PROJECTIONS AND CONSTRAINTS II

M.J. BAINES

NUMERICAL ANALYSIS REPORT 2/86

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

1. INTRODUCTION

In an earlier report (15/85)* we showed that the Moving Finite Element (MFE) method in one dimension may be regarded as a two-step procedure, consisting first of a projection into each local element subspace, followed by an explicit transfer of the elementwise information thus generated onto the nodes. We also showed that in the case of MFE in higher dimensions, or for fixed finite elements (FFE) in any number of dimensions, the second step leads to a solution which is inconsistent with the continuity of the underlying function and a further projection is needed. For details see the original report.

In the 'Constraints' section of the same report it was shown how to incorporate a linear constraint on the nodal motions into the method, giving an elaborate formula for the solution in such a situation. In particular, the FFE method can be regarded as MFE with the nodal positions constrained not to move, and the FFE equations are regained in this case.

In the present report we give a simplified account of the incorporation of constraints into the method, and link the idea in with the projections themselves. Later we explore the possibilities of the use of constraints in two practical situations, those of convection diffusion and of non-linear heat flow.

2. CONSTRAINTS

All the methods discussed here (including the FFE method and unconstrained and constrained MFE methods) start, in the case of the partial differential equation

$$u_t = Lu \quad , \tag{2.1}$$

*see Baines (1985c)

with the projection of Lv (where v is the piecewise linear finite element approximation to u) into each local element subspace. Suppose that the element k supports $d+1$ element basis functions $\phi_{k\nu}$, $\nu = 1,2,\ldots,d+1$, where d is the dimensionality of the problem. Projection of Lv into the subspace S_{ϕ_k} spanned by the $\phi_{k\nu}$ leads to the approximation

$$\sum_{\nu} \dot{w}_{k\nu}\phi_{k\nu} \tag{2.2}$$

where, in the case of an L_2 projection, the $\dot{w}_{k\nu}$ are given by the normal equations

$$C_k\underline{\dot{w}}_k = \underline{b}_k \quad , \tag{2.3}$$

say, where $C_k = \{C_{k\mu\nu}\}$, $C_{k\mu\nu} = \langle\phi_{k\mu},\phi_{k\nu}\rangle$, $\underline{\dot{w}}_k = \{\dot{w}_{k\nu}\}$ and $\underline{b}_k = \{b_{k\nu}\}$, $b_{k\nu} = \langle\phi_{k\nu},Lv\rangle$. Globally

$$C\underline{\dot{w}} = \underline{b} \tag{2.4}$$

where $C = \{C_k\}$, $\underline{\dot{w}}^T = \{\underline{w}_k\}$, $\underline{b}^T = \{\underline{b}_k\}$. The matrix C is block diagonal in $(d+1) \times (d+1)$ blocks.

The $\underline{\dot{w}}_k$'s give information about the element motions (see Baines (1985b)). Transfer of this information on to the nodes is effected simply in <u>one-dimensional</u> MFE by solving, for each node j, the pair of equations

$$\left.\begin{aligned} \dot{a}_j - m_{j-\frac{1}{2}}\dot{s}_j &= w_{j-\frac{1}{2},2} \\ \dot{a}_j - m_{j+\frac{1}{2}}\dot{s}_j &= w_{j+\frac{1}{2},1} \end{aligned}\right\} \tag{2.5}$$

where $(\dot{s}_j,\dot{a}_j)$ are the components of the nodal velocity, and $m = v_x$ is the (constant) slope of the approximation in an element. These equations can be written in the matrix form

$$M_j \dot{\underline{y}}_j = \underline{w}_j \tag{2.6}$$

where $\dot{\underline{y}}_j^T = (\dot{a}_j, \dot{s}_j)$, etc. Globally we have

$$M\dot{\underline{y}} = \dot{\underline{w}} \tag{2.7}$$

where $\dot{\underline{y}}^T = \{\dot{\underline{y}}_j^T\}$, $\underline{w}^T = \{\underline{w}_j^T\}$. Combining (2.4) and (2.7) yields

$$CM\dot{\underline{y}} = \underline{b} \tag{2.8}$$

or

$$M^T CM\dot{\underline{y}} = M^T\underline{b} \quad . \tag{2.9}$$

The point of the last operation is to achieve the form (2.9), which is the standard form of the MFE equations (see Miller (1981), Wathen & Baines (1985)) with the structure of the mass matrix made clear. This matrix structure (in one dimension) has the remarkable property that, when preconditioned by the inverse of the diagonal, its eigenvalues are simply $\frac{1}{2}$ and $\frac{3}{2}$ (repeated N+1 times, where N+1 is the number of nodes) (see Wathen & Baines (1985), Wathen (1985)). The matrix form M^TCM may be regarded as "assembling" the elementwise matrices in C by the MFE assembly matrix M.

