LOCAL MOVING FINITE ELEMENTS

M.J. BAINES

NUMERICAL ANALYSIS REPORT 9/85

This work forms part of the research programme of the Institute for Computational Fluid Dynamics at the Universities of Oxford and Reading.

1. INTRODUCTION

Computational Fluid Dynamics covers a wide range of modelling activity. It includes complex codes based on traditional methods for difficult and pressing engineering problems and, at the other end of the spectrum, the developmer of numerical methodology stimulated by the desire for flexible and efficient new algorithms. It is with the latter area that this report is concerned, in particular with the development of algorithms which adapt in some way to the flow field. By considering moving grids we overcome some of the limitations peculiar to fixed grids while raising other issues. Here we describe and develop the Moving Finite Element (MFE) technique in relation to hyperbolic conservation laws.

We begin in one dimension by summarising the MFE method for piecewise linear approximations as originally presented in (Miller, 1981) but without his use of penalty functions: see (Wathen & Baines, 1985). We then describe an alternative formulation, based on element basis functions, which is equivalent to the Miller formulation but involves only local projections as in (Baines, 1985). From this formulation we obtain ordinary differential equations in time for the evolution of each element segment in terms of its velocity and angular speed. The determination of the nodal velocities from the motion of the segments (also local) is then described, and the resulting method is used to derive the decomposition of the standard MFE matrix.

We then describe singularities of the method. In the rather rare event of the nodes becoming collinear (parallelism) one of the local matrices becomes singular, and a special non-local procedure is needed which is however simple to implement. The only other singularity of the method occurs when nodes overtake one another. This arises if the time step (in the numerical integration of the semi-discrete problem) is so large as to destroy the accuracy. However it may also arise (correctly) in hyperbolic problems when a shock forms. In that case the appropriate jump conditions may be used to continue the solution.

Time-stepping strategies are then discussed in relation to accuracy and practical results. So far the forward Euler explicit method has been found to be sufficient, as in (Wathen, 1984, Johnson 1984). Results are given for two non-linear examples, the inviscid Burgers' equation (for which the MFE method gives the exact solution) and the Buckley-Leverett equation. Both programs run in BASIC on a home computer.

Generalising to systems of equations we describe a method which uses separate grids for each conserved variable given by (Baines and Wathen,1985). Resul are given for a shock tube problem. Again a small microcomputer is sufficient to obtain a reasonable approximate solution to this problem.

The method has considerable potential in two dimensions see (Wathen, 1984). However care is required in that the time derivative v_t of the piecewise linear continuous function v in two dimensions now belongs only to a subspace of the space S_ϕ spanned by the element basis functions. Hence a constrained projection is considered which gives the standard Miller method. The approach again demonstrates that the MFE matrix has a decomposition similar to that in one dimension. A local segment velocity can also be derived which assists in determining the rules for shock formulation. Results are given for a non-linear problem in two dimensions, a generalisation of the Buckley-Leverett equation.

2. STANDARD MFE APPROACH

We first describe the standard Moving Finite Element method as presented in (Miller,1981). We begin by approximating u in the scalar equation

$$u_t + f(u)_x = 0 \tag{2.1}$$

by the piecewise linear function v given by

$$v = \sum_j a_j \alpha_j \tag{2.2}$$

where the a_j are nodal coefficients and the α_j are linear basis functions as shown in Figure 1(a).

Fig 1 → HERE

For fixed finite elements a_j depends on time t and α_j depends on the space co-ordinate x. In the MFE method α_j depends also on the time t through the nodal co-ordinates s_i $(i = 0,1,\ldots,N+1)$. In order to study the solution of the conservation law (2.1) we differentiate v with respect to time. Since t appears twice in the expression for v we have

$$v_t = \sum_j (\dot{a}_j \alpha_j + a_j \dot{\alpha}_j) \tag{2.3}$$

where the dot denotes differentiation with respect to time. Now, by the chain rule,

$$\dot{\alpha}_j = \sum_i \dot{s}_i \frac{\partial \alpha_j}{\partial s_i} \tag{2.4}$$

so that (2.3) becomes

$$v_t = \sum_j \dot{a}_j \alpha_j + \sum_j a_j \sum_i \dot{s}_i \frac{\partial \alpha_j}{\partial s_i} \tag{2.5}$$

By interchanging the order of summation we obtain

$$v_t = \sum_j (\dot{a}_j \alpha_j + \dot{s}_j \beta_j) , \tag{2.6}$$

where

$$\beta_j = \sum_i a_i \frac{\partial \alpha_i}{\partial s_j} \tag{2.7}$$

is a second type basis function dependent not only on t and on s_i but also on $\underline{a}$, the vector of nodal coefficients. From (2.2) we may write

$$\beta_j = \frac{\partial v}{\partial s_j} \tag{2.8}$$

which in the case of piecewise linear basis functions can be written (after some manipulation as in Baines (1985) in the form

$$\beta_j = -m\alpha_j \quad , \tag{2.9}$$

where m is the local slope of v (see also (Lynch,1981)).

Thus the second type basis function β_j has components which are multiples of α_j and have the same support as α_j. (The result is true for linear basis functions in any number of dimensions. Diagrams illustrating the α_j and β_j in one dimension are shown in Figure 1.

Hence in one dimension

$$v = \sum_j (\dot{a}_j\alpha_j + \dot{s}_j\beta_j) \tag{2.10}$$

belongs to the space of piecewise linear discontinuous functions $S_{\alpha\beta}$, say.

If $f(v)_x$ also belongs to $S_{\alpha\beta}$ we immediately obtain from (2.1) ordinary differential equations for a_j and s_j in time. For example if

$$f(u) = \tfrac{1}{2}u^2 \tag{2.11}$$

the ODE's are

$$\dot{a}_j = 0 \qquad \dot{s}_j = a_j \quad , \tag{2.12}$$

which are readily integrated to give a_j and s_j. In this particular case the trajectories of the nodes are the characteristics of the partial differential equation as in (Wathen,1984). This is an exceptional case however.

More generally we may project $f(v)_x$ into the space $S_{\alpha\beta}$ by minimising the L_2 norm of the residual

$$\| v_t + f(v)_x \|_2 \tag{2.13}$$

over the variables $\dot{a}_j$ and $\dot{s}_j$. This leads to the double set of Galerkin equations

$$\left.\begin{aligned} \langle \alpha_j,\ v_t + f(v)_x \rangle &= 0 \\ \langle \beta_j,\ v_t + f(v)_x \rangle &= 0 \end{aligned}\right\} \quad \forall_j \tag{2.14}$$

Substituting for v from (2.10) we obtain the matrix system of ODE's

$$A\dot{\underline{y}} = \underline{g} \tag{2.15}$$

where

$$\dot{\underline{y}}^T = [\ldots,\dot{a}_j,\dot{s}_j,\ldots] \tag{2.16}$$

$$A = \{A_{ij}\}\ , \qquad A_{ij} = \begin{bmatrix} <\alpha_i,\alpha_j> & <\alpha_i,\beta_j> \\ <\beta_i,\alpha_j> & <\beta_i,\beta_j> \end{bmatrix} \tag{2.17}$$

and $$\underline{g} = \{g_i\}\ , \qquad g_i = <\begin{matrix}\alpha_i \\ \beta_i\end{matrix}\ , -f(v)_x> \ . \tag{2.18}$$

The matrix A is symmetric block 2x2 tri-diagonal, positive semi-definite depending on $\underline{y}$. Equation (2.15) gives the MFE equations as derived in (Miller,19 without the use of penalty functions.

It has been shown in (Morton,1982) that the MFE equations (2.15) carry the best least squares fit to the exact solution in the case of a scalar conservation law.

