

First published in

Mika Juntunen and Rolf Stenberg. 2009. Nitsche's method for general boundary conditions. *Mathematics of Computation*, volume 78, number 267, pages 1353-1374. doi:10.1090/S0025-5718-08-02183-2.

Published by the American Mathematical Society.

© 2009 American Mathematical Society (AMS)

Reprinted with permission from American Mathematical Society.

## NITSCHÉ'S METHOD FOR GENERAL BOUNDARY CONDITIONS

MIKA JUNTUNEN AND ROLF STENBERG

ABSTRACT. We introduce a method for treating general boundary conditions in the finite element method generalizing an approach, due to Nitsche (1971), for approximating Dirichlet boundary conditions. We use Poisson's equations as a model problem and prove a priori and a posteriori error estimates. The method is also compared with the traditional Galerkin method. The theoretical results are verified numerically.

### 1. INTRODUCTION

In his classical paper [6] Nitsche discusses techniques for incorporating Dirichlet boundary conditions in the finite element approximation of the model Poisson problem: find  $u$  such that

$$(1.1) \quad -\Delta u = f \quad \text{in } \Omega,$$

$$(1.2) \quad u = u_0 \quad \text{on } \Gamma = \partial\Omega.$$

Before introducing his technique he discusses the penalty method, i.e. the Ritz approximation to the “perturbed” problem in which the Dirichlet boundary condition (1.2) is replaced by the condition

$$(1.3) \quad \frac{\partial u}{\partial n} = \frac{1}{\epsilon}(u_0 - u) \quad \text{on } \Gamma,$$

where  $\epsilon > 0$  is a small parameter. He points out the drawbacks of this approach, i.e. nonconformity, which requires a coupling of the penalty parameter to the mesh size, and the possible ill-conditioning of the discrete system when the penalty parameter is too small (see [2] for a recent survey on this).

If instead of the Dirichlet problem we consider the problem with the boundary condition (1.3), then the solution to the continuous problem converges to the solution of the Dirichlet problem when  $\epsilon \rightarrow 0$ . For the finite element discretization the discrete problem gets more ill-conditioned when  $\epsilon$  approaches zero. In the limit  $\epsilon = 0$ , we have to switch to some other way of imposing the Dirichlet condition, like the conventional approach or Nitsche's technique. The following question arises quite naturally: can we extend Nitsche's method so that it can be used for the whole range of boundary conditions  $\epsilon \geq 0$ ? The purpose of this paper is to give a

---

Received by the editor October 17, 2007, and, in revised form, May 21, 2008.

2000 *Mathematics Subject Classification*. Primary 65N30.

This work was supported by the Finnish National Graduate School in Engineering Mechanics, by the Academy of Finland, and TEKES, the National Technology Agency of Finland.

positive answer to this question. We will consider general boundary conditions and extend Nitsche’s method to cover the whole class of problems.

The outline of the paper is as follows. In the next section we derive the method and show that it is consistent. In Section 3 we prove the ellipticity and derive the a priori error estimates. Section 4 is devoted to the a posteriori error estimates. For the a posteriori estimate we show that it gives both an upper bound and a lower bound to the error. In Section 5 we give a summary of the error analysis of the traditional finite element method. Finally, in Section 6, we show numerical applications of the proposed method and the error estimates and compare them to those obtained with the traditional method.

2. THE METHOD AND ITS CONSISTENCY

We consider the following problem:

$$(2.1) \quad -\Delta u = f \quad \text{in } \Omega,$$

$$(2.2) \quad \frac{\partial u}{\partial n} = \frac{1}{\epsilon}(u_0 - u) + g \quad \text{on } \Gamma,$$

where  $\Omega$  is a bounded domain with polygonal boundary,  $f \in L^2(\Omega)$ ,  $u_0 \in H^{1/2}(\Gamma)$ ,  $g \in L^2(\Gamma)$  and  $\epsilon \in \mathbb{R}$ ,  $0 \leq \epsilon \leq \infty$ . The limiting values of the parameter  $\epsilon$  give the pure Dirichlet and Neumann problems, respectively, i.e.

$$(2.3) \quad \epsilon \rightarrow 0 \Rightarrow u = u_0 \quad \text{on } \Gamma, \quad \epsilon \rightarrow \infty \Rightarrow \frac{\partial u}{\partial n} = g \quad \text{on } \Gamma.$$

For simplicity we consider a shape regular finite element partitioning  $\mathcal{T}_h$  of the domain  $\Omega \subset \mathbb{R}^N$ ,  $N = 2, 3$ , into simplices, i.e. triangles or tetrahedra. This partitioning induces a mesh, denoted by  $\mathcal{G}_h$ , on the boundary  $\Gamma$ . By  $K \in \mathcal{T}_h$  we denote an element of the mesh and by  $E$  we denote one edge or face in  $\mathcal{G}_h$ . By  $h_K$  we denote the diameter of the element  $K \in \mathcal{T}_h$  and by  $h_E$  we denote the diameter of  $E \in \mathcal{G}_h$ . We also define

$$h := \max\{h_K : K \in \mathcal{T}_h\}$$

and

$$V_h := \{v \in H^1(\Omega) : v|_K \in \mathcal{P}_p(K) \quad \forall K \in \mathcal{T}_h\},$$

where  $\mathcal{P}_p(K)$  is the space of polynomials of degree  $p$ . The method is now defined as follows. Here  $\gamma$  is a positive parameter that has to be bounded from above; see Theorem 3.2 below.

**Nitsche’s method.** Find  $u_h \in V_h$  such that

$$(2.4) \quad \mathcal{B}_h(u_h, v) = \mathcal{F}_h(v) \quad \forall v \in V_h,$$

where

$$(2.5) \quad \mathcal{B}_h(u, v) = (\nabla u, \nabla v)_\Omega + \sum_{E \in \mathcal{G}_h} \left\{ -\frac{\gamma h_E}{\epsilon + \gamma h_E} \left[ \left\langle \frac{\partial u}{\partial n}, v \right\rangle_E + \left\langle u, \frac{\partial v}{\partial n} \right\rangle_E \right] + \frac{1}{\epsilon + \gamma h_E} \left\langle u, v \right\rangle_E - \frac{\epsilon \gamma h_E}{\epsilon + \gamma h_E} \left\langle \frac{\partial u}{\partial n}, \frac{\partial v}{\partial n} \right\rangle_E \right\}$$

and

$$(2.6) \quad \mathcal{F}_h(v) = (f, v)_\Omega + \sum_{E \in \mathcal{G}_h} \left\{ \frac{1}{\epsilon + \gamma h_E} \langle u_0, v \rangle_E - \frac{\gamma h_E}{\epsilon + \gamma h_E} \langle u_0, \frac{\partial v}{\partial n} \rangle_E + \frac{\epsilon}{\epsilon + \gamma h_E} \langle g, v \rangle_E - \frac{\epsilon \gamma h_E}{\epsilon + \gamma h_E} \langle g, \frac{\partial v}{\partial n} \rangle_E \right\}.$$

Next we prove the consistency of the proposed method.

**Lemma 2.1.** *The solution  $u$  of the equations (2.1)–(2.2) satisfies*

$$(2.7) \quad \mathcal{B}_h(u, v) = \mathcal{F}_h(v) \quad \forall v \in V_h.$$

*Proof.* Multiplying the differential equation (2.1) with  $v \in V_h$ , integrating over the domain  $\Omega$ , and using Green's formula leads to

$$(2.8) \quad (\nabla u, \nabla v)_\Omega - \langle \frac{\partial u}{\partial n}, v \rangle_\Gamma = (f, v)_\Omega.$$

Next, multiplying the boundary condition (2.2) by  $v$ , and integrating over an element  $E$ , we have

$$(2.9) \quad \epsilon \langle \frac{\partial u}{\partial n}, v \rangle_E + \langle u, v \rangle_E = \langle u_0, v \rangle_E + \epsilon \langle g, v \rangle_E.$$

This gives

$$(2.10) \quad \sum_{E \in \mathcal{G}_h} \frac{1}{\epsilon + \gamma h_E} \left\{ \epsilon \langle \frac{\partial u}{\partial n}, v \rangle_E + \langle u, v \rangle_E \right\} = \sum_{E \in \mathcal{G}_h} \frac{1}{\epsilon + \gamma h_E} \left\{ \langle u_0, v \rangle_E + \epsilon \langle g, v \rangle_E \right\}.$$

Similarly, we obtain

$$(2.11) \quad \sum_{E \in \mathcal{G}_h} -\frac{\gamma h_E}{\epsilon + \gamma h_E} \left\{ \epsilon \langle \frac{\partial u}{\partial n}, \frac{\partial v}{\partial n} \rangle_E + \langle u, \frac{\partial v}{\partial n} \rangle_E \right\} = \sum_{E \in \mathcal{G}_h} -\frac{\gamma h_E}{\epsilon + \gamma h_E} \left\{ \langle u_0, \frac{\partial v}{\partial n} \rangle_E + \epsilon \langle g, \frac{\partial v}{\partial n} \rangle_E \right\}.$$

The equation (2.7) is now the sum of equations (2.8), (2.10), and (2.11). □

The method has two parameters, the stability parameter  $\gamma$  and the problem dependent parameter  $\epsilon$  in the boundary condition. By choosing  $\gamma = 0$  in (2.4) we get:

**The traditional method.** Find  $u_h \in V_h$  such that

$$(2.12) \quad (\nabla u_h, \nabla v)_\Omega + \frac{1}{\epsilon} \langle u_h, v \rangle_\Gamma = (f, v)_\Omega + \frac{1}{\epsilon} \langle u_0, v \rangle_\Gamma + \langle g, v \rangle_\Gamma \quad \forall v \in V_h.$$

This may become ill-conditioned when  $\epsilon > 0$  is small. We will return to this method in Section 5 below.

