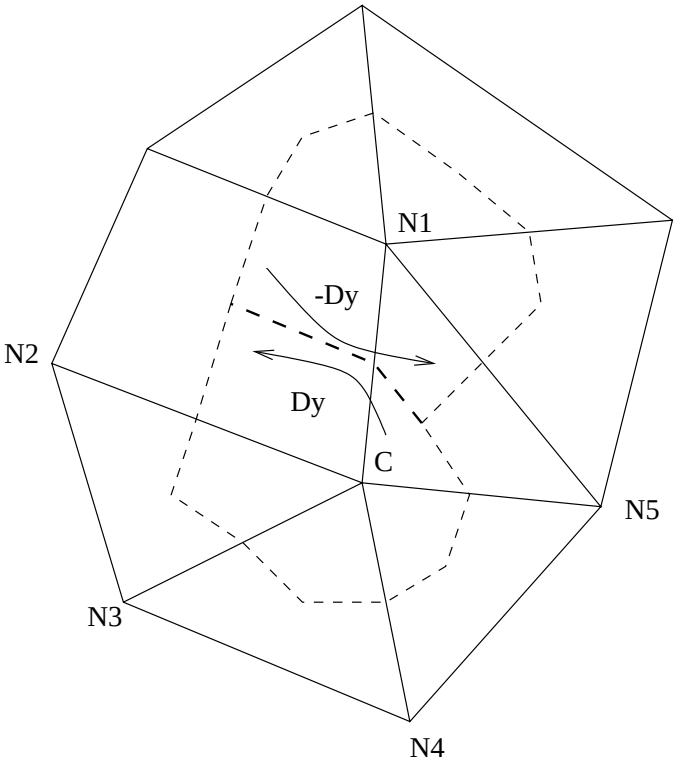
FOI-R–0121–SE
May 2001
ISSN 1650-1942

Scientific Report

Carl Adamsson, Karl Forsberg and Jan Nordström

Finite volume methods, unstructured meshes and strict stability



Division of Aeronautics, FFA
SE-172 90 STOCKHOLM

FOI – Swedish Defence Research Agency
Division of Aeronautics, FFA
SE-172 90 STOCKHOLM

FOI-R–0121–SE
May 2001
ISSN 1650-1942

Scientific Report

Carl Adamsson, Karl Forsberg and Jan Nordström

Finite volume methods, unstructured meshes and strict stability

Division of Aeronautics, FFA
SE-172 90 STOCKHOLM

Abstract

The unstructured node centered finite volume method is analyzed and it is shown that it can be interpreted in the framework of summation by parts operators. It is also shown that introducing boundary conditions weakly produces strictly stable formulations.

Contents

1 Introduction

In computational fluid dynamics as well as computational electromagnetics, finite volume methods (FVM) formulated on unstructured grids are widely used to handle complex geometries. In [4], it is shown that strictly stable finite volume methods on structured grids can be constructed from so called summation by parts operators (SBP–operators). These operators satisfy a discrete summation by part rule which mimic the integration by parts rule in the continuous case.

In this work it is shown that one particular node centered FVM, for unstructured grids, can be incorporated in the terminology of SBP–operators. A weak procedure to introduce boundary conditions is shown to produce energy estimates that lead to strict stability. This method is equivalent to the standard penalty procedure, called SAT (simultaneous approximation term) [2], often used for high order finite difference operators of SBP–character.

2 Analysis

2.1 The model problem

The model equation considered in this paper is of the form:

$$u_t + Au_x + Bu_y = 0, \quad (x, y) \in \Omega \subset \mathbb{R}^2 \tag{1}$$

with suitable boundary and initial conditions. In (1), u is the vector of unknowns and A and B are constant, symmetric, square matrices.

The energy method (see for example [3]) will be used to investigate stability. For the scalar equation $u_t + u_x = 0, 0 \leq x \leq 1, u(0,t) = g(t)$, this gives

$$\frac{d}{dt}\int_0^1 u^2 dx = \frac{d}{dt}||u||^2 = u(0)^2 - u(1)^2 = g^2 - u(1)^2. \tag{2}$$

The boundary condition keep the potential growth under control, i.e there is an upper limit for the growth rate.

For equation (1), the energy method gives

$$\begin{aligned} &\frac{d}{dt}\iint_\Omega u^2 dxdy = \frac{d}{dt}||u||^2 = -\oint_{\partial\Omega} u^T Audy + \oint_{\partial\Omega} u^T Budx = \\ &= -\oint_{\partial\Omega} u^T \left(A\hat{x} + B\hat{y}\right) u \cdot \hat{n} ds \end{aligned} \tag{3}$$

with the use of Greens formula and the symmetry of A and B. In (3), $\hat{n}$ is the outward pointing unit normal to $\partial\Omega$, $\hat{x}$ and $\hat{y}$ are the unit vectors in the x– and y–directions and ds is the infinitesimal arc length element counted counter clockwise around Ω.

The number of boundary conditions at any point on the boundary is the least number that make $(A\hat{x} + B\hat{y}) \cdot \hat{n}$ positive semidefinite. When referring to the model problem in this report it is assumed that the boundary conditions in (1) are such that this is true. In the examples in section 3 this will be explicitly shown to hold.

2.2 SBP operators

First consider the scalar equation $u_t + u_x = 0$ and its semi discrete approximation

$$\boldsymbol{u}_t + D_x \boldsymbol{u} = 0, \quad D_x = P^{-1}Q_x \tag{4}$$

where P is symmetric and positive definite and $Q_x + Q_x^T = \text{diag}[-1, 0, \cdots, 0, 1]$. D_x is said to be a summation–by–parts (SBP) operator. With SBP operators it is possible to mimic the energy estimate (2) in the semi discrete case. Multiplying (4) by $\boldsymbol{u}^T P$, transposing and adding gives

$$\frac{d}{dt}(\boldsymbol{u}^T P \boldsymbol{u}) = \frac{d}{dt}||\boldsymbol{u}||_P^2 = u_1^2 - u_N^2. \tag{5}$$

The symmetry and positive definiteness of P guarantees that it can be used as a norm. The result (5) should be compared to (2).

Next we discretize (1) in space by introducing the vector $\boldsymbol{u}$ of length $N = dn$ where d is the number of unknowns in (1) and n is the number of grid-points. The elements of $\boldsymbol{u}$ are organized such that the first n elements are the discretization of the first variable in u, the elements $n + 1 \ldots 2n$ are the discretization of the second variable and so on. Furthermore we introduce discrete x– and y– derivative operators D_x and D_y. Equation (1) can now formally be written

$$\boldsymbol{u}_t + (A \otimes D_x)\boldsymbol{u} + (B \otimes D_y)\boldsymbol{u} = 0 \tag{6}$$

where $\otimes$ is the Kronecker product.

For 2D-equations on unstructured grids, a generalized SBP-concept will be used. We aim for Q_x and Q_y to be such that

$$\boldsymbol{u}^T(Q_x + Q_x^T)\boldsymbol{u} \approx \oint_{\partial\Omega} u^2 dy, \quad \boldsymbol{u}^T(Q_y + Q_y^T)\boldsymbol{u} \approx -\oint_{\partial\Omega} u^2 dx. \tag{7}$$

If (7) holds, the discrete energy method will lead to an approximation that corresponds to equation (3).

2.3 The finite volume method

In a node centered FVM on an unstructured grid, the unknowns are associated with the nodes in the grid. The control-volumes that constitute the dual grid are defined as follows. Each control-volume is a polygon with its vertices at the centers of gravity of the surrounding triangles (or quadrilaterals) and at the midpoints of the grid–sides, see figure 1.

