

MOVING FINITE ELEMENTS:
PROJECTIONS AND CONSTRAINTS

M.J. BAINES

NUMERICAL ANALYSIS REPORT 15/85

The work reported here forms part of the research programme of the
Institute for Computational Fluid Dynamics at Oxford and Reading.

MOVING FINITE ELEMENTS:
PROJECTIONS AND CONSTRAINTS

M.J. BAINES

INTRODUCTION

When the Moving Finite Element (MFE) method was first introduced and implemented by Miller & Miller (1981) and Miller (1981) on parabolic problems, penalty functions were used to inhibit the onset of singularities. This use of penalty functions was done in a completely implicit manner, the closer the solution to singularity the higher the penalty to be added. The resulting method needs a stiff solver to carry through the time stepping.

With the realisation that the MFE method without penalty functions can be regarded as a local method and that explicit time stepping is sufficient in this case (Wathen & Baines (1985), Baines (1985a,b)) the possibility arose of using MFE as a fast economic solver for hyperbolic problems. Nevertheless there do appear to be situations where node overtaking limits time steps in an impractical way and constraints may be envisaged as a way of speeding up the solution.

In this report we discuss the implementation of constraints in the MFE method in an explicit manner, giving generalised forms of the MFE equations which preserve the matrix structure observed by Wathen & Baines (1985). One possible constraint is the fixing of nodes, yielding the Galerkin fixed rule (FFE) method. We begin then with a section on projections describing the interconnection of the MFE and FFE methods. This is followed by the section on constraints. The two sections are largely independent however.

1. PROJECTIONS

Consider finite element methods for obtaining an approximate solution v to the solution u of the equation

$$u_t = Lu \tag{1.1}$$

(i) The semi-discrete Finite Element Method with fixed nodes (FFE) may be taken as consisting of two steps. First the residual

$$\|v_t - Lv\|_2 \tag{1.2}$$

is minimised over the coefficients \hat{a}_j in the approximation v , i.e. in

$$v = \sum a_j \alpha_j, \quad v_t = \sum \dot{a}_j \alpha_j, \tag{1.3}$$

where α_j is as shown in Fig. 1(a) and $\dot{a}_j = \frac{\partial a_j}{\partial t}$. This leads to the

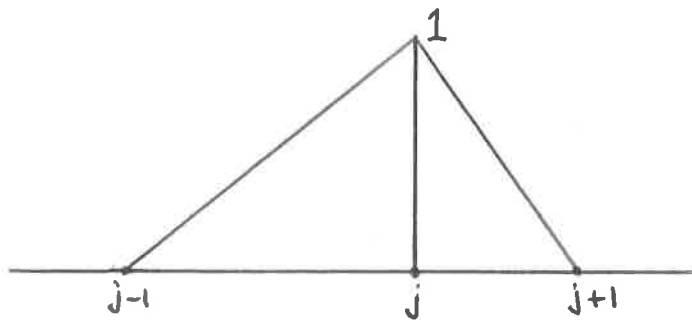


FIG. 1(a) : The α_j function

Galerkin equations

$$\langle \alpha_i, \sum_j \dot{a}_j \alpha_j - Lv \rangle = 0, \tag{1.4}$$

a set of ordinary differential equations (ODE's) in time for a_j .

Secondly, these ODE's are solved by a finite difference time-stepping algorithm.

Equations (1.4) can be written in matrix form as

$$A_\alpha \underline{\dot{a}} = \underline{g}_\alpha \tag{1.5}$$

where $\underline{\dot{a}} = [\dots, \dot{a}_j, \dots]^T$, $\underline{g}_\alpha = [g_i^\alpha]$, $g_i^\alpha = \langle \alpha_i, Lv \rangle$, $A_\alpha = [A_{ij}^\alpha]$
 $A_{ij}^\alpha = \langle \alpha_i, \alpha_j \rangle$. The matrix A_α is symmetric positive definite tridiagonal but its inverse is full. Hence $\dot{a}_i = \{ (A_\alpha^{-1} \underline{g}_\alpha)_i \}$ depends globally on contributions from Lv in all parts of the region.

The minimisation can be regarded as a projection of Lv into the space P_α spanned by the α_j .

(ii) The semi-discrete moving finite element method (MFE) has a corresponding structure. This time the residual (1.2) is minimised over both sets of coefficients \dot{a}_j and \dot{s}_j in the approximation

$$v = \sum a_j \alpha_j \quad v_t = \sum (\dot{a}_j \alpha_j + \dot{s}_j \beta_j) \tag{1.6}$$

where β_j is as shown in Fig. 1(b) and $\dot{s}_j = \frac{\partial s_j}{\partial t}$ (c.f. Fig. 1).

