INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 35, 623–636 (1992)

A NOTE ON THE GRID MOVEMENT INDUCED BY MFE*

P. A. ZEGELING AND J. G. BLOM

CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

SUMMARY

Moving-grid methods in one space dimension have become popular for solving several kinds of parabolic and hyperbolic partial differential equations (PDEs) involving fine scale structures such as steep moving fronts and emerging steep layers, pulses, shocks, etc.. In two space dimensions, however, application of moving-grid methods is less trivial than in 1D. For some methods, e.g. those based on equidistributing principles, it is not even clear how to extend the underlying grid selection procedure to 2D. The moving-finite-element (MFE) method does not suffer from this drawback; its mathematical extension to 2D is trivial. However, because of the intrinsic coupling between the discretization of the PDE and the grid selection, the application of MFE, as for any other method, is not without difficulties. In this paper we describe the node movement induced by MFE for various PDEs and we indicate some problems concerning the grid structure that can result from this movement

1. INTRODUCTION

During the last decade, moving-grid methods in one space dimension have become popular for solving several kinds of parabolic and hyperbolic partial differential equations (PDEs) involving fine scale structures such as steep moving fronts, emerging steep layers, pulses, shocks, etc.. Moving-grid methods use non-uniform space grids, and move the grid continuously in the space–time domain while the discretization of the PDE and the grid selection procedure are intrinsically coupled. Examples are provided by the moving-finite-element (MFE) method of Miller,[11,13] and by the moving-finite-difference (MFD) method discussed in Verwer *et al.*[18] (see also references therein). The latter is, in contrast to MFE, restricted to problems in one space dimension.

In two space dimensions, however, application of moving-grid methods is less trivial than in 1D. For instance, there are many possibilities to treat the one-dimensional boundary and to discretize the spatial domain, each having their own difficulties for specific PDEs. Therefore, 2D moving-grid methods have mostly been applied only to special types of PDEs. The MFE method,[7,9,12] considering its general approach, allows in principle a large class of problems to be dealt with. However, because of the intrinsic coupling between the discretization of the PDE and the grid selection, the application of MFE, as for any other moving-grid method, is not without difficulties. The main difficulty we are referring to is the threat of grid distortion. Grid distortion

* This work has been carried out in connection with a joint CWI/Shell project on 'Adaptive Grids'. For this project Paul Zegeling has received support from the 'Netherlands Foundation for the Technical Sciences' (STW) (Contract No. CWI 59.0922)

0029–5981/92/160623–14$12.00

Received 5 November 1990
Revised 20 November 1991

can occur in many different ways due to the quite complex solution behaviour of 2D-evolution problems. For example, sharp layer regions could develop and propagate through the domain, or rotating pulses could emerge and die out again. The purpose of this note is to describe the node movement induced by MFE for various PDEs and to indicate some problems concerning the grid structure that can result from this movement.

A standard way of describing moving-grid methods is the introduction of a transformation of the three dependent variables x, y (space) and t (time) into new variables ξ, η and τ (usually one chooses $t = \tau$). The effect of the transformation may be to stretch the co-ordinates in a steep region, so that the transformed derivatives are small compared with the old ones. Of course, many of the difficulties that the spatial discretization yields are now shifted to the problem of how to define the mapping. After having applied the transformation, we obtain the so-called Lagrangian form of the PDE. Within this new formulation the time-derivatives of the spatial variables x and y appear. It is clear that, before using a numerical scheme to discretize the model, one has to define extra equations for these quantities. There are various approaches to take care of this. First, one can use a 2D extension of the equidistribution principle, see, e.g. Brackbill and Saltzman[6] or Dwyer.[10] This idea is either very difficult to work out and to implement, owing to the complicated structure of the formulas, or, in a simpler form, it can be applied only to a small class of models. Second, one can use the method of characteristics. This method can, however, be applied only to certain scalar hyperbolic equations or systems having a common convective velocity. For general systems in 2D the use of this method is problematical if possible at all. We would like to focus our attention on the MFE- method, which defines the transformation in terms of a residual minimization. For scalar hyperbolic equations MFE is related to the method of characteristics (see, e.g. Baines[2, 3]). This link with the characteristics of the PDE is very useful in one dimension. For in that case all 'disturbances', i.e. shocks, pulses, etc., can merely follow the characteristics. So, once the user has located the grid points at the right positions, the characteristics do the rest. This has the advantage that MFE needs very few points to follow such solutions. In two dimensions it may work properly as well, for the same reasons (see e.g. Miller[12] or Carlson and Miller[7]). However, in some situations one has to be very careful in applying this method. We will illustrate this with some examples. For parabolic equations the node movement induced by MFE is less understood. For 1D scalar equations one can derive asymptotic relations for the node movement and for the node distribution, indicating that for parabolic equations MFE strives after an equidistribution of second and first order derivatives. An example gives some indication that these results possibly also hold in 2D.

The paper is divided into four sections. In Section 2 we briefly describe MFE in two space dimensions, its relation to the method of characteristics for hyperbolic equations and results on the grid movement that can be derived for the parabolic case. Section 3 contains two examples of hyperbolic PDEs with a typical solution behaviour. For these two examples it is shown that MFE yields a severely distorted grid, although the computed solution remains accurate. However, this distortion can lead to a breakdown of the numerical time-stepping procedure. Section 3 also contains an example of a parabolic equation for which MFE strives after a transformation equidistributing second order derivatives. Finally, Section 4 is devoted to some conclusions.

2. THE MOVEMENT OF THE NODES IN MFE

Let us consider the scalar PDE

$$\frac{\partial u}{\partial t} = L(u), \quad (x, y) \in \Omega, \quad t > 0 \tag{1}$$

with initial and boundary conditions

$$u|_{t=0} = u^0(x, y), \quad (x, y) \in \Omega$$
$$B(u, \nabla u)|_{\partial\Omega} = g(t), \quad t > 0$$

where u^0 and g are given functions, and L represents a differential operator involving only spatial derivatives up to second order. In general, the solution $u(x, y, t)$ of (1) may have a very complex behaviour. Even for a restricted situation (a scalar linear PDE with simple boundary conditions), one can have severely varying u-values in space (x, y) and time t. Some examples in this context are steep moving fronts and emerging and rotating pulses.