The FVM is obtained by integrating the equation over the control volumes. Start with the scalar equation $u_t + u_x = 0$ and integrate over a control volume, Ω_C , to get:

$$\iint_{\Omega_C} u_t dxdy + \iint_{\Omega_C} u_x dxdy = \iint_{\Omega_C} u_t dxdy + \oint_{\partial\Omega_C} u dy = 0. \tag{8}$$

The semi discrete approximation of equation (8) can be written

$$P\boldsymbol{u}_t + Q\boldsymbol{u} = 0 \tag{9}$$

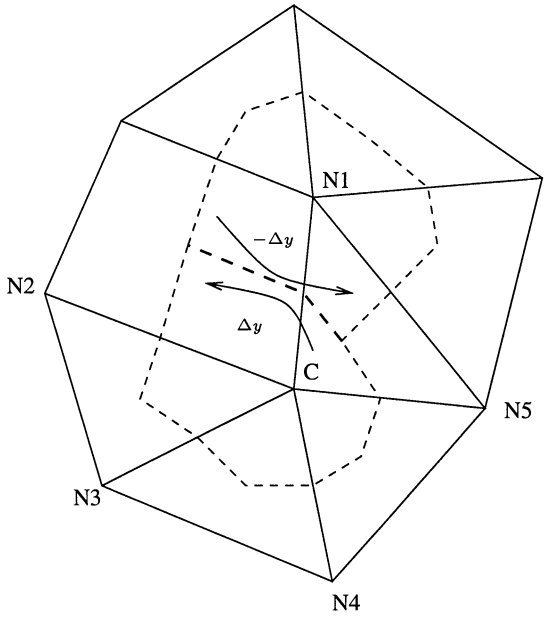
where P is a diagonal matrix with the control volumes on the diagonal and Q_x associates with each node an approximation to the line integral of u around the boundary of the control volume. The approximation to this line integral, the fluxes, are computed as follows.

Consider a node in the interior of the mesh with index C. The flux is the integral

$$\text{flux} = \oint_{\partial\Omega_C} u dy \tag{10}$$

where Ω_C is the dual grid cell that belongs to the node C. The node C has neighbours which indices N_i, see figure 1. Note that each of the neighbours can be

Figure 1. Part of the grid and the dual grid.



associated in a one–to–one manner with the two sides of the polygon $\partial\Omega_C$ that have a common vertex at the side connecting the neighbour to the node C. Each neighbour will contribute to the flux with one term. This term is the mean value of u_C and u_{N_i} times Δy over the corresponding dual grid side. This can formally be written:

$$\text{flux} = \sum_i \frac{u_C + u_{N_i}}{2}\Delta y_i = \sum_i u_C \frac{\Delta y_i}{2} + \sum_i u_{N_i}\frac{\Delta y_i}{2} \tag{11}$$

where the sum goes over all neighbours to the point C. Not considering the boundary of the domain, (11) leads to

$$\begin{cases} Q_{CC} &= \sum_i \frac{\Delta y_i}{2} = 0 \qquad \text{(closed loop)} \\ Q_{CN_i} &= \frac{\Delta y_i}{2} = -Q_{N_iC} \end{cases} \tag{12}$$

i.e the matrix Q is skew symmetric in the interior.

Let us consider the case where no boundary condition (b.c) is necessary. The flux through the boundary edge is calculated as the node value at the boundary node, u_B, times the corresponding Δy_B, see figure 2. Formally:

$$\begin{aligned} \text{flux} = &\sum_i \frac{u_B + u_{N_i}}{2}\Delta y_i + u_B \Delta y_B = \\ = &u_B \Delta y_B + \sum_i u_B \frac{\Delta y_i}{2} + \sum_i \frac{u_{N_i}}{2}\Delta y_i. \end{aligned} \tag{13}$$

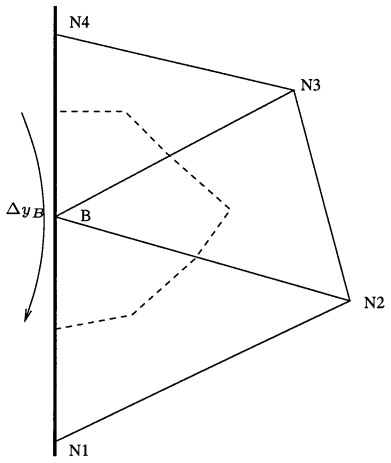
Note that the first sum is not over a closed loop. From figure 2 we obtain

$$\sum_i \Delta y_i = -\Delta y_B. \tag{14}$$

Thus we have

$$\text{flux} = \sum_i u_{N_i}\frac{\Delta y_i}{2} + u_B \frac{\Delta y_B}{2} \tag{15}$$

Figure 2. The geometry at the boundary.



which leads to

$$\begin{cases} Q_{BB} &= \frac{\Delta y_B}{2} \\ Q_{BN_i} &= \frac{\Delta y_i}{2} = -Q_{N_iB} \end{cases}. \tag{16}$$

Let us now consider the case with b.c, $u_i = g_i$ at the boundary. Even though we know the u-values at the boundary a priori from the b.c we do *not* remove those points from the scheme. Instead we impose the b.c weakly. The fluxes, using (14), become:

$$\text{flux} = \sum_i \frac{u_B + u_{N_i}}{2}\Delta y_i + g_B\Delta y_B = \sum_i u_{N_i}\frac{\Delta y_i}{2} - u_B\frac{\Delta y_B}{2} + g_B\Delta y_B. \tag{17}$$

Clearly the elements in Q become

$$\begin{cases} Q_{BB} &= -\frac{\Delta y_B}{2} \\ Q_{BN_i} &= \frac{\Delta y_i}{2} = -Q_{N_iB}. \end{cases} \tag{18}$$

The results (12), (16) and (18) yield

$$P\boldsymbol{u}_t + Q\boldsymbol{u} + b = 0. \tag{19}$$

In (19) we have introduced

$$b_i = \begin{cases} g_i\Delta y_i & \text{at the boundary with b.c} \\ 0 & \text{otherwise} \end{cases}. \tag{20}$$

2.4 Generalization to the full problem

The generalization of (19) in the interior is straight forward. The discretisation of (1) leads to equation (6). In (6) we use $D_x = P^{-1}Q_x$ and $D_y = P^{-1}Q_y$ where P, Q_x and Q_y are derived as in the previous section. The generalization of (19) becomes

$$(I \otimes P)\boldsymbol{u}_t + (A \otimes Q_x)\boldsymbol{u} + (B \otimes Q_y)\boldsymbol{u} + b = 0 \tag{21}$$

where b is a vector of the same length as $\boldsymbol{u}$, i.e of length $N = nd$ where d is the number of unknowns and n is the number of grid points. ($\boldsymbol{u}$ is organized as in section 2.2). The energy method applied to (21) gives

$$\begin{aligned}\frac{d}{dt}||\boldsymbol{u}||^2_{I\otimes P} = \\ = -\boldsymbol{u}^T(A\otimes(Q_x+Q_x^T))\boldsymbol{u} - \boldsymbol{u}^T(B\otimes(Q_y+Q_y^T))\boldsymbol{u}^T + 2\boldsymbol{u}^T b\end{aligned} \tag{22}$$

where the symmetry of A and B has been used. Note that the energy depends only on the symmetric part of Q_x and Q_y and the boundary data. As seen in section 2.3 the symmetric parts are zero in the interior of the grid; the symmetric elements are introduced at the boundary.

For each row of $A\otimes Q_x$ (and $B\otimes Q_y$) that corresponds to a boundary point, the boundary flux must be calculated. In a typical time stepping procedure with no b.c, we calculate the flux using the local value of the variable exactly as in the previous section. This will produce a symmetric element as in equation (16) in Q_x (or Q_y) and thus in $A\otimes Q_x$ (or $B\otimes Q_y$). With boundary conditions in the form $u_i = g_i, i = 1..n$ we use the boundary data to calculate the flux as described above. This will also produce a symmetric element (see equation (18)) in $A\otimes Q_x$ with the boundary data in the b–vector.

If a combination of boundary data and local values are used to calculate the flux, it must lead to an energy estimate for the discrete problem. There is not always a unique way to do this. One method that always work is to substitute boundary data for the ingoing characteristic variables and use local data for the outgoing ones. This method is illustrated in the examples below.

There is a standard penalty procedure to introduce b.c in a stable way when working with SBP–operators called SAT (Simultaneous Approximation Term), see for example [2]. In the present case, when the matrix P is diagonal, the SAT procedure is equivalent to the weak procedure described above. Equation (17) can be written

$$\text{flux} = \sum_i u_{N_i}\frac{\Delta y_i}{2} + u_B\frac{\Delta y_B}{2} + \underbrace{(g_B - u_B)\Delta y_B}_{b_B}. \tag{23}$$

The relation (23) is a penalty formulation, where the matrix elements in Q are given by (16).

2.5 Observations

There are other methods to introduce b.c than the weak method described in the previous paragraph. One can, for example, remove the boundary points from the scheme altogether and satisfy the b.c exactly. This reduces the size of the system (19), but sometimes introduces stability problems, see [4], [1].