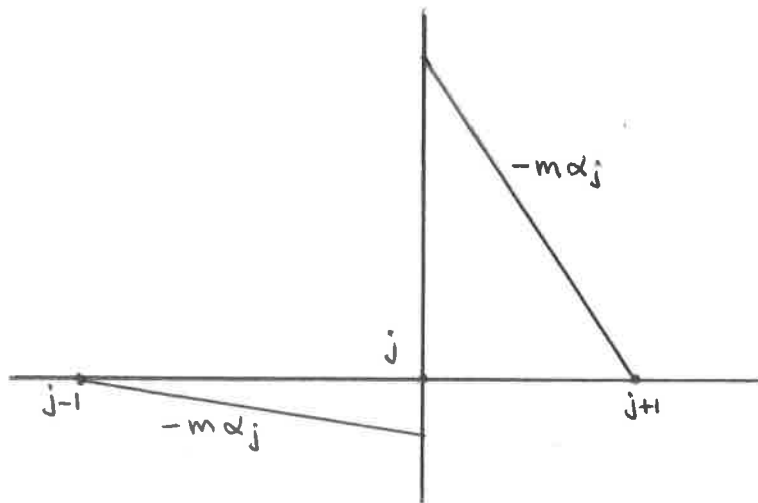


FIG. 1(b) : The β_j function

[The result $\sum_j a_j \alpha_j = \sum_j \dot{s}_j \beta_j$ appears in several places in the MFE literature. (See e.g. Miller, Lynch)]. $\beta_j = -m \alpha_j$ where m is the local slope v_x of the approximate solution. We now have the double system of

Galerkin equations

$$\begin{aligned} \langle \alpha_i, \sum (\dot{a}_j \alpha_j + \dot{s}_j \beta_j) - Lv \rangle &= 0 \\ \langle \beta_i, \sum (\dot{a}_j \alpha_j + \dot{s}_j \beta_j) - Lv \rangle &= 0 \end{aligned} \quad (1.7)$$

a set of ODE's in time for a_j, s_j . These ODE's are again solved by a finite difference time-stepping scheme, usually forward Euler.

Equations (1.7) can be written in matrix form as

$$A \dot{\underline{y}} = \underline{g} \quad (1.8)$$

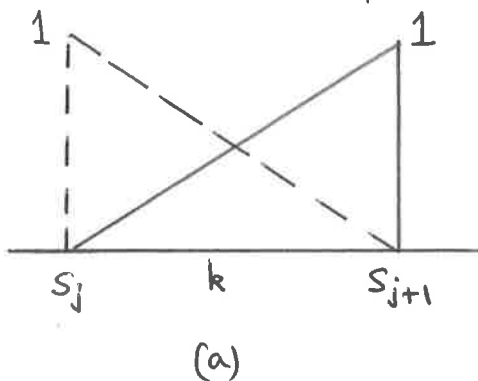
where $\dot{\underline{y}} = \{\dots, \dot{a}_j, \dot{s}_j, \dots\}^T$, $\underline{g} = \{\dots, g_i^\alpha, g_i^\beta, \dots\}$, $g_i^\beta = \langle \beta_i, Lv \rangle$,

$A = [A_{ij}]$, $A_{ij} = \begin{bmatrix} \langle \alpha_i, \alpha_j \rangle & \langle \alpha_i, \beta_j \rangle \\ \langle \beta_i, \alpha_j \rangle & \langle \beta_i, \beta_j \rangle \end{bmatrix}$. The matrix A is symmetric non-

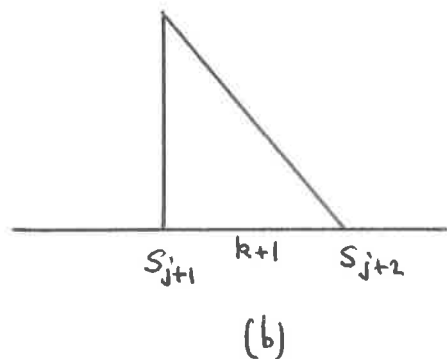
negative definite tridiagonal in 2×2 blocks and its inverse has the same structure. Hence $\dot{y}_i = \{(A^{-1})_i\}$ depends locally only on contributions Lv from elements adjoining the point i .

The minimisation may be regarded as a projection of Lv into the space $P_{\alpha\beta}$ spanned by the α_j and β_j ; since α_j and β_j are time dependent, $P_{\alpha\beta}$ varies with time.

(iii) In one dimension the space $P_{\alpha\beta}$ is also spanned by the basis functions ϕ_k (see Fig. 2). The projection of Lv into $S_{\alpha\beta}$ is the same as the projection of Lv into S_ϕ the space spanned by the ϕ 's. The corresponding



The functions ϕ_{k1} (dotted line) and ϕ_{k2}



The function $\phi_{k+1,1}$

Galerkin equations are

$$\langle \phi_k, \sum \dot{w}_\ell \phi_\ell - Lv \rangle = 0, \tag{1.9}$$

where $v_t = \sum \dot{a}_j \alpha_j + \dot{s}_j \beta_j = \sum \dot{w}_k \phi_k$. (1.10)

Equations (1.9) can be written in matrix form as

$$C_k \dot{w}_k = \underline{b}_k \quad \forall k \tag{1.11}$$

where $\underline{\dot{w}}_k^T = (\dot{w}_{k1}, \dot{w}_{k2})^T$, $\underline{b}_k^T = (b_{k1}, b_{k2})^T$, $b_{km} = \langle \phi_{km}, Lv \rangle$, and

$$C_k = \begin{bmatrix} \langle \phi_{k1}, \phi_{k1} \rangle & \langle \phi_{k1}, \phi_{k2} \rangle \\ \langle \phi_{k2}, \phi_{k1} \rangle & \langle \phi_{k2}, \phi_{k1} \rangle \end{bmatrix} \quad \text{(see Fig. 2). For each } k \text{ equation (1.11)}$$

is a 2x2 system, trivially solved to give $\dot{w}_k = C_k^{-1} \underline{b}_k$.

Combining together equations (1.11), write $C = \text{diag} [C_k]$,

$$\underline{\dot{w}}^T = [\underline{\dot{w}}_k^T], \quad \underline{b}^T = [\underline{b}_k^T], \text{ giving}$$

$$C \underline{\dot{w}} = \underline{b} \tag{1.12}$$

C is symmetry non-negative definite diagonal in 2x2 blocks with the trivial inversion $\underline{\dot{w}} = C^{-1} \underline{b}$.