A common approach in handling these phenomena is to introduce a transformation which maps the variables x, y and t into new variables ξ, η and τ. Such a transformation can be defined as, e.g.

$$\begin{aligned} x &= x(\xi, \eta, \tau) \\ y &= y(\xi, \eta, \tau) \\ t &= \tau \\ u(x, y, t) &= v(\xi, \eta, \tau) \end{aligned} \tag{2}$$

Applied to the left-hand side of equation (1) this gives

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial \tau} - u_x \frac{\partial x}{\partial \tau} - u_y \frac{\partial y}{\partial \tau} \tag{3}$$

and additionally equations for x and y must be defined. The effect of the transformation may be to stretch the co-ordinates in a steep region in space so that, for example, u_ξ and u_η are small in contrast with u_x and u_y. This type of transformation is strived after by methods which equidistribute first or higher order derivatives of the solution. Another effect of the transformation may be to decrease the $\partial v/\partial \tau$ as is done by the method of characteristics and by the finite difference method of Petzold,[15] in 1D. Of course, when using a transformation, most difficulties are shifted to the problem of how to define and carry out the mapping. The moving-finite-element (MFE) method can, in some cases, also be shown to underly a transformation of variables.[3] Below we will discuss this method and in particular the node movement induced.

2.1. *Description of MFE*

MFE restricts v, x and y to U, X and Y from a finite-dimensional subspace. The MFE-approximations are piecewise linear on a, in our case, hexagonally connected, triangularization of Ω

$$\begin{aligned} v \approx U &= \sum_{j \in J} U_j(\tau)\alpha_j(\xi, \eta) \\ x \approx X &= \sum_{j \in J} X_j(\tau)\alpha_j(\xi, \eta) \\ y \approx Y &= \sum_{j \in J} Y_j(\tau)\alpha_j(\xi, \eta) \end{aligned} \tag{4}$$

where J is the set of indices of the grid points and α_j are the standard piecewise linear hat functions. Substituted in the PDE (1), (3), this approximation gives in general a non-zero residual

R, defined by

$$R\left(\frac{\partial U}{\partial \tau}, \frac{\partial X}{\partial \tau}, \frac{\partial Y}{\partial \tau}\right) = \frac{\partial U}{\partial \tau} - U_x \frac{\partial X}{\partial \tau} - U_y \frac{\partial Y}{\partial \tau} - L(U) \tag{5}$$

A least-squares minimization is performed on R with respect to the unknowns $\partial U_i/\partial \tau$, $\partial X_i/\partial \tau$ and $\partial Y_i/\partial \tau$, yielding a system of implicit ODEs

$$\begin{aligned} \langle R, \alpha_i \rangle &= 0 \\ \langle R, -U_x \alpha_i \rangle &= 0 \\ \langle R, -U_y \alpha_i \rangle &= 0, \quad \forall i \in J \end{aligned} \tag{6}$$

where $\langle , \rangle$ is the usual L^2-inner product on Ω (for an elaboration of (6), see Reference 3). This ODE-system must be integrated numerically to obtain the required fully discretized solution. It is known that this system may become very stiff. For integration in time, therefore, a suitable stiff ODE-solver must be used to cover all possibilities.

In practical applications, regularization terms (penalty functions) will be added before the minimization procedure is carried out. These penalties prevent the parametrization of $\dot{U}$, $\dot{X}$ and $\dot{Y}$ becoming degenerate (see Reference 9). Further, they produce forces on the grid movement to prevent the triangles from getting too thin or from losing their orientation. In our experiments we use the penalty functions as defined in Reference 9, but in this section we will not discuss their influence on the grid movement, since the penalties are not the 'driving forces' behind the movement.

Although the gradient weighted version of MFE (see, e.g. References 7 and 20) is more robust than MFE for steep solutions, the phenomena observed below will be essentially the same for GWMFE.

2.2. *Relation of MFE with the method of characteristics*

Only a few theoretical properties of the resulting ODE-system (6) are known. One important property is the relation of MFE, in both 1D and 2D, with the method of characteristics for the scalar hyperbolic PDE with

$$L(u) = -\beta_1(u, x, y, t)\frac{\partial u}{\partial x} - \beta_2(u, x, y, t)\frac{\partial u}{\partial y} \tag{7}$$

It is easy to derive that for β_1 and β_2 linear in u, x, y, while setting aside boundary effects, the ODE-system (6) is equivalent to

$$\begin{aligned} \dot{X}_i &= \beta_1(U_i, X_i, Y_i, t) \\ \dot{Y}_i &= \beta_2(U_i, X_i, Y_i, t) \\ \dot{U}_i &= 0, \quad i \in J \end{aligned} \tag{8}$$

This simple formulation holds for non-linear β_1 and β_2 in 1D as well (see, e.g. Baines[2, 3]). So, the ODE-system is identical to the discretized system of characteristic ODEs for the PDE (1). In the case that

$$L(u) = -\beta_1 \frac{\partial u}{\partial x} - \beta_2 \frac{\partial u}{\partial y} + \varepsilon \Delta u \tag{9}$$

one can expect, that for small ε (and assuming that the parametrization of $\dot{U}$, $\dot{X}$ and $\dot{Y}$ is not degenerate), MFE results in a grid movement more or less the same as (8). (In one dimension one

can even quantify the perturbation of the characteristics produced by the diffusion term (see below).)

In 1D this relation with the characteristics is very useful. For, in that case, shocks and pulses have only one degree of freedom to move: they propagate along the characteristic curves of the PDE. In many cases in two space dimensions, this characteristic behaviour is also very beneficial (cf. References 8 and 12). However, there are some situations in 2D for which this behaviour will give problems. The main purpose of this note is to illustrate this. We will discuss some of these problems in Section 3.

