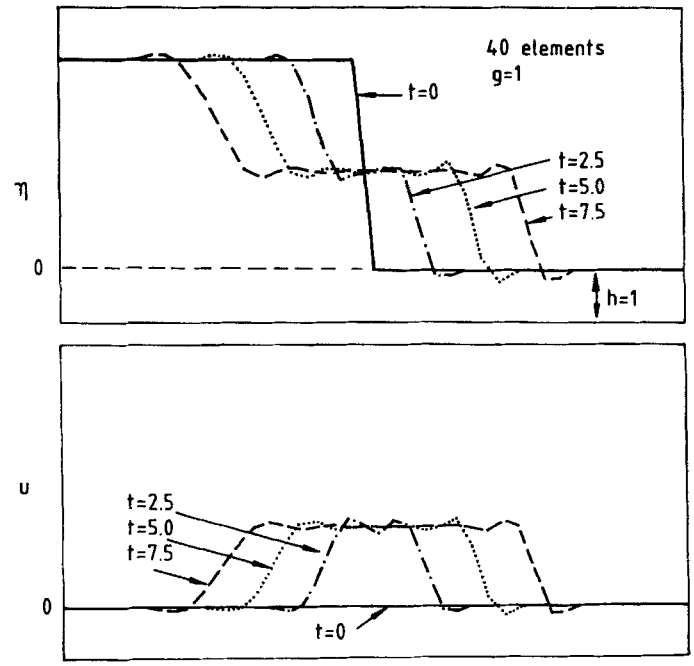


Figure 3. Breaking of a dam

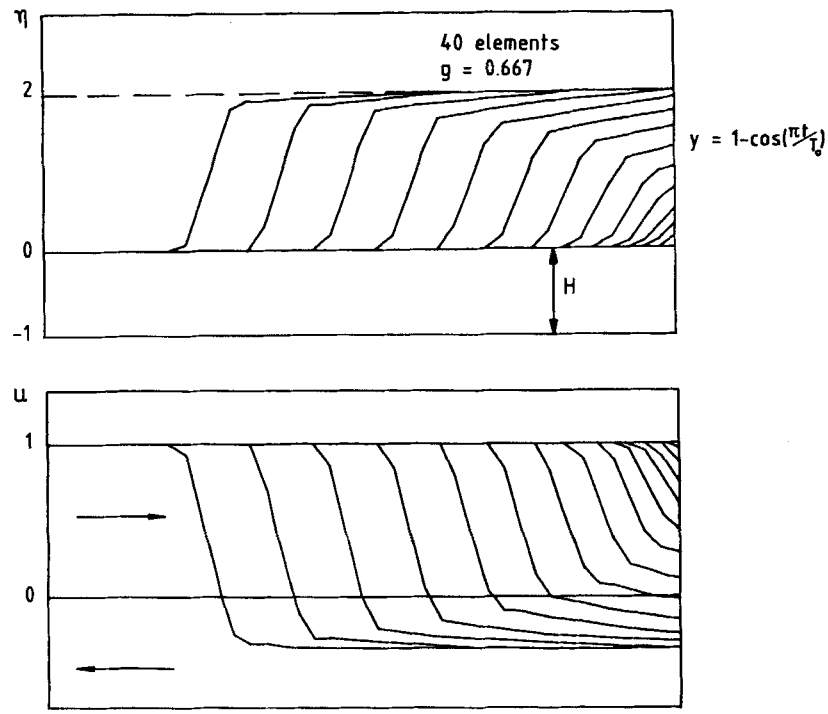


Figure 4. Outflow of a river with bore formation due to tidal motion

In Figure 4 the successive solutions are shown at intervals of $\Delta t_p = 20\Delta t = 5$. As the bore steepens, the velocity at the right end reaches its correct value of $u = -0{\cdot}333$, thus simulating an *inflow*.

6.2 *Compressible flows*

6.2.1 *Isothermal flow in a nozzle.* The equations governing one-dimensional isothermal flow in a nozzle with varying cross-section $a(x)$ are of the form of equation (20) with