In higher dimensions a similar structure is obtained. The difference lies in the matrix M, which is now rectangular. Because of the inequality between the total number of elements and the total number of nodes in higher dimensional meshes the vector $\dot{\underline{w}}$ is longer than $\dot{\underline{y}}$, so that (2.7) does not have a $\dot{\underline{y}}$ solution for general $\dot{\underline{w}}$. Indeed, a constraint is now needed on $\dot{\underline{w}}$ to make it lie in the range space of M. The way solutions are obtained is to use the special form (2.9), which is equivalent to minimising

$$M\dot{\underline{y}} - \dot{\underline{w}} \tag{2.10}$$

in the L_2 norm with weight $C^{\frac{1}{2}}$. This is the second projection referred to in the introduction. The resulting equation (2.9) gives again the standard MFE equations. This time we can show (Wathen (1985)) that the eigenvalues of the matrix M^TCM, when preconditioned by the inverse of the diagonal, lie in the interval $(\frac{1}{2}, 1 + \frac{d}{2})$, where d is the dimension of the problem. The pre-conditioned matrix is readily inverted using the conjugate gradient method.

This summary of the MFE method and its solution procedure is needed so that we can incorporate (fairly) general constraints into the method. In the above there was in general no valid $\dot{\underline{y}}$ corresponding to $\dot{\underline{w}}$ in (2.7) and a restricted set (2.9) of equations (2.8) had to be solved. Consider now the FFE method from this point of view. In this case $\dot{\underline{y}} = \{\dot{a}_j\}$ since the $\dot{s}_j = 0 \ \forall j$. Even more strongly there is now no valid $\dot{\underline{y}}$ corresponding to $\dot{\underline{w}}$ and for a solution for $\dot{\underline{y}}$ we require $\dot{\underline{w}}$ to lie in the range space of L where $L = \{L_j\}$ and $L_j = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ("half" of the matrix M_j). The solution is obtained by solving

$$L^TCL\dot{\underline{y}} = L^T\underline{b} \tag{2.11}$$

(c.f. (2.9)), a restricted set of equations based on minimising $L\dot{\underline{y}} - \dot{\underline{w}}$ in the L_2 norm with weight $C^{\frac{1}{2}}$. The equations (2.11) are the standard FFE equations.

We turn now to more general constraints. Suppose that some elements of $\dot{\underline{y}}$ are constrained to be linear functions of other elements of $\dot{\underline{y}}$. Then we may write

$$\dot{\underline{y}} = R\dot{\underline{y}}^* \tag{2.12}$$

where $\dot{\underline{y}}^*$ contains the unconstrained elements of $\dot{\underline{y}}$. An example might be the constraint that one of the $\dot{y}_j$ should move with the average of the speed of its neighbours. If $\dot{\underline{y}}^*$ can be evaluated, all the elements of

$\dot{\underline{y}}$ are obtained from (2.12). Substituting (2.12) into (2.7) we obtain

$$MR\dot{\underline{y}}^{*} = \dot{\underline{w}} \tag{2.13}$$

and, by analogy with previous arguments, for a solution $\dot{\underline{y}}^{*}$ to exist the vector $\dot{\underline{w}}$ must lie in the range space of MR, which is not true in general. A further projection is therefore again necessary and this time we minimise

$$||MR\dot{\underline{y}}^{*} - \dot{\underline{w}}|| \tag{2.14}$$

in the L_2 norm with weight $C^{\frac{1}{2}}$. This leads to the equations

$$(MR)^{T} C\, MR\dot{\underline{y}}^{*} = (MR)^{T}\underline{b} \tag{2.15}$$

(c.f. (2.9)). The particular advantage of this form is that the eigenvalue property mentioned above goes over to these equations (see Appendix). Therefore, using pre-conditioned conjugate gradients (see Wathen (1985)), (2.15) can be readily solved for $\dot{\underline{y}}^{*}$, and (2.12) used to generate $\dot{\underline{y}}$.

