A Stable and Accurate Summation-by-Parts Finite Volume Formulation of the Laplacian Operator

Magnus Svärd and Jan Nordström

24th January 2003

Abstract

Our objective is to analyse a commonly used edge based finite volume approximation of the Laplacian and construct an accurate and stable way to implement boundary conditions. Of particular interest are general unstructured grids where the strength of the finite volume method is fully utilised.

As a model problem we consider the heat equation. We analyse the Cauchy problem in one and several space dimensions and we prove stability on unstructured grids. Next, the initial-boundary value problem is considered and a scheme is constructed in a summation-by-parts framework. The boundary conditions are imposed in a stable and accurate manner, using a penalty formulation.

Numerical computations of the wave equation in two-dimensions are performed, verifying stability and order of accuracy for structured grids. However, the results are not satisfying for unstructured grids. Further investigation reveals that the approximation is not consistent for general unstructured grids. However, grids consisting of equilateral polygons recover the convergence.

1 Introduction

Edge based finite volume (FV) approximations are widely used in computational fluid dynamics.(see [1, 2, 3, 4, 5, 6, 7, 8]) The main reason is that the algorithm works equally well on structured as well as unstructured grids. The algorithm only needs information about what nodes connect to each other. This property is called grid transparency by Haselbacher et al in [1] and is essential for the efficiency and applicability of the scheme.

Boundary conditions can be implemented strongly or weakly and often both of these types are used at the same boundary for different terms in the equation. Further, the choice of boundary conditions at one boundary may differ from the choice at other boundaries. This leads to a scheme which is hard to analyse in terms of stability and consistency and there are certainly cases where these notions are not met. If stability and accuracy are not fulfilled, we can not use the Lax-Richtmyer's equivalence theorem (see [9]) and thus by no means claim that the solution is an approximation of the true one. The objective of this work is to provide a systematic treatment of the boundary conditions for the Laplacian such that small modifications of already existing algorithms yield stability. The significance of the formulation of the boundary conditions in practical calculations was already pointed out by Mavriplis in [5] for the Euler equations and the same should apply to the viscous terms.

In [6], Nordström et al considered a first derivative approximation such that convective terms can be implemented in a stable and accurate manner. A Laplacian term is required in the approximation of the incompressible Navier-Stokes equations. In the compressible case, the Laplacian enters when the equations are simplified using a specific thin layer approximation where the cross derivative terms are neglected. (c.f. articles by Mavriplis et al. [2, 3, 4]) and also by Haselbacher et al [1])

We will develop a boundary treatment for the Laplacian which is compatible with the boundary treatment developed in [6]. The idea is to construct the finite volume approximation of the Laplacian as a summation-by-parts operator (SBP). The theory was developed in [10] and [11]. This makes it possible to treat the boundary using a Simultaneous Approximation Term technique (SAT), developed by Carpenter et al in [12, 13] and Nordström et al in [14, 15] for high order methods on structured grids. These two notions yield an energy estimate that leads to stability.

The contents of this report are divided as follows; in section 2 the general finite volume approximation is derived; section 3 shows the SBP and SAT technique to make clear what we aim for; in section 4 a one-dimensional SBP-FV operator is derived; a generalisation to multidimensional space is done in section 5. Section 6 shows some numerical computations using the new SBP operators and in section 7 conclusions are drawn.

2 The Finite Volume Approximation

The simplest time dependent equation containing the Laplacian is the heat equation,

$$u_t = \Delta u \tag{1}$$

which we will use as our model problem when constructing the discrete operator. The approximation derived below is used in [2, 3, 4] and also discussed in [1]. Integration of (1) over the domain D yields,

$$\int_D u_t dv = \int_{\partial D} \frac{\partial u}{\partial N} ds, \tag{2}$$

where Gauss' theorem is used. N denotes the outward pointing unit normal vector such that $\frac{\partial u}{\partial N} = u_N = grad(u) \cdot N$. Let D be a volume V_i. Further, let V_i be an n-sided polygon with sides ds_{in}. At last, let $(u_N)_{in}$ denote the normal outward pointing derivative at ds_{in}. A straightforward approximation of (2) would be,

$$V_i \cdot (u_t)_i = \sum_n (u_N)_{in} ds_{in}. \tag{3}$$

In order to approximate $(u_N)_{in}$ we have to define V_i and ds_{in}.

Given any grid, let r_i denote a grid point and V_i be the corresponding volume. V_i is defined as the volume inside the so called dual grid around r_i. The dual grid is in turn defined as the straight lines drawn between the centres of mass of the cells with r_i as a vertex and the midpoints of the edges from r_i, see figure 1.

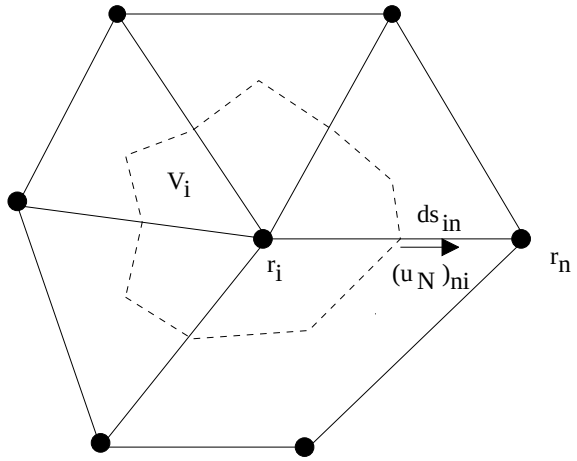

Figure 1: A generic 2D grid. Solid lines are the grid lines and dashed lines corresponds to the dual grid.

Further, ds_{in} is defined as the sum of the length of the "centre of mass-midpoint-centre of mass" lines passing over one edge (see figure 1). Let $r_{ij} = |r_i - r_j|$. Finally, let N_i denote the set of indices of points being neighbours to r_i. Then $(u_N)_{in}$ can be approximated by,

$$(u_N)_{in} = \frac{u_n - u_i}{r_{ni}}, \quad n \in N_i.$$

Summing up, for an interior point, r_i, the scheme becomes,

$$V_i \cdot (u_t)_i = \sum_{n \in N_i} \frac{u_n - u_i}{r_{ni}} ds_{in}. \tag{5}$$

Without boundaries, this scheme could, for all interior points, be summarised on matrix form as,

$$Pu_t = Q_\Delta u, \tag{6}$$

where P is a diagonal matrix with V_i on the diagonal. u_t and u are vectors with components $(u_t)_i$ and u_i respectively. Q_Δ is a matrix that we will return to in forthcoming sections.

3 The Energy Method and Penalty Technique

Consider equation (1). If the energy method is applied to this equation we obtain, using Green's formula (The continuous solution is denoted v to distinguish it from the discrete solution vector u.),

$$\frac{1}{2}(\|v\|^2)_t = \int_D v v_t dv = \int_D v \Delta v dV = -\int_D |\nabla v|^2 dV + \int_{\partial D} v \frac{\partial v}{\partial N} ds. \tag{7}$$

If proper boundary conditions are given this yields that the growth of the norm is bounded, i.e. the problem is well posed. Let $v_N + \alpha v = g(t, x, y)$ on the boundary. By adding,

$$-\int_{\partial D} v \left(\frac{\partial v}{\partial N} + \alpha v - g(t) \right) ds = 0,$$

to the right hand side of (7), we obtain,

$$\frac{1}{2}(\|v\|^2)_t = -\int_D |\nabla v|^2 dV + \int_{\partial D} (v g(t) - \alpha v^2) ds,$$

implying well posedness for $g(t, x, y) = 0$ and $\alpha \geq 0$.

Remark The requirement $g(t) = 0$ can be relaxed by transforming the problem into one with zero boundary data and non-zero forcing function. Or, one can use the estimate $vg - \alpha v^2 \leq g^2/\alpha$ for $\alpha > 0$ which leads directly to well posedness.

If it is possible to mimic the above derivation for the discrete formulation it will lead to stability. The discrete scheme (6) should fulfill the following analogous derivation,

$$(\|u\|_P^2)_t = 2u^T P u_t = u^T (Q_\Delta + Q_\Delta^T) u. \tag{9}$$

P is a positive semidefinite matrix such that the first equality sign defines a discrete norm. This norm is the same as in [6] which is a necessary condition if we are to obtain energy estimates when first order terms are added to the equation. The norm, P, is defined as the diagonal matrix with V_i on the diagonal, thus being positive definite.

The right hand side should, in order to mimic (7), consist of a negative definite part (or at least semidefinite), and a term similar to $\int_{\partial D} u u_N$. In [13] the following form was suggested,

$$P u_t = Q_\Delta u = (A + DS)u, \tag{10}$$

and we get,

$$(\|u\|_P^2)_t = u^T (A + DS + A^T + (DS)^T) u = u^T (A + A^T) u + 2u^T DSu. \tag{11}$$

A has to be negative semidefinite (Here, definiteness refers to the symmetric part of the matrix.), S some approximation of the outward normal derivative and D a matrix that picks the boundary points.

Since the interior accuracy is second order the boundary condition must be implemented at least to that order, i.e. S must be a second order approximation of the normal derivative. Finally, $P^{-1} Q_\Delta u$ should approximate Δu at every point, including the boundary. However, the order of accuracy is allowed to be one order less at a finite number of points (see [16, 17]). To sum up, the conditions are,

$$\begin{aligned}
P &> 0, && (12)\\
A &\leq 0, && (13)\\
Su &= \frac{\partial u}{\partial N} + \mathcal{O}(\Delta r^2), && (14)\\
P^{-1} Q_\Delta u &= \Delta u + \mathcal{O}(\Delta r^2, \Delta r), && (15)\\
Du &= u_{boundary}. && (16)
\end{aligned}$$

With $P_{ii} = V_i$, condition (12) is fulfilled. If the remaining conditions are fulfilled, we can continue the discrete analogous derivation, with $DSu + \alpha Du = G$ and $\alpha \geq 0$ at the boundaries, and add a term $\sigma(DSu + \alpha Du - G)$ to (10), where G denotes a vector with components equal to $g(t, x, y)$ at boundary points and zero elsewhere. Then (11) becomes,

$$(\|u\|_P^2)_t = u^T(A + A^T)u + 2u^T DSu + 2\sigma u^T(DSu + \alpha Du - G). \tag{17}$$

With $\sigma = -1$, the nondefinite boundary term in (11) precisely cancels and we have stability for $G = 0$ also at the boundary points.