The relation (19) (or (21)) is a linear system of ordinary differential equations. If the vector b contains known boundary data, i.e not the unknowns of the system, the solution can be written

$$u = -e^{-P^{-1}Qt}\left(f - \int_0^t e^{P^{-1}Qt}b(t)dt\right) \tag{24}$$

which is stable when $P^{-1}Q$ is negative semidefinite, independently of b.

The FVM discussed here can be shown to be equivalent to a finite element method (FEM) in the interior of the grid if the grid consist of triangles only. The FEM is a modification of the classical variant obtained by using piecewise linear 'tent–functions'. The FEM can be written in the form (9) where

$$P_{ij} = \int_{\Omega} \phi_i \phi_j dxdy, \quad Q_{ij} = \int_{\Omega} \phi_i \frac{\partial}{\partial x} \phi_j dxdy. \tag{25}$$

Here ϕ_i is the piecewise linear function that is 1 in node i and 0 in all other nodes. The Q–matrix above can be shown to be identical to the Q–matrix from the FVM. To make also the P–matrix identical, the FEM has to be lumped, i.e the P in (25) must be changed to $P^l_{ij} = \delta_{ij} \sum_j P_{ij}$. (where δ_{ij} is the Kronecker δ–symbol).

3 Examples

The following definitions will be used in all examples below. The domain of computation will be $\Omega \subset \mathbb{R}^2$. The boundary of Ω will be denoted $\partial\Omega$. Define Γ_1 and Γ_2 to be such that $\Gamma_1 \cup \Gamma_2 = \partial\Omega$ and Γ_1 is the part of $\partial\Omega$ where $\hat{x} \cdot \hat{n} < 0$ where $\hat{n}$ is the outward pointing unit normal to $\partial\Omega$. This of course implies that $\hat{x} \cdot \hat{n} > 0$ on Γ_2. Moreover if $dl = (dx, dy)$ is the infinitesimal tangent vector to $\partial\Omega$ counted counter clockwise around Ω we have that $dy < 0$ on Γ_1 and $dy > 0$ on Γ_2.

3.1 The one way wave equation

Let us start with

$$u_t + u_x = 0, \quad (x, y) \in \Omega \subset \mathbb{R}^2, \quad u|_{\Gamma_1} = g. \tag{26}$$

Equation (3) for this problem reduces to

$$\frac{d}{dt} \iint_{\Omega} u^2 dxdy = - \oint_{\partial\Omega} u^2 dy = - \int_{\Gamma_1} g^2 dy - \int_{\Gamma_2} u^2 dy. \tag{27}$$

The definition of Γ_2 implies that we have an energy estimate.

We obtain the discrete energy rate by multiplying equation (19) by $\boldsymbol{u}^T$, transposing and adding. The result is

$$\frac{d}{dt}(\boldsymbol{u}^T P \boldsymbol{u}) = -\boldsymbol{u}^T (Q + Q^T)\boldsymbol{u} - 2\boldsymbol{u}^T b. \tag{28}$$

By using the expressions (16) and (18) for the matrix elements of Q we have

$$\boldsymbol{u}^T (Q + Q^T)\boldsymbol{u} = \sum_{\Gamma_2} u_i^2 \Delta y_i - \sum_{\Gamma_1} u_i^2 \Delta y_i \tag{29}$$

where Γ_1 is the boundary with b.c. Furthermore, we have

$$2\boldsymbol{u}^T b = 2 \sum_{\Gamma_1} u_i g_i \Delta y_i. \tag{30}$$

The relations (29), (30) inserted in (28) leads to

$$\frac{d}{dt}(\boldsymbol{u}^T P \boldsymbol{u}) = \sum_{\Gamma_1} (g_i^2 - (u_i - g_i)^2)|\Delta y_i| - \sum_{\Gamma_2} u_i^2 |\Delta y_i| \tag{31}$$

where we have used that $\Delta y_i < 0$ on Γ_1 and $\Delta y_i > 0$ on Γ_2. This is completely similar to (27) (plus a negative definite term). With use of the definition of strict stability introduced in [4], the approximation (19) of problem (26) is strictly stable.

As mentioned in section 2.5 above, the b.c could be introduced by injection, i.e by removing the boundary-points from the scheme. This technique results in a reduced system with a b–vector containing boundary-data and a Q–matrix where rows and columns corresponding to the points on the inflow boundary (Γ_1) are

Figure 3. The spectrum of $P^{-1}Q$ using weak formulation, on an unstructured mesh with 49 nodes.

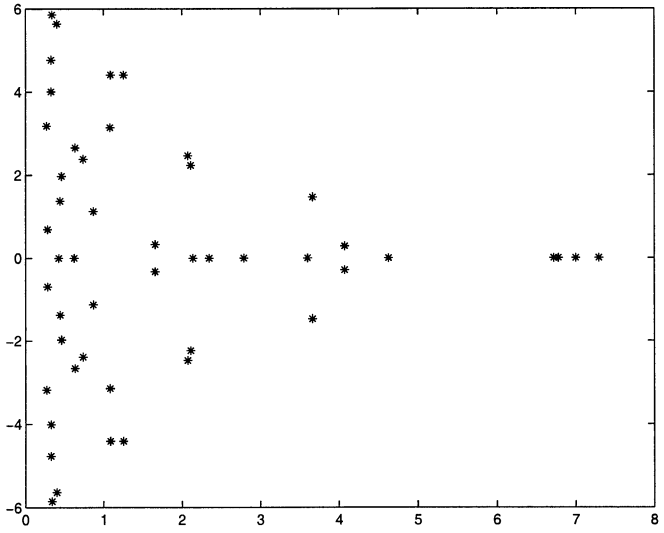
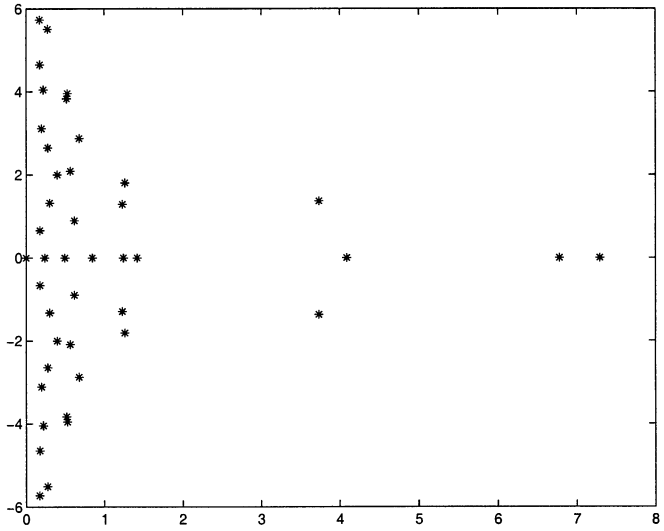


Figure 4. The spectrum of $P^{-1}Q$ using injection, on an unstructured mesh with 49 nodes.



removed. Denote the vector $\boldsymbol{u}$ and matrix P of this new system with $\tilde{\boldsymbol{u}}$ and $\tilde{P}$ respectively. In this case the energy estimate becomes

$$\frac{d}{dt}(\tilde{\boldsymbol{u}}^T \tilde{P} \tilde{\boldsymbol{u}}) = -2 \sum_{\tilde{\Gamma}_1} u_i g_i |\Delta y_i| - \sum_{\Gamma_2} u_i^2 |\Delta y_i| \tag{32}$$

where $\tilde{\Gamma}_1$ is the set of points that are neighbours to the points in Γ_1. That is, points 'one layer in' from the boundary now couples to boundary data. With this technique there is no way to show strict stability in the sense used in [4]. However, stability is given by (32) and (24). In figures 3 - 6 the spectra for both the injection method and the weak method are shown on two grids with different number of nodes. In both cases all eigenvalues have positive real parts which confirms that both methods are stable.

Figure 5. The spectrum of $P^{-1}Q$ using weak formulation, on an unstructured mesh with 105 nodes.