2.3. Node movement for parabolic equations

Theoretically, little is known about the grid movement in 2D induced by MFE when applied to parabolic PDEs. In one space dimension, however, it is possible to get some insight by examining specific PDEs. Thrasher and Sepehrnoori[16] have derived expressions for the so-called asymptotic node velocity and density for the transport equation in 1D. These expressions are obtained by letting the number of grid points in an arbitrary subinterval tend to infinity using the concept of asymptotic grading functions (see below). Here, we will analyse in an analogous way the node distribution in an asymptotic sense for the scalar PDE

$$\frac{\partial u}{\partial t} = \mu \frac{\partial^2 u}{\partial x^2} + F(x, u, t) \tag{10}$$

where F can also contain spatial derivatives of u.

Let us first write down the MFE-equations (6) in one space dimension (without penalty terms) for the PDE (10). These are

$$\sum_j \langle \alpha_i, \alpha_j \rangle \dot{U}_j + \langle \alpha_i, -U_x \alpha_j \rangle \dot{X}_j = \langle \alpha_i, \mu u_{xx} + F \rangle \tag{11a}$$

$$\sum_j \langle -U_x \alpha_i, \alpha_j \rangle \dot{U}_j + \langle -U_x \alpha_i, -U_x \alpha_j \rangle \dot{X}_j = \langle -U_x \alpha_i, \mu u_{xx} + F \rangle, \quad i = 1, \ldots, N \tag{11b}$$

where N denotes the total number of moving-grid points. Before continuing, we must make some assumptions regarding smoothness of the variables and the rate of convergence. For more details we refer to Reference 16.

Let $\mathscr{F}$ be defined by

$$\frac{\mathrm{d}\mathscr{F}}{\mathrm{d}x} = \frac{\partial \mathscr{F}}{\partial x} + \frac{\partial \mathscr{F}}{\partial u} \frac{\partial u}{\partial x} := F \tag{12}$$

Let the spatial domain be defined as $[x_L, x_R]$ and let $[A, B]$ be an arbitrary subinterval, with time-dependent endpoints $A(t)$ and $B(t)$, not including any exceptional points, i.e. points which have a zero curvature or a zero asymptotic node density. Define $j = j(N)$ and $k = k(N)$ so that X_{j-1} is the first node and X_{k+1} the last node, respectively, in $[A, B]$. Define h to be $\max_{j \leq i \leq k+1} \{h_i\}$ and $h_i = X_i - X_{i-1}$. We then need the following assumptions:

(I) convergence and smoothness of the transformation in (2)

(1) $x(\xi, \tau)$ is a continuous asymptotic grading function, i.e.

$$\lim_{N \to \infty} X_{i(N)}(\tau) = \lim_{N \to \infty} x(i(N)/N, \tau)$$

(2) x has continuous third derivatives (except possibly at a finite number of exceptional points).

(II) smoothness of u and F

The function $\mathscr{F}$ in (12) and the exact solution $u(x, t)$ have continuous third derivatives.

(III) rate of convergence

$$X_i = x(i/N) + O(1/N), \quad \text{for } j-1 \leqslant i \leqslant k+1$$

$$\dot{X}_i = x_\tau(i/N) + O(1/N), \quad \text{for } j-1 \leqslant i \leqslant k+1$$

$$X_i - X_{i-1} = x(i/N) - x((i-1)/N) + O(1/N^2), \quad \text{for } j \leqslant i \leqslant k+1$$

$$\dot{X}_i - \dot{X}_{i-1} = x_\tau(i/N) - x_\tau((i-1)/N) + O(1/N^2), \quad \text{for } j \leqslant i \leqslant k+1$$

$$U_i = u(X_i) + O(1/N), \quad \text{for } j-1 \leqslant i \leqslant k+1$$

$$\dot{U}_i = u_t(X_i) + u_x(X_i)\dot{X}_i + O(1/N), \quad \text{for } j-1 \leqslant i \leqslant k+1$$

$$U_i - U_{i-1} = u(X_i) - u(X_{i-1}) + O(1/N^2) \quad \text{for } j \leqslant i \leqslant k+1$$

$$\dot{U}_i - \dot{U}_{i-1} = u_t(X_i) - u_t(X_{i-1}) + u_x(X_i)\dot{X}_i - u_x(X_{i-1})\dot{X}_{i-1} + O(1/N^2) \quad \text{for } j \leqslant i \leqslant k+1$$

From these assumptions the following relations can be derived

(i) $h_i = x_\xi(i/N)/N + O(1/N^2) = O(1/N)$
(ii) $m_i := (U_i - U_{i-1})/h_i = u_x(X_i) + O(h) = u_x(X_{i-1}) + O(h)$
(iii) $\dot{U}_i^+ := \dot{U}_i - m_{i+1}\dot{X}_i = u_t(X_i) + O(h)$, $\dot{U}_i^- := \dot{U}_i - m_i\dot{X}_i = u_t(X_i) + O(h)$
(iv) $\dot{m}_i = u_{x\tau}(X_i) + O(h)$
(v) $\dot{X}_i$ is bounded
(vi) $O(1/N)$ and $O(h)$ are interchangeable.

Working out the inner products and applying Taylor expansions of $\mathscr{F}$ around U_i and X_i, equations (11a) and (11b) can be combined to obtain

$$\begin{aligned} &h_{i+1}(\dot{U}_{i+1}^- + 2\dot{U}_i^+) - h_i(2\dot{U}_i^- + \dot{U}_{i-1}^+) \\ &\quad - 6\left\{\frac{h_{i+1}-h_i}{2}\mathscr{F}_{x_i} + \frac{m_{i+1}h_{i+1} - m_ih_i}{2}\mathscr{F}_{u_i} + \frac{m_{i+1}h_{i+1}^2 + m_ih_i^2}{3}\mathscr{F}_{xu_i}\right. \\ &\quad \left. + \frac{h_{i+1}^2 + h_i^2}{6}\mathscr{F}_{xx_i} + \frac{m_{i+1}^2h_{i+1}^2 + m_i^2h_i^2}{6}\mathscr{F}_{uu_i}\right\} = O(h^3) \end{aligned} \tag{13}$$

