

AN UNCONDITIONALLY STABLE MIXED DISCONTINUOUS GALERKIN METHOD

Mika Juntunen

Rolf Stenberg



AN UNCONDITIONALLY STABLE MIXED DISCONTINUOUS GALERKIN METHOD

Mika Juntunen

Rolf Stenberg

Mika Juntunen, Rolf Stenberg: *An unconditionally stable mixed discontinuous Galerkin method*; Helsinki University of Technology, Institute of Mathematics, Research Reports A531 (2007).

Abstract: *For the model Poisson problem we propose a method combining the discontinuous Galerkin method with a mixed formulation. In the method independent and fully discontinuous basis functions are used both for the scalar unknown and its flux. The continuity requirement is imposed by Nitsche's technique [7]. In the implementation the flux is eliminated by local condensing. We show that the method is stable and optimally convergent for all positive values of the stability parameter. We also perform an a posteriori error analysis. The theoretical results are verified by numerical computations.*

AMS subject classifications: 65N30, 65N55

Keywords: Mixed method, discontinuous Galerkin method, stability, Nitsche's method

Correspondence

mika.juntunen@tkk.fi, rolf.stenberg@tkk.fi

ISBN 978-951-22-8942-4
ISSN 0784-3143
Teknillinen Korkeakoulu 2007, Espoo, Finland

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, FI-02015 TKK, Finland
email:math@tkk.fi <http://math.tkk.fi/>

1 Introduction

The purpose of this paper is to introduce and analyze a simple discontinuous Galerkin (DG) finite element method. The method differs from the (by now) standard DG method in that it is based on a mixed formulation in which the flux variable is taken as an independent unknown, fully discontinuous between elements. This flux is an auxiliary unknown that is condensed at each element at a negligible cost. The advantage of this approach is that it yields a stable method for all positive values of the stability parameter. We recall that for the standard DG the lower bound is given by a constant in a discrete trace inequality, cf. e.g. [7, 8]. We have been led to this formulation from our previous work on Galerkin-Least-Squares methods for the Stokes problem [5], where a similar phenomena occur.

Our method is similar to the Bassi-Rebay method [2], which has been analyzed in [4], but appears to be more straight forward both in analyzing and implementation. The method and its a priori analysis is probably also covered by the general theory developed in [1, 3], but by focusing on this particular method we are able to perform a concise and transparent analysis, both a priori and a posteriori.

The outline of the article is as follows: in the next section we introduce the model problem and derive the variational form of the DG method. Sections 3 and 4 are devoted to the a priori and a posteriori error analysis, respectively. Finally, in Section 5 we give numerical results.

2 The Model problem and the variational form

Our model problem is the mixed form of the Poisson equation, which we intend to solve with a discontinuous Galerkin method. The continuity in the variational formulation is imposed weakly using the Nitsche method. For simplicity we restrict ourselves to two dimensions.

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with a piecewise smooth boundary $\partial\Omega$. With \mathcal{T}_h we denote the mesh, i.e. the partitioning of Ω into triangles. With $\mathcal{E}_{\partial\Omega}$ we denote the edges of the triangles that lie on the boundary $\partial\Omega$ and with \mathcal{E}_{int} we denote the internal edges of the mesh. We assume that the boundary $\partial\Omega$ is split into two non-overlapping parts Γ_D and Γ_N . The edges on the boundary are grouped into those on the Dirichlet and Neumann part, respectively, i.e. $\mathcal{E}_{\partial\Omega} = \mathcal{E}_D \cup \mathcal{E}_N$. In addition, we denote with h_T the diameter of the element $T \in \mathcal{T}_h$ and with h_E the diameter of $E \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_{\partial\Omega}$. For the mesh we assume that there exists $C_1, C_2 > 0$ such that

$$C_1 h_E \leq h_T \leq C_2 h_E \quad \forall E \subset \partial T, \quad \forall T \in \mathcal{T}_h.$$

We solve the problem

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= u_0 && \text{on } \Gamma_D, \\ \nabla u \cdot \mathbf{n} &= g && \text{on } \Gamma_N, \end{aligned} \tag{1}$$

in which the load $f \in L^2(\Omega)$, $u_0 \in H^{1/2}(\Gamma_D)$ and $g \in L^2(\Gamma_N)$. Instead of solving the equations (1) directly, we pose the problem in an equivalent mixed form

$$\begin{aligned} \boldsymbol{\sigma} - \nabla u &= 0 & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{\sigma} + f &= 0 & \text{in } \Omega, \\ u &= u_0 & \text{on } \Gamma_D, \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= g & \text{on } \Gamma_N. \end{aligned} \tag{2}$$

Next we derive a discrete form for the equations (2). We begin with the definition of the finite element spaces:

$$\begin{aligned} V_h &:= \{v \in L^2(\Omega) \mid v|_T \in \mathcal{P}_k(T) \ \forall T \in \mathcal{T}_h\}, \\ W_h &:= \{\mathbf{v} \in [L^2(\Omega)]^2 \mid \mathbf{v}|_T \in [\mathcal{P}_{k-1}(T)]^2 \ \forall T \in \mathcal{T}_h\}, \end{aligned} \tag{3}$$

in which $\mathcal{P}_k(T)$ denotes the polynomials of order k on T . Multiplying the first equation in (2) with a test function $\boldsymbol{\tau} \in W_h$ and integrating over the domain Ω yields

$$(\boldsymbol{\sigma}, \boldsymbol{\tau})_\Omega - (\nabla u, \boldsymbol{\tau})_\Omega = 0. \tag{4}$$

Multiplying the equation in the middle of (2) with a test function $v \in V_h$ and integrating by parts we get

$$\begin{aligned} (-f, v)_\Omega &= \sum_{T \in \mathcal{T}_h} (\nabla \cdot \boldsymbol{\sigma}, v)_T = \sum_{T \in \mathcal{T}_h} \{-(\boldsymbol{\sigma}, \nabla v)_T + \langle \boldsymbol{\sigma} \cdot \mathbf{n}, v \rangle_{\partial T}\} \\ &= \sum_{T \in \mathcal{T}_h} -(\boldsymbol{\sigma}, \nabla v)_T + \sum_{E \in \mathcal{E}_{\text{int}}} \langle \{\boldsymbol{\sigma} \cdot \mathbf{n}\}, \llbracket v \rrbracket \rangle_E \\ &\quad + \sum_{E \in \mathcal{E}_D} \langle \boldsymbol{\sigma} \cdot \mathbf{n}, v \rangle_E + \sum_{E \in \mathcal{E}_N} \langle g, v \rangle_E, \end{aligned} \tag{5}$$

in which we have employed the continuity of the normal component of the flux and denoted

$$\begin{aligned} \{\boldsymbol{\sigma} \cdot \mathbf{n}\} &:= \frac{1}{2}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2) \cdot \mathbf{n}_1, \\ \llbracket v \rrbracket &:= v_1 - v_2. \end{aligned}$$

Above the subindexes denote the functions on triangles T_1 and T_2 sharing an edge E and \mathbf{n}_1 denotes the outward pointing normal vector of T_1 . Neither the Dirichlet boundary condition nor the continuity is imposed in the solution spaces. Therefore, we have need to enforce them in the variational form. Since the correct solution u is continuous and fulfils $u|_{\Gamma_D} = u_0$, it holds

$$\sum_{E \in \mathcal{E}_{\text{int}}} -\frac{\gamma}{h_E} \langle \llbracket u \rrbracket, \llbracket v \rrbracket \rangle_E = 0 \quad \text{and} \tag{6}$$

$$\sum_{E \in \mathcal{E}_D} -\frac{\gamma}{h_E} \langle u, v \rangle_E = \sum_{E \in \mathcal{E}_D} -\frac{\gamma}{h_E} \langle u_0, v \rangle_E, \tag{7}$$

in which we have introduced the positive stability parameter $\gamma > 0$. Equation (6) enforces the continuity and equation (7) the Dirichlet boundary condition. The model problem is symmetric, and thus it is logical to maintain this also in the variational form. Once again due to the continuity and the Dirichlet boundary conditions we have

$$\sum_{E \in \mathcal{E}_{\text{int}}} \langle \{\boldsymbol{\tau} \cdot \mathbf{n}\}, \llbracket u \rrbracket \rangle_E = 0 \quad \text{and} \quad (8)$$

$$\sum_{E \in \mathcal{E}_D} \langle \boldsymbol{\tau} \cdot \mathbf{n}, u \rangle_E = \sum_{E \in \mathcal{E}_D} \langle \boldsymbol{\tau} \cdot \mathbf{n}, u_0 \rangle_E. \quad (9)$$

Combining the equations (5)–(9) yields the variational form of the problem.