Remark Also in the discrete case the requirement $G = 0$ can be relaxed. Either by transforming the problem to one with homogeneous boundary conditions or by estimating the indefinite term directly, see the remark above.

The resulting scheme would be,

$$Pu_t = Q_\Delta u + \sigma(DSu + \alpha u - G). \tag{18}$$

4 One-Dimensional Operator

In one space dimension (1D) we consider the equation $u_t = u_{xx}$. At first we will prove stability for the Cauchy problem before we proceed to consider boundaries. Finally, an alternative derivation of the approximation is done.

4.1 The Initial Value Problem

Consider the equation $u_t = u_{xx}$ on the real line. The dual grid around x_i is simply the interval $[x_{i+1/2}, x_{i-1/2}] = [\frac{x_i + x_{i-1}}{2}, \frac{x_{i+1} + x_i}{2}]$. The scheme is derived as,

$$\int_{x_{i-1/2}}^{x_{i+1/2}} u_t dx = \int_{x_{i-1/2}}^{x_{i+1/2}} u_{xx} dx = (u_x)_{i+1/2} - (u_x)_{i-1/2},$$

yielding,

$$\frac{x_{i+1} - x_{i-1}}{2}(u_t)_i = \frac{u_{i+1} - u_i}{x_{i+1} - x_i} - \frac{u_i - u_{i-1}}{x_i - x_{i-1}} = \tag{19}$$
$$\frac{1}{x_i - x_{i-1}} u_{i-1} - \frac{x_{i+1} - x_{i-1}}{(x_{i+1} - x_i)(x_i - x_{i-1})} u_i + \frac{1}{x_{i+1} - x_i} u_{i+1}.$$

in the interior.

We note that the scheme (19) can be stated on matrix form as,

$$\begin{pmatrix} \ddots & \ddots & \ddots & \ddots & \cdots & & \\ \cdots & 0 & 0 & P_{ii} & 0 & \cdots & \\ & \cdots & 0 & 0 & P_{i+1,i+1} & 0 & \cdots \\ & & \cdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ (u_t)_{i-1} \\ (u_t)_i \\ (u_t)_{i+1} \\ (u_t)_{i+2} \\ \vdots \end{pmatrix} =$$

$$\begin{pmatrix} \ddots & \ddots & \ddots & \ddots & \cdots & & & \\ \cdots & 0 & q_{i,i-1} & q_{ii} & q_{i,i+1} & 0 & & \cdots \\ & \cdots & 0 & q_{i+1,i} & q_{i+1,i+1} & q_{i+1,i+2} & 0 & \cdots \\ & & \cdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots \\ u_{i-1} \\ u_i \\ u_{i+1} \\ u_{i+2} \\ \vdots \end{pmatrix},$$

where,

$$\begin{aligned} P_{ii} &= \frac{x_{i+1} - x_{i-1}}{2}, \\ P_{i+1,i+1} &= \frac{x_{i+2} - x_i}{2}, \end{aligned}$$

$$\begin{aligned} q_{i,i-1} &= \frac{1}{x_i - x_{i-1}}, \\ q_{i,i} &= \frac{-(x_{i+1} - x_{i-1})}{(x_{i+1} - x_i)(x_i - x_{i-1})}, \\ q_{i,i+1} &= \frac{1}{x_{i+1} - x_i}, \\ q_{i+1,i} &= \frac{1}{x_{i+1} - x_i}, \\ q_{i+1,i+1} &= \frac{-(x_{i+2} - x_i)}{(x_{i+2} - x_{i+1})(x_{i+1} - x_i)}, \\ q_{i+1,i+2} &= \frac{1}{x_{i+2} - x_{i+1}}. \end{aligned}$$

This is the $Pu_t = Q_\Delta u$ formulation, though the matrices are infinite. Note that P is positive definite as $x_{i+1} > x_i$. Further, Q_Δ is symmetric. Next, we construct the Gerschgorin circles of Q_Δ (see for example [18]) as,

$$C_i = \{z : |z - q_{ii}| \leq \sum_{j \in \mathbf{Z}(j \neq i)} |q_{ij}|\}, \quad i \in \mathbf{Z}.$$

Consider row i. Then,

$$\sum_{j\in\mathbf{Z}(j\neq i)} |q_{ij}| = |q_{i,i-1}| + |q_{i,i+1}| = q_{i,i-1} + q_{i,i+1} = \tag{21}$$
$$\frac{1}{x_i - x_{i-1}} + \frac{1}{x_{i+1} - x_i} = \frac{(x_{i+1} - x_{i-1})}{(x_{i+1} - x_i)(x_i - x_{i-1})} = -q_{ii}.$$

Thus $C_i = \{z : |z - q_{ii}| \leq |q_{ii}|\}$ where q_{ii} is negative yielding, together with symmetry, that Q_Δ is negative semidefinite.

Remark The last equality of (21) is in fact obvious since if $P^{-1}Q_\Delta$ is a second derivative approximation, the row sums have to be zero.

4.2 The Initial–Boundary Value Problem

At the boundary, x_0, we construct the dual grid as the interval $[x_0, x_{1/2}] = [x_0, \frac{x_1+x_0}{2}]$. The dual grid at the boundary inserted in (19) yields the boundary approximation,

$$\frac{x_1 - x_0}{2}(u_t)_0 = \frac{u_1 - u_0}{x_1 - x_0} - \left(\frac{\partial u}{\partial x}\right)_0. \tag{22}$$

We observe that equation (22) closely resembles (10). In fact, since DS only affects Q_Δ at the boundary we know that $A = Q_\Delta$ in the interior of the domain. Assuming that $-(u_x)_0$ corresponds to DSu we let $\frac{u_1-u_0}{x_1-x_0}$ be a part of Au. Then A would be,

$$\begin{pmatrix} -\frac{1}{x_1-x_0} & \frac{1}{x_1-x_0} & 0 & 0 & \dots \\ \frac{1}{x_1-x_0} & \frac{-(x_2-x_0)}{(x_2-x_1)(x_1-x_0)} & \frac{1}{x_2-x_1} & 0 & \dots \\ & \ddots & \ddots & \ddots & \end{pmatrix}$$

Denoting by $a_{ij}, i = 0...\infty, j = 0...\infty$ the $i : th, j : th$ element of A. The Gerschgorin circle $C_0 = \{z : |z - a_{00}| \leq |a_{00}|\}$ lies in the left complex half plane. Also, note that $a_{01} = a_{10}$ such that A is symmetric. Thus, A is negative semidefinite.

Next, $(u_x)_0$ has to be chosen such that it is a second order approximation and the complete operator $P^{-1}Q_\Delta$ is at least a first order accurate second derivative at x_0.

By making the ansatz,

$$au_0 + bu_1 + cu_2 = u_x|_0 + \mathcal{O}(\Delta x^2), \tag{23}$$

we get by Taylor expansions:

$$a = \frac{(x_2 - x_0)^2 - (x_1 - x_0)^2}{(x_1 - x_2)(x_1 - x_0)(x_2 - x_0)} \tag{24}$$

$$b = \frac{-(x_2 - x_0)}{(x_1 - x_2)(x_1 - x_0)} \tag{25}$$

$$c = \frac{x_1 - x_0}{(x_1 - x_2)(x_2 - x_0)}. \tag{26}$$

It is easy to see that in the equidistant case this is the usual one-sided second order 3 point scheme. Next, it is straightforward to show with Taylor expansions that,

$$(u_t)_0 = \frac{2}{x_1 - x_0}\left(\frac{u_1 - u_0}{x_1 - x_0} - (au_0 + bu_1 + cu_2)\right) = (u_{xx})_0 + \mathcal{O}(\Delta x).$$

Concluding that in the 1D case it is possible to construct a finite volume approximation on the form (10) with the SBP property. We can summarise the result in the following proposition.

Proposition 4.1 *The discretisation (5) of (1) in the interior of the domain together with (22), (23) and (24)-(26) at the boundary, and the use of the penalty method for imposing boundary conditions lead to a stable and second order accurate finite volume discretision of the heat equation in one space dimension.*

4.3 An Auxiliar Observation

The derivation of the 1D-SBP operator in the previous section does not lend itself immediately to a generalisation in several space dimensions (multiD) on unstructured grids. The reason is that the Taylor expansions made to construct the first derivative are not easily done on a generic unstructured grid. Hence, another approach has to be adopted.

Consider the finite volume approximation at x_0 in equation (22). Now, define a new dual grid for x_0 as $[x_0, x_{3/2})$, which yields a new finite volume approximation:

$$\frac{x_2 + x_1 - 2x_0}{2}(u_t)_0 = \frac{u_2 - u_1}{x_2 - x_1} - \left(\frac{\partial u}{\partial x}\right)_0. \tag{28}$$

The key to construct the multiD-SBP is the following observation. Consider (22) and (28) as a system of equations in the unknowns $(u_x)_0$ and $(u_t)_0$ and

solve for those. The solution is given by (23) and (24)-(26). This gives us a hint on how to construct the multidimensional counterpart. The idea is to make a new dual grid as a sum of the dual grids at the boundary point and one step away from the boundary and solve the system of equations.

5 Multidimensional Operator

We will first show that Q_Δ as given in section 2 is a symmetric negative semidefinite matrix for the initial value problem. Next, we construct the boundary approximation such that stability is achieved for the initial–boundary value problem. Finally, we prove that the desired accuracy of the approximation is obtained.