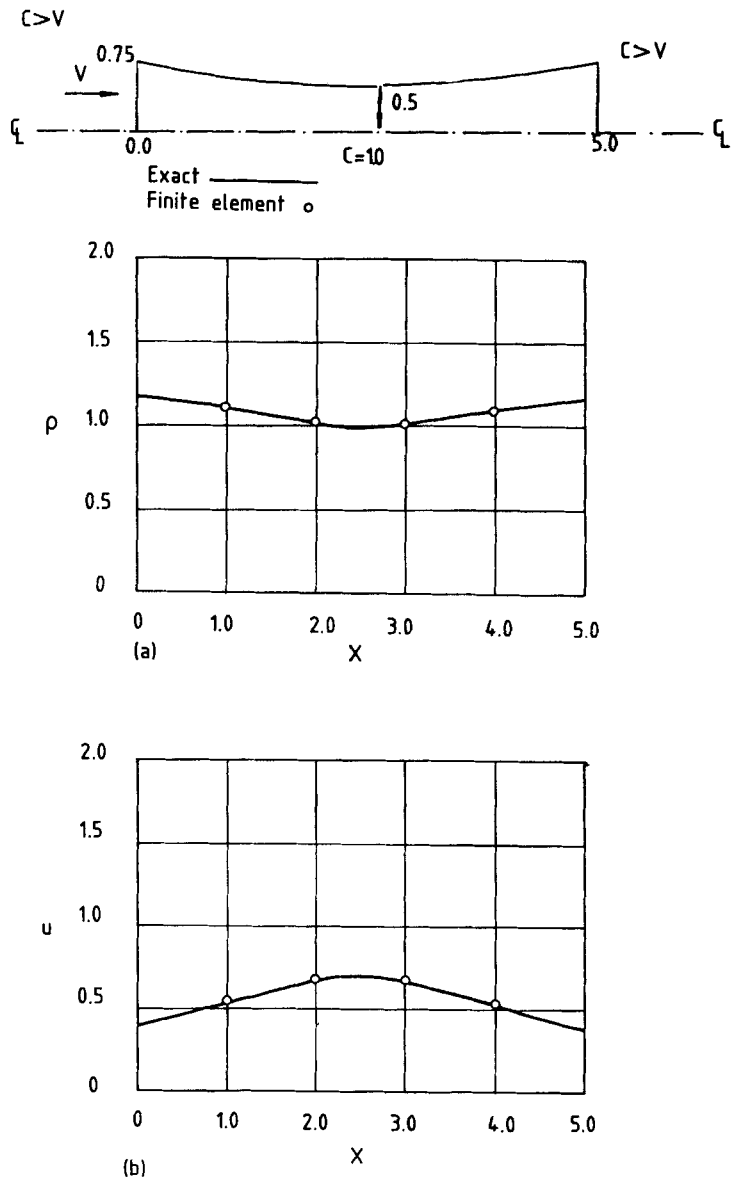


Figure 5. Isothermal flow in a nozzle: subsonic inflow–subsonic outflow without shock

$$\mathbf{u}=\begin{bmatrix}\rho a\\ \rho a u\end{bmatrix};\quad \mathbf{F}=\begin{bmatrix}\rho a u\\ c^2\rho a+\rho a u^2\end{bmatrix};\quad \mathbf{S}=\begin{bmatrix}0\\ \rho c^2\dfrac{da}{dx}\end{bmatrix}$$

$$\mathbf{A}=\begin{bmatrix}0 & 1\\ c^2-u^2 & 2u\end{bmatrix};\quad \mathbf{C}=\begin{bmatrix}0 & 0\\ \dfrac{c^2}{a}\dfrac{da}{dx} & 0\end{bmatrix}, \tag{35}$$

where ρ is the density, u the velocity and c the constant speed of sound. The problem considered has been studied previously by Tezduyar and Hughes,[23] who used

$$a=1{\cdot}0+(x-2{\cdot}5)^2/12{\cdot}5,\qquad 0\leqslant x\leqslant 5$$

and $c=1$. The calculated variations of density and velocity through the nozzle at the final steady

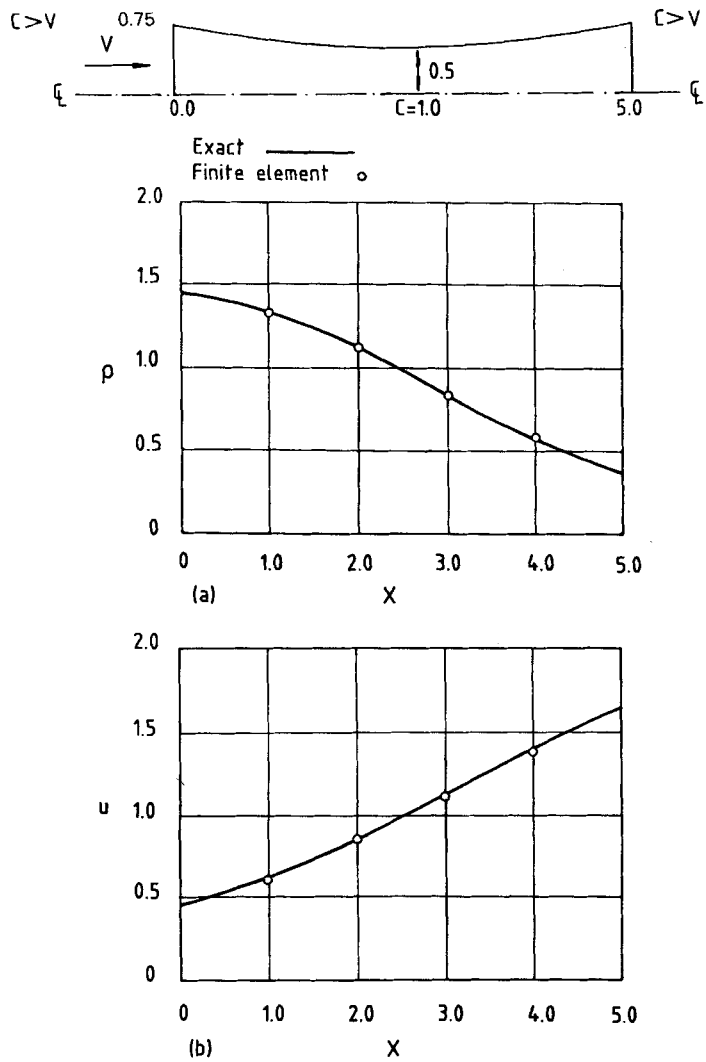


Figure 6. Isothermal flow in a nozzle: subsonic inflow–supersonic outflow without shock

state (which was attained after 500 explicit time-steps) are shown in Figures 5–7. In every case 40 equal elements were used.