To obtain \dot{a}_j, \dot{s}_j from \dot{w}_k , use (1.10). Comparison of coefficients gives

$$\left. \begin{aligned} \dot{a}_j - m_L \dot{s}_j &= \dot{w}_{L2} \\ \dot{a}_j - m_R \dot{s}_j &= \dot{w}_{R1} \end{aligned} \right\} \tag{1.13}$$

(see Fig. 3), which can be written in matrix form as

$$M_j \dot{y}_j = \dot{w}_j \tag{1.14}$$

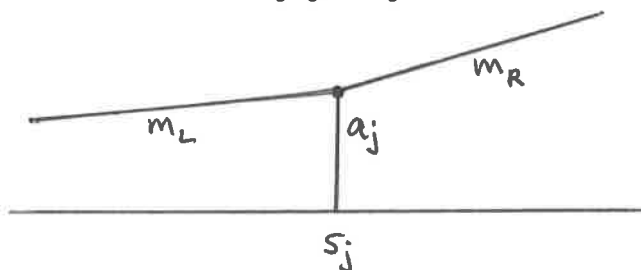


FIG. 3 : Node-element relationship

where $\underline{\dot{y}}_j^T = (\dot{a}_j, \dot{s}_j)^T$, $\underline{\dot{w}}_j^T = (\dot{w}_{L2}, \dot{w}_{R1})^T$, and $M_j = \begin{bmatrix} 1 & -m_L \\ -1 & -m_R \end{bmatrix}$.

For each j equation (1.14) is a 2x2 system, trivially solved to give

$$\underline{\dot{y}}_j = M_j^{-1} \underline{\dot{w}}_j.$$

Combining together equations (1.14), write $M = \text{diag}\{M_j\}$, $\underline{\dot{y}}^T = \{\underline{\dot{y}}_j\}$, $\underline{\dot{w}}^T = \{\underline{\dot{w}}_j^T\}$ giving

$$M \underline{\dot{y}} = \underline{\dot{w}} \quad (1.15)$$

M is diagonal in 2x2 blocks with the trivial inversion $\underline{\dot{y}} = M^{-1} \underline{\dot{w}}$.

Equations (1.11) and (1.15) constitute an alternative, equivalent, way of obtaining (1.8). It follows that, since $\underline{g} = M^T \underline{b}$,

$$M^T C M \underline{\dot{y}} = M^T C \underline{\dot{w}} = M^T \underline{b} = \underline{g} \quad (1.16)$$

so that $A = M^T C M$, c.f. (1.8). Then

$$\underline{\dot{y}} = M^{-1} C^{-1} M^{-T} \underline{g} \quad (1.17)$$

exhibiting the local dependence of $\underline{\dot{y}}$ on Lv . For future reference, note that equations (1.16) include the normal equations for the least squares minimisation of $M \underline{\dot{y}} - \underline{\dot{w}}$ with weight $C^{\frac{1}{2}}$.

(iv) Returning to the FFE method, it has been shown by Wathen (1985) that the mass matrix A_α has the decomposition

$$A_\alpha = L^T C L \quad (1.18)$$

where $L = \text{diag}\{L_j\}$ and $L_j = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$: note that L is rectangular.

The equations (1.5), namely

$$L^T C L \underline{\dot{a}} = \underline{g}_\alpha = L^T \underline{b} = L^T C \underline{\dot{w}} \quad (1.19)$$

are the normal equations for the least squares solution of the overdetermined system.

$$L \underline{a} = \underline{\dot{w}} \quad (1.20)$$

with weight $C^{\frac{1}{2}}$. The FFE Galerkin equations can therefore be regarded as a

projection of Lv into the space S_ϕ (c.f. (1.12)) followed by such a least squares solution. It is this last step which converts the local nature of the method into something global.

The presence of the weight $C^{\frac{1}{2}}$ is also important. Since C is proportional to the size of element, small elements have small weight in the least squares minimisation. As a result, mesh refinement yields lower returns.

Both the non-local difficulty and the mesh refinement difficulty are overcome if a weighting C^0 is used (see below).

(v) A useful observation for MFE in one dimension, following from (1.13) is that, since

$$\left. \begin{aligned} \dot{a}_j - m_k \dot{s}_j &= \dot{w}_{k1} \\ \text{and} \quad \dot{a}_{j+1} - m_k \dot{s}_{j+1} &= \dot{w}_{k2} \end{aligned} \right\} \quad (1.21)$$

the speed of the nodes $j, j+1$ perpendicular to the k th segment are $\dot{w}_{k1} \cos \theta_k, \dot{w}_{k2} \cos \theta_k$, respectively (see Fig. 4). Moreover we have, by addition and subtraction,

$$\dot{a}_{j+1} + \dot{a}_j - m_k (\dot{s}_{j+1} + \dot{s}_j) = \dot{w}_{k1} + \dot{w}_{k2} = (1 \ 1) \underline{\dot{w}}_k \quad (1.22)$$

$$\dot{a}_{j+1} - \dot{a}_j - m_k (\dot{s}_{j+1} - \dot{s}_j) = \dot{w}_{k2} - \dot{w}_{k1} = (1 \ -1) \underline{\dot{w}}_k \quad (1.23)$$

Using $m_k = \tan \theta_k = \frac{a_{j+1} - a_j}{s_{j+1} - s_j}$ (see Fig. 4), we obtain

$$\dot{\eta}_k = \frac{1}{2} (\dot{a}_{j+1} + \dot{a}_j) \cos \theta_k - \frac{1}{2} (\dot{s}_{j+1} + \dot{s}_j) \sin \theta_k = \frac{1}{2} \cos \theta_k (1 \ 1) \underline{\dot{w}}_k \quad (1.24)$$

and

$$\underline{\dot{m}}_k = \frac{1}{(s_{j+1} - s_j)} (1 \ -1) \underline{\dot{w}}_k \quad (1.25)$$

where $\dot{\eta}_k$ is the lateral speed of the mid point of the segment k .