3. LOCAL APPROACH

We now adopt instead a local elementwise approach. The function v_t in the space of piecewise linear discontinuous functions may be re-parameterised in the form

$$v_t = \sum_j (\dot{a}_j\alpha_j + \dot{s}_j\beta_j) = \sum_k (w_{k1}\phi_{k1} + w_{k2}\phi_{k2}) \tag{3.1}$$

using element basis functions ϕ_{ki} as shown in Figure 2. The w_{ki} are the coefficients of the ϕ_{ki} in the expansion which can be related to the $\dot{a}_j,\dot{s}_j$ (see below). Denote by S_ϕ the space spanned by the basis functions ϕ_{ki}: this is the same space as $S_{\alpha\beta}$ in the one-dimensional case.

Again, if $f(v)_x$ belongs to S_ϕ ($\equiv S_{\alpha\beta}$) we obtain w_{k1}, w_{k2} at once and consequently $\dot{a}_j$ and $\dot{s}_j$. More generally we may again project $f(v)_x$ into S_ϕ. This can be done locally within each element by minimising the local

element L_2 norm

$$\| v_t + f(v)_x \|_2 \tag{3.2}$$

over w_{k1}, w_{k2}. We thus obtain alternative Galerkin equations

$$\begin{aligned} \langle \phi_{k1}, v_t + f(v)_x \rangle &= 0 \\ \langle \phi_{k2}, v_t + f(v)_x \rangle &= 0 \end{aligned} \tag{3.3}$$

for each element which can be written in the form

$$C_k \underline{w}_k = \underline{b}_k \quad , \tag{3.4}$$

a 2x2 system. In (3.4)

$$\underline{w}_k = [w_{k1}, w_{k2}]^T \tag{3.5}$$

$$C_k = \begin{bmatrix} \langle \phi_{k1}, \phi_{k1} \rangle & \langle \phi_{k1}, \phi_{k2} \rangle \\ \langle \phi_{k2}, \phi_{k1} \rangle & \langle \phi_{k2}, \phi_{k2} \rangle \end{bmatrix} \tag{3.6}$$

and

$$\underline{b}_k = \left\langle \begin{matrix} \phi_{k1} \\ \phi_{k2} \end{matrix} , -f(v)_x \right\rangle \quad . \tag{3.7}$$

To relate the w_k's to $\dot{a}_j$ and $\dot{s}_j$ we use the fact that

$$\alpha_j = \phi_{j-\frac{1}{2},2} + \phi_{j+\frac{1}{2},1} \qquad \beta_j = -m_{j-\frac{1}{2}} \phi_{j-\frac{1}{2},2} - m_{j+\frac{1}{2}} \phi_{j+\frac{1}{2},1} \quad , \tag{3.8}$$

where elements $j-\frac{1}{2}$, $j+\frac{1}{2}$ are adjacent to node j, as in Figure 2. Then using (3.1) we have the correspondence

$$\left.\begin{aligned} \dot{a}_j - m_{j-\frac{1}{2}} \dot{s}_j &= w_{j-\frac{1}{2},2} \\ \dot{a}_j - m_{j+\frac{1}{2}} \dot{s}_j &= w_{j+\frac{1}{2},1} \end{aligned}\right\} \ \forall j \quad . \tag{3.9}$$

This can be written in the form

$$M_j \dot{\underline{y}}_j = \underline{w}_j \quad , \tag{3.10}$$

another 2x2 system, where

$$\dot{y}_j^T = [\dot{a}_j, \dot{s}_j] \tag{3.11}$$

$$M_j = \begin{bmatrix} 1 & -m_{j-\frac{1}{2}} \\ 1 & -m_{j+\frac{1}{2}} \end{bmatrix} \tag{3.12}$$

$$\underline{w}_j^T = [w_{j-\frac{1}{2},2},\ w_{j+\frac{1}{2},1}] \ . \tag{3.13}$$

Since both the Miller method and the local elementwise method minimise the same residual in the same space the MFE equations derived from the two methods must be identical. It follows that, by writing

$$C = \begin{bmatrix} \ddots & & 0 \\ & C_k & \\ 0 & & \ddots \end{bmatrix} \ , \qquad M = \begin{bmatrix} \ddots & & 0 \\ & M_j & \\ 0 & & \ddots \end{bmatrix} \ , \tag{3.14}$$

we have from (2,15), (2.16), (3.11), (3.10), (3.4), (3.7), (3.8) and (2.18) the decomposition of the global MFE matrix A

$$A = M^T C M \ , \tag{3.15}$$

where both C and M are diagonal block 2x2 matrices (apart from end effects).

Thus we have shown that the MFE method consists of finding a straight line best fit to $f(v)_x$ in each element together with a local mapping from the elementwise velocity description (in terms of w's) to the nodal velocities.

A particular property shared by the methods is conservation. Since both the α_j's and the ϕ_{ki}'s are a partition of the unit function, summation of the first of (2.14) or (3.3) gives

$$\int_{s_0}^{s_{N+1}} [v_t + f(v)_x] dx = 0 \ , \tag{3.16}$$

from which we may deduce that

$$\frac{d}{dt}\left[\int_{s_0}^{s_{N+1}} v \ dx\right] = -\left[f(v)\right]_{s_0}^{s_{N+1}} \tag{3.17}$$

and consequently that $\int_{s_0}^{s_{N+1}} v \ dx$ is constant in time apart from boundary effects.

Boundary conditions may be imposed locally on the elements adjacent to boundaries as in (Baines,1985). In the case of a Neumann condition at the end $j = 0$ we have

$$\dot{s}_0 = 0 \tag{3.18}$$

which, in conjunction with the second of (3.9), gives

$$\dot{a}_0 = w_{\frac{1}{2},1} \tag{3.19}$$

and hence the motion of the boundary node.

If the boundary condition at the end $j = 0$ is Dirichlet then, because we cannot impose both

$$\dot{s}_0 = 0, \quad \dot{a}_0 = 0 \tag{3.20}$$

simultaneously and preserve the projection, a special constrained projection has to be carried out in the end element. The result is that

$$w_{\frac{1}{2},1} = 0, \quad w_{\frac{1}{2},2} = 3\, b_{\frac{1}{2},2}/(s_1 - s_0) \ , \tag{3.21}$$

where $b_{\frac{1}{2},2}$ is the second of (3.7) for the end element. This is consistent with (3.20).

4. LOCAL ELEMENT MOTION

We now show that the motion of a local segment of the approximating piecewise linear function may be obtained entirely from the local projection step (3.4).

Let V_k be the velocity of the mid-point of the segment in the direction perpendicular to the segment and let θ_k be the angle between the segment and the x-axis, as shown in Figure 3. Then the equations (3.9) can be written in the form

$$\left.\begin{aligned} \dot{a}_{k-\frac{1}{2}} \cos\theta_k - \dot{s}_{k-\frac{1}{2}} \sin\theta_k &= w_{k1} \cos\theta_k \\ \dot{a}_{k+\frac{1}{2}} \cos\theta_k - \dot{s}_{k+\frac{1}{2}} \sin\theta_k &= w_{k2} \cos\theta_k \end{aligned}\right\} \forall k \tag{4.1}$$

where nodes $k-\frac{1}{2}$, $k+\frac{1}{2}$ are the ends of element k, as in Figure 3. The left hand sides of (4.1) are the velocities $V_{k-\frac{1}{2}}$, $V_{k+\frac{1}{2}}$ (due to the motion of the single element k only) of the ends of the element k in Figure 3 at right angles to the element. (The full velocity of a node will be a combination of two such element end velocities from adjacent segments).