The next step is summaton of equation (13) on the interval $[A, B]$ from $i = j$ to k. This yields

$$\begin{aligned} &h_{k+1}(\dot{U}_{k+1}^- + 2\dot{U}_k^+) - h_j(2\dot{U}_j^- + \dot{U}_{j-1}^+) \\ &\quad - 3\{h_{k+1}\mathscr{F}_{x_k} - h_j\mathscr{F}_{x_j} + m_{k+1}h_{k+1}\mathscr{F}_{u_k} - m_jh_j\mathscr{F}_{u_j}\} \\ &\quad + \sum_{i=j+1}^{k}\left[h_i(-\dot{U}_i^- + \dot{U}_{i-1}^+) - 6\left\{-\frac{h_i}{2}(\mathscr{F}_{x_i} - \mathscr{F}_{x_{i-1}}) - \frac{m_ih_i}{2}(\mathscr{F}_{u_i} - \mathscr{F}_{u_{i-1}})\right.\right. \\ &\quad \left.\left. + \frac{m_ih_i^2}{3}(\mathscr{F}_{xu_i} + \mathscr{F}_{xu_{i-1}}) + \frac{h_i^2}{6}(\mathscr{F}_{xx_i} + \mathscr{F}_{xx_{i-1}}) + \frac{m_i^2h_i^2}{6}(\mathscr{F}_{uu_i} + \mathscr{F}_{uu_{i-1}})\right\}\right] = O(h^2) \end{aligned} \tag{14}$$

Using assumption II, the relations (ii), (iii), (iv) and (v), and substituting the PDE (10) we arrive at a discrete formulation of the distribution of the nodes

$$\begin{aligned} &3\mu[h_{k+1}u_{xx}(X_k) - h_ju_{xx}(X_j)] \\ &\quad - \sum_{i=j+1}^{k}\{h_i^2[\mu u_{xxx}(X_i) + u_{xx}(X_i)(\dot{X}_i + \mathscr{F}_{u_i})]\} = O(h^2) \end{aligned} \tag{15}$$

Applying (i), multiplying by N and letting $N \to \infty$ results in the continuous form

$$3\mu\left[\frac{u_{xx}}{\xi_x}\right]_A^B = \int_A^B \frac{\mu u_{xxx} + u_{xx}[\dot{x} + \mathscr{F}_u]}{\xi_x} \mathrm{d}x \tag{16}$$

Finally, differentiating results in the asymptotic node movement

$$\dot{x} = -\mathscr{F}_u + \mu\left(2\frac{u_{xxx}}{u_{xx}} - 3\frac{\xi_{xx}}{\xi_x}\right) \tag{17}$$

Note that relation (17) is valid only in intervals without points with a zero asymptotic node density ($\xi_x = 0$) or with zero curvature ($u_{xx} = 0$). However, Baines[3] stated that grid points cannot pass a point with zero curvature, with the consequence that grid points are confined to regions between two zero-curvature points (the so-called anti-cluster property of singular points). In fact, the grid points are even repulsed from the singular points.

If $\mu \neq 0$, then equation (17) can be integrated to obtain the asymptotic node distribution

$$\xi_x = K(t)|u_{xx}|^{2/3} \exp\left(\frac{1}{3\mu}\int(-\mathscr{F}_u - \dot{x})\mathrm{d}x\right) \tag{18}$$

For $\mu = 0$ and $F = F(u, t)$ equation (17) means that a grid point will propagate along the characteristic $\dot{x} = -\mathscr{F}_u$ and is *not* dependent on the grid distribution. This is the situation as described in Section 2.2. For $\mu \neq 0$ one can easily derive asymptotic node distributions for restricted choices of F. For instance, the node distribution of the so-called shifting pulse, which we used as an example in Reference 20, once every point travels with the velocity of the pulse, is given by

$$\xi_x = K(t)|u_{xx}|^{2/3} \tag{19}$$

which can be derived from (18) provided that $\dot{x} = \dot{x}(t)$ and $F = F(x, t)$. So once every point travels with the velocity of the pulse, the nodes should be distributed by MFE according to (19) and the plots in Reference 20 show indeed that MFE approximately equidistributes some power of the second derivative of the solution.

For convection–diffusion equations like (9), one can derive from (18), assuming that $\dot{x} = 0$, $u_t = 0$ and $F = F(u, t)$, a steady-state distribution

$$\xi_x = K|u_{xx}|^{2/3}|u_x|^{1/3} \tag{20}$$

indicating that in this case a combination of first and second order derivatives is equidistributed.

3. NUMERICAL EXAMPLES

In our numerical experiments we have solved the implicit ODE-system (6) with the (implicit) BDE-integrator SPGEAR of the SPRINT package[4] in the usual way. This means a.o. that the resulting algebraic system is solved by a modified Newton process.

3.1. Example I ('Anisotropy')

Our first example is an anisotropic wavefront (see Whitham, Reference 19, p. 254). In short, anisotropy means that a difference exists between the directions of the characteristic curves of the PDE (the movement of the 'fluid'-particles) and the movement of the wavefront. This phenomenon can not occur in one space dimension. In 2D, anisotropy may give rise to a distorted MFE grid, eventually leading to a breakdown of the numerical time-stepping procedure.

Probably the best way to illustrate this effect is by giving a PDE-example. Consider, for this purpose,

$$\frac{\partial u}{\partial t} = -\beta_1 \frac{\partial u}{\partial x} - \beta_2 \frac{\partial u}{\partial y} + \varepsilon \Delta u \tag{21}$$

on the domain $\Omega = (0, 1) \times (0, 1)$, with

$$\beta_1 = u$$
$$\beta_2 = (\tfrac{3}{2} - u)$$
$$u|_{t=0} = u_{\text{exact}}|_{t=0}$$
$$u|_{\partial\Omega} = u_{\text{exact}}|_{\partial\Omega}$$

and

$$u_{\text{exact}} = \frac{3}{4} - \frac{1}{4} \frac{1}{1 + \exp\left[\dfrac{-4x + 4y - t}{32\varepsilon}\right]}$$

The exact solution of this model problem (a scalar version of the system in Reference 5, p. 89) describes a wave-front with a steep transition area of thickness $O(\varepsilon)$ that moves, under an angle of 135° with the positive x-axis, from the middle of Ω to the upper left corner. For $\varepsilon \downarrow 0$ the transition area becomes steeper, and for $\varepsilon = 0$ a pure hyperbolic situation is created with a discontinuous moving shock.