For the stabilized method with  $\gamma > 0$  we obtain, in the limit  $\epsilon = 0$ ,

$$(2.13) \quad (\nabla u_h, \nabla v)_\Omega - \langle \frac{\partial u_h}{\partial n}, v \rangle_\Gamma - \langle u_h, \frac{\partial v}{\partial n} \rangle_\Gamma + \sum_{E \in \mathcal{G}_h} \frac{1}{\gamma h_E} \langle u_h, v \rangle_E = (f, v)_\Omega - \langle u_0, \frac{\partial v}{\partial n} \rangle_\Gamma + \sum_{E \in \mathcal{G}_h} \frac{1}{\gamma h_E} \langle u_0, v \rangle_E \quad \forall v \in V_h,$$

which is Nitsche’s method [7] applied to the Dirichlet problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= u_0 \quad \text{on } \Gamma. \end{aligned}$$

This is also exactly how the Dirichlet boundary conditions are treated in the Interior Penalty Discontinuous Galerkin method; cf. [1].

When  $\epsilon \rightarrow \infty$  the problem to be solved is the pure Neumann problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ \frac{\partial u}{\partial n} &= g \quad \text{on } \Gamma, \end{aligned}$$

which is approximated by

$$\begin{aligned} (2.14) \quad & (\nabla u_h, \nabla v)_\Omega - \sum_{E \in \mathcal{G}_h} \gamma h_E \left\langle \frac{\partial u_h}{\partial n}, \frac{\partial v}{\partial n} \right\rangle_E \\ &= (f, v)_\Omega + \langle g, v \rangle_\Gamma - \sum_{E \in \mathcal{G}_h} \gamma h_E \left\langle g, \frac{\partial v}{\partial n} \right\rangle_E. \end{aligned}$$

This is the variational form of the Neumann problem with the extra terms

$$- \sum_{E \in \mathcal{G}_h} \gamma h_E \left\langle \frac{\partial u_h}{\partial n}, \frac{\partial v}{\partial n} \right\rangle_E \quad \text{and} \quad - \sum_{E \in \mathcal{G}_h} \gamma h_E \left\langle g, \frac{\partial v}{\partial n} \right\rangle_E,$$

which do not affect the consistency of the method. Note, that the Neumann problem requires that the data satisfy

$$(2.15) \quad (f, 1)_\Omega + \langle g, 1 \rangle_\Gamma = 0,$$

and this condition is not violated in our formulation.

### 3. STABILITY AND A PRIORI ERROR ESTIMATES

In the stability and error analysis we will use the following mesh-dependent norms

$$(3.1) \quad \|v\|_h^2 := \|\nabla v\|_{L^2(\Omega)}^2 + \sum_{E \in \mathcal{G}_h} \frac{1}{\epsilon + h_E} \|v\|_{L^2(E)}^2$$

and

$$(3.2) \quad \|v\|_h^2 := \|v\|_h^2 + \sum_{E \in \mathcal{G}_h} h_E \left\| \frac{\partial v}{\partial n} \right\|_{L^2(E)}^2.$$

In the subspace  $V_h$  these two norms are equivalent. This follows from the well-known estimate below.

**Lemma 3.1.** *There is a positive constant  $C_I$  such that*

$$(3.3) \quad \sum_{E \in \mathcal{G}_h} h_E \left\| \frac{\partial v}{\partial n} \right\|_{L^2(E)}^2 \leq C_I \|\nabla v\|_{L^2(\Omega)}^2 \quad \forall v \in V_h.$$

For the formulation we have the following stability result. Here and in what follows,  $C$  denotes a generic positive constant independent of both the mesh parameter  $h$  and the parameter  $\epsilon$ .

**Theorem 3.2.** *Suppose that  $0 < \gamma < 1/C_I$ . Then there exists a positive constant  $C$  such that*

$$(3.4) \quad \mathcal{B}_h(v, v) \geq C \|v\|_h^2 \quad \forall v \in V_h.$$

*Proof.* First, the Schwarz inequality gives

$$(3.5) \quad \begin{aligned} \mathcal{B}_h(v, v) &= (\nabla v, \nabla v)_\Omega + \sum_{E \in \mathcal{G}_h} \left\{ -\frac{\gamma h_E}{\epsilon + \gamma h_E} \left[ \left\langle \frac{\partial v}{\partial n}, v \right\rangle_E + \left\langle v, \frac{\partial v}{\partial n} \right\rangle_E \right] \right. \\ &\quad \left. + \frac{1}{\epsilon + \gamma h_E} \langle v, v \rangle_E - \frac{\epsilon \gamma h_E}{\epsilon + \gamma h_E} \left\langle \frac{\partial v}{\partial n}, \frac{\partial v}{\partial n} \right\rangle_E \right\} \\ &\geq \|\nabla v\|_{L^2(\Omega)}^2 + \sum_{E \in \mathcal{G}_h} \left\{ -2 \frac{\gamma h_E}{\epsilon + \gamma h_E} \left\| \frac{\partial v}{\partial n} \right\|_{L^2(E)} \|v\|_{L^2(E)} \right. \\ &\quad \left. + \frac{1}{\epsilon + \gamma h_E} \|v\|_{L^2(E)}^2 - \frac{\epsilon \gamma h_E}{\epsilon + \gamma h_E} \left\| \frac{\partial v}{\partial n} \right\|_{L^2(E)}^2 \right\}. \end{aligned}$$

Next, using Young's inequality, with  $\delta > 0$ , we get

$$(3.6) \quad \mathcal{B}_h(v, v) \geq \left( 1 - \frac{1}{\delta} \frac{C_I \gamma^2 h_E}{\epsilon + \gamma h_E} - \frac{C_I \epsilon \gamma}{\epsilon + \gamma h_E} \right) \|\nabla v\|_{L^2(\Omega)}^2 + C \frac{1 - \delta}{\epsilon + \gamma h_E} \|v\|_{L^2(\Gamma)}^2.$$

The second term is positive if  $1 - \delta > 0$  and the first term is positive if

$$(3.7) \quad 1 - \frac{1}{\delta} \frac{C_I \gamma^2 h_E}{\epsilon + \gamma h_E} - \frac{C_I \epsilon \gamma}{\epsilon + \gamma h_E} = \frac{1}{\epsilon + \gamma h_E} \left( \epsilon (1 - C_I \gamma) + \gamma h_E \left( 1 - \frac{C_I \gamma}{\delta} \right) \right) > 0.$$

Hence, we choose  $\delta$  such that  $C_I \gamma < \delta < 1$ . The choice is possible due to the assumption  $\gamma < 1/C_I$ . This shows that  $\mathcal{B}_h(v, v) \geq C \|v\|_h^2$  with  $C > 0$  independent of  $\epsilon$  and  $h$ .  $\square$

In the rest of the paper we will assume that the stability requirement is satisfied, i.e. we make the following assumption.

**Assumption 3.3.** The real parameter  $\gamma$  satisfies  $0 < \gamma < C_I$ .

For the a priori estimate we need the following well-known interpolation estimate.

**Lemma 3.4.** *Suppose that  $u \in H^s(\Omega)$ , with  $3/2 < s \leq p + 1$ . Then it holds that*

$$(3.8) \quad \inf_{v \in V_h} \| |u - v| \|_h \leq C h^{s-1} \|u\|_{H^s(\Omega)}.$$

We then have

**Theorem 3.5.** *For  $u \in H^s(\Omega)$ , with  $3/2 < s \leq p + 1$  it holds that*

$$(3.9) \quad \|u - u_h\|_h \leq C h^{s-1} \|u\|_{H^s(\Omega)}.$$

*Proof.* From the consistency and coercivity, i.e. Lemma 2.1 and Theorem 3.2, we get

$$(3.10) \quad \|u_h - v\|_h^2 \leq C \mathcal{B}_h(u_h - v, u_h - v) \leq C \mathcal{B}_h(u - v, u_h - v) \quad \forall v \in V_h.$$

Using the continuity of the bilinear form and the two norms  $\| \cdot \|_h$  and  $||| \cdot |||_h$  we have the bound

$$(3.11) \quad \mathcal{B}_h(u - v, u_h - v) \leq C |||u - v|||_h \|u_h - v\|_h \quad \forall v \in V_h.$$

Combining equations (3.10) and (3.11) we have

$$(3.12) \quad \|u_h - v\|_h \leq C \|u - v\|_h \quad \forall v \in V_h$$

and the assertion follows by triangle inequality and Lemma 3.4 above. □

#### 4. A POSTERIORI ERROR ESTIMATE

In this section we introduce a residual based a posteriori error estimator for the problem. We will prove that this gives both an upper and a lower bound for the error.

