THE UNIVERSITY OF READING

The Solution of Steady PDEs on Adjustable Meshes in Multidimensions Using Local Descent Methods

by

M.J. Baines

Numerical Analysis Report 7/2000

The University of Reading
P O Box 220
Reading RG6 6AX
Berkshire, UK

DEPARTMENT OF MATHEMATICS

The Solution of Steady PDEs on Adjustable Meshes in Multidimensions Using Local Descent Methods

M.J.Baines

1 Introduction

The method of lines (MOL) is a technique for solving partial differential equations (PDEs) in which the parameters of a space discretization of the PDE are advanced in time through the solution of an ODE system, normally using a software package. The method can be applied to time-dependent or steady PDES, in the latter case via convergence in pseudo time. Iterative procedures are in any case necessary for steady nonlinear PDEs. The MOL has reached a high degree of sophistication, as evidenced elsewhere in this volume, and has produced impressive results.

The purpose of this chapter is to discuss the introduction of mesh movement for steady PDEs in multidimensions, using mesh locations as additional parameters. In this way we may seek an optimal mesh at the same time as finding a converged solution on that mesh. When the mesh locations are included in the parameters of the space discretization, an extended system of ordinary differential equations (ODEs) or differential algebraic equations (DAEs) is normally obtained which includes both mesh and solution parameters in a coupled way. Integration of these equations may then be carried out using the MOL, although there is a wide variety of approaches.

Adaptation via mesh movement is known as r-refinement. The underlying idea is that the numerical solution of PDEs, particularly those that have rough solutions, should use all available resources and the mesh is one of these resources. In r-refinement the solution is adaptively improved by mesh relocation, normally using a fixed amount of resource. In its simplest form the number of nodes remains unchanged and provided there is no change in connectivity there is a fixed data structure.

There are however a number of special difficulties with algorithms which involve mesh movement. Firstly, there is generally no information within the problem about how the mesh should be moved and so prescription of the movement is very much in the hands of the algorithm designer. Although there is sometimes an obvious choice for the mesh velocity, for example in Lagrangian fluid codes where the mesh is moved with the velocity of the fluid, there is in general no physically identifiable choice. Secondly, many PDEs are derived from the application of a physical principle, for example conservation of mass,

in a fixed frame of reference. If the frame of reference moves, the PDE must be modified and may not retain the physical properties on which it is based. Thirdly, mesh movement algorithms are essentially nonlinear and exhibit a high degree of complexity.

Another major difficulty is the possibility of mesh tangling. In one dimension this simply means node overtaking. In two dimensions, to take an example, a moving triangulation in which a node of a triangle crosses an opposite side may lead to the breakdown of a method, either because of singularity at the point of crossing or because of the inability of the method to function on an invalid triangulation. Therefore constraints are often built in to a method to avoid tangling.

There are two main techniques for the movement of nodes. The most well-known technique in this area is that of equidistribution but we shall be mostly concerned with techniques that use optimization, since they are valid in multi-dimensions. The two techniques overlap in some formulations. Equidistribution is a one-dimensional concept although there have recently been some significant advances in generalizing the idea to two dimensions. Links with an approximate form of multidimensional equidistribution are described in the penultimate Section of teh Chapter. We shall consider two kinds of functional to be minimized, normally associated with two different types of PDE. The first is a class of variational principles which generate PDEs of Euler-Lagrange type, which of their nature are of second order. The second is the L_2 norm of the residual associated with a discretization of the PDE, which can be used for first order equations and systems. Both finite element and finite volume discretizations will be discussed.

The existence of a functional allows at least two different approaches to generate solutions (and meshes). In the more standard approach the full (augmented) ODE system of normal equations can be solved by the MOL, which we shall refer to as the global approach. In the other approach we use a descent method on the functional, which can be implemented in a local manner (node by node) sweeping through the mesh, which we shall refer to as the local approach.

An early moving mesh method was the Moving Finite Element (MFE) method ([1],[5],[4]), which uses piecewise linear finite elements and generates the augmented ODE system from minimization of the L_2 norm of the residual of the PDE over the time derivatives of both the nodal positions and the solution parameters. The method is a Galerkin method in which the test functions span the space of the Lagrangian time derivatives. The method is truly multidimensional and has had some notable successes, particularly for parabolic problems. However it also showed up one of the difficulties in using piecewise linear approximation in a moving node context, namely an indeterminacy when the solution and the mesh are simultaneously trying to represent a linear manifold. This problem occurs when the local curvature of the solution manifold vanishes. For this reason (as well as the difficulty of mesh tangling) regularizing terms were added to the L_2 norm in the MFE method with a number of adjustable parameters. However the effectiveness of the method was found to

depend crucially on the manipulation of these parameters and the method has not found favour with practitioners.

Nevertheless, for a class of *steady* problems derived from variational principles, the MFE method gives an optimal solution and the time-dependent form may be used as an iterative method to drive solutions to steady state (see Section 2 below).

We commence this chapter by considering the role of the MFE method in the context of optimization. Although originally formulated as an L_2 minimization the method is not an optimization method in the usual sense but simply an extended weak form of the PDE. On the other hand, for steady equations of variational type, it has been shown in [2] that the weak forms correspond to the optimization of a minimization principle in a discrete space.

The MFE philosophy incorporates a global MOL approach to the solution of the normal equations. An example of this approach for the steady MFE method is given in [2]. However, if a functional is available, there is the alternative of sweeping through the mesh using local descent methods. This is the central theme of this Chapter. Such an approach to optimization using minimization principles is given in [3] and is described in Section 3. A possible finite volume formulation is also proposed.

Section 4 of the chapter is devoted to least squares minimization, of particular relevance to first order equations and systems. The Least Squares MFE (LSMFE) method [11] and a corresponding finite volume method [7] are described, which use global and local approaches to the solution procedure, respectively. For the important case of conservation laws the finite volume procedure may be extended to systems [15], and a description of this technique forms Section 5 of the Chapter.

Finally, there is a section on the links with equidistribution (in one dimension) and approximate equidistribution (in higher dimensions), and a summary section.

2 Moving Finite Elements

The Moving Finite Element (MFE) method ([1],[5],[4]) for the time dependent PDE

$$u_t = Lu, \tag{1}$$

where u is a function of $\underline{x}$ and t and L is a space operator, is a semi-discrete moving mesh finite element method in which the node locations are allowed to depend on time. It is based on two weak forms of the PDE which can be derived from the minimization of the L_2 norm of the residual over the time derivatives of the parameters.

The approximate solution U is an explicit function of the $\underline{X}_j(t)$ (the nodal positions) of the form

$$U = \sum_j U_j \psi_j(\underline{x}) \tag{2}$$

where U_j are coefficients and the $\psi_j(\underline{x})$ are piecewise linear basis functions. Using the result

$$\frac{\partial U}{\partial \underline{X}_j} = (-\nabla U)\,\psi_j \tag{3}$$

(see e.g. [5]) the derivative of U with respect to t becomes

$$U_t = \frac{\partial U}{\partial t}|_{moving \underline{X}} = \frac{\partial U}{\partial t}|_{fixed \underline{X}} + \sum_j \frac{\partial U}{\partial \underline{X}_j}.\frac{d\underline{X}_j}{dt}$$

$$= \frac{d\hat{U}}{dt} + \sum_j (-\nabla U)\,\psi_j.\frac{d\underline{X}_j}{dt}$$

$$= \dot{U} - \nabla U.\,\dot{\underline{X}} \tag{4}$$

where the independent $\hat{U}$ and $\underline{X}$ functions have time derivatives

$$\dot{U} = \frac{d\hat{U}}{dt} = \sum_j \frac{dU_j}{dt}\psi_j, \qquad \dot{\underline{X}} = \frac{d\underline{X}}{dt} = \sum_j \frac{d\underline{X}_j}{dt}\psi_j \tag{5}$$

which are taken to be continuous functions, corresponding to the evolution of a continuous piecewise linear approximation.

From (1) and (4) minimisation of the square of the L_2 residual $\|U_t - LU\|^2_{L_2}$ over the coefficients $\dot{U}_j, \dot{\underline{X}}_j$ then takes the form

$$\min_{\dot{U}_j, \dot{X}_j} \left\| \dot{U} - \nabla U.\,\dot{\underline{X}} - LU \right\|^2_{L_2} \tag{6}$$

and, using (5), gives the MFE or extended Galerkin equations

$$\left\langle \psi_j, \dot{U} - \nabla U.\,\dot{\underline{X}} - LU \right\rangle = 0 \tag{7}$$

$$\left\langle (-\nabla U)\,\psi_j, \dot{U} - \nabla U.\,\dot{\underline{X}} - LU \right\rangle = 0 \tag{8}$$

Substituting for $\dot{U}$ and $\dot{\underline{X}}$ from (5) gives a nonlinear system of ODEs for U_j and $\underline{X}_j$ containing an extended MFE mass matrix. The system may be solved globally for the unknowns U_j and $\underline{X}_j$ by a stiff ODE package, as in the MOL.