In Figure 5 the results are depicted for the case of subsonic inflow–subsonic outflow with no shock, comparing them with the exact solution. The time-step size was $\Delta t = 0{\cdot}07$ and niter = 1 (see equation (30)). Figure 6 shows the results for the case of subsonic inflow–supersonic outflow with no shock, comparing them with the exact solution. The time-step size was $\Delta t = 0{\cdot}046$ and niter = 1 (see equation (30)). No sensible solutions could be obtained for the case of subsonic inflow–subsonic outflow with a shock. Here an artificial viscosity term was added to stabilize the solution. The form adopted is due to Lapidus,[24] and replaces the quantities resulting from equation (27) by smoothed values according to:

$$\hat{\mathbf{U}}_{si}^{n+1} = \hat{\mathbf{U}}_i^{n+1} - C_v \cdot \Delta t \cdot \int h^2 \left| \frac{\partial v^{n+1}}{\partial x} \right| \frac{\partial \mathbf{u}^{n+1}}{\partial x} \frac{\partial N_i}{\partial x} \mathrm{d}x \tag{36}$$

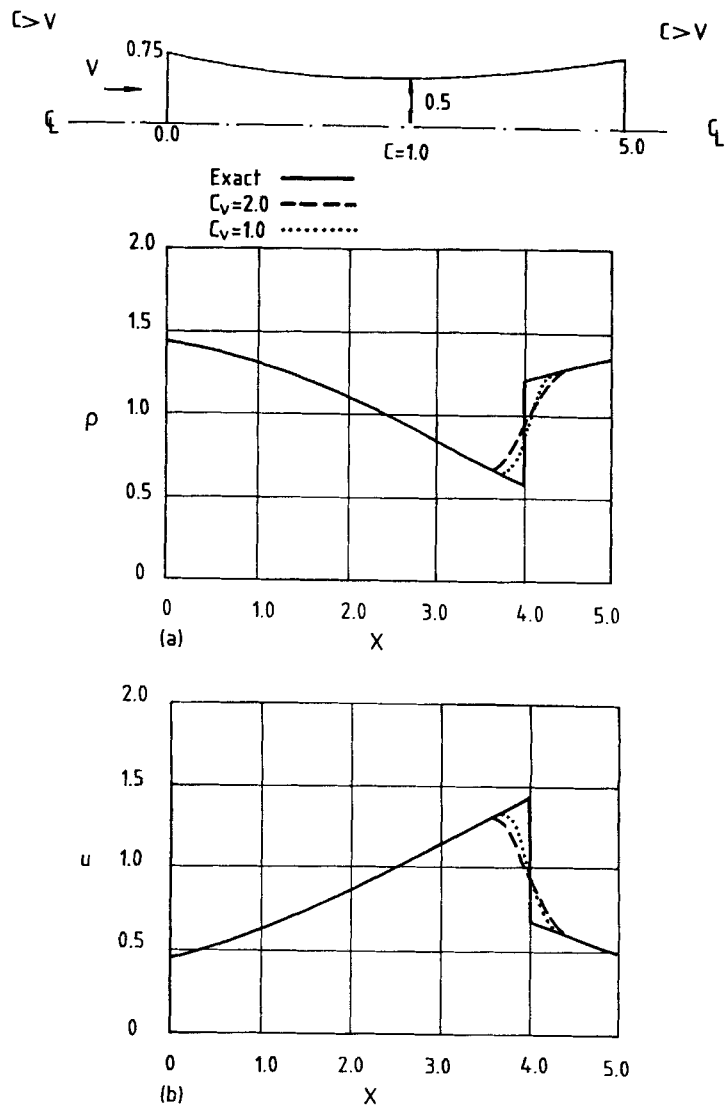


Figure 7. Isothermal flow in a nozzle: Subsonic inflow–subsonic outflow with shock

where C_v is a constant. Notice that the 'artificial' diffusion is proportional to the gradient of the velocity v.

The steady-state results obtained for $C_v = 1, 2$ and a time step $\Delta t = 0{\cdot}04$ are compared with the exact solution in Figure 7.

6.2.2. *Riemann shock tube problem.* The one-dimensional Euler equations for compressible flow can be expressed as (20) with:

$$\mathbf{u} = \begin{bmatrix} \rho \\ \rho u \\ \rho e \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho u e + u p \end{bmatrix}; \quad \mathbf{S} = \mathbf{0}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -u^2(3-\gamma)/2 & u(3-\gamma) & (\gamma-1) \\ (\gamma-1)u^3 - \gamma u e & \gamma e - 3(\gamma-1)u^2/2 & \gamma u \end{bmatrix}; \quad \mathbf{C} = \mathbf{0} \tag{37}$$

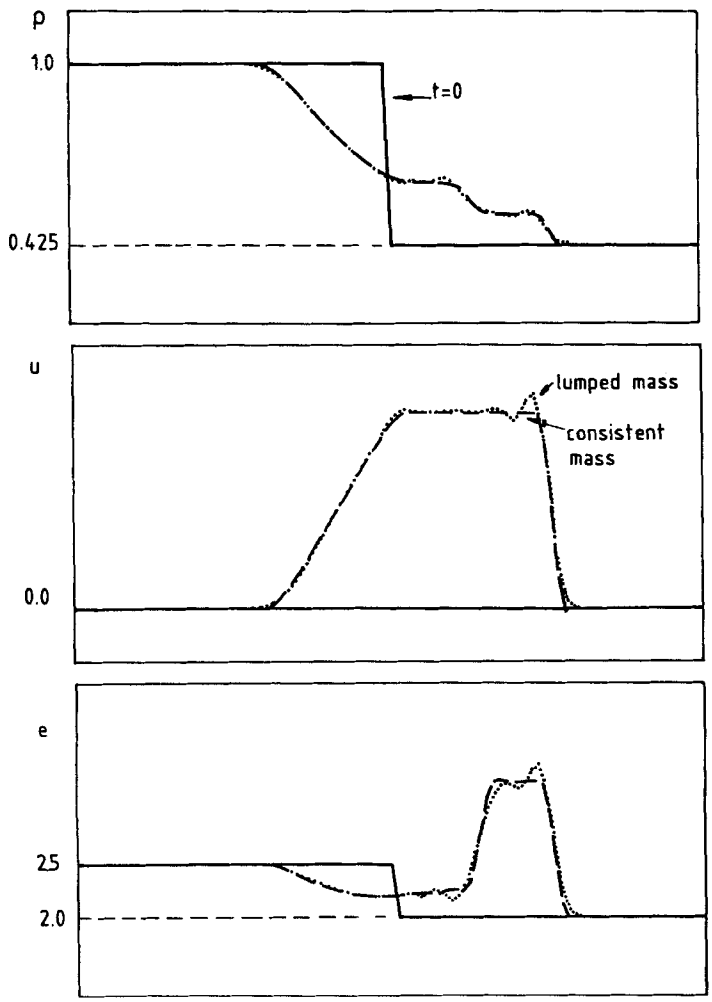


Figure 8. Riemann shock tube problem. Solutions at time $t = 14{\cdot}35$:... niter = 1 'Lax–Wendroff':----- niter = 3 'consistent mass'