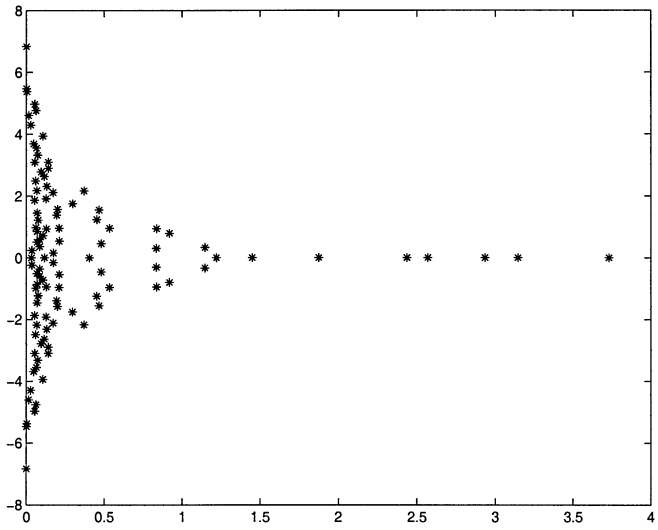
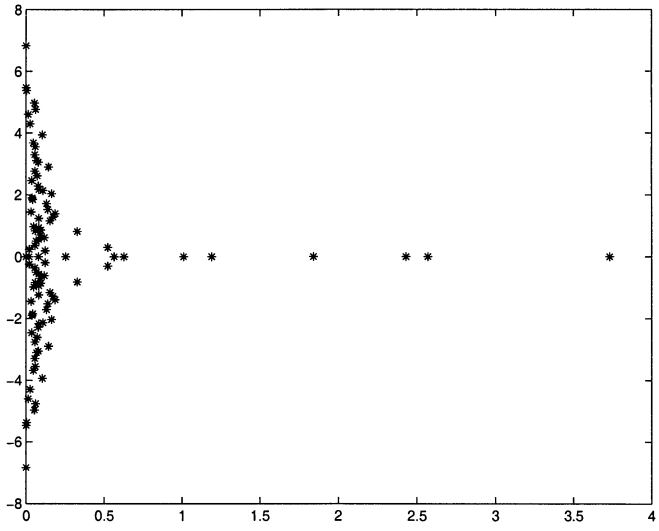


Figure 6. The spectrum of $P^{-1}Q$ using injection, on an unstructured mesh with 105 nodes.



3.2 A system of equations

Consider the 1D Maxwell equations

$$\begin{pmatrix} E \\ H \end{pmatrix}_t + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} E \\ H \end{pmatrix}_x = 0, \quad (x,y) \in \Omega \subset \mathbb{R}^2 \tag{33}$$

with b.c

$$E|_{\partial\Omega} = 0. \tag{34}$$

The energy rate becomes

$$\frac{d}{dt} \iint_{\Omega} E^2 + H^2 dxdy = - \oint_{\partial\Omega} 2EHdy = 0. \tag{35}$$

We introduce discrete representations $\boldsymbol{E}$ and $\boldsymbol{H}$ for the unknowns E and H. If we first consider the system *without* b.c we get

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix}_t + \begin{pmatrix} 0 & Q_f \\ Q_f & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} = 0. \tag{36}$$

Where Q_f uses the matrix elements (16). By introducing the b.c in E we get

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix}_t + \begin{pmatrix} 0 & Q_f \\ Q_b & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} = 0 \tag{37}$$

where Q_b uses the matrix elements (18). The b–vector will be zero because the boundary data is zero.

The discrete energy rate becomes

$$\begin{aligned} \frac{d}{dt}(\boldsymbol{E}^T P \boldsymbol{E} + \boldsymbol{H}^T P \boldsymbol{H}) &= -\boldsymbol{E}^T (Q_f + Q_f^T)\boldsymbol{H} - \boldsymbol{H}^T (Q_b + Q_b^T)\boldsymbol{E} = \\ &= -\sum_{\partial\Omega} E_i H_i \Delta y_i + \sum_{\partial\Omega} H_i E_i \Delta y_i = 0 \end{aligned} \tag{38}$$

Thus we have

$$\frac{d}{dt}(||\boldsymbol{E}||_P^2 + ||\boldsymbol{H}||_P^2) = 0 \tag{39}$$

which means that the discrete energy rate corresponds exactly to the continuous case, i.e the approximation is strictly stable, see [4].

Also in this case the method of injection of b.c works fine. Removing the boundary points means that all symmetric elements of the Q–matrix disappear leaving a pure skew symmetric Q. With zero boundary data, the b–vector will also be zero. Therefore, the energy growth rate is exactly zero in this case too.

3.2.1 Characteristic boundary conditions

Again, consider equation (33) with characteristic b.c

$$\begin{cases} E + H|_{\Gamma_1} = f \\ E - H|_{\Gamma_2} = g \end{cases} \tag{40}$$

The energy equation (3) can, after some algebra, be written

$$\begin{aligned}\frac{d}{dt}\iint_{\Omega} E^2+H^2 dy &= -\oint_{\partial\Omega} 2EH dy = \\ &= -\int_{\Gamma_1}\frac{1}{2}\left(f^2-(E-H)^2\right)dy-\int_{\Gamma_2}\frac{1}{2}\left((E+H)^2-g^2\right)dy\end{aligned} \tag{41}$$

where the definitions of Γ_1 and Γ_2 show that we have an energy estimate. The discrete approximation *without* b.c becomes

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix}_t + \begin{pmatrix} 0 & Q_f \\ Q_f & 0 \end{pmatrix}\begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} = 0 \tag{42}$$

where Q_f uses the matrix elements (16). An arbitrary point at the boundary is updated according to

$$\begin{aligned} P_{BB}\frac{d}{dt}E_B + H_B\Delta y_B + \sum_i \frac{H_B+H_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}H_B + E_B\Delta y_B + \sum_i \frac{E_B+E_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{43}$$

Now we have to decide how to use the b.c to calculate the boundary flux.

Note that the b.c are in the characteristic variables μ and ν, defined as

$$\begin{cases} \mu = E+H \\ \nu = E-H \end{cases} \Leftrightarrow \begin{cases} \frac{1}{2}(\mu+\nu) = E \\ \frac{1}{2}(\mu-\nu) = H \end{cases}. \tag{44}$$

To find a suitable linear combination of boundary data and local data (see section 2.4), we introduce the b.c in terms of the ingoing characteristic variable and transform back to E and H to get

$$\begin{cases} E = \frac{1}{2}(f+\nu) = \frac{1}{2}(f+E-H) \\ H = \frac{1}{2}(f-\nu) = \frac{1}{2}(f-(E-H)) \end{cases} \tag{45}$$

on Γ_1 and

$$\begin{cases} E = \frac{1}{2}(\mu+g) = \frac{1}{2}(E+H+g) \\ H = \frac{1}{2}(\mu-g) = \frac{1}{2}(E+H-g) \end{cases} \tag{46}$$

on Γ_2. The boundary fluxes using (45), (46) becomes

$$\begin{aligned} P_{BB}\frac{d}{dt}E_B + \overbrace{\frac{1}{2}(f_B-E_B+H_B)\Delta y_B}^{1} + \sum_i \frac{H_B+H_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}H_B + \underbrace{\frac{1}{2}(f_B+E_B-H_B)\Delta y_B}_{2} + \sum_i \frac{E_B+E_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{47}$$

on Γ_1 and

$$\begin{aligned} P_{BB}\frac{d}{dt}E_B + \overbrace{\frac{1}{2}(-g_B+E_B+H_B)\Delta y_B}^{3} + \sum_i \frac{H_B+H_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}H_B + \underbrace{\frac{1}{2}(g_B+E_B+H_B)\Delta y_B}_{4} + \sum_i \frac{E_B+E_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{48}$$

on Γ_2. The formulations (47), (48) are obtained if one uses the matrix elements (18) and introduce a right hand side vector b containing the expressions 1, 2, 3 and 4 at the corresponding boundary points.