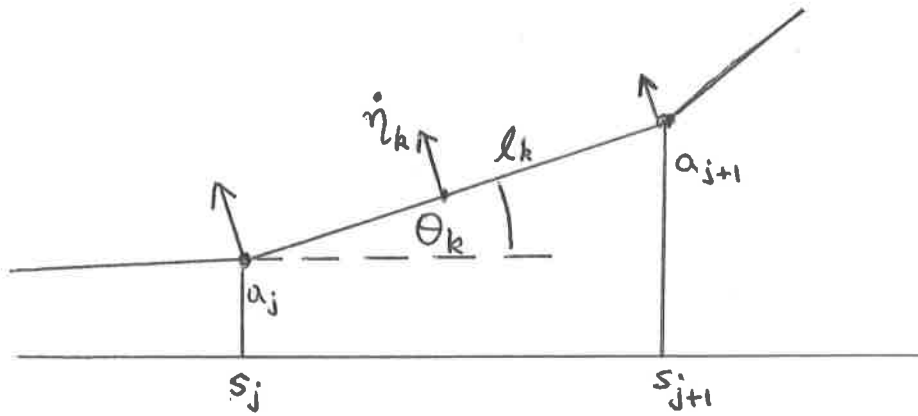


FIG. 4 : k'th element segment motion

Since both $(1 \ 1)^T$ and $(1 \ -1)^T$ are eigenvectors of C_k with eigenvalues $s_{j+1} - s_j$ and $\frac{1}{6}(s_{j+1} - s_j)$, respectively, we obtain from (1.11),

$$\begin{aligned} \dot{\eta}_k &= \frac{\cos\theta_k}{s_{j+1} - s_j} (1 \ 1)^T \underline{b}_k \\ &= \frac{1}{l_k} \langle 1, Lv \rangle \\ &= \frac{1}{l_k} \int_{s_j}^{s_{j+1}} Lv \, dx \end{aligned} \quad (1.26)$$

where $l_k \cos\theta_k = s_{j+1} - s_j$ (see Fig. 4) and

$$\begin{aligned} \dot{m}_k &= \frac{6}{(s_{j+1} - s_j)^2} (1 \ -1)^T \underline{b}_k \\ &= \frac{6}{(s_{j+1} - s_j)^2} \int_{s_j}^{s_{j+1}} \psi_k Lv \, dx \end{aligned} \quad (1.27)$$

where $\psi_k = \phi_{k1} - \phi_{k2}$.

Thus the segment moves perpendicular to itself with speed $\dot{w}_{k1} \cos\theta_k$ at one end and $\dot{w}_{k2} \cos\theta_k$ at the other, equivalent to a mid-point speed $\dot{\eta}_k$ and rotation \dot{m}_k as given above. It has been shown by Baines (1985b) that, for a hyperbolic scalar conservation law, the above properties give smooth transitions to and from entropy respecting shocks and expansions.

(vi) The above analysis is for one dimension. In higher dimensions, however, the structures is the same, but with one major difference. In MFE the space $P_{\alpha\beta}$ (now $P_{\alpha\beta}$) is not identical with the space P_{ϕ} (see §1(iii)). This is because in higher diemnsions there are generally more elements surrounding a node than there are nodes at the corners of an element. The consequence is that, although we can write down equation (1.11) (with the elements now having more than 2 components), and invert C_k as before, the M_j of equation (1.14) is no longer square and (1.14) is an overdetermined system. Nevertheless if we take a least squares solution of (1.14) with $C^{\frac{1}{2}}$ weighting we obtain the standard MFE equations (1.16) (c.f. the FFE description of §1(iv)).

A consequence and criticism of the MFE method in higher dimensions is the same as that pointed out at the end of §1(iv), namely, that the method is no longer local and that mesh refinement gives diminishing returns. A remedy for both criticisms is found in dropping the $C^{\frac{1}{2}}$ weighting in the least squares step, giving the equations

$$M^T M \dot{\underline{y}} = M^T \dot{\underline{w}} \quad (1.28)$$

which invert to give

$$\dot{\underline{y}} = (M^T M)^{-1} M^T \dot{\underline{w}} \quad (1.29)$$

a local solution.

The result that the speed of the segment normal to itself can be calculated in the form (1.26) follows from the fact that M^T contains a row of 1's referring to the nodes at the corner of each element. Thus, taking that particular row in (1.28) yields the sum of the rows of

$$M \dot{\underline{y}} = \dot{\underline{w}} \quad (1.30)$$

which refer to a particular element. Within (1.28) therefore, we have the

equation

$$\sum_v \dot{a}_{kv} - m \sum_v \dot{s}_{kv} - n \sum_v \dot{t}_{kv} = \sum_v \dot{w}_{kv} \quad (1.31)$$

where $(\dot{s}_{kv}, \dot{t}_{kv})$ is the velocity of node v in element k (c.f. (1.22)). Since such a row of 1's is a left eigenvector of C we have as in (iv) the result that

$$\dot{\eta}_k = \frac{1}{A_k} \int L^v ds \quad (1.32)$$

where A_k is the "area" of element k . The result is true for any number of dimensions, and also (because of the eigenvector property) holds for the standard MFE equations (1.16) in higher dimensions.