The basic method has intrinsic singularities, however. If the gradients ∇U have components whose values are equal in adjacent elements (dubbed "parallelism" in [1]), the system of equations (7)/(8) becomes singular and must be regularized in some way. If the area of an element vanishes (i.e. a triangle

becomes degenerate) the system again becomes singular and special action is required. In Miller's MFE method penalty functions are added to the L_2 norm of the residual in (6) (see [1],[4],[5]).

Although a full understanding of the MFE method is incomplete, in the steady limit the resulting mesh has a significant optimal property. We now consider this limit.

2.1 MFE in the Steady State Limit

In many cases the MFE method may be used to generate weak forms for the approximate solution of the *steady* PDE

$$Lu = 0 \tag{9}$$

by driving the MFE solutions to convergence in pseudo-time, although not always. For scalar first order PDEs the MFE method is known to move the nodes with characteristic speeds [5] which do not generally settle down to a steady state.

From (6) the MFE method in the steady case implements the minimization

$$\min_{\dot{U}_j, \dot{\underline{X}}_j} \|LU\|_{L_2}^2 \tag{10}$$

and the steady state solution satisfies the weak forms

$$\langle \begin{pmatrix} 1 \\ -\nabla U \end{pmatrix} \psi_j, LU \rangle = 0 \tag{11}$$

Although $\dot{U}$ and $\dot{\underline{X}}$ no longer appear in LU, the minimization is over the span of their time derivatives] which is the space spanned by the functions $\{\psi_j, (-\nabla U)\,\psi_j\}$.

In order to describe the optimal property of the steady MFE method we recall the origin of PDEs of Euler-Lagrange type.

2.2 Minimization Principles and Weak Forms

A standard result in classical analysis is that minimization of the functional

$$J(F) = \int F(u, \nabla u) d\underline{x} \tag{12}$$

over a suitable class of functions yields the PDE

$$Lu = -\frac{\partial F}{\partial u} + \nabla . \frac{\partial F}{\partial \nabla u} = 0 \tag{13}$$

By a similar argument minimization of the functional (12) over the finite dimensional space spanned by the $\{\psi_j(\underline{x})\}$ (i.e. over approximations of the form (2)) on a fixed mesh yields the weak form

$$\left\langle \psi_j, \frac{\partial F}{\partial U} \right\rangle + \left\langle \nabla\psi_j, \frac{\partial F}{\partial \nabla U} \right\rangle = 0 \tag{14}$$

2.3 An Optimal Property of the Steady MFE Equations

It has been shown in [2] that minimization of the functional (12) over functions in the MFE approximation space, spanned by $\psi_j(\underline{x}), -\nabla U)\psi_j(\underline{x})$, yields the steady MFE equations (11). Hence the steady MFE equations provide an optimal U and $\underline{X}$ for variations of the functional (12) within this space.

These weak forms are

$$\frac{\partial}{\partial U_j}\int F(U,\nabla U)d\underline{x} = \left\langle \psi_j, \frac{\partial F}{\partial U} \right\rangle + \left\langle \nabla\psi_j, \frac{\partial F}{\partial \nabla U} \right\rangle = 0 \tag{15}$$

as in (14) and

$$\frac{\partial}{\partial \underline{X}_j}\int F(U,\nabla U)d\underline{x} = \left\langle \psi_j, \frac{\partial F}{\partial \underline{x}} \right\rangle + \left\langle \nabla\psi_j, \left(F - \nabla U.\frac{\partial F}{\partial \nabla U}\right) \right\rangle = 0 \tag{16}$$

where the identity

$$\nabla.\left(F - \nabla U.\frac{\partial F}{\partial \nabla U}\right) = \frac{\partial F}{\partial \underline{x}} + \frac{\partial F}{\partial U}\nabla U - \left(\nabla.\frac{\partial F}{\partial \nabla U}\right)\nabla U \tag{17}$$

has been used to derive a form of (16) which is formally suitable for piecewise linear approximation. In carrying out the integration by parts to arrive at (16) we have used the fact that the continuous piecewise linear finite element basis function ψ_j vanishes on the boundary of the patch. The result in [2] is that (15) and (16) are identical to the weak forms (11).

It may be possible to use the time-dependent MFE method as an iterative procedure to generate locally optimal meshes in the steady state. This approach has been used in [2] to generate optimal solutions with variable nodes to a number of examples of PDEs of Euler-Lagrange type. A partially regularized form of the MFE method is used in order to avoid singular behaviour and a global MOL solver employed to extract the solution. For further details see [2]. However, since the iteration need not be time accurate, the MFE mass matrix may, if desired, be replaced by any positive definite matrix.

We now come to the central theme of this chapter, which is to consider the role of descent methods in generating solutions of problems of this type using a local approach.

3 A Local Approach to Variational Principles

Since minimization principles provide a functional to monitor and reduce it is possible to take advantage of standard optimization procedures in generating local minima. For example, procedures based on descent methods give the freedom to use a local approach to iteration, which significantly reduces the complexity of problems involving mesh movement. First we recall the nature of descent methods.

3.1 Descent Methods

Descent methods are based upon the property that the first variation of a functional J with respect to a vector variable $\underline{Y}$,

$$\delta J = \frac{\partial J}{\partial \underline{Y}} \delta \underline{Y} = \underline{g}^T \delta \underline{Y} \tag{18}$$

say, is negative when

$$\delta \underline{Y} = -\tau \underline{g} = -\tau \frac{\partial J}{\partial \underline{Y}} \tag{19}$$

for a sufficiently small positive relaxation parameter τ, and therefore reduces J. Choice of τ is normally governed by a line search or a local quadratic model.

The left hand side of (19) may be preconditioned by any positive definite matrix. The Hessian gives the full Newton approach but may be approximated in various ways.

In the present context of r-adaptivity a local approach is advantageous which consists of updating the unknowns one node at a time (scalar $\underline{Y}$), using only local information. Moreover, U_j and $\underline{X}_j$ may be updated sequentially, which permits close control of the mesh movement. The updates may be carried out in a block (Jacobi iteration) or sequential (Gauss-Seidel) manner. Descent methods of this type have been used by Tourigny and Baines [6] and Tourigny and Hulsemann [3] in the L_2 case and by Roe [7] and Baines and Leary [8] in the discrete case.

First we mention a local approach to L_2 best fits with adjustable nodes.

3.2 A Local Approach to Best Fits

A minimization based on a local approach was used in [9] to generate algorithms to determine best *discontinuous* piecewise constant and piecewise linear L_2 fits to a given function in one or two dimensions, with adjustable nodes. The convergence of the one-dimensional algorithm was subsequently investigated in [6] and the method shown to reduce the L_2 norm of the residual error monotonically. The two-dimensional algorithm was modified in [6] and a procedure for improving the connectivity introduced. Convergence of the method was also considered in [6] for successively (globally) refined meshes.

A special feature of these algorithms is their local nature, the nodal and solution updates being carried out a node at a time within sweeps through the mesh. This approach not only reduces the complexity of the problem but also allows mesh tangling to be avoided relatively easily using a limiter (see e.g. [19]). Moreover, owing to the existence of a functional to minimize, edge swapping and node removal are readily incorporated.

We now turn to the approximation of PDEs and consider minimization principles using the local approach.

3.3 Direct Optimization using Minimization Principles

An early attempt to include mesh adaptation into a minimization principle was due to Delfour et al [10] who sought a finite element solution with free nodes for a variational formulation of an elliptic PDE but found significant problems with the complexity of the equations and with mesh tangling.

More recently, an iterative algorithm with variable nodes for the finite element solution of minimization problems has been described in [3] using the local approach. The criterion is that the mesh should be such that a variational "energy functional" evaluated at the finite element approximation is reduced. Each node is treated separately in sweeping through the mesh. The nodal positions are updated by a steepest descent procedure, during which a sequence of *local* finite element problems is solved, each involving very few degrees of freedom. The order of sweeping through the mesh is based on the size of the local residuals. The method is applied to a variety of minimization principles in two dimensions.

We describe the essence of the method here. A finite element approximation U is sought to optimize a convex energy functional of the form (12) in a subspace V_h such that

$$J(U) = \min_{V \in V_h(\Delta)} J(V) \tag{20}$$

where V_h is the set of piecewise linear functions defined on a triangulation Δ which is allowed to deform. The minimization is conceived in terms of solving a sequence of local problems on patches of triangles $\{T_j\}$ surrounding node j (see fig. 1) and sweeping through the mesh. Each local approximation $\hat{U}$ is computed using the normal equations

$$\int_{\{T_j\}} \left(\frac{\partial F}{\partial \hat{U}}(\hat{U}, \nabla \hat{U}) \psi_j + \frac{\partial F}{\partial \hat{\nabla} U}(\hat{U}, \nabla \hat{U}) . \nabla \psi_j \right) d\underline{x} = 0 \tag{21}$$

$\forall \psi_j \in V_h$ (cf. (14)). In a local patch such as that in fig. 1 the computation of $\hat{U}$ at node j on a fixed mesh involves the determination of only a single unknown, which can be carried out cheaply using only local information on the patch.