Let
$$m_k = \frac{a_{k+\frac{1}{2}} - a_{k-\frac{1}{2}}}{s_{k+\frac{1}{2}} - s_{k-\frac{1}{2}}} \ . \tag{4.2}$$

Then these end velocities may be written

$$\begin{bmatrix} V_{k-\frac{1}{2}} \\ V_{k+\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} w_{k1} \\ w_{k2} \end{bmatrix} \cos\theta_k = \frac{1}{\sqrt{(1+m_k^2)}} C_k^{-1} \begin{bmatrix} b_{k1} \\ b_{k2} \end{bmatrix} = \frac{1}{\sqrt{(1+m_k^2)}} C_k^{-1} \begin{bmatrix} \langle \phi_{k1}, f(v)_x \rangle \\ \langle \phi_{k2}, f(v)_x \rangle \end{bmatrix} \tag{4.3}$$

using (3.4) and (3.7). Since $[1,1]^T$ is an eigenvector of the symmetric matrix C_k with eigenvalue $\frac{1}{2}(s_{k+\frac{1}{2}} - s_{k-\frac{1}{2}}) = \frac{1}{2}\Delta s_k$, we obtain

$$[1,1] \begin{bmatrix} V_{k-\frac{1}{2}} \\ V_{k+\frac{1}{2}} \end{bmatrix} = \frac{-1}{\sqrt{(1+m_k^2)}} \frac{2}{\Delta s_k} [1,1] \begin{bmatrix} \langle \phi_{k1}, f(v)_x \rangle \\ \langle \phi_{k2}, f(v)_x \rangle \end{bmatrix}$$

$$= \frac{-1}{\sqrt{(1+m_k^2)}} \frac{2}{\Delta s_k} \langle 1, f(v)_x \rangle = \frac{-2}{\Delta s_k \sqrt{(1+m_k^2)}} \int_{s_{k-\frac{1}{2}}}^{s_{k+\frac{1}{2}}} f(v)_x dx \tag{4.4}$$

Hence $$V_k = \tfrac{1}{2}(V_{k-\frac{1}{2}} + V_{k+\frac{1}{2}}) = \frac{-1}{\sqrt{(1+m_k^2)}} \frac{1}{\Delta s_k} \{(f(v))_{k+\frac{1}{2}} - (f(v))_{k-\frac{1}{2}}\} \tag{4.5}$$

which gives the normal velocity of the mid-point of the segment in Figure 3. Putting $\Delta f_k = \{(f(v))_{k+\frac{1}{2}} - (f(v))_{k-\frac{1}{2}}\}$ this may be written in the alternative forms

$$V_k = -\frac{\Delta f_k}{\Delta s_k} \cos\theta_k = -\frac{\Delta f_k}{PQ} = -\frac{\Delta f_k}{\Delta v_k} \sin\theta_k \tag{4.6}$$

where $\Delta v_k = \Delta a_k = a_{k+\frac{1}{2}} - a_{k-\frac{1}{2}}$ and PQ are as shown in Figure 3. Thus we have the important result that the speed of the mid-point of the segment in Figure 3 in the direction normal to the segment is consistent with the local average wave speed $\Delta f_k / \Delta v_k$ in the element.

Subtracting pairs of the equations (3.9) we obtain another important result, namely,

$$\dot{a}_{k+\frac{1}{2}} - \dot{a}_{k-\frac{1}{2}} - m_k(\dot{s}_{k+\frac{1}{2}} - \dot{s}_{k-\frac{1}{2}}) = w_{k2} - w_{k1} \tag{4.7}$$

or, using (4.2),

$$\frac{dm_k}{dt} = \frac{1}{\Delta s_k} [-1\ 1] \underline{w}_k \tag{4.8}$$

Since $[-1\ 1]$ is also an eigenvector of C_k with eigenvalue $\frac{1}{6}\Delta s_k$ we obtain

$$\frac{dm_k}{dt} = -\frac{6}{[\Delta s_k]^2}[-1\ 1]\begin{bmatrix}\langle\phi_{k1}, f(v)_x\rangle \\ \langle\phi_{k2}, f(v)_x\rangle\end{bmatrix} = -\frac{6}{[\Delta s_k]^2}\int_{s_{k-\frac{1}{2}}}^{s_{k+\frac{1}{2}}}(\phi_{k2}-\phi_{k1})f(v)_x dx \qquad (4.9)$$

This leads to the alternative forms

$$\frac{dm_k}{dt} = \frac{12}{[\Delta s_k]^2}(\hat{f}-\bar{f}) = f_{xx}(\eta_k) = m_k^2 f''(\eta) \qquad (4.10)$$

where

$$\hat{f} = \frac{1}{\Delta s_k}\int_{s_{k-\frac{1}{2}}}^{s_{k+\frac{1}{2}}} f(v)dx, \quad \bar{f} = \tfrac{1}{2}((f(v))_{k-\frac{1}{2}} + (f(v))_{k+\frac{1}{2}})$$

$$\text{and } \eta_k \in (s_{k-\frac{1}{2}}, s_{k+\frac{1}{2}}). \qquad (4.11)$$

Hence we have the result that the rate of change of the slope of the solution in an element is equal to the second space derivative of the flux function. In other words the solution segment rotates in response to the local convexity of f. Another form of this result is

$$\frac{d\theta_k}{dt} = \frac{m_k^2}{1+m_k^2} f''(\eta_k). \qquad (4.12)$$

As the segments move the intersections (nodes) also move, giving the nodal velocities. Note that the segments have lengths that vary with time. The movement of a node is thus the locus of the intersection of adjacent elements (see Figure 4).

The results (4.6) and (4.10) show clearly the element behaviour in terms of the flux function.

5. SUMMARY AND PARALLELISM

We give now a summary of the local elementwise method which is complete provided that none of the matrices involved are singular. We then go on to discuss the treatment of such singularities.

Summary

(1) Evaluate $\underline{b}_k = \langle\phi_k, -f(v)_x\rangle$

(2) Solve $C_k\underline{w}_k = \underline{b}_k$ for $\underline{w}_k$ $\quad \forall$ elements k

(3) Pair off the w's nodewise

(4) Solve $M_j \dot{\underline{y}}_j = \underline{w}_k$ for $\dot{\underline{y}}_j = [\dot{a}_j, \dot{s}_j]^T$ $\forall$ nodes j.

Note that the method involves the inversion of 2x2 matrices only: little storage is required and the algorithm can be run on a small micro computer.

The method may break down if any of the matrices C_k or M_j are singular. Singularity of M_j corresponds to collinearity of nodes, usually described by the term parallelism. This occurs when

$$m_{k-\frac{1}{2}} = m_{k+\frac{1}{2}} \tag{5.1}$$

in (3.12). Singularity of C_k arises only if $\Delta s_k = 0$ (see (3.6)).

Consider singularity of M_j first. In this case we find that we obtain inconsistent solutions of the pairs of equations (3.10). As a result we can no longer solve equation (3.10) locally. The remedy is to temporarily fix any parallel nodes, solve over a patch consisting of these nodes and their neighbours and relocate the parallel nodes in some averaged way as in (Wathen & Baines,1985; Baines,1985). It is convenient to return to the α,β basis by combining equations (3.4) in staggered pairs. Then in the event of parallelism at a single node we retain the combination of these equations corresponding to the basis function α_j and replace the second (corresponding to β_j) by

$$\dot{s}_j^* = 0 \ . \tag{5.2}$$

where the * refers to the special solution under construction.
This gives

$$\frac{1}{6}(\Delta_{j-\frac{1}{2}}s)w_{j-\frac{1}{2},1} + \frac{1}{3}(\Delta_{j-\frac{1}{2}}s)w_{j-\frac{1}{2},2} + \frac{1}{3}(\Delta_{j+\frac{1}{2}}s)w_{j+\frac{1}{2},1} + \frac{1}{6}(\Delta_{j+\frac{1}{2}}s)w_{j+\frac{1}{2},2}$$

$$= b_{j-\frac{1}{2},2} + b_{j+\frac{1}{2},1} \tag{5.3}$$

Since now $\dot{s}_j^* = 0$ we have $\dot{a}_j^* = w_{j-\frac{1}{2},2} = w_{j+\frac{1}{2},1}$ from (3.9) in the reduced problem which yields

$$\dot{a}_j^* = \frac{b_{j-\frac{1}{2},2} + b_{j+\frac{1}{2},1} - \frac{1}{6}\{(\Delta_{j-\frac{1}{2}}s)w_{j-\frac{1}{2},1} + (\Delta_{j+\frac{1}{2}}s)w_{j+\frac{1}{2},2}\}}{\frac{1}{3}(\Delta_{j-\frac{1}{2}}s + \Delta_{j+\frac{1}{2}}s)} \tag{5.4}$$

$$\dot{s}_j^* = 0$$

as the solution for the modified system with the parallel node fixed.