Formulation (8) reveals that the method of characteristics, and, to a great extent MFE as well, at first will send the grid points to the upper right corner of the domain. This can be seen very easily by writing out the equations (8) for this case ($\varepsilon = 0$):

for $y > x + t$

$$\dot{X}_i = U_i(t) \approx \tfrac{3}{4}$$
$$\dot{Y}_i = \tfrac{3}{2} - U_i(t) \approx \tfrac{3}{4} \tag{22a}$$

and

for $y < x + t$

$$\dot{X}_i = U_i(t) \approx \tfrac{1}{2}$$
$$\dot{Y}_i = \tfrac{3}{2} - U_i(t) \approx 1 \tag{22b}$$

The characteristic movement from $t = 0$ until $t = 1$ is pictured in Figure 1. This grid movement will lead to a coarse grid in the lower left corner of Ω, since all grid points are moved to the upper right corner. Further, at later points in time, a congestion of grid points near the upper side of the domain Ω will arise, owing to the boundary effects. Since, in that area, the relative distance between the nodes will become very small, the penalty functions should keep the points from moving into each other and thus the ease with which the ODE-system (6) can be solved (if at all) will become very dependent upon the correct choice of the penalty functions. It could easily result in a drastic drop of performance caused only by inadmissible triangle orientations during the Newton process. It must be noted, however, that for $\varepsilon \downarrow 0$ MFE will resemble the method of characteristics more and more, resulting in an almost exact solution in each grid point. In Figure 2 a boundary layer of points is shown, obtained by applying MFE to problem (21) with $\varepsilon = 5 \times 10^{-3}$ and a uniform starting grid of 11×11 moving grid points. At $t \approx 1{\cdot}02$ the computa-

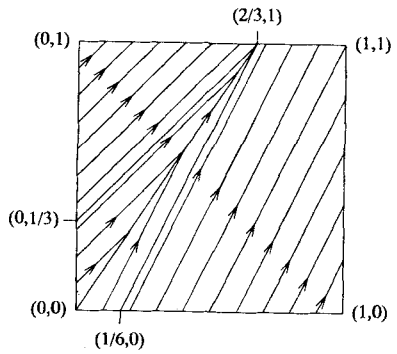

Figure 1. Node movement by characteristics from $t = 0$ to $t = 1$

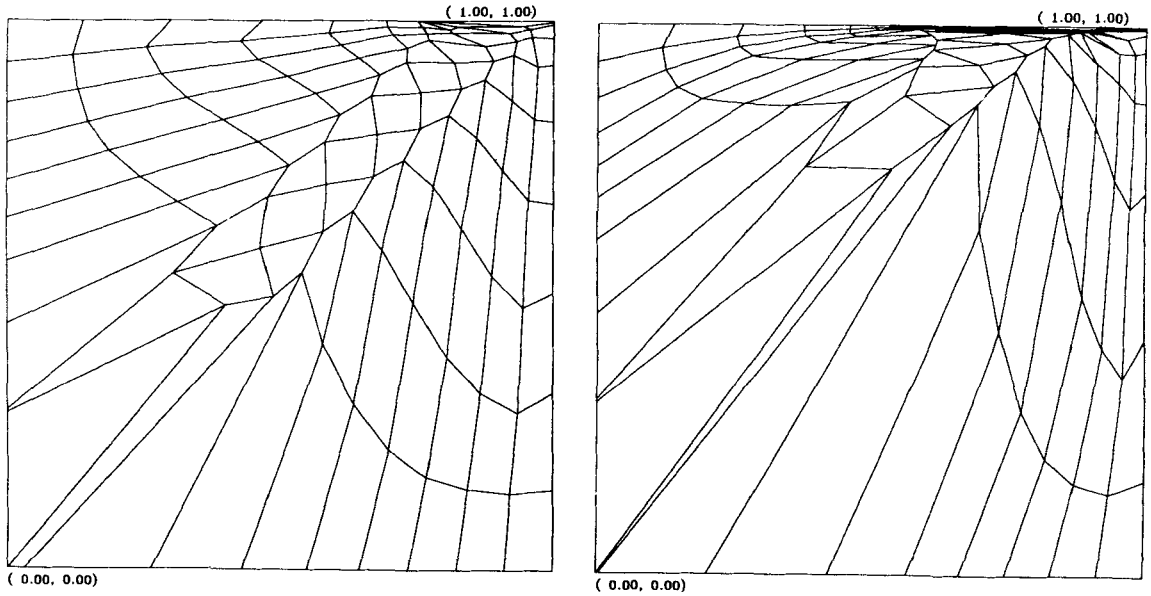

Figure 2. MFE-grid for Example I at $t = 0{\cdot}5$ and $1{\cdot}0$. Dividing each quadrilateral by the diagonal from upper left to lower right gives the MFE-triangles

tional process breaks down because of the unacceptable triangle orientations. This could be prevented by taking larger penalty values, resulting in a less accurate solution.

It is obvious that for these situations a procedure to delete and create nodes could be added to MFE to prevent a congestion of grid points and to keep the finite element approximation of the solution accurate enough. Also, for this special case, a solution to eliminate the anisotropy in the PDE could be found. One could think of applying a transformation to the PDE that describes a rotation of the variables over an angle $\phi = 135°$. In the new variables the characteristic curves and the direction of the normal to the wavefront would coincide (the anisotropy would then cease to exist). In general situations, however, it is, *a priori*, unclear how to choose ϕ, especially as ϕ could even be time-dependent. So far, it has not been possible to reformulate MFE in a proper way to generate such a transformation automatically.