For the proof we will use a mesh  $\mathcal{T}_{h/2}$  obtained from  $\mathcal{T}_h$  by dividing each simplex into  $2^N$ ,  $N = 2, 3$ , equal simplices. The corresponding mesh induced on  $\Gamma$  will be denoted by  $\mathcal{G}_{h/2}$ . By  $V_{h/2}$  we denote the finite element subspace on the refined mesh and  $u_{h/2} \in V_{h/2}$  is the corresponding finite element solution. By  $\mathcal{I}_h$  and  $\mathcal{I}_{h/2}$  we denote the collection of interior edges/faces of elements in  $\mathcal{T}_h$  and  $\mathcal{T}_{h/2}$ , respectively.

The local error indicator is defined as

$$(4.1) \quad \begin{aligned} E_K(u_h)^2 &= h_K^2 \|\Delta u_h + f\|_{L^2(K)}^2 + h_E \left\| \left[ \left[ \frac{\partial u_h}{\partial n} \right] \right] \right\|_{L^2(\partial K \cap \mathcal{I}_h)}^2 \\ &+ \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) + u_h - u_0 \right\|_{L^2(\partial K \cap \Gamma)}^2. \end{aligned}$$

In our analysis we use the following saturation assumption [4].

**Assumption 4.1.** Assume there exists  $\beta < 1$  such that

$$(4.2) \quad \|u - u_{h/2}\|_{h/2} \leq \beta \|u - u_h\|_h,$$

where  $u_{h/2}$  is the solution on the mesh  $\mathcal{T}_{h/2}$ . The mesh  $\mathcal{T}_{h/2}$  is derived by splitting the elements of the mesh  $\mathcal{T}_h$ .

We then have the following result.

**Theorem 4.2.** *Under the Assumptions 3.3 and 4.1 it holds that*

$$(4.3) \quad \|u - u_h\|_h \leq C \left( \sum_{K \in \mathcal{T}_h} E_K(u_h)^2 \right)^{1/2}.$$

*Proof. Step 1.* By the triangle inequality we have

$$(4.4) \quad \|u_{h/2} - u_h\|_{h/2} \geq \|u - u_h\|_{h/2} - \|u - u_{h/2}\|_{h/2} \geq \|u - u_h\|_h - \beta \|u - u_h\|_h$$

and as a consequence of the saturation assumption we have

$$(4.5) \quad \|u - u_h\|_h \leq \frac{1}{1 - \beta} \|u_{h/2} - u_h\|_{h/2}.$$

Hence, it is sufficient to bound  $\|u_{h/2} - u_h\|_{h/2}$ . To this end we use the stability. By Theorem 3.2 there exists  $v \in V_{h/2}$  such that

$$(4.6) \quad \|v\|_{h/2} = 1 \quad \text{and} \quad C \|u_{h/2} - u_h\|_{h/2} \leq \mathcal{B}_{h/2}(u_{h/2} - u_h, v).$$

Let  $\tilde{v} \in V_h$  be the Lagrange interpolate of  $v \in V_{h/2}$ . By scaling arguments one obtains

$$\begin{aligned}
 & \sum_{K \in \mathcal{T}_{h/2}} \left\{ h_K^{-2} \|v - \tilde{v}\|_{L^2(K)}^2 + h_E^{-1} \|v - \tilde{v}\|_{L^2(\partial K)}^2 \right\} \\
 (4.7) \quad & + \sum_{E \in \mathcal{G}_{h/2}} \left\{ \frac{1}{\epsilon + h_E} \|v - \tilde{v}\|_{L^2(E)}^2 + h_E \left\| \frac{\partial(v - \tilde{v})}{\partial n} \right\|_{L^2(E)}^2 \right\} \\
 & \leq C \|v\|_{h/2}^2 \leq C.
 \end{aligned}$$

To simplify the notation we define  $w := v - \tilde{v}$  and the above estimate gives

$$\begin{aligned}
 (4.8) \quad & \left( \sum_{K \in \mathcal{T}_{h/2}} h_K^{-2} \|w\|_{L^2(K)}^2 \right)^{1/2} \leq C, \quad \left( \sum_{K \in \mathcal{T}_{h/2}} h_E^{-1} \|w\|_{L^2(\partial K)}^2 \right)^{1/2} \leq C \\
 & \text{and} \quad \left( \sum_{E \in \mathcal{G}_{h/2}} h_E \left\| \frac{\partial w}{\partial n} \right\|_{L^2(E)}^2 \right)^{1/2} \leq C.
 \end{aligned}$$

In (4.6) we split the right-hand side into two parts:

$$(4.9) \quad \mathcal{B}_{h/2}(u_{h/2} - u_h, v) \leq \mathcal{B}_{h/2}(u_{h/2} - u_h, w) + \mathcal{B}_{h/2}(u_{h/2} - u_h, \tilde{v}) =: W_1 + W_2.$$

We will bound the terms  $W_1$  and  $W_2$  separately.

*Step 2.* Since  $w \in V_{h/2}$ , it holds that

$$(4.10) \quad \mathcal{B}_{h/2}(u_{h/2}, w) = \mathcal{F}_{h/2}(w),$$

and we have

$$\begin{aligned}
 (4.11) \quad W_1 &= \mathcal{F}_{h/2}(w) - \mathcal{B}_{h/2}(u_h, w) \\
 &= \left\{ (f, w)_\Omega - (\nabla u_h, \nabla w)_\Omega + \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left\langle \frac{\partial u_h}{\partial n}, w \right\rangle_E \right\} \\
 &+ \sum_{E \in \mathcal{G}_{h/2}} \frac{1}{\epsilon + \gamma h_E} \left[ \langle u_0 - u_h, w \rangle_E + \epsilon \langle g, w \rangle_E \right] \\
 &+ \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left[ \langle u_h - u_0, \frac{\partial w}{\partial n} \rangle_E + \epsilon \left\langle \frac{\partial u_h}{\partial n} - g, \frac{\partial w}{\partial n} \right\rangle_E \right].
 \end{aligned}$$

Integrating by parts on each  $K \in \mathcal{T}_{h/2}$  gives

$$\begin{aligned}
 (4.12) \quad (f, w)_\Omega - (\nabla u_h, \nabla w)_\Omega &= \sum_{K \in \mathcal{T}_{h/2}} (f + \Delta u_h, w)_K - \sum_{E \in \mathcal{T}_{h/2}} \left\langle \left[ \frac{\partial u_h}{\partial n} \right], w \right\rangle_E \\
 &- \sum_{E \in \mathcal{G}_{h/2}} \left\langle \frac{\partial u_h}{\partial n}, w \right\rangle_E.
 \end{aligned}$$

Rearranging terms we thus have

$$(4.13) \quad W_1 = R_1 + R_2 + R_3,$$



with

$$(4.14) \quad R_1 = \sum_{K \in \mathcal{T}_{h/2}} (f + \Delta u_h, w)_K - \sum_{E \in \mathcal{I}_{h/2}} \left\langle \left[ \frac{\partial u_h}{\partial n} \right], w \right\rangle_E,$$

$$(4.15) \quad R_2 = \sum_{E \in \mathcal{G}_{h/2}} \left( \frac{\gamma h_E}{\epsilon + \gamma h_E} - 1 \right) \left\langle \frac{\partial u_h}{\partial n}, w \right\rangle_E \\ + \sum_{E \in \mathcal{G}_{h/2}} \frac{1}{\epsilon + \gamma h_E} \left[ \langle u_0 - u_h, w \rangle_E + \epsilon \langle g, w \rangle_E \right],$$

and

$$(4.16) \quad R_3 = \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left[ \langle u_h - u_0, \frac{\partial w}{\partial n} \rangle_E + \epsilon \left\langle \frac{\partial u_h}{\partial n} - g, \frac{\partial w}{\partial n} \right\rangle_E \right].$$

The first term is estimated using Schwarz inequality and (4.8)

$$(4.17) \quad R_1 \leq \sum_{K \in \mathcal{T}_{h/2}} \|f + \Delta u_h\|_{L^2(K)} \|w\|_{L^2(K)} + \sum_{E \in \mathcal{I}_{h/2}} \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)} \|w\|_{L^2(E)} \\ \leq \left( \sum_{K \in \mathcal{T}_{h/2}} h_K^2 \|f + \Delta u_h\|_{L^2(K)} \right)^{1/2} \left( \sum_{K \in \mathcal{T}_{h/2}} h_K^{-2} \|w\|_{L^2(K)} \right)^{1/2} \\ + \left( \sum_{E \in \mathcal{I}_{h/2}} h_E \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)} \right)^{1/2} \left( \sum_{E \in \mathcal{I}_{h/2}} h_E^{-1} \|w\|_{L^2(E)} \right)^{1/2} \\ \leq C \left[ \left( \sum_{K \in \mathcal{T}_{h/2}} h_K^2 \|f + \Delta u_h\|_{L^2(K)} \right)^{1/2} + \left( \sum_{E \in \mathcal{I}_{h/2}} h_E \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)} \right)^{1/2} \right].$$