(vii) Another approach to dealing with non-locality is (e.g. in two dimensions) to abandon triangles in the basic grid and work rather with linear functions or polygons. By projecting L^v into each polygonal element to obtain a best plane fit, we obtain an approximation v_t to u_t which can be used to move the plane approximation v . As the planes move the polygons are distorted and the grid also moves. Another way of thinking about this is to imagine straight lines moving around in two dimensions where points had moved in one dimension. This approach was suggested by Prof. P.L. Roe.

(viii) For systems of m equations there are m sets of local equations of the type (1.12) but it may be convenient to work with a single grid, in which case a further projection may be necessary. (If each component has its own grid a projection will still be necessary in more than one dimension).

Using a superfix μ to indicate the μ 'th component, we have

$$C^{\mu} \underline{\dot{w}}^{\mu} = \underline{\dot{b}}^{\mu}$$

and in one dimension, if a single grid is to be used, the $\underline{\dot{w}}^{\mu}$ have to be mapped on to \dot{a}^{μ} and a single \dot{s} . The appropriate overdetermined set of equations is

$$M_j \begin{bmatrix} \dot{a}_j^1 \\ \vdots \\ \dot{a}_j^m \\ \dot{s}_j \end{bmatrix} = \begin{bmatrix} \dot{w}_j^1 \\ \vdots \\ \dot{w}_j^m \end{bmatrix}, \quad (1.33)$$

where M_j is rectangular, and we may solve then in a least square sense, giving

$$M_j^T M_j \begin{bmatrix} \dot{a}_j^1 \\ \vdots \\ \dot{a}_j^m \\ \dot{s}_j \end{bmatrix} = M_j^T \begin{bmatrix} \dot{w}_j^1 \\ \vdots \\ \dot{w}_j^m \end{bmatrix} \quad (1.34)$$

(c.f. (1.28)) which determines the vector of unknowns in the form

$$\begin{bmatrix} \dot{a}_j^1 \\ \vdots \\ \dot{a}_j^m \\ \dot{s}_j \end{bmatrix} = (M_j^T M_j)^{-1} M_j^T \begin{bmatrix} \dot{w}_j^1 \\ \vdots \\ \dot{w}_j^m \end{bmatrix} \quad (1.35)$$

2. CONSTRAINTS

We now discuss the topic of Moving Finite Elements with constraints.

The FFE method is an example of the MFE method with constraints.

These constraints may be expressed as

$$\dot{s}_j = 0 \quad (2.1)$$

or, in the formulation of §1(v), as

$$\dot{w}_{jL} = \dot{w}_{jR} \quad (2.2)$$

The MFE method in more than one dimension may also be regarded as a constrained method, where the $\underline{\dot{w}}$ must lie in the range space of M . These and other possible (linear) constraints are now investigated in a general way.

(i) Suppose that the residual

$$\|v_t - Lv\|_2 \quad (2.3)$$

is to be minimised subject to the linear constraint

$$Z^T \underline{\dot{w}} = \underline{e} \quad (2.4)$$

(e.g. (2.2) above). Using (1.10) we can write

$$\frac{1}{2} \|v_t - Lv\|_2^2 = \frac{1}{2} \underline{\dot{w}}^T C \underline{\dot{w}} - \underline{\dot{w}}^T \underline{b} + \text{terms independent of } \underline{\dot{w}}. \quad (2.5)$$

Thus, to incorporate (2.4) we use a Lagrange multiplier $\underline{\lambda}$, minimising

$$\frac{1}{2} \underline{\dot{w}}^T C \underline{\dot{w}} - \underline{\dot{w}}^T \underline{b} + \underline{\lambda}^T (Z^T \underline{\dot{w}} - \underline{e}) \quad (2.6)$$

This gives the pair of equations

$$C \underline{\dot{w}} = \underline{b} - Z \underline{\lambda} \quad (2.7)$$

$$Z^T \underline{\dot{w}} = \underline{e}. \quad (2.8)$$

To eliminate $\underline{\lambda}$ invert C in (2.7) and multiply by Z^T : this gives

$$Z^T \underline{\dot{w}} = Z^T C^{-1} (\underline{b} - Z \underline{\lambda}) = \underline{e} \quad (2.9)$$

or

$$Z^T C^{-1} Z \underline{\lambda} = Z^T C^{-1} \underline{b} - \underline{e} \quad (2.10)$$

Solving for $\underline{\lambda}$ gives

$$\underline{\lambda} = (Z^T C^{-1} Z)^{-1} (Z^T C^{-1} \underline{b} - \underline{e}) \quad (2.11)$$

and substitution into (2.7) gives

$$C \underline{\dot{w}} = \underline{b} - Z (Z^T C^{-1} Z)^{-1} (Z^T C^{-1} \underline{b} - \underline{e}), \quad (2.12)$$

which is the modified version of (1.12) in the presence of the constraints (2.4).

Note that the numerical inversion of $Z^T C^{-1} Z$ can be simply achieved using the pre-conditioned conjugate gradient method. The spectrum of $Z^T C^{-1} Z$, preconditioned by the inverse of its diagonal entries, has the same support as the λ 's in

$$|C^{-1} - \lambda C_D^{-1}| = 0 \quad , \quad (2.13)$$

where $C_D^{-1} = \text{diag}\{C^{-1}\}$, which are simple to construct and have a compact support (see Wathen (1985)).

