Approximate Solutions of Partial Differential Equations on Optimal Meshes using Variational Principles

M.J.Baines

Numerical Analysis Report 7/96
Department of Mathematics,
University of Reading, RG6 2AX

Abstract

We review recent advances in the numerical approximation of solutions to differential and algebraic equations on optimal grids in one or more dimensions using variational principles. A unified approach to the problem is presented which clarifies the link between the different variational conditions that have been obtained. The description incorporates iterative algorithms which have the property that the functional tends monotonically to a limit. One such algorithm is related to the Moving Finite Element method used as an iterative solver while another generates solutions via a sequence of purely local problems. The latter approach has the advantage that mesh tangling can be avoided. The extension to time dependent problems and systems of equations is discussed.

1. Introduction

In this report we discuss the generation of both optimal meshes and optimal approximations for a range of differential and algebraic equations governed by variational principles. Currently the main approaches to variable meshes are via element subdivision [18], equidistribution [19] and techniques for maximising grid quality [20]. In each of these cases a separate solver for the approximate solution is required. The idea of seeking both approximations and nodal positions that minimise a discrete energy for a given number of elements has been considered for example in [17], where the main difficulties are seen to be the complexity of the algorithm and the problem of mesh tangling. However this approach has seen a number of recent developments. Jimack [8] has used the Moving Finite Element (MFE) method to obtain numerical approximations and grids with an optimal property, while Baines, Tourigny and Hulsemann [5], [13],[16] have generated numerical approximations and grids via sequences of local problems, again with an optimal property. In this report we shall review and compare both approaches from a variational point of view, seeking a unified approach to the two methods.

Variational problems of interest to the numerical analyst include those of minimising an error norm in the approximation of functions [1] and minimising a convex functional (an "energy") in the approximate solution of elliptic PDEs, both within finite-dimensional subspaces [2]. The variational principles of fluid mechanics and gasdynamics whereby equations of motion can be generated by finding stationary values [3],[15] are also of interest since they can be used to generate finite-dimensional approximations to the dependent variables [4]. Optimal meshes can be generated in each of these applications using the techniques reviewed here.

The layout of the paper is as follows. In Section 1 some standard variational analysis is described in which we emphasise a monotonicity property of the functional for certain choices of the variations. The analysis includes the situation when stretching of the abscissa is allowed. In Section 2 finite-dimensional approximations are introduced and in Section 3 the analysis is extended to include stretching of the mesh, giving rise to coupled variational conditions for the numerical approximation and the mesh. In Section 4 the problem is approached in a different way using constrained variations and new conditions obtained which are local in character. Section 5 is concerned with iterative methods for the solution of the variational equations which exploit the monotonicity properties of the functional and the local character of the equations established in the previous sections. One of these is an application of the MFE method used as an iterative solver while another depends on a sequence of local problems. In Section 6 the methods are generalised to include discrete time stepping methods for unsteady problems, including a modified application of MFE .The extension to several dependent variables is given in Section 7. Finally, in Section 8 we summarise the results of the discussion.

1.1. Variational analysis

Let $u(x)$ be a function twice differentiable in the space variable x and let $F(x, u, u_x)$ be a twice differentiable function of its arguments. Define the functional

$$\mathcal{I}(u) = \int_a^b F(x, u, u_x) dx \tag{1.1}$$

and consider the first variation

$$\begin{aligned} \delta\mathcal{I} &= \int_a^b \left(\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_x}\delta u_x \right) dx \\ &= \int_a^b \left(\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u_x} \right) \delta u dx \end{aligned} \tag{1.2}$$

assuming that $\delta u(a) = \delta u(b) = 0$.

We emphasise two properties. First, by Lagrange's lemma, if $\delta\mathcal{I} = 0$ for all δu in the neighbourhood of $u = u^*$ then $\mathcal{I}$ is stationary and u^* satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u_x} = 0. \tag{1.3}$$

Secondly, by choosing the u variation such that

$$\delta u = \left(-\frac{\partial F}{\partial u} + \frac{d}{dx}\frac{\partial F}{\partial u_x} \right) \delta\tau \tag{1.4}$$

within (a, b), where $\delta\tau$ is a positive constant, we have from (1.2)

$$\delta\mathcal{I} = -\int_a^b (\delta u)^2 dx \delta\tau^{-1} \leq 0 \tag{1.5}$$

with zero only if $\delta u = 0$, i.e. only if $u = u^*$. Hence $\mathcal{I}$ is strictly decreasing under the variations (1.4), in which case it is stationary and $u = u^*$ satisfying (1.3) .

In higher dimensions, if $u(\underline{x})$ is a function twice differentiable in the components of the space variable $\underline{x}$ and $F(\underline{x}, u, \underline{\nabla} u)$ is a twice differentiable function of its arguments in some domain Ω of $\underline{x}$ space,

$$\mathcal{I}(u) = \int_\Omega F(\underline{x}, u, \underline{\nabla} u) d\Omega \tag{1.6}$$

which has first variation

$$\delta\mathcal{I} = \int_\Omega \left(\frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial \underline{\nabla} u}\underline{\nabla}\delta u \right) d\Omega$$

$$= \int_\Omega \left(\frac{\partial F}{\partial u} - \underline{\nabla}. \frac{\partial F}{\partial \underline{\nabla} u} \right) \delta u d\Omega \tag{1.7}$$

provided that $\delta u = 0$ on the boundary $\partial\Omega$ of Ω.

If $\delta\mathcal{I} = 0$ for all δu in the neighbourhood of u^* then $\mathcal{I}$ is stationary and the $u = u^*$ satisfies the Euler-Lagrange equation

$$\frac{\partial F}{\partial u} - \underline{\nabla}. \frac{\partial F}{\partial \underline{\nabla} u} = 0. \tag{1.8}$$

Moreover, if the δu variations are chosen to be

$$\delta u = \left(-\frac{\partial F}{\partial u} + \underline{\nabla}. \frac{\partial F}{\partial \underline{\nabla} u} \right) \delta\tau \tag{1.9}$$

within Ω (where $\delta\tau$ is positive), then

$$\delta\mathcal{I} = -\int_\Omega \delta u^2 d\Omega \delta\tau^{-1} \leq 0, \tag{1.10}$$

zero only if $\delta u = 0$. Hence $\mathcal{I}(u)$ is strictly decreasing under the variations (1.9) unless $\delta u = 0$ in which case it is stationary and $u = u^*$.

If the functional $\mathcal{I}(u)$ is bounded below it approaches a limit under these variations at which $\delta\mathcal{I} = 0$, corresponding to a solution $u = u^*$ of the steady state Euler-Lagrange equation (1.3) or (1.8). In particular there is a unique limit if the functional $\mathcal{I}(u)$ is a strictly convex function of u.

1.2. Examples

(i) The classic example of a functional $\mathcal{I}(u)$ which is convex and therefore bounded below in this way is given by the function

$$F(u, \underline{\nabla} u) = u^2 + (\underline{\nabla} u)^2 \tag{1.11}$$

for which $\mathcal{I}(u)$ is the Sobolev norm

$$\mathcal{I}(u) = a(u, u) + b(u, u) \tag{1.12}$$

where

$$a(u, v) = \int_\Omega \underline{\nabla} u . \underline{\nabla} v d\Omega, \qquad b(u, v) = \int_\Omega uv d\Omega.$$

The stationary value u^* corresponding to $\delta\mathcal{I} = 0\ \forall \delta u$ satisfies

$$\underline{\nabla}^2 u^* - u^* = 0. \tag{1.13}$$

Moreover, if the δu variations are chosen to be

$$\delta u = \left(\underline{\nabla}^2 u - u \right) \delta\tau,$$

then $\mathcal{I}(u)$ is non-increasing, stationary only if $\delta u = 0$ in which case $u = u^*$ and (1.13) holds.

(ii) For the more general quadratic example

$$F(\underline{x}, u, \underline{\nabla} u) = p(\underline{x})\left(\underline{\nabla} u\right)^2 + q(\underline{x})u^2 - 2r(\underline{x})u \tag{1.14}$$

the equation for the stationary value of u^* is the self-adjoint equation

$$-\underline{\nabla}.\left(p(\underline{x})\underline{\nabla} u^*\right) + q(\underline{x})u^* = r(\underline{x}) \tag{1.15}$$

while the δu for which $\mathcal{I}(u)$ is non-increasing is

$$\delta u = \left\{\underline{\nabla}.\left(p(\underline{x})\underline{\nabla} u\right) - q(\underline{x})u + r(\underline{x})\right\}\delta\tau.$$

(iii) If F is independent of u_x or $\underline{\nabla} u$, equation (1.3) or (1.8) becomes the nonlinear algebraic equation

$$\frac{\partial F}{\partial u} = 0 \tag{1.16}$$

while the variation for which $\mathcal{I}$ is non-increasing is

$$\delta u = -\frac{\partial F}{\partial u}\delta\tau.$$

An example from Shallow Water flow in a channel (in which u is the depth) is [4]

$$F(x, u) = B(x)\left(\frac{Q^2}{2u} - \frac{1}{2}gu^2 + E(x)u\right) \tag{1.17}$$

where Q, g are positive constants and $B(x), E(x)$ are given (breadth and energy) functions. The function F is convex if $u^3 - Q^2/g > 0$ (supercritical) and concave if $u^3 - Q^2/g < 0$ (subcritical) switching when this quantity passes through zero. The stationary function u^* satisfies the algebraic equation

$$-\frac{Q^2}{2u^{*2}} - gu^* + E(x) = 0 \tag{1.18}$$

and the choice of δu for which $\mathcal{I}(u)$ is non-increasing is

$$\delta u = B(x)\left(\frac{Q^2}{2u^2} + gu - E(x)\right)\delta\tau$$

(in the supercritical case).

In all cases boundary conditions may be incorporated in F as required.

1.3. Stretching of the abscissae

Suppose now that the x variable also participates in the variations, giving simultaneous variations $\delta^L u$ and $\delta^L x$. By analogy with 'differentiation following the motion' the chain rule gives

$$\delta^L u = \delta u + u_x \delta^L x, \tag{1.19}$$

i.e δu is the Eulerian displacement which is converted into the Lagrangian displacement $\delta^L u$ by the addition of the $u_x \delta^L x$ term. Then with $\mathcal{I}(u)$ defined as in (1.1) the first variation becomes

$$\delta\mathcal{I} = \int_a^b \left(\frac{\partial F}{\partial u} - \frac{d}{dx}\frac{\partial F}{\partial u_x} \right) (\delta^L u - u_x \delta^L x) dx. \tag{1.20}$$

The right hand side of (1.20) vanishes when the stationary solutions u^* and x^* satisfy (1.3).

Under variations defined by

$$\delta^L u = \left\{ -\frac{\partial F}{\partial u} + \frac{d}{dx}\frac{\partial F}{\partial u_x} \right\} \delta\tau, \qquad \delta^L x = \left\{ \left(-\frac{\partial F}{\partial u} + \frac{d}{dx}\frac{\partial F}{\partial u_x} \right) (-u_x) \right\} \delta\tau$$

where $\delta\tau > 0$, (1.20) becomes

$$\delta\mathcal{I} = -\int_a^b \left\{ \left(\delta^L u\right)^2 + \left(\delta^L x\right)^2 \right\} dx \delta\tau^{-1} \leq 0. \tag{1.21}$$

Hence $\mathcal{I}(u)$ is non-increasing, stationary only if $\delta^L u = \delta^L x = 0$ in which case $u = u^*$ and $x = x^*$ satisfying (1.3).

1.4. Higher dimensions

In higher dimensions the form of (1.20) is

$$\delta\mathcal{I} = \int_\Omega \left(\frac{\partial F}{\partial u} - \underline{\nabla}\frac{\partial F}{\partial \underline{\nabla} u} \right) \left(\delta^L u - \underline{\nabla} u . \delta^L \underline{x}\right) d\Omega \tag{1.22}$$

so that $\delta\mathcal{I}$ vanishes when the stationary solution $u = u^*$ and $\underline{x} = \underline{x}^*$ satisfy (1.8).

Under the variations

$$\delta^L u = \left\{ -\frac{\partial F}{\partial u} + \underline{\nabla} . \frac{\partial F}{\partial \underline{\nabla} u} \right\} \delta\tau, \qquad \delta^L \underline{x} = \left\{ \left(-\frac{\partial F}{\partial u} + \underline{\nabla} . \frac{\partial F}{\partial \underline{\nabla} u} \right) (-\underline{\nabla} u) \right\} \delta\tau$$

it follows from (1.22) that

$$\delta\mathcal{I} = -\int_\Omega \left\{ \left(\delta^L u\right)^2 + \left(\delta^L \underline{x}\right)^2 \right\} d\Omega \delta\tau^{-1} \leq 0, \tag{1.23}$$

zero only if $\delta^L u = \delta^L \underline{x} = 0$ in which case $u = u^*$ and $\underline{x} = \underline{x}^*$ satisfying (1.8).