Adding the terms in  $R_2$ , using Schwarz inequality and the estimate (4.8) gives

$$(4.18) \quad R_2 = \sum_{E \in \mathcal{G}_{h/2}} \frac{1}{\epsilon + \gamma h_E} \langle u_0 - u_h + \epsilon g - \epsilon \frac{\partial u_h}{\partial n}, w \rangle_E \\ \leq \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2} \\ \cdot \left( \sum_{E \in \mathcal{G}_{h/2}} h_E^{-1} \|w\|_{L^2(E)}^2 \right)^{1/2} \\ \leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2}.$$

For the third term we similarly get

$$\begin{aligned}
 R_3 &= \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left[ \left\langle u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right), \frac{\partial w}{\partial n} \right\rangle_E \right] \\
 &\leq \gamma \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| \left\langle u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\| \right\|^2 \right)^{1/2} \\
 (4.19) \quad &\cdot \left( \sum_{E \in \mathcal{G}_{h/2}} h_E \left\| \frac{\partial w}{\partial n} \right\|_E^2 \right)^{1/2} \\
 &\leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| \left\langle u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\| \right\|^2 \right)^{1/2}.
 \end{aligned}$$

Now we have bounded the term  $W_1$ , i.e. we have

$$(4.20) \quad W_1 \leq C \left( \sum_{K \in \mathcal{T}_{h/2}} E_K(u_h)^2 \right)^{1/2}.$$

*Step 3.* Next, we prove the same upper bound to term  $W_2$  of equation (4.9). To obtain the upper bound we need the following bounds:

$$(4.21) \quad \|\tilde{v}\|_{h/2} + \left( \sum_{E \in \mathcal{G}_{h/2}} h_E \left\| \frac{\partial \tilde{v}}{\partial n} \right\|_{L^2(E)}^2 \right)^{1/2} \leq C \|v\|_{h/2} \leq C,$$

which follow from (4.7) and (4.6). Below, for clarity, we will denote by  $E$  an element in  $\mathcal{G}_{h/2}$  and by  $F$  an element in  $\mathcal{G}_h$ . Using the relation  $\mathcal{B}_h(u_h, \tilde{v}) - \mathcal{F}_h(\tilde{v}) = 0$  and rearranging terms we obtain

$$\begin{aligned}
 W_2 &= \mathcal{F}_{h/2}(\tilde{v}) - \mathcal{B}_{h/2}(u_h, \tilde{v}) = \mathcal{F}_{h/2}(\tilde{v}) - \mathcal{F}_h(\tilde{v}) + \mathcal{B}_h(u_h, \tilde{v}) - \mathcal{B}_{h/2}(u_h, \tilde{v}) \\
 &= \left[ \sum_{E \in \mathcal{G}_{h/2}} \frac{1}{\epsilon + \gamma h_E} \langle u_0 - u_h + \epsilon g, \tilde{v} \rangle_E \right] \\
 &\quad + \left[ - \sum_{F \in \mathcal{G}_h} \frac{1}{\epsilon + \gamma h_F} \langle u_0 - u_h + \epsilon g, \tilde{v} \rangle_F \right] \\
 (4.22) \quad &\quad + \left[ \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left\langle \frac{\partial u_h}{\partial n}, \tilde{v} \right\rangle_E - \sum_{F \in \mathcal{G}_h} \frac{\gamma h_F}{\epsilon + \gamma h_F} \left\langle \frac{\partial u_h}{\partial n}, \tilde{v} \right\rangle_F \right] \\
 &\quad + \left[ \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \left\langle u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right), \frac{\partial \tilde{v}}{\partial n} \right\rangle_E \right] \\
 &\quad + \left[ - \sum_{F \in \mathcal{G}_h} \frac{\gamma h_F}{\epsilon + \gamma h_F} \left\langle u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right), \frac{\partial \tilde{v}}{\partial n} \right\rangle_F \right] \\
 &= T_1 + T_2 + T_3 + T_4 + T_5.
 \end{aligned}$$

Since  $u_h$  has the same values on both meshes  $\mathcal{T}_{h/2}$  and  $\mathcal{T}_h$ , we can write the term  $T_3$  as follows:

$$\begin{aligned}
 T_3 &= \left[ \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{\epsilon + \gamma h_E} \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_E - \sum_{F \in \mathcal{G}_h} \frac{\gamma h_F}{\epsilon + \gamma h_F} \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_F \right] \\
 (4.23) \quad &= \left[ \sum_{E \in \mathcal{G}_{h/2}} \left( \frac{\gamma h_E}{\epsilon + \gamma h_E} - 1 \right) \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_E - \sum_{F \in \mathcal{G}_h} \left( \frac{\gamma h_F}{\epsilon + \gamma h_F} - 1 \right) \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_F \right] \\
 &= - \sum_{E \in \mathcal{G}_{h/2}} \frac{\epsilon}{\epsilon + \gamma h_E} \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_E + \sum_{F \in \mathcal{G}_h} \frac{\epsilon}{\epsilon + \gamma h_F} \langle \frac{\partial u_h}{\partial n}, \tilde{v} \rangle_F.
 \end{aligned}$$

Next, adding  $T_1$ ,  $T_2$  and  $T_3$ , and using the fact that  $h_F = 2h_E$ , for  $E \subset F$ , gives

$$\begin{aligned}
 T_1 + T_2 + T_3 &= \sum_{E \in \mathcal{G}_{h/2}} \frac{1}{\epsilon + \gamma h_E} \langle u_0 - u_h + \epsilon(g - \frac{\partial u_h}{\partial n}), \tilde{v} \rangle_E \\
 &\quad - \sum_{F \in \mathcal{G}_h} \frac{1}{\epsilon + \gamma h_F} \langle u_0 - u_h + \epsilon(g - \frac{\partial u_h}{\partial n}), \tilde{v} \rangle_F \\
 &= \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{(\epsilon + \gamma h_E)(\epsilon + 2\gamma h_E)} \langle u_0 - u_h + \epsilon(g - \frac{\partial u_h}{\partial n}), \tilde{v} \rangle_E \\
 (4.24) \quad &\leq C \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^{3/2}} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)} \\
 &\quad \cdot \frac{1}{(\epsilon + \gamma h_E)^{1/2}} \|\tilde{v}\|_{L^2(E)} \\
 &\leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E^2}{(\epsilon + \gamma h_E)^3} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2} \|\tilde{v}\|_h \\
 &\leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2}.
 \end{aligned}$$

The terms  $T_4$  and  $T_5$  of the equation (4.22) are the same terms on different meshes and the proofs are exactly the same for both of them. For brevity we show the proof only for  $T_4$

$$\begin{aligned}
 T_4 &\leq C \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E h_E^{-1/2}}{\epsilon + \gamma h_E} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)} h_E^{1/2} \left\| \frac{\partial \tilde{v}}{\partial n} \right\|_{L^2(E)} \\
 &\leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{\gamma h_E}{(\epsilon + \gamma h_E)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2} \\
 (4.25) \quad &\quad \cdot \left( \sum_{E \in \mathcal{G}_{h/2}} h_E \left\| \frac{\partial \tilde{v}}{\partial n} \right\|_{L^2(E)}^2 \right)^{1/2} \\
 &\leq C \left( \sum_{E \in \mathcal{G}_{h/2}} \frac{h_E}{(\epsilon + \gamma h_E)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(E)}^2 \right)^{1/2},
 \end{aligned}$$

where last line follows from the bound of the interpolant; see equation (4.21). For  $T_5$  we get

$$(4.26) \quad T_5 \leq C \left( \sum_{F \in \mathcal{G}_h} \frac{h_F}{(\epsilon + \gamma h_F)^2} \left\| u_h - u_0 + \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) \right\|_{L^2(F)}^2 \right)^{1/2}.$$

Now we have also bounded the term  $W_2$ , i.e. we have

$$(4.27) \quad W_2 \leq C \left[ \left( \sum_{K \in \mathcal{T}_{h/2}} E_K(u_h)^2 \right)^{1/2} + \left( \sum_{K \in \mathcal{T}_h} E_K(u_h)^2 \right)^{1/2} \right].$$

Since  $u_h \in V_h$  has the same values on both  $\mathcal{T}_{h/2}$  and  $\mathcal{T}_h$  we have

$$(4.28) \quad \sum_{K \in \mathcal{T}_{h/2}} E_K(u_h)^2 \leq C \sum_{K \in \mathcal{T}_h} E_K(u_h)^2.$$

The assertion now follows by combining (4.5), (4.6), (4.9), (4.20) and (4.27).

Let us next discuss the estimator. When  $\epsilon = 0$ , i.e. for the pure Dirichlet problem, we get

$$(4.29) \quad \begin{aligned} E_K(u_h)^2 &= h_K^2 \|\Delta u_h + f\|_{L^2(K)}^2 + \sum_{E \subset \partial K \cap \mathcal{I}_h} h_E \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)}^2 \\ &+ \sum_{E \subset \partial K \cap \Gamma} \frac{1}{h_E} \|u_h - u_0\|_{L^2(E)}^2, \end{aligned}$$

which is the estimator of Nitsche's method for the Dirichlet boundary value problem; see [3]. Note also that the error is measured in the norm

$$(4.30) \quad \|v\|_h^2 = \|\nabla v\|_{L^2(\Omega)}^2 + \sum_{E \in \mathcal{G}_h} \frac{1}{h_E} \|v\|_{L^2(E)}^2.$$

The other limit,  $\epsilon \rightarrow \infty$ , leads to

$$(4.31) \quad \begin{aligned} E_K(u_h)^2 &= h_K^2 \|\Delta u_h + f\|_{L^2(K)}^2 + \sum_{E \subset \partial K \cap \mathcal{I}_h} h_E \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)}^2 \\ &+ \sum_{E \subset \partial K \cap \Gamma} h_E \left\| \frac{\partial u_h}{\partial n} - g \right\|_{L^2(E)}^2, \end{aligned}$$

which is the traditional a posteriori estimator of the Neumann problem with the error measured in the  $H^1(\Omega)$ -seminorm

$$(4.32) \quad \|v\|_h^2 = \|\nabla v\|_{L^2(\Omega)}^2.$$

These remarks show that the a posteriori estimate holds for all values of the parameter  $\epsilon$ , even the limit values give the correct and numerically stable a posteriori estimate.