5.1 Stability of the Cauchy Problem

Note that although the figures are 2D for clarity, the forthcoming derivations are not restricted to a specific number of space dimensions. In the original edge based finite volume formulation, one only has to keep track of the neighbours of each point. The locations of those, being in the plane or in 3D space, do not matter.

Consider the Cauchy problem in a multidimensional space. The scheme is given by equations (5) and (6). As stated in section 2, P is positive definite. The remaining issue is the structure of Q_Δ. Since the grid might be unstructured it is not possible to give Q_Δ explicitly, but properties are still obtainable. Consider the right hand side of equation (5) and recollect that N_i is the set of indices of neighbours to r_i. Then we have,

$$\sum_{n\in N_i} \frac{u_n - u_i}{r_{ni}} ds_{in} = \sum_{n\in N_i} \frac{ds_{in}}{r_{ni}} u_n - u_i \sum_{n\in N_i} \frac{ds_{in}}{r_{ni}} = \sum_{n\in N_i} q_{in} u_n + q_{ii} u_i, \qquad (29)$$

implying that,

$$q_{ii} = -\sum_{n\in N_i} \frac{ds_{in}}{r_{ni}}, \qquad (30)$$

$$q_{in} = \frac{ds_{in}}{r_{ni}}, \quad \text{if } n \in N_i. \qquad (31)$$

Define $q_{ij} = 0$ whenever $j \notin N_i$. Further, let q_{ij} be the $i:th, j:th$ component of Q_Δ. Since we now consider a Cauchy problem, equation (29) applies to all $i \in \mathbf{Z}$. Thus equation (29) defines $Q_\Delta u$, for all i.

Proposition 5.1 *The Cauchy problem, (1) discretised using (5) is stable.*

Proof Since $r_{ij} = r_{ji}$ and $ds_{ij} = ds_{ji}$ we will have using (31) that $q_{ij} = q_{ji}$ for all i, j, concluding that Q_Δ is symmetric. In equation (31) it is seen that $q_{in} > 0$ if $n \in N_i$ (or else $q_{in} = 0$) since $r_{in} > 0$ and $ds_{in} > 0$. By the same reason equation (30) yields that $q_{ii} < 0$. Consider the i:th row in Q_Δ, then

$$\sum_{j \in \mathbf{Z}, j \neq i} q_{ij} = \sum_{n \in N_i} q_{in} = -q_{ii}, \tag{32}$$

the last equality is due to equations (30) and (31). Again, we construct the Gerschgorin circles for Q_Δ as,

$$C_i = \{z : |z - q_{ii}| \leq \sum_{j \in \mathbf{Z}, j \neq i} |q_{ij}|\}.$$

Using (32), $q_{ij} = |q_{ij}|$ for $i \neq j$ and $-q_{ii} = |q_{ii}|$ yields,

$$C_i = \{z : |z - q_{ii}| \leq |q_{ii}|\}.$$

Since $q_{ii} \leq 0$ we obtain that C_i is in the left half plane for all i, i.e. that all the eigenvalues of Q_Δ are negative or zero. Since Q_Δ is symmetric we can conclude that it is negative semidefinite. The fact that $P = P^T > 0$ concludes the proof. ∎

Remark The equality (32) is obvious since $P^{-1}Q_\Delta$ is an approximation of the Laplacian which implies that the row sums are equal to zero.

5.2 Stability of the Initial-Boundary Value Problem

Consider a boundary point r_b. The dual grid around r_b is defined similarly as for interior points with one exception. The side facing out of the domain is denoted ds_{bb} and goes from the midpoints of r_b and its boundary neighbours, through r_b, i.e. it is a strip of the domain boundary (see figure 2).

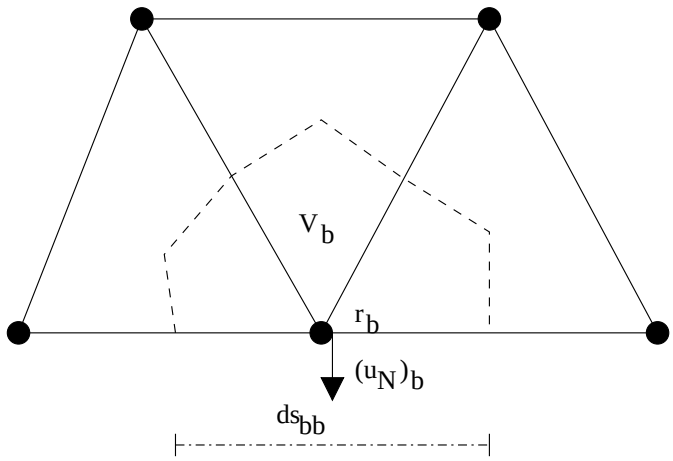

Figure 2: A generic and the corresponding dual grid at the boundaries.

Equation (5) at that point is,

$$V_b \cdot (u_t)_b = \sum_{n \in N_b} \frac{u_n - u_b}{r_{nb}} ds_{bn} + (u_N)_b ds_{bb}. \tag{35}$$

$(u_N)_b$ denotes the outward facing normal derivative to ds_{bb} (see figure 2). Again, note the similarity of (10) and (35). Suppose that $(u_N)_b$ corresponds to DSu. Rewrite (35) as,

$$\begin{gathered} \sum_{n \in N_b} \frac{u_n - u_b}{r_{nb}} ds_{bn} + (u_N)_b ds_{bb} = \\ \sum_{n \in N_b} \frac{ds_{bn}}{r_{nb}} u_n - u_b \sum_{n \in N_b} \frac{ds_{bn}}{r_{nb}} + (u_N)_b ds_{bb} = \\ (u_N)_b ds_{bb} + \sum_{n \in N_b} a_{bn} u_n + u_b a_{bb}. \end{gathered} \tag{36}$$

Equation (36) gives a definition of A at positions representing neighbours and the boundary point itself. We extend the definition with $a_{bj} = 0$ for $j \notin N_b \cup \{b\}$ and $b \in B$, where B denotes the set of indices of boundary points. The interior scheme is not altered allowing us to define $a_{ij} = q_{ij}$ where q_{ij} is defined in the previous subsection for $i \notin B$. With A defined for all points in the domain, using (36) and (31), we see that $a_{bn} = a_{nb}$. Thus A is symmetric.

Since ds_{bn} and r_{nb} are positive for all $n \in N_b$, a_{bn} is positive for those n. Denote by C the set of all point indices. We will now consider the off

diagonal row sum of A for row b,

$$\sum_{i \in C \setminus \{b\}} a_{bi} = \sum_{n \in N_b} a_{bn} = -a_{bb},$$

by (36). Also, $a_{bb} < 0$. The Gerschgorin circle,

$$C_b = \{z : |z - a_{bb}| \leq \sum_{i \in C \setminus \{b\}} |a_{bi}|\} = \{z : |z - a_{bb}| \leq |a_{bb}|\},$$

lies in the left half plane. The Gerschgorin circles of the inner points precisely equal the Cauchy case. Thus all circles of A lie in the left half plane yielding that the eigenvalues of A are all negative or zero. Together with symmetry of A we conclude that A is negative semidefinite. This is summarised in the following proposition.

Proposition 5.2 *The discretisation (5) of (1) in the interior and (35) at the boundary, together with the penalty method, is a stable approximation of the heat equation.*

Proof The preceeding derivation and the penalty technique of section 3, i.e. equations (9)-(11),(17) and (18) prove the proposition.

Now we turn to the construction of $(u_N)_b$. In subsection 4.2 we derived the difference approximation at the boundary using Taylor expansions. On a general grid this approach is not realistic. Instead we use the observation made in subsection 4.3 to construct a new dual grid around r_b. To do so, we need to define some sets before proceeding.

Let B_b denote the set of indices of points that are neighbours of r_b and being boundary points, i.e. $B_b = \{r_i : i \in N_b \cap B\}$. Let M_b be the set of indices of neighbours not being boundary points, $M_b = N_b \setminus B_b$. Further, define a set $I_b = M_b \cup \{b\}$. Finally, we define for $i \in I_b$, $L_{bi} = N_i \setminus I_b$. (See figure 3 and 4.)

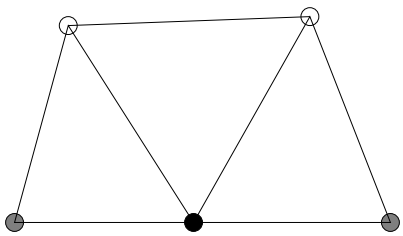

Figure 3: The black point is the index set $\{b\}$. The grey points are B_b and the white are M_b. I_b is the set of white and black points.

In subsection 4.3 we extended the dual grid to include one more point. In that case this could only be done by extension in the one existing direction. Now we might extend the volume in as many directions as dimensions of the space. However, we are interested in second order accuracy in the normal direction. Therefore, we will only extend the volume to interior points and not include more boundary points. Hence, the dual volume V_b' is defined $V_b' = V_b + \sum_{n \in M_b} V_n = \sum_{n \in I_b} V_n$. The new dual grid is then the boundary of V_b'. Also observe that the sets of points L_{bi} all lie outside V_b' and the set I_b lies inside. (See figure 4.)

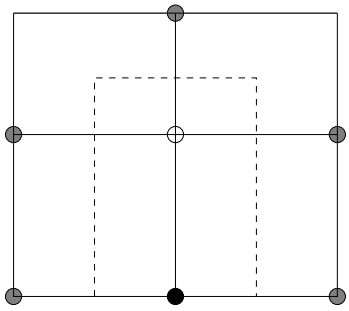

Figure 4: The black point is the index set $\{b\}$. I_b is the set of white and black points. The grey points are the sets $L_{bi}, i \in I_b$. The dashed line is the boundary of V_b'.