Apart from the contraints which are built in to the MFE method in dimension greater than one and those for FFE, we mention now some others.

(i) Boundary conditions

The effects of the outside world on a problem (known as boundary conditions) give constraints on the elements of $\dot{\underline{y}}$. For example a Dirichlet condition at a fixed boundary s_0 gives two constraints $\dot{a}_0 = \dot{s}_0 = 0$, which hold for all time. These can be modelled by (2.12).

(ii) Parallelism

If $m_{j-\frac{1}{2}} = m_{j+\frac{1}{2}}$ in (2.5) or (2.6), M in (2.7) becomes singular and a special procedure is needed. It has become standard to temporarily fix $\dot{s}_j = 0$, solve a reduced system of equations and to

add a multiple of the null space of M to force $\dot{s}_j$ to take some average of the speeds of its neighbours. Suppose we reverse this procedure. First force $\dot{s}_j$ to be the prescribed average. This determines R in (2.12). Then, solving (2.15) with (2.12), automatically contrives the constraint and avoids the singularity.

(iii) Shocks

More generally the constraints are prescribed (and are non-zero) so that (2.12) takes the form

$$\underline{\dot{y}} = R\underline{\dot{y}}^* + \underline{s} \tag{2.16}$$

In that case (2.15) becomes

$$(MR)^T C(MR)\underline{\dot{y}}^* = (MR)^T \underline{b} - (MR)^T CM\underline{s} \tag{2.17}$$

A special example of the inhomogeneous form (2.16) arises when nodes overtake in a scalar hyperbolic problem. The approximate solution v has the form of an overturned manifold, and this unphysical behaviour is restricted by applying a constraint at the moment of overturning (see Baines (1985b)). The simplest description of the constraint is that the rotation of the element is stopped or "frozen" at this instant. In practice the constraint equations (in one dimension) are

$$\dot{s}_j = \dot{s}_{j+1} = \frac{f(a_{j+1}) - f(a_j)}{a_{j+1} - a_j} \tag{2.18}$$

corresponding to the usual jump condition for shocks. This constraint is of the form (2.16) with $R=0$. In higher dimensions however the full form of (2.16) is required.

(iv) Steep fronts.

In problems with diffusion present node overtaking and shocks are non-physical and some form of constraint must be devised. Indeed, although the MFE method, being close to a method of characteristics (see Baines (1985a), Wathen (1984)), is particularly suited to advection processes, it is in a sense unsuited to diffusion processes. Where advection and diffusion occur simultaneously, however, there is a case for using the MFE method with constraints. This was done by Miller (1981) in his original presentation of MFE using penalty functions. The approach here is somewhat different.

A direct treatment of second derivatives (from diffusion terms) is difficult because of the piecewise linear nature of the underlying approximation v. Even within the usual finite element inner products a mollifying or recovery argument is needed to justify the existence of the integrals (see Johnson (1985)). Added to this is the special difficulty that the nodes cluster in regions of high curvature where second derivatives are important. When the additional inaccuracies of finite difference time stepping are included there is little confidence in the node overtaking phenomena that occur. One way out is to take very small time steps but there is evidence that this is an over-restrictive artefact of the method.

An obvious constraint is to limit the nodal movement in such a way that some average spacing is maintained. Clearly this cannot be done over the whole range (without going back to FFE) and some selectivity is required. Where advection terms dominate diffusion terms the flexibility of MFE is clearly demanded, but where diffusion terms become dominant (in regions of high curvature)

there is a good case for constraining the nodes. It is therefore suggested that the ratio of these terms is monitored and, for nodes where diffusion dominates, constraints are imposed. The type of constraint suggested is the same as that imposed in the case of parallelism, namely, that a constrained node is moved at a speed which is the average of its neighbours. Where a patch of constrained nodes occurs, each node of the patch can be moved at a speed which is a proportionate average of the nearest unconstrained nodes. Thus in one dimension a typical $\dot{s}_j$ will then be of the form

$$\dot{s}_j = \frac{\lambda \dot{s}_{jL} + \mu \dot{s}_{jR}}{\lambda + \mu} \quad , \tag{2.19}$$

where s_{jL} and s_{jR} are the co-ordinates of the nearest unconstrained nodes to the left and right and