2. The Finite-Dimensional Case

Suppose now that the function $u(x)$ of Section 1 is approximated in one dimension by the finite-dimensional function $U(x)$ which is expanded in terms of a finite number of basis functions $\psi_j(x)$ $(j = 1, 2, ..., J)$ as

$$U(x) = \sum_{j=1}^{J} U_j \psi_j(x). \tag{2.1}$$

Then, assuming that $\psi_j(x)$ is piecewise differentiable, the first variation of $\mathcal{I}(U)$, defined as in (1.1) with u replaced by U, is

$$\delta\mathcal{I} = \sum_{j=1}^{J} \frac{\partial \mathcal{I}}{\partial U_j}\delta U_j = \sum_{j=1}^{J} \delta U_j \int_a^b \left(\frac{\partial F}{\partial U}\psi_j(x) + \frac{\partial F}{\partial U_x}\psi_j'(x)\right) dx. \tag{2.2}$$

If $\delta\mathcal{I}$ vanishes for all δU_j then U is a stationary solution U^* satisfying the weak form

$$\int_a^b \left(\frac{\partial F}{\partial U}\psi_j(x) + \frac{\partial F}{\partial U_x}\psi_j'(x)\right) dx = 0 \tag{2.3}$$

of the Euler-Lagrange equation (cf. (1.3)).

2.1. Monotonicity preserving variations

Under the variations

$$\delta U_j = \int_a^b \left(-\frac{\partial F}{\partial U}\psi_j(x) - \frac{\partial F}{\partial U_x}\psi_j'(x)\right) dx\delta\tau \tag{2.4}$$

$\forall j$ it follows from (2.2) that

$$\delta\mathcal{I} = -\sum_{j=1}^{J} \delta U^2 \delta\tau^{-1} \leq 0.$$

Alternatively, using the Galerkin form, if δU is given by

$$\int_a^b \delta U \psi_i(x) dx = \int_a^b \left(-\frac{\partial F}{\partial U}\psi_i(x) - \frac{\partial F}{\partial U_x}\psi_i'(x)\right) dx\delta\tau \tag{2.5}$$

$\forall i$ (2.2) becomes

$$\delta\mathcal{I} = -\sum_{j=1}^{J} \delta U_j \int_a^b \psi_j(x)\delta U dx = -\sum_{j=1}^{J}\sum_{i=1}^{J} \delta U_j \left\{\int_a^b \psi_i(x)\psi_j(x) dx\right\} \delta U_i \delta\tau^{-1} \leq 0 \tag{2.6}$$

$\forall i$. Provided that the quadratic form in (2.6) is positive definite, it then follows in both cases that $\delta\mathcal{I} \leq 0$, zero only if $\delta U_j = 0$ in which case $U = U^*$ satisfying (2.3).

The quadratic form in (2.6) is positive definite if the mass matrix $A = \{A_{ij}\}$, where

$$A_{ij} = \left\{ \int_a^b \psi_j(x)\psi_i(x)dx \right\}, \tag{2.7}$$

is positive definite. If the $\psi_j(x)$ are the once differentiable piecewise linear finite element hat functions, so that U in (2.1) is the piecewise linear finite element approximation with nodal values $U_j(t)$, then the mass matrix A is positive definite [6].

2.2. Higher dimensions

In higher dimensions, if $F = F(\underline{x}, U, \underline{\nabla}U)$ and

$$u \sim U = \sum_{j=1}^{J} U_j \psi_j(\underline{x}), \tag{2.8}$$

where the $\psi_j(\underline{x})$ are piecewise linear basis functions on linear simplexes such as triangles or tetrahedra then as in (2.2)

$$\delta\mathcal{I} = \sum_{j=1}^{J} \frac{\partial\mathcal{I}}{\partial U_j}\delta U_j = \sum_{j=1}^{J} \delta U_j \int_\Omega \left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_j(\underline{x}) \right) d\Omega. \tag{2.9}$$

If $\delta\mathcal{I}$ vanishes for all δU_j in the neighbourhood of U^* then $\mathcal{I}$ is stationary and $U = U^*$ satisfying the weak form

$$\int_\Omega \left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_j(\underline{x}) \right) d\Omega = 0 \tag{2.10}$$

(cf. (1.8)).

Also, under the variations

$$\delta U_j = \int_\Omega \left(-\frac{\partial F}{\partial U}\psi_j(\underline{x}) - \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_j(\underline{x}) \right) d\Omega\delta\tau \tag{2.11}$$

or, in the Galerkin form, with δU given by

$$\int_\Omega \delta U \psi_i(\underline{x}) d\Omega = \int_\Omega \left(-\frac{\partial F}{\partial U}\psi_i(\underline{x}) - \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_i(\underline{x}) \right) d\Omega\delta\tau \tag{2.12}$$

$\forall i$, then since the matrix $A = \{A_{ij}\}$ where

$$A_{ij} = \left\{ \int_\Omega \psi_i(\underline{x})\psi_j(\underline{x}) d\Omega \right\} \tag{2.13}$$

is positive definite [6], we may deduce in the same way from (2.9) that $\delta\mathcal{I} \leq 0$, zero only if $\delta U_j = 0$ in which case U is a stationary function U^* satisfying (2.10).

2.3. Algebraic Forms

To put these results in matrix-vector form, denote by $\mathbf{U}, \mathbf{U}^*$ the vectors of coefficients U_j, U_j^* and by $\mathbf{b}(\mathbf{U})$ the vector of coefficients of the $b_j(U)$ defined by

$$b_j(U) = -\int_a^b \left(\frac{\partial F}{\partial U} \psi_j(x) + \frac{\partial F}{\partial U_x} \psi_j'(x) \right) dx, \tag{2.14}$$

so that, from (2.2),

$$\delta \mathcal{I} = -\delta \mathbf{U}^T \mathbf{b}(\mathbf{U}) \tag{2.15}$$

and the weak form (2.3) of the Euler-Lagrange equation becomes

$$\mathbf{b}(\mathbf{U}^*) = \mathbf{0}. \tag{2.16}$$

It follows from (2.15) that $\mathbf{b}(\mathbf{U})$ may be regarded as a search direction for the minimisation of $\mathcal{I}$ over $\delta \mathbf{U}$ by the method of steepest descent (see Section 5). More specifically, if δU is given by (2.4) or (2.11), i.e. in matrix form

$$\delta \mathbf{U} = \mathbf{b}(\mathbf{U}) \delta \tau, \tag{2.17}$$

we have from (2.15) that

$$\delta \mathcal{I} = - \|\delta \mathbf{U}\|^{\mathbf{2}} \delta \tau^{-1} \tag{2.18}$$

and $\mathcal{I}$ is non-increasing, stationary only if $\delta \mathbf{U} = \mathbf{0}$ in which case $U = U^*$ satisfying (2.16). Similarly, if in the Galerkin form δU is given by (2.5) or (2.12), i.e. in matrix form

$$A \delta \mathbf{U} = \mathbf{b}(\mathbf{U}) \delta \tau, \tag{2.19}$$

where A is the matrix with elements (2.7) or (2.13), it follows that

$$\delta \mathcal{I} = -\delta \mathbf{U}^T \mathbf{b}(\mathbf{U}) \delta \tau^{-1} = -\delta \mathbf{U}^T A \mathbf{U} \delta \tau^{-1} \tag{2.20}$$

and again, provided that A is positive definite, $\mathcal{I}$ is non-increasing, stationary only if $\delta \mathbf{U} = \mathbf{0}$ in which case $\mathbf{U} = \mathbf{U}^*$ again satisfying (2.16).

Equations (2.18) and (2.20) are equally valid as equations which possess the appropriate limit. Indeed, all that is required is an equation of the form of (2.20) with any positive definite A. The unit matrix is the one used in (2.18) but this choice lacks the scaling properties which are possessed by (2.20) [6],[7]. A good compromise is to replace A by its diagonal in (2.20). We therefore also consider variations for which

$$D \delta \mathbf{U} = \mathbf{b}(\mathbf{U}) \delta \tau \tag{2.21}$$

(cf. (2.19)) where $D = diag\{A\}$, which may be thought of as being brought about by tampering with the test function $\psi_i(x)$ on the left hand side of (2.5), which is permissible if only the limit is sought.

2.4. A special case

In the particular case where the function F is independent of u_x or $\underline{\nabla} u$ with $F_{uu} > 0$ we have an additional property. We can then show that if the components of the matrix A in (2.20) are weighted by F_{uu} the monotonicity property of $\mathcal{I}$ is maintained and the order of the method increased, provided that the matrix (2.13) is positive definite.

To see this return to (2.12), now in the form

$$\int_\Omega \delta U \psi_i(\underline{x}) d\Omega = -\int_\Omega \frac{\partial F}{\partial U} \psi_i(\underline{x}) d\Omega \delta\tau,$$

and weight the left hand side by the function F_{UU} giving

$$\int_\Omega F_{UU} \delta U \psi_i(\underline{x}) d\Omega = -\int_\Omega \frac{\partial F}{\partial U} \psi_i(\underline{x}) d\Omega \delta\tau \tag{2.22}$$

which leads to the weak form

$$\sum_j \left(\int_\Omega \psi_i(\underline{x}) F_{UU} \psi_j(\underline{x}) d\Omega \right) \delta U_j = -\int_\Omega \frac{\partial F}{\partial U} \psi_i(\underline{x}) d\Omega \delta\tau \tag{2.23}$$

$\forall i$. If the matrix (2.13) is positive definite, then by the convexity of F the weighted matrix

$$H_{ij} = \left\{ \int_\Omega \psi_i(\underline{x}) F_{UU} \psi_j(\underline{x}) d\Omega \right\}$$

is also positive definite. In addition we have the bonus that the descent step is approximately second order, because (2.22) is a weak form of Newton's method

$$-F_{UU} \delta U = \frac{\partial F}{\partial U} \tag{2.24}$$

for the solution of $\frac{\partial F}{\partial U} = 0$.

3. Adaptivity in the Finite-Dimensional Case

We consider now adaptive mesh methods which are finite-dimensional versions of the variational methods with stretched abscissae considered in Section 1.3. In one dimension let X_j ($j = 1, 2, ..., J$) be the coordinates of J nodes, across which the approximations U or U_x may have reduced continuity, which are to be varied in a Lagrangian manner in addition to the variations of the abscissa described in Section 1.3. Then, in one dimension, taking $\mathcal{I}(u)$ to be defined as in Section 1.1 with u, x replaced by U, X and taking account of the variation in the nodal points, we have (cf. (1.20) and (2.2))

$$\delta \mathcal{I} = \delta \sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} F(X, U, U_x) dx$$

$$= \sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \left(\frac{\partial F}{\partial U} \delta U_j + \frac{\partial F}{\partial U_x} \delta U_x \right) dx + \sum_{j=1}^{J-1} \left(F_j^- - F_j^+ \right) \delta^L X_j \tag{3.1}$$

where K is the number of elements ($= J+1$) and $F^{\pm}$ refer to the values of F on either side of node j. Therefore, using (1.19),

$$\delta \mathcal{I} = \sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \left\{ \frac{\partial F}{\partial U} (\delta^L U - U_x \delta^L X) + \frac{\partial F}{\partial U_x} \left(\delta^L U - U_x \delta^L X \right)_x \right\} dx - \sum_{j=1}^{J} [F]_j \, \delta^L X_j \tag{3.2}$$

where $[F]_j$ denotes the jump in F in crossing the point j.