The energy method applied to the equation

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix}_t + \begin{pmatrix} 0 & Q_b \\ Q_b & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{H} \end{pmatrix} + b = 0 \tag{49}$$

leads to (Q_b uses the matrix elements (18))

$$\begin{aligned}
\frac{d}{dt}(\boldsymbol{E}^T P \boldsymbol{E} + \boldsymbol{H}^T P \boldsymbol{H}) = & \overbrace{+\sum_{\partial\Omega} 2E_i H_i \Delta y_i}^{\text{diag. elements from (18)}} \quad - \\
& - 2\sum_{\Gamma_1} E_i \frac{1}{2}(f_i - E_i + H_i)\Delta y_i - 2\sum_{\Gamma_1} H_i \frac{1}{2}(f_i + E_i - H_i)\Delta y_i - \\
& - 2\sum_{\Gamma_2} E_i \frac{1}{2}(E_i + H_i - g_i)\Delta y_i - 2\sum_{\Gamma_2} H_i \frac{1}{2}(E_i + H_i + g_i)\Delta y_i = \\
& = -\sum_{\Gamma_1} \left((E_i + H_i)f_i - E_i^2 - H_i^2\right)\Delta y_i - \\
& - \sum_{\Gamma_2} \left(-(E_i - H_i)g_i + E_i^2 + H_i^2\right)\Delta y_i = \\
& = -\sum_{\Gamma_1} \left\{ \frac{1}{2}\left(f_i^2 - (E_i - H_i)^2\right) - \frac{1}{2}\left(f_i - (E_i + H_i)\right)^2 \right\}\Delta y_i - \\
& - \sum_{\Gamma_2} \left\{ \frac{1}{2}\left((E_i + H_i)^2 - g_i^2\right) + \frac{1}{2}\left((E_i - H_i) - g_i\right)^2 \right\}\Delta y_i
\end{aligned} \tag{50}$$

which should be compared to the continuous energy estimate (41). The definitions of Γ_1 and Γ_2 show that (50) gives an energy estimate. As in section 3.1 there are small dissipative terms but no terms that could lead to instability. The approximation (49) of (33) and (40) is strictly stable in the sense introduced in [4]. Note also that (49) is a penalty formualtion.

The method of injection of boundary data could be used here too in a straight forward way.

3.2.2 The weak formulation in penalty form

The diagonalized form of the problem in section 3.2 reads

$$\begin{pmatrix} \mu \\ \nu \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \nu \end{pmatrix}_x = 0, \quad (x, y) \in \Omega \subset \mathbb{R}^2 \tag{51}$$

$$(\mu - \nu)|_{\partial\Omega} = 0. \tag{52}$$

The energy equation (3) becomes

$$\frac{d}{dt}\iint_{\Omega} \mu^2 + \nu^2 dxdy = -\oint_{\partial\Omega} \mu^2 - \nu^2 dy = 0. \tag{53}$$

The discrete approximation *without* b.c becomes

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix}_t + \begin{pmatrix} Q_f & 0 \\ 0 & -Q_f \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix} = 0 \tag{54}$$

where Q_f uses the matrix elements (16). This means that an arbitrary point, B , on the boundary is updated according to

$$\begin{aligned} P_{BB}\frac{d}{dt}\mu_B + \mu_B \Delta y_B + \sum_i \frac{\mu_B + \mu_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}\nu_B - \nu_B \Delta y_B - \sum_i \frac{\nu_B + \nu_{N_i}}{2}\Delta y_i = 0. \end{aligned} \tag{55}$$

Again we have to find a way to calculate the boundary flux. One way is to divide the boundary in a left part Γ_1 and a right part Γ_2 and update the boundary according to

$$\begin{aligned} P_{BB}\frac{d}{dt}\mu_B \overbrace{+\nu_B \Delta y_B}^{1} + \sum_i \frac{\mu_B + \mu_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}\nu_B \underbrace{-\nu_B \Delta y_B}_{2} - \sum_i \frac{\nu_B + \nu_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{56}$$

on Γ_1 and

$$\begin{aligned} P_{BB}\frac{d}{dt}\mu_B \overbrace{+\mu_B \Delta y_B}^{3} + \sum_i \frac{\mu_B + \mu_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}\nu_B \underbrace{-\mu_B \Delta y_B}_{4} - \sum_i \frac{\nu_B + \nu_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{57}$$

on Γ_2. Since μ propagates in the $\hat{x}$–direction and ν in the $-\hat{x}$–direction this means that we prescribe the ingoing characteristic variable. Putting the expressions 1, 2, 3 and 4 in the b–vector at the corresponding points and using the matrix elements (18) gives the discrete energy rate

$$\begin{aligned} \frac{d}{dt}(||\boldsymbol{\mu}||_P^2 + ||\boldsymbol{\nu}||_P^2) = +\sum_{\partial\Omega}(\mu_i^2 - \nu_i^2)\Delta y_i - \\ - 2\sum_{\Gamma_1}(\mu_i\nu_i - \nu_i^2)\Delta y_i - 2\sum_{\Gamma_2}(\mu_i^2 - \nu_i\mu_i)\Delta y_i = \\ = -\sum_{\Gamma_1}(\mu_i - \nu_i)^2|\Delta y_i| - \sum_{\Gamma_2}(\mu_i - \nu_i)^2|\Delta y_i|. \end{aligned} \tag{58}$$

We see that this method gives small dissipative terms. The b.c could also be handled by updating all boundary points according to

$$\begin{aligned} P_{BB}\frac{d}{dt}\mu_B + \frac{1}{2}(\mu_B + \nu_B)\Delta y_B + \sum_i \frac{\mu_B + \mu_{N_i}}{2}\Delta y_i = 0 \\ P_{BB}\frac{d}{dt}\nu_B - \frac{1}{2}(\mu_B + \nu_B)\Delta y_B - \sum_i \frac{\nu_B + \nu_{N_i}}{2}\Delta y_i = 0 \end{aligned} \tag{59}$$

Proceeding as before this gives

$$\begin{aligned}\frac{d}{dt}(||\boldsymbol{\mu}||_P^2 + ||\boldsymbol{\nu}||_P^2) = +\sum_{\partial\Omega}\mu_i^2\Delta y_i - \sum_{\partial\Omega}\nu_i^2\Delta y_i - \\ -2\sum_{\partial\Omega}\frac{1}{2}\left(\mu_i\nu_i + \mu_i^2\right)\Delta y_i + 2\sum_{\partial\Omega}\frac{1}{2}\left(\nu_i^2 + \mu_i\nu_i\right)\Delta y_i = 0\end{aligned} \tag{60}$$

Now the discrete energy corresponds exactly to the continuous one.

The two methods above can be viewed as penalty procedures. To see this write equation (59) as

$$\begin{aligned}&P_{BB}\frac{d}{dt}\mu_B + \overbrace{\frac{1}{2}\left(\mu_B + \nu_B\right)\Delta y_B - \mu_B\Delta y_B}^{1} + \mu_B\Delta y_B + \\ &+\sum_i \frac{\mu_B + \mu_{N_i}}{2}\Delta y_i = 0 \\ &P_{BB}\frac{d}{dt}\nu_B \underbrace{-\frac{1}{2}\left(\mu_B + \nu_B\right)\Delta y_B + \nu_B\Delta y_B}_{2} - \nu_B\Delta y_B - \\ &-\sum_i \frac{\nu_B + \nu_{N_i}}{2}\Delta y_i = 0\end{aligned} \tag{61}$$

If we now put the expressions 1 and 2 in the right hand side vector, b, this means that we are again using the matrix elements (16). This procedure corresponds to the usual way of introducing b.c by penalty. The scheme now looks like

$$\begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\nu}\end{pmatrix}_t + \begin{pmatrix} Q_f & 0 \\ 0 & -Q_f \end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\nu}\end{pmatrix} + b = 0 \tag{62}$$

where

$$b = \begin{cases} \frac{1}{2}(\nu_B - \mu_B) & \text{at boundary points} \\ 0 & \text{otherwise} \end{cases}. \tag{63}$$

The boundary method in equations (56) and (57) can be rewritten in the same way resulting in

$$b = \begin{cases} \nu_B - \mu_B & \text{for } \mu\text{–points on } \Gamma_1 \text{ and } \nu\text{–points on } \Gamma_2 \\ 0 & \text{otherwise} \end{cases}. \tag{64}$$

To use the method of injection, take the μ–points on Γ_1 and the ν–points on Γ_2 out of the system and inject ν on Γ_1 and μ on Γ_2. This gives

$$\begin{aligned}&\frac{d}{dt}(||\boldsymbol{\mu}'||_{P'}^2 + ||\boldsymbol{\nu}'||_{P'}^2) = \\ &-\begin{pmatrix}\boldsymbol{\mu}'^T & \boldsymbol{\nu}'^T\end{pmatrix}\begin{pmatrix} Q_f' & 0 \\ 0 & -Q_f' \end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}'\\ \boldsymbol{\nu}'\end{pmatrix} + 2\begin{pmatrix}\boldsymbol{\mu}'^T & \boldsymbol{\nu}'^T\end{pmatrix} b = \\ &= -\sum_{\Gamma_2}\mu_i^2\Delta y_i + \sum_{\Gamma_1}\nu_i^2\Delta y_i - \sum_{\tilde{\Gamma_2}}\mu_i\nu_{i'}\Delta y_i + \sum_{\tilde{\Gamma_1}}\nu_i\mu_{i'}\Delta y_i\end{aligned} \tag{65}$$

where $\tilde{\Gamma_1}$ and $\tilde{\Gamma_2}$ are the sets of points that are neigbours to the points in Γ_1 and Γ_2 respectively and i' is the index of the node on the boundary that is a neigbour

Figure 7. An unstructured mesh with 23 nodes.