Method. Find $(u_h, \boldsymbol{\sigma}_h) \in V_h \times W_h$ such that

$$a(u, \boldsymbol{\sigma}; v, \boldsymbol{\tau}) = \mathcal{L}(v, \boldsymbol{\tau}) \quad \forall (v, \boldsymbol{\tau}) \in V_h \times W_h, \quad (10)$$

in which

$$\begin{aligned} a(u, \boldsymbol{\sigma}; v, \boldsymbol{\tau}) := & \sum_{T \in \mathcal{T}_h} [(\boldsymbol{\sigma}, \boldsymbol{\tau})_T - (\nabla u, \boldsymbol{\tau})_T - (\boldsymbol{\sigma}, \nabla v)_T] \\ & + \sum_{E \in \mathcal{E}_{\text{int}}} [\langle \{\boldsymbol{\sigma} \cdot \mathbf{n}\}, \llbracket v \rrbracket \rangle_E + \langle \{\boldsymbol{\tau} \cdot \mathbf{n}\}, \llbracket u \rrbracket \rangle_E] \\ & + \sum_{E \in \mathcal{E}_D} [\langle \boldsymbol{\sigma} \cdot \mathbf{n}, v \rangle_E + \langle \boldsymbol{\tau} \cdot \mathbf{n}, u \rangle_E] \\ & - \sum_{E \in \mathcal{E}_D} \frac{\gamma}{h_E} \langle u, v \rangle_E - \sum_{E \in \mathcal{E}_{\text{int}}} \frac{\gamma}{h_E} \langle \llbracket u \rrbracket, \llbracket v \rrbracket \rangle_E \end{aligned} \quad (11)$$

and

$$\mathcal{L}(v, \boldsymbol{\tau}) := (-f, v)_\Omega - (g, v)_{\Gamma_N} + \sum_{E \in \mathcal{E}_D} \langle \boldsymbol{\tau} \cdot \mathbf{n}, u_0 \rangle_E - \sum_{E \in \mathcal{E}_D} \frac{\gamma}{h_E} \langle u_0, v \rangle_E. \quad (12)$$

By the derivation of the variational form it is clear that the proposed method is consistent, i.e. solution to the equations (2) is also the solution to the variational equation (10).

The energy norm of the variational problem is

$$\begin{aligned} \|\|v, \boldsymbol{\tau}\|\|^2 := & \sum_{T \in \mathcal{T}_h} \left[\|\boldsymbol{\tau}\|_{L^2(T)}^2 + \|\nabla v\|_{L^2(T)}^2 \right] \\ & + \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|\llbracket v \rrbracket\|_{L^2(E)}^2 + \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|v\|_{L^2(E)}^2. \end{aligned} \quad (13)$$

Note that the energy norm is mesh dependent. In order to prove the method to be continuous and elliptic in the energy norm we need the following estimate (which is easily proved by scaling).

Lemma 2.1. *There exists a positive constant C_I such that*

$$h_E \|\boldsymbol{\tau}\|_{L^2(\partial T)}^2 \leq C_I \|\boldsymbol{\tau}\|_{L^2(T)}^2 \quad \forall \boldsymbol{\tau} \in W \text{ and } \forall T \in \mathcal{T}_h. \quad (14)$$

With Lemma 2.1 it is straightforward to show that the proposed bilinear form $a(\cdot, \cdot; \cdot, \cdot)$ and the linear functional $\mathcal{L}(\cdot, \cdot)$ are continuous in the energy norm $\|\| \cdot, \cdot \|\|$.

3 The stability and the a priori error estimates

In this section we show that the method is stable for all positive values of the stability parameter γ .

Theorem 3.1. *There exists a positive constant C such that*

$$\sup_{(v, \boldsymbol{\tau}) \in V_h \times W_h} \frac{a(u, \boldsymbol{\sigma}; v, \boldsymbol{\tau})}{\|(v, \boldsymbol{\tau})\|} \geq C \|(u, \boldsymbol{\sigma})\| \quad \forall (u, \boldsymbol{\sigma}) \in V_h \times W_h. \quad (15)$$

Proof. First, we note that

$$a(u, \boldsymbol{\sigma}; -u, \boldsymbol{\sigma}) = \sum_{T \in \mathcal{T}_h} \|\boldsymbol{\sigma}\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}_{\text{int}}} \frac{\gamma}{h_E} \|\llbracket u \rrbracket\|_{L^2(E)}^2 + \sum_{E \in \mathcal{E}_D} \frac{\gamma}{h_E} \|u\|_{L^2(E)}^2. \quad (16)$$

Next, we choose $\boldsymbol{\kappa} \in W_h$ such that $\boldsymbol{\kappa} = \nabla u$, which yields

$$(\boldsymbol{\kappa}, \nabla u)_T = \|\nabla u\|_{L^2(T)}^2 \quad \text{and} \quad \|\boldsymbol{\kappa}\|_{L^2(T)} \leq \|\nabla u\|_{L^2(T)}. \quad (17)$$

Then by the Schwarz inequality we get

$$\begin{aligned} a(u, \boldsymbol{\sigma}; 0, -\boldsymbol{\kappa}) &= \sum_{T \in \mathcal{T}_h} [-(\boldsymbol{\sigma}, \boldsymbol{\kappa})_T + (\nabla u, \boldsymbol{\kappa})_T] - \sum_{E \in \mathcal{E}_{\text{int}}} \langle \{\boldsymbol{\kappa} \cdot \mathbf{n}\}, \llbracket u \rrbracket \rangle_E \\ &\quad - \sum_{E \in \mathcal{E}_D} \langle \boldsymbol{\kappa} \cdot \mathbf{n}, u \rangle_E \\ &\geq \sum_{T \in \mathcal{T}_h} \left[\|\nabla u\|_{L^2(T)}^2 - \|\boldsymbol{\sigma}\|_{L^2(T)} \|\boldsymbol{\kappa}\|_{L^2(T)} \right] \\ &\quad - \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{2} \left[h_E^{1/2} \|\boldsymbol{\kappa}_1 \cdot \mathbf{n}_1\|_{L^2(E)} h_E^{-1/2} \|\llbracket u \rrbracket\|_{L^2(E)} \right. \\ &\quad \left. + h_E^{1/2} \|\boldsymbol{\kappa}_2 \cdot \mathbf{n}_1\|_{L^2(E)} h_E^{-1/2} \|\llbracket u \rrbracket\|_{L^2(E)} \right] \\ &\quad - \sum_{E \in \mathcal{E}_D} h_E^{1/2} \|\boldsymbol{\kappa} \cdot \mathbf{n}\|_{L^2(E)} h_E^{-1/2} \|\llbracket u \rrbracket\|_{L^2(E)}. \end{aligned} \quad (18)$$