(ii) The constraint of fixing nodes in the MFE method to obtain the FFE method can be expressed as (2.2) or, alternatively, in the form that $\dot{\underline{w}}$ should lie in the range space of L (see (1.18)). The constraint takes the form

$$Z^T \dot{\underline{w}} = 0 \quad (2.14)$$

or

$$Z^T L = 0 \quad (L^T Z = 0) \quad , \quad (2.15)$$

Replacing (2.8) by the second of (2.15) yields, from (2.7),

$$L^T C \dot{\underline{w}} = L^T \underline{b} \quad (2.16)$$

or, since $\dot{\underline{w}} = L \dot{\underline{a}}$ and $L^T \underline{b} = \underline{g}_\alpha$, we obtain

$$L^T C L \dot{\underline{a}} = \underline{g}_\alpha \quad (2.17)$$

as before (eqn. (1.19)). In this case Z is an array of vectors of the form $[1 \ -1]^T$ and, if used in (2.12), produces a $\dot{\underline{w}}$ which is consistent with solving the rectangular system

$$L \dot{\underline{a}} = \dot{\underline{w}} \quad . \quad (2.18)$$

Likewise the constraint of forcing $\dot{a}_j = 0 \quad \forall j$ may be expressed as

$$Z^T \dot{\underline{w}} = 0 \quad (2.19)$$

or

$$Z^T L_m = 0 \quad (L_m^T Z = 0) \quad (2.20)$$

where L_m consists of vectors $[-m_L \ -m_R]^T$ and Z consists of vectors $[m_R \ -m_L]^T$. This yields the equations

$$L_m^T C L_m = \underline{g}_\beta \quad (2.21)$$

where $\underline{g}_\beta = \{g_i^\beta\}$, $g_i^\beta = \langle \beta_i, L v \rangle$, c.f. (2.17). If Z is used directly in (2.12) the result is a $\dot{\underline{w}}$ which is consistent with the rectangular system

$$L_m \dot{\underline{s}} = \dot{\underline{w}} \quad (2.22)$$

In both of these examples the alternative procedure of using (2.14) or (2.19) to define Z and substitution of Z into (2.12) produces the same results by eradicating the other equations.

(iii) To minimise $\|v_t - L v\|_2^2$ subject to a constraint directly expressed in terms of \underline{y} , say $\Gamma^T \dot{\underline{y}} = \underline{d}$, we again use a Lagrange multiplier $\underline{\lambda}$ and minimise

$$\frac{1}{2} \dot{\underline{y}}^T A \dot{\underline{y}} - \dot{\underline{y}}^T \underline{g} + \underline{\lambda}^T (\Gamma^T \dot{\underline{y}} - \underline{d}) \quad (2.23)$$

c.f. (2.6). This gives the pair of equations

$$A \dot{\underline{y}} = \underline{g} - \Gamma \underline{\lambda} \quad (2.24)$$

$$\Gamma^T \dot{\underline{y}} = \underline{d} \quad (2.25)$$

To eliminate $\underline{\lambda}$ invert A and multiply by Γ^T : this gives

$$\Gamma^T \dot{\underline{y}} = \Gamma^T A^{-1} (\underline{g} - \Gamma \underline{\lambda}) = \underline{d} \quad (2.26)$$

or

$$\Gamma^T A^{-1} \Gamma \underline{\lambda} = \Gamma^T A^{-1} \underline{g} - \underline{d} \quad (2.27)$$

giving

$$\underline{\lambda} = (\Gamma^T A^{-1} \Gamma)^{-1} (\Gamma^T A^{-1} \underline{g} - \underline{d}) \quad (2.27)$$

Substituting for $\underline{\lambda}$ gives

$$A \dot{\underline{y}} = \underline{g} - \Gamma (\Gamma^T A^{-1} \Gamma)^{-1} (\Gamma^T A^{-1} \underline{g} - \underline{d}) \quad (2.28)$$

This is the modified equation for $\underline{\dot{y}}$ in the presence of a constraint.

Numerical inversion of A can be done using pre-conditioned conjugate gradients, which is again rapid since the spectrum of $D^{-1}A$ is compact. [This is proved by Wathen (1985) by considering $A - \lambda D = M^T(C - \lambda C_D)M$, where C_D is the matrix of diagonal elements of C]. Here it is also necessary to invert

$$B = \Gamma^T A^{-1} \Gamma = (A^{-1} \Gamma)^T A (A^{-1} \Gamma) = (M A^{-1} \Gamma)^T C (M A^{-1} \Gamma). \quad (2.29)$$

By the same argument as used by Wathen the spectrum of B , preconditioned in the standard way, satisfies the same bounds as those of A .

(iv) The same equations are obtained if an unconstrained projection is followed by a constrained least-squares minimisation.