The new dual volume, V_b', will, as in 4.3, define a new finite volume approximation at r_b,

$$V_b' \cdot (u_t)_b = \sum_{i \in I_b} \sum_{j \in L_{bi}} \frac{u_j - u_i}{r_{ji}} ds_{ij} + (u_N)_b ds_{bb}. \tag{38}$$

This expression is just an application of the fundamental equation (5). Equation (38) together with (35) constitute a system of equations in $(u_t)_b$ and $(u_N)_b$ which has a unique solution if $V_b' \neq V_b$. Notice that in one dimension this is precisely the operator previously derived. Now, we could solve for $(u_N)_b$ to get an explicit expression but it will be more productive to first examine the order of accuracy. So far, we have shown that the scheme is stable and given a suggestion on how to close it at the boundaries.

5.3 Order of Accuracy

At first, we will show that (38) and (35) yield a consistent first order difference approximation of the Laplacian at the boundary provided that the internal

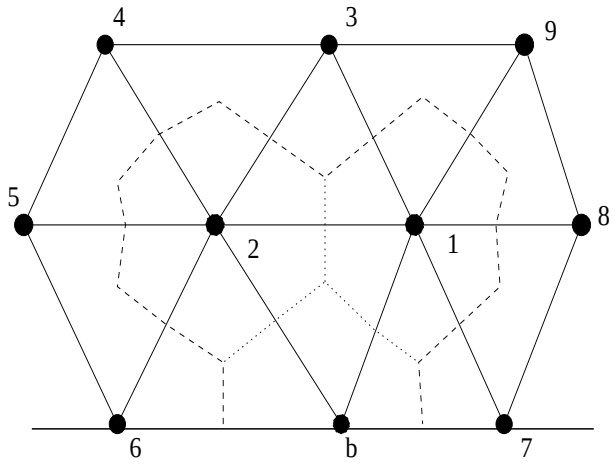

Figure 5: A generic grid near the boundary. The dashed line represents V_b'. The dotted lines are the internal boundaries that separate V_1, V_2 and V_b.

scheme is first order accurate. Recall that first order is sufficient at the boundary to obtain global second order accuracy. Combining (38) and (35) gives,

$$V_b' \cdot (u_t)_b - \sum_{i \in I_b} \sum_{j \in L_{bi}} \frac{u_j - u_i}{r_{ji}} ds_{ij} = V_b \cdot (u_t)_b - \sum_{n \in N_b} \frac{u_n - u_b}{r_{nb}} ds_{bn}. \qquad (39)$$

The objective is to reformulate this equation such that $(u_t)_b$ equals an approximation of the Laplacian at r_b. This will be done by adding the internal normal derivative contributions to finally obtain a linear combination of approximate Laplacians at neighbouring points. The resulting approximate Laplacians will be interpreted at r_b using Taylor expansions.

We will add normal derivative contributions between points inside V_b'. That is derivatives between I_b and $I_b \cap N_i$. However, if r_k is in $I_b \cap N_i$ it is also in I_b. That is, at each internal edge there will be two contributions of equal magnitude but opposite sign, indicating that adding these terms does not alter the equation.

Consider figure 5. In this case, the point $r_1 \in I_b$ and $I_b \cap N_1 = \{2, b\} \subset I_b$. Thus, at the internal boundary between r_2 and r_1 we can add,

$$\frac{u_2 - u_1}{r_{21}} ds_{12} + \frac{u_1 - u_2}{r_{12}} ds_{21} = 0.$$

Analogously, two terms may be added between r_1 and r_b as well as between r_2 and r_b.

Formally, this is stated and proved in the following lemma.

Lemma 5.3 *For a point r_b, $b \in B$ the following holds,*

$$\sum_{i \in I_b} \sum_{j \in N_i \cap I_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} = 0. \tag{41}$$

Proof To make the proof easier to follow some parentheses refering to the specific case of figure 5 are added. Use $I_b = M_b \cup \{b\}$ and that $M_b \cap \{b\} = \emptyset$ such that the second sum can be split on the left hand side of (41) (M_b=$\{1,2\}$) to obtain,

$$\sum_{i \in I_b} \sum_{j \in N_i \cap I_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} = \tag{42}$$
$$\sum_{i \in I_b} \left(\sum_{j \in N_i \cap M_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} + \sum_{j \in N_i \cap \{b\}} \frac{u_j - u_i}{r_{ji}} ds_{ij} \right) = RHS$$

Again, use $I_b = M_b \cup \{b\}$ to split the sum on the right hand side of (42),

$$RHS = \sum_{i \in M_b} \left(\sum_{j \in N_i \cap M_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} + \sum_{j \in N_i \cap \{b\}} \frac{u_j - u_i}{r_{ji}} ds_{ij} \right) +$$
$$\sum_{j \in N_b \cap M_b} \frac{u_j - u_b}{r_{jb}} ds_{bj} + \sum_{j \in N_b \cap \{b\}} \frac{u_j - u_b}{r_{jb}} ds_{bj}$$

The next step of the proof uses, $N_b \cap \{b\} = \emptyset$ ($\{1,2,6,7\} \cap \{b\} = \emptyset$). For any $j \in N_i \cap \{b\}$ where $i \in M_b$ yields $j = b$ (for example $j \in N_1 \cap \{b\} = \{7,8,9,3,2,b\} \cap \{b\} = \{b\}$). Finally, we will need that $N_b \cap M_b = M_b$. These relations yield,

$$RHS = \sum_{i \in M_b} \sum_{j \in N_i \cap M_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} + \sum_{i \in M_b} \frac{u_b - u_i}{r_{bi}} ds_{ib} + \sum_{j \in M_b} \frac{u_j - u_b}{r_{jb}} ds_{bj},$$

where the last two terms cancel each other. Thus, (42) can be written,

$$\sum_{i \in I_b} \sum_{j \in N_i \cap I_b} \frac{u_j - u_i}{r_{ji}} ds_{ij} = \sum_{i \in M_b} \sum_{j \in N_i \cap M_b} \frac{u_j - u_i}{r_{ji}} ds_{ij}.$$

Finally, using $r_{ij} = r_{ji}$, $ds_{ij} = ds_{ji}$ and $N_i \cap M_b = M_b \setminus \{i\}$ for $i \in M_b$ (for example $i = 1 \in M_b = \{1,2\}$ yields $N_1 \cap M_b = \{7,8,9,3,2,b\} \cap \{1,2\} =$

$\{2\} = M_b \setminus \{1\}$) , (42) becomes,

$$\sum_{i\in I_b}\sum_{j\in N_i\cap I_b}\frac{u_j-u_i}{r_{ji}}ds_{ij} = \sum_{i\in M_b}\sum_{j\in N_i\cap M_b}\frac{u_j-u_i}{r_{ji}}ds_{ij} = \sum_{i\in M_b}\sum_{j\in M_b\setminus\{i\}}\frac{u_j-u_i}{r_{ji}}ds_{ij} =$$

$$\sum_{i\in M_b}\left((u_i-u_i)\frac{ds_{ii}}{r_{ii}} + \sum_{j\in M_b\setminus\{i\}}\frac{(u_j-u_i)ds_{ij}}{r_{ji}}\right) = \sum_{i\in M_b}\sum_{j\in M_b}(u_j-u_i)\frac{ds_{ij}}{r_{ji}} = 0,$$

where $ds_{ii}/r_{ii} = 1$ $i \in M_b$, is defined merely as a symbol. ∎

The next observation is that if we add all the internal contributions we obtain the sum of the finite volume approximations for each dual volume inside V_b' except at the boundary point b. This is used when proving the next lemma. Again, some parentheses are added to clarify the set algebra using figure 5 as an example.

Lemma 5.4 *Equation (39) is, at least, a first order approximation of (1) at r_b. On Cartesian grids it is second order accurate.*

Proof Using lemma 5.3 we subtract (41) from the left hand side of (39).

$$V_b'\cdot(u_t)_b - \sum_{i\in I_b}\sum_{j\in L_{bi}}\frac{u_j-u_i}{r_{ji}}ds_{ij} - \sum_{i\in I_b}\sum_{j\in N_i\cap I_b}\frac{u_j-u_i}{r_{ji}}ds_{ij} = V_b\cdot(u_t)_b - \sum_{n\in N_b}\frac{u_n-u_b}{r_{nb}}ds_{bn}. \tag{43}$$

Since $L_{bi} = N_i \setminus I_b = N_i \setminus (N_i \cap I_b)$ we obtain that $L_{bi} \cup (N_i \cap I_b) = N_i$.

(For example, $L_{b1} = N_1 \cap I_b = \{b,7,8,9,3,2\} \setminus \{b,1,2\} = \{7,8,9,3\}$. Or, $L_{b1} = N_1 \setminus (N_1 \cap I_b) = \{b,7,8,9,3,2\} \setminus (\{b,7,8,9,3,2\} \cap \{b,1,2\}) = \{b,7,8,9,3,2\} \setminus \{b,2\} = \{7,8,9,3\}$)

Using this in (43),

$$V_b'\cdot(u_t)_b - \sum_{i\in I_b}\sum_{j\in N_i}\frac{u_j-u_i}{r_{ji}}ds_{ij} = V_b\cdot(u_t)_b - \sum_{n\in N_b}\frac{u_n-u_b}{r_{nb}}ds_{bn}.$$

The term $\sum_{i\in I_b}\sum_{j\in N_i}\frac{u_j-u_i}{r_{ji}}ds_{ij}$ is almost a sum of Laplacians of the points interior of V_b'. The only troublesome part is the boundary point

where the normal derivative out of the domain is missing. However, this part is found as the last term of (43). Then, by combining these two terms, using $I_b \setminus \{b\} = M_b$, we obtain,

$$V_b' \cdot (u_t)_b - \sum_{i \in M_b} \sum_{j \in N_i} \frac{u_j - u_i}{r_{ji}} ds_{ij} = V_b \cdot (u_t)_b.$$

The sum over M_b is a sum of the approximate Laplacians of interior points, the boundary point excluded.