For $\delta > 0$ Lemma 2.1, (17) and the Young's inequality give

$$\begin{aligned} a(u, \boldsymbol{\sigma}; 0, -\boldsymbol{\kappa}) &\geq \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 - \frac{1}{2\delta} \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 - \frac{\delta}{2} \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 \\ &\quad - \frac{C_I \delta}{2} \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 - \frac{1}{2\delta} \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|\llbracket u \rrbracket\|_{L^2(E)}^2 \\ &\quad - \frac{C_I \delta}{2} \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 - \frac{1}{2\delta} \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u\|_{L^2(E)}^2 \\ &= \left(1 - \delta \left(\frac{1}{2} + C_I \right) \right) \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 - \frac{1}{2\delta} \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|\llbracket u \rrbracket\|_{L^2(E)}^2 \\ &\quad - \frac{1}{2\delta} \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u\|_{L^2(E)}^2. \end{aligned} \quad (19)$$

Choosing $\delta < 2/(1 + 2C_I)$ yields

$$\begin{aligned} a(u, \boldsymbol{\sigma}; 0, -\boldsymbol{\kappa}) &\geq -C_1 \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + C_2 \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 \\ &\quad - C_3 \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|[u]\|_{L^2(E)}^2 - C_4 \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u\|_{L^2(E)}^2, \end{aligned} \quad (20)$$

with positive constants $C_1, C_2, C_3,$ and C_4 independent of the stability parameter γ . Using the linearity and combining the equations (16) and (20) we get

$$\begin{aligned} a(u, \boldsymbol{\sigma}, -u, \boldsymbol{\sigma} - \epsilon \boldsymbol{\kappa}) &\geq (1 - \epsilon C_1) \|\boldsymbol{\sigma}\|_{L^2(\Omega)}^2 + \epsilon C_2 \sum_{T \in \mathcal{T}_h} \|\nabla u\|_{L^2(T)}^2 \\ &\quad + (\gamma - \epsilon C_3) \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|[u]\|_{L^2(E)}^2 + (\gamma - \epsilon C_4) \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u\|_{L^2(E)}^2. \end{aligned} \quad (21)$$

Choosing the parameter ϵ such that

$$\epsilon > 0, \quad \epsilon < \frac{1}{C_1}, \quad \epsilon < \frac{\gamma}{C_3} \quad \text{and} \quad \epsilon < \frac{\gamma}{C_4}, \quad (22)$$

the inequality (21) gives

$$a(u, \boldsymbol{\sigma}, -u, \boldsymbol{\sigma} - \epsilon \boldsymbol{\kappa}) \geq C_5 \|\|u, \boldsymbol{\sigma}\|\|^2, \quad (23)$$

with a constant $C_5 > 0$. By the definition of $\boldsymbol{\kappa}$ it is clear that

$$\|\| -u, \boldsymbol{\sigma} - \epsilon \boldsymbol{\kappa} \|\| \leq C_6 \|\|u, \boldsymbol{\sigma}\|\|. \quad (24)$$

Substituting the equations (23) and (24) into the left hand side of the equation (15) proves the claim. \square

From the stability and consistency we directly get the a priori estimate. The lower bound $s > 3/2$ is needed in order that $\boldsymbol{\sigma} \cdot \boldsymbol{n} \in L^2(E)$ for all $E \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_{\partial\Omega}$.

Theorem 3.2. *For $u \in H^s(\Omega)$, with $3/2 < s \leq k + 1$ it holds*

$$\|\|u - u_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|\| \leq Ch^{s-1} \|u\|_{H^s(\Omega)}. \quad (25)$$

From above see that the the difference of this method compared to the standard discontinuous Galerkin method is that the lower bound (i.e. zero) is readily available. Let us discuss the implementation of the method a little further. The form of the discrete equations is the following

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{U} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{F} \end{bmatrix}, \quad (26)$$

where $\boldsymbol{\Sigma}$ and \mathbf{U} are the degrees of freedom for $\boldsymbol{\sigma}_h$ and u_h , respectively. Eliminating $\boldsymbol{\Sigma}$, yields the system of equations for \mathbf{U} :

$$(\mathbf{C} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B}) \mathbf{U} = \mathbf{F}. \quad (27)$$

Since the matrix \mathbf{A} corresponds to the part $\sum_{T \in \mathcal{T}_h} (\boldsymbol{\sigma}, \boldsymbol{\tau})_T$ in the bilinear form it is inverted element by element (i.e. condensed) and the cost of this is negligible. For triangular elements the situation is even simpler. An orthogonalization of the basis functions on the reference element gives orthogonal functions on the real element and in this case \mathbf{A} is diagonal. Further, it should be noted that the stability of the method implies that the matrix in (27) is positively definite. The conclusion is hence, that this method is implemented very similarly to the standard discontinuous Galerkin method, but with the advantage that the stability is ensured for all values of the stability parameter.

4 The a posteriori error estimate

In this section we introduce and prove the following a posteriori error estimate for the method.

Theorem 4.1. *There exists a positive constant C such that*

$$\| \|u - u_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \| \| \leq C \left(\sum_{T \in \mathcal{T}_h} \eta_T^2 \right)^{1/2}, \quad (28)$$

in which

$$\begin{aligned} \eta_T^2 &:= h_T^2 \|\nabla \cdot \boldsymbol{\sigma}_h + f\|_{L^2(T)}^2 + \|\boldsymbol{\sigma}_h - \nabla u_h\|_{L^2(T)}^2 \\ &+ h_E \|\llbracket \boldsymbol{\sigma}_h \cdot \mathbf{n} \rrbracket\|_{L^2(\partial T \cap \mathcal{E}_{int})}^2 + \frac{1}{h_E} \|\llbracket u_h \rrbracket\|_{L^2(\partial T \cap \mathcal{E}_{int})}^2 \\ &+ h_E \|\boldsymbol{\sigma} \cdot \mathbf{n} - g\|_{L^2(\partial T \cap \mathcal{E}_N)}^2 + \frac{1}{h_E} \|u_h - u_0\|_{L^2(\partial T \cap \mathcal{E}_D)}^2. \end{aligned} \quad (29)$$

In the proof of Theorem 4.1 we need the following Helmholtz decomposition, cf. [6].

Theorem 4.2. *For every vector $\boldsymbol{\tau} \in [L^2(\Omega)]^2$, with $\boldsymbol{\tau} \cdot \mathbf{n} = g$ on Γ_N , there exists $\psi \in H^1(\Omega)$, with $\psi|_{\Gamma_D} = 0$, and $q \in H^1(\Omega)/\mathbb{R}$, with $\mathbf{curl} q \cdot \mathbf{n}|_{\Gamma_N} = 0$, such that*

$$\boldsymbol{\tau} = \nabla \psi + \mathbf{curl} q \quad \text{and} \quad \|\boldsymbol{\tau}\|_{L^2(\Omega)}^2 = \|\nabla \psi\|_{L^2(\Omega)}^2 + \|\mathbf{curl} q\|_{L^2(\Omega)}^2. \quad (30)$$

The \mathbf{curl} operator, used in the Theorem 4.2, is defined as

$$\mathbf{curl} v := \begin{pmatrix} -\frac{\partial v}{\partial x_2} \\ \frac{\partial v}{\partial x_1} \end{pmatrix}, \quad (31)$$

when $v \in H^1(\Omega)$ and $\Omega \subset \mathbb{R}^2$. We define the tangent to an edge $E \in \mathcal{E}_{int} \cup \mathcal{E}_{\partial\Omega}$ by

$$\mathbf{t} := \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} -n_2 \\ n_1 \end{pmatrix}, \quad (32)$$

in which $\mathbf{n} = (n_1, n_2)$ denotes the outer normal vector of the edge E . The operator $\nabla \times$ is defined by

$$\nabla \times \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} := \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}. \quad (33)$$