Assuming now that $[F]_j \delta^L X = [F \delta^L X]_j$, as is the case for piecewise linear approximation, the last term may also be written as the integral

$$\sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \frac{d}{dx} \left(F \delta^L X \right) dx$$

and the $\delta^L X$ terms in (3.2) are

$$\sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \left\{ U_x \left(-\frac{\partial F}{\partial U} \delta^L X - \frac{\partial F}{\partial U_x} (\delta^L X)_x \right) - U_{xx} \frac{\partial F}{\partial U_x} \delta^L X + \frac{d}{dx} \left(F \delta^L X \right) \right\} dx$$

or

$$\sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \left\{ \left(F \delta^L X \right)_x + F_x \delta^L X - U_x \frac{\partial F}{\partial U_x} (\delta^L X)_x \right\} dx. \tag{3.3}$$

Hence, from (3.2) and (3.3),

$$\delta \mathcal{I} = \sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \left\{ \frac{\partial F}{\partial U} \delta^L U + \frac{\partial F}{\partial U_x} (\delta^L U)_x + \frac{\partial F}{\partial X} \delta^L X + \left(F - U_x \frac{\partial F}{\partial U_x} \right) (\delta^L X)_x \right\} dx. \tag{3.4}$$

3.1. Ritz expansions

It is convenient to expand each of the functions $u \sim U$ and $x \sim X$ (in the Lagrangian frame) in terms of the same set of piecewise linear basis functions $\psi_j(x)$, as

$$U = \sum_{j=1}^{J} U_j \psi_j(x), \qquad X = \sum_{j=1}^{J} X_j \psi_j(x) \tag{3.5}$$

Then (3.4) becomes

$$\delta \mathcal{I} = \sum_{j=1}^{J} \int_{X_{j-1}}^{X_{j+1}} \left\{ \left(\frac{\partial F}{\partial U} \psi_j(x) + \frac{\partial F}{\partial U_x} \psi_j'(x) \right) \delta^L U_j \right.$$

$$+\left(\frac{\partial F}{\partial X}\psi_j(x)+\left(F-U_x\frac{\partial F}{\partial U_x}\right)\psi_j'(x)\right)\delta^L X_j\Big\}\,dx \tag{3.6}$$

(cf. [5]) and if $\mathcal{I}$ is stationary we obtain the weak forms

$$\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U}\psi_j(x)+\frac{\partial F}{\partial U_x}\psi_j'(x)\right)dx=0, \tag{3.7}$$

and

$$\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial X}\psi_j(x)+\left(F-U_x\frac{\partial F}{\partial U_x}\right)\psi_j'(x)\right)dx=0 \tag{3.8}$$

$\forall j$ of the Euler-Lagrange equation (1.3) for the simultaneous solution of the stationary values U_j^* and the X_j^*.

3.2. Monotonicity preserving variations in 1-D

If we define

$$b_j(U,X)=-\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U}\psi_j(x)+\frac{\partial F}{\partial U_x}\psi_j'(x)\right)dx, \tag{3.9}$$

$$c_j(U,X)=-\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial X}\psi_j(x)+\left(F-U_x\frac{\partial F}{\partial U_x}\right)\psi_j'(x)\right)dx, \tag{3.10}$$

then under the variations

$$\delta^L U_j=b_j(U,X)\delta\tau, \tag{3.11}$$

$$\delta^L X_j=c_j(U,X)\delta\tau \tag{3.12}$$

$\forall j$, we find from (3.6) that, as in (1.21), $\delta\mathcal{I}\leq 0$, zero only if $\delta^L U_j=\delta^L X_j=0$ in which case $U=U^*$ and $X=X^*$ satisfying (3.7) and (3.8).

Similarly, if $\delta^L U,\delta^L X$ satisfy the Galerkin forms

$$\int_{X_{j-1}}^{X_{j+1}}\delta U\psi_i(x)dx=b_i(U,X)\delta\tau, \tag{3.13}$$

$$\int_{X_{j-1}}^{X_{j+1}}(-U_x)\delta U\psi_i(x)dx=c_i(U,X)\delta\tau \tag{3.14}$$

$\forall i$, where $\delta U=\delta^L U-U_x\delta^L X$, then δI has the same property, provided that the resulting quadratic form in (3.6) is positive definite.

3.3. Algebraic forms in the adaptive case

For the matrix-vector forms in the adaptive case, write also $\mathbf{X}, \mathbf{X}^*$ as the vectors of coefficients X_j, X_j^*. Then setting

$$\delta^L \mathbf{X} = (\delta^L X_1, \delta^L X_2, ..., \delta^L X_J)$$

and

$$\mathbf{b}(\mathbf{U},\mathbf{X}) = \{b_j(U,X)\}, \qquad \mathbf{c}(\mathbf{U},\mathbf{X}) = \{c_j(U,X)\}$$

and using (3.9),(3.10) it follows from (3.6) that

$$\delta \mathcal{I} = -\mathbf{b}(\mathbf{U},\mathbf{X})\delta^L U - \mathbf{c}(\mathbf{U},\mathbf{X})\delta^L \mathbf{X}. \tag{3.15}$$

Further, introducing the composite notation

$$\mathbf{Y} = \{U_1, X_1, U_2, X_2, ..., U_J, X_J\}^T, \tag{3.16}$$

$$\mathbf{Y}^* = \{U_1^*, X_1^*, U_2^*, X_2^*, ..., U_J^*, X_J^*\}^T,$$

$$\delta^L \mathbf{Y} = \{\delta^L U_1, \delta^L X_1, \delta^L U_2, \delta^L X_2, ..., \delta^L U_J, \delta^L X_J\}^T,$$

$$\mathbf{g}(\mathbf{Y}) = \{b_1, c_1, b_2, c_2, ..., b_J, c_J\}^T, \tag{3.17}$$

equation (3.15) may be written concisely in the form

$$\delta \mathcal{I} = -\left(\delta^L \mathbf{Y}\right)^T \mathbf{g}(\mathbf{Y}). \tag{3.18}$$

Thus if $\delta \mathcal{I} = 0$ for all $\delta^L U$ and $\delta^L X$ surrounding U^*, X^*, the algebraic forms of the equations for the stationary values $\mathbf{U}^*$ and $\mathbf{X}^*$ are, from (3.15),

$$\mathbf{b}(\mathbf{U}^*,\mathbf{X}^*) = \mathbf{c}(\mathbf{U}^*,\mathbf{X}^*) = \mathbf{0},$$

or in the composite notation, from (3.18),

$$\mathbf{g}(\mathbf{Y}^*) = \mathbf{0}. \tag{3.19}$$

It is clear from (3.18) that $\mathbf{g}(\mathbf{Y})$ is a search direction for the minimisation of $\mathcal{I}$ over $\delta^L \mathbf{Y}$ by the method of steepest descent.

Choosing the variations to be

$$\delta^L \mathbf{Y} = \mathbf{g}(\mathbf{Y})\delta\tau \tag{3.20}$$

ensures from (3.18) that

$$\delta \mathcal{I} = -\left\|\delta^L \mathbf{Y}\right\|^2 \Delta\tau^{-1} \leq 0.$$

Alternatively, the Galerkin forms (3.13),(3.14) give rise to the matrix system

$$A(\mathbf{Y})\delta^L \mathbf{Y} = \mathbf{g}(\mathbf{Y})\delta\tau, \tag{3.21}$$

where $A(\mathbf{Y})$ is the mass matrix with components

$$\begin{pmatrix} \int_{X_{j-1}}^{X_{j+1}} \psi_i(x)\psi_j(x)dx & \int_{X_{j-1}}^{X_{j+1}}(-U_x)\psi_i(x)\psi_j(x)dx \\ \int_{X_{j-1}}^{X_{j+1}}(-U_x)\psi_i(x)\psi_j(x)dx & \int_{X_{j-1}}^{X_{j+1}}(-U_x)^2\psi_i(x)\psi_j(x)dx \end{pmatrix} \tag{3.22}$$

which is block tridiagonal. Then we have from (3.18) that

$$\delta\mathcal{I} = -(\delta^L\mathbf{Y})^T A(\mathbf{Y})(\delta^L\mathbf{Y})\delta\tau^{-1} \tag{3.23}$$

which is non-positive provided that $A(\mathbf{Y})$ is positive definite. In both cases $\mathcal{I}$ is then non-increasing, stationary only if $\delta^L\mathbf{Y} = 0$ in which case $\mathbf{Y} = \mathbf{Y}^*$ satisfying (3.19).

A third alternative to (3.20) and (3.21) is to define the variations by

$$D(\mathbf{Y})\delta^L\mathbf{Y} = \mathbf{g}(\mathbf{Y})\delta\tau, \tag{3.24}$$

where $D(\mathbf{Y})$ is the 2×2 (block) diagonal of $A(\mathbf{Y})$ given by (3.22), corresponding to modified basis functions on the left hand side of (3.13),(3.14), which facilitates inversion while keeping the right scaling properties [6].

The preconditioned form $D(\mathbf{Y})^{-1}A(\mathbf{Y})$ is also useful for the purpose of inverting the mass matrix $A(\mathbf{Y})$ in (3.21) [6],[7].

3.4. Higher Dimensions

In higher dimensions (3.1) becomes

$$\delta\mathcal{I} = \sum_{k=1}^{K} \int_{\Delta_k} \left(\frac{\partial F}{\partial U}\delta U + \frac{\partial F}{\partial \underline{\nabla} U}.\delta\underline{\nabla} U \right) d\Omega + \sum_{j=1}^{J} \frac{\partial}{\partial \underline{X}_j} \int_{\Delta_k} F d\Omega \tag{3.25}$$

where $\int_{\Delta_k}$ denotes integration over the elements k and where K is the total number of elements, J the total number of nodes. Then, using the higher dimensional form of (1.19), the K sum in (3.25) becomes

$$\sum_{k=1}^{K} \int_{\Delta_k} \left\{ \frac{\partial F}{\partial U}(\delta^L U - \underline{\nabla} U.\delta^L \underline{X}) + \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}(\delta^L U - \underline{\nabla} U.\delta^L \underline{X}) \right\} d\Omega. \tag{3.26}$$

Now let U and $\underline{X}$ belong to the space S_1 of piecewise linear functions, in which case $\underline{\nabla} U$ is a constant vector. Using the Ritz expansion $\delta^L \underline{X} = \sum \delta^L \underline{X}_j \psi_j(\underline{x})$, where the $\psi_j(\underline{x})$ are piecewise linear (pyramid type) basis functions, $\delta\mathcal{I}$ becomes

$$\sum_{j=1}^{J} \int_{\Delta_{kj}} \left\{ \frac{\partial F}{\partial U}(\delta^L U_j - \underline{\nabla} U.\delta^L \underline{X}_j)\psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_j(\underline{x})(\underline{\nabla}\delta^L U_j - \underline{\nabla} U.\delta^L \underline{X}_j) \right\} d\Omega$$

$$+ \sum_{j=1}^{J} \int_{\partial\Delta_{kj}} F\psi_j(\underline{x})\delta^L \underline{X}_j.\underline{n}_k ds, \tag{3.27}$$

carrying out the differentiation with respect to X_j [13], where Δ_{kj} is the patch of elements surrounding node j with boundary $\partial\Delta_{kj}$.

Since

$$\int_{\partial\Delta_{kj}} F\psi_j(\underline{x})\underline{n}_k ds = \sum_{j=1}^{J}\int_{\Delta_{kj}} \underline{\nabla}\left(F\psi_j(\underline{x})\right) d\Omega,$$

the $\delta^L\underline{X}_j$ terms in (3.27) are

$$\sum_{j=1}^{J}\int_{\Delta_{kj}} \left\{-\frac{\partial F}{\partial U}(\underline{\nabla}U.\delta^L\underline{X}_j)\psi_j(\underline{x}) - \left(\frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right)\left(\underline{\nabla}U.\delta^L\underline{X}_j\right)\right.$$
$$\left. + \underline{\nabla}\left(F\psi_j(\underline{x})\right).\delta^L\underline{X}_j\right\} d\Omega. \qquad (3.28)$$

Hence from (3.27) and (3.28) we have

$$\delta\mathcal{I} = \sum_{j=1}^{J}\int_{\Delta_{kj}} \left\{\left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right)\left(\delta^L U_j - \underline{\nabla}U.\delta^L\underline{X}_j\right)\right.$$
$$\left. +\delta^L\underline{X}_j.\underline{\nabla}\left(F\psi_j(\underline{x})\right)\right\} d\Omega. \qquad (3.29)$$

The functions U^* and X^* corresponding to a stationary value of $\mathcal{I}$ therefore satisfy the weak forms

$$\int_{\Delta_{kj}} \left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right) d\Omega = 0 = -b_j(U,X), \qquad (3.30)$$

and

$$\int_{\Delta_{kj}} \left\{-\left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right)\underline{\nabla}U + \underline{\nabla}\left(F\psi_j(\underline{x})\right)\right\} d\Omega = 0 = -\underline{c}_j(U,X), \qquad (3.31)$$

say. Differentiating the $\underline{\nabla}(F\psi_j(\underline{x}))$ term, the latter equation becomes

$$\int_{\Delta_{kj}} \left\{F\underline{\nabla}\psi_j(\underline{x}) + \frac{\partial F}{\partial\underline{X}}\psi_j(\underline{x}) - \left(\frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right)\underline{\nabla}U\right\} d\Omega = 0 \qquad (3.32)$$

since locally $\nabla^2 U = 0$.

3.5. Monotonicity preserving variations in higher dimensions

The choices of variations corresponding to (3.11) and (3.12) are

$$\delta^L U_j = b_j(U,\underline{X})\delta\tau \quad \text{and} \quad \delta^L\underline{X}_j = \underline{c}_j(U,\underline{X})\delta\tau \qquad (3.33)$$

$\forall j$, where $b_j(U,X), \underline{c}_j(U,X)$ are given by (3.30),(3.31), while the Galerkin weak forms corresponding to (3.13) and (3.14) are

$$\int_{\Delta_{kj}} \delta U\psi_i(\underline{x})d\Omega = b_i(U,\underline{X})\delta\tau \qquad (3.34)$$

and

$$\int_{\Delta_{kj}} \delta U(-\underline{\nabla}U)\psi_i(\underline{x})d\Omega = \underline{c}_i(U,\underline{X})\delta\tau \tag{3.35}$$

$\forall i$, where $\delta U = \delta^L U - \underline{\nabla}U.\delta^L \underline{X}$. In either case $\mathcal{I}$ is non-increasing provided that the relevant quadratic form is positive definite.