where ρ is the density, u the velocity, e the specific total energy, γ the ratio of specific heats, and the pressure p satisfies

$$p = (\gamma - 1)\rho[e - u^2/2] \tag{38}$$

The problem analysed is that of a shock tube in which initially a diaphragm separates fluid in a low pressure chamber from fluid in a high pressure chamber. At $t = 0$, the diaphragm is removed and the resulting flow is characterized by the presence of a shock wave, contact discontinuity and rarefaction wave. The particular problem considered is the one due to Sod[25] which has also been studied by Baker[26] in the context of finite elements. The initial configuration is given by

$$\left.\begin{aligned} \rho &= 1{\cdot}0 \\ u &= 0{\cdot}0 \\ p &= 1{\cdot}0 \end{aligned}\right\} x < 50; \quad \left.\begin{aligned} \rho &= 0{\cdot}125 \\ u &= 0{\cdot}0 \\ p &= 0{\cdot}1 \end{aligned}\right\} x > 50$$

The problem was solved using 100 equal elements, a time step of $\Delta t = 0{\cdot}205$ and a Lapidus constant $c_v = 1{\cdot}0$. Figure 8 shows the profiles for ρ, u and at time $t = 14{\cdot}35$ (70 time steps) for niter = 1 ('Lax–Wendroff') and niter = 3. The influence of the complete mass is readily seen, giving less overshoot across the shock fronts.

7. EXTENSION TO HIGHER-DIMENSIONAL PROBLEMS

In more than one dimension an equation of the form:

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}^i}{\partial x^i} = \mathbf{S}(\mathbf{u}) \tag{39}$$

can be solved in a similar fashion. Equation (23) is now replaced by:

$$\begin{aligned} \mathbf{u}^{n+1} &= \mathbf{u}^n + \Delta t\left(\mathbf{S} - \frac{\partial \mathbf{F}^i}{\partial x^i}\right)\Big|^n + \frac{\Delta t^2}{2}\left[\mathbf{C}\left(\mathbf{S} - \frac{\partial \mathbf{F}^i}{\partial x^i}\right) - \frac{\partial}{\partial x^i}\left[\mathbf{A}^i\left(\mathbf{S} - \frac{\partial \mathbf{F}^j}{\partial x^j}\right)\right]\right]\Big|^n \\ &= \mathbf{u}^n + \Delta t\left(\mathbf{S} - \mathbf{A}^i\frac{\partial \mathbf{u}}{\partial x^i}\right)\Big|^n + \frac{\Delta t^2}{2}\left[\mathbf{C}\left(\mathbf{S} - \mathbf{A}^i\frac{\partial \mathbf{u}}{\partial x^i}\right) - \frac{\partial}{\partial x^i}\left[\mathbf{A}^i\left(\mathbf{S} - \mathbf{A}^j\frac{\partial \mathbf{u}}{\partial \mathbf{x}^j}\right)\right]\right]\Big|^n \end{aligned} \tag{40}$$

where

$$\mathbf{A}^i = \frac{\partial \mathbf{F}^i}{\partial \mathbf{u}} \tag{41}$$

Specifically, for the problem of advection of a scalar variable u with a solenoidal velocity field $\mathbf{v} = (A^1, A^2, A^3)$, equation (10) can be written as:

$$u^{n+1} = u^n - \Delta t(\mathbf{v}\cdot\nabla u)|^n + \frac{\Delta t^2}{2}\nabla\cdot(\mathbf{v}\otimes\mathbf{v})\cdot\nabla\mathbf{u} \tag{42}$$

the analogy with the streamline upwinding[4] or balancing dissipation[5] being obvious.

The Galerkin method for spatial discretization has been applied to equation (40) and several examples have been studied in 2-D using linear triangular elements. Obviously with equal ease other elements could be employed. We have chosen to use the simple triangle for ease of mesh generation and refinement and for computational efficiency. As in many problems steep fronts are developed, higher order elements are not recommended.

8. EXAMPLES IN 2-D

8.1 *Scalar advection*

As a first example we studied the classical problem of the advection of a cone in a rotating fluid. The problem has been studied extensively in the past.[4,13,14,15,17,26–28] For the computation a criss-cross mesh was chosen (the alignment of elements plays no crucial role) with $C_v = 0$, and a complete rotation was accomplished in 200 time steps. The initial configuration is shown in Figure 9; Figure 10 shows the solution after 1 rotation for niter = 1 ('Lax–Wendroff), or lumped