$$\lambda = s_{jR} - s_j, \quad \mu = s_j - s_{jL} \quad . \tag{2.20}$$

3. EXAMPLES

The form of the monitor function will differ from problem to problem. We give here two examples, those of advection-diffusion and of non-linear heat flow. For the advection diffusion equation

$$u_t + uu_x = \varepsilon u_{xx} \tag{3.1}$$

the advection and diffusion terms are uu_x and u_{xx}, respectively. These may be approximated by $u_j \bar{m}_j$ and $\varepsilon(m_{j+\frac{1}{2}} - m_{j-\frac{1}{2}})$ where $\bar{m}_j = \frac{1}{2}(m_{j-\frac{1}{2}} + m_{j+\frac{1}{2}})$. The ratio

$$\left| \frac{\varepsilon(m_{j+\frac{1}{2}} - m_{j-\frac{1}{2}})}{u_j \bar{m}_j} \right| \tag{3.2}$$

is therefore a monitor of the strength of diffusion versus advection. If this monitor exceeds a certain tolerance, the node is designated a constrained node and (2.19) applied.

It has been shown by Herbst (1981) that, for a solution with a shock structure with a jump J, the finite element solution satisfies

$$\varepsilon \sum_j m_j^2 = J^3 \quad . \tag{3.3}$$

This relation can be used to demonstrate that, for ε small and fixed J, large slopes are required in the solution which, if the nodes are not sufficiently close, manifest themselves in oscillations. Thus in applying constraints it is important to allow at least one pair of nodes to approach to within $\Delta s_{min.} = J/m_{max.}$ where $m_{max.}$ is the maximum slope. But the maximum slope, by (3.3), is not greater than $\sqrt{(J^3/\varepsilon)}$ and hence $s_{min.}$ does not need to be less than $\sqrt{(\varepsilon/J)}$. If this minimum spacing is reached by a pair of nodes during the calculation, they can be constrained at that spacing thereafter and they will always be able to take up the greatest slope that can possibly arise so that oscillations need never occur.

The second example considered is that of the non-linear heat flow equation

$$\begin{aligned} u_t &= (uu_x)_x \\ &= u_x^2 + uu_{xx} \quad . \end{aligned} \tag{3.4}$$

Here the "advection" term is u_x^2 and the diffusion term is uu_{xx}, so that the monitor corresponding to (3.2) is

$$\left| \frac{u_j (m_{j+\frac{1}{2}} - m_{j-\frac{1}{2}})}{\overline{m}_j^2} \right| \quad . \tag{3.5}$$

It is interesting to note that in both cases the test is an inequality of the form

$$|m_{j+\frac{1}{2}} - m_{j-\frac{1}{2}}| > K(u_j, \bar{m}_j) \tag{3.6}$$

where K is some function. The parallelism inequality on the other hand is

$$|m_{j+\frac{1}{2}} - m_{j-\frac{1}{2}}| < k \tag{3.7}$$

where k is some tolerance. The point of view developed here therefore suggests that the curvature, or change in slope, should be neither too large nor too small in applications of the MFE method and that constraints (of the same form - averaging the nodal speeds) should be applied in both cases.

One further point is that, in a steady state, advection and diffusion in (3.1) are balanced. It is natural to expect a fully constrained , i.e. stationary, grid in this case. The monitor (3.2) should therefore always provide for a co-ordinate constraint when it is close to unity. Of course there is the implication here that linear diffusion problems should not be treated by MFE, but by FFE.

The constraints may in all cases be taken to be of the form (2.19) which, when written in the form (2.12), provides the R which is used in solving (2.15).

4. EXTENSION TO SYSTEMS

For systems of equations the main question in MFE is whether to use separate grids for each component or to use a single grid. We demonstrate that single grid methods are constrained versions of the separate grid method.

Suppose there are three components to the system so that three sets of $\dot{\underline{w}}$'s may be calculated from three equations of type (2.4). Transfer of this information on to the nodes is effected by (2.5) or its generalisation. Taking separate grids gives three such forms of (2.5) (or (2.7)). However, three grids are technically cumbersome to handle (although easy to solve for the $\dot{\underline{y}}$'s). A simplification suggested in Baines (1985a) and implemented in Baines & Wathen (1986), Edwards (1985), is to use a single grid determined by some average. This seems appropriate when the system is physically such that shocks, for example, occur simultaneously in all three components (the Euler equations).