The algebraic forms are the same as in Section 3.3 with $\mathbf{X}$ in (3.15) replaced by $\underline{\mathbf{X}}$. The mass matrix then takes the form $A = \{A_{ij}\}$ where the blocks A_{ij} are

$$\begin{pmatrix} \int_{\Delta_{kj}} \psi_i(\underline{x})\psi_j(\underline{x})d\Omega & \int_{\Delta_{kj}}(-\underline{\nabla}U)\psi_i(\underline{x})\psi_j(\underline{x})d\Omega \\ \int_{\Delta_{kj}}(-\underline{\nabla}U)\psi_i(\underline{x})\psi_j(\underline{x})d\Omega & \int_{\Delta_{kj}}(\underline{\nabla}U)^2\psi_i(\underline{x})\psi_j(\underline{x})d\Omega \end{pmatrix}. \tag{3.36}$$

3.6. Examples

If $\mathcal{I}$ is the convex functional

$$\mathcal{I}(u) = \int_\Omega \left(u^2 + (\underline{\nabla}u)^2\right) d\Omega \tag{3.37}$$

(cf. (1.11)), the weak forms (3.30) and (3.32) become the familiar

$$\int_{\Delta_{kj}} (U\psi_j(\underline{x})d\Omega + \underline{\nabla}U.\underline{\nabla}\psi_j(\underline{x}))\, d\Omega = 0 \tag{3.38}$$

and the less familiar

$$\int_{\Delta_{kj}} \left\{\frac{1}{2}(U^2 + (\underline{\nabla}U)^2)\underline{\nabla}\psi_j(\underline{x}) - (\underline{\nabla}U.\underline{\nabla}\psi_j(\underline{x}))\underline{\nabla}U\right\} d\Omega = 0 \tag{3.39}$$

$\forall j$, to be solved simultaneously for U^* and X^*.

For the least squares best fit functional

$$\mathcal{I}(u) = \int_\Omega (u - f(\underline{x}))^2\, d\Omega \tag{3.40}$$

(3.30) reduces to

$$\int_{\Delta_{kj}} (U - f(\underline{x}))\psi_j d\Omega = 0 \tag{3.41}$$

$\forall j$, showing that U^* is the best piecewise linear fit to $f(\underline{x})$ on the optimal grid in the L_2 norm, while for the best fit functional on the H^1semi-norm

$$\mathcal{I}(u) = \int_\Omega (\underline{\nabla}u - \underline{\nabla}g(\underline{x}))^2\, d\Omega \tag{3.42}$$

(3.30) becomes

$$\int_{\Delta_{kj}} (\underline{\nabla}U - \underline{\nabla}g(\underline{x})).\underline{\nabla}\psi_j(\underline{x})d\Omega = 0, \tag{3.43}$$

$\forall j$, i.e. U^* is the best piecewise linear fit to $g(\underline{x})$ on the optimal grid in the H^1 semi-norm.

If $\mathcal{I}(u)$ is the shallow water functional with F as in (1.17), then (3.30) and (3.31) reduce to the nonlinear algebraic equations

$$\int_{X_{j-1}}^{X_{j+1}} B(x)\left(-\frac{Q^2}{U^{*2}} + gU^* + E(x)\right)\psi_j(x)dx = 0 \tag{3.44}$$

and

$$\int_{X_{j-1}}^{X_{j+1}} (B(x)\psi_j(x))'\left(\frac{Q^2}{2U^*} - \frac{1}{2}gU^{*2} - E(x)U^*\right)dx = 0 \tag{3.45}$$

$\forall j$. Boundary conditions can easily be incorporated (see [16]).

Iterations for the simultaneous solution of such equations with the property that $\mathcal{I}$ is a non-increasing function are considered in Section 5 below. However, there is an alternative interpretation of the variations which gives rise to local problems, which we describe first.

4. Constrained Approximation

We now demonstrate a different approach to the use of the first variation in generating optimal solutions and meshes, applying constraints to the $\mathcal{I}$ variations so as to be able to carry out Eulerian variations and nodal point variations separately. Care is required when doing the nodal variations to avoid interfering with the current approximation to the solution. The discussion will be confined to linear and constant finite element approximation on simplexes.

Thus in one dimension, from (3.2) we may write

$$\delta\mathcal{I} = \sum_{k=1}^{K}\int_{X_{k-1}}^{X_k}\left\{\frac{\partial F}{\partial U}(\delta^L U - U_x\delta^L X) + \frac{\partial F}{\partial U_x}(\delta^L U_x - U_x\delta^L X_x)\right.$$
$$\left. - \frac{\partial F}{\partial U_x}U_{xx}\delta^L X\right\}dx - \sum_{j=1}^{J}[F]_j\delta^L X_j. \tag{4.1}$$

Substituting piecewise linear Ritz expansions (3.5) we obtain

$$\delta\mathcal{I} = \sum_{j=1}^{J}\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U}\psi_j(x) + \frac{\partial F}{\partial U_x}\psi_j'(x)\right)(\delta^L U_j - U_x\delta^L X_j)dx - \sum_{j=1}^{J}[F]_j\delta^L X_j \tag{4.2}$$

since locally $U_{xx} = 0$.

Suppose now that the $\mathcal{I}$ variations are constrained by fixing the mesh, so that $\delta X = 0$. Then the stationary function U satisfies the weak form

$$\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U}\psi_j(x) + \frac{\partial F}{\partial U_x}\underline{\nabla}\psi_j'(x)\right)dx = 0 \tag{4.3}$$

$\forall j$. We shall refer to this step as stage 1.

Now in a separate step constrain the $\mathcal{I}$ variations by freezing the current U approximation, in which case $\delta U_j = U_x \delta X_j$ corresponding to U remaining on its graph (or its extrapolation) as X varies. Then from (4.2) the stationary function X satisfies

$$[F]_j = 0 \tag{4.4}$$

which we refer to as stage 2.

The solution of equations (4.3) and (4.4) may be sought via an alternating sequence of approximations converging to a limit. First equation (4.3) may be solved for U with X fixed and then (4.4) solved for X with U locally frozen, i.e. constrained by variations $\delta U_j = U_x \delta X_j$. The procedure is then repeated to convergence. Each stage has the property that the functional $\mathcal{I}$ is non-increasing (see [13],[16]).

It is clear that equation (4.4) is a local form in which the X_j may be solved separately for each j. On the other hand equation (4.3) is global, giving a set of linear equations in which all the U_j are coupled together. However, in [16] this stage is also made purely local by observing that the property $\delta I \leq 0$ may be preserved when (4.3) is replaced by a local problem on the patch of elements surrounding node j in which only U_j is allowed to vary.

From (4.2) the Galerkin steepest descent step for (4.3) is

$$\int_{X_{j-1}}^{X_{j+1}} \delta U \psi_i(x) dx = -\int_{X_{j-1}}^{X_{j+1}} \left(\frac{\partial F}{\partial U} \psi_i(x) + \frac{\partial F}{\partial U_x} \underline{\nabla} \psi_i'(x) \right) dx \delta\tau. \tag{4.5}$$

4.1. Linear discontinuous approximation

In the case where F is independent of u_x we may exploit equation (4.4) by considering approximation of U by *discontinuous* linear functions, although we shall still require that X remains continuous. The introduction of discontinuous U's allows solutions with jumps.

The expansion of U may then be written

$$U = \sum_{k=1}^{K} \sum_{\nu} W_{k\nu} \phi_{k\nu}(x), \tag{4.6}$$

say, where $\phi_{k\nu}$ are the local discontinuous basis functions and $W_{k\nu}$ are the corner values of U in each element, k referring to an element and ν to the 'corners' of the element. The one-dimensional form of $\delta\mathcal{I}$ then becomes, from (4.2),

$$\sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} \sum_{\nu} \delta^L W_{k\nu} \phi_{k\nu}(x) dx - \sum_{j=1}^{J} \int_{X_{j-1}}^{X_{j+1}} \frac{\partial F}{\partial U_x} \delta^L X_j U_x \psi_j(x) dx - \sum_{j=1}^{J} [F]_j \delta^L X_j. \tag{4.7}$$

As in the previous section we consider variations which are constrained in two ways, first by fixing X_j so that the stationary value U satisfies the weak form

$$\int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} \phi_{k\nu}(x) dx = 0 \tag{4.8}$$

$\forall k$ and $\nu = 1, 2$ (cf. (4.3)), and secondly by constraining U to lie on its own graph (or its extrapolation), so that $\delta U_j = U_x \delta X_j$, so that

$$[F]_j = 0 \tag{4.9}$$

$\forall j$ again. Equations (4.8) and (4.9) are both local problems because the former can be solved for the unknown coefficients $W_{k\nu}$ of U element by element. Note that one solution of (4.9) is that U is continuous (notwithstanding the discontinuous approximation space) [5],[13].

The Galerkin steepest descent method for (4.8) is

$$\int_{X_{k-1}}^{X_k} \delta U \phi_{k\nu}(x) dx = - \int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} \phi_{k\nu}(x) dx \tag{4.10}$$

$\forall k, \nu$. If $F_{UU} > 0$ then as in Section 2.3 this term may be inserted into the integrand on the left hand side of (4.10) to increase the accuracy of the descent method.

4.2. Piecewise constant approximation

We can repeat the above analysis using piecewise constant approximation. In this case U is expanded in terms of the piecewise constant basis functions, $\pi_k(x)$ say, in the form

$$U = \sum_{k=1}^{K} W_k \pi_k(x). \tag{4.11}$$

Equation (4.2) then becomes

$$\delta \mathcal{I} = \sum_{k=1}^{K} \int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} \pi_k(x) \delta^L W_k dx - \sum_{j=1}^{J} [F]_j \delta^L X_j. \tag{4.12}$$

In the two separate constrained stages the approximations U, X therefore respectively satisfy the local weak forms

$$\int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} \pi_k(x) dx = \int_{X_{k-1}}^{X_k} \frac{\partial F}{\partial U} dx = 0, \tag{4.13}$$

$\forall k$, corresponding to variations $\delta^L W_k$ constrained by $\delta^L X_j = 0$ (fixed mesh), and

$$[F]_j = 0 \tag{4.14}$$

$\forall j$, corresponding to variations $\delta^L X_j$ constrained by $\delta^L W_k = 0$ (in the case of piecewise constant U).

4.3. Higher dimensions

In higher dimensions consider the expression for $\delta\mathcal{I}$ in (3.26) which may be written as

$$\sum_{k=1}^{K}\int_{\Delta_k}\left(\frac{\partial F}{\partial U}(\delta^L U-\underline{\nabla}U.\delta^L\underline{X})+\frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\delta^L U+\left(\underline{\nabla}.\frac{\partial F}{\partial\underline{\nabla}U}\right)\underline{\nabla}U.\delta^L\underline{X}\right)d\Omega$$

$$-\sum_{j=1}^{J}\int_{\partial\Delta_{kj}}(\underline{\nabla}U.\delta^L\underline{X})\frac{\partial F}{\partial\underline{\nabla}U}.\underline{n}_k ds \qquad (4.15)$$

using integration by parts. Substituting the Ritz approximations (3.5) we obtain from (3.27)

$$\delta\mathcal{I}=\sum_{j=1}^{J}\int_{\Delta_{kj}}\left\{\left(\frac{\partial F}{\partial U}(\delta^L U_j-\underline{\nabla}U.\delta^L\underline{X}_j)+\left(\underline{\nabla}.\frac{\partial F}{\partial\underline{\nabla}U}\right)(\underline{\nabla}U.\delta^L\underline{X}_j)\right)\psi_j(\underline{x})\right.$$

$$\left.+\frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\delta^L U_j\right\}d\Omega+\sum_{j=1}^{J}\int_{\partial\Delta_{kj}}\left\{F\delta^L\underline{X}_j-(\underline{\nabla}U.\delta^L\underline{X}_j)\frac{\partial F}{\partial\underline{\nabla}U}\right\}.\underline{n}_k\psi_j(\underline{x})ds. \qquad (4.16)$$

Now, applying the fixed grid constraint $\delta^L\underline{X}=0$, we have from (4.16) that U satisfies

$$\int_{\Delta_{kj}}\left(\frac{\partial F}{\partial U}\psi_j(\underline{x})+\frac{\partial F}{\partial\underline{\nabla}U}.\underline{\nabla}\psi_j(\underline{x})\right)d\Omega=0 \qquad (4.17)$$

$\forall j$ as in equation (3.30), while under the frozen solution constraint $\delta^L U_j=\underline{\nabla}U.\delta^L X_j$ the equation for X is

$$\int_{\Delta_{kj}}\left(\underline{\nabla}.\frac{\partial F}{\partial\underline{\nabla}U}\right)\psi_j(\underline{x})\underline{\nabla}U d\Omega+\int_{\partial\Delta_{kj}}\left\{F\underline{n}_k-\left(\frac{\partial F}{\partial\underline{\nabla}U}.\underline{n}_k\right)\underline{\nabla}U\right\}\psi_j(\underline{x})ds=0 \qquad (4.18)$$

$\forall j$. The U variations in (4.18) move along the graph (in two dimensions on the discontinuous planes) of U or its extrapolation as $\underline{X}$ is varied [5]. If $\underline{\nabla}.\frac{\partial F}{\partial\underline{\nabla}U}=0$ (4.18) reduces to

$$\int_{\partial\Delta_{kj}}\left\{F\underline{n}_k-\left(\frac{\partial F}{\partial\underline{\nabla}U}.\underline{n}_k\right)\underline{\nabla}U\right\}\psi_j(\underline{x})ds=0 \qquad (4.19)$$

$\forall j$ which is true in particular if F is independent of X and U since locally $\nabla^2 U=0$.