Finally, we prove the efficiency of the a posteriori estimate. For the proof we use and adopt established techniques using test functions with local support. We let  $\Psi_E$  be the  $N$ -th degree polynomial which has the support  $\omega_E$  on the element with  $E$  as an edge/face and is normalized such that  $0 \leq \Psi_E \leq 1 = \max \Psi_E$ . For the edges we also need an extension operator  $\mathcal{E}$  from the edge  $E$  to the elements sharing  $E$ , i.e.

$$\mathcal{E} : L^2(E) \rightarrow L^2(\omega_E).$$

On the boundary  $\partial\Omega$  we assume that  $\Psi_E$  and  $\mathcal{E}$  operate in the obvious way, i.e. they only extend towards the interior of the domain  $\Omega$ . For the bubble functions and the extension operator the following estimates hold; see e.g. [8].

**Lemma 4.3.** *Let  $\mathcal{T}_h$  be a shape-regular mesh. Then there exists  $C > 0$  such that*

$$(4.33) \quad \|\Psi_E^{1/2} p_E\|_{L(E)} \geq C \|p_E\|_{L^2(E)},$$

$$(4.34) \quad Ch_K^{1/2} \|p_E\|_{L^2(E)} \leq \|\Psi_E \mathcal{E} p_E\|_{L^2(K)} \leq Ch_K^{1/2} \|p_E\|_{L^2(E)},$$

$$(4.35) \quad \|\nabla(\Psi_E \mathcal{E} p_E)\|_{L^2(K)} \leq Ch_K^{-1} \|\Psi_E \mathcal{E} p_E\|_{L^2(K)},$$

for all  $p_E \in \mathcal{P}_p(E)$ ,  $K \in \mathcal{T}_h$  and  $E \subset \partial K$ .

We now have the following local bounds.

**Theorem 4.4.** *The elementwise estimator  $E_K(u_h)$ , defined in equation (4.1), also fulfills*

$$(4.36) \quad \begin{aligned} E_K(u_h)^2 \leq C & \left( |u - u_h|_{H^1(\omega_K)}^2 + h_K^2 \|f - f_h\|_{L^2(\omega_K)}^2 + \sum_{E \subset \partial K \cap \Gamma} \frac{1}{\epsilon + h_E} \|u - u_h\|_{L^2(E)}^2 \right. \\ & \left. + \sum_{E \subset \partial K \cap \Gamma} \frac{h_E}{(\epsilon + h_E)^2} \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)}^2 \right), \end{aligned}$$

where  $f_h$ ,  $u_{0,h}$  and  $g_h$  are approximations in  $V_h$  of the given data, and  $\omega_K$  is the domain of element  $K$  and all elements sharing an edge/face with  $K$ .

*Proof.* We will consider the upper bound for each term of the estimator  $E_K(u_h)$ , equation (4.1), separately.

For the terms  $R_K := \Delta u_h + f$  and  $R_E := \llbracket \frac{\partial u_h}{\partial n} \rrbracket$  we have the well-known bounds [8]:

$$(4.37) \quad h_K \|R_K\|_{L^2(\omega_K)} \leq C \left( |u - u_h|_{H^1(\omega_K)} + h_K \|f - f_h\|_{L^2(\omega_K)} \right)$$

and

$$(4.38) \quad h_E^{1/2} \|R_E\|_{L^2(E)} \leq C \left( |u - u_h|_{H^1(\omega_E)} + h_K \|f - f_h\|_{L^2(\omega_E)} \right).$$

Therefore, we only give the proof for the last term

$$(4.39) \quad R_\Gamma = \epsilon \left( \frac{\partial u_h}{\partial n} - g \right) + u_h - u_0.$$

We denote

$$R_{\Gamma,\text{red}} = \epsilon \left( \frac{\partial u_h}{\partial n} - g_h \right) + u_h - u_{0,h}, \quad \hat{w}_\Gamma = \Psi_E \mathcal{E} R_{\Gamma,\text{red}} \quad \text{and} \quad w_\Gamma = \Psi_E R_{\Gamma,\text{red}}.$$

With the triangle inequality we get

$$(4.40) \quad \|R_\Gamma\|_{L^2(E)} \leq \|R_{\Gamma,\text{red}}\|_{L^2(E)} + \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)}.$$

Lemma 4.3 and the identities

$$(R_K, \hat{w}_\Gamma)_K = (\nabla(u - u_h), \nabla \hat{w}_\Gamma)_K + \left\langle \frac{\partial}{\partial n} (u_h - u), \hat{w}_\Gamma \right\rangle_E,$$

and

$$\epsilon \left( \frac{\partial u}{\partial n} - g \right) + (u_0 - u) = 0,$$

lead to

$$\begin{aligned}
 (4.41) \quad & C \|R_{\Gamma,\text{red}}\|_{L^2(E)}^2 \leq \|\Psi_E^{1/2} R_{\Gamma,\text{red}}\|_{L^2(E)}^2 = \langle R_{\Gamma,\text{red}}, w_\Gamma \rangle_E \\
 & = \epsilon \langle \frac{\partial}{\partial n} (u_h - u), \hat{w}_\Gamma \rangle_E + \langle u_h - u, w_\Gamma \rangle_E + \langle \epsilon(g - g_h) + u_0 - u_{0,h}, w_\Gamma \rangle_E \\
 & = \epsilon (R_K, \hat{w}_\Gamma)_K + \epsilon (\nabla(u_h - u), \nabla \hat{w}_\Gamma)_K + \langle u_h - u, w_\Gamma \rangle_E \\
 & \quad + \langle \epsilon(g - g_h) + u_0 - u_{0,h}, w_\Gamma \rangle_E \\
 & \leq \epsilon \|R_K\|_{L^2(K)} \|\hat{w}_\Gamma\|_{L^2(K)} + \epsilon \|\nabla(u - u_h)\|_{L^2(K)} \|\nabla \hat{w}_\Gamma\|_{L^2(K)} \\
 & \quad + \|u_h - u\|_{L^2(E)} \|w_\Gamma\|_{L^2(E)} + \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)} \|w_\Gamma\|_{L^2(E)} \\
 & \leq C \left( \epsilon h_K^{1/2} \|R_K\|_{L^2(K)} + \epsilon h_K^{-1/2} |u - u_h|_{H^1(K)} + \|u - u_h\|_{L^2(E)} \right. \\
 & \quad \left. + \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)} \right) \|R_{\Gamma,\text{red}}\|_{L^2(E)}.
 \end{aligned}$$

Multiplying equation (4.41) with  $\frac{h_E^{1/2}}{\epsilon + h_E}$  and using the bound (4.37) for  $\|R_K\|_{L^2(K)}$  gives

$$\begin{aligned}
 (4.42) \quad & \frac{h_E^{1/2}}{\epsilon + h_E} \|R_{\Gamma,\text{red}}\|_{L^2(E)} \\
 & \leq C \left( \frac{\epsilon}{\epsilon + h_E} |u - u_h|_{H^1(K)} + \frac{\epsilon h_E}{\epsilon + h_E} \|f - f_h\|_{L^2(K)} \right. \\
 & \quad \left. + \frac{h_E^{1/2}}{\epsilon + h_E} \|u - u_h\|_{L^2(E)} + \frac{h_E^{1/2}}{\epsilon + h_E} \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)} \right) \\
 & \leq C \left( |u - u_h|_{H^1(K)} + h_K \|f - f_h\|_{L^2(K)} \right. \\
 & \quad \left. + \frac{1}{(\epsilon + h_E)^{1/2}} \|u - u_h\|_{L^2(E)} + \frac{h_E^{1/2}}{\epsilon + h_E} \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)} \right).
 \end{aligned}$$

Combining equations (4.40) and (4.42) gives the following bound to  $R_\Gamma$ :

$$\begin{aligned}
 (4.43) \quad & \frac{h_E^{1/2}}{\epsilon + h_E} \|R_\Gamma\|_{L^2(E)} \leq C \left( |u - u_h|_{H^1(K)} + h_K \|f - f_h\|_{L^2(K)} \right. \\
 & \quad \left. + \frac{1}{(\epsilon + h_E)^{1/2}} \|u - u_h\|_{L^2(E)} + \frac{h_E^{1/2}}{\epsilon + h_E} \|\epsilon(g - g_h) + u_0 - u_{0,h}\|_{L^2(E)} \right).
 \end{aligned}$$

All terms in equation (4.1) are now bounded separately, hence combining equations (4.37), (4.38), and (4.43) completes the proof.  $\square$