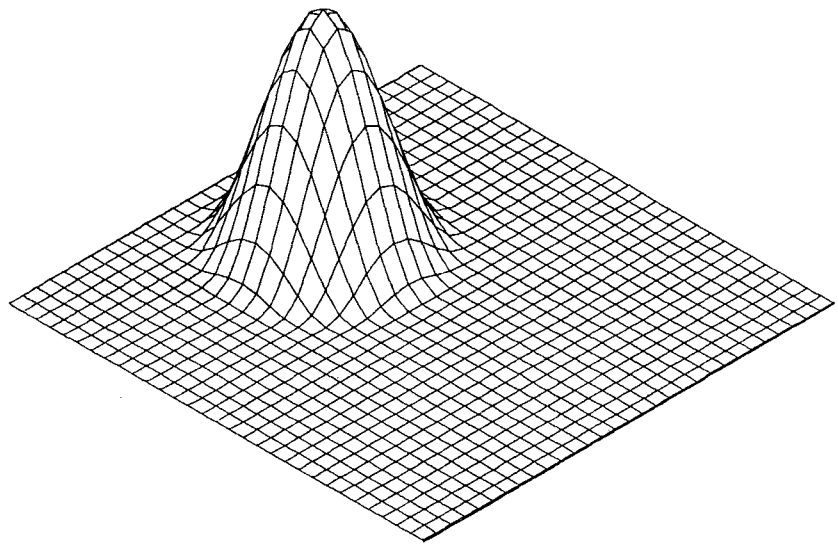


Figure 9. Advection of a cone in a rotating fluid: initial configuration

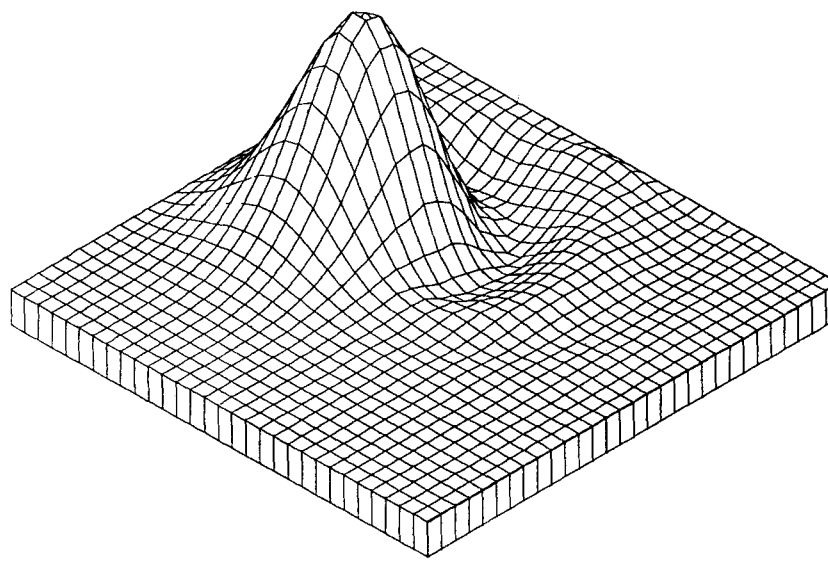


Figure 10. Advection of a cone in a rotating fluid: solution after 1 rotation for niter = 1

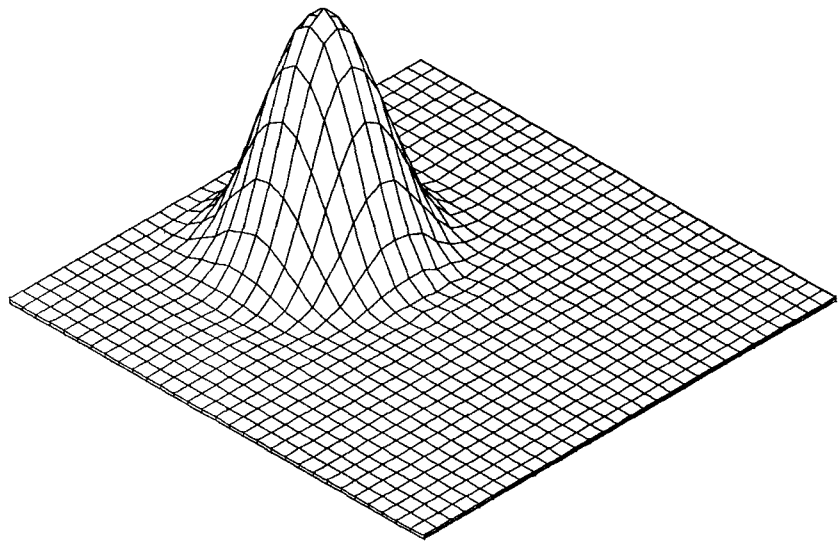


Figure 11. Advection of a cone in a rotating fluid: solution after 1 rotation for niter = 3

mass FEM) and Figure 11 for niter = 3. A very good solution is thus obtained for the case of consistent mass, although no big systems of equations have to be solved or stored. The cone height was reduced by 2 per cent and the maximum undershoot was 1·5 per cent. This shows that the method attains an accuracy comparable or even higher than that achieved by implicit procedures[4,26,27] or characteristics-based methods.[13,14,15,28]

8.2. *Non-Linear shallow water equations in 2-D*

The 2-D non-linear shallow water equations can be written in the form of equation (39) with:

$$\mathbf{u} = \begin{bmatrix} H+\eta \\ u(H+\eta) \\ v(H+\eta) \end{bmatrix}; \quad \mathbf{F}^1 = \begin{bmatrix} u(H+\eta) \\ u^2(H+\eta)+\frac{g}{2}(H+\eta)^2 \\ uv(H+\eta) \end{bmatrix}; \quad \mathbf{F}^2 = \begin{bmatrix} v(H+\eta) \\ uv(H+\eta) \\ v^2(H+\eta)+\frac{g}{2}(H+\eta)^2 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 0 \\ gH_{,x}(H+\eta) \\ gH_{,y}(H+\eta) \end{bmatrix}$$

We considered the problem of a channel with uniform depth $H = 1$, in which a sudden contraction of width occurs. The inflow was prescribed as supercritical, with a Froude number of $Fr = 2{\cdot}5$.

The spatial discretization is shown in Figure 12(a).

The solutionwas obtained after approximately 500 times steps and is depicted in Figures 12(b) and 12(c), showing the height distribution and the velocity field. Observe that although the mesh is quite coarse the solution depicts correctly the oblique wave originating in front of the contraction.