If F is independent of $\underline{\nabla}U$ (4.18) also reduces to

$$\int_{\partial\Delta_{kj}}F\underline{n}_k\psi_j(\underline{x})ds=0 \qquad (4.20)$$

$\forall j$. In this case we may approximate U by *discontinuous* linear functions (with X still piecewise linear continuous). Putting

$$U = \sum_{k=1}^{K} \sum_{\nu} W_{k\nu} \phi_{k\nu}(\underline{x})$$

the conditions (4.17) and (4.18) then reduce to the truly local problems

$$\int_{\Delta_k} \frac{\partial F}{\partial U} \phi_{k\nu}(\underline{x}) d\Omega = 0 \tag{4.21}$$

$\forall k, \nu$, and (4.20) again $\forall j$. In the case of approximation by piecewise constant functions the conditions are

$$\int_{\Delta_{kj}} \frac{\partial F}{\partial U} d\Omega = 0 \tag{4.22}$$

$\forall j$ and (4.20).

In each of these cases the procedure of solving for U and X alternately ensures that $\mathcal{I}$ is a non-increasing function, stationary only when $\delta U = 0$, $\delta \underline{X} = 0$. In this context stage (4.17) may also be replaced by a local problem while still preserving the monotonicity property $\delta I \leq 0$, as in [16]. In this case equation (4.17) is solved for U_j only on the local patch Δ_{kj} with the remaining U_i fixed $(i \neq j)$.

It is clearly possible to combine both continuous and discontinuous approximation in practice, in order to model a line of discontinuity, say.

4.4. Algebraic forms

Let $\mathbf{b}(\mathbf{U}) = \{b_i(\mathbf{U})\}$ be defined as in (3.30) and let $\Theta_i(\mathbf{X})$ be

$$\int_{\Delta_k} \left(\underline{\nabla} \cdot \frac{\partial F}{\partial \underline{\nabla} U} \right) \left(\underline{\nabla} U . \underline{\nabla} \psi_i(\underline{x}) \right) d\Omega - \int_{\partial \Delta_{kj}} \left\{ F \underline{n}_k - \left(\frac{\partial F}{\partial \underline{\nabla} U} . \underline{n}_k \right) \underline{\nabla} U \right\} \psi_i(\underline{x}) ds \tag{4.23}$$

(see (4.18)), or in one dimension

$$\Theta_i(\mathbf{X}) = [F]_i. \tag{4.24}$$

Then defining $\boldsymbol{\Theta}(\mathbf{X}) = \{\Theta_i(\mathbf{X})\}$ (4.2) may be written

$$\delta \mathcal{I} = -\mathbf{b}(\mathbf{U})^T \delta \mathbf{U} - \boldsymbol{\Theta}(\mathbf{X})^T \delta \mathbf{X}, \tag{4.25}$$

vanishing when

$$\mathbf{b}(\mathbf{U}) = \mathbf{0} \tag{4.26}$$

and

$$\boldsymbol{\Theta}(\mathbf{X}) = \mathbf{0}. \tag{4.27}$$

The Galerkin form of the solution procedure for (4.26) is (3.21) whilst a correponding steepest descent step for (4.27) is

$$\delta\mathbf{X} = \boldsymbol{\Theta}(\mathbf{X})\delta\tau. \tag{4.28}$$

In the case of discontinuous approximation let

$$\mathbf{W} = (W_{11}, W_{12}, ..., W_{21}, W_{22}, ..., W_{K1}, W_{K2}, ...)^T$$

where K is the number of elements. Then, defining

$$d_{k\nu}(\mathbf{W}) = -\int_{\Delta_{kj}} \frac{\partial F}{\partial U}\phi_{k\nu}(\underline{x})d\Omega \tag{4.29}$$

(see (4.21)) and $\mathbf{d}(\mathbf{W}) = \{d_{k\nu}(\mathbf{W})\}$, an algebraic form for (4.7) is

$$\delta\mathcal{I} = -\mathbf{d}(\mathbf{W})^T\delta\mathbf{W} - \boldsymbol{\Theta}(\mathbf{X})^T\delta\mathbf{X}. \tag{4.30}$$

Equations (4.8) and (4.9) then become

$$\mathbf{d}(\mathbf{W}) = \mathbf{0} \tag{4.31}$$

and (4.27) again.

A Galerkin form of steepest descent for the solution of (4.31) is

$$E\delta\mathbf{W} = \mathbf{d}(\mathbf{W})\delta\tau \tag{4.32}$$

where E is the positive semi-definite local elementwise mass matrix having diagonal blocks

$$E_{kk} = \int_{\Omega} \phi_{k\nu}(\underline{x})\phi_{k\mu}(\underline{x})d\Omega, \tag{4.33}$$

where both ν, μ run over the corners of the element k.

In the piecewise constant case the algebraic forms are unaltered except that (4.29) is replaced by

$$d_k(\mathbf{W}) = -\int_{\Delta_k} \frac{\partial F}{\partial U}d\Omega \tag{4.34}$$

(cf. (4.22)) and E is then a diagonal matrix.

4.5. Relationship between algebraic forms in the linear case

The relationship (1.19) between Lagrangian and Eulerian variations, when applied at each point of the grid may be written algebraically as [7]

$$\mathcal{X}\delta\mathbf{Y} = \delta\mathbf{W} \tag{4.35}$$

where $\mathcal{X}$ is a block diagonal matrix whose blocks are

$$\begin{pmatrix} \mathbf{1} & -(\underline{\nabla} U_j^T) \end{pmatrix} \tag{4.36}$$

where $\mathbf{1}$ is a vector of ones and $\underline{\nabla}U_j^T$ is a vector of gradients in elements surrounding the node j; in one dimension (4.36) is

$$\begin{pmatrix} 1 & -(U_x)_j \\ 1 & -(U_x)_{j+1} \end{pmatrix}. \tag{4.37}$$

In conjunction with (4.32) this leads to

$$E\mathcal{X}\delta\mathbf{Y} = \mathbf{d}(\mathbf{W})\delta\tau$$

and hence we find that

$$\mathcal{X}^T E\mathcal{X}\delta\mathbf{Y} = \mathcal{X}^T\mathbf{d}(\mathbf{W})\delta\tau \tag{4.38}$$

from which, by comparison with (3.21), we can identify the relationships

$$A(\mathbf{Y}) = \mathcal{X}^T E\mathcal{X} \tag{4.39}$$

and

$$\mathbf{g}(\mathbf{Y}) = \mathcal{X}^T\mathbf{d}(\mathbf{W}). \tag{4.40}$$

Moreover $D(\mathbf{Y})$, the diagonal of $A(\mathbf{Y})$, is given by

$$D(\mathbf{Y}) = \mathcal{X}^T diag\{E\}\mathcal{X}. \tag{4.41}$$

In the relationship between the algebraic forms in the piecewise constant case the blocks (4.37) of the matrix $\mathcal{X}$ are replaced by the single column $\mathbf{1}$.

4.6. Examples

If $\mathcal{I}$ is the convex functional (3.37), the weak forms (4.3) and (4.4) become the familiar

$$\int_{\Delta_{kj}} (U\psi_j(\underline{x})d\Omega + \underline{\nabla}U.\underline{\nabla}\psi_j(\underline{x}))\, d\Omega = 0 \tag{4.42}$$

and the less familiar

$$\left[U^2 + (\underline{\nabla}U)^2\right]_j = 0 \tag{4.43}$$

$\forall j$, where in calculating X from (4.43) U is constrained to remain locally on its graph or the extrapolation of its graph.

For the best fit functionals (3.40) and (3.42) U is still the best fit on the optimal grid in the appropriate norm.

If $\mathcal{I}(U)$ is the shallow water functional with F as in (1.17), then (3.30) and (3.31) reduce to the nonlinear algebraic equations (3.44) and

$$\left[B(x)\left(\frac{Q^2}{2U} - \frac{1}{2}gU^2 - E(x)U\right)\right]_j = 0 \tag{4.44}$$

$\forall j$ where again in (4.44) U is constrained to remain locally on its graph or the extrapolation of its graph..

Iterations for the sequential solution of such equations with the property that $\mathcal{I}$ is a non-increasing function are considered in Section 5 below.

5. Iteration Procedures

The property of each of the previous sections, that under certain conditions $\mathcal{I}(u)$ is non-increasing and tends to a limit, holds only in an asymptotic sense. In order to obtain a practical algorithm for reaching the limit, a finite displacement is necessary which may invalidate this property. However, the monotonicity property of the variation can still be used as a guide in the construction of iterative steepest descent algorithms which may converge (in the sense that $\mathcal{I}$ converges) to a stationary point and hence to a corresponding solution of the weak form of the Euler-Lagrange equation.

There are two types of iteration suggested by the analysis. In the stretched abscissa approach of Section 3 the iterations for U and X are carried out simultaneously, while in the constrained variation approach of Section 4 the iterations are sequential. As we shall see, in the latter case the local nature of the normal equations means that inexpensive algorithms can be constructed and used with advantage.

In the case of fixed grid approximation using the finite element method of Section 2 the aim is to solve the weak form (2.10) which in its algebraic form is (2.16). From (2.15) it follows that $\mathbf{b}(\mathbf{U})$ gives a steepest descent direction for the minimisation of $\mathcal{I}$. A steepest descent iteration based on the Galerkin form (2.19) is

$$A(\mathbf{U}^{p+1} - \mathbf{U}^{p}) = \mathbf{b}(\mathbf{U})^{p}\delta\tau \tag{5.1}$$

$(p = 1, 2, ...)$ or its diagonal variant

$$D(\mathbf{U}^{p+1} - \mathbf{U}^{p}) = \mathbf{b}(\mathbf{U})^{p}\delta\tau, \tag{5.2}$$

these iterations effectively altering the descent direction in order to improve the conditioning. To choose $\delta\tau$ we may seek a minimum of the function $\mathbf{b}(\mathbf{U})^{p}$ along the descent line before repeating with a new descent direction. The process is continued until a minimum of $\mathcal{I}$ is reached to within a satisfactory tolerance.

It is well known that steepest descent iterations converge progressively more slowly as the limit is approached. Since it is desirable to be able to take as large a pseudo-time step $\delta\tau$ as possible, consistent with reaching convergence, the standard approach is to accelerate the convergence by switching to Newton's method when possible. Newton's iteration is

$$-J^{p}(\mathbf{U}^{p+1} - \mathbf{U}^{p}) = \mathbf{b}(\mathbf{U})^{p} \tag{5.3}$$

where J is the Jacobian matrix

$$J = \frac{\partial \mathbf{b}(\mathbf{U})}{\partial \mathbf{U}}. \tag{5.4}$$

Although the iterations (5.1),(5.2) guarantee the reduction of $\mathcal{I}$ for sufficiently small $\delta\tau$, in general (5.3) does not. Switching to Newton therefore may require a check that $\mathcal{I}$ is decreasing: if that fails we should consider returning to steepest descent (see [13]). (Exceptionally, we have seen that in the case of F independent of u_x or $\underline{\nabla} u$ with $F_{uu} > 0$ the weak form (2.23) of the iteration (5.3) does result in a monotonic $\mathcal{I}$.)

If F is quadratic the matrix equation (4.26) is linear and it is straight forward to find $\mathbf{U}^*$ such that $\mathbf{b}(\mathbf{U}^*) = \mathbf{0}$. However, in all the adaptive cases, as well as (1.17), equation (4.26) is nonlinear.

We now apply these arguments to the problem of generating optimal grids with optimal approximations in practice.