The null space of the singular matrix M_j is spanned by the vector $[m,1]^T$ (where $m_{j-\frac{1}{2}} = m_{j+\frac{1}{2}} = m$) and an appropriate multiple of this vector may be added to satisfy an externally imposed averaged velocity or position.

If several nodes become parallel simultaneously a number of equations of the type (5.3) will occur and it may be necessary to solve a tri-diagonal system if the nodes are adjacent to one another. This can be avoided by a different approach which leads to an explicit solution for the modified system.

In the approach above the system of equations

$$C\underline{w} = \underline{b} \tag{5.5}$$

or

$$CM\underline{y} = \underline{b} \tag{5.6}$$

with C given by (3.14) is pre-multiplied by M^T to give the α,β inner-product equations. Suppose that we pre-multiply this equation by M^TC^{-1} (assuming C non-singular). Then we obtain different linear combinations of the equations (5.6) but they still have the property (through the presence of the M^T matrix) that in the event of parallelism the equations remain consistent. We therefore again delete one of each of the offending pairs of equations by removing the second column of each relevant M_j (see (3.14)). Call the resulting matrix M^*. Then, if we also impose $\dot{s}^* = 0$ as before we end up solving

$$M^{*T}C^{-1}CM^*\dot{\underline{y}}^* = M^{*T}C^{-1}\underline{b} \quad , \tag{5.7}$$

where

$$\dot{\underline{y}}^* = (\ldots,\dot{a}_j^*,\dot{s}_j^*,\ldots), \tag{5.8}$$

which gives

$$\dot{\underline{y}}^* = (M^{*T}M^*)^{-1}M^{*T}C^{-1}\underline{b} \quad , \tag{5.9}$$

where $M^{*T}M^*$ is a block diagonal non-singular matrix. This is an explicit particular solution of the MFE equations in the event of (possibly multiple) parallelism, to which the null space of the matrix M can be added as before to locate the parallel nodes as desired.

Parallelism is a rather rare occurrence and is usually associated with the curvature of the solution changing sign during the evolution, or perhaps with the evolution to a steady state.

We go now to consider singularity of C_k. Since this type of singularity is linked with node overtaking as a consequence of time stepping we discuss it in the next section in association with time integration.

6. TIME STEPPING AND NODE OVERTAKING

The MFE method is semi-discrete and gives rise to ordinary differential equations in time which require integration to obtain the full solution.

We have already seen that in the case of the inviscid Burgers' equation the nodes move along characteristics, while for the general scalar hyperbolic law the least squares best fit to the exact solution is carried asymptotically for small time steps. Approximate time stepping will degrade the latter property if the time step is too large, while the former property will be lost for more general flux functions.

It has been found that using the Euler explicit forward difference method is sufficient in the examples tried so far by Wathen, 1984, Johnson, 1984. In no case do implicit methods give any advantage. We therefore use

$$\begin{pmatrix} a_j^{n+1} \\ s_j^{n+1} \end{pmatrix} = \underline{y}_j^{n+1} = \underline{y}_j^n + \Delta t\, M_j^{-1}\, \underline{w}_j = \begin{pmatrix} a_j^n \\ s_j^n \end{pmatrix} + \Delta t\, M_j^{-1}\, \underline{w}_j \qquad (6.1)$$

(see (3.10)).

As far as the choice of Δt is concerned we still require an accuracy criterion. Algorithms for accuracy are not well developed but in view of the simplicity of the method we can afford to be generous in taking a trial and error approach. A possible algorithm compares the result of one MFE step with that of two half steps and continues halving the step until the difference between the two results is acceptable, as in (Baines, 1985).

A major difficulty with time stepping is that the nodes may overtake one another if Δt is not small enough, and this gives a restriction on the time step. The restriction is that Δt should not be greater than the smallest time $(\Delta t)_0$ taken for any node to catch up with its neighbour. This is easily calculated if Euler time stepping is used. However, time accuracy is also lost so Δt should not be too large.

In problems whose solution is expected to be smooth we expect nodes to merge when they overtake. Because of time inaccuracy this may not occur in practice and we have found that a practical time step is obtained by taking half $(\Delta t)_0$. For hyperbolic problems which admit shocks however we expect discontinuities to form and we can take advantage of node overtaking to model shocks in an effective way.

If it is not known in advance whether a shock is forming or not we can test the slope of the segment which is tending to zero. If this remains finite then there is no shock forming.

As the separation of nodes goes to zero the element segment becomes vertical (parallel to the u-axis) and from (4.6) with $\theta_k \to \pi/2$ the normal velocity of the mid-point of the segment tends smoothly to the shock speed, at least in the semi-discrete case. The shock speed may then be "frozen", i.e. imposed on both nodes of the shocked element: this acts as an internal boundary condition and the solution on the adjacent elements to left and right may proceed separately. The procedure is also feasible when nodes run into shocks or when shocks overtake shocks, in which cases a node can be deleted.

From equation (4.12) we see that the angular speed of the segment is non-zero when the shock forms $(m_k \to \infty)$ so that there is a change of state at this instant. The manner in which the shock forms in the MFE method is consistent with the (Oleinik,1958) entropy condition in the semi-discrete case. Moreover this is also true for expansions.

Accuracy in space is of course determined by the number of nodes used to represent the solution and this is decided when carrying out the projection of the initial data into the piecewise linear space. Less obviously it has been found that it is crucial how these nodes are distributed in space, equidistribution of $(u'')^{\frac{1}{2}}$ being an effective choice: see (Herbst,1982)

We show numerical results for two scalar problems, the inviscid Burgers' equation and the Buckley-Leverett equation. These are shown in Figure 5. The first gives the exact solution for this (convex) flux function while the second gives an approximate solution (in the case of a non-convex flux function).

7. EXTENSIONS TO 1-D SYSTEMS

In extending the above ideas to systems of conservation laws we are faced with an immediate decision. Should we work with separate nodal coefficients and a common mesh or should we give each component of the system its own mesh with individual nodal coefficients and co-ordinates?

Where discontinuous features are expected to occur simultaneously for all components of the system, as in the Euler equations, there is an argument for using a common mesh. However the useful algebraic structure of previous sections is only preserved when each component is given its own mesh, (but see below).

One possible strategy is to use a single mesh whose movement is determined by a single preferred component of the system as in (Baines & Wathen,1985). For example, in the Euler equations we might choose the density as the most significant component and use that to drive the nodes. The remaining components are then determined on a prescribed moving mesh.

The difficulty here is that the flux function in the density equation is the momentum, which is itself piecewise linear on the same mesh as the density. By (4.10) it follows that the slopes of the element segments of the density do not change with time, which is much too restrictive.