New local nodal positions $\underline{\hat{X}} = \underline{X}_j^{new}$ are sought within the iteration, with the corresponding solution $\hat{U} = U^{new}$, such that $J(U)$ is reduced. A steepest

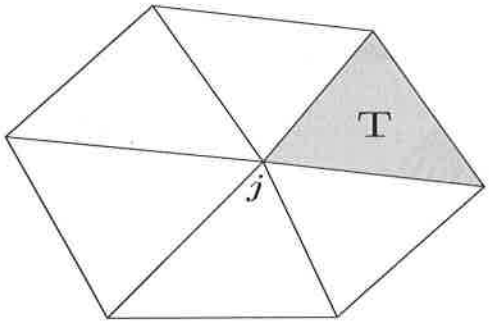

Figure 1: A local patch of elements surrounding node j.

descent method is used. Thus if $\underline{X}_j$ is an interior node, a new mesh location is sought along the line given by

$$\underline{X}_j^{new} = \underline{X}_j - \tau \frac{\partial J}{\partial \underline{X}_j}, \tag{22}$$

where the relaxation factor τ is chosen by a line search. Although this requires the solution of a sequence of finite element problems of the form (21), these are small local problems.

The sequence of nodal updates is carried out in a Gauss-Seidel manner. Edge swapping is interleaved with the grid movement algorithm, an edge being swapped if it leads to a lower energy. *Global* mesh refinement is also included in the algorithm, refinement taking place after the algorithm for the current number of nodes had converged. That is, starting from a coarse mesh the algorithm is used to optimise the mesh: this optimal mesh is then uniformly refined to provide the starting mesh for the next refinement level. The possible occurrence of degenerate triangles is overcome by the use of a node deletion algorithm. Convergence to the "global" solution relies on the sweeps through the mesh.

Convergence rates in a test problem on Laplace's equation involving a re-entrant corner showed that, whereas convergence on a uniform mesh was sub-optimal, the approximation on the adapted meshes constructed in this way converged at the optimal rate. Fig. 2 shows two of the meshes obtained using the algorithm with successive globally refined meshes. There is an exact solution of this problem of the form $r^{2/7} sin(\frac{2}{7})\theta$ (for full details see [3]).

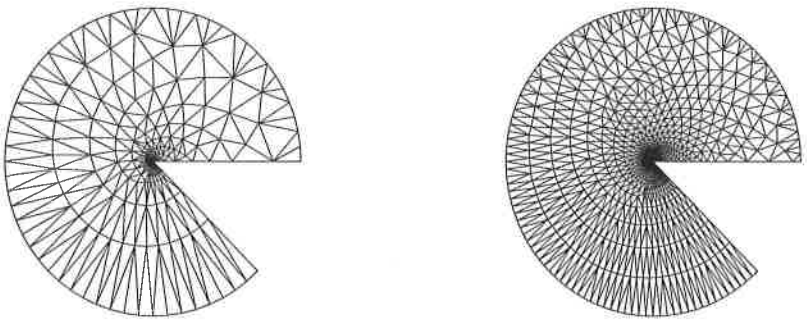

Figure 2: Two meshes for the the Re-entrant Corner problem.

3.4 A Discrete Variational Principle

A discrete form of (12) (obtained by quadrature) is

$$J_d(F) = \sum_T S_T \overline{F(U, \nabla U)}_T \tag{23}$$

where the suffix T runs over all the triangles of the mesh, S_T is the area of trangle T, and the overbar denotes the average value of the argument over the vertices of the triangle T.

Differentiation of (23) with respect to U_j gives (see [22])

$$\frac{\partial J_d}{\partial U_j} = \sum_{T_J} \left[\frac{1}{3} S_T \left(\frac{\partial F}{\partial U} \right)_j + \frac{\partial \overline{F}}{\partial (\nabla U)_T} . \underline{n}_j \right] \tag{24}$$

where $\underline{n}_j$ is the inward normal to the side opposite node j scaled by the length of that side. Setting this gradient to zero gives the finite volume weak form corresponding to the finite element weak form (14).

Differentiation with respect to $\underline{X}_j$ gives (see [17])

$$\frac{\partial I_d}{\partial \underline{X}_j} = \sum_{T_j} \left[\frac{1}{3} S_T \left(\frac{\partial \overline{F}}{\partial \underline{x}} \right)_T + \left(-\frac{\partial \overline{F}}{\partial U_y}, \frac{\partial \overline{F}}{\partial U_x} \right)_j \Delta U_j \right.$$

$$\left. + \sum_{T_j} \left(\frac{\partial \overline{F}}{\partial \nabla U} . \nabla U - F \right)_T \underline{n}_j \right] \tag{25}$$

Setting this gradient to zero gives the companion weak form to (24), corresponding to the second finite element weak form (16).

We may approximate $(\nabla U)_T$ by the gradient of the linear interpolation between the corner values of U in the triangle T, given by (see [13])

$$(\nabla U)_T = \left(\frac{-\sum' U\Delta Y}{\sum' X\Delta Y}, \frac{\sum' U\Delta X}{-\sum' Y\Delta X} \right) = \left(\frac{\sum' Y\Delta U}{\sum' X\Delta Y}, \frac{-\sum' X\Delta U}{\sum' Y\Delta X} \right) \qquad (26)$$

where the sums $\sum'$ run over the vertices of the triangle T and $\Delta X, \Delta Y, \Delta U$ denote the increments in the values of X, Y, U taken anticlockwise across the side of T *opposite* the corner concerned (see fig.1). This is the same as the piecewise linear approximation used in the finite element case. In the same notation as above the area of the triangle T is

$$S_T = \frac{1}{2}\sum{}' X\Delta Y = -\frac{1}{2}\sum{}' Y\Delta X. \qquad (27)$$

The expressions (24) and (25) with ∇U given by (26) provide gradients for a steepest descent method for minimizing (23). Examples of this approach will be given in the next section.

We turn now to Least Squares methods, which extend the applicability of the techniques already described to first order equations and systems.

4 Least Squares Methods

The steady MFE and local minimization techniques are valid and useful if there exists a minimization principle for the PDE, but are not available in other cases where no such principle exists, in particular for first order equations and systems. (Euler-Lagrange equations, by their nature, are of at least second order.) However, the same minimization techniques can also be applied to least squares methods, where the "energy functional" is the square of the norm of the residual. We describe two such methods, one of MFE type exploiting the optimal property of the steady MFE method described in section 2.3 and the other arising from a finite volume approach, similar to that in section 3.4.

In this section we shall assume that L is a first order space operator, depending on $\underline{x}, u$ and ∇u only.

4.1 Least Squares Moving Finite Elements

Although the L_2 norm (6) was minimized in formulating the MFE method described in Section 2, this minimization is only carried out over the velocities $\dot{U}_j$ and $\dot{\underline{X}}_j$ and is thus not a true minimization at the fully discrete level. Variations in U_j and $\underline{X}_j$ are treated as independent of those in $\dot{U}_j$ and $\dot{\underline{X}}_j$ and are ignored. It can therefore be seen from (6) that the method is simply a linear least squares problem, generating the weak forms (7) and (8). It is more useful

to say that the minimization is carried out in the space spanned by the basis functions $\{\psi_j, (-\nabla U)\psi_j\}$.

By contrast, a full minimization of the L_2 norm in (10) may be carried out in the steady case over the nodal coordinates $\underline{X}_j$ and the coefficients U_j. This is the approach of the recent Least Squares Moving Finite Element (LSMFE) method [11]. This is a nonlinear least squares problem so only a local minimum can be expected.

Consider then the minimisation of (10) over these parameters, which leads to the two weak forms

$$\left\langle LU, \left(\frac{\partial}{\partial U_j} LU\right)\right\rangle = \left\langle LU, (-\nabla U)\frac{\partial}{\partial U_j}(LU)\right\rangle + \oint \frac{1}{2}(LU)^2 \psi_j \underline{\hat{n}} ds = 0 \quad (28)$$

($\underline{\hat{n}}$ is the unit normal) which may be written

$$\left\langle \frac{\partial (LU)^2}{\partial U}, \psi_j \right\rangle + \left\langle \frac{\partial (LU)^2}{\partial \nabla U}, \nabla \psi_j \right\rangle = 0 \quad (29)$$

and

$$\left\langle \frac{\partial (LU)^2}{\partial \underline{x}}, \psi_j \right\rangle + \left\langle \left((LU)^2 - (\nabla U).\frac{\partial (LU)^2}{\partial \nabla U}\right), \nabla \psi_j \right\rangle = 0, \quad (30)$$

where the identity (17) has been used with $F = \frac{1}{2}(LU)^2$.

Referring back to (15)/(16) we see that equations (29)/(30) are the *steady* MFE equations for the PDE

$$0 = u_t = -\frac{\partial (Lu)^2}{\partial u} + \nabla.\frac{\partial (Lu)^2}{\partial \nabla u} \quad (31)$$

which corresponds to the Euler-Lagrange equation for the minimisation of the Least Squares functional $\|Lu\|^2_{L_2}$. For the LSMFE method the optimal property holds just as for the variational method in section 3.