In this case we take the tripled M_j matrix (2.6) and use the constraint

$$\dot{\underline{y}} = R\,\dot{\underline{y}}^* \tag{4.1}$$

where $$\dot{\underline{y}} = \{\ldots;\dot{a}_j^{(1)},\dot{s}_j^{(1)},\dot{a}_j^{(2)},\dot{s}_j^{(2)},\dot{a}_j^{(3)},\dot{s}_j^{(3)};\ldots\} \tag{4.2}$$

and $$\dot{\underline{y}}^* = \{\ldots;\dot{a}_j^{(1)},\dot{a}_j^{(2)},\dot{a}_j^{(3)},\dot{s}_j^{*};\ldots\} \tag{4.3}$$

the unconstrained unknowns being those in $\dot{\underline{y}}^*$. The form of R_j (the local block of R) is

$$R_j = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \tag{4.4}$$

and that of $(MR)_j$ is

$$(MR)_j = \begin{bmatrix} 1 & 0 & 0 & -m_L^{(1)} \\ 0 & 0 & 0 & -m_R^{(1)} \\ 0 & 1 & 0 & -m_L^{(2)} \\ 0 & 0 & 0 & -m_R^{(2)} \\ 0 & 0 & 1 & -m_L^{(3)} \\ 0 & 0 & 0 & -m_R^{(3)} \end{bmatrix} \tag{4.5}$$

c.f. Baines (1985a).

This leads to the method described in Baines & Wathen (1986) and Edwards (1985).

5. LOCAL MFE

Asin Baines (1985), Baines & Wathen (1986), Edwards (1985) a purely local MFE method may be generated by replacing the C on the left hand side of (2.9) by Δ where Δ is a matrix diagonal in blocks Δ_k, each block consisting of the unit matrix I_k multiplied by the size (length, area or volume) of the element k (c.f. (2.3)). The presence of Δ_k preserves the conservation property of the method while keeping the connections with other elements local. For example, in one dimension

$$C_k = \tfrac{1}{2}\Delta s_k \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \qquad \Delta_k = \Delta s_k \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.1}$$

This corresponds to carrying out the minimisation of the L_2 norm of (2.10) with the weight $C^{\frac{1}{2}}$ replaced by $\Delta_k^{\frac{1}{2}}$. (It is also the result of a Petrov-Galerkin formulation (see Baines & Wathen (1986)).

Writing

$$\Gamma_k = \Delta_k C_k^{-1} \quad , \quad \Delta = \text{diag}\ \{\Delta_k\}, \quad \Gamma = \text{diag}\ (\Gamma_k), \tag{5.2}$$

where Γ_k is purely numerical, we find that the MFE equations (2.9) become

$$M^T \Delta M \dot{\underline{y}} = M^T \Gamma^{-1} \underline{b} \tag{5.3}$$

Constraints work in the same way as in Section 2. If some elements of $\dot{\underline{y}}$ are constrained leaving the unconstrained elements in $\dot{\underline{y}}^*$, we write as before

$$\dot{\underline{y}} = R\dot{\underline{y}}^* \qquad (MR)^T \Delta MR\dot{\underline{y}}^* = (MR)^T \Gamma^{-1} \underline{b} \tag{5.4}$$

and the solution via the preconditioned conjugate gradient method is as efficient as before.

6. TIME STEPPING

In the FFE method, time stepping in the fully discrete method is a matter of accuracy and stability, the stability often being triggered by oscillatory behaviour. In the unconstrained MFE method time stepping is concerned with accuracy but also with node overtaking. The application of constraints is intended to overcome the limitation of miniscule time steps which are forced to prevent node overtaking, but full constraints (FFE) may lead to equally small time steps for stability reasons. It may be possible to build this into the monitor.

7. CONCLUSION

We have shown that prescribed linear constraints may easily be incorporated into the Moving Finite Element method in a consistent way. The technique may be described as constrained assembly. In particular, parallelism, boundary conditions and shocks may all be described as constraints.

If diffusion is a prominent mechanism in the equation to be solved these constraints can be used to couple nodal motions, thus providing a more rapid way of transferring information across the solution region than would be the case with the standard localised method.

REFERENCES

BAINES, M.J. (1985a) "On the Approximate Solution of Partial Differential Equations by the Moving Finite Element Method". Num. Anal. Rpt. 1/85, Dept. of Mathematics, Univ. of Reading.