We therefore consider here a model which uses a different mesh for each component of the system. As a result of the independence of each mesh we can easily solve the MFE equations in the manner of earlier sections, once the right hand sides have been set up. Thus the main new feature is the quadrature in equation (3.7) which links the components of the system through the evaluation of $\underset{\sim}{f}(\underset{\sim}{v})$. (In one dimension the elements can be sub-divided suitably with an elementary quadrature over each sub-element.) The only other difficulty arises in the shock modelling. A feature of a shock in gasdynamics is that components shock simultaneously and this is not guaranteed in the numerical method. An additional device is therefore needed in general to ensure that when a shock occurs it is simultaneous in the appropriate components.

We can devise yet another algorithm which uses only a single grid by mapping the result of the separate mesh method above onto a single mesh. This can be done by a least squares projection which replaces (3.10) by

$$M_j^T M_j \dot{\underline{y}}_j = M_j^T \underline{w}_j \qquad (7.1)$$

as in (Baines, 1985).

We give numerical results for the well known shock tube problem used as a basis for comparisons by (Sod,4978). These are shown in Figs. 6a and 6b.

8. THE METHOD IN HIGHER DIMENSIONS

One of the most promising aspects of the MFE method is its straightforward generalisation to higher dimensions. We consider here the two dimensional scalar conservation law

$$u_t + f_x + g_y = 0 \,. \tag{8.1}$$

We again approximate the function u by a piecewise linear function v given by equation (2.2), where the basis functions are now two dimensional "pyramid" functions as shown in Figure 7. The time derivative v_t of the function v now however belongs to a subspace $S_{\alpha\beta}$ of the space of piecewise linear discontinuous functions S_ϕ on the two dimensional mesh, because it has to correspond to a continuous v and not all members of S_ϕ do so. This is because there are generally more elements surrounding a node than there are nodes at vertices of an element (see Figure 8).

From a local elementwise point of view we can readily calculate the $\underline{b}_k$ in the two dimensional generalisation of equation (3.4), which is now a 3x3 system. But in order to obtain the nodal velocities, which are evaluated from the union of equations (3.10), we require $\underline{w}$, the union of the $\underline{w}_k$'s, to lie in the range space of M (see (3.14)), which is now rectangular.

It is therefore necessary to constrain the projection which leads to the vector $\underline{w}$. The result of this constrained projection (as in Baines, 1985) is that $\underline{w}$ satisfies

$$M^T C \underline{w} = M^T \underline{b} \tag{8.2}$$

which, since

$$M^T \underline{b} = \underline{g} \tag{8.3}$$

leads to the familiar form

$$M^T C M \dot{\underline{y}} = \underline{g} \,, \tag{8.4}$$

i.e. it is equivalent to the standard method (see for example (Wathen & Baines, 1985).

Hence the decomposition

$$A = M^T C M \tag{8.5}$$

holds in higher dimensions, but although the matrix C is square 2x2 block diagonal (using elementwise numbering) the matrix M is rectangular. For example, in two dimensions using nodewise numbering M takes the form N where

$$N = \text{diag}\,\{N_j\} \qquad N_j = \begin{bmatrix} 1 & -m_{j1} & -n_{j1} \\ 1 & -m_{j2} & -n_{j2} \\ 1 & -m_{jI} & -n_{jI} \end{bmatrix} \tag{8.6}$$

where m_{ji}, n_{ji} are the slopes v_x, v_y of the function v in the x and y directions within the element i(for $i = 1,2,\ldots,I$).

Using a permutation matrix Q to map between the elementwise numbering and the nodewise numbering we obtain the consistent decomposition

$$A = N^T Q^T C Q N \quad , \tag{8.7}$$

where $M = QN$.

It has been shown by Wathen (1984) using this decomposition that if D is the matrix consisting of diagonal blocks of the MFE matrix A then the eigenvalues of the matrix $D^{-1}A$ lie in the real interval

$$[\tfrac{1}{2},\ 1 + \frac{d}{2}] \quad , \tag{8.8}$$

where d is the number of dimensions.* There is therefore every reason for using a conjugate gradient method with D^{-1} as preconditioner to invert the MFE matrix A.

It is interesting to note that the MFE equations (8.4) also arise from first carrying out an unconstrained projection for $\underline{w}$ into the space S_ϕ and then doing a least squares projection of the equation

$$N\dot{\underline{y}} = \underline{w} \tag{8.9}$$

weighted by the matrix $C^{\frac{1}{2}}Q$ as in (Baines,1985).

---

* The result also holds for the usual fixed finite element consistent mass matrix (see Wathen, 1985).

By a calculation similar to that in Section 4 we again find as in (Baines,1985) that, when $\underline{w}$ lies in the range space of M, the velocity V_k of the centroid of the triangular segment of the solution in the direction perpendicular to the segment is given by equation (4.6), where now

$$\Delta f_k = \text{the outward flux of } (f,g) \text{ across the element boundary} , \tag{8.10}$$

$$\Delta v_k = \text{the vertical profile of the element } k, \tag{8.11}$$

and the angle θ_k is given by

$$\tan \theta_k = |\underline{\nabla} v| . \tag{8.12}$$

There is also a corresponding result on the segment rotation similar to equation (4.10) but apart from noting that the convexity of the flux function is again involved we do not set down the details.

9. SINGULARITIES IN HIGHER DIMENSIONS

Because of the decomposition (8.5) the singularities of the method in higher dimensions correspond to those in one dimension, namely, they are singularity of C when the element size goes to zero, and rank deficiency of M when nodes become coplanar : see (Wathen & Baines, 1985).

In the latter case of parallelism the remedy is the same as in one dimension, namely to remove the (linearly dependent) equations causing the parallelism and to solve a reduced system, adding a suitable multiple of the null space at the end. Some care is required with the larger blocks occurring in higher dimensions: it is safest to transform the local system to upper triangular form, which avoids problems of ill-conditioning which can arise if an arbitrary equation is omitted.

Singularity of C arises when the size of an element goes to zero as a result of a node running into the opposite side of a triangle. Even though we do not have a proved best fit property in more than one dimension we may still conjecture that this occurrence corresponds to the formation of a shock. Technical difficulties arise in determining how the shocked triangle should move subsequently but we have the following lead.

We recall the result in Section 8 on the normal velocity V_K of the centroid of the triangle segment perpendicular to itself. It can be seen that this velocity tends to the local shock speed as the segment becomes vertical in the sense that the triangle sweeps out "mass" at the correct rate. It appears that we should therefore impose this shock speed on all points of the vertical line through the centroid of the triangle in the subsequent motion with the shock in place. This still allows possible rotation of the triangle about the line through the centroid which must be determined from the solution of the rest of the system. We do not go into the details here.

We show results for one problem, a generalisation of the Buckley-Leverett equation in one dimension. The governing equation is

$$u_t + \underline{\nabla} \cdot \left[\frac{u^2}{u^2 + \frac{1}{2}(1-u)^2} \right] = 0 \qquad (8.13)$$

and Figure 9 shows both the initial data and the solution at a later time. It is worth noting that even though the triangles of the mesh become highly distorted there is no ill-conditioning of the MFE matrix.

10. CONCLUSION

In this report we have shown that the MFE method in one dimension is a local method which gives the motion of individual element segments by means of a local straight line best fit. Moreover the segment movement can be found in terms of elementary properties of the flux function. As a result non-linear problems involving scalar conservation laws can be solved simply on a small computer. Shocks are particularly well resolved by the method, the entropy rules being respected. Two ways of dealing with systems are described.

In higher dimensions the local nature of the method is modified in that an additional least squares fit is required to map the local element motion onto the nodewise velocities. The technical problem of dealing with shocks is assisted by the consistency of the element segment velocity with a local wave speed.

11. ACKNOWLEDGEMENTS

I should like to thank Professor K.W. Morton and Dr. A.J. Wathen for many illuminating discussions.

12. REFERENCES

BAINES, M.J. (1985) "On Approximate Solutions of Time-Dependent Partial Differential Equations by the Moving Finite Element Method", Numerical Analysis Report 1/85, Department of Mathematics, University of Reading.