It is natural to solve the nonlinear system of equations (29)/(30) using the MFE time-stepping method, but any other convenient iteration can be used. In [11] the mass matrix of the MFE method is replaced by a Laplacian regularization matrix.

4.2 Properties of the LSMFE Method

The LSMFE method has the following properties:

- The weak forms (29)/ (30) arising from these variations correspond to equations (15)/(16) with F given by $\frac{1}{2}(LU)^2$ and therefore have the optimal property.

- In the LSMFE tests carried out in [11] on scalar first order steady equations the nodes no longer move with characteristic speeds but instead move to regions of high curvature. This is to be expected because the least squares procedure embeds the original first order equation in the *second order* equation (31) and it is already known that, for Laplace's equation in one dimension, the final positions of the nodes in the MFE steady limit asymptotically equidistribute a $\frac{2}{5}$ power of the second derivative of $|u|$ [12].

- A third property is only stated here. In the particular case where LU takes the form of a divergence of a continuous function, a modification of the result discussed in section 6.3 below shows that, *asymptotically*, minimisation of $\|LU\|^2_{L_2}$ is equivalent to an equidistribution of LU over each element in the particular sense described there. For example in the case where

$$Lu = \nabla.(\underline{a}u) \tag{32}$$

with constant $\underline{a}$, and u is approximated by the continuous piecewise linear function U, the LSMFE method asymptotically equidistributes the piecewise constant residual $LU = \nabla.(\underline{a}U)$ in each element. Within a coupled iteration scheme this ensures that convergence proceeds in a relatively uniform manner.

An illustration of the results of the method for a 2-D circular advection problem is given in fig. 4 of section 5, the mesh being very similar to that given by the least squares finite volume method described there.

We now consider minimization of least squares functionals with moving nodes in which the approximations are of finite volume type.

4.3 Minimization of Discrete Norms

The use of *discrete* norms with area weighting may be regarded as a simple quadrature of the L_2 norm used previously.

Let $\overline{LU}_T$ be the average residual in triangle T and the norm be the weighted sum over triangles of the average residual, viz.

$$\|LU\|^2_d = \sum_T S_T \overline{LU}^2_T \tag{33}$$

(cf. quadrature of the square of the L_2 norm of LU). Here the suffix T runs over all the triangles of the region, S_T is the area of triangle T given by (27) and $\overline{LU}_T$ is the average value of the residual LU over the vertices of T.

This norm coincides with the L_2 norm if LU is constant on each triangle. For then

$$\|LU\|^2_{L_2} = \int_\Omega LU^2 d\underline{x} = \sum_T \int_T LU^2 d\underline{x} = \sum_T \overline{LU}^2_T \int_T d\underline{x}$$

$$= \sum_T S_T \overline{LU}_T^2 = \|LU\|_d^2 \tag{34}$$

If the area weighting in (33) is omitted this link is lost.

The form (33) may be rewritten as a sum over nodes j, namely

$$\|LU\|_d^2 = \frac{1}{3} \sum_j \sum_{T_j} S_T \overline{LU}_T^2 \tag{35}$$

where T_j runs over the patch of trangles abutting node j (cf. fig. 1).

When triangles become degenerate, however, the gradient is unbounded and the norm of the gradient is also unbounded. But by squaring the weight S_T in the norm in (33) S_T^2 a second norm

$$|||LU|||_d^2 = \sum_T S_T^2 \overline{LU}_T^2 \tag{36}$$

is obtained which is always well-defined, even on degenerate triangles.

Two moving mesh methods based on discrete Least Squares minimisation will be described here, that of Roe [7] for scalar PDEs which is applicable to problems with continuous solutions, and that of Baines, Leary and Hubbard [15] for systems of conservation laws which includes an algorithm which aligns triangle edges with shocks when flows containing such discontinuities are modelled.

4.4 Least Squares Finite Volumes

In [7] Roe included the nodal positions in the minimization of the discrete Least Squares norm (33) or (35) of the residual of a first order PDE. Roe uses the idea of a *fluctuation* instead of the residual [13], which we now define.

The fluctuation is defined as the integral of the residual in a triangle T

$$\phi_T = \int_T LU \, d\underline{x}, \tag{37}$$

or in its finite volume form, using quadrature,

$$\phi_T = S_T \overline{LU} \tag{38}$$

The discrete least squares norm of the residual, from (33) and (35), is therefore given by

$$\|LU\|_d^2 = \sum_T \frac{\phi_T^2}{S_T} = \frac{1}{3} \sum_j \sum_{T_j} \frac{\phi_T^2}{S_T} \tag{39}$$

With the alternative weight S_T^2 in (36) the discrete least squares norm of the residual is simply the l_2 norm of the fluctuation,

$$|||LU|||_d^2 = \|\phi\|_{l_2}^2 = \sum_T \phi_T^2 = \frac{1}{3} \sum_j \sum_{T_j} \phi_T^2 \tag{40}$$

By including mesh variables in the least squares minimization of (39) Roe in [7] alleviated the counting problems with the use of a fixed grid where, even though the norm of the residuals over a patch may vanish, the element residuals do not, leading to an unsatisfactory solution (see [8]). When nodal positions are included in the minimization process the number of degrees of freedom is increased and at convergence the *element* fluctuations are driven close to zero and a much improved solution is obtained.

A steepest descent method was used in [7] in which local updates of the solution and the mesh were made with a safe value of the relaxation parameter τ. The convergence of the algorithm is extremely slow but can be improved by using a more sophisticated line search [14]. However, what enormously improves the convergence rate is an updating mechanism which does not come from a full least squares descent method but which takes updates only from the upwind direction [8].

As an illustration consider the scalar two-dimensional advection equation

$$\underline{a}(\underline{x}).\nabla u = 0 \tag{41}$$

Then the fluctuation may be written

$$\phi_e = -\frac{1}{2}\sum_{ei=1}^{3}\left(a_{ei}U_{ei}\Delta Y_{ei} - b_{ei}U_{ei}\Delta X_{ei}\right) \tag{42}$$

where $\underline{a} = (a, b) = (a(X_{ei}, Y_{ei}), b(X_{ei}, Y_{ei}))$.

The steady state residual is

$$\underline{a}(\underline{x}).\nabla U \tag{43}$$

and, from (38),

$$\phi_T = \frac{(\overline{a}.\nabla U)^2}{S_T} \tag{44}$$

Then the derivatives of (39) with respect to U_j and $\underline{X}_j$ reduce to

$$\left\langle \underline{\overline{a}}.\nabla U, \underline{\overline{a}}.\frac{\partial(\nabla U)}{\partial U_j}\right\rangle_d = \left\langle \underline{\overline{a}}.\nabla U, \frac{\partial(\underline{\overline{a}}.\nabla U)}{\partial \underline{X}_j}\right\rangle_d + \frac{1}{2}\sum_{\{T_j\}}\overline{(\underline{a}.\nabla U_T)}^2\frac{\partial S_T}{\partial \underline{X}_j} = 0, \tag{45}$$

subject to boundary conditions, where from (26) and (27)

$$\frac{\partial(\nabla U)_{T_j}}{\partial U_j} = \frac{1}{2S_{T_j}}\underline{n}_j \tag{46}$$

$$\frac{\partial(\underline{\overline{a}}.\nabla U)_{T_j}}{\partial \underline{X}_j} = \Delta U_j\begin{pmatrix} -\overline{b} \\ \overline{a}\end{pmatrix} - \frac{1}{2\left(S_{T_j}\right)}(\nabla U)_{T_j} \tag{47}$$

Recall that $\underline{n}_j$ is the inward normal to the side of the triangle opposite node j scaled by the length of that side and ΔU_j is the increment in U across that side, taken anticlockwise (see fig 1).

Equations (45) may therefore be written

$$\sum_{T_j} (\underline{\bar{a}}.\nabla U)_T \, (\underline{\bar{a}}.\underline{n})_T = 0 \tag{48}$$

and

$$\sum_{T_j} \left\{ (\underline{\bar{a}}.\nabla U)_T \, \Delta U_j \begin{pmatrix} -b \\ a \end{pmatrix} - \frac{1}{2} (\underline{a}.\nabla U)_T^2 \, \underline{n} \right\} = 0 \tag{49}$$

We observe that (48) is identical to (29)) when $LU = \underline{a}.\nabla U$, noting that ∇U is constant and $\nabla \phi = S_T^{-1} \underline{n}$. However, (49) does not reduce to (30) even when $\underline{a}$ is constant.

We show an example below taken from [7].

4.5 Example

Let $\underline{a}(\underline{x}) = (y, -x)$ in a rectangle $-1 \leq x \leq 1, 0 \leq y \leq 1$. Then the solution of (41) is a semicircular annulus swept out by the initial data, here chosen to be

$$U = \begin{cases} 1 & -0.6 \leq x \leq -0.5 \\ 0 & otherwise \end{cases} \tag{50}$$

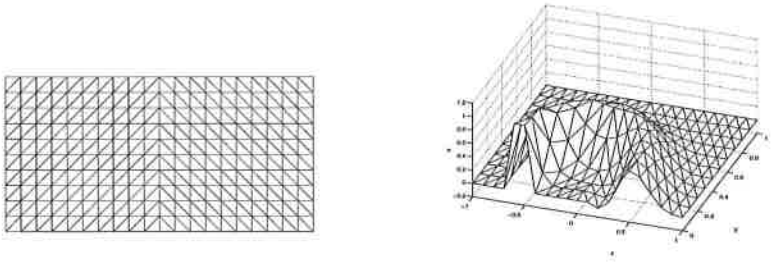

Figure 3: Initial mesh and solution for the Circular Advection problem.