Solving for $(u_t)_b$ gives,

$$(u_t)_b = \frac{1}{\sum_{n \in M_b} V_n} \left(\sum_{i \in M_b} \sum_{j \in N_i} \frac{u_j - u_i}{r_{ji}} ds_{ij} \right). \tag{45}$$

Using equation (5) we see that,

$$\frac{1}{V_i} \sum_{n \in N_i} \frac{u_n - u_i}{r_{ni}} ds_{in} = (\tilde{\Delta} u)_i,$$

where $(\tilde{\Delta} u)_i$ denotes the finite volume approximation of Δu at r_i. Thus equation (45) becomes,

$$(u_t)_b = \frac{1}{\sum_{n \in M_b} V_n} \left(\sum_{i \in M_b} V_i (\tilde{\Delta} u)_i \right). \tag{47}$$

Since $(\tilde{\Delta} u)_i$ is an approximation of $(\Delta u)_i$ it is straightforward to show, using Taylor expansions, that $(\tilde{\Delta} u)_i = (\Delta u)_b + \mathcal{O}(r_{ib})$ for $i \in N_b$. Denote by h_b the maximum of r_{ib} such that $i \in M_b$. Then (47) yields,

$$(u_t)_b = \frac{1}{\sum_{n \in M_b} V_n} \left(\sum_{i \in M_b} V_i (\tilde{\Delta} u)_b \right) + \mathcal{O}(h_b) = \tilde{\Delta} u_b + \mathcal{O}(h_b), \tag{48}$$

i.e. a first order approximation of the Laplacian at r_b.

On a Cartesian grid, first and second order accuracy for the second and first derivative approximations respectively, follows directly from subsection 4.2. The only difference is that the volumes being line segments in the 1D case are multiplied by a (possibly multidimensional) width of the volume. This width is the same along a coordinate line (c.f. figure 4). It is straightforward, although tedious to prove second order accuracy on curvilinear structured grids, so we omit that part here. ∎

The usual interior finite volume approximation (5) is assumed to be at least a first order accurate approximation at u_i on an arbitrary grid. For an arbitrary grid we have just proven first order accuracy of (35), i.e. the boundary approximation of the Laplacian. It is likely that $(u_N)_b$ is a second order approximation of the normal derivative at the boundary. However, to prove order of accuracy of $(u_N)_b$ a specific grid has to be considered and Taylor expansions applied. Thus, we make the following assumption and check it, either numerically or theoretically, for each grid.

Assumption 5.5 *If the internal approximation (5) is first order accurate then $(u_N)_b$ given by (35) and(38) is second order accurate.*

Remark If the grid is Cartesian we have already proven the assumption true in 4.2. Other types of grids are considered below by using Taylor expansions.

When implementing this scheme one should use (45) multiplied by V_b as the approximation of the Laplacian at the boundary. To implement the boundary conditions we will also need $(u_N)_b$. This is most easily obtained by combining (45) and (35) obtaining,

$$(u_N)_b ds_{bb} = \frac{V_b}{\sum_{n\in M_b} V_n}\left(\sum_{i\in M_b}\sum_{j\in N_i}\frac{u_j-u_i}{r_{ji}}ds_{ij}\right) - \sum_{n\in N_b}\frac{u_n-u_b}{r_{nb}}ds_{bn}. \quad (49)$$

Theorem 5.6 *The discretisation (5) in the interior and (35),(49) at the boundary, together with the penalty method for imposition of boundary conditions and assumption 5.5, is a stable and at least a first order accurate approximation of the heat equation.*

Proof Proposition 5.2 and lemma 5.4 together with the derivation in section 3 using the penalty technique, i.e. (9)-(11),(17) and (18) prove the theorem. ∎

5.4 Corners

An additional comment has to be made on corners of the computational domain. The suggested boundary approximation of the normal derivative might not be possible to construct at corners in the absence of an edge from the corner point to some inner point. Even in cases where there are such an edge, the suggested approximation is not always optimal. The reason is that the normal derivative at the corner is not defined. Rather one should construct two different normal derivatives, one for each side of the corner and impose

the boundary conditions using two penalty terms. In the computations on the unit square in section 6.2 we have used the 1D approximation, derived in subsection 4.2, along the boundary. (See figure 6.)

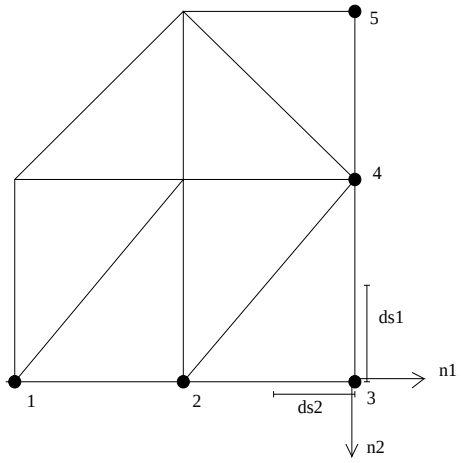

Figure 6: A generic grid at a corner. $n1$ and $n2$ are normal directions.

In figure 6 the points $1, 2$ and 3 are used in the 1D approximation (23), (24)-(26) to approximate the normal derivative in the $n1$ direction for the $ds1$ part of the dual grid. Correspondingly the points $3, 4$ and 5 are used in the $n2$ direction for $ds2$. Thus, it is also clear that if there are different boundary conditions on different boundaries the above procedure handles that at the corner cell.

Another approach that is preferable when the corner angle is different from $\pi/2$ would be to use the approximation of the Laplacian corresponding to points 2 and 4 and calculate DS for $ds1$ and $ds2$ in a similar manner as in equation (49) where the inner scheme approximation is replaced by the boundary approximation of the Laplacian at the points 2 and 4 respectively. That is,

$$(u_N)_3 ds_{33} = \frac{V_3}{V_2 + V_4} \left(\sum_{i \in M_3} \sum_{j \in N_i} \frac{u_j - u_i}{r_{ji}} ds_{ij} + (u_N)_2 ds_{22} + (u_N)_4 ds_{44} \right) - \sum_{n \in N_3} \frac{u_n - u_3}{r_{n3}} ds_{3n}.$$

This would also result in, at least, a first order approximation and it can also handle different boundary conditions on $ds1$ and $ds2$ (c.f figure 6). This is easily seen if $(u_N)_3 ds1 \cdot n1 = g_1(t) ds1 \cdot n1$ and $(u_N)_3 ds2 \cdot n2 = g_2(t) ds2 \cdot n2$ then $(u_n)_3 ds_{33} = g_1(t) ds1 + g_2(t) ds2$ is used in the penalty term. In section 6.5 this technique is applied.

6 Numerical Experiments

The Laplacian was derived using the heat equation as a model problem but the choice of equation is not essential. As an example, we will do numerical computations using the two-dimensional wave equation. The example will also illustrate the use of the SAT boundary treatment.

6.1 The Wave Equation

Consider,

$$u_{tt} = \Delta u + F(x, y, t), \quad 0 \leq x, y \leq 1, \quad 0 \leq t \leq T, \tag{50}$$

with the solution $u(x, y, t)$. In the computations below we use $u(x, y, 0) = sin(x)sin(y)$ and $u_t(x, y, 0) = 0$ as initial data. Further, the boundary data are,

$$\begin{aligned}
u_x &= g_{x0} = sin(y)cos(t) && at \quad x = 0,\\
u_x &= g_{x1} = cos(1)sin(y)cos(t) && at \quad x = 1,\\
u_y &= g_{y0} = sin(x)cos(t) && at \quad y = 0,\\
u_y &= g_{y1} = cos(1)sin(x)cos(t) && at \quad y = 1.
\end{aligned}$$

We also use $F(x, y) = sin(x)sin(y)cos(t)$. The problem has the exact solution $u(x, y, t) = sin(x)\ sin(y)\ cos(t)$.

To prove well posedness of the problem above we begin by making a transformation of variables such that $w = u - f$. Differentiate to obtain $w_x = u_x - f_x$ and $w_y = u_y - f_y$. Now, choose $f_x(i, y, t) = g_{xi}$ at $x = i$, $i = 0, 1$ and $f_y = g_{yj}$ at $y = j$, $j = 0, 1$. By inserting the transformation above in equation (50) we obtain,

$$\begin{aligned}
w_{tt} &= \Delta w + \tilde{F}(x, y, t), \\
\tilde{F}(x, y, t) &= \Delta f(x, y, t) - f(x, y, t)_{tt} + F(x, y, t)
\end{aligned} \tag{51}$$

with homogeneous boundary conditions. We apply the energy method by multiplying (51) by $2w_t$ and integrating. Then,

$$\begin{aligned}
\frac{d}{dt}(\|w_t\|^2 + \|w_x\|^2 + \|w_y\|^2) \leq \\
2\int_0^1 [w_t w_x]_0^1 dy + 2\int_0^1 [w_t w_y]_0^1 dx + 2\eta\|w_t\|^2 + \frac{2}{\eta}\|\tilde{F}\|^2 \leq \\
2\eta\|w_t\|^2 + \frac{2}{\eta}\|\tilde{F}\|^2,
\end{aligned}$$

where $\|\cdot\|$ is the L^2 norm on the unit square and $\eta > 0$. We can conclude that the problem is well posed.

Discretise equation (50) using the operator derived in section 5. G denotes the vector with $g_{x0}, g_{x1}, g_{y0}, g_{y1}$ at positions corresponding to the boundary points and $F(t)$ is a vector containing the values of $F(x, y, t)$ projected onto the grid. Then,

$$v_{tt} = P^{-1}Q_\Delta v + \sigma P^{-1}(DSv - G) + F(t), \tag{52}$$

on a grid with $n \times n$ points. The corners are discretised as described in subsection 5.4. The time integration is done using a three stage explicit Störmer method. (See for example [19].)