Proof. (of Theorem 4.1) Since the exact solution is continuous and fulfils the boundary conditions, we get

$$\begin{aligned} \|u - u_h, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2 &= \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)}^2 + \sum_{T \in \mathcal{T}_h} \|\nabla u - \nabla u_h\|_{L^2(T)}^2 \\ &+ \sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|[[u_h]]\|_{L^2(E)}^2 + \sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u_0 - u_h\|_{L^2(E)}^2. \end{aligned} \quad (34)$$

The two last terms of the equation (34) already belong to the indicator η_T , therefore we only need to estimate the first two terms. We begin with the first term. The definition of the norm and Theorem 4.2 yield

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} &= \sup_{\boldsymbol{\tau} \in [L^2(\Omega)]^2} \frac{(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau})_{\Omega}}{\|\boldsymbol{\tau}\|_{L^2(\Omega)}} \\ &\leq \sup_{\psi} \frac{(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla \psi)_{\Omega}}{\|\nabla \psi\|_{L^2(\Omega)}} + \sup_q \frac{(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{curl} q)_{\Omega}}{\|\nabla q\|_{L^2(\Omega)}}. \end{aligned} \quad (35)$$

Next we turn our attention to the first term in equation (35). Since $\psi \in H^1(\Omega)$ and $\psi|_{\Gamma_D} = 0$, there exists a continuous and piecewise linear Clément interpolation $I_h \psi$ that vanishes on the boundary Γ_D and fulfils

$$\sum_{T \in \mathcal{T}_h} h_T^{-1} \|\psi - I_h \psi\|_{L^2(T)} + \sum_{E \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_N} h_E^{-1/2} \|\psi - I_h \psi\|_{L^2(E)} \leq C \|\nabla \psi\|_{L^2(\Omega)}. \quad (36)$$

From equation

$$a(0, \boldsymbol{\sigma}_h; I_h \psi, 0) = \mathcal{L}(I_h \psi, 0) \quad (37)$$

we get, using integration by parts,

$$\sum_{T \in \mathcal{T}_h} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla I_h \psi)_T = 0. \quad (38)$$

The orthogonality above yields

$$\begin{aligned} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla \psi)_{\Omega} &= \sum_{T \in \mathcal{T}_h} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla(\psi - I_h \psi))_T \\ &= \sum_{T \in \mathcal{T}_h} [-(\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h), \psi - I_h \psi)_T + ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}, \psi - I_h \psi)_{\partial T}] \\ &= \sum_{T \in \mathcal{T}_h} (\nabla \cdot \boldsymbol{\sigma}_h + f, \psi - I_h \psi)_T + \sum_{E \in \mathcal{E}_{\text{int}}} ([[(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) \cdot \mathbf{n}], \psi - I_h \psi)]_E \\ &+ \sum_{E \in \mathcal{E}_N} \langle g - \boldsymbol{\sigma}_h \cdot \mathbf{n}, \psi - I_h \psi \rangle_E. \end{aligned} \quad (39)$$

Applying the Schwarz inequality and (36) gives

$$\begin{aligned}
& (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \nabla \psi)_\Omega \\
& \leq \sum_{T \in \mathcal{T}_h} h_T \|\nabla \cdot \boldsymbol{\sigma}_h + f\|_{L^2(T)} h_T^{-1} \|\psi - I_h \psi\|_{L^2(T)} \\
& \quad + \sum_{E \in \mathcal{E}_{\text{int}}} h_E^{1/2} \|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]\|_{L^2(E)} h_E^{-1/2} \|\psi - I_h \psi\|_{L^2(E)} \\
& \quad + \sum_{E \in \mathcal{E}_N} h_E^{1/2} \|g - \boldsymbol{\sigma}_h \cdot \mathbf{n}\|_{L^2(E)} h_E^{-1/2} \|\psi - I_h \psi\|_{L^2(E)} \\
& \leq C \left\{ \sum_{T \in \mathcal{T}_h} h_T^2 \|\nabla \cdot \boldsymbol{\sigma}_h + f\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}_{\text{int}}} h_E \|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]\|_{L^2(E)}^2 \right. \\
& \quad \left. + \sum_{E \in \mathcal{E}_N} h_E \|g - \boldsymbol{\sigma}_h \cdot \mathbf{n}\|_{L^2(E)}^2 \right\}^{1/2} \|\nabla \psi\|_{L^2(\Omega)}.
\end{aligned} \tag{40}$$

Now, the first term in the equation (35) is bounded by the indicator η_T .

Next, we consider the second term. For the function $q \in H^1(\Omega)/\mathbb{R}$ we construct a piecewise linear interpolate $\pi_h q$ in the following way. Since $\mathbf{curl} q \cdot \mathbf{n}|_{\Gamma_N} = 0$, it follows that $q|_{\Gamma_N}$ is a constant. On Γ_N we thus assign this constant value to $\pi_h q$. For all other vertices we use the Clément construction. The following interpolation estimate holds.

$$\begin{aligned}
& \sum_{T \in \mathcal{T}_h} h_T^{-1} \|q - \pi_h q\|_{L^2(T)} + \sum_{E \in \mathcal{E}_{\text{int}} \cup \mathcal{E}_D} h_E^{-1/2} \|q - \pi_h q\|_{L^2(E)} \\
& \quad + \|\mathbf{curl}(q - \pi_h q)\|_{L^2(\Omega)} \leq C \|\mathbf{curl} q\|_{L^2(\Omega)}.
\end{aligned} \tag{41}$$

From the definition of the variational form yields

$$\sum_{T \in \mathcal{T}_h} (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{curl} \pi_h q)_T = a(0, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h; 0, \mathbf{curl} \pi_h q) = 0, \tag{42}$$

which leads to

$$\begin{aligned}
& (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \mathbf{curl} q)_\Omega = \sum_{T \in \mathcal{T}_h} (\nabla u - \boldsymbol{\sigma}_h, \mathbf{curl}(q - \pi_h q))_T \\
& = \sum_{T \in \mathcal{T}_h} [(\nabla u - \nabla u_h, \mathbf{curl}(q - \pi_h q))_T + (\nabla u_h - \boldsymbol{\sigma}_h, \mathbf{curl}(q - \pi_h q))_T] \\
& := R_1 + R_2.
\end{aligned} \tag{43}$$

Integrating by parts and using the result $\nabla \times \nabla v = 0$ in T we get

$$\begin{aligned}
R_1 & = \sum_{T \in \mathcal{T}_h} [-(\nabla \times \nabla(u - u_h), q - \pi_h q)_T + (\nabla(u - u_h) \cdot \mathbf{t}, q - \pi_h q)_{\partial T}] \\
& = \sum_{E \in \mathcal{E}_{\text{int}}} \langle [\nabla(u - u_h) \cdot \mathbf{t}], q - \pi_h q \rangle_E + \sum_{E \in \mathcal{E}_{\partial\Omega}} \langle \nabla(u - u_h) \cdot \mathbf{t}, q - \pi_h q \rangle_E \\
& = \sum_{E \in \mathcal{E}_{\text{int}}} \langle [\nabla u_h \cdot \mathbf{t}], q - \pi_h q \rangle_E + \sum_{E \in \mathcal{E}_{\partial\Omega}} \langle \nabla(u - u_h) \cdot \mathbf{t}, q - \pi_h q \rangle_E \\
& := S_1 + S_2.
\end{aligned} \tag{44}$$