BAINES, M.J. and WATHEN, A.J. (1985) "Moving Finite Element Modelling of Compressible Flow", Numerical Analysis Report 4/85, Department of Mathematics, University of Reading.

HERBST, B.M. (1982) "Moving Finite Element Methods for the Solution of Evolution Equations", Ph.D. Thesis, University of the Orange Free State, South Africa.

JOHNSON, I.W. (1984) "The Moving Finite Element Method for the Viscous Burgers' Equation", Numerical Analysis Report 3/84, Department of Mathematics, University of Reading.

LYNCH, D.R. (1982) "Unified Approach to Simulation on Deforming Elements with Application to Phase Change Problems", J. Comput. Phys. 47, 387-411.

MILLER, K. (1981) "Moving Finite Elements, Part I" (with R.N. Miller), SIAM J. Numer. Anal. 18, 1019-1032: "Moving Finite Elements, Part II", SIAM J. Numer. Anal. 18, 1033-1057.

MORTON, K.W. (1982) Private Communication. See also (Wathen, 1984).

OLEINIK, O.A. (1957) "Discontinuous Solutions of Nonlinear Differential Equations", Amer. Math. Soc. Trans. Ser. 2, No. 26, 95-172.

SOD, G.A. "A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws", J. Comput. Phys. 27, 1-31.

WATHEN, A.J. (1984) "Moving Finite Elements and Oil Reservoir Modelling", Ph.D. Thesis, University of Reading.

WATHEN, A.J. (1985) To appear.

WATHEN, A.J. and BAINES, M.J. (1985) "On the Structure of the Moving Finite Element Equations", IMAJNA (to appear).

Figure Captions

Figure 1 Basis functions (a) α_j and (b) β_j

Figure 2 Basis functions ϕ_{k1} and ϕ_{k2} (a) on an element, (b) around a node.

Figure 3 Local segment angle and velocities

Figure 4 Nodal movement as locus of segment intersection

Figure 5 MFE solutions for the scalar non-linear equations

$$u_t + f_x = 0 \quad .$$

(a) $f(u) = \frac{1}{2}u^2$ (Burgers')

(b) $f(u) = \dfrac{u^2}{u^2 + \frac{1}{2}(1-u)^2}$ (Buckley- Leverett)

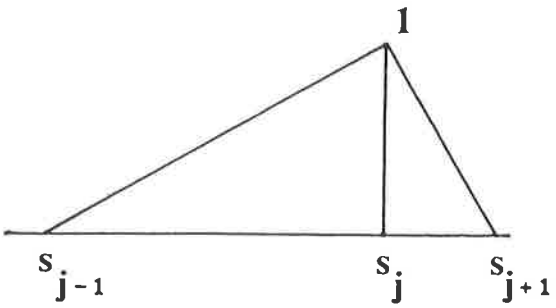
with initial data as in (Wathen & Baines, 1985).

Figure 6 MFE solutions for the Sod shock tube problem at $t = 0.144$ from initial data at $t = 0.1$ as in (Baines and Wathen, 1985).

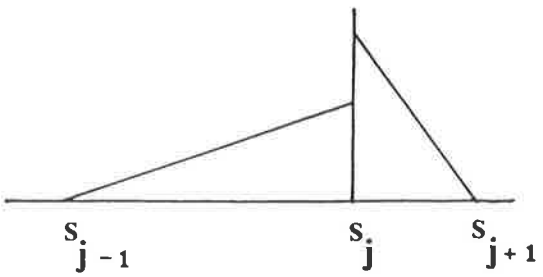
Figure 7 Basis functions (a) α_j and (b) β_j in two dimensions.

Figure 8 Nodes/elements in two dimensions.

Figure 9 MFE solution for the two dimensional equation (8.13): initial data and solution (isoplots) as in (Wathen, 1984).



(a)



(b)