The unconstrained projection gives $C \underline{\dot{w}} = \underline{b}$. The unconstrained least-squares minimisation of $\|C^{\frac{1}{2}}(M \underline{\dot{y}} - \underline{\dot{w}})\|$ gives $A \underline{\dot{y}} = \underline{g}$ where $A = M^T C M$. If the latter is constrained by $\Gamma^T \underline{\dot{y}} = \underline{d}$, introduce a Lagrange multiplier $\underline{\lambda}$ and minimise

$$\frac{1}{2} \|C^{\frac{1}{2}}(M \underline{\dot{y}} - \underline{\dot{w}})\|^2 + \underline{\lambda}^T (\Gamma^T \underline{\dot{y}} - \underline{d}) \quad (2.30)$$

This gives the pair of equations

$$M^T C M \underline{\dot{y}} - M^T C \underline{\dot{w}} + \Gamma \underline{\lambda} = 0 \quad (2.31)$$

$$\Gamma^T \underline{\dot{y}} = \underline{d} \quad (2.32)$$

Using $M^T C M = A$, $C \underline{\dot{w}} = \underline{b}$, $M^T \underline{b} = \underline{g}$, the first of these becomes

$$A \underline{\dot{y}} - \underline{g} + \Gamma \underline{\lambda} = 0. \quad (2.33)$$

So we get the same pair of equations as before.

(v) Suppose we drop the $C^{\frac{1}{2}}$ weighting making everything local. We then get the pair of equations

$$M^T M \underline{\dot{y}} - M^T \underline{\dot{w}} + \Gamma \underline{\lambda} = 0 \quad (2.34)$$

$$\Gamma^T \underline{\dot{y}} = 0 \quad (2.35)$$

Invert $M^T M$ and multiply by Γ^T : this gives

$$\Gamma^T \underline{\dot{y}} = \Gamma^T (M^T M)^{-1} (M^T \underline{\dot{w}} - \Gamma \underline{\lambda}) = \underline{d} \quad (2.36)$$

or

$$\Gamma^T (M^T M)^{-1} \Gamma \underline{\lambda} = \Gamma^T (M^T M)^{-1} M^T \underline{\dot{w}} - \underline{d}$$

giving

$$\underline{\lambda} = (\Gamma^T (M^T M)^{-1} \Gamma)^{-1} (\Gamma^T (M^T M)^{-1} M^T \underline{\dot{w}} - \underline{d}) \quad (2.37)$$

Substituting for $\underline{\lambda}$ gives

$$M^T M \underline{\dot{y}} = M^T \underline{\dot{w}} - \Gamma (\Gamma^T (M^T M)^{-1} \Gamma)^{-1} (\Gamma^T (M^T M)^{-1} M^T \underline{\dot{w}} - \underline{d}) \quad (2.38)$$

the modified equation for a new $\underline{\dot{y}}$ in the presence of the constraint.

[If there is no constraint the equation is $M^T M \underline{\dot{y}} = M^T \underline{\dot{w}}$, c.f. (1.34)].

Inversion of $M^T M$ is straightforward. For $B_0 = \Gamma^T (M^T M)^{-1} \Gamma$, write $B_0 = ((M^T M)^{-1} \Gamma)^T (M^T M)^{-1} \Gamma$: the spectrum of B_0 matches that of $M^T M$, see Wathen (1985).

(vi) An important application of constraints is in the formation of two-dimensional shocks. In one dimension, when nodes overtake in a hyperbolic problem such as a conservation law, a shock is formed by maintaining the speed $\dot{\eta}_k$ of (1.26) but reducing the \dot{m}_k of (1.27) instantaneously to zero. In practice two of the components of $\underline{\dot{w}}$ in (1.12) are lost but these are replaced by the appropriate shock conditions, which derive from (1.26):

if $Lv = -f_x$, we have

$$\dot{\eta}_k = \frac{-1}{\ell_k} \int_{s_j}^{s_{j+1}} f_x dx = \frac{f_{j+1} - f_j}{a_{j+1} - a_j} \quad (2.39)$$

the usual jump condition. Thus we impose the constraints

$$\dot{s}_{j+1} = \dot{s}_j = \frac{f_{j+1} - f_j}{a_{j+1} - a_j} \quad (2.40)$$

which may be written (trivially) in the form (2.32). However, the

formalism of section (v) is unnecessary since the constraints can be expressed

explicitly. Note that two equations are lost (for the $\dot{\underline{w}}_k$) and two gained (from (2.40)).

In two dimensions the equivalent of node overtaking is a triangular segment becoming vertical (see Fig. 5). As already pointed out in section 1(vi) the velocity $\dot{\eta}_k$ of the centroid of the triangle perpendicular to its

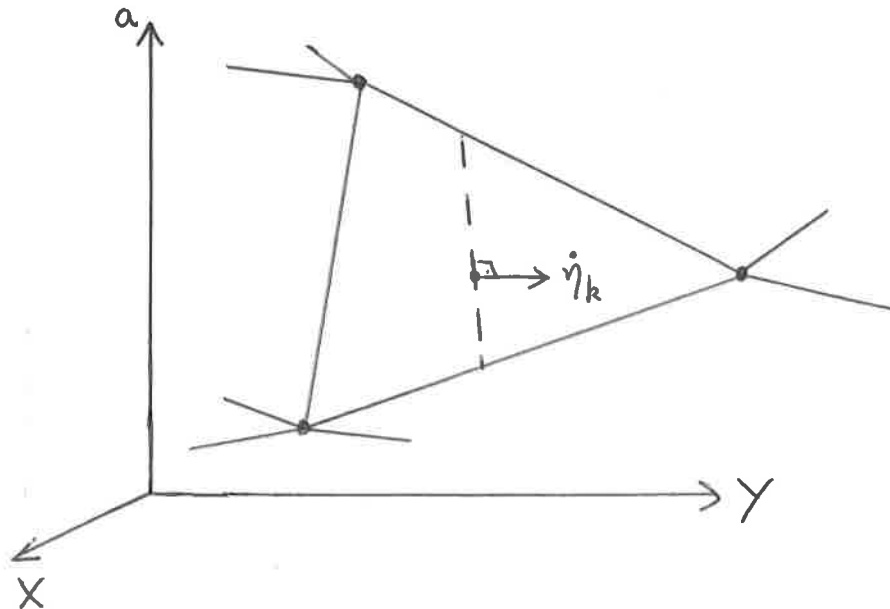


FIG. 5 : Triangular segment becoming vertical

plane is given by (1.32). If this is maintained when the triangle becomes vertical, and its angular velocity making it overturn is reduced to zero instantaneously then, by analogy with the one dimensional case (2.39) above, the triangle sweeps out mass in accordance with the conservation law. Specifically, for the problem