The Schwarz inequality for sums and the equation (36) lead to

$$\begin{aligned} S_1 &\leq \left(\sum_{E \in \mathcal{E}_{\text{int}}} h_E \|\llbracket \nabla u_h \cdot \mathbf{t} \rrbracket\|_{L^2(E)}^2 \right)^{1/2} \left(\sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|q - \pi_h q\|_{L^2(E)}^2 \right)^{1/2} \\ &\leq C \left(\sum_{E \in \mathcal{E}_{\text{int}}} \frac{1}{h_E} \|\llbracket u - u_h \rrbracket\|_{L^2(E)}^2 \right)^{1/2} \|\mathbf{curl} q\|_{L^2(\Omega)}. \end{aligned} \quad (45)$$

Above, we have also used the estimate

$$\|\llbracket \nabla u \cdot \mathbf{t} \rrbracket\|_{L^2(E)} \leq C \frac{1}{h_E} \|\llbracket u \rrbracket\|_{L^2(E)}. \quad (46)$$

By the equation (45) the term S_1 is bounded by the indicator η_T . Since $q - \pi_h q = 0$ on Γ_N , the estimate (41) gives

$$S_2 \leq C \left(\sum_{E \in \mathcal{E}_D} \frac{1}{h_E} \|u_h - u_0\|_{L^2(E)}^2 \right) \|\mathbf{curl} q\|_{L^2(\Omega)}, \quad (47)$$

since both q and $\pi_h q$ vanish on Γ_N . Combining the equation (45) and (47) shows that the term R_1 is bounded by the indicator η_T . The Schwarz inequality for sums yields

$$\begin{aligned} R_2 &\leq C \left(\sum_{T \in \mathcal{T}_h} \|\nabla u_h - \boldsymbol{\sigma}_h\|_{L^2(T)}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} \|\mathbf{curl}(q - \pi_h q)\|_{L^2(T)}^2 \right)^{1/2} \\ &\leq C \left(\sum_{T \in \mathcal{T}_h} \|\nabla u_h - \boldsymbol{\sigma}_h\|_{L^2(T)}^2 \right)^{1/2} \|\mathbf{curl} q\|_{L^2(\Omega)}. \end{aligned} \quad (48)$$

Now, we have proved that

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)}^2 \leq C \sum_{T \in \mathcal{T}_h} \eta_T^2, \quad (49)$$

and we still need to bound the second term in equation (34). The equation (49) and the definition of η_T lead to

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \|\nabla(u - u_h)\|_{L^2(T)}^2 &= \sum_{T \in \mathcal{T}_h} \|\boldsymbol{\sigma} - \nabla u_h\|_{L^2(T)}^2 \\ &\leq \sum_{T \in \mathcal{T}_h} \left[\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(T)}^2 + \|\boldsymbol{\sigma}_h - \nabla u_h\|_{L^2(T)}^2 \right] \leq C \sum_{T \in \mathcal{T}_h} \eta_T^2. \end{aligned} \quad (50)$$

Combining the equations (34), (49), and (50) completes the proof. \square

Next we give the lower bound estimate. The claim follows from standard techniques, see [9], and the proof is omitted here.

Theorem 4.3. *There exist a positive constant C such that*

$$\eta_T^2 \leq C \left(|u - u_h|_{H^1(\omega_T)}^2 + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\omega_T)}^2 + h_T^2 \|f - f_h\|_{L^2(\omega_T)}^2 + \frac{1}{h_T} \|u - u_h\|_{L^2(\partial T)}^2 + h_T \|g - g_h\|_{L^2(\partial T \cap \mathcal{E}_N)}^2 + \frac{1}{h_T} \|u_0 - u_{0,h}\|_{L^2(\partial T \cap \mathcal{E}_D)}^2 \right). \quad (51)$$

Above we denote with ω_T the union of T and all the elements sharing an edge with T . With f_h , g_h and $u_{0,h}$ we denote the projections of the given data to the discrete space.

5 Numerical results

In this section we investigate the numerical performance of the Nitsche method. We show that the Nitsche method has the optimal convergence rate with respect to the mesh size h . After that we test the adaptive refinement based on the a posteriori error estimate. In all the computations, if not otherwise stated, the stability parameter is set to $\gamma = 1$. A choice which would produce unstable Nitsche method for the non-mixed problem.

For simplicity we choose the unit square as the computational domain; $\Omega = (0, 1) \times (0, 1)$. To have a problem with typical corner singularities we choose our exact solution to be, in polar coordinates,

$$u(r, \theta) = r^\beta \sin(\beta\theta),$$

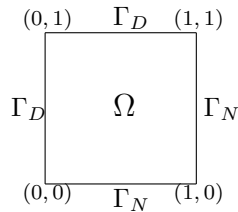
with the parameter $\beta > 0$. With β we can control the regularity of the solution, namely

$$u \in H^{\beta+1-\delta}(\Omega),$$

for all $\delta > 0$. The chosen exact solution u is harmonic ($f = 0$) and we compute the boundary conditions from it, i.e. we define

$$u_0 = u(r, \theta) \quad \text{and} \quad g = \frac{\partial u(r, \theta)}{\partial n} \quad \text{on } \partial\Omega.$$

Our model problem is:



$$\begin{aligned} \boldsymbol{\sigma} - \nabla u &= 0 & \text{on } \partial\Omega \\ \nabla \cdot \boldsymbol{\sigma} &= 0 & \text{on } \partial\Omega \\ u &= u_0 & \text{on } \Gamma_D \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= g & \text{on } \Gamma_N \end{aligned}$$

The convergence results are computed with parameter values $\beta = 0.7, 1.3$ and 2.3 . With this choice the solution belongs to $u \in H^{1.7-\delta}(\Omega)$, $H^{2.3-\delta}(\Omega)$ and $H^{3.3-\delta}(\Omega)$, respectively. Figure 1 shows the solutions for the chosen values of β with both linear and parabolic elements on a mesh of size $h = 0.25$.

In Figure 2 we show the convergence of the error in the energy norm $\|\cdot, \cdot\|$ for both linear and parabolic elements, with different values of β and using a uniform mesh refinement. Both methods perform as expected by the analytical results. Note that the linear elements cannot take advantage of the regularity beyond $u \in H^2(\Omega)$. The numerical values of the slopes are given in the legends of the figure.

Next we test the adaptive mesh refinement based on the a posteriori error distribution. On each step we refine the elements that have larger error than the average elementwise error. The elementwise errors and the average elementwise error are given by the a posteriori error estimator. Figure 3 shows the first three adaptive mesh refinement for linear elements with $\beta = 0.7$. The first mesh has the size $h = 0.25$. We see that the error indicator notices the singularity at the origin and refines there, but that the error at the origin is still dominant after two refinements. In Figure 4 is the same computation with parabolic elements. Again the error singularity at the origin dominates the error.

Figure 5 shows the three adaptive refinement for linear elements and $\beta = 2.3$. We see that the origin is not the dominant part here, instead the error indicator notices the large changes at the boundaries and refines there. In Figure 6 we show the mesh refinements for parabolic elements. Now the origin is again the dominant part of the error since the parabolic elements are able to capture the large but smooth changes at the boundaries with larger elements. Notice also the scales of the error when comparing to Figure 5.

In Figures 3–6 we also have the estimated error and the exact error in the energy norm. We see that both diminish at the same speed, as predicted by the theory.

Acknowledgements

This work has been supported by the Finnish National Graduate School in Engineering Mechanics and TEKES, The National Technology Agency of Finland (project KOMASI, decision number 210622). This paper was completed when the second author was visiting Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie and he would like to thank Prof. Vivette Girault for the kind invitation and the hospitality.

References

- [1] D. ARNOLD, F. BREZZI, B. COCKBURN, AND L. MARINI, *Unified analysis of discontinuous galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.
- [2] F. BASSI AND S. REBAY, *High-order accurate discontinuous finite element solution of the 2D Euler equations*, J. Comput. Phys., 138 (1997), pp. 251–285.