3.2. Example II ('Grid rotation')

Our second example, copied from Reference 14, is concerned with the fact that in 2D an unwanted rotation of the grid can occur. To illustrate this, consider

$$\frac{\partial u}{\partial t} = -\beta_1 \frac{\partial u}{\partial x} - \beta_2 \frac{\partial u}{\partial y} \tag{23}$$

on the domain $\Omega = (-0{\cdot}5, 1{\cdot}5 \times (-0{\cdot}5, 1{\cdot}5)$, with

$$\beta_1 = +\pi(y - \tfrac{1}{2})$$

$$\beta_2 = -\pi(x - \tfrac{1}{2})$$

$$u|_{t=0} = \exp(-80[(x - \tfrac{1}{2})^2 + (y - \tfrac{3}{4})^2])$$

and

$$u|_{\partial\Omega} = 0$$

Although the boundary conditions is mathematically not consistent with the initial condition, it is expected that this will give no problems in numerical computations, since the difference is less than the machine precision.

The exact solution describes a pulse that moves around in circles with a constant speed. During this movement the shape of the pulse does not change. The characteristic curves are circles with centre $(\frac{1}{2}, \frac{1}{2})$, which can be derived immediately from (8) and (23):

$$\dot{U}_i = 0$$

and

$$(X_i - \tfrac{1}{2})^2 + (Y_i - \tfrac{1}{2})^2 = r_i^2, \quad 0 < r_i < 1$$

with $\dot{r}_i = 0$, $\forall_i$.

In contrast with the previous example, the movement of the grid points might be called ideal. They follow the steep parts of the solution in an optimal way and MFE benefits by this property, resulting in a good approximation. However, since we fixed the corner points of the square, the grid will exhibit an unwanted spiral structure. This occurs when the pulse has moved down to the lower region of Ω. A consequence of this effect is the so-called line tangling, a 2D version of node crossing in 1D. The numerical procedure will break down whenever this occurs, again because of inadmissible triangle orientations during the Newton process. In this case, however, larger penalty values can only delay but not prevent the breakdown. We show this spiral effect in Figure 3, where we pictured the grids produced by MFE, at various time values. The starting grid consists again of 11×11 points of which 5×5 are distributed uniformly around the cone in $(0{\cdot}25, 0{\cdot}75) \times (0{\cdot}5, 1{\cdot}0)$. At $t \approx 1{\cdot}52$ the computation breaks down. Again the MFE-approximation in each grid point is rather accurate and the performance of MFE in the time-stepping process is satisfying until the spiral structure leads to line tangling.

Note that in this case intermediate grid rezoning or annihilation and creation of points, based only on the accuracy of the MFE-approximation, would be no cure for the grid distribution problem. Of course, there are some other means to check this effect, again for this special case. First, one could allow the grid points on the boundary to move with the internal points (i.e. 'move around the corner'). For this problem, for instance, it is easy to replace Ω initially by a circular domain and allow free movement of the boundary nodes. The grid then produces no longer spirals, but is congruent with the initial grid and rotated with the characteristic velocity, and the problem is solved without any trouble. Another trick to avoid that the numerical procedure