### 5. THE TRADITIONAL METHOD

In this section we give a short review of the error analysis of the traditional finite element method: Find  $u_h \in V_h$  such that

$$(5.1) \quad (\nabla u_h, \nabla v)_\Omega + \frac{1}{\epsilon} \langle u_h, v \rangle_\Gamma = (f, v)_\Omega + \frac{1}{\epsilon} \langle u_0, v \rangle_\Gamma + \langle g, v \rangle_\Gamma \quad \forall v \in V_h.$$

We denote

$$h_\Gamma = \max_{E \in \mathcal{G}_h} h_E.$$

Then the standard technique for error estimation together with an interpolation estimate in the  $L^2(\Gamma)$ -norm (cf. [5]) gives:

**Theorem 5.1.** For  $u \in H^s(\Omega)$ , with  $1 < s \leq p + 1$  it holds that

$$(5.2) \quad \|\nabla(u - u_h)\|_{L^2(\Omega)} + \epsilon^{-1/2}\|u - u_h\|_{L^2(\Gamma)} \leq Ch^{s-1}(1 + h_\Gamma^{1/2}\epsilon^{-1/2})\|u\|_{H^s(\Omega)}.$$

From this estimate it is seen that the a priori estimate is optimal if  $h_\Gamma \leq C\epsilon$ . Note also that (for a quasiuniform mesh) the condition number of the method is

$$(5.3) \quad \kappa = \mathcal{O}(h^{-2} + (\epsilon h)^{-1}).$$

Hence, the natural  $\mathcal{O}(h^{-2})$  condition number for a second order equation is obtained when  $\epsilon \geq Ch$ .

Next, we will show that the same condition is needed for the a posteriori estimates to be optimal. By the standard technique [8] we obtain

**Theorem 5.2.** It holds that

$$(5.4) \quad \|\nabla(u - u_h)\|_{L^2(\Omega)} + \epsilon^{-1/2}\|u - u_h\|_{L^2(\Gamma)} \leq C\left(\sum_{K \in \mathcal{G}_h} E_{t,K}(u_h)^2\right)^{1/2},$$

with

$$(5.5) \quad \begin{aligned} E_{t,K}(u_h)^2 &= h_K^2 \|\Delta u_h + f\|_{L^2(K)}^2 + \sum_{E \subset \partial K \cap \mathcal{I}_h} h_E \left\| \left[ \frac{\partial u_h}{\partial n} \right] \right\|_{L^2(E)}^2 \\ &+ \sum_{E \subset \partial K \cap \Gamma} h_E \left\| \frac{\partial u_h}{\partial n} - g + \frac{1}{\epsilon}(u_h - u_0) \right\|_{L^2(E)}^2. \end{aligned}$$

When the data  $u_0$  is approximated by  $u_{0,h}$  we get from the last term

$$(5.6) \quad \begin{aligned} h_E^{1/2} \left\| \frac{\partial u_h}{\partial n} - g + \frac{1}{\epsilon}(u_h - u_0) \right\|_{L^2(\partial K \cap \Gamma)} &\leq h_E^{1/2} \left\| \frac{\partial u_h}{\partial n} - g + \frac{1}{\epsilon}(u_h - u_{0,h}) \right\|_{L^2(\partial K \cap \Gamma)} \\ &+ h_E^{1/2} \epsilon^{-1} \|u_{0,h} - u_0\|_{L^2(\partial K \cap \Gamma)}. \end{aligned}$$

From above it can be seen that in order to have an estimate uniformly valid with respect to  $\epsilon$  the condition  $h_E \leq C\epsilon$  has to be satisfied. The same condition is needed for the optimality of the following lower bound.

**Theorem 5.3.** The elementwise estimator  $E_{t,K}(u_h)$ , defined in equation (5.5), also fulfills

$$(5.7) \quad \begin{aligned} E_{t,K}(u_h)^2 &\leq C \left\{ |u - u_h|_{H^1(\omega_K)}^2 + h_K^2 \|f - f_h\|_{L^2(\omega_K)}^2 \right. \\ &+ \sum_{E \subset \partial K \cap \Gamma} h_E \epsilon^{-2} \left( \|u - u_h\|_{L^2(E)}^2 + \|u_0 - u_{0,h}\|_{L^2(E)}^2 \right) \\ &\left. + \sum_{E \subset \partial K \cap \Gamma} h_E \|g - g_h\|_{L^2(E)}^2 \right\}, \end{aligned}$$

where  $f_h$ ,  $u_{0,h}$  and  $g_h$  are approximations in  $V_h$  of the given data, and  $\omega_K$  is the domain of element  $K$  and all elements sharing an edge/face with  $K$ .

*Proof.* Clearly, it is only the boundary term that has not been treated in the earlier proofs. We let

$$(5.8) \quad R_\Gamma = \frac{\partial u_h}{\partial n} - g + \frac{1}{\epsilon}(u_h - u_0)$$

and

$$(5.9) \quad R_{\Gamma,\text{red}} = \frac{\partial u_h}{\partial n} - g_h + \frac{1}{\epsilon}(u_h - u_{0,h}).$$

We have

$$(5.10) \quad \|R_\Gamma\|_{L^2(E)} \leq \|R_{\Gamma,\text{red}}\|_{L^2(E)} + \|g - g_h\|_{L^2(E)} + \epsilon^{-1}\|u_0 - u_{0,h}\|_{L^2(E)}.$$

Let

$$\hat{w}_\Gamma = \Psi_E \mathcal{E} R_{\Gamma,\text{red}} \quad \text{and} \quad w_\Gamma = \Psi_E R_{\Gamma,\text{red}}.$$

Using Lemma 4.3 and the identities (with  $R_K$  defined as as before)

$$(R_K, \hat{w}_\Gamma)_K = (\nabla(u - u_h), \nabla \hat{w}_\Gamma)_K + \left\langle \frac{\partial}{\partial n}(u_h - u), \hat{w}_\Gamma \right\rangle_E$$

and

$$\frac{\partial u}{\partial n} - g + \frac{1}{\epsilon}(u - u_0) = 0,$$

gives

$$(5.11) \quad \begin{aligned} C\|R_{\Gamma,\text{red}}\|_{L^2(E)}^2 &\leq \|\Psi_E^{1/2} R_{\Gamma,\text{red}}\|_{L^2(E)}^2 = \langle R_{\Gamma,\text{red}}, w_\Gamma \rangle_E \\ &= \left\langle \frac{\partial}{\partial n}(u_h - u), \hat{w}_\Gamma \right\rangle_E - \langle g - g_h, w_\Gamma \rangle_E \\ &\quad + \epsilon^{-1} \langle u - u_h, w_\Gamma \rangle_E - \epsilon^{-1} \langle u_0 - u_{0,h}, w_\Gamma \rangle_E \\ &= (R_K, \hat{w}_\Gamma)_K + (\nabla(u_h - u), \nabla \hat{w}_\Gamma)_K - \langle g - g_h, w_\Gamma \rangle_E \\ &\quad + \epsilon^{-1} \langle u - u_h, w_\Gamma \rangle_E - \epsilon^{-1} \langle u_0 - u_{0,h}, w_\Gamma \rangle_E \\ &\leq \|R_K\|_{L^2(K)} \|\hat{w}_\Gamma\|_{L^2(K)} + \|\nabla(u - u_h)\|_{L^2(K)} \|\nabla \hat{w}_\Gamma\|_{L^2(K)} \\ &\quad + \|g - g_h\|_{L^2(E)} \|w_\Gamma\|_{L^2(E)} + \epsilon^{-1} \|u - u_h\|_{L^2(E)} \|w_\Gamma\|_{L^2(E)} \\ &\quad + \epsilon^{-1} \|u_0 - u_{0,h}\|_{L^2(E)} \|w_\Gamma\|_{L^2(E)} \\ &\leq C \left( h_K^{1/2} \|R_K\|_{L^2(K)} + h_K^{-1/2} \|u - u_h\|_{H^1(K)} + \|g - g_h\|_{L^2(E)} \right. \\ &\quad \left. + \epsilon^{-1} \|u - u_h\|_{L^2(E)} + \epsilon^{-1} \|u_0 - u_{0,h}\|_{L^2(E)} \right) \|R_{\Gamma,\text{red}}\|_{L^2(E)}. \end{aligned}$$

Hence, we have

$$(5.12) \quad \|R_{\Gamma,\text{red}}\|_{L^2(E)} \leq C \left( h_K^{1/2} \|R_K\|_{L^2(K)} + h_K^{-1/2} \|u - u_h\|_{H^1(K)} + \|g - g_h\|_{L^2(E)} \right. \\ \left. + \epsilon^{-1} \|u - u_h\|_{L^2(E)} + \epsilon^{-1} \|u_0 - u_{0,h}\|_{L^2(E)} \right),$$

which, together with (5.10) proves the assertion. □

From here we see that the estimator is sharp, i.e. it holds that

$$(5.13) \quad C \left( \sum_{K \in \mathcal{G}_h} E_{t,K}(u_h)^2 \right)^{1/2} \leq \|\nabla(u - u_h)\|_{L^2(\Omega)} + \epsilon^{-1/2} \|u - u_h\|_{L^2(\Gamma)},$$

when  $h_E \leq C\epsilon$ .