- [3] F. BREZZI, B. COCKBURN, L. MARINI, AND E. SÜLI, *Stabilization mechanism in discontinuous galerkin finite element methods*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3293–3310.
- [4] F. BREZZI, G. MANZINI, D. MARINI, P. PIETRA, AND A. RUSSO, *Discontinuous Galerkin approximations for elliptic problems*, Numer. Methods Partial Differential Equations, 16 (2000), pp. 365–378.
- [5] L. P. FRANCA AND R. STENBERG, *Error analysis of Galerkin least squares methods for the elasticity equations*, SIAM J. Numer. Anal., 28 (1991), pp. 1680–1697.
- [6] V. GIRAULT AND P. RAVIART, *Finite element methods for Navier-Stokes equations, theory and algorithms*, Springer-Verlag, 1986.
- [7] J. NITSCHKE, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, 36 (1970/71), pp. 9–15.
- [8] R. STENBERG, *On some techniques for approximating boundary conditions in the finite element method*, J. Comput. Appl. Math., 63 (1995), pp. 139–148. International Symposium on Mathematical Modelling and Computational Methods Modelling 94 (Prague, 1994).
- [9] R. VERFÜRTH, *A Review of a posteriori error estimation and adaptive mesh refinement techniques*, Wiley, 1996.

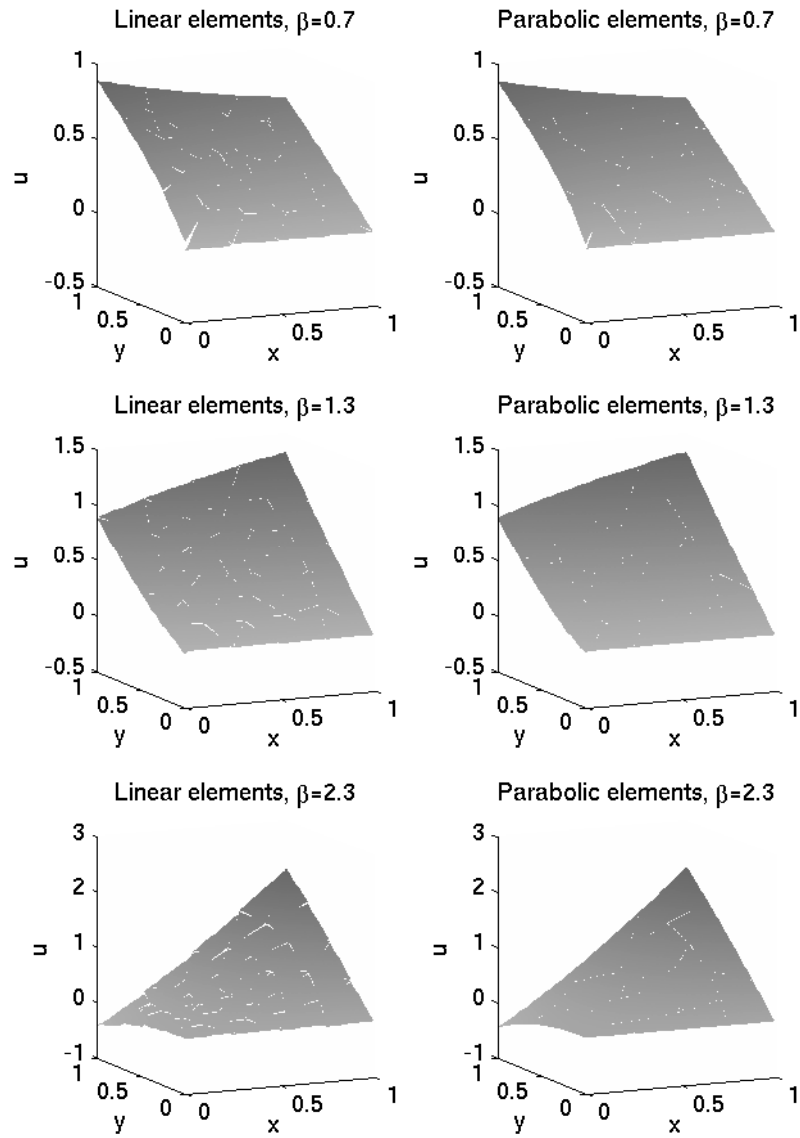


Figure 1: Solutions to the model problem with different values of the parameter β . On the left are the solutions with linear elements and on the right with parabolic elements. From top to bottom β has values 0.7, 1.3 and 2.3. The mesh is of size $h = 0.25$.

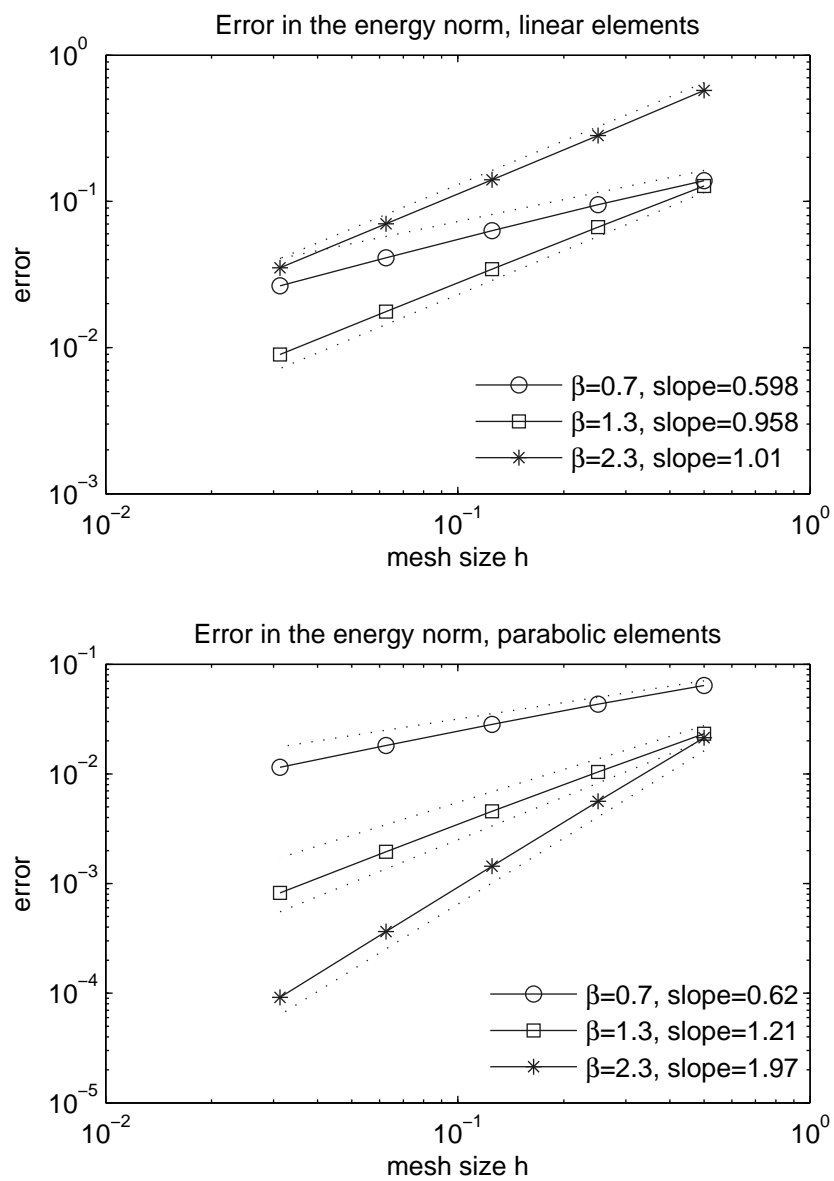


Figure 2: Convergence of the error in the energy norm in uniform mesh refinement for different values of β . The dotted lines are reference convergence rates of $Ch^{0.7}$, Ch , $Ch^{1.3}$ and Ch^2 . The numerical values of the slopes are in the legend.