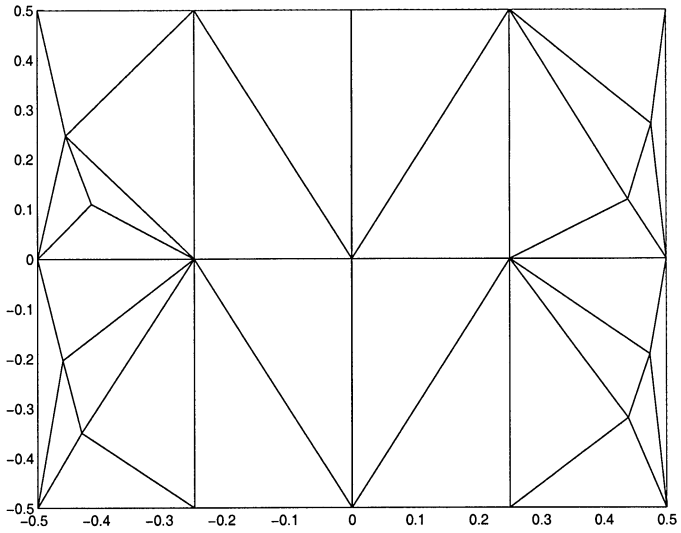
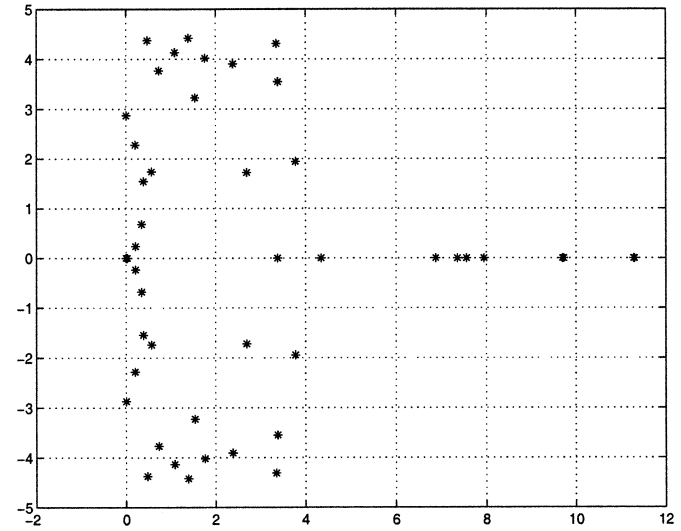


Figure 8. The spectrum of the method defined by equations (56) and (57), on an unstructured mesh with 23 nodes. min(Re(λ_i))=0



to the node with index i (if such a node exists). μ', ν', P' and Q'_f are the vectors and matrices where the points mentioned above are removed.

The energy rate contains indefinite cross-terms which means that we cannot show strict stability [4]. The spectrum using injection is given in figures 9 and 12. We see that on one grid the method of injection gives some eigenvalues with negative real part and thus an unstable scheme. The mesh with 23 nodes is shown in figure 7. The mesh with 169 nodes is of the same type as the mesh in figure 14. In figures 8 and 11 the spectrum of the method defined by the equations (56) and (57) is shown. The spectrum for the method (59) is shown in figures 10 and 13. The last two spectra are purely imaginary because the energy rate is zero.

3.2.3 Convergence

It can be shown that the spectrum of the continuous problem (51), (52) consists of the points $0, \pm\pi i, \pm 2\pi i, \ldots$. The discrete spectra in section 3.2.2 should converge to the continuous spectrum when the number of nodes in the mesh increases. To investigate this, the unit square was discretized as in figures 14 and 15 and the smallest distance to the points πi and $2\pi i$ was plotted against the number of

Figure 9. The spectrum of $P^{-1}Q$ with the method of injection, on an unstructured mesh with 23 nodes. min(Re(λ_i))=-0.105

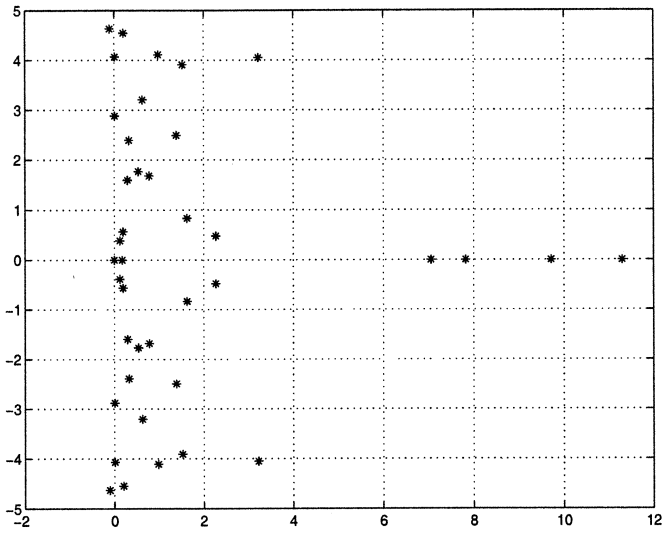


Figure 10. The spectrum of $P^{-1}Q$ with the method defined by equation (59), on an unstructured grid with 23 nodes. min(Re(λ_i))=0

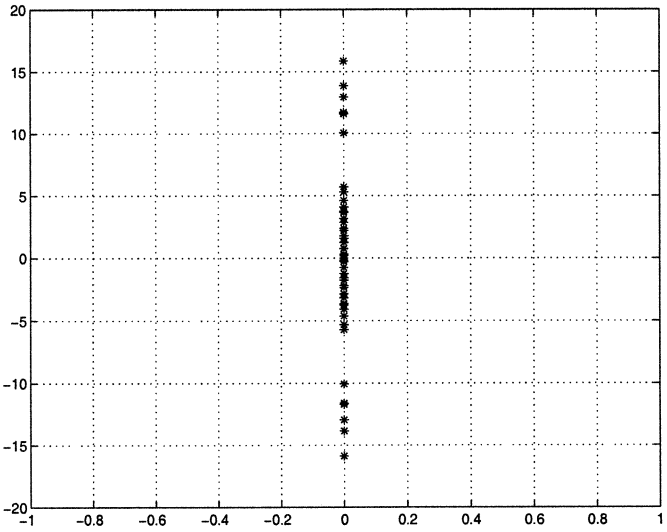


Figure 11. The eigenspectrum of $P^{-1}Q$ with the method defined by equations (56) and (57) , computed on an unstructured mesh with 169 nodes. min(Re(λ_i))=0

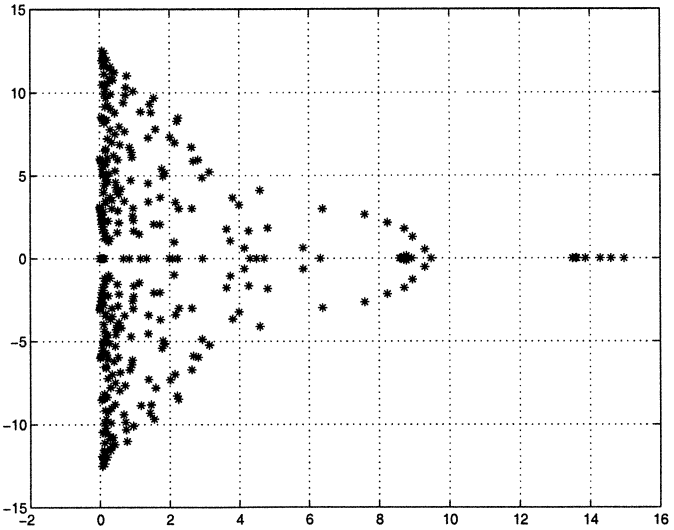


Figure 12. The spectrum of the method of injection, on an unstructured mesh with 169 nodes. $\min(\mathrm{Re}(\lambda_i))=0$