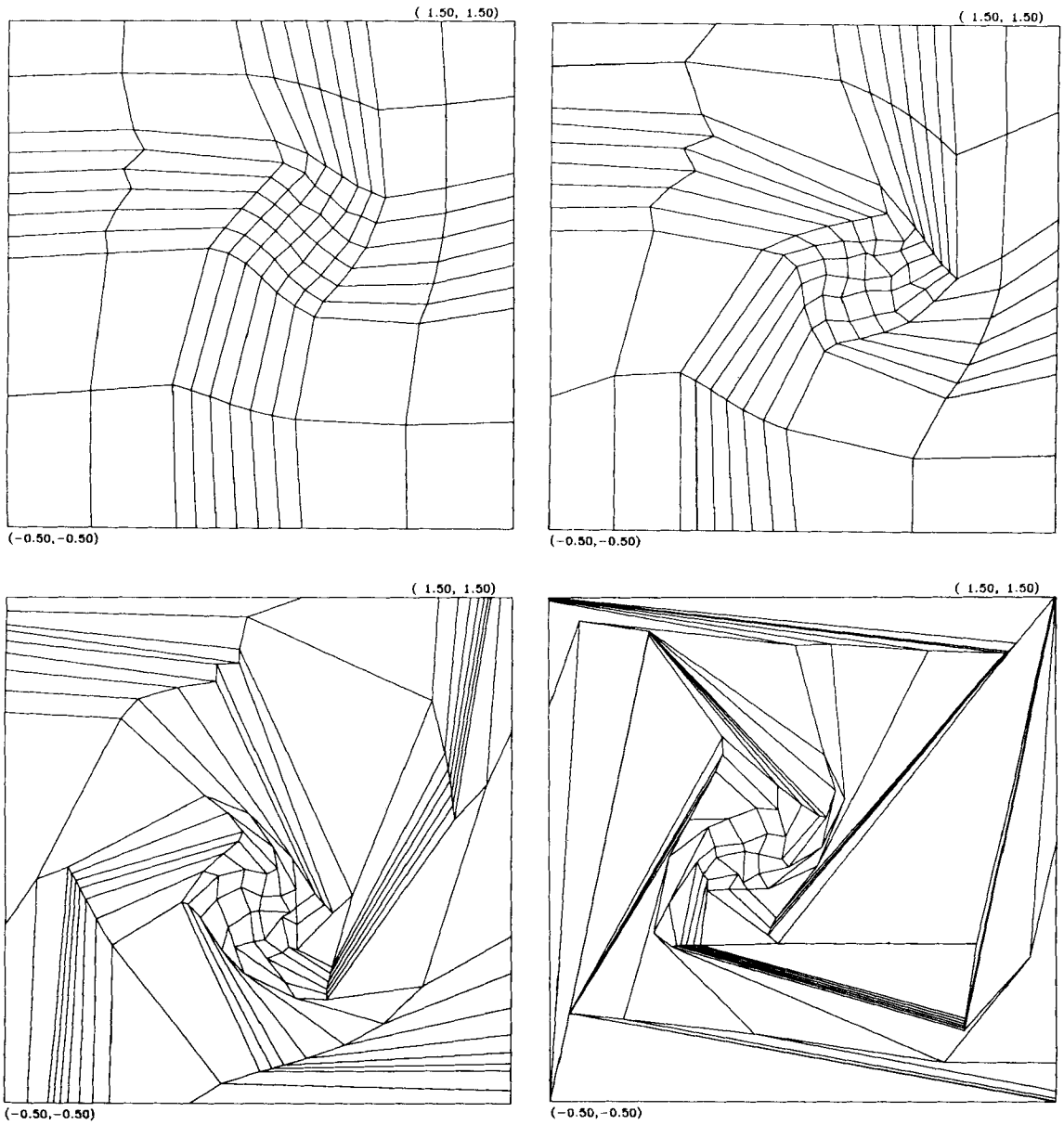

Figure 3. MFE-grid for Example II at $t = 0{\cdot}25, 0{\cdot}5, 1{\cdot}0$ and $1{\cdot}5$. Dividing each quadrilateral by the diagonal from upper left to lower right gives the MFE-triangles

breaks down is described by Mueller and Carey.[14] They add an extra penalty term to the method, which brings on an anti-rotation to the grid movement. This regularization term, however, has only a limited working: with any choice of the constant, appearing in the penalty, there remains some point of time for which the line tangling takes place. With larger penalty values the method would merely collapse at a later moment in the time-integration. But larger penalty values also result in a worse resolution of the pulse, yielding larger errors during the computation.

3.3. *Example III ('Parabolic pulse')*

In the two previous examples we encountered difficulties in applying MFE because of its strong relation with the method of characteristics for hyperbolic equations. Next, we give an example of a PDE with an exact solution very similar to that of model (23), but now the PDE has a parabolic character. It has already been treated by several authors,[1, 17] and is defined by

$$\frac{\partial u}{\partial t} = \Delta u + f(x, y, t) \tag{24}$$

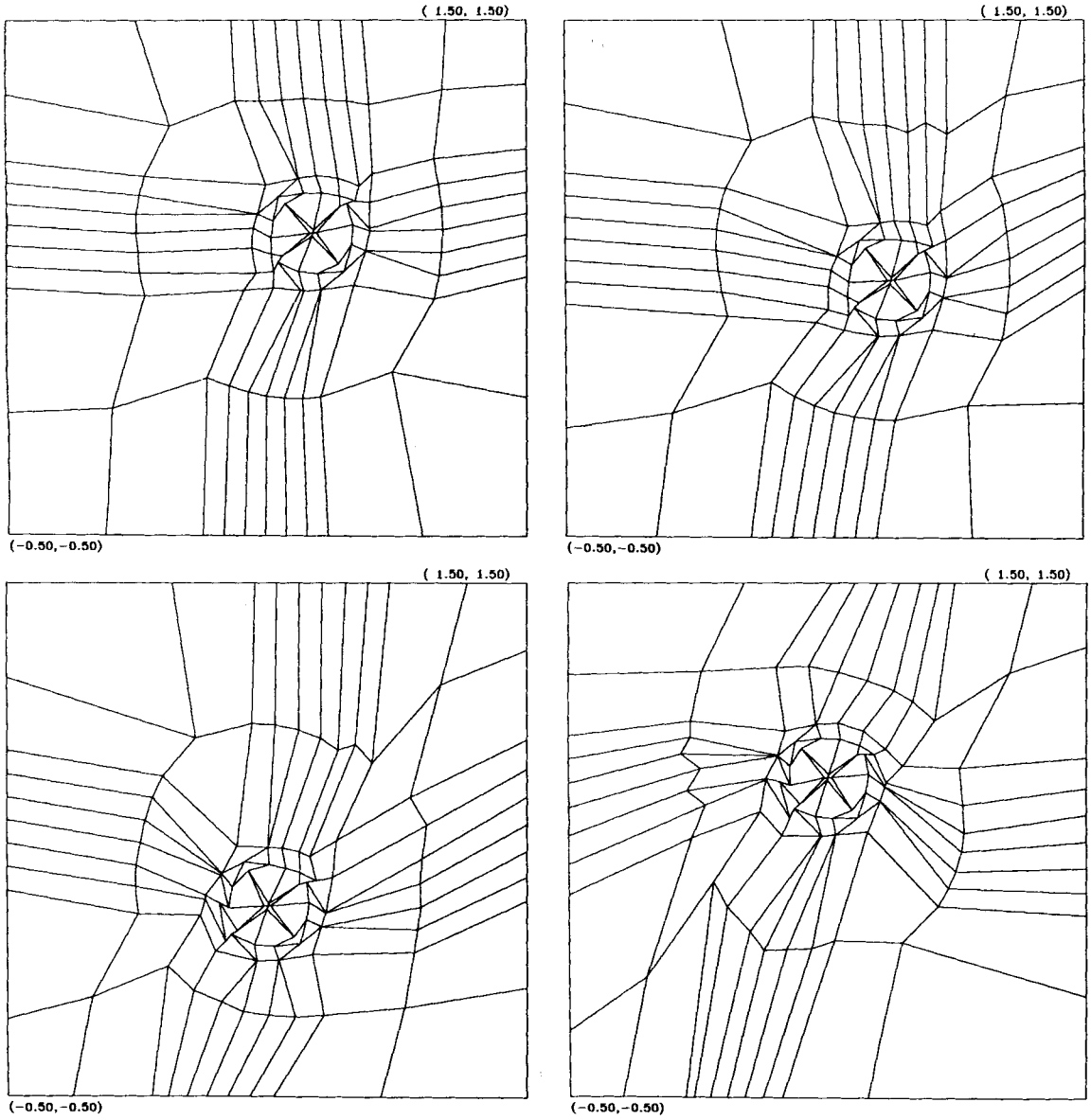

Figure 4. MFE-grid for Example III at $t = 0{\cdot}25$, $0{\cdot}5$, $1{\cdot}0$ and $2{\cdot}0$. Dividing each quadrilateral by the diagonal from upper left to lower right gives the MFE-triangles