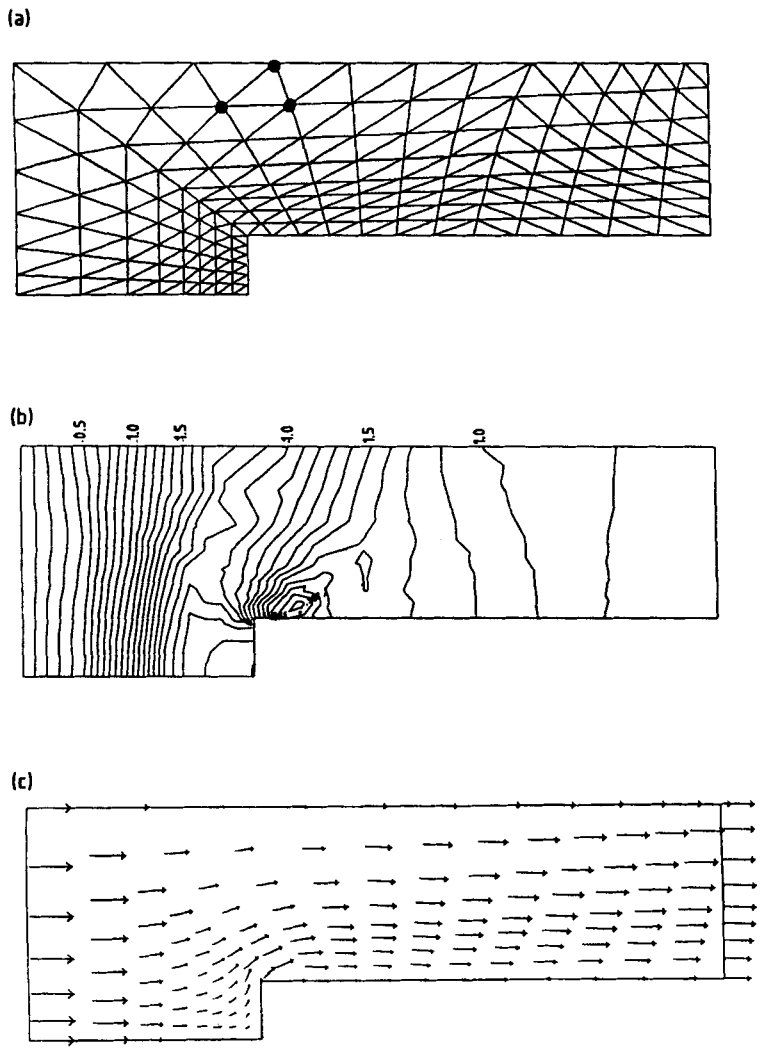


Figure 12. Supercritical flow ($Fr = 2{\cdot}5$) in a narrowing channel: (a) spatial discretization: (b) η (Total depth $= \eta + 1{\cdot}0$); (c) velocity field

8.3. *Euler equations in 2-D*

The 2-D Euler equations can be written in the form of equation (39) with

$$\mathbf{u} = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho e \end{bmatrix}; \quad \mathbf{F}^1 = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(\rho e + p) \end{bmatrix}; \quad \mathbf{F}^2 = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(\rho e + p) \end{bmatrix}$$

$$p = (\gamma - 1)\rho[e - (u^2 + v^2)/2]$$

We considered the flow around a parabolic arc in a wind-tunnel. The problem statement is described in Reference 23. The inflow Mach number was given by $Ma = 0{\cdot}84$, producing a local supersonic

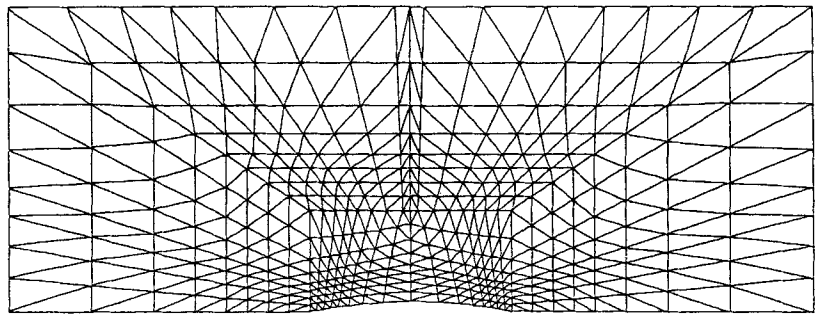


Fig. 11(a)

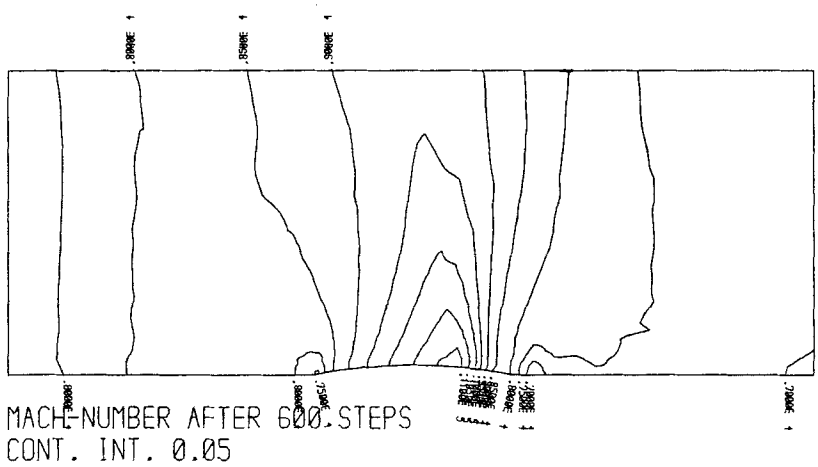


Fig. 11(b)