To prove stability we apply the energy method to the semidiscrete equation (52). As in the continuous case we make a variable transformation $DS\tilde{v} = DSv - G$ to homogenise the problem. Thus, we choose a vector function $\tilde{g}(t)$ such that $DS\tilde{g} = G$, yielding the transformation of variables $\tilde{v} = v - \tilde{g}$. Inserting this in (52) yields,

$$\frac{d}{dt}\|\tilde{v}_t\|_P^2 - \frac{d}{dt}(2\tilde{v}^T A\tilde{v}) = 2\tilde{v}_t^T DS\tilde{v} + 2\sigma\tilde{v}_t^T(DS\tilde{v} - \mathbf{0}) + 2\tilde{v}_t^T P\tilde{F},$$
$$\tilde{F}(t) = F(t) + P^{-1}Q_\Delta\tilde{g} - \tilde{g}_{tt},$$

where $\mathbf{0}$ is a vector with 0 on all positions. Since A is symmetric and negative definite we achieve stability with $\sigma = -1$.

6.2 Cartesian Grid

With these settings we compute the solution from $t = 0$ to $t = 1$. The L^2-error is calculated and the order of accuracy is deduced for structured Cartesian grids. A typical grid is shown in figure 7. At the corners we apply first derivative approximations derived in subsection 4.2 for $(u_N)_{corner}$. For further details, see subsection 5.4. The convergence of the solution as the grid is refined from $N = 5 \times 5$ points to $N = 50 \times 50$ points is shown in figure 8. A good agreement with the reference line representing second order convergence is achieved. The numbers plotted in figure 8 are given in table 1. $\|\cdot\|_2$ is the l^2-norm scaled by the grid size Δx, i.e. $\|a\|_2 = \Delta x\sqrt{\sum_{i=1}^N a_i^2}$ where Δx is the side length of a cell. A small Δt is used to minimise the influence of the time integration scheme, $\Delta t = 0.1(\Delta x)^2$.

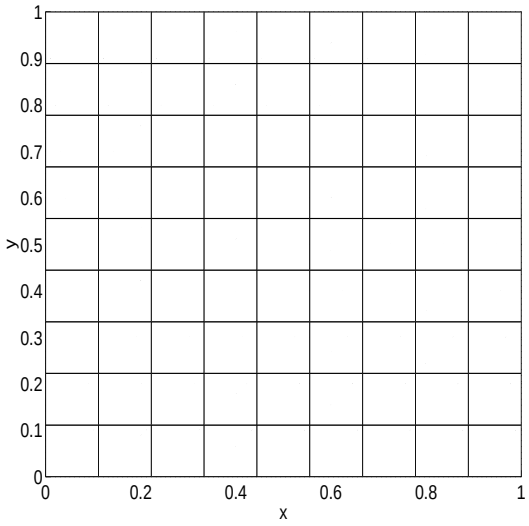

Figure 7: A Cartesian grid with 10×10 points.

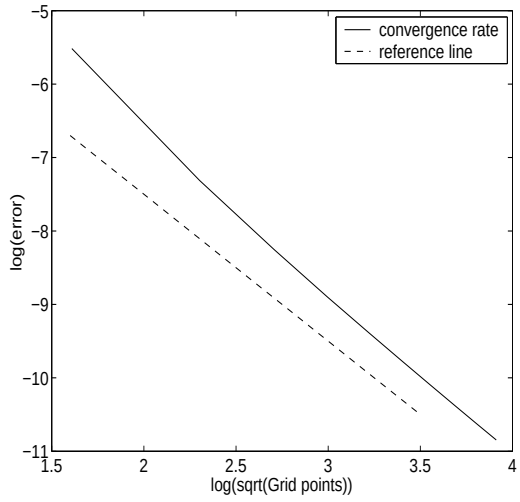

Figure 8: Convergence for Cartesian grids. The solid line is the convergence data for the finite volume method. The x-axis are $\log(\sqrt{N})$ where N is the number of grid points. The y-axis is $\log(\|error(x, y)\|_2)$. The dashed line is a reference line with slope -2 representing second order convergence.

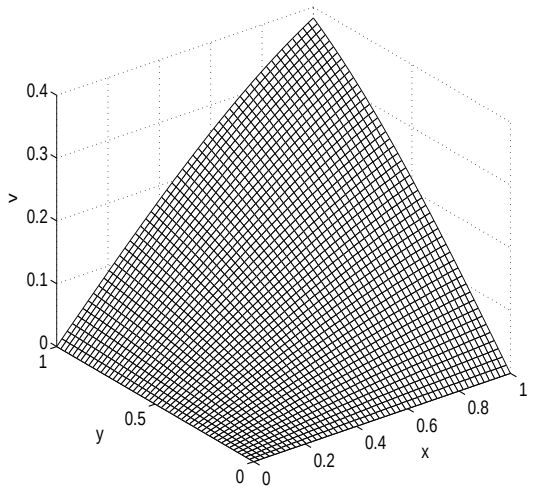

Figure 9: Numerical solution at $T = 1$ on the 50×50 points Cartesian grid.

$\sqrt{N}$	$\log(\sqrt{N})$	$\log(\|error(x,y)\|_2)$	max(eig(A))
5	1.6094	-5.5160	-0.0293e-14
10	2.3026	-7.3127	-0.0198e-14
15	2.7081	-8.2562	0.0584e-14
20	2.9957	-8.8977	-0.0414e-14
25	3.2189	-9.3821	-0.0446e-14
30	3.4012	-9.7730	0.1251e-14
35	3.5553	-10.0995	0.1479e-14
40	3.6889	-10.3803	0.1159e-14
45	3.8067	-10.6267	0.0247e-14
50	3.9120	-10.8458	-0.06032e-14

Table 1: Convergence data and the maximum eigenvalue of A for the Cartesian case.

To confirm that the method is stable the largest eigenvalue of A is computed numerically for the different mesh sizes, displayed in table 1. These are all zero to round off errors. The solution on the 50×50 points mesh is plotted in figure 9.

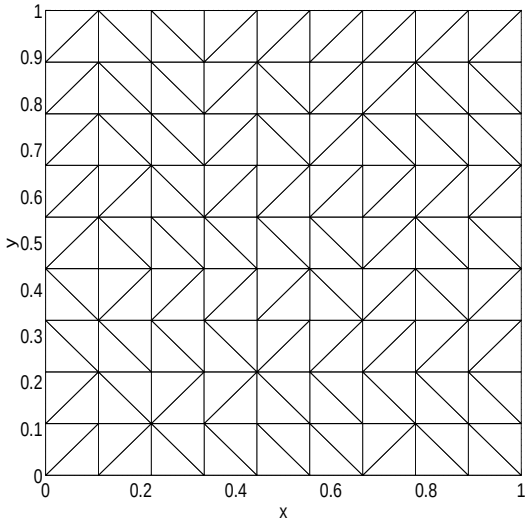

Figure 10: A triangulated Cartesian grid with 10×10 points.

6.3 Triangulated Cartesian Grid

Next, we turn to an unstructured case. We triangulate the Cartesian grids previously used (see figure 10) and compute the solution of the same problem defined in 6.1. Δx and Δt are the same as in the previous computations.

As in the Cartesian case the eigenvalues of A (table 2) shows that the scheme is stable. The convergence results are displayed in figure 11 and table 2.

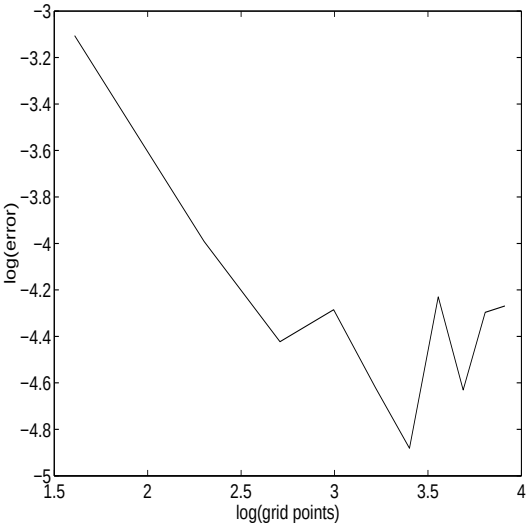

Figure 11: Convergence for the triangulated Cartesian grids.

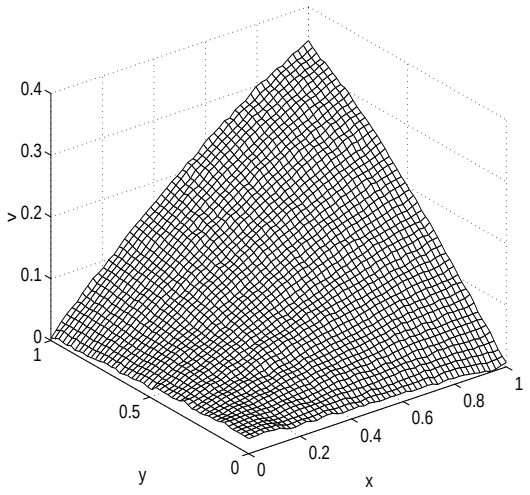

Figure 12: Numerical solution for the 50×50 points triangulated Cartesian grid at $T = 1$.

$\sqrt{N}$	$\log(\sqrt{N})$	$\log(\|error(x,y)\|_2)$	max(eig(A))
5	1.6094	-3.1057	0.0003e-13
10	2.3026	-3.9919	-0.0113e-13
15	2.7081	-4.4227	0.0182e-13
20	2.9957	-4.2852	0.0154e-13
25	3.2189	-4.6201	-0.0183e-13
30	3.4012	-4.8813	0.0026e-13
35	3.5553	-4.2300	0.1079e-13
40	3.6889	-4.6305	0.0338e-13
45	3.8067	-4.2963	-0.1545e-13
50	3.9120	-4.2693	-0.0289e-13

Table 2: Convergence data for the triangulated Cartesian case.

In figure 11 only the first few grid refinements result in a better approximation, thereafter the error behaves in a random fashion. Since the scheme is stable (see table 2) it can not be consistent. Also, the errors are everywhere large compared to the Cartesian case. The bad quality of the solution is highlighted in figure 12. A thorough analysis of the consistency will be made in the next subsection.

6.4 Consistency Analysis for the Triangulated Cartesian Mesh

Consider the triangular grid shown in figure 13.