Results are shown in figs. 3 and 4 for a fixed and moving mesh, respectively, taken from [15].

As expected the solution on a fixed mesh is poor. However, when the mesh takes part in the minimization the norm is driven down to machine accuracy. The redistribution effected by the least squares minimisation forces global conservation and equidistributes ϕ amongst the triangles (see Section 6.3) leading to more uniform convergence. Cell edges have approximately aligned with characteristics in regions of non-zero ϕ, allowing a highly accurate solution to be

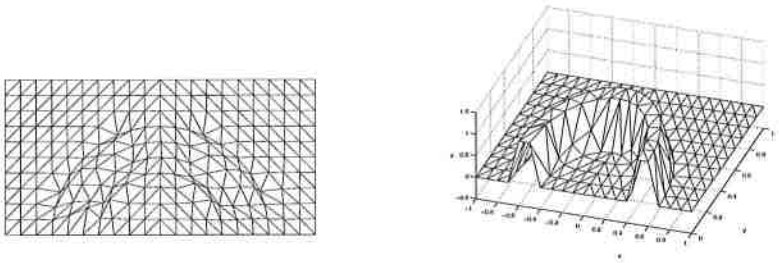

Figure 4: Final mesh and solution for the Circular Advection problem.

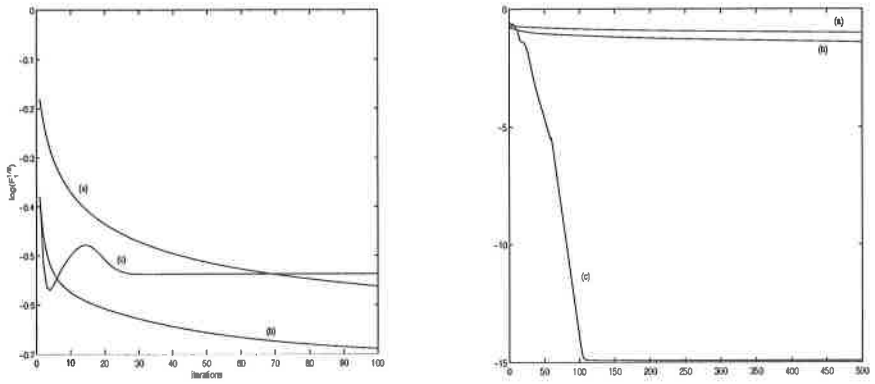

Figure 5: Comparison of convergence histories.

obtained. Essentially the same final mesh is obtained by the LSMFE method of Section 3.

The left hand graph in fig. 5 shows the convergence of the solution updating procedure on the fixed mesh using (a) steepest descent with a global relaxation factor $\tau = 0.5$, (b) optimal local updates using a quadratic model, (c) optimal local updates over downwind cells only. Convergence is much improved in (b) and (c). Even though (c) is not monotonic it converges very quickly, albeit to a higher value of the functional, due to the nature of the procedure.

The convergence rates obtained when the nodes are allowed to move are shown in the right hand graph in fig. 5. The iteration is started from the converged solution on the fixed mesh and uses (a) steepest descent with the global relaxation factors $\tau = 0.5$ for the solution and $\tau = 0.01$ for the meshpoints, (b) a line search using a Newton iteration, (c) a line search with updates over downwind cells only.

A small amount of mesh smoothing was included in (b) and (c). In particular, (b) became stuck in a local minimum if more iterations are used. Node locking was a problem with the full least squares approach. Node removal or steepest descent updates may be used to alleviate this problem ([6],[3]) but when tried in [14] still took over 1000 iterations so were not competitive when compared to the upwinding approach, which yielded the best result.

Discrete least squares solutions of the Stokes Problem have been considered in [21]. Here the two different discrete norms (39) and (40) were compared, one with the area weighting and one without, but little difference was seen in the results.

5 Conservation Laws by Least Squares

Finally we describe an advance in locating shocks in the solution of systems of nonlinear hyperbolic equations. In [15] the least squares minimizationn technique has been used for systems, combining shock capturing techniques with shock fitting.

A general system of conservation laws is of the form

$$div\underline{\mathbf{f}}(\mathbf{u}) = 0 = \underline{A}(\mathbf{u}).\underline{\nabla}\mathbf{u} \tag{51}$$

where $\underline{A}$ is a vector of the Jacobian matrices $(A, B)^T$. The integral form is

$$\oint_\Gamma \underline{\mathbf{f}}(\mathbf{u}).\underline{\hat{n}}d\Gamma = 0 \tag{52}$$

where $\underline{\hat{n}}$ is the inward facing unit normal.

It is assumed that $\underline{\mathbf{f}}$ is approximated by a piecewise linear function $\underline{\mathbf{F}}$. Then the fluctuation in triangle T is defined to be

$$\mathbf{\Phi}_T = -\int_T div\underline{\mathbf{F}}d\underline{x} = \oint_{\partial T} \underline{\mathbf{F}}.\underline{\hat{n}}ds \tag{53}$$

where ∂T is the perimeter of T. The average residual is also defined as

$$\overline{\mathbf{R}}_T = \frac{1}{S_T} \oint_{\partial T} \underline{\mathbf{F}}.\hat{\underline{n}} ds = \frac{\mathbf{\Phi}_T}{S_T} \tag{54}$$

where S_T is the area of triangle T.

Since $\underline{\mathbf{F}}$ is assumed to be linear in the triangle a trapezium rule quadrature can be used to write the fluctuation in triangle e, from (53), as

$$\mathbf{\Phi}_e = \frac{1}{2} \left\{ (\underline{\mathbf{F}}_{e1} + \underline{\mathbf{F}}_{e2}).\underline{n}_{e3} + (\underline{\mathbf{F}}_{e2} + \underline{\mathbf{F}}_{e3}).\underline{n}_{e1} + (\underline{\mathbf{F}}_{e3} + \underline{\mathbf{F}}_{e1}).\underline{n}_{e2}) \right\}, \tag{55}$$

where $\underline{n}_{ei}$ $(i = 1, 2, 3)$ is the inward unit normal to the i^{th} edge of triangle e (opposite the vertex ei), as shown in fig. 6, multiplied by the length of that edge.

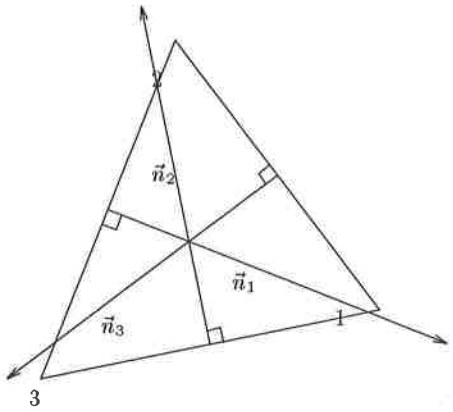

Figure 6: A general triangular cell e

It is easy to verify that, for any triangle,

$$\underline{n}_{e1} + \underline{n}_{e2} + \underline{n}_{e3} = 0, \tag{56}$$

so the fluctuation (55) may be written as

$$\mathbf{\Phi}_e = -\frac{1}{2} \left\{ \underline{\mathbf{F}}_{e1}.\underline{n}_{e1} + \underline{\mathbf{F}}_{e2}.\underline{n}_{e2} + \underline{\mathbf{F}}_{e3}.\underline{n}_{e3} \right\} \tag{57}$$

or, since $\underline{n}_{ei} = (\Delta Y_{ei}, -\Delta X_{ei})$, as

$$\mathbf{\Phi}_e = -\frac{1}{2} \sum_{ei=1}^{3} (\mathbf{F}_{ei} \Delta Y_{ei} - \mathbf{G}_{ei} \Delta X_{ei}) \tag{58}$$

(cf. (42), where $\underline{\mathbf{F}} = (\mathbf{F}, \mathbf{G})$ and $(\Delta_{e1} X, \Delta_{e1} Y) = (X_{e2} - X_{e3}, Y_{e2} - Y_{e3})$ denotes the difference in X taken across the side opposite node $e1$ in an anticlockwise

sense (with similar definitions for $(\Delta_{e2}X, \Delta_{e2}Y)$ and $(\Delta_{e3}X, \Delta_{e3}Y)$) (see fig. 6). A useful dual form of the fluctuation is obtained by rewriting (58) as

$$\mathbf{\Phi}_e = \frac{1}{2}\sum_{ei=1}^{3}\left(Y_{ei}\Delta\mathbf{F}_{ei} - X_{ei}\Delta\mathbf{G}_{ei}\right) \tag{59}$$

We aim to set the fluctuations $\mathbf{\Phi_e}$ to zero in order to minimize a vector form of (35).