5.1. One step iterations

These are iterations using the theory of Section 3 in which we seek limits U^* and X^* by varying U and X simultaneously. The aim is to solve (3.9),(3.10), i.e.

$$\mathbf{g}(\mathbf{Y}^*) = \mathbf{0}, \tag{5.5}$$

for $\mathbf{Y}^*$. A possible method of solution is to use one of the Galerkin forms

$$A(\mathbf{Y}^p)(\mathbf{Y}^{p+1} - \mathbf{Y}^p) = \mathbf{g}(\mathbf{Y}^p)\delta\tau, \tag{5.6}$$

$$D(\mathbf{Y}^p)(\mathbf{Y}^{p+1} - \mathbf{Y}^p) = \mathbf{g}(\mathbf{Y}^p)\delta\tau, \tag{5.7}$$

where $\delta\tau > 0$, for which $\mathcal{I}$ possesses the monotonic descent property provided that $A(\mathbf{Y})$ is positive definite.

It was noted in Section 3.4 that the matrix $A(\mathbf{Y})$ in equation (5.6) is the MFE mass matrix [9],[7]. Indeed, the iteration (5.6) is an explicit time discretisation of the standard MFE equation [7]

$$A(\mathbf{Y})\frac{d\mathbf{Y}}{dt} = \mathbf{g}(\mathbf{Y}). \tag{5.8}$$

The link is seen clearly from (3.2) which in the piecewise linear case, apart from boundary terms, can be put into the form

$$\delta\mathcal{I} = \sum_{j=1}^{J}\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U} - \frac{d}{dx}\frac{\partial F}{\partial U_x}\right)(\delta^L U_j - U_x\delta^L X_j)\psi_j(x)dx \tag{5.9}$$

which in MFE notation is equivalent to

$$\frac{d\mathcal{I}}{dt} = \sum_{j=1}^{J}\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U} - \frac{d}{dx}\frac{\partial F}{\partial U_x}\right)(\dot{U}_j\ \alpha_j(x) + \dot{X}_j\ \beta_j(x))dx. \tag{5.10}$$

As observed by Miller [9] and confirmed by Jimack [11], in many cases equation (5.5) gives a local optimal $\mathbf{Y}$ (i.e. solution plus grid) for minimisations governed by (1.1). The solutions may be reached by iterations of the form (5.6) or (5.7). However, the MFE-type matrices $A(\mathbf{Y})$ and $D(\mathbf{Y})$ are only positive semi-definite because two rows may be identical if components of $\underline{\nabla}U$ coincide in adjacent elements [7]. In that event $\mathcal{I}$ fails to decrease and a modification is required. In the approach of Miller [9] and Carlson and Miller [10] penalty terms are added in order to prevent $A(\mathbf{Y})$ becoming singular. These authors subsequently apply a Newton method to the nonlinear equation arising from the implicit time discretisation of (5.8) in the form

$$-\left(J^{n+1}\mathbf{Y}\right)^q\left(\left(\mathbf{Y}^{n+1}\right)^{q+1}-\mathbf{Y}^n\right)=\left(A(\mathbf{Y}^{n+1})(\mathbf{Y}^{n+1}-\mathbf{Y}^p)-\Delta t\mathbf{g}(\mathbf{Y}^{n+1})\right)^q \tag{5.11}$$

where $J(\mathbf{Y})$ is here the Jacobian matrix of the right hand side vector. This approach has no obvious monotonicity property, however.

A regularised version of the MFE method with the monotonicity property intact has been used by Jimack [11] to drive $\mathbf{g}(\mathbf{Y})$ to zero via (5.6) with a regualrised mass matrix. Although the path to the limit is altered by regularisation the limit is unaffected. In this way locally optimal solutions and meshes are obtained for a particular example of a self-adjoint problem (cf. (1.14)). The results show the feasibility of deriving optimal grids and solutions in this way although the size of $\delta\tau$ required to avoid tangling of the grid means that the method converges rather slowly.

A feature of solutions in one dimension is the generation of grids which asymptotically equidistribute properties of the solution U in the limit of large numbers of nodes [7]. For example, in the case of Poisson's equation $u_{xx} = f_{xx}$, for which

$$\mathcal{I}(u) = \int \left(\frac{1}{2}u_x^2 + f(x)u\right) dx \tag{5.12}$$

(cf. (1.14)) the steady state grid equidistributes $|f|^{2/3}$ and there is a corresponding result in the case of the convection-diffusion equation [7]. It is also known [8] that the resulting grid in the case of Poisson's equation is optimal in any number of dimensions in the sense that U is to the best L_2 fit with adjustable nodes by piecewise linear functions to the steady state solution in the H^1 semi-norm. A similar result holds in the case of best fits to continuous functions by piecewise linears in the L_2 norm [10]. In the asymptotic limit of large numbers of nodes, therefore, there is an equivalence between best fits and the equidistribution principle. Away from this limit the equivalence is only approximate, although it may be used to generate grids which give a first guess for the iterative algorithms discussed in this section.

5.2. Constrained iterations

We have seen in Section 4 that an optimal solution for U and X may also be sought via separate constrained minimisations of the functional $\mathcal{I}$, each of which has the property that the value of the functional is non-increasing. The approach leads naturally to an alternating two-stage iteration scheme on these pairs of equations, that is we solve (4.17) (or (4.3)) for U with nodal positions held fixed, and then solve (4.18) (or (4.9)) for X_j keeping U on the piecewise linear graph of the current solution. The two stages are then repeated to convergence.

The equations can sometimes be solved outright but where that is not possible each stage may be replaced by a single steepest descent step which preserves the monotonic behaviour of $\mathcal{I}$. Since the resulting problems are local, the stages of the iteration are cheap and a bonus of the steepest descent approach is that $\delta\tau$ can be chosen to control the behaviour of the grid [13],[16].

If in one dimension F is a function of U only with a unique minimum then there exists an ordering property for the nodes of the grid under the two-stage iterations. From (4.8) we deduce that $\frac{\partial F}{\partial U}$ vanishes twice in the interval (X_{j-1}, X_j) so that the function is stationary at two interior points. Since F has a unique minimum it takes the same value at all these interior points. It follows that the jump $[F]_j$ in F considered as a function of x is positive at the rightmost such interior point in the interval (X_{j-1}, X_j) and negative at the leftmost such point in the interval (X_j, X_{j+1}). Hence $[F]_j$ vanishes at least once in the interval between these two points, at $x = x_V$ say. By choosing x_V as the new nodal position we ensure that all such new nodal points belong to disjoint intervals and the grid remains ordered. The new nodal positions may be easily found, for example by bisection.

Unfortunately there is no such ordering property in higher dimensions and the mesh is prone to tangling. Indeed mesh tangling is a fundamental difficulty in grid optimisation of this kind. However, one of the benefits of the local conditions obtained as a result of the constrained approach is that the updates can be organised in such a way as to avoid grid tangling. As explained in [16], by careful choice of $\delta\tau$ in the steepest descent method and taking a "Gauss-Seidel" apprach to the updating of the nodal positions, it is possible to avoid mesh tangling altogether.

5.3. Best fits using direct minimisation

In [5] and [13] the problem of obtaining best fits to continuous functions with adjustable nodes is treated by the methods of Section 5.2 using piecewise linear and piecewise constant approximation both in one dimension and in two dimensions on triangles. In the one-dimensional case both the linear system (4.3) and the jump condition (4.9) may be solved outright. In the piecewise linear case the stages are

$$\int_{X_{k-1}}^{X_k} (U - f(x))\phi_{k,\nu}(x)dx = 0 \qquad (5.13)$$

$\forall k, \nu = 1, 2$, to be solved for U on the fixed grid, and

$$[(U - f(x))^2]_j = 0 \qquad (5.14)$$

$\forall j$ to be solved for X with U confined to lie on the graph of the current approximation (or its extrapolation). The procedure used in [5],[13] is to solve (5.13) for U and then to solve (5.14) for X (picking the solution which moves the node the least distance), always reducing $\mathcal{I}$, and repeating to convergence. The converged solution gives continuity of U almost everywhere with no tangling of the grid. Convergence of the iteration may be accelerated by Newton's method as long as the monotonicity property of $\mathcal{I}$ is preserved..

The approach reveals relationships between successive element sizes for optimal piecewise linear approximation (superseding the asymptotic equidistribution properties of Section 5.1) given by

$$\frac{1}{\Delta X_{j-h}} \int_{X_{j-1}}^{X_j} (-2\phi_{j-\frac{1}{2},L} + 4\phi_{j-\frac{1}{2},R}) (f(x) - f(X_j))\, dx$$

$$= \frac{\pm 1}{\Delta X_{j+h}} \int_{X_j}^{X_{j+1}} (4\phi_{j+\frac{1}{2},L} - 2\phi_{j+\frac{1}{2},R}) (f(x) - f(X_j))\, dx, \qquad (5.15)$$

$\forall j$, where $\Delta X_{j-h} = X_j - X_{j-1}$.The corresponding relationship for piecewise constant approximation is

$$\frac{1}{\Delta X_{j-h}} \int_{X_{j-1}}^{X_j} (f(x) - f(X_j))\, dx = \frac{\pm 1}{\Delta X_{j+h}} \int_{X_j}^{X_{j+1}} (f(x) - f(X_j))\, dx. \qquad (5.16)$$

In the two-dimensional case it is possible to solve outright only the linear system (4.17) for U, in the form

$$\int_{\Delta_k} (U - f(\underline{x}))\phi_{k,\nu}(\underline{x}) d\Omega = 0 \qquad (5.17)$$

$\forall k, \nu$. Equation (4.19) for X, namely,

$$\int_{\partial \Delta_{kj}} (U - f(\underline{x}))^2 \psi_j(\underline{x}) \underline{n}_k ds = 0 \qquad (5.18)$$

cannot be solved outright and in [13] is replaced by the steepest descent iteration (4.28) with a limiter on $\delta\tau$ to avoid mesh tangling.

There are special problems with convergence in two dimensions in that the error cannot easily be driven down to machine accuracy because of the inflexibility of the grid topology. However, in [13] by using an edge-swapping routine [14] and a technique for small element removal this is achieved in a number of examples.

5.4. Application to the shallow water equations

In [4],[15] variational methods have been applied to the problem of approximating both the depth and the grid in channel flows governed by the steady shallow water equations. The variational principles are used in several different ways. The functional

$$\frac{1}{2g}\int_a^b B(x)\left(E-\frac{1}{2}u_x^2\right)^2 dx, \tag{5.19}$$

where u is the velocity potential, is used to generate approximations and adaptive grids by a one step iteration via Newton's method. Similarly, (1.17), where U is now the depth, is used in [4],[15] to obtain approximations to depths with a single jump, using piecewise linear approximation with a single discontinuity at one moving node, as well as piecewise linear local approximation with all nodes varying.

6. Time Discretisation

All the problems considered so far have been independent of time. As is well known, time dependent variational principles can invoke difficulties with the direct application of boundary conditions at future time boundaries. However, if the time derivative is already discretised the generalisation of the variational principles in Section 2 is straightforward.

For example, the extended Euler-Lagrange equation

$$\frac{\partial u}{\partial t} = -\frac{\partial F}{\partial u} + \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} u} \tag{6.1}$$

discretised in time using implicit Euler finite differences is

$$\frac{u^{n+1}-u^n}{\Delta t} = \left(-\frac{\partial F}{\partial u} + \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} U}\right)^{n+1} . \tag{6.2}$$

Writing $u^{n+1} = u(x)$, we may extend the function $F(\underline{x}, u, \underline{\nabla} u)$ in (1.1) to

$$G(\underline{x}, u, \underline{\nabla} u) = \frac{1}{2}\frac{(u-u^n)^2}{\Delta t} + F(\underline{x}, u, \underline{\nabla} u), \tag{6.3}$$

defining also the functional $\mathcal{J}(u)$ as

$$\mathcal{J}(u) = \int_\Omega G(\underline{x}, u, \underline{\nabla} u) d\Omega = \int_\Omega F(\underline{x}, u, \underline{\nabla} u) d\Omega + \frac{1}{2\Delta t}\int_\Omega (u-u^n)^2 d\Omega. \tag{6.4}$$

If $\mathcal{I}$ is a convex functional of u then so is $\mathcal{J}$.

The first variation of $\mathcal{J}(u)$ is then

$$\begin{aligned}\delta\mathcal{J} &= \int_\Omega \left(\frac{\partial G}{\partial u}\delta u + \frac{\partial G}{\partial \underline{\nabla} u}.\underline{\nabla}\delta u\right) d\Omega \\ &= \int_\Omega \left\{\frac{\partial G}{\partial u} - \underline{\nabla}.\left(\frac{\partial G}{\partial \underline{\nabla} u}\right)\right\}\delta u d\Omega, \end{aligned} \tag{6.5}$$

vanishing for all δu when $u = u^*$ satisfies

$$\frac{\partial G}{\partial u} - \underline{\nabla}.\left(\frac{\partial G}{\partial \underline{\nabla} u}\right) = 0 \tag{6.6}$$

which, using (6.3), is readily seen to be equivalent to (6.2).