$$u_t + \underline{\nabla} \cdot \underline{f}(u) = 0 \tag{2.41}$$

we have from (1.32)

$$\dot{\eta}_k = - \frac{1}{A_k} \int \underline{\nabla} \cdot \underline{f}(u) ds_k \tag{2.42}$$

where ds_k is measured perpendicular to the triangular segment k . When the triangle becomes vertical mass is swept out at a rate $A_k \dot{\eta}_k$ which, from (2.42), is consistent with the conservation law (2.41).

In practice three of the components of $\dot{\underline{w}}$ in (1.12) are lost and are replaced by suitable constraints. One of these constraints is that the speed of the centroid of the triangle is continued at the formation of the shock; the other is that no subsequent overturning takes place, i.e. that the vertical line through the centroid remains vertical. Since two constraints replace three equations the system is no longer explicit and the formalism of section (v) comes into play.

From Fig. 6 we see what the constraints actually are. If the

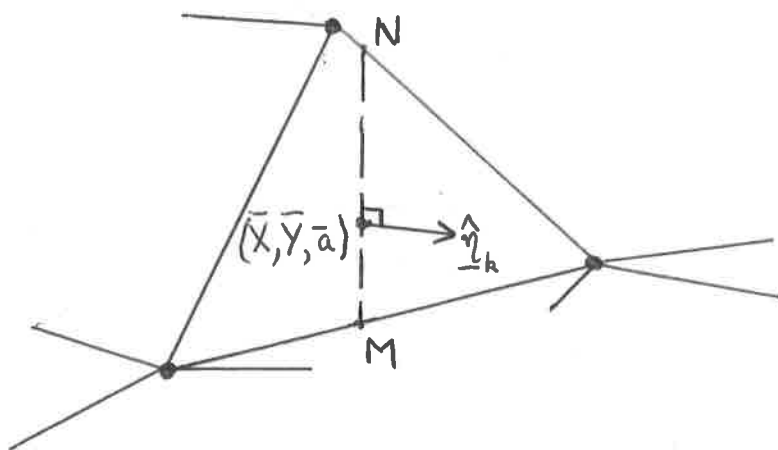


FIG. 6 : Shocked (vertical) triangular segment

co-ordinates of the vertices of the triangle are (X_i, Y_i, a_i) ($i=1,2,3$) and those of the centroid are $(\bar{X}, \bar{Y}, \bar{a})$ then one constraint is

$$(\dot{\bar{X}}, \dot{\bar{Y}}) \cdot \hat{n}_k = \text{constant}, \quad (2.43)$$

equal to its value just before the shock forms (where \hat{n}_k is the unit vector perpendicular to the plane of the triangle). The other constraint, which is longer to write down, is that the line MN remains vertical. Together, these constraints supply the $\dot{\Gamma} \underline{y} = \underline{d}$ of (2.32).

CONCLUSION

Here we summarise the results in this report.

We compared the fixed finite element (FFE) method with the moving finite element (MFE) method in terms of projections. In particular, we noted that in one dimension for a scalar equation the MFE method is local but the FFE method is global. Also in higher dimensions both methods are global but there are possible local versions, the crux of the matter being the transfer of elementwise projection information onto the nodes. If this is done with the standard weighting a global method results, but it can be done in such a way that a local method results. Moreover with the standard weighting mesh refinement is an uphill struggle since the weight is proportional to the size of the element.

The local form of the MFE method in one dimension gives a clear picture of what happens at shocks and even in higher dimensions there is a smooth transition to the shock.

The second part of the report showed how explicit constraints may be built into the MFE method, both in the local and global formulations. We observed that although the constrained form looks complicated its solution make good use of the pre-conditioned conjugate gradient method, as for the unconstrained global MFE. Finally an application of constraints to the formation of two-dimensional shocks was considered in some detail.

REFERENCES

- BAINES, M.J. (1985a) On Approximate Solutions of Time-Dependent Partial Differential Equations by the Moving Finite Element Method, Numerical Analysis Report 1/85, Dept. of Maths., University of Reading.
- BAINES, M.J. (1985b) Local Adaptive Moving Finite Elements, in Numerical Methods for Fluid Dynamics, Proc. Conf. Reading, April 1985, (O.U.P.), to appear.
- 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. & MILLER, R. (1981) Moving finite elements, part I. SIAM J. Num. Analysis 18, 1019-1032.
- MILLER, K. (1981) Moving finite elements, part II. SIAM J. Num. Analysis 18, 1033-1057.
- WATHEN, A.J. (1985) Attainable Eigenvalue Bounds for the Galerkin Mass Matrix, submitted to IMAJNA.
- WATHEN, A.J. & BAINES, M.J. (1985) On the structure of the Moving Finite-element Equations. IMAJNA 5, 161-182.