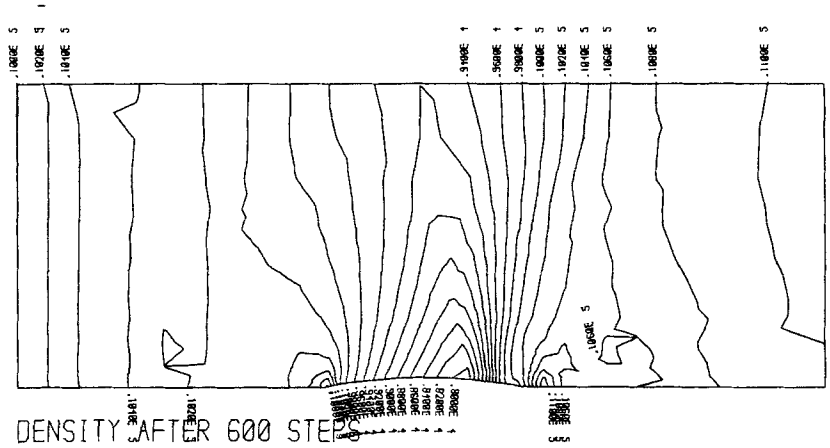


Fig. 11(c)

Figure 13. Parabolic arc in a wind-tunnel ($Ma = 0{\cdot}84$): (a) spatial discretization; (b) Mach number contours; (c) pressure contours

field over the aerofoil. Figure 13(a) shows the spatial discretization, Figure 13(b) the Mach number contours and Figure 13(c) the contours of the pressure field. The results were in good agreement with those given by Tezduyar and Hughes.[23]

9. CONCLUDING REMARKS

The examples illustrate that the procedures suggested in the present paper lead to an accurate and versatile method of dealing with hyperbolic, non-linear systems using f.e.m. procedures. Indeed in several of the examples results much superior to those attainable with equivalent finite difference methods (currently dominating this field) have been obtained. The method expressed is relatively simple to use and in this paper we only treated explicit time-marching methods which are economical and can be used on relatively small computers. Obviously it can be adapted to implicit forms but in this context (finite velocity of propagation of disturbances) these do not appear attractive. Although this paper concentrates on examples which are purely hyperbolic and contain no natural dissipation (e.g. diffusion) a direct inclusion of this does not present any difficulty (see Appendix I).

Although the solution presented overcomes many of the difficulties encountered, much yet remains to be done. Two directions of improvement are currently being explored. The first deals with the inherent problems of discontinuities at shocks. Here it has been frequently suggested that local mesh refinement or use of discontinuous shape functions may be effective.

The second is the question of economy for meshes with rapidly varying element size or for systems with order of magnitude different eigenvalues. Here technologies of simultaneous use of different time steps may be the best way to follow.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. J. Donea for many interesting and stimulating discussions during the course of this work.

Mr. Löhner kindly acknowledges the support of the Studienstiftung des Deutschen Volkes and K. Morgan the support of NASA Langley Research Center via grant NSG-1321.

APPENDIX I: INTRODUCTION OF DIFFUSION

If the equation to be solved reads:

$$u_{,t} + f_{,x} = (vu_{,x})_{,x} \tag{43}$$

we can follow two lines:

(a) make a formal Taylor expansion for Δu as in the hyperbolic case and neglect higher order derivatives (this has been interpreted as an operator splitting[17])

(b) transform (43) to Lagrangian co-ordinates, move the mesh, and reinterpolate neglecting in the Taylor series expansion again the derivatives of order $k > 2$.

In either case, the algorithm gives for the Euler time-stepping scheme:

$$\Delta u = -\Delta t f_{,x} + \Delta t\left(vu_{,x} + \frac{\Delta t}{2} A f_{,x} \right)_{,x} \tag{44}$$

We note that nothing changes in the 'hyperbolic part' of the equation. The inclusion of a source term $S \neq S(u)$ was omitted for the sake of clarity, but can be introduced without problems.[16]

APPENDIX II: SYSTEMS OF EQUATIONS IN 1-D

Let us consider the general form:

$$\mathbf{u}_{,t} + \mathbf{A}\mathbf{u}_{,x} = \mathbf{0} \tag{45}$$

Suppose $\mathbf{A} = \text{const.}$ for simplicity. We can decompose $\mathbf{A}$ using

$$\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1} \tag{46}$$

with

$\mathbf{X}$ = matrix of eigenvectors

$\mathbf{\Lambda}$ = diagonal matrix of eigenvalues

Transforming the variables from $\mathbf{u}$ to $\mathbf{v} = \mathbf{X}^{-1}\mathbf{u}$ we obtain:

$$\mathbf{v}_{,t} + \mathbf{\Lambda}\mathbf{v}_{,x} = \mathbf{0} \tag{47}$$

i.e. n decoupled equations of the form

$$v^i{}_{,t} + \lambda_i v^i_{,x} = 0 \tag{48}$$

Now we can solve each of these equations as in the case of a single equation (See section 2), and obtain:

$$\Delta v^i = -\Delta t \lambda_i v^i_{,x}|^n + \frac{\Delta t^2}{2}(\lambda_i^2 v^i_{,x})_{,x}|^n \tag{49}$$

or simply

$$\Delta \mathbf{v} = -\Delta t \mathbf{\Lambda}\mathbf{v}_{,x}|^n + \frac{\Delta t^2}{2}(\mathbf{\Lambda}^2\mathbf{v}_{,x})_{,x}|^n \tag{50}$$

This general result can be transformed back to the original ($\mathbf{u}$) variables, giving:

$$\Delta \mathbf{u} = -\Delta t \mathbf{A}\mathbf{u}_{,x}|^n + \frac{\Delta t^2}{2}(\mathbf{A}^2\mathbf{u}_{,x})_{,x}|^n \tag{51}$$

$$= -\Delta t \mathbf{F}_{,x}|^n + \frac{\Delta t^2}{2}(\mathbf{A}\mathbf{F}_{,x})_{,x}|^n$$

i.e. the general form required. If $\mathbf{A} = \mathbf{A}(u, x, t)$ the analysis becomes much more cumbersome.