If the variations in u are chosen to be

$$\begin{aligned}\delta u &= \left(-\frac{\partial G}{\partial u} + \underline{\nabla}.\left(\frac{\partial G}{\partial \underline{\nabla} u}\right)\right)\delta\tau \\ &= \left(\frac{-(u-u^n)}{\Delta t} - \frac{\partial F}{\partial u} + \underline{\nabla}.\left(\frac{\partial F}{\partial \underline{\nabla} u}\right)\right)\delta\tau \end{aligned} \tag{6.7}$$

where $\delta\tau > 0$ then

$$\delta\mathcal{J} = -\int_\Omega \delta u^2 d\Omega \delta\tau \leq 0, \tag{6.8}$$

being zero only if $\delta u = 0$, i.e. when $u = u^*$ satisfies

$$\frac{(u^* - u^n)}{\Delta t} = -\frac{\partial F}{\partial u} + \underline{\nabla}.\left(\frac{\partial F}{\partial \underline{\nabla} u}\right) \tag{6.9}$$

where u^* is the required solution u^{n+1} of (6.2). From (6.3), if $\mathcal{I}$ is bounded below then so is $\mathcal{J}$. Hence $\mathcal{J}$ tends to a limit at which $\delta u = 0$ in which case $u = u^*$ satisfying (6.9).

For example, if $F(\underline{x}, u, \underline{\nabla} u)$ is given by (1.11) then $G(\underline{x}, u, \underline{\nabla} u)$ is

$$G(\underline{x}, u, \underline{\nabla} u) = \frac{1}{2\Delta t}(u - u^n)^2 + u^2 + (\underline{\nabla} u)^2. \tag{6.10}$$

The first variation of $\mathcal{J}$ vanishes when

$$\frac{u - u^n}{\Delta t} = -u + \underline{\nabla}^2 u. \tag{6.11}$$

and the variations δu which make $\mathcal{J}$ non-increasing are

$$\delta u = \left(-\frac{(u-u^n)}{\Delta t} - u + \underline{\nabla}^2 u\right)\delta\tau. \tag{6.12}$$

The approach is easily generalised to Crank-Nicolson and even explicit time stepping, in which case the problem reduces to that of best fits (see Section 5.3).

This procedure gives no information about the way in which to discretise the differential equation in time but simply concerns the convergence of the solution of the implicit equation resulting from the implicit time-stepping.

The independent nature of the extension allows all previous cases to be incorporated, including both fixed and adapting finite elements. In the finite-dimensional case $\mathcal{J}$ becomes

$$\mathcal{J} = \int_\Omega \left\{\frac{1}{2\Delta t}(U - U^n)^2 + F(\underline{x}, U, \underline{\nabla} U)\right\} d\Omega. \tag{6.13}$$

For example, in the fixed grid linear finite element case the first variation vanishes when

$$\int_\Omega \left\{ \left(\frac{(U-U^n)}{\Delta t} + \frac{\partial F}{\partial U} \right) \psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U} . \underline{\nabla} \psi_j(\underline{x}) \right\} d\Omega = 0 \tag{6.14}$$

which is the appropriate weak form of (6.2). Moreover, $\mathcal{J}$ is a non-increasing function when δU is chosen to satisfy the Galerkin form

$$\int \delta U \psi_j(\underline{x}) d\Omega = \int_\Omega \left\{ - \left(\frac{U-U^n}{\Delta t} + \frac{\partial F}{\partial U} \right) \psi_j(\underline{x}) - \frac{\partial F}{\partial \underline{\nabla} U} . \underline{\nabla} \psi_j(\underline{x}) \right\} d\Omega \delta\tau \tag{6.15}$$

(cf. (2.12)). Similarly, in the adaptive case the first variation of $\mathcal{J}$ vanishes when U,X satisfy (6.14) and, from (3.32),

$$\int_{\Delta_{kj}} \left\{ F \underline{\nabla} \psi_j(\underline{x}) + \left(\frac{\partial F}{\partial \underline{X}} \psi_j(\underline{x}) - \left(\frac{\partial F}{\partial \underline{\nabla} U} . \underline{\nabla} \psi_j(\underline{x}) \right) \underline{\nabla} U \right) \right\} d\Omega = 0. \tag{6.16}$$

If F is independent of $\underline{\nabla} u$ (6.14) reduces to

$$\int_\Omega \left(\frac{(U-U^n)}{\Delta t} + \frac{\partial F}{\partial U} \right) \psi_j(\underline{x}) d\Omega = 0 \tag{6.17}$$

and if $F_{uu} > 0$ we may use the modified Galerkin steepest descent iteration

$$\int F_{UU} \delta U \psi_j(\underline{x}) d\Omega = \int_\Omega \left(- \frac{U-U^n}{\Delta t} - \frac{\partial F}{\partial U} \right) \psi_j(\underline{x}) d\Omega \delta\tau$$

instead of (6.15), which preserves the monotonicity property and turns the iteration into a weak form of Newton's method.

The MFE and GWMFE methods of Miller [9],[7] present another way of treating time dependent problems in which the nodes are moved automatically by seeking finite-dimensional approximations satisfying

$$\min_{u_t} \int (u_t - L(u))^2 d\Omega. \tag{6.18}$$

A finite difference method such as implicit Euler is then used to perform the time integration. The method has been extensively developed to include penalty functions (to prevent node tangling), implicit time stepping and fast matrix solvers.

The above considerations prompt another look at the Moving Finite Element method. By discretising the time derivative in (6.18) before carrying out the minimisation it is possible for iterations of the form (6.2) to maintain the monotonicity preserving property of $\mathcal{J}$. Thus, for the PDE (6.1), the minimisation (6.18) becomes

$$\min_U \int \left\{ \left(\frac{U-U^n}{\Delta t} \right) + \frac{\partial F}{\partial U} - \underline{\nabla} . \frac{\partial F}{\partial \underline{\nabla} U} \right\}^2 d\Omega \tag{6.19}$$

and the weak forms (6.14) and (6.16) may be solved by a Galerkin steepest descent iteration whose algebraic form is

$$A(\mathbf{Y})(\mathbf{Y}^{p+1} - \mathbf{Y}^{p}) = \mathbf{g}_e(\mathbf{Y}), \tag{6.20}$$

where

$$\mathbf{g}_e(\mathbf{Y}) = \mathbf{g}(\mathbf{Y}) - \mathbf{h}(\mathbf{Y}), \tag{6.21}$$

$\mathbf{h}(\mathbf{Y})$ being defined by

$$\mathbf{h}(\mathbf{Y}) = \left\{ \int_\Omega (U - U^n)\psi(\underline{x})d\Omega, \quad \int_\Omega (U - U^n)\underline{\nabla} U \psi(\underline{x})d\Omega \right\}^T . \tag{6.22}$$

This steepest descent method (with regularised $A(\mathbf{Y})$) ensures that the functional in (6.19) is non-increasing except at a stationary point. $A(\mathbf{Y})$ may be replaced by $D(\mathbf{Y})$ in (6.20) without affecting the monotonicity property of $\mathcal{I}$.

For algebraic problems with F independent of $\underline{\nabla} u$ and $F_{uu} > 0$, we may weight the elements of the matrix $D(\mathbf{Y})$ by F_{UU} to obtain a higher order steepest descent method.

As an example consider the PDE

$$u_t = \underline{\nabla}^2 u \tag{6.23}$$

discretised in time by the implicit Euler method as

$$\frac{U^{n+1} - U^n}{\Delta t} = \underline{\nabla}^2 U^{n+1}. \tag{6.24}$$

The corresponding variational principle is

$$\min_U \int_\Omega \left\{ \frac{(U - U^n)^2}{2\Delta t} + (\underline{\nabla} U)^2 \right\} d\Omega \tag{6.25}$$

where $U = U^{n+1}$, and the weak forms (6.14) and (6.16) become

$$\int_\Omega \left\{ \frac{(U - U^n)}{\Delta t}\psi_j(\underline{x}) + \underline{\nabla} U . \underline{\nabla} \psi_j(\underline{x}) \right\} d\Omega = 0 \tag{6.26}$$

and

$$\int_{\Delta_{kj}} \left\{ \left(\frac{(U - U^n)^2}{2\Delta t} + (\underline{\nabla} U)^2 \right) \underline{\nabla} \psi_j(\underline{x}) - (\underline{\nabla} U . \underline{\nabla} \psi_j(\underline{x}))\, \underline{\nabla} U \right\} d\Omega = 0 \tag{6.27}$$

which taken together constuitute $\mathbf{g}_e(\mathbf{Y}) = \mathbf{0}$. Then the iteration (6.20) may be used to generate the solution U at time $n + 1$.

7. Several Dependent Variables

The investigation carried out so far in this report has concerned only scalar problems. Most problems contain several dependent variables, however, so we shall here outline the theory for the case of two dependent variables. If $u(x), v(x)$ are functions twice differentiable in the space variable x and $F(x, u, v, \underline{\nabla} u, \underline{\nabla} v)$ is a once differentiable function of its arguments, the functional of two variables corresponding to (1.1) is

$$\mathcal{I}(u,v) = \int_\Omega F(\underline{x}, u, v, \underline{\nabla} u, \underline{\nabla} v) d\Omega \tag{7.1}$$

and the stationary functions $u = u^*, v = v^*$ satisfy the equations

$$\frac{\partial F}{\partial u} - \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} u} = 0, \qquad \frac{\partial F}{\partial v} - \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} v} = 0. \tag{7.2}$$

If the x variable also participates in the variations, giving simultaneous variations $\delta u, \delta v$ and $\delta \underline{x}$, the chain rule gives

$$\delta^L u = \delta u + \underline{\nabla} u.\delta^L \underline{x}, \qquad \Delta v = \delta v + \underline{\nabla} v.\delta^L \underline{x} \tag{7.3}$$

and the first variation of $\mathcal{I}$ is

$$\int_\Omega \left\{ \left(\frac{\partial F}{\partial u} - \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} u} \right) \left(\delta^L u - \underline{\nabla} u.\delta^L \underline{x} \right) + \left(\frac{\partial F}{\partial v} - \underline{\nabla}.\frac{\partial F}{\partial \underline{\nabla} v} \right) \left(\delta^L v - \underline{\nabla} v.\delta^L \underline{x} \right) \right\} d\Omega. \tag{7.4}$$

Expanding each of the functions $u \sim U, v \sim V$ and $\underline{x} \sim \underline{X}$ in terms of the same piecewise linear basis functions $\psi_j(\underline{x})$ as

$$U = \sum_{j=1}^{J} U_j \psi_j(\underline{x}), \qquad V = \sum_{j=1}^{J} V_j \psi_j(\underline{x}), \qquad \underline{X} = \sum_{j=1}^{J} \underline{X}_j \psi_j(\underline{x}) \tag{7.5}$$

and with $\mathcal{I}(u)$ is defined as in (7.1) with $u, v, \underline{x}$ replaced by $U, V, \underline{X}$, it can be shown that $\delta\mathcal{I}$ vanishes for all $\delta^L U_j, \delta^L V_j$ and $\delta^L \underline{X}_j$ if $U = U^*, V = V^*, \underline{X} = \underline{X}^*$ satisfying the weak forms

$$\int_\Omega \left(\frac{\partial F}{\partial U} \psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U} . \underline{\nabla} \psi_j(\underline{x}) \right) d\Omega = -a(U, \underline{X}) = 0 \tag{7.6}$$

$$\int_\Omega \left(\frac{\partial F}{\partial V} \psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} V} . \underline{\nabla} \psi_j(\underline{x}) \right) d\Omega = -b(U, \underline{X}) = 0 \tag{7.7}$$

and

$$\int_\Omega \left\{ - \left(\underline{\nabla}_{\underline{X}} F \right) \psi_j(\underline{x}) - F \underline{\nabla} \psi_j(\underline{x}) \right.$$

$$+\left(\frac{\partial F}{\partial \underline{\nabla} U}.\underline{\nabla}\psi_j(\underline{x})\right)\underline{\nabla} U+\left(\frac{\partial F}{\partial \underline{\nabla} V}.\underline{\nabla}\psi_j(\underline{x})\right)\underline{\nabla} V\Bigg\} d\Omega=-c(U,\underline{X})=0 \quad (7.8)$$

say, $\forall j$.