FIG. 1

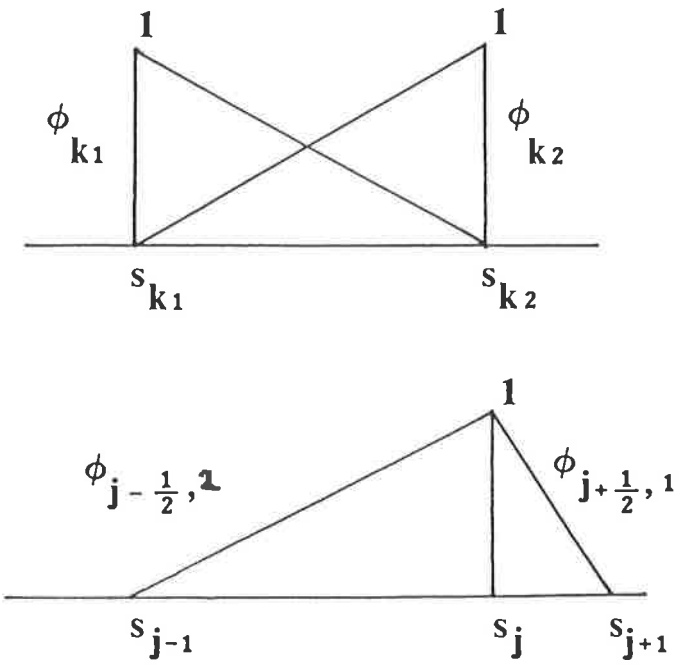


FIG. 2

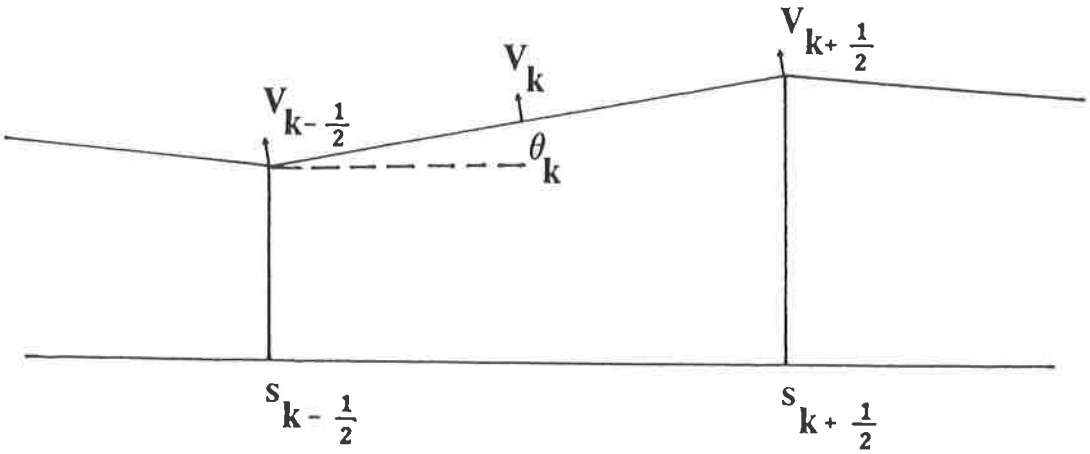


FIG. 3

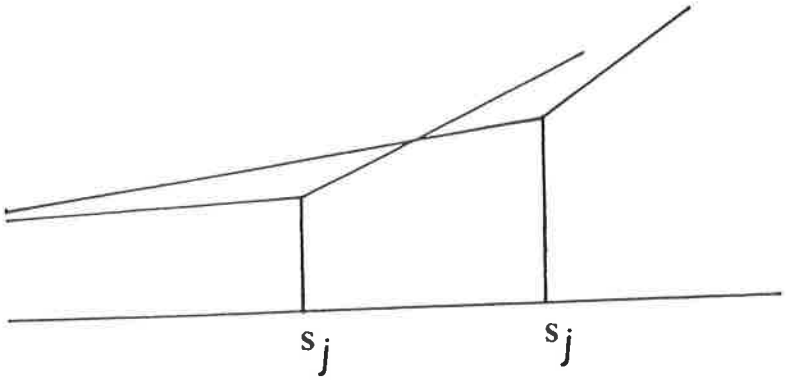
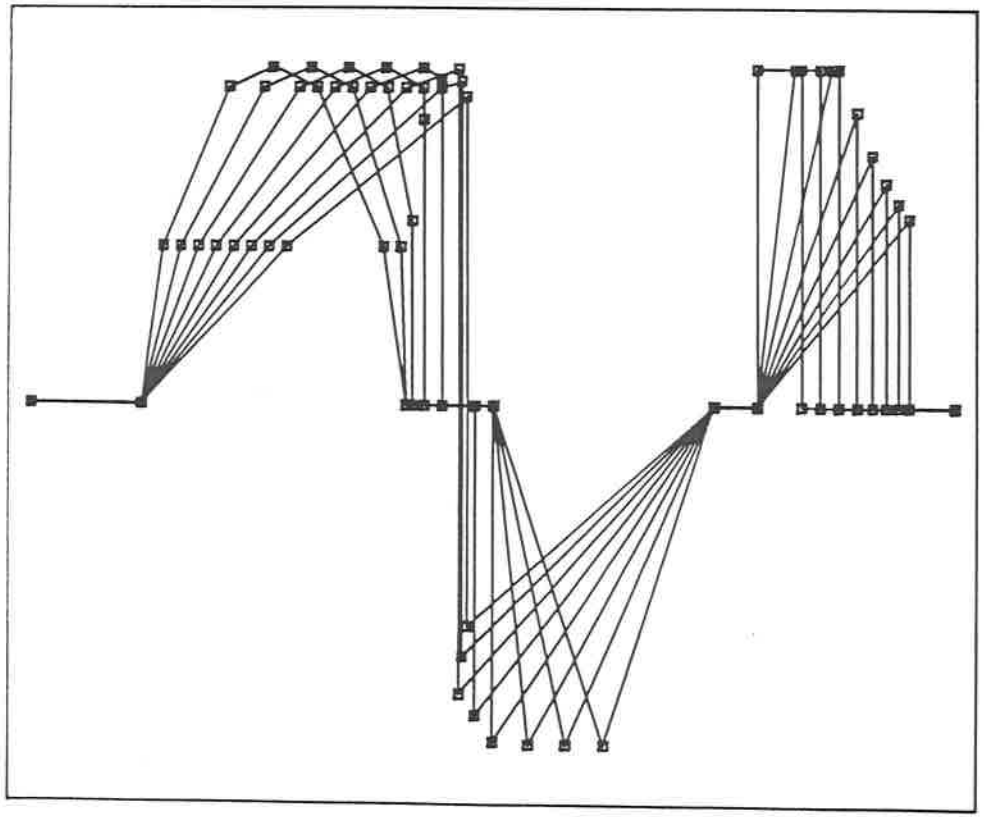
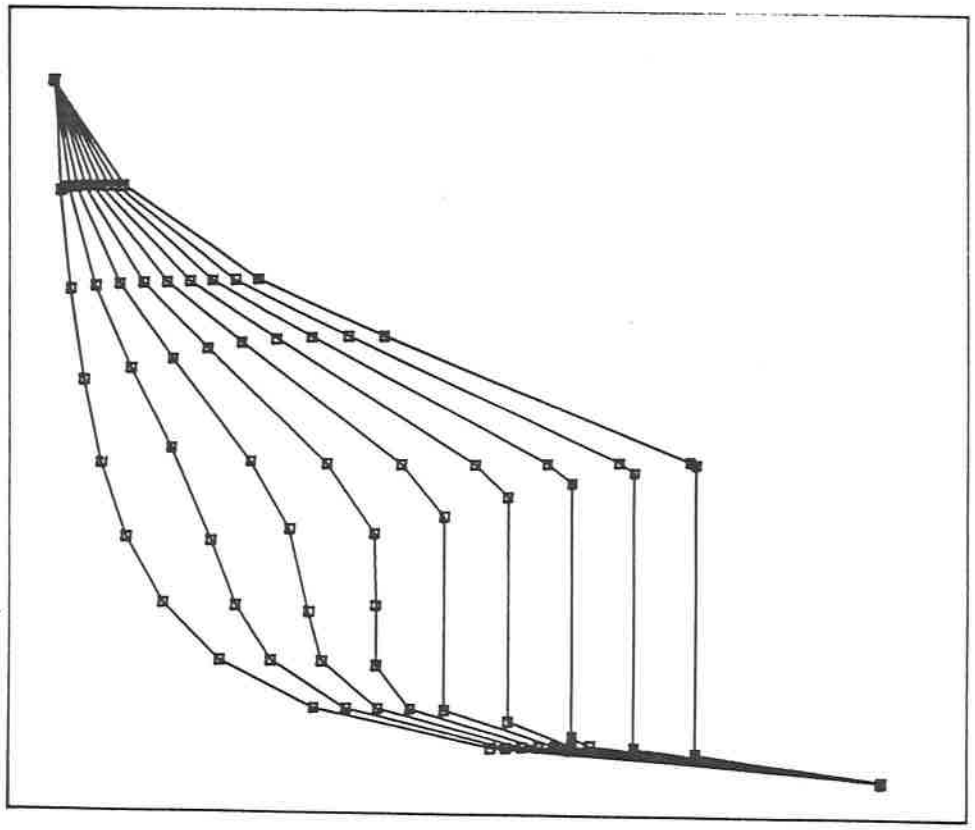


FIG. 4



(a)



(b)

fig. 5

CALCOMP EUROPE PRINTED IN U.K.