Two special systems of interest are the Shallow Water equation system and the Euler equations of gasdynamics. Details of these systems are given in e.g. [14].

5.1 Use of Degenerate Triangles

In the presence of shocks least squares methods give inaccurate solutions which are unacceptable. In [15] a way of combatting this problem is shown which is to divide the region into two domains and introduce degenerate triangles at the interface. The least squares method with moving nodes is then used to adjust the position of the discontinuity, as in shock fitting methods.

An initial approximate solution to the equations can be found by any standard method. In [15] a multidimensional upwinding shock capturing scheme is used. An initial discontinuous solution is then constructed by introducing degenerate (vertical) triangles in the regions identified as shocks, using a shock identification technique. This step is carried out manually in [15] although degenerate triangles can be added automatically using techniques that exist in the shock fitting literature (see for example [16]). The corners of the degenerate triangles are designated as shocked nodes and these form an internal boundary, on either side of which the least squares method is applied in the two smooth regions where it is known to perform well. The position of the discontinuity is then improved by minimising a shock functional which is derived from (40).

Consider the interface shown in figure 7. The fluctuations $\mathbf{\Phi}_{d1}$ and $\mathbf{\Phi}_{d2}$ in adjacent degenerate triangles d_1 and d_2 on the edges $i = (i_L, j_L)$ and $j = (i_R, j_R)$ are, from (57),

$$\mathbf{\Phi}_{d_1} = -\frac{1}{2}\left[\underline{\mathbf{F}}_i\right].\underline{n}_{i_L} \qquad\qquad \mathbf{\Phi}_{d_2} = -\frac{1}{2}\left[\underline{\mathbf{F}}_j\right].\underline{n}_{j_R} \tag{60}$$

respectively, where the square bracket denotes the jump across the discontinuity.

Then

$$\mathbf{\Phi}_{d_1}^T\mathbf{\Phi}_{d_1} + \mathbf{\Phi}_{d_2}^T\mathbf{\Phi}_{d_2} = \frac{1}{4}\left\{\left([\underline{\mathbf{F}}_i].\underline{n}_{i_L}\right)^T\left([\underline{\mathbf{F}}_i].\underline{n}_{i_L}\right) + \left([\underline{\mathbf{F}}_j].\underline{n}_{j_L}\right)^T\left([\underline{\mathbf{F}}_j].\underline{n}_{j_L}\right)\right\} \tag{61}$$

and the functional

$$\sum_{T\in\Omega_D}\mathbf{\Phi}_T^T\mathbf{\Phi}_T \tag{62}$$

(cf. (40)) is minimized to improve the position of the shock, where Ω_D is the set of degenerate triangles. This norm is always bounded, even at shocks where $\mathbf{U}$ is discontinuous. On the other hand, the average residual, given by (39), is not bounded at shocks.

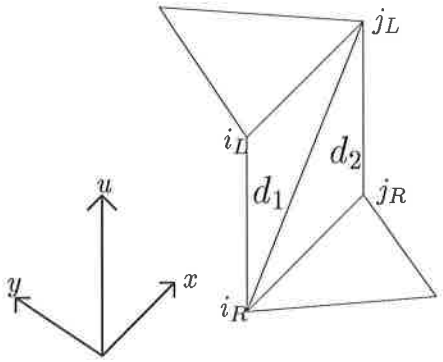

Figure 7: Degenerate triangles d_1, d_2.

A recent result of Nishikawa, Rad and Roe [28] somewhat surprisingly suggests that the capability of fluctuation splitting methods to capture characteristics or shocks depends on the quadrature used in defining the fluctuation.

A descent least squares method is used on (62) to move the shocked nodes into a more accurate position. The procedure is interleaved with a descent least squares method on (39) for the smooth solution on either side.

When updating the nodal positions $\underline{X}_{i_L}$ and $\underline{X}_{i_R}$ it is required that they have the same update (so that the cell remains degenerate). The update comes from minimization with respect to their common position vector. Degenerate quadrilaterals can be used instead of degenerate triangles.

We reproduce here two results using this technique, taken from [15].

5.2 Numerical Results for Discontinuous Solutions

Results are shown from two problems which exhibit discontinuous solutions, one for the Shallow Water equations and the other for the Euler equations of gasdynamics.

The Shallow Water equations system can be used to describe the problem of a transcritical constricted channel flow which exhibits a hydraulic jump in the constriction. The computational domain represents a channel of length 3 metres and width 1 metre with a 10% bump in the middle third. The freestream Froude number is defined to be $F_\infty = 0.55$, the freestream depth is $h_\infty = 1m$ (and the freestream velocity is given by $(u_\infty, v_\infty) = (1.72, 0)$.) An initial solution is found by the Elliptic-Hyperbolic Lax-Wendroff multidimensional upwinding scheme of

Mesaros and Roe, see [17]. The hydraulic jump is then located and degenerate quadrilaterals added at the approximate position of the shock. The best position of the shock is then sought using a least squares descent method with degenerate triangles, moving the nodes to improve the position of the shock.

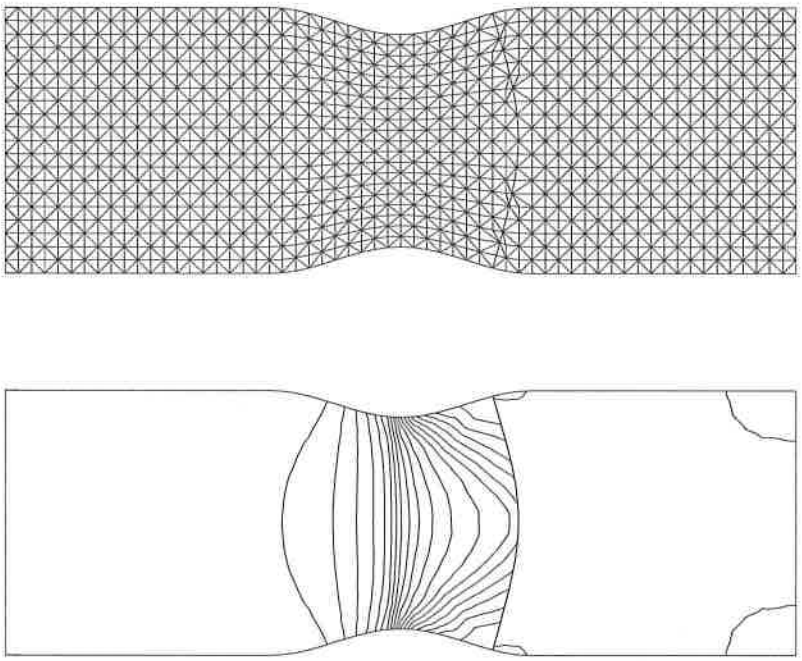

Figure 8: Mesh and Height Contours for the Shallow Water example.

Results are shown in fig. 8 which shows the meshes and solution depth contours obtained. A bow-shaped hydraulic jump which is strongest at the boundaries is predicted. This agrees with the solution obtained using a shock capturing solution on a very fine mesh. Here it is achieved sharply at very much less cost.

In the second example the Euler equations of gasdynamics are considered written in conserved variables.

The example chosen exhibits the shock fitting capabilities of the method for a purely supersonic flow which has an exact solution [18]. The computational domain is of length 3 metres and width 1 metre. Supersonic inflow boundary conditions, given by

$$\underline{U}(0,y) \quad = \quad (1.0, 2.9, 0, 5.99073)^t$$

$$\underline{U}(x,1) \;=\; (1.69997, 4.45280, -0.86073, 9.87007)^t, \qquad (63)$$

are imposed on the left and upper boundaries, respectively. At the right hand boundary supersonic outflow conditions are applied, while the lower boundary is treated as a solid wall.

The boundary conditions are chosen so that the shock enters the top left hand corner of the region at an angle of 29^o to the horizontal and is reflected by a flat plate on the lower boundary. The flow in regions away from shocks is constant. The same strategy is employed as in the previous example, with the results shown in figure 9 where the mesh and the density are shown. The solution has a shock which comes in from the top left hand at an angle of 29.2^o to the horizontal and is virtually constant apart from the discontinuities, in close agreement with the analytic solution.