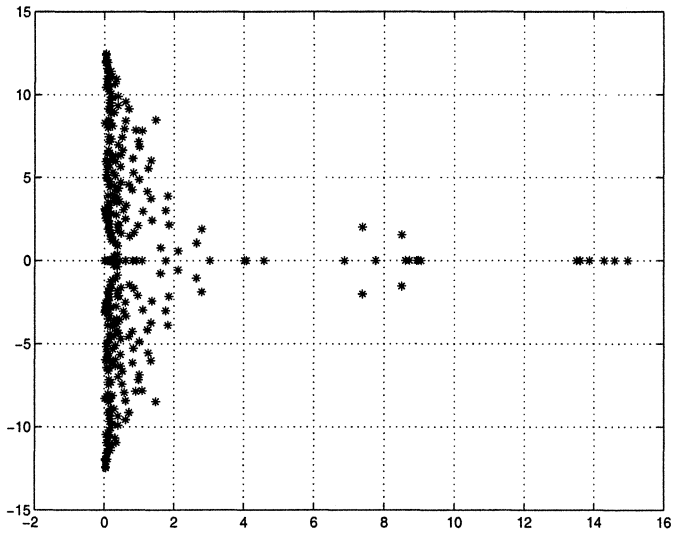


Figure 13. The spectrum of the method defined by equation (59), on an unstructured mesh with 169 nodes. $\min(\mathrm{Re}(\lambda_i))=0$

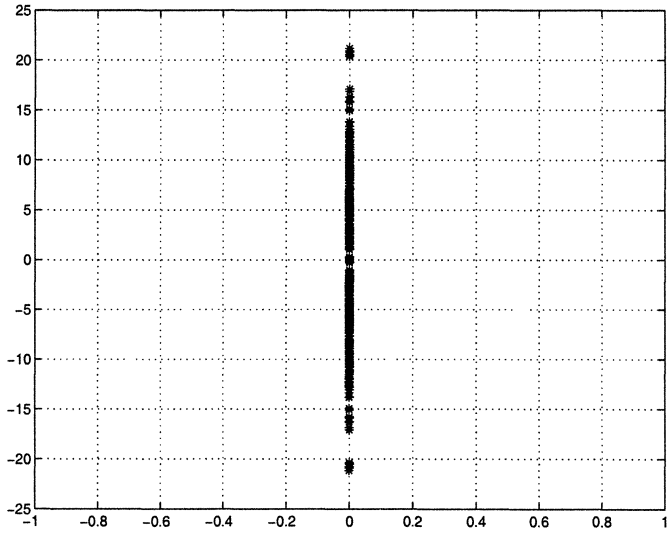


Figure 14. The mesh, more points in x–direction than in y–direction.

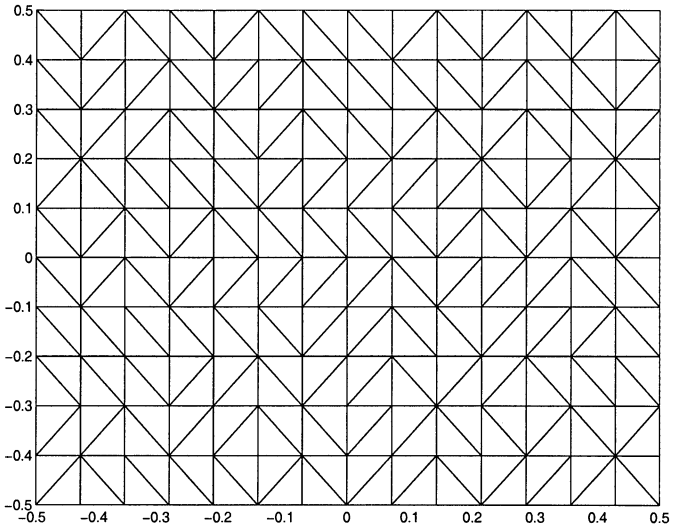
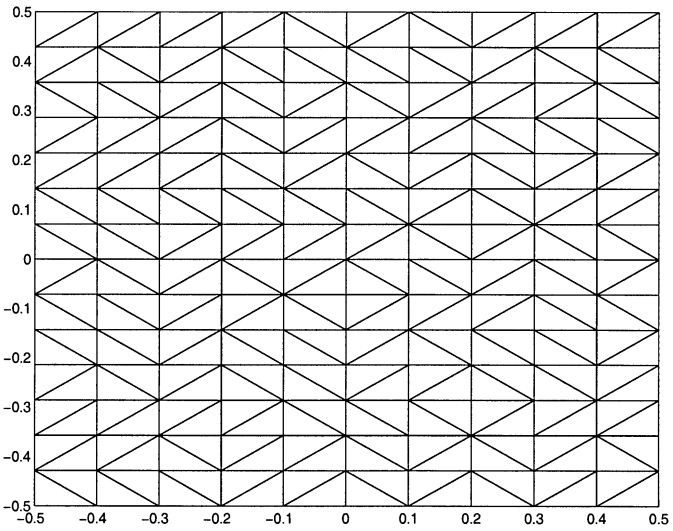


Figure 15. The mesh, more points in y–direction than in x–direction.



nodes, see figures 16-19. (The point 0 was omitted because it is contained in all spectra). Two cases were tested; the convergence when the new nodes were introduced in the x–direction only, and the convergence when new nodes were introduced in the y–direction only. The plots 16-19 show that we have a second order accurate scheme. The refinement in the y–direction does not lead to more accurate solutions since the problem (33) contains x–derivatives only. (In figures 16 and 17 two points deviate a lot from the others. The reason for that deviation is the degenerated meshes obtained from Matlab.)

3.2.4 The most general case

Consider the following problem, which contains all previous examples

$$\begin{pmatrix} \mu \\ \nu \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \nu \end{pmatrix}_x = 0, \quad (x,y) \in \Omega \subset \mathbb{R}^2$$
$$\begin{cases} (\mu - \alpha\nu)|_{\Gamma_1} = g \\ (\nu - \beta\mu)|_{\Gamma_2} = f \end{cases} \tag{66}$$

Figure 16. Convergence against the point πi. New nodes are introduced in the x–direction only. The dashed line is a reference line with slope -2.

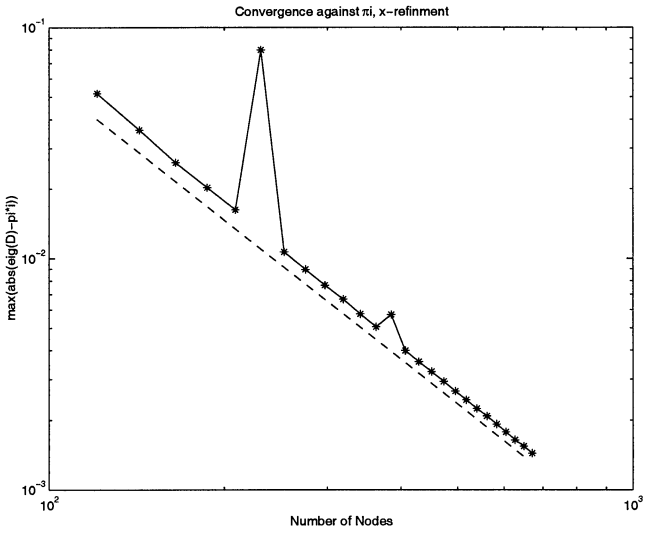


Figure 17. Convergence against the point $2\pi i$. New nodes are introduced in the x–direction only. The dased line is a reference line with slope -2.

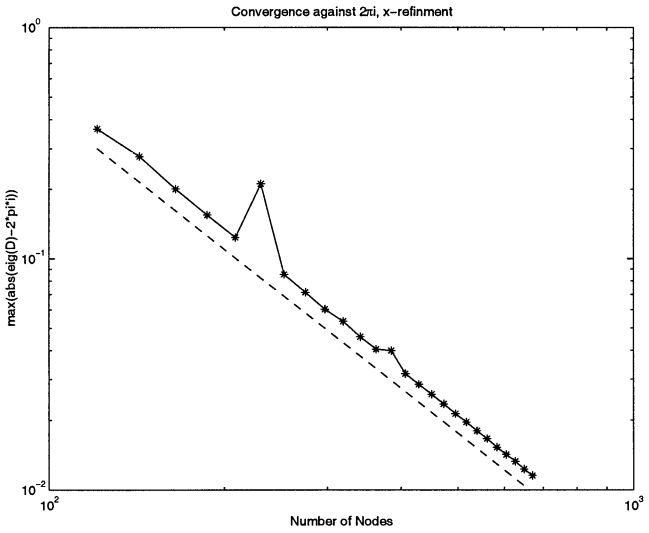
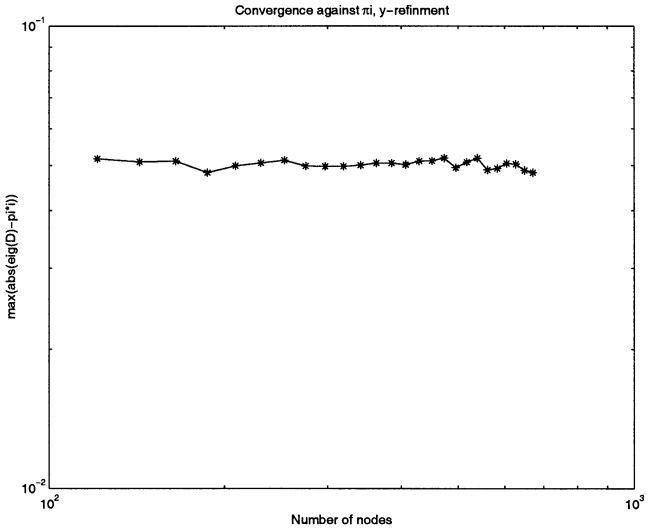


Figure 18. Convergence against the point πi. New nodes are introduced in the y–direction only.