6. NUMERICAL EXAMPLES

In this section we report on numerical studies for the following problem

$$(6.1) \quad \begin{array}{l} \begin{array}{c} \begin{array}{c} \uparrow y \\ \Gamma_R \\ \Omega \\ \downarrow x \end{array} \end{array} \end{array} \quad \begin{array}{l} -\Delta u = 0 \\ \frac{\partial u}{\partial n} = \frac{1}{\epsilon}(u_0 - u) + g \\ u = 0 \end{array} \quad \begin{array}{l} \text{in } \Omega, \\ \text{on } \Gamma_R, \\ \text{on } \partial\Omega \setminus \Gamma_R, \end{array}$$

where

$$\Omega = \{(x, y) \mid x \in (0, 1), y \in (0, 3/10)\} \quad \text{and} \quad \Gamma_R = \{(x, y) \mid y = 3/10, x \in [0, 1]\}.$$

In order to get a nontrivial problem with a known exact solution we proceed in the following way. On  $\Gamma_R$  we let  $u_0$  be the  $n$ -th partial sum of the Fourier series of the function

$$\tilde{u}_0(x) = \begin{cases} 1 & \frac{3}{10} \leq x \leq \frac{7}{10}, \\ 0 & \text{otherwise,} \end{cases}$$

i.e.

$$u_0 = \sum_{k=1}^n U_k \sin(k\pi x),$$

with

$$(6.2) \quad U_k = 2 \frac{\cos(\frac{7}{10}k\pi) - \cos(\frac{3}{10}k\pi)}{k\pi}.$$

The solution to our problem is then equal to the solution of the Dirichlet problem, with  $u|_{\Gamma_D} = 0$  and  $u|_{\Gamma_R} = u_0$ . By standard Fourier techniques we then obtain

$$u(x, y) = \sum_{k=1}^n U_k \frac{\sinh(k\pi y) \sin(k\pi x)}{\sinh(3/10k\pi)}.$$

This is also the solution to our model problem (6.1) when we choose

$$g = \sum_{k=1}^n k\pi U_k \frac{\sinh(3/10k\pi)}{\cosh(3/10k\pi)} \sin(k\pi x).$$

By our definition, the exact solution is independent of the parameter  $\epsilon$  appearing in the boundary condition. With this we are able to extract the effect of the parameter  $\epsilon$  on the method rather than on the problem.

For all the computations in this paper we fix the number of Fourier coefficients to 21. Figure 1 shows this solution and we see how the regularity decreases near  $\Gamma_R$ . In all the computations the stability parameter appearing in the formulation is chosen as  $\gamma = 0.1$ .

Since the mathematical analysis seen earlier in this paper already establishes the a priori convergence results, we do not show any of the usual convergence graphs. Instead, we directly investigate the difference between the traditional method and Nitsche's approach.

First we show figures of the distribution of the error estimators  $E_K(u_h)$  and  $E_{t,K}(u_h)$  for a fixed mesh with different values of the parameter  $\epsilon$ . In Figure 2 we see the estimator distributions on mesh size  $h = 0.15$  and with  $\epsilon = 1, 0.1, 0.01$ .

We immediately notice that the traditional error estimator  $E_{t,K}(u_h)$  is highly dependent on the value of  $\epsilon$ . Also the proposed estimator  $E_K(u_h)$  grows as the  $\epsilon$  diminishes but the effect is much smaller. The analytical a posteriori results predict that the traditional method should perform well if the mesh size  $h$  is of the same order as  $\epsilon$  or smaller. This can be seen in Figure 2; for the traditional estimator the mesh is suited only for the first value of  $\epsilon$ .

In Figure 3 we show again the distributions of the estimators with the same values of  $\epsilon$ , but now for the mesh size  $h = 0.04$ . With this choice we expect the traditional estimator to perform well with the two larger values of the parameter  $\epsilon$ . Again both methods perform as expected, Nitsche's approach is unaffected by the  $\epsilon$  and the traditional method performs well for the values of  $\epsilon$  that are larger than the mesh size. From these figures it is clear that the boundary estimator of the traditional method cannot perform well with small values of  $\epsilon$ . Obviously, the problems of the traditional method arise from the boundary error estimator since the interior parts of the estimators are the same.

Next we test how the elementwise estimators  $E_K(u_h)$  and  $E_{t,K}(u_h)$  perform in adaptive mesh refinement. We refine until the error estimate, i.e. the sum of local estimators, is below the given tolerance. An element  $K$  is refined if

$$E_K(u_h)^2 \text{ or } E_{t,K}(u_h)^2 > \frac{(\text{tolerance})^2}{\text{number of elements}}.$$

All the adaptive computations have the same starting mesh with size  $h = 0.2$  and the same convergence tolerance. In Figure 4 we see the final meshes of the adaptive computations for both Nitsche's and the traditional method using different values of the parameter  $\epsilon$ . We notice that Nitsche's method produces almost the same mesh regardless of  $\epsilon$  which is natural since the exact solution is independent of  $\epsilon$ .

On the other hand, the traditional method needs more degrees of freedom as the  $\epsilon$  diminishes. For larger values of  $\epsilon$  both methods detect the regions at the boundary where the solution changes rapidly. For smaller values of  $\epsilon$ , the traditional estimator over-emphasizes the boundary error and is no longer able to detect the steep parts. Instead, the estimator sees error on the whole boundary and therefore refines on the whole boundary.

Finally, in Figure 5, we show the condition number of the system matrix for Nitsche's and the traditional method as a function of  $\epsilon$ . We notice that the condition number of the traditional method increases as equation (5.3) predicts. On the other hand, the condition number of Nitsche's method stays bounded for fixed  $h$ . For this reason the traditional method may cause trouble for iterative solvers such as multigrid method. In our two-dimensional computations incomplete Cholesky conjugate gradient (ICCG) methods have, however, performed well.

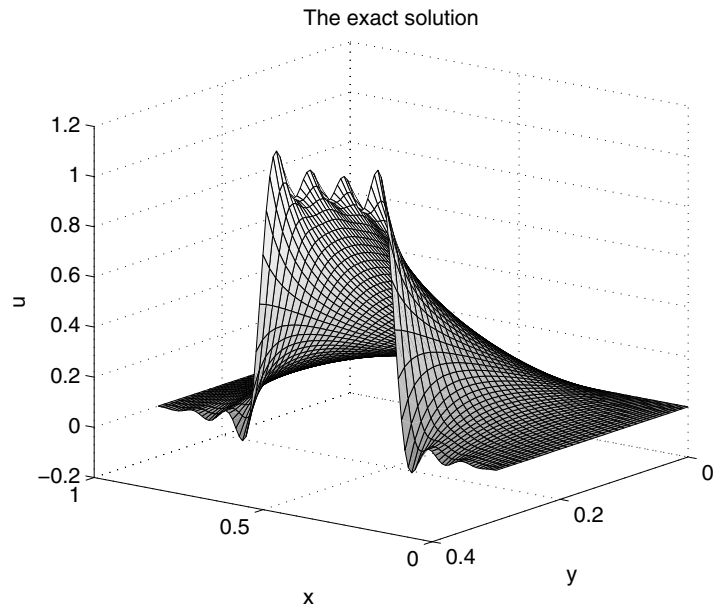


FIGURE 1. The exact solution to the model problem with 21 terms on the boundary data. Recall that the design of the model problem is such that the solution is independent of the boundary condition parameter  $\epsilon$ .

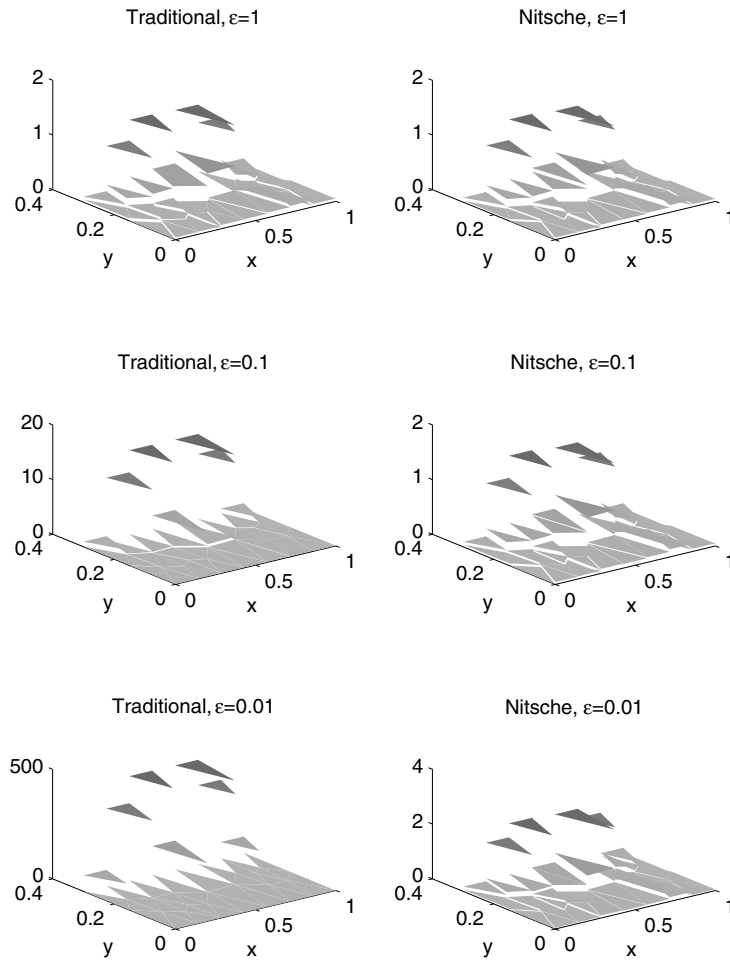


FIGURE 2. Distribution of the error estimators with different values of the boundary parameter  $\epsilon$ . On the left we have the traditional estimator and on the right the Nitsche estimator. From top to bottom  $\epsilon$  has values 1, 0.1 and 0.01. The mesh has size  $h = 0.15$ . Notice the scales and how dramatically the traditional estimator depends on  $\epsilon$ .