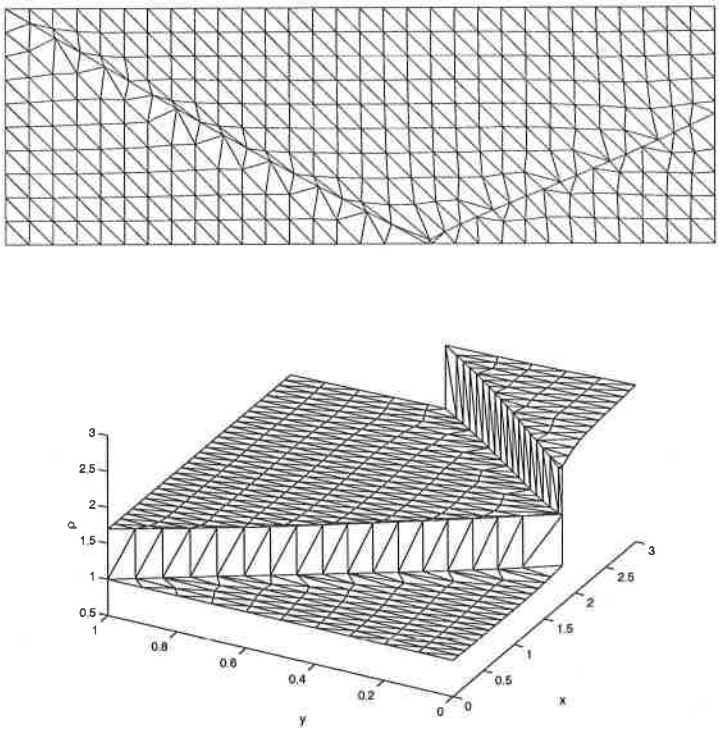

Figure 9: Mesh and Density for the Euler equations example.

The final section in this chapter highlights the links between the minimization procedures discussed previously and the ideas of equidistribution.

6 Links with Equidistribution

The well-known equidistribution principle (EP) in one dimension involves locating meshpoints such that some measure of a function is equalised over each subinterval [22]. In one dimension, denoting by x and ξ the physical and computational coordinates, respectively, define a coordinate transformation

$$x = x(\xi) \qquad \xi \in [0,1] \tag{64}$$

with fixed end points $x(0) = a, x(1) = b$, say. The computational coordinates are given by

$$\xi_i = \frac{i}{N} \quad , \quad i = 0, 1, ..., N \tag{65}$$

where N is the number of mesh points

A positive monitor function function $M(u)$ is chosen that provides some desired measure of the solution u to be equidistributed. The integral form of the EP is then given by

$$\int_0^{x(\xi)} M(u)dx = \xi\theta \tag{66}$$

where $\theta = \int_0^1 M(u)dx$. Differentiating (66) twice with respect to ξ gives the alternative differential form

$$\frac{\partial}{\partial\xi}\left(M(u)\frac{\partial x}{\partial\xi}\right) = 0. \tag{67}$$

In practice a discretized form of (67)

$$M_{j-1/2}(x_j - x_{j-1}) = M_{j+1/2}(x_{j+1} - x_j) \tag{68}$$

may be solved subject to the boundary conditions $x(0) = a, x(1) = b$.

The method has become very popular in many contexts. However, different monitor functions are often required for different purposes [22]).

Since the monitor function depends on u, which depends in turn on x, an iteration procedure is needed to solve (68). More specifically, the monitor function depends on the solution of the PDE, so equation (68) should be thought of as just one step in an iterative algorithm for both the mesh and the solution.

In iteratively solving (68) a single step of an iteration for each equation may be generated and the mesh iterations interleaved with the iterations for solving the PDE. These iteration steps may be chosen to involve only one node at a time (so that the iteration is tantamount to a sweep through the mesh) and then we have a sequence of local problems as in sections 3-5 above.

6.1 Approximate Multidimensional Equidistribution

Equidistribution was conceived as a technique for approximation in one dimension. Nevertheless there have recently been important developments in multidimensional equidistribution, (see [23],citeBaines99b). However, we shall discuss only approximate generalizations to higher dimensions here, since these are simple to implement and relate to other ideas in this Chapter.

A formal generalization of (68) is

$$\sum_{e\in\{T_j\}} M_e(\overline{\underline{x}}_e^n - \underline{x}_j^n) = 0 \tag{69}$$

where $\overline{\underline{x}}_e^n$ is the centroid of triangle e, M_e is a weight and $\{T_j\}$ is the set of triangles surrounding node j. (see fig. 1). Although the formula is convex, there are examples of meshes in which mesh tangling can take place in this case [19].

The formula (69) is not a statement of equidistribution, but it does have an interpolatory status, intuitively equidistributing in the two limits

(a) if $M_e = constant\ \forall e$, $\underline{x}_j$ is the average of the centroids of the surrounding triangles,

(b) if one M_e dominates, say M_E, then the nodes cluster towards the centroid of the element E.

The positions of the mesh vertices may also be interpreted as the solution of the least squares minimization problem (see [24], [25])

$$\min_{\underline{x}_j} \sum_e M_e(\underline{x}_j - \underline{x}_{ej})^2, \tag{70}$$

6.2 A Local Approach to Approximate Equidistribution

Again, the use of iterative techniques facilitates a local approach in which the iteration may be carried out a node at a time, sweeping through the mesh. Consider the iteration in which mesh points are moved to weighted averages of the positions of centroids of adjacent cells.

In one dimension an iteration for the solution of (68) is of the form

$$x_j^{n+1} = \frac{M_{j-\frac{1}{2}}(x_j^n + x_{j-1}^n) + M_{j+\frac{1}{2}}(x_j^n + x_{j+1}^n)}{2(M_{j-\frac{1}{2}} + M_{j+\frac{1}{2}})}, \tag{71}$$

which is convex, convergent, and does not allow mesh tangling [19]. In two dimensions a corresponding iteration for (69) is

$$\underline{x}_j^{n+1} = \frac{\sum_{e\in\{T_j\}} M_e \overline{\underline{x}}_e^n}{\sum_{e\in\{T_j\}} M_e} \tag{72}$$

6.3 Approximate Equidistribution and Conservation

A link between discrete least squares and equidistribution is described in [20] where it is shown that least squares minimisation of the residual of the divergence of a vector field is equivalent to that of a least squares measure of "equidistribution" of the residual.

The conservation law (51) is considered where $\underline{\mathbf{u}}$ is approximated by the continuous approximation $\underline{\mathbf{U}}$. The fluctuation $\mathbf{\Phi_e}$ is defined as in (53) and the average residual as in (54). Then the following identity holds.

$$\left(\sum_{i=1}^{N} S_i\right)\left(\sum_{i=1}^{N} \bar{\mathbf{R}}_T^T S_T \bar{\mathbf{R}}_T\right)$$

$$= \left(\sum_{i=1}^{N} \mathbf{\Phi_i}\right)^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\bar{\mathbf{R}}T_i - \bar{\mathbf{R}}T_j)^T S_{T_i} S_{T_j}(\bar{\mathbf{R}}_{T_i} - \bar{\mathbf{R}}_{T_j}) \tag{73}$$

Now

$$\sum_{i=1}^{N} S_{T_i} = \Omega \tag{74}$$

is equal to the total area of the union of the triangles Ω and may be taken to be constant. Moreover, by definition,

$$\sum_{i=1}^{N} \mathbf{\Phi_i} = -\sum_{\mathbf{i=1}}^{\mathbf{N}} \int_{\mathbf{e}} \mathbf{div}\underline{\mathbf{F}}(\underline{\mathbf{u}})\mathbf{d}\underline{\mathbf{x}} = \oint_{\partial\mathbf{\Omega}} \underline{\mathbf{F}}(\underline{\mathbf{u}}).\mathbf{d}\underline{\mathbf{s}} \tag{75}$$

by internal cancellation, which is independent of interior values of $\underline{\mathbf{F}}$ and interior mesh locations.

We may then write (73) as

$$\Omega\|\bar{\mathbf{R}}\|_{l_2}^2 = \left(\oint_{\partial\Omega} \underline{\mathbf{F}}.ds\right)^2 + |||\bar{\mathbf{R}}|||_{eq}^2 \tag{76}$$

where

$$\|\bar{\mathbf{R}}\|_{l_2}^2 = \sum_{i=1}^{N} \bar{\mathbf{R}}_T^T S_T \bar{\mathbf{R}}_T \tag{77}$$

corresponding to (39) and where

$$|||\bar{\mathbf{R}}|||_{eq}^2 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\bar{\mathbf{R}}_{T_i} - \bar{\mathbf{R}}_{\mathbf{j}_T})^T S_{T_i} S_{T_j}(\bar{\mathbf{R}}_{T_i} - \bar{\mathbf{R}}_{T_j}) \tag{78}$$

which is a measure of equidistribution of the average residual $\bar{\mathbf{R}}_T$.

A similar result can be derived for the norm (40). The identity

$$N\|\mathbf{\Phi}\|_{\mathbf{l_2}}^2 \equiv \left(\oint_{\partial\Omega} \underline{\mathbf{F}}.\mathbf{d}\underline{\mathbf{s}}\right)^2 + |||\mathbf{\Phi}|||_{\mathbf{eq}}^2 \tag{79}$$

holds, where N is the number of cells and

$$\|\mathbf{\Phi}\|_{\mathbf{l_2}}^2 = \sum_{\mathbf{i=1}}^{\mathbf{N}} \mathbf{\Phi}_{\mathbf{T_i}}^{\mathbf{T}} \mathbf{\Phi}_{\mathbf{T_i}} \tag{80}$$

proportional to (40), and

$$|||\mathbf{\Phi}|||_{\mathbf{eq}}^2 = \frac{\mathbf{1}}{\mathbf{2}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{j=1}}^{\mathbf{N}} (\mathbf{\Phi}_{\mathbf{T_i}} - \mathbf{\Phi}_{\mathbf{T_j}})^{\mathbf{T}} (\mathbf{\Phi}_{\mathbf{T_i}} - \mathbf{\Phi}_{\mathbf{T_j}}) \tag{81}$$

If we allow only interior mesh points to be varied, then (75) is a fixed quantity and in any minimization procedure the two norms (77) and (78) (or (80) and (81)) will be minimized simultaneously. The minimization of (77) (or (80)) (corresponding to finding a least squares approximation to the solution of (51)) is equivalent to minimising a measure of equidistribution over the triangles in the sense of (78) or (81). This result holds in any number of dimensions and in an iterative context encourages convergence to take place in a uniform way.