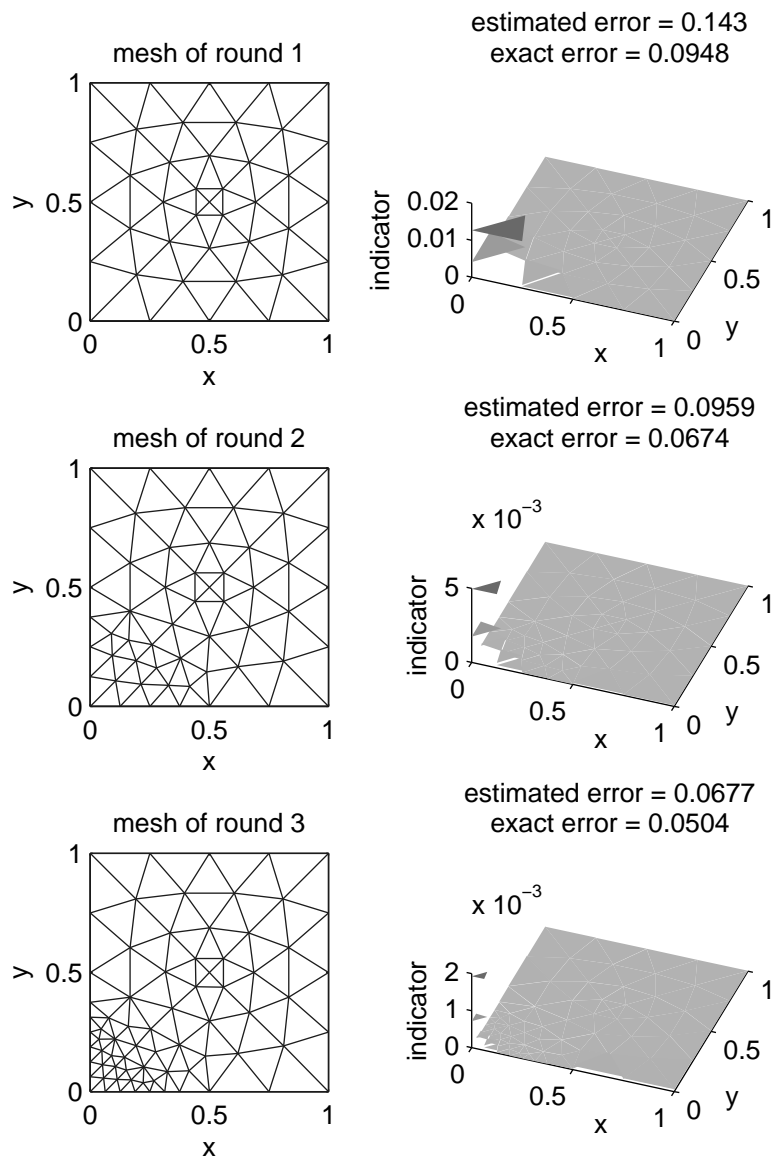


Figure 3: The first three meshes in the adaptive refinement with linear elements and $\beta = 0.7$. On the left the mesh and on the right the distribution of the a posteriori error. In the titles on the right we give the estimated and the exact error in the energy norm. Here the non-regularity at the origin dominates the error.

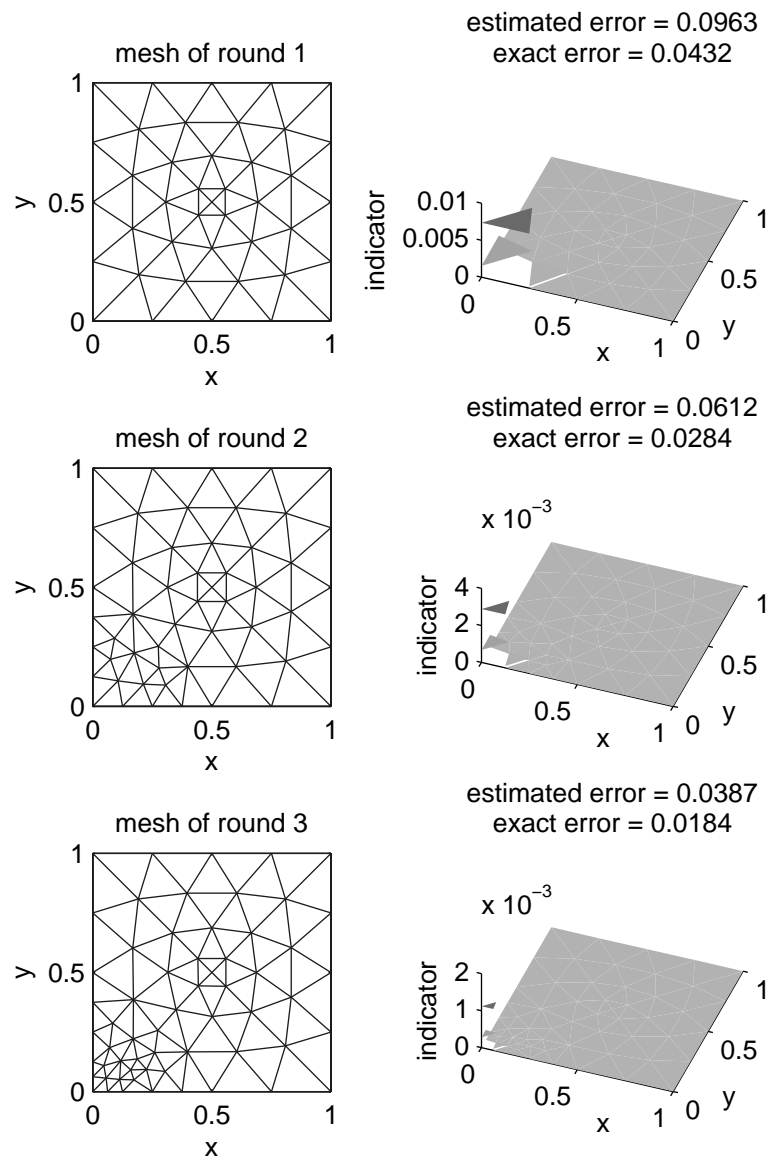


Figure 4: The first three meshes in the adaptive refinement with parabolic elements and $\beta = 0.7$. On the left the mesh and on the right the distribution of the a posteriori error. In the titles on the right we give the estimated and the exact error in the energy norm. Here the non-regularity at the origin dominates the error.