on the domain $\Omega = (-0{\cdot}5, 1{\cdot}5) \times (-0{\cdot}5, 1{\cdot}5)$, with

$$u|_{t=0} = u_{\text{exact}}|_{t=0}$$

$$u|_{\partial\Omega} = u_{\text{exact}}|_{\partial\Omega}$$

The source $f(x, y, t)$ is chosen so that the exact solution is

$$u_{\text{exact}} = \exp(-80[(x - r(t))^2 + (y - s(t))^2])$$

where

$$r(t) = (2 + \sin(\pi t))/4, \quad s(t) = (2 + \cos(\pi t))/4$$

This solution is a rotating pulse and thus very similar to the solution of Example II. However, in contrast with the hyperbolic Example II, the grid points do not move according to a principle like (8). In particular, MFE, applied to (24), shows no spiral effect. The points are not stuck to their position on the pulse and the grid structure remains more or less the same during the time-stepping. This is illustrated in Figure 4, where we pictured the grid at several points in time. Although the error of the approximation is higher than in Example II (this can be repaired by increasing the number of points), the procedure does not break down because of grid tangling. On the contrary, once the grid has been forced around the cone, the time-stepping process is satisfying, although the penalty choice is also of influence in this case.

Finally, noteworthy is that the concentration of triangles in regions with large second order derivatives indicates a similar equidistribution behaviour as stated in Section 2.3 for one dimension.

4. CONCLUSIONS

For hyperbolic or strongly convection dominated convection–diffusion equations, the grid points are moved by MFE in a way similar to the method of characteristics. This results in a very good approximation of the solution but sometimes also in distorted grids, because the grid movement is independent of the grid distribution. Such grids then eventually cause the numerical time-stepping to fail. A procedure to delete and create points could in some cases be a remedy, but will, on the other hand, complicate the method considerably.

For scalar parabolic equations one can show that in 1D the MFE-movement of a grid point does depend on the grid distribution. MFE approximates a transformation striving after equidistribution of derivatives of the solution. An example showed that possibly this remains valid also in 2D.

REFERENCES

1. S. Adjerid and J. E. Flaherty, 'A local refinement finite element method for two-dimensional parabolic systems', *SIAM J. Sci. Stat. Comp.*, **9**, 792–811 (1988).
2. M. J. Baines, 'Moving finite elements and approximate Legendre transformations', *Numerical Analysis Report 5/89*, University of Reading, 1989.
3. M. J. Baines, 'An analysis of the moving finite element procedure', *SIAM J. Numer. Anal.*, **28**, 1323–1349 (1991).
4. M. Berzins and R. M. Furzeland, 'A user's manual of SPRINT—A versatile software package for solving systems of algebraic, ordinary and partial differential equations: Part 1—Algebraic and ordinary differential equations', *Report TNER.85.058*, Thornton Research Centre, Shell Research Ltd., U.K., 1985.
5. J. H. M. ten Thije Boonkkamp, 'The numerical computation of time-dependent, incompressible fluid flow, *Ph.D. Thesis*, Universiteit van Amsterdam, 1988.
6. J. U. Brackbill and J. S. Saltzman, 'Adaptive zoning for singular problems in two dimensions', *J. Comp. Phys.*, **46**, 342–368 (1982).
7. N. Carlson and K. Miller, 'Gradient weighted moving finite elements in two dimensions', in D. L. Dwoyer *et al.* (eds.), *Finite Elements Theory and Application*, Springer-Verlag, Berlin, 1988, pp. 151–164.

8. M. J. Djomehri, S. K. Doss, R. J. Gelinas and K. Miller, 'Applications of the moving finite element method for systems in 2D', *Preprint*, 1985 (submitted to *J. Comp. Phys.*).
9. M. J. Djomehri and K. Miller, 'A moving finite element code for general systems of PDE's in 2-D', *Report PAM-57*, Center for Pure and Applied Mathematics, University of California, Berkeley, 1981.
10. H. A. Dwyer, 'A discussion of some criteria for the use of adaptive gridding', in I. Babuška *et al.* (eds.), *Adaptive Computational Methods for PDEs*, SIAM, Philadelphia, 1983, pp. 111–122.
11. K. Miller, 'Moving finite elements II', *SIAM J. Numer. Anal.*, **18**, 1033–1057 (1981).
12. K. Miller, 'Recent results on finite element methods with moving nodes', in I. Babuška *et al.* (eds.), *Accuracy Estimates and Adaptive Refinements in Finite Element Computations*, Wiley, New York, 1986, pp. 325–328.
13. K. Miller and R. N. Miller, 'Moving finite elements I' , *SIAM J. Numer. Anal.*, **18**, 1019–1032 (1981).
14. A. C. Mueller and G. F. Carey, 'Continuously deforming finite elements,' *Int. j. numer. methods eng.*, **21**, 2099–2126 (1985).
15. L. R. Petzold, 'Observations on an adaptive moving grid method for one-dimensional systems of partial differential equations', *Appl. Numer. Math.*, **3**, 347–360 (1987).
16. R. Thrasher and K. Sepehrnoori, 'On equidistributing principles in moving finite element methods', *J. Comp. Appl. Math.*, **16**, 309–318 (1986).
17. R. A. Trompert and J. G. Verwer, 'A static-regridding method for two-dimensional parabolic partial differential equations', *Appl. Numer. Math.*, **8**, 65–90 (1991).
18. J. G. Verwer, J. G. Blom, R. M. Furzeland and P. A. Zegeling, 'A moving-grid method for one-dimensional PDEs based on the method of lines', in J. E. Flaherty *et al.* (eds.), *Adaptive Methods for Partial Differential Equations*, SIAM, Philadelphia, 1989, pp. 160–175.
19. G. B. Whitham, *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.
20. P. A. Zegeling and J. G. Blom, 'An evaluation of the gradient-weighted moving-finite-element method in one space dimension', *Report NM-R9006*, Centre for Mathematics and Computer Science (CWI), Amsterdam, 1990 (to appear in *J. Comp. Phys.*).