Also, under the choice of variations $\delta U, \delta V, \delta \underline{X}$ given by the Galerkin forms

$$\int_\Omega \delta U \psi_i(\underline{x}) d\Omega = a(U,\underline{X})\delta\tau \quad (7.9)$$

$$\int_\Omega \delta V \psi_i(\underline{x}) d\Omega = b(U,\underline{X})\delta\tau \quad (7.10)$$

$$\int_\Omega \{(-\underline{\nabla} U)\delta U + (-\underline{\nabla} V)\delta V\}\, \psi_i(\underline{x}) d\Omega = c(U,\underline{X})\delta\tau \quad (7.11)$$

$\forall i$, where

$$\delta U = \delta^L U - \underline{\nabla} U.\delta^L \underline{X}, \qquad \delta V = \delta^L V - \underline{\nabla} V.\delta^L \underline{X}$$

as before, it follows that

$$\delta \mathcal{I} = -\sum_{i,j}\int_\Omega \left(\delta^L U_j - \underline{\nabla} U.\delta^L \underline{X}_j\right)\psi_i(\underline{x})\psi_j(\underline{x})\left(\delta^L U_j - \underline{\nabla} U.\delta^L \underline{X}_j\right) d\Omega\Delta\tau^{-1}$$

$$-\sum_{i,j}\int_\Omega \left(\delta^L V_j - \underline{\nabla} V.\delta^L \underline{X}_j\right)\psi_i(\underline{x})\psi_j(\underline{x})\left(\delta^L V_j - \underline{\nabla} V.\delta^L \underline{X}_j\right) d\Omega\Delta\tau^{-1} \le 0 \quad (7.12)$$

which gives the same monotonicity property as before provided that the quadratic forms arising in (7.12) are positive definite. The solutions of equations (7.6), (7.7) and (7.8) correspond to the simultaneous solutions approach of Section 3.4.

7.1. Constrained approximation

In one dimension, constraining the variations as in section 4, the stationary values U_j, V_j satisfy the weak forms

$$\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial U}\psi_j(x)+\frac{\partial F}{\partial U_x}\underline{\nabla}\psi_j'(x)\right)dx=0, \quad (7.13)$$

$$\int_{X_{j-1}}^{X_{j+1}}\left(\frac{\partial F}{\partial V}\psi_j(x)+\frac{\partial F}{\partial V_x}\underline{\nabla}\psi_j'(x)\right)dx=0 \quad (7.14)$$

$(j=1,2,...,J)$ with the X_j fixed, while the X_j satisfy

$$[F]_j = 0 \quad (7.15)$$

$\forall j$, where in the latter case variations in U or V are constrained by $\delta U_j = U_x \delta X_j, \delta V_j = V_x \delta X_j$. In the case where F is independent of $\underline{\nabla} U$ and $\underline{\nabla} V$, discontinuous linear approximation for U and V yields the local forms

$$\int_{X_{k-1}}^{X_k}\frac{\partial F}{\partial V}\phi_{k\nu}(x)dx=0, \qquad \int_{X_{k-1}}^{X_k}\frac{\partial F}{\partial V}\phi_{k\nu}(x)dx=0 \quad (7.16)$$

for U, V and (7.15) again for X.

In higher dimensions the conditions are

$$\int_{\Delta_{kj}} \left(\frac{\partial F}{\partial U} \psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U} . \underline{\nabla} \psi_j(\underline{x}) \right) d\Omega = 0, \tag{7.17}$$

$$\int_{\Delta_{kj}} \left(\frac{\partial F}{\partial V} \psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} V} . \underline{\nabla} \psi_j(\underline{x}) \right) d\Omega = 0 \tag{7.18}$$

$\forall j$ for U, V, while under the frozen solution constraints $\delta^L U_j = \underline{\nabla} U . \delta^L X_j$, $\delta^L V_j = \underline{\nabla} V . \delta^L X_j$, we have

$$\begin{aligned} & \int_{\Delta_{kj}} \left\{ \left(\underline{\nabla} . \frac{\partial F}{\partial \underline{\nabla} U} \right) \psi_j(\underline{x}) \underline{\nabla} U + \left(\underline{\nabla} . \frac{\partial F}{\partial \underline{\nabla} V} \right) \psi_j(\underline{x}) \underline{\nabla} V \right\} d\Omega \\ & + \int_{\partial \Delta_{kj}} \left\{ F \underline{n}_k - \left(\frac{\partial F}{\partial \underline{\nabla} U} . \underline{n}_k \right) \underline{\nabla} U - \left(\frac{\partial F}{\partial \underline{\nabla} V} . \underline{n}_k \right) \underline{\nabla} V \right\} \psi_j(\underline{x}) ds = 0 \end{aligned} \tag{7.19}$$

$\forall j$ for X. The U, V variations in (4.18) move on the graphs (in two dimensions on the discontinuous planes) of U, V as $\underline{X}$ is varied.

If F is independent of $\underline{\nabla} U, \underline{\nabla} V$, and if U, V are approximated by piecewise discontinuous functions, then the U_j and V_j are given by the equations

$$\int_{\Omega} \frac{\partial F}{\partial U} \phi_{k\nu}(\underline{x}) d\Omega = 0, \qquad \int_{\Omega} \frac{\partial F}{\partial V} \phi_{k\nu}(\underline{x}) d\Omega = 0 \tag{7.20}$$

$\forall k, \nu$, and the X_j are given by

$$\int_{\partial \Delta_{kj}} F \underline{n}_k \psi_j(\underline{x}) ds = 0 \tag{7.21}$$

under the constraints that U, V remain on their graphs as X_j is varied.

An example from Shallow Water theory is as follows [15]. The equations of incompressible irrotational quasi one-dimensional flow in a channel may be generated by the principle

$$\min_{Q,d,\phi} \int B(x) \left(\frac{Q^2}{2d} - \frac{1}{2} g d^2 + E(x) d - \phi' Q \right) dx \tag{7.22}$$

where Q is the mass flow, d is the depth and ϕ is the velocity potential. The weak forms of the variational conditions are

$$\int_{X_{j-1}}^{X_{j+1}} B(x) \left(\frac{Q}{d} - \phi' \right) \psi_j(x) dx = 0, \tag{7.23}$$

$$\int_{X_{j-1}}^{X_{j+1}} B(x) \left(-\frac{Q^2}{d^2} - g d + E(x) \phi' \right) \psi_j(x) dx = 0. \tag{7.24}$$

and

$$\int_{X_{j-1}}^{X_{j+1}} B(x)Q'\psi_j(x)dx = 0 \tag{7.25}$$

to be solved for Q, d and ϕ, while that for the grid X is either

$$\int_{X_{j-1}}^{X_{j+1}} (B(x)\psi_j(x))' \left(\frac{Q^2}{2d} - \frac{1}{2}gd^2 + E(x)d - \phi' Q\right) dx = 0 \tag{7.26}$$

in the case of the simultaneous solution approach of Section 3, or

$$\left[\frac{Q^2}{2d} - \frac{1}{2}gd^2 + E(x)d - \phi' Q\right]_j = 0 \tag{7.27}$$

when the constraints are applied in a sequential way, as in Section 4.

8. Conclusions

In this report we have reviewed recent advances in numerical variational techniques which generate optimal grids as well as optimal solutions. A unified approach to the problem of finding generalised Euler-Lagrange equations has been described together with iterative schemes which have the property that the functionals behave monotonically. One of the equation sets obtained is related to the MFE method, to be solved for U and X simultaneously. The other, constrained, set leads to local problems solved for U and X sequentially, which can be implemented in such a way as to avoiding mesh tangling. Both methods are viable in any number of dimensions,

In Section 3 the generalised variational conditions obtained in one dimension were given by (3.7),(3.8) and in higher dimensions by (3.30),(3.31). In Section 4 the conditions were (4.3),(4.4) in one dimension and (4.17),(4.18) in higher dimensions. Since the first of each pair of equations coincide, there is an equivalence between the second equation of each pair. That is to say

$$\int_{X_{j-1}}^{X_{j+1}} \left(\frac{\partial F}{\partial X}\psi_j(x) + \left(F - U_x\frac{\partial F}{\partial U_x}\right)\psi_j'(x)\right) dx = 0 \tag{8.1}$$

is equivalent to

$$[F]_j = 0 \tag{8.2}$$

and

$$\int_{\Delta_{kj}} \left\{-\left(\frac{\partial F}{\partial U}\psi_j(\underline{x}) + \frac{\partial F}{\partial \underline{\nabla} U}\cdot\underline{\nabla}\psi_j(\underline{x})\right)\underline{\nabla}U + \underline{\nabla}\left(F\psi_j(\underline{x})\right)\right\} d\Omega = 0 \tag{8.3}$$

is equivalent to

$$\int_{\Delta_{kj}} \left(\underline{\nabla}\cdot\frac{\partial F}{\partial \underline{\nabla} U}\right)\psi_j(\underline{x})\underline{\nabla}U d\Omega + \int_{\partial\Delta_{kj}} \left\{F\underline{n}_k - \left(\frac{\partial F}{\partial \underline{\nabla} U}\cdot\underline{n}_k\right)\underline{\nabla}U\right\}\psi_j(\underline{x})ds = 0. \tag{8.4}$$

To make this point clearer consider the manner in which these equations are obtained. In Section 4 there is an implicit decoupling between the pairs of equations due to the application of the constraints. In Section 3 the pairs of equations are sufficiently similar to formally be solved simultaneously although this is not essential. They could just as feasibly be solved sequentially for U and X, as in Section 4 where there is an implicit decoupling between the pairs of equations due to the application of the constraints. It is in this sense that there is an equivalence between the second of each pair of equations above.

Thus, the similarity depends entirely on the strategy of first constraining X to be fixed, giving rise to the first equation of each pair, and then constraining U to be either constant or to lie on its graph as X varies. In the first case we are led to (3.8) or (3.31) while in the second to (4.4) or (4.18). In each case the latter form is easier to work with because of its local character. In particular, as we have seen in Section 4, the local forms provide the flexibility to control grid behaviour and to avoid mesh tangling.

The approach may be extended to time-dependent problems discretised in a finite difference manner and to problems involving several dependent variables. Implementation will be carried out in a later report.

9. References

[1] **Atkinson, K.E.**, *Numerical Analysis*, Wiley (1989).

[2] **Fortin, M. and Glowinski, R.**, *Methodes de Lagrangien Augmente*, Dunod, Paris (1994).

[3] **Seliger, R.L. and Whitham, G.B.**, Variational Principles in Continuum Mechanics. Proc. Roy. Soc. A, 305, 1-25 (1968).

[4] **Wakelin, S.L., Baines, M.J. and Porter, D**. The Use of Variational Principles in Determining Approximations to Continuous and Discontinuous Shallow Water Flows. Numerical Analysis Report 13/95, Department of Mathematics, University of Reading.

[5] **Baines, M.J.**, 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).

[6] **Wathen, A.J.**, Mesh-independent spectra in the moving finite element equations. SIAM J.Num.Anal., 23, 797-814 (1986).

[7] **Baines, M.J.**, *Moving Finite Elements.* Clarendon Press, Oxford (1994).

[8] **Jimack, P.K.**, A Best Approximation Property of the Moving Finite Element Method. School of Computer Studies Research Report 93.35, University of Leeds. (To appear in SIAM J. Numer. An.)

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

[10] **Carlson, N.N. and Miller, K.** Design and Application of a Gradient-Weighted Moving Finite Element Code, Parts 1 and 2, technical reports 236 and 237, Department of Mathematics, Purdue University (to appear in SIAM J Num An).

[11] **Jimack, P.K.**, On the Stability of the Moving Finite Element Method for a Class of Parabolic Partial Differential Equations. Proceedings of Paris Conference (to appear).

[12] **Baines, M.J.**, On the relationship between Moving Finite Elements and Best L_2 Fits to Continuous Functions with Adjustable Nodes. J.Numer.PDEs, 10, 191-203 (1994).

[13] **Tourigny, Y. and Baines, M.J.**, Analysis of an Algorithm for Generating Locally Optimal Meshes for L_2 Approximation by Discontinuous Piecewise Polynomials. Math Comp (to appear).

[14] **Lawson, C. L.**, Software for C^1 Interpolation, in *Mathematical Software III*, J.R.Rice (ed.), Academic Press (1977), pp 161-194.

[15] **Wakelin, S.L.**, Variational Principles and the Finite Element Method for Channel Flows. PhD thesis. University of Reading, UK (1993).

[16] **Tourigny, Y. and Hulsemann, F.**, A New Moving Mesh Algorithm for the Finite Element Solution of Variational Problems (submitted to Math.Comp.)

[17] **Delfour, M., Payre, G. and Zolesio, J.P.**, An Optimal Triangulation for Second Order Elliptic Problems. Comput Meths in Appl Mech Engrg, 50, 231-261 (1985).

[18] **Johnson, C & Hansbo, P.**, Adaptive Finite Element Methods in Computational Mechanics, Comput Meths in App Mech Engrg, 101, 143-181 (1992).

[19] **Huang, W., Ren, Y. and Russell, R.D.**, Moving Mesh Methods based on the Equidistribution Principle, SIAM J Numer An, 31, 709-721 (1994).

[20] **Knupp, P. & Steinberg, S.**, *Fundamentals of Grid Generation*, CRC Press (1994).