Finite volume methods of the type discussed here may not give very accurate solutions. However, as far as the mesh is concerned, high accuracy is not crucial. A finite volume approach may therefore be sufficiently accurate for the mesh locations but for a higher order solution a more sophisticated method, such as high order finite elements or multidimensional upwinding ([26],[27]), may be required for the solution on the optimal mesh.

7 Summary

The MFE method is a Galerkin method extended to include node movement. For the PDE

$$Lu = -\frac{\partial F}{\partial u} + \nabla . \frac{\partial F}{\partial \nabla u} = 0 \tag{82}$$

the steady MFE equations provide a local optimum for the variational problem

$$\min_{U_j, \underline{X}_j} \int F(U, \nabla U) d\underline{x} \tag{83}$$

in a piecewise linear approximation space with moving nodes.

Solutions of such PDEs may also be obtained by direct minimization of (83). A local approach is possible which is advantageous in reducing the complexity

of the mesh location procedure and in applying constraints which preserve the integrity of the mesh. An approach of this kind was described in section 3.

The Least Squares Moving Finite Element method (LSMFE) is a least squares method for steady first order PDEs which includes node movement. In the steady state the LSMFE equations for $Lu = 0$ are equivalent to the steady MFE weak forms for the PDE

$$-\frac{\partial (Lu)^2}{\partial u} + \nabla . \left(\frac{\partial (Lu)^2}{\partial \nabla u} \right) = 0 \tag{84}$$

and therefore provide a local minimum for the variational problem

$$\min_{U_j, \underline{X}_j} \int (LU)^2 \, d\underline{x} \tag{85}$$

Moreover, if LU is the divergence of a continuous flux function then the flux across element boundaries is *asymptotically* equidistributed over the elements.

A least squares finite volume fluctuation distribution scheme with mesh movement, given by Roe in [7] is an adaptive mesh method based on minimisation of a weighted l_2 norm of the residual of a steady first order PDE over the solution and the mesh. It also uses a local approach and a steepest descent algorithm. It lacks the optimal property of LSMFE but has the property that, if LU is the divergence of a continuous flux function, then the flux across element boundaries is equidistributed over the elements in the sense of (78) or (81), thus proceeding to the steady limit in a uniform way.

For scalar problems convergence can be greatly accelerated by carrying out the iterations in an upwind manner.

For problems with discontinuities the mesh movement technique enables improvement of the location of the discontinuity in a manner akin to shock fitting. By minimising a measure of the fluctuation in degenerate triangles an initially approximate position of the shock can be manoeuvred into an accurate position. The descent least squares method may be used on either side of the shock to gain good approximations in the smooth regions of the flow.

References

[1] K.Miller. *Moving Finite Elements I (with R.N.Miller) and II.* SIAM J.Num.Anal. 18,1019-1057.(1981).

[2] P.K.Jimack. *Local Minimisation of Errors and Residuals using the Moving Finite Element Method.* University of Leeds Report 98.17, School of Computer Science.(1998).

[3] Y.Tourigny and F.Hulsemann. *A New Moving Mesh Algorithm for the Finite Element Solution of Variational Problems.* SIAM J Num Anal. 34, 1416-1438.(1998).

[4] N.N.Carlson and K.Miller. *Design and Application of a Gradient Weighted Moving Finite Element Method I: in One Dimension. II: in Two Dimensions.* SIAM J Sci Comp, 19, 728-798.(1998).

[5] M.J.Baines. *Moving Finite Elements.* Oxford University Press.(1994).

[6] Y.Tourigny and M.J.Baines. *Analysis of an Algorithm for Generating Locally Optimal Meshes for L_2 Approximation by Discontinuous Piecewise Polynomials.* Math. Comp., 66, 623-650.(1998).

[7] P.L.Roe. *Compounded of Many Simples*: In Proceedings of Workshop on Barriers and Challenges in CFD, ICASE, NASA Langley, August 1996,(Ventakrishnan, Salas and Chakravarthy (eds.)), pp 241-. Kluwer (1998).

[8] M.J.Baines and S.J.Leary. *Fluctuation and Signals for Scalar Hyperbolic Equations on Adjustable Meshes.* Com. Num. Meth. Eng.,15, 877-886 (1999).

[9] M.J.Baines. *Algorithms for Optimal Discontinuous Piecewise Linear and Constant L_2 Fits to Continuous Functions with Adjustable Nodes in One and Two Dimensions.* Math.Comp. 62, 645-669.(1994).

[10] M.Delfour et al. *An Optimal Triangulation for Second Order Elliptic Problems.* Comput.Meths.Applied Mech.Engrg., 50, 231-261 (1985).

[11] K.Miller and M.J.Baines. *Least Squares Moving Finite Elements.* OUCL report 98/06. Oxford University Computing Laboratory (1998).

[12] G.F.Carey and H.T.Dinh. *Grading Functions and Mesh Distribution.* SIAM J Num An, 22, 1028-1050 (1985).

[13] H.Deconinck, P.L.Roe and R.Struijs. A Multidimensional *Generalisation of Roe's Flux Difference Splitter for the Euler Equations.* Computers and Fluids, 22, 215 (1993).

[14] S.J.Leary. *Least Squares Methods with Adjustable Nodes for Steady Hyperbolic PDEs.* PhD Thesis, Department of Mathematics, University of Reading, UK (1999).

[15] M.J.Baines, S.J.Leary and M.E.Hubbard. *Multidimensional Least Squares Fluctuation Distribution Schemes with Adaptive Mesh Movement for Steady Hyperbolic Equations.* (2000) (submitted to SIAM J Sci Stat Comp.). See also by the same authors *A Finite Volume Method for Steady Hyperbolic Equations* . In Proceedings of Conference on Finite Volumes for Complex Applications II, (R. Vilsmeier, F. Benkhaldoun and D. Hanel (eds.)) Duisburg, July 1999, pp.787-794, Hermes (1999).

[16] J.Y.Trepanier, M.Paraschivoiu, M.Reggio, R.Camarero, *A Conservative Shock Fitting Method on Unstructured Grids*, J.Comp.Phys., 126, pp 421-433 (1996).

[17] L.M.Mesaros and P.L.Roe. *Multidimensional fluctuation splitting schemes based on decomposition methods.* Proceedings of the 12th AIAA CFD Conference, San Diego (1995).

[18] H.Yee, R.F.Warming and A. Harten. *Implicit Total Variation Diminishing (TVD) Schemes for Steady State Calculations.* J.Comp.Phys., 57, 327-366 (1985).

[19] M.J.Baines and M.E Hubbard. *Multidimensional Upwinding with Grid Adaptation.* In Numerical Methods for Wave Propagation (E.F.Toro and J.F.Clarke (eds.)). Kluwer (1998).

[20] M.J.Baines. *Least-Squares and Approximate Equidistribution in Multi-dimensions.* Numerical Methods for Partial Differential Equations, 15, pp.605-615 (1999).

[21] H.Nishikawa. *The Discrete Least Squares Method for 2D Stokes Flow.* Technical Report (unpublished), Department of Aerospace Engineering, University of Michigan (1997).

[22] E.A. Dorfi, and L.O'C. Drury. *Simple Adaptive Grids for 1D Initial Value Problems.* J.Comput.Phys., 69, 175-195 (1987).

[23] W.Huang and R.D.Russell. *Moving Mesh Strategy based upon a Gradient Flow Equation for Two-Dimensional Problems.* SIAM J. Sci Stat. Comput., 20, 998, (1999).

[24] D.Ait-Ali-Yahia, W.G.Habashi, A.Tam, M.G.Vallet and M.Fortin.*A Directionally Adaptive Finite Element Method for High Speed Flows.* Int. J. for Num. Meths. in Fluids 23, pp673-690.(1996).

[25] J.A.Mackenzie. *A Moving Mesh Finite Element Method for the Solution of Two-Dimensional Stefan Problems.* Technical Report 99/26, Department of Mathematics, University of Strathclyde, UK (1999).

[26] C.Johnson. *Finite Element Methods for Partial Differential Equations.* Cambridge University Press.(1993).

[27] M.E.Hubbard. *Multidimensional Upwinding and Grid Adaptation for Conservation Laws,* PhD Thesis, Department of Mathematics, University of Reading, UK (1996).

[28] H.Nishikawa, M.Rad and P.L.Roe. *Grids and Solutions for Residual Distribution.* Private Communication (2000).