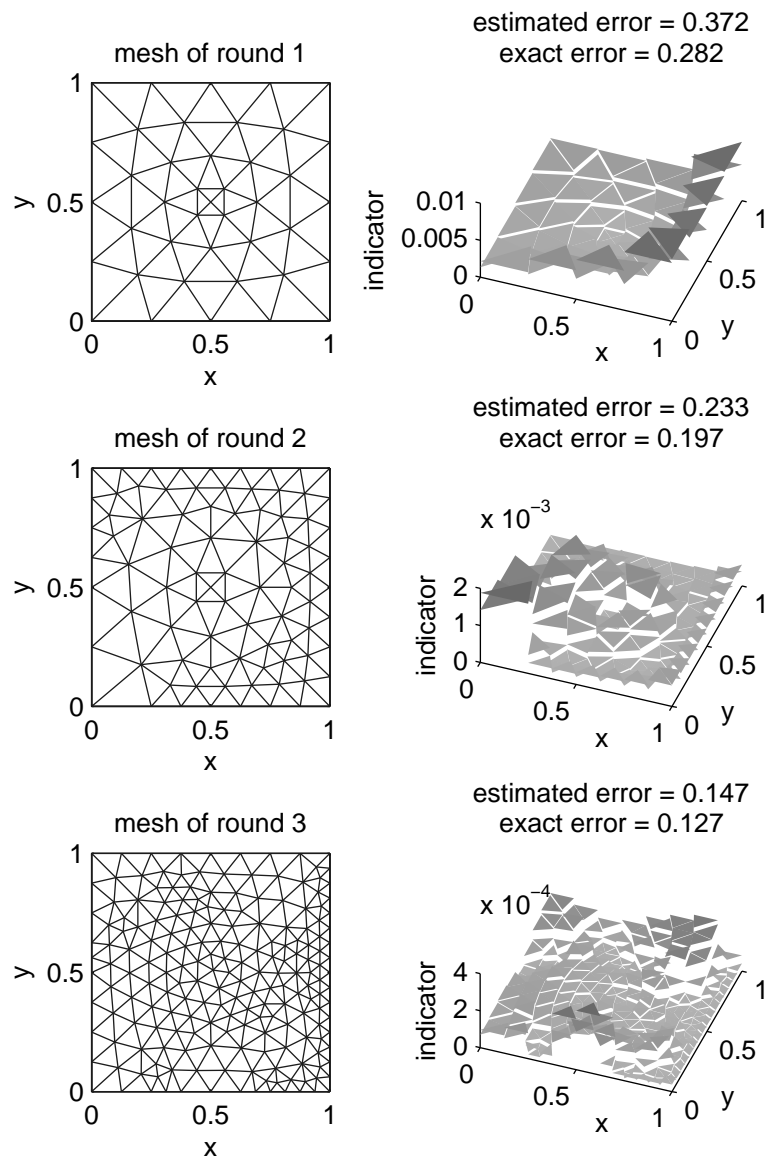


Figure 5: The first three meshes in the adaptive refinement with linear elements and $\beta = 2.3$. On the left the mesh and on the right the distribution of the a posteriori error. In the titles on the right we give the estimated and the exact error in the energy norm. Now, the large changes at the boundaries dominate the error.

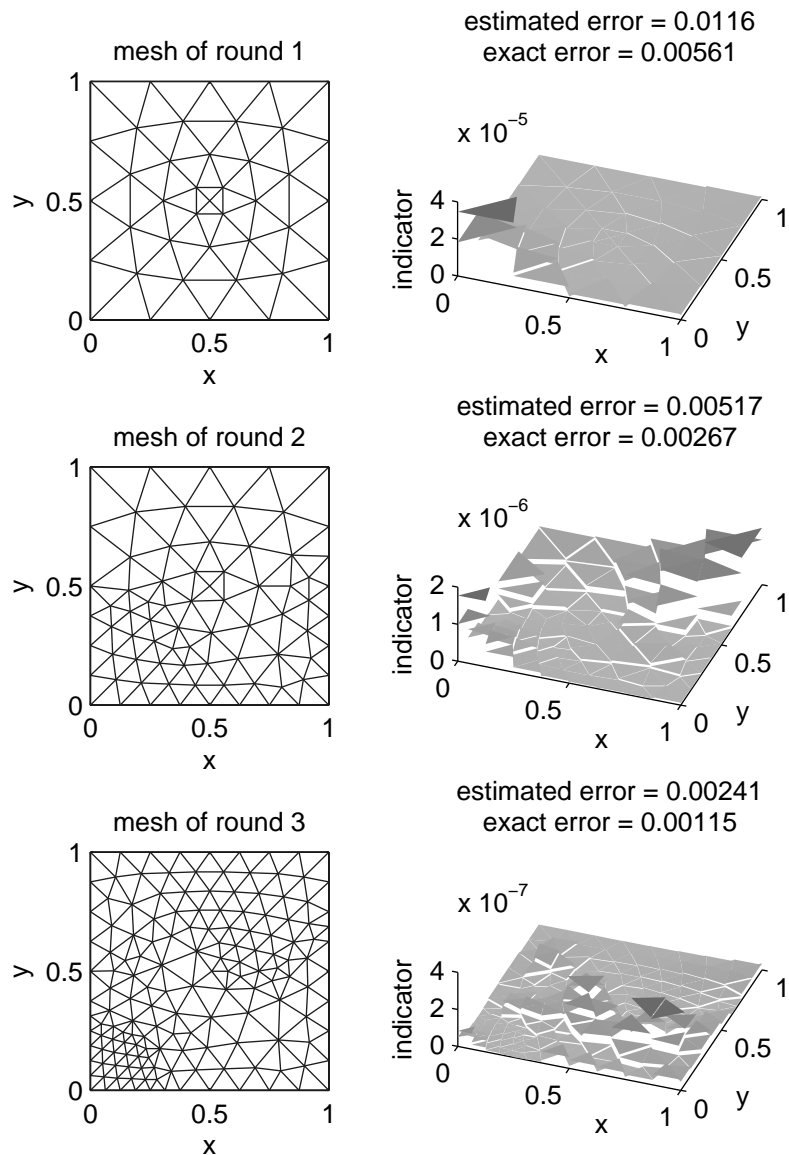


Figure 6: The first three meshes in the adaptive refinement with parabolic elements and $\beta = 2.3$. On the left the mesh and on the right the distribution of the a posteriori error. In the titles on the right we give the estimated and the exact error in the energy norm. Parabolic elements capture the large but smooth changes at the boundaries and the singularity at the origin dominates the error.

(continued from the back cover)

- A534 Jarkko Niiranen
A priori and a posteriori error analysis of finite element methods for plate models
October 2007
- A533 Heikki J. Tikanmäki
Edgeworth expansion for the one dimensional distribution of a Lévy process
September 2007
- A532 Tuomo T. Kuusi
Harnack estimates for supersolutions to a nonlinear degenerate equation
September 2007
- A530 Mika Juntunen , Rolf Stenberg
Nitsches Method for General Boundary Conditions
October 2007
- A529 Mikko Parviainen
Global higher integrability for nonlinear parabolic partial differential equations
in nonsmooth domains
September 2007
- A528 Kalle Mikkola
Hankel and Toeplitz operators on nonseparable Hilbert spaces: further results
August 2007
- A527 Janos Karatson , Sergey Korotov
Sharp upper global a posteriori error estimates for nonlinear elliptic variational
problems
August 2007
- A526 Beirao da Veiga Lourenco , Jarkko Niiranen , Rolf Stenberg
A family of C^0 finite elements for Kirchhoff plates II: Numerical results
May 2007
- A525 Jan Brandts , Sergey Korotov , Michal Krizek
The discrete maximum principle for linear simplicial finite element approxima-
tions of a reaction-diffusion problem
July 2007

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The reports are available at <http://math.tkk.fi/reports/> .

The list of reports is continued inside the backcover.

- A539 Aly A. El-Sabbagh , F.A. Abd El Salam , K. El Nagaar
On the Spectrum of the Symmetric Relations for The Canonical Systems of Differential Equations in Hilbert Space
December 2007
- A538 Aly A. El-Sabbagh , F.A. Abd El Salam , K. El Nagaar
On the Existence of the selfadjoint Extension of the Symmetric Relation in Hilbert Space
December 2007
- A537 Teijo Arponen , Samuli Piipponen , Jukka Tuomela
Kinematic analysis of Bricard's mechanism
November 2007
- A536 Toni Lassila
Optimal damping set of a membrane and topology discovering shape optimization
November 2007
- A535 Esko Valkeila
On the approximation of geometric fractional Brownian motion
October 2007

ISBN 978-951-22-8942-4

ISSN 0784-3143

Teknillinen Korkeakoulu 2007, Espoo, Finland