BAINES, M.J. (1985b) "Locally Adaptive Moving Finite Elements". Num. Anal. Rpt. 9/85, Dept. of Mathematics, Univ. of Reading.

BAINES, M.J. (1985c) "Projections and Constraints". Num. Anal. Rpt. 15/85, Dept. of Mathematics, Univ. of Reading.

BAINES, M.J. & WATHEN, A.J. (1986) "Moving Finite Element Modelling of Compressible Flow", submitted to Applied Numerical Mathematics.

EDWARDS, M.G. (1985) "The Mobile Element Method", Num. Anal. Rpt. 20/85, Dept. of Mathematics, Univ. of Reading.

HERBST, B.M. (1981) "Moving Finite Element Methods for the Solution of Evolution Equations", Ph.D. Thesis, University of the Orange Free State.

JOHNSON, I.W. (1985) "Moving Finite Element Methods for Diffusion Problems in One and Two Dimensions". Num. Anal. Rpt. 12/85, Dept. of Mathematics, Univ. of Reading.

MILLER, K. & MILLER, R. (1981) "Moving Finite Elements Parts I and II", SIAM J. Numer. Anal., 18, 1019-1057.

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

WATHEN, A.J. (1985) "Attainable Eigenvalue Bounds for the Galerkin Mass Matrix", submitted to IMA J. Numer. Anal.

WATHEN, A.J. & BAINES, M.J. (1985) "On the Structure of the Moving Finite Element Equations", IMA J. Numer. Anal. 5, 161-182.

APPENDIX

Here we follow the argument of Wathen (1985) to show that the eigenvalue spectrum of R^TAR (see (2.19)), when preconditioned by its $(d+1) \times (d+1)$ diagonal blocks, lies in the real interval $\left(\frac{1}{2}, 1 + \frac{d}{2}\right)$.

Let

$$R^TAR = R^TM^TCMR = (MR)^TCMR \tag{A1}$$

where C is a blocked diagonal matrix (with $(d+1) \times (d+1)$ blocks). Let D be the matrix of diagonal blocks of R^TAR and consider the matrix

$$R^TAR - \lambda D = (MR)^TCMR - \lambda D. \tag{A2}$$

Now, since C is a rank 1 matrix (see (2.3) et seq.) we can write the diagonal D as

$$D = (MR)^TD_CMR \tag{A3}$$

where D_C is the diagonal of C. Thus, from (A2) and (A3),

$$R^TAR - \lambda D = (MR)^T(C-\lambda D_C)MR \quad . \tag{A4}$$

If λ is an eigenvalue of $D^{-1}R^TAR$ there exists from (A4) an $\underline{x} \neq 0$ such that

$$0 = \underline{x}^T(R^TAR - \lambda D)\underline{x} \tag{A5}$$

$$= (MR\underline{x})^T(C-\lambda D_C)MR\underline{x}$$

$$= \underline{z}^T(C-\lambda D_C)\underline{z} \tag{A6}$$

where

$$\underline{z} = MR\underline{x} \tag{A7}$$

is not zero since MR has full column rank. So if λ is an eigenvalue of $D^{-1}R^TAR$ then $C - \lambda D_C$ is singular or indefinite. If therefore we can show that $C - \lambda D_C$ is positive definite for

$\lambda < \lambda_{min.}$ and negative definite for $\lambda > \lambda_{max.}$, then $\lambda \in [\lambda_{min}, \lambda_{max}]$. It remains to investigate the definiteness of (A6) which equals

$$\sum_k \underline{z}_k^T (C - \lambda D_C)_k \underline{z}_k \quad , \tag{A8}$$

where $\underline{z}^T = \{\underline{z}_k^T\}$ and $C - \lambda D_C = \{(C - \lambda D_C)_k\}$, and it is sufficient to consider each term of (A8).

For example, in the case of one dimensional linear elements (see (5.1)) we need only consider

$$\underline{z}_k^T \begin{bmatrix} \frac{1}{3}(1-\lambda) & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3}(1-\lambda) \end{bmatrix} \underline{z}_k \tag{A9}$$

and this leads to $\lambda_{min} = \frac{1}{2}$, $\lambda_{max} = \frac{3}{2}$. See Wathen (1985) for further details.