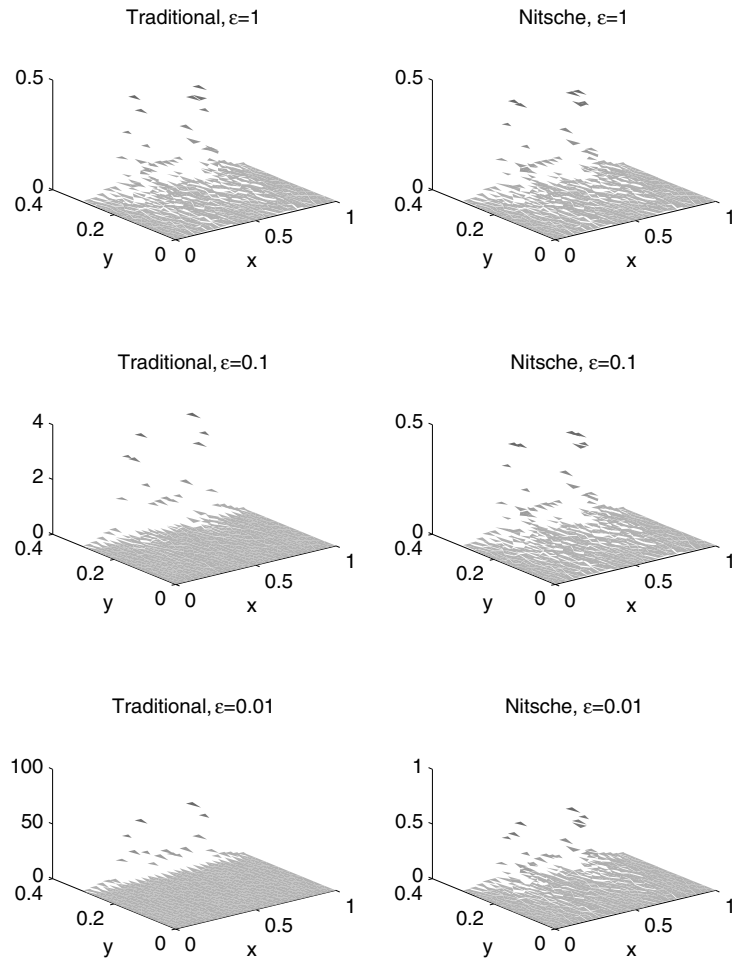


FIGURE 3. Distribution of the error estimators with different values of the boundary parameter  $\epsilon$ . On the left we have the traditional estimator and on the right the Nitsche estimator. From top to bottom  $\epsilon$  has values 1, 0.1 and 0.01. The mesh has size  $h = 0.04$ . Notice the scales.

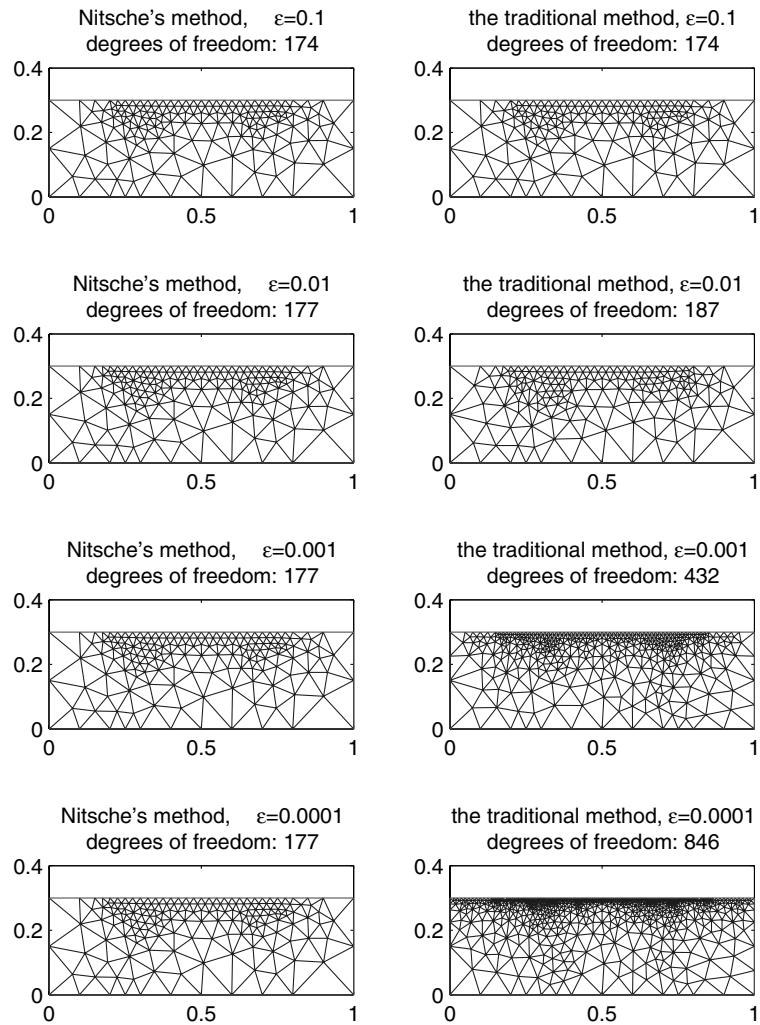


FIGURE 4. The final meshes of the adaptive refinement that fulfill the given tolerance. On the left meshes of Nitsche's method and on the right meshes of the traditional method. Notice that the traditional method is unable to detect the difficult parts of the solution with small  $\epsilon$ . Recall that the exact solution does not depend on  $\epsilon$ .

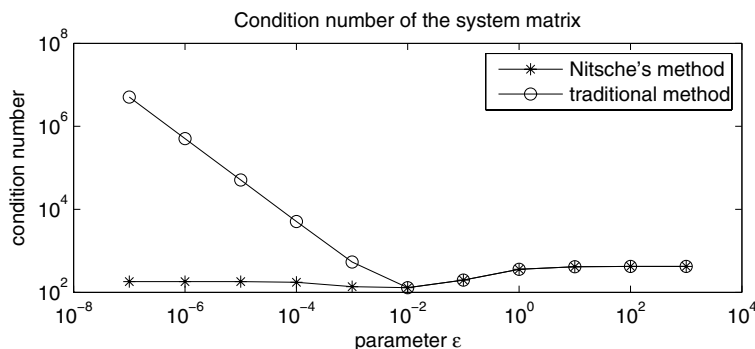


FIGURE 5. The condition number of the system matrix as a function of  $\epsilon$  for fixed  $h$  and for both Nitsche's and the traditional method. Notice the growth of condition number in the traditional method; see equation (5.3).

#### REFERENCES

1. Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2001/02), no. 5, 1749–1779 (electronic). MR1885715 (2002k:65183)
2. Ivo Babuška, Uday Banerjee, and John E. Osborn, *Survey of meshless and generalized finite element methods: a unified approach*, Acta Numer. **12** (2003), 1–125. MR2249154
3. Roland Becker, Peter Hansbo, and Rolf Stenberg, *A finite element methods for domain decomposition with non-matching grids*, Mathematical Modelling and Numerical Analysis **37** (2003), no. 2, 209–225. MR1991197 (2004e:65129)
4. D. Braess and R. Verfürth, *A posteriori error estimator for the Raviart-Thomas element*, SIAM J. Numer. Anal. **33** (1996), 2431–2444. MR1427472 (97m:65201)
5. Philippe G. Ciarlet, *The finite element methods for elliptic problems*, second ed., North-Holland, 1987. MR0520174 (58:25001)
6. J.A. Nitsche, *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), 9–15. MR0341903 (49:6649)
7. Rolf Stenberg, *Mortaring by a method of J.A. Nitsche*, Computational Mechanics; New Trends and Applications, S. Idelsohn, E. Oñate and E. Dvorkin (Eds.) (CIMNE, Barcelona, Spain, 1998). MR1839048
8. Rüdiger Verfürth, *A review of a posteriori error estimation and adaptive mesh refinement techniques*, Wiley, 1996.

INSTITUTE OF MATHEMATICS, HELSINKI UNIVERSITY OF TECHNOLOGY, P. O. BOX 1100, 02015 TKK, FINLAND

*E-mail address:* mika.juntunen@tkk.fi

INSTITUTE OF MATHEMATICS, HELSINKI UNIVERSITY OF TECHNOLOGY, P. O. BOX 1100, 02015 TKK, FINLAND