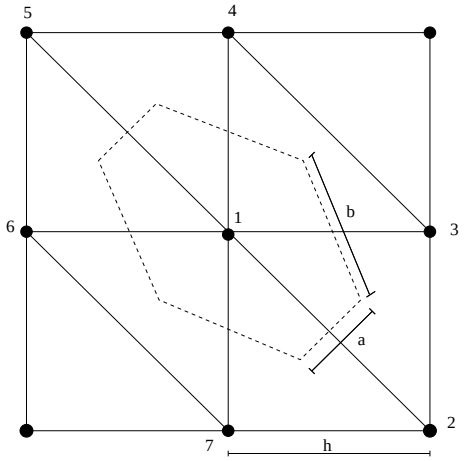

Figure 13: A triangular grid where the dual grid is the dashed line, $b = h\sqrt{5}/3$ and $a = h\sqrt{2}/3$.

We will analyse consistency for the centre point of this grid. The scheme is,

$$V_1(u_1)_t = \sum_{i=2}^{7} \frac{u_i - u_1}{r_{i1}} ds_{i1} \tag{53}$$

$V_1 = h^2$ is the area inside the dashed line. By Taylor expansions we obtain for the point 2.

$$\frac{u_2 - u_1}{\sqrt{2}h} = \frac{u_1 + hu_x + hu_y + \frac{1}{2}(h^2 u_{xx} + 2h^2 u_{xy} + h^2 u_{yy}) - u_1 + \mathcal{O}(h^3)}{\sqrt{2}h},$$

where the derivatives are taken at point 1. A similar derivation for point 5 leads to,

$$\frac{u_2 - u_1}{\sqrt{2}h} ds_{12} + \frac{u_5 - u_1}{\sqrt{2}h} ds_{15} = \frac{u_2 - u_1}{\sqrt{2}h} \frac{\sqrt{2}h}{3} + \frac{u_5 - u_1}{\sqrt{2}h} \frac{\sqrt{2}h}{3} = \frac{h^2 u_{xx} + 2h^2 u_{xy} + h^2 u_{yy} + \mathcal{O}(h^3)}{3}.$$

In the same manner we obtain for points 3,6 and 7,4.

$$\frac{u_3 - u_1}{h} ds_{13} + \frac{u_6 - u_1}{h} ds_{16} = \frac{\sqrt{5}}{3}(h^2 u_{xx} + \mathcal{O}(h^3)),$$
$$\frac{u_7 - u_1}{h} ds_{17} + \frac{u_4 - u_1}{h} ds_{14} = \frac{\sqrt{5}}{3}(h^2 u_{yy} + \mathcal{O}(h^3)).$$

Summing up, using (53) and $V_1 = h^2$ yields,

$$(u_1)_t = \frac{1 + \sqrt{5}}{3} \Delta u + \frac{2}{3} u_{xy} + \mathcal{O}(h),$$

i.e. an $\mathcal{O}(1)$ error. This is stated in theorem 6.1.

Remark As mentioned in the introduction, the Laplacian enters the compressible Navier–Stokes equations when neglecting the cross derivative terms using a specific thin layer approximation. As was shown above the cross derivative term reenters the scheme with a weight depending on the grid. If the elements of the grid do not have the same geometry everywhere, different equations will be solved on different parts of the domain. This explains why the solution in figure 12 looks oscillatory in the interior of the domain. The diagonals that triangulate the Cartesian grid are not all oriented in the same direction (see figure 10).

Theorem 6.1 *On a general grid the approximation (5) is inconsistent.*

Remark As is seen in the derivation above, the theorem does not imply inconsistency on all grids.

6.5 Computations on Equilateral Triangular Grids

As is shown in the analysis and computations above, the internal scheme is not consistent for the triangulated Cartesian grid. It is also easily seen that the main problem is the cross derivative term that does not cancel. On a grid consisting of equilateral polygons the scheme will be consistent. Furthermore, for such grids assumption 5.5 will be fulfilled and thus theorem 5.6 applies.

To show convergence, the scheme is run on an equilateral triangular domain, which is split into smaller equilateral triangles. The same PDE, the wave equation in 2D, with the same forcing function was used. However,

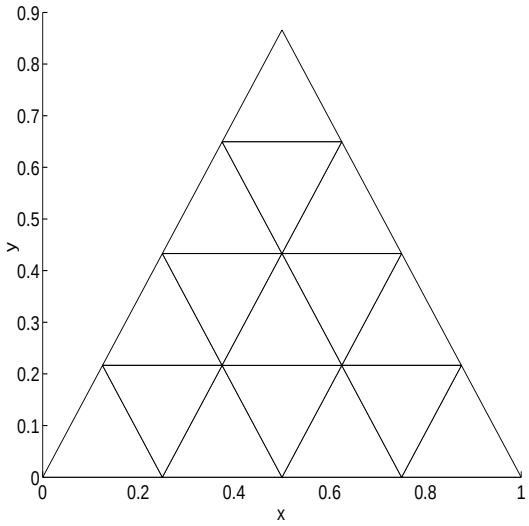

Figure 14: An equilaterally triangulated grid.

the boundary conditions have to be changed to obtain the same solution as before. In this case,

$$g_{y0} = -sin(x)cos(t) \quad x = [0,1], y = 0$$

$$g_{yx0} = \big(-\frac{\sqrt{3}}{2}cos(x)sin(y) + \frac{1}{2}sin(x)cos(y)\big)cos(t),$$

$$x \in [0, \frac{1}{2}], y = \sqrt{3}x$$

$$g_{yx1} = \big(\frac{\sqrt{3}}{2}cos(x)sin(y) + \frac{1}{2}sin(x)cos(y)\big)cos(t),$$

$$x \in [\frac{1}{2}, 1], y = \sqrt{3}(1-x)$$

Here, Δx is taken to be the side of a triangular cell and $\Delta t = 0.1(\Delta x)^2$ as before. A similar stability derivation as that of subsection 6.1 applies here as well. The 15 points grid is shown in figure 14. This grid is refined by inscribing an equilateral triangle into each cell yielding three times the number of triangles. In table 3 and figure 15 the convergence results are displayed.

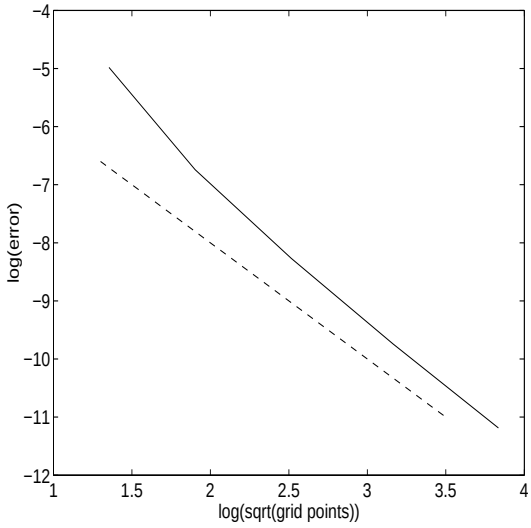

Figure 15: Convergence rate for equilaterally triangulated grids. The dashed line is a reference with slope -2, corresponding to second order accuracy.

N	$\sqrt{N}$	$\log(\sqrt{N})$	$\log(\|error(x,y)\|_2)$	max(eig(A))
15	3.8730	1.3540	-4.9815	0.2385e-15
45	6.7082	1.9033	-6.7421	0.3154e-15
153	12.3693	2.5152	-8.2681	-0.1866e-15
561	23.6854	3.1649	-9.7422	-0.9259e-15
2145	46.3141	3.8354	-11.1877	-0.5086e-15

Table 3: Convergence data and maximum eigenvalue of A for the equilaterally triangulated case.

As is seen in figure 15 and table 3 convergence is recovered even to second order accuracy. Thus the internal scheme is second order accurate and the boundary treatment is at least first order accurate. In fact, the convergence is larger than second order to begin with which indicates that the boundary treatment is second order since one order higher convergence is obtained at the boundary. (c.f [20]) However, as the boundary points become a smaller part of the total amount of points the convergence decreases to second order. Also in this example, the maximum eigenvalue of A is computed and shown in table 3. These are all zero to round off errors, implying that the scheme is stable. In figure 16 the solution at $T = 1$ is displayed.

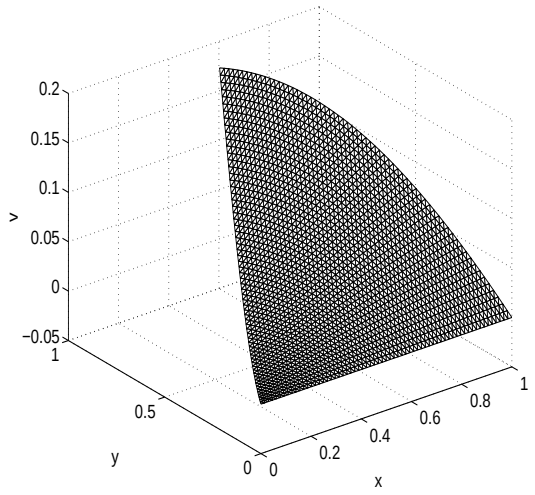

Figure 16: Numerical solution for the 2145 points equilatarally triangulated grid at $T = 1$.

6.6 A Comment on Possible Grids

It has been shown that the scheme works satisfactory on Cartesian as well as equilateral triangular grids which are both equilateral polygons. In fact, the scheme will work on any equilateral polygon. In the case of equilateral triangles the domain was also chosen to be a triangle. One might of course have a domain that consists of several equilateral triangles. However, it is not possible to alter the shape of the cell triangles near the boundary since neither the Laplacian nor the normal derivative would be consistently approximated. Further, the scheme becomes inconsistent if it is built of different equilateral polygons since at internal boundaries between domains of different elements the scheme would be inconsistent.