CS1917

8450 AS

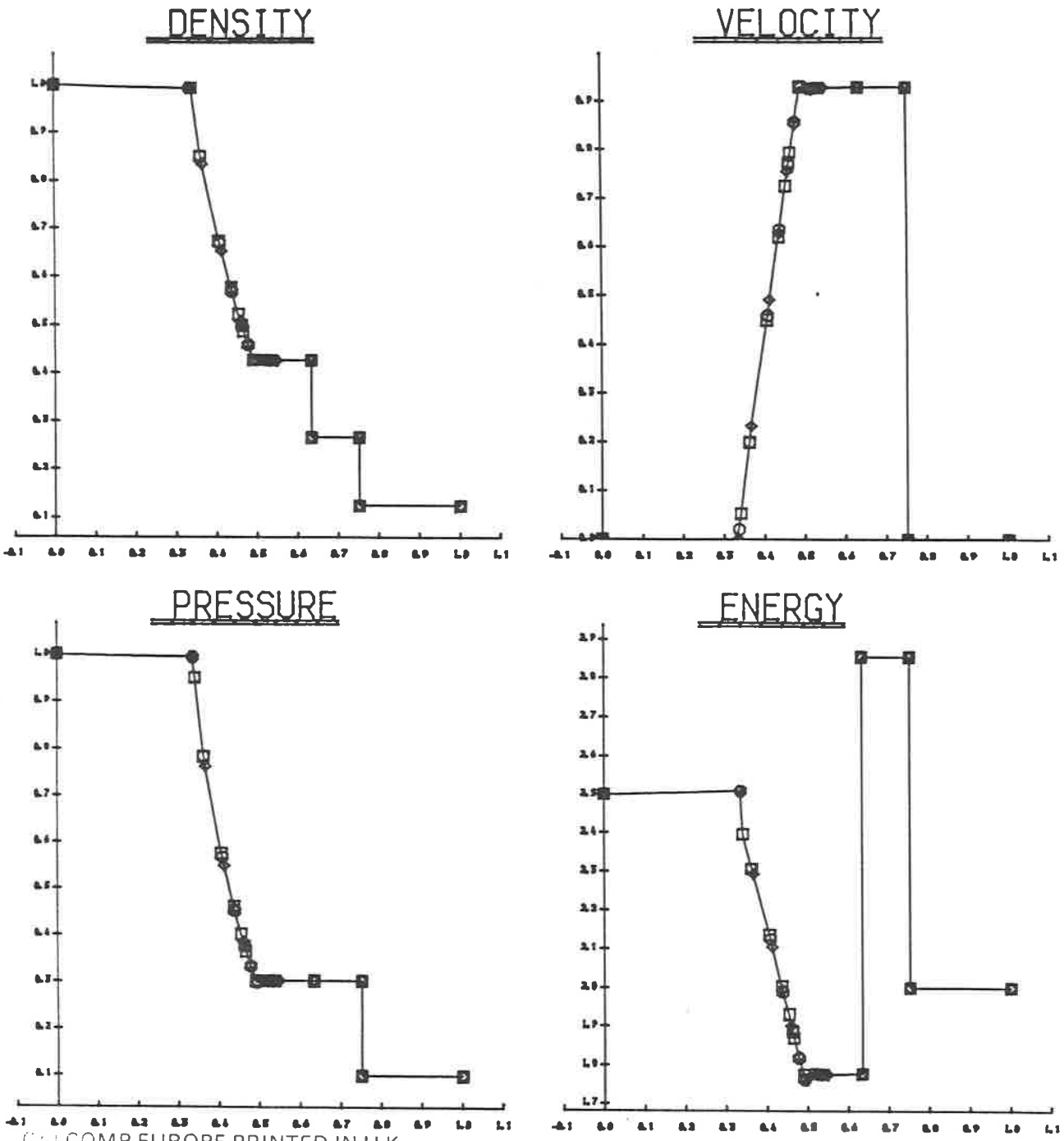


fig. 6a

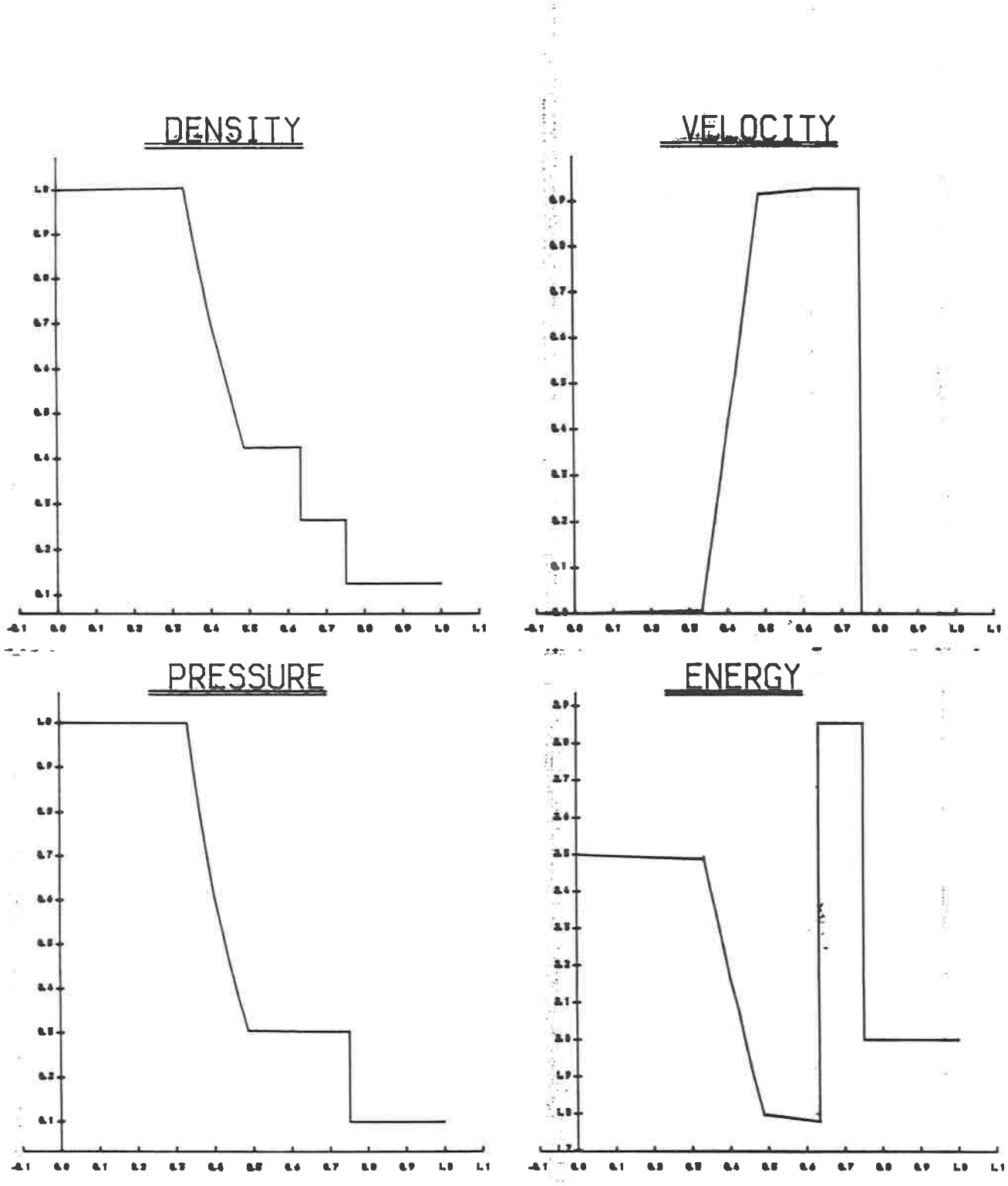
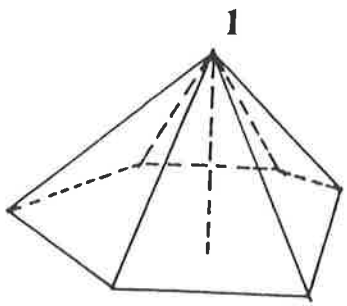
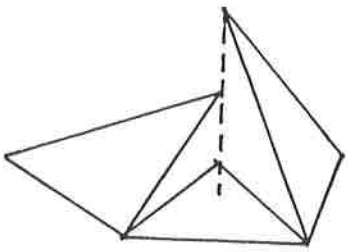


Fig. 6b



(a)



(b)

FIG. 7

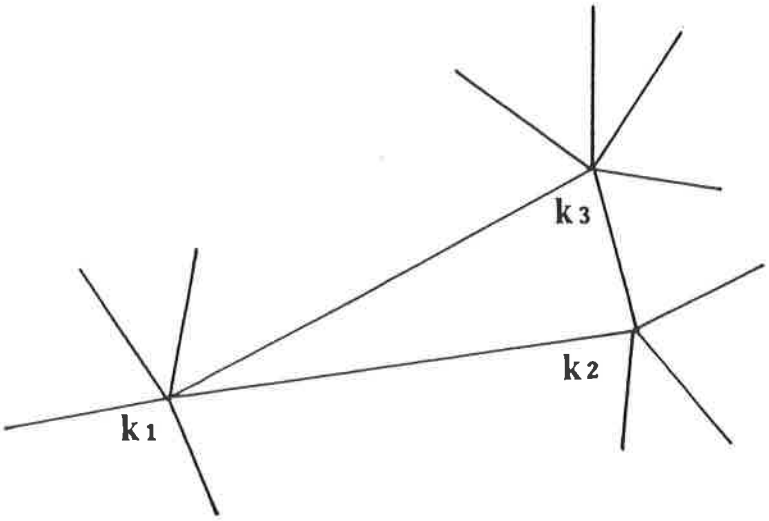


FIG. 8

fig.9