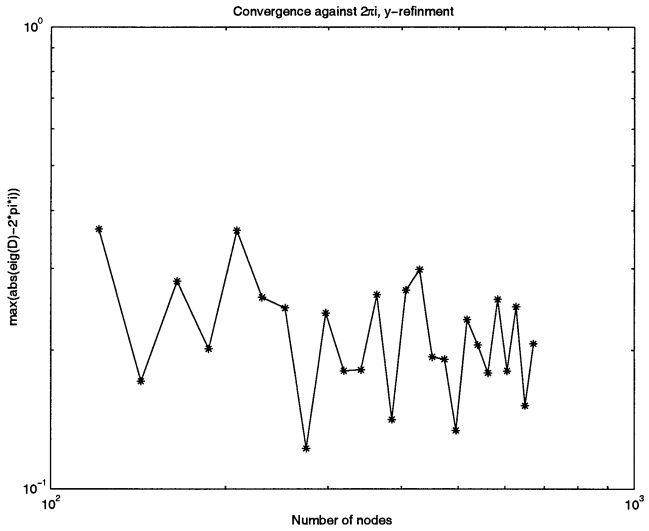


Figure 19. Convergence against the point $2\pi i$. New nodes are introduced in the y–direction only.

where α and β are fixed parameters. The energy equation (3) becomes

$$\begin{aligned}
&\iint_{\Omega} \mu^2 + \nu^2 dxdy = -\oint_{\partial\Omega} \mu^2 - \nu^2 dy = \\
&= + \int_{\Gamma_1} g^2 + (\alpha^2 - 1)\left(\left(\nu - \frac{\alpha}{\alpha^2-1} g\right)^2 - \frac{\alpha^2}{(\alpha^2-1)^2} g^2\right)|dy| - \\
&- \int_{\Gamma_2} -f^2 + (1-\beta^2)\left(\left(u - \frac{\beta}{1-\beta^2} f\right)^2 - \frac{\beta^2}{(1-\beta^2)^2} f^2\right)|dy|
\end{aligned} \tag{67}$$

where the definitions of Γ_1 and Γ_2 have again been used ($|dy|$ denotes the absolute value of the differential dy). This gives an energy estimate if $|\alpha| \leq 1$ and $|\beta| \leq 1$. (In fact, by changing the norm one can show that an energy estimate exists whenever $|\alpha\beta| \leq 1$, see [1]).

On Γ_1 we update μ by $g + \alpha\nu$ and on Γ_2 we update ν by $f + \beta\mu$. This gives the following discrete energy rate:

$$\begin{aligned}
&\frac{d}{dt}(||\boldsymbol{\mu}||_P^2 + ||\boldsymbol{\nu}||_P^2) = \\
&= \sum_{\Gamma_1} \Bigg\{ -(\mu_i - (g_i + \alpha\nu_i))^2 + \\
&\quad + g_i^2 + (\alpha^2 - 1)\left(\left(\nu_i + \frac{\alpha}{\alpha^2-1} g_i\right)^2 - \frac{\alpha^2}{(\alpha^2-1)^2} g_i^2\right)\Bigg\} |\Delta y_i| - \\
&- \sum_{\Gamma_2} \Bigg\{ + (\nu_i - (f_i + \beta\mu_i))^2 - \\
&\quad - f_i^2 - (1-\beta^2)\left(\left(\mu_i + \frac{\beta}{1+\beta^2} f_i\right)^2 - \frac{\beta^2}{(1+\beta^2)^2} f_i^2\right)\Bigg\} |\Delta y_i|
\end{aligned} \tag{68}$$

which is the continuous energy rate plus a negative definite term, i.e the approximation is strictly stable in the sense defined in [4].

If we set $\alpha = \beta = 1$ and $f = g = 0$ we have the method defined by equations (56) and (57) again.

4 Conclusions

It has been shown that it is possible to generalize the concept of SBP–operators to two space dimensions and FVM on unstructured grids.

The method analyzed in this work with boundary conditions imposed weakly lead to energy estimates and strict stability if one specifies the ingoing characteristic variable. Furthermore the method of injection of boundary conditions does not always lead to energy estimates and sometimes results in an unstable scheme.

To introduce boundary conditions weakly is shown to be equivalent to the standard penalty procedure(SAT) for boundary conditions used together with SBP–operators.

References

[1] Mark H. Carpenter, Jan Nordström, and David Gottlieb. A stable and conservative interface treatment of arbitrary spatial accuracy. *Journal of Computational Physics*, pages 148:341–365, 1999.

[2] M.H. Carpenter, D. Gottlieb, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems:methodology and applications to high-order compact schemes. *Journal of Computational Physics*, pages 111:220–236, 1994.

[3] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Method.* John Wiley & Sons, 1995.

[4] Jan Nordström and Martin Björck. Finite volume approximations and strict stability for hyperbolic problems. *Applied Numerical Mathematics*. To appear.

Issuing organisation	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Division of Aeronautics, FFA SE-172 90 STOCKHOLM	FOI-R–0121–SE	Scientific Report
	Month year May 2001	Project number E840250
	Customers code 3. Aeronautical Research	
	Research area code 7. Vehicles	
	Sub area code 73. Aeronautical Research	

Author(s)	Project manager
Carl Adamsson, Karl Forsberg and Jan Nordström	Jan Nordström
	Approved by Torsten Berglind Head, Computational Aerodynamics Department
	Scientifically and technically responsible Jan Nordström

Report title

Finite volume methods, unstructured meshes and strict stability

Abstract

The unstructured node centered finite volume method is analyzed and it is shown that it can be interpreted in the framework of summation by parts operators. It is also shown that introducing boundary conditions weakly produces strictly stable formulations.

Keywords

finite volume method, summation by part, strict stability, unstructured

Further bibliographic information

ISSN	Pages	Language
ISSN 1650-1942	37	English
	Price Acc. to price list	
	Security classification Unclassified	

Utgivare	Rapportnummer, ISRN	Klassificering
Totalförsvarets Forskningsinstitut – FOI Avdelningen för Flygteknik, FFA SE-172 90 STOCKHOLM	FOI-R–0121–SE	Vetenskaplig Rapport
	Månad år Maj 2001	Projektnummer E840250
	Verksamhetsgren 3. Flygteknisk forskning	
	Forskningsområde 7. Bemannade och obemannade farkoster	
	Delområde 73. Flygteknisk forskning	
Författare Carl Adamsson, Karl Forsberg och Jan Norström	Projektledare Jan Nordström	
	Godkänd av Torsten Berglind Chef, Institutionen för Beräkningsaerodynamik	
	Tekniskt och/eller vetenskapligt ansvarig Jan Nordström	

Rapporttitel

Finita volymmetoder, strikt stabilitet och ostrukturerade nät

Sammanfattning

Den ostrukturerade, nod–centrerade finita volymmetoden analyseras och vi visar att den kan behandlas med hjälp av summation by parts–operatorer. Vi visar också att strikt stabila formuleringar erhålls genom att introducera randvillkor svagt.

Nyckelord

finit volymmetod, summation by part, strikt stabilitet, ostrukturerad

Övriga bibliografiska uppgifter

ISSN	Antal sidor	Språk
ISSN 1650-1942	37	Engelska
Distribution enligt missiv	Pris Enligt prislista	
	Sekretess Öppen	