6.7 A Comparison with Finite Elements

The triangulated Cartesian grid used in 6.3 resulted in an inconsistent finite volume scheme. In the equilateral triangular case the finite volume scheme coincides with a mass lumped finite element scheme. To show the differences between finite volume and finite element schemes we derive the mass lumped finite element approximation with linear basis functions applied to the grid in figure 13 and obtain,

$$h^2(u_1)_t = -v_4 - v_3 + 4u_1 - u_7 - u_6.$$

The left hand side, $h^2(u_1)_t$, is the result of the mass lumping and is precisely equal to $V_1(u_1)_t$ obtained with the finite volume technique. However, the right hand side coincides with the ordinary finite difference scheme on the corresponding Cartesian grid. Compared to the finite volume case different weights on all the neighbouring points are obtained and especially the cross derivative contribution is cancelled since u_5 and u_2 do not enter the scheme.

6.8 A Different Finite Volume Approximation

To further explore the finite volume discretisations of viscous terms we will briefly discuss a discretisation proposed in [8] and also used in [7]. This discretisation allows for the computation of all viscous terms in the Navier-Stokes equations but it is not as compact as the Laplace approximation previously discussed. The scheme is basically an application of the edge based first derivative approximation twice. (The first derivative approximation is discussed in [6] where it is shown to be an SBP operator.) In [8] they conclude that this scheme do not take into account the closest neighbours possibly yielding poor smoothing properties. (Compare the second derivative approximation that is obtained by applying a centered finite difference scheme twice.) Accordingly, they propose an augmented form that also uses the neighbours.

We will analyse consistency of the first derivative applied twice without this augmentation to see whether this can be used as an approximation of the Laplacian on more general grids. It certainly has the disadvantage that it is wider and thus not as computational efficient. Recalling that N_i is the neighbours of the point indexed i and denoting by $(u_x)_i$ the approximation of the x-derivative at the same point we have that,

$$(u_x)_i = \frac{1}{V_i} \sum_{j \in N_i} \frac{u_j + u_i}{2} \Delta y_j, \tag{56}$$

where Δy_j is the difference (with sign) in the y direction going anti-clockwise along ds_{ij}. If approximation (56) is carried out for point 1 in figure 13, we would obtain with $V_1 = h^2$,

$$(u_x)_1 = \frac{1}{V_1} \sum_{j \in N_1} \frac{u_j + u_1}{2} \Delta y_j = u(x_1, y_1)_x + \mathcal{O}(h^2). \tag{57}$$

Further, if the grid in figure 13 is extended in all directions we obtain second order approximations similar to equation (57) of $(u_x)_i$ at all the points $2 - 7$.

Next, applying (56) for point 1 again, yields a first order approximation of $(u_{xx})_i$.

$$(u_{xx})_1 = \frac{1}{V_1} \sum_{j \in N_1} \frac{(u_x)_j + (u_x)_1 + \mathcal{O}(h^2)}{2} \Delta y_j = u(x_1, y_1)_{xx} + \mathcal{O}(h) \qquad (58)$$

However, if the grid is not extended precisely as figure 13, for example if the location of the points 2 and 5 are slightly disturbed or perhaps the diagonal between 1 and 5 is replaced by a diagonal between 4 and 6, a term of the form hu_{xy} term will enter the Taylor expansion of (57). As an example consider point 1 with the proposed diagonal altered. (Compare also to figure 10 where the orientaion of the diagonals are somewhat arbitrary.). Then,

$$(u_x)_1 = u(x_1, y_1)_x + const \cdot hu_{xy} + \mathcal{O}(h^2). \qquad (59)$$

Next, suppose that this shift of diagonal orientation appears further away from point 1 such that 5 has an error of the form (59). Then $(u_x)_i$, $i \in \{2, 3, 4, 6, 7\}$ are approximated to second order and $(u_x)_5$ only to first order. Applying (56) again to compute $(u_{xx})_1$ yields,

$$(u_{xx})_1 = \frac{1}{V_1} \sum_{j \in N_1} \frac{(u_x)_j + (u_x)_1 + \mathcal{O}(h)}{2} \Delta y_j = u(x_1, y_1)_{xx} + \mathcal{O}(1). \qquad (60)$$

The last equality is due to the fact that the error of order $\mathcal{O}(h)$ only appears in one of the $(u_x)_i$:s. Thus, there is no way of cancelling that error. Then by multiplying by $\frac{\Delta y_5}{V_1} \sim \frac{1}{h}$ yields an error of $\mathcal{O}(1)$. This example shows that unless the grid has a high degree of regularity the scheme becomes inconsistent. The above considerations can be summarised in the following theorem.

Theorem 6.2 *On a general grid the application of the first derivative approximation (56) twice to approximate a second derivative results in an inconsistent scheme.*

Remark Note that the above theorem does not implicate that the scheme is inconsistent on all grids. As shown above, with certain restrictions imposed on the grid, the scheme is consistent.

As mentioned above, in [8] they propose an augmented form of the above approximation. However, the augmentation assumes implicitly that the above approximation is consistent. Thus, the same restriction of possible grids applies here as well. Further, even if the grid has the required properties, the issue of implementing boundary conditions is left. Without analysing

the scheme in detail we can conclude that in the finite difference case, a second order SBP operator for the first derivative will give an inconsistent second derivative approximation at the boundary if applied twice and the same might happen in this case.

7 Conclusions

We have interpreted a basic finite volume algorithm in the interior domain (see [1, 2, 3, 4]) as a summation-by-parts formulation and constructed an SBP boundary treatment. The new approximation works equally well in any number of space dimensions. This formulation together with the penalty technique for implementing boundary conditions lead to a stable and consistent scheme. The boundary treatment guarantees that global second order accuracy is achieved on structured grids and on grids consisting only of equilateral polygons which is verified by numerical computations.

The strength of the finite volume scheme is its applicability on unstructured grids where the rather heavy construction of data structures is compensated by the simpler and possibly automatic construction of grids around complex geometries and efficient implementation of the scheme (cf [2, 3, 4, 5, 7, 6, 1]). These were the main reasons to construct a stable boundary implementation. However, the use of unstructured grids is limited by the condition that the elements have to be equilateral polygons to achieve consistency.

To give a broader view of finite volume approximations of the Laplacian we also shortly discussed the approach taken in [8] and [7]. A consistency analysis showed that similar restrictions on the grid applies to this method as well. The conclusion is that when viscous flows are computed it seems that extra care has to be taken when constructing grids.

Future work would be to construct an algorithm that fulfils as many of the following conditions as possible. It should have no grid constraints; it should keep the edge based structure for efficiency and grid transparancy; it should be in an SBP framework for stability, systematic implementation of boundary conditions and compatibility with the first derivative.

References

[1] A. Haselbacher, J.J. McGuirk, and G.J. Page. Finite volume discretization aspects for viscous flows on mixed unstructured grids. *AIAA Journal*, 37(2), Feb. 1999.

[2] D.J. Mavriplis and V. Venkatakrishnan. A unified multigrid solver for the Navier-Stokes equations on mixed element meshes. Technical report, Institute for Computer Applications in Science and Engineering, 1995.

[3] D.J Mavriplis. Multigrid strategies for viscous flow solvers on anisotropic unstructured meshes. *J. Comput. Physics*, 145, 1998.

[4] D.J. Mavriplis and D.W. Levy. Transonic drag prediction using an unstructured multigrid solver. Technical report, Institute for Computer Applications in Science and Engineering, 2002.

[5] D.J Mavriplis. Accurate multigrid solution of the Euler equations on unstructured and adaptive meshes. *AIAA Journal*, 28(2), Feb. 1990.

[6] Jan Nordström, Karl Forsberg, Carl Adamsson, and Peter Eliasson. Finite volume methods, unstructured meshes and strict stability for hyperbolic problems. *To appear in Applied Numerical Mathematics*, 2003.

[7] T. Gerhold, O. Friedrich, and J.Evans. Calculation of complex three-dimensional configurations employing the *DLR-τ-Code*. In *AIAA paper 97-0167*, 1997.

[8] J.M. Weiss, J.P. Maruszewski, and W.A. Smith. Implicit solution of preconditioned Navier-Stokes equations using algebraic multigrid. *AIAA Journal*, 37(1), Jan. 1999.

[9] P.D. Lax and R.D.Richtmyer. Survey of the stability of linear finite difference equations. *Comm. on Pure and Applied Math.*, IX, 1956.

[10] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. *Mathematical Aspects of Finite Elements in Partial Differential Equations., Academic Press, Inc.*, 1974.

[11] H.-O. Kreiss and G. Scherer. On the existence of energy estimates for difference approximations for hyperbolic systems. Technical report, Dept. of Scientific Computing, Uppsala University, 1977.

[12] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. The stability of numerical boundary treatments for compact high-order finite difference schemes. *J. Comput. Phys.*, 108(2), 1994.

[13] M. H. Carpenter, J. Nordström, and D. Gottlieb. A stable and conservative interface treatment of arbitrary spatial accuracy. *J. Comput. Phys.*, 148, 1999.

[14] J. Nordström and M. H. Carpenter. Boundary and interface conditions for high-order finite-difference methods applied to the Euler and Navier-Stokes equations. *J. Comput. Phys.*, 148, 1999.

[15] J. Nordström and M. H. Carpenter. High-order finite difference methods, multidimensional linear problems, and curvilinear coordinates. *J. Comput. Phys.*, 173, 2001.

[16] B. Gustafsson. The convergence rate for difference approximations to mixed initial boundary value problems. *Math. Comp.*, 29(130):396–406, Apr. 1975.

[17] B. Gustafsson. The convergence rate for difference approximations to general mixed initial boundary value problems. *SIAM J. Numer. Anal.*, 18(2):179–190, Apr. 1981.

[18] R.A. Horn and C.R. Johnson. *Matrix analysis.* Cambridge University Press, 1990.

[19] E. Hairer, S.P. Nørsett, and G.Wanner. *Solving Ordinary Differential Equations I.* Springer Verlag, 1993.

[20] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time dependent problems and difference methods.* John Wiley & Sons, Inc., 1995.