REFERENCES

1. I. Christie, D. F. Griffiths, A. R. Mitchell and O. C. Zienkiewicz, 'Finite element methods for second order differential equations with significant first derivatives', *Int. J. Num. Meth. Eng.*, **10**, 1389–1396 (1976).
2. J. C. Heinrich, P. S. Huyekorn, O. C. Zienkiewicz and A. R. Mitchell, 'An "upwind" finite element scheme for two-dimensional convective transport equation', *Int. J. Num. Meth. Eng.*, **11**, 131–143 (1977).
3. O. C. Zienkiewicz and J. C. Heinrich, 'The finite element method and convection problems in fluid mechanics', in D. C. Zienkiewicz, J. T. Oden, M. Morandi-Cecchi and C. Taylor (eds), *Finite Elements in Fluids* Vol. 3, H. H. Gallager, Wiley, 1978 Chapter 1, pp. 1–22.
4. T. J. Hughes and A. Brooks, 'A multidimensional upwind scheme with no cross wind diffusion', pp. 19–36 in T. J. R. Hughes (ed.) *Finite Element Methods for Convection Dominated Flows*, ASME, New York, AMD, December 1979, Vol. 34, pp. 19–36.
5. D. W. Kelly, S. Nakazawa, O. C. Zienkiewicz and J. C. Heinrich, 'A note on anisotropic balancing dissipation in finite element method approximation to convective diffusion problems', *Int. J. Num. Meth. Eng*, **15**, 1705–1711 (1980).
6. C. Johnson 'On the convergence of a mixed finite element method for plate bending problems', *Numer. Math.*, **21**, 43–62 (1973).
7. I. Babuska and W. G. Szymczak, 'An error analysis of finite element methods applied to convection diffusion problems', *Technical Report BN-962*, Inst. for Phys. Science and Technology, University of Maryland, 1981.

8. W. H. Raymond and A. Gardner, 'Selective lumping in a Galerkin method for solving wave problems with variable grids', *Monthly Weather Rev.*, **104,** 1583–1590 (1976).
9. A. N. Brooks and T. J. R. Hughes 'Streamline upwind/Petrov Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations', *Comp. Meth. Appl. Mech. Eng.*, **30,** (1982).
10. T. J. R. Hughes R. L. Taylor and W. Kanoknukulchai 'A finite element method for large displacement contact and impact problems', in K. J. Bathe, J. T. Oden and W. Wunderlich (eds), *Formulations and Computational Algorithm in Finite Element Analysis* MIT, Boston, 1977, chap. 16.
11. J. C. Heinrich and E. Envia, 'Finite element techniques in transport phenomena', private communication, 1982.
12. R. A. Adey and C. A. Brebbia, 'Finite element solution of effluent dispersion', in C. A. Brebbia and J. J. Connor (eds) *Numerical Methods in Fluid Dynamics* Pentech Press, London 1974, pp. 325–354.
13. B. Ibler 'Resolution des equations de Navier–Stokes par une methode d'elements finis', *EDF Report HE 041/81/15*, 1981.
14. M. Bercovier, O. Pironneau and V. Sastri, 'Finite elements and characteristics for some parabolic–hyperbolic problems', *Appl. Math. Modelling*, **7**, (2) 89–96 (1983).
15. J. P. Benque, G. Labadie and J. Ronat, 'A new finite element method for the Navier–Stokes equations coupled with a temperature equation', to appear in *Int. J. Num. Meth. Fluids.*
16. O. C. Zienkiewicz, 'Finite elements in fluid mechanics—a decade of progress', To appear in *F. E. in Fluids*, Vol. V.
17. J. Donea 'A Taylor–Galerkin method for convective transport problems', *Int. J. Num. Meth. Eng.*, **20**, 101–119 (1984).
18. P. J. Roache, '*Computational Fluid Mechanics*, Hermosa publishers, Albuquerque.
19. O. C. Zienkiewicz, *The Finite Element Method*, 3rd edn McGraw-Hill, 1977.
20. O. C. Zienkiewicz and K. Morgan, *Finite Elements and Approximation*, Wiley, 1983.
21. S. Nakazawa, D. W. Kelly, O. C. Zienkiewicz, I. Christie and M. Kawahara, 'An analysis of explicit finite element approximations for the shallow water wave equations', *Proceedings of Third International Conference on Finite Elements in Flow Problems*, Banff, 1980, Vol. 2, pp. 1–7.
22. M. Kawahara, H. Hirano, K. Tsubota and K. Inagaki, 'Selective lumping finite element method for shallow water flow', *Int. J. Num. Meth. Fluids*, **2,** 89–112 (1982).
23. T. E. Tezduyar and T. J. R. Hughes, 'Development of time-accurate finite element techniques for first-order hyperbolic systems with particular emphasis on the compressible Euler equations', *Report*, Stanford University, 1983.
24. A. Lapidus 'A detached shock calculation by second-order finite differences', *J.C.P.*, **2,** 154–177 (1967).
25. G. Sod. 'A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws', *J.C.P.*, **27,** 1–31 (1978).
26. A. J. Baker, *Finite Element Computational Fluid Mechanics*, McGraw-Hill, New York, 1983.
27. P. M. Gresho, R. L. Lee and R. L. Sani, 'Advection dominated flows, with emphasis on the consequences of mass lumping', in R. H. Gallager, O. C. Zienkiewics, J. T. Oden, M. Morandi-Cecchi and C. Taylor (eds) Finite Elements in Fluids, Vol. III Wiley, 1978, pp. 335–350.
28. J. Glass and W. Rodi, 'A higher order numerical scheme for scalar transport', *Comp. Meth. Appl. Mech. Eng.* **31** 337–358 (1982).