

# Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method

Erik Burman,<sup>a</sup> Peter Hansbo<sup>b</sup>

<sup>a</sup>*Department of Mathematics, University of Sussex, Falmer, BN1 9RF UK*

<sup>b</sup>*Department of Mathematical Sciences, Chalmers University of Technology and  
University of Gothenburg, SE-41296 Göteborg, Sweden*

---

## Abstract

We extend the classical Nitsche type weak boundary conditions to a fictitious domain setting. An additional penalty term, acting on the jumps of the gradients over element faces in the interface zone, is added to ensure that the conditioning of the matrix is independent of how the boundary cuts the mesh. Optimal a priori error estimates in the  $H^1$  and  $L^2$  norms are proved as well as an upper bound on the condition number of the system matrix.

*Key words:* interior penalty, fictitious domain, finite element.

---

## 1 Introduction

In a previous paper [2], we introduced and analyzed a Lagrange multiplier method for handling Dirichlet boundary conditions weakly on a mesh cut by the boundary. We showed optimal order error estimates in the  $H^1$ - and  $L_2$ -norms for our elliptic model problem, Poisson's equation. In this paper we propose the alternative of using Nitsche's method [7] to apply boundary conditions on a boundary to which the mesh is not fitted. This method has the advantage of avoiding the use of multipliers and yields a symmetric positive definite discrete system **for the case at hand**.

If we let  $\Omega \subset \mathbb{R}^2$  denote the computational domain and  $\mathcal{T}_h \subset \mathbb{R}^2$ , the computational mesh consisting of triangles  $K$ , we only assume **that  $\Omega$  is covered by the mesh and that all elements have at least a part inside  $\Omega$** :

- $\Omega \subset \mathcal{T}_h$ .

- $K \cap \Omega \neq \emptyset$  for all  $K \in \mathcal{T}_h$ .

The analysis is inspired by the unfitted finite element method proposed by Hansbo and Hansbo [5], which was limited to **the corresponding interface problem** and did not cover the case of Dirichlet boundaries.

In this paper we offer a solution valid in the case of a piecewise affine approximation space. The idea is to add a penalty term on the gradient jumps in the boundary zone. This penalty term acts also on the part of the elements that are outside the domain, where the solution has no physical significance. The rationale for this ghost penalty term is that it ensures coercivity over the whole computational mesh by extending the  $H^1$  control of the solution over the whole mesh, i.e., also outside the physical domain  $\Omega$ . A related approach was used in [1] to ensure inf-sup stability for an unfitted finite element method for incompressible elasticity.

The outline of this paper is as follows: in the next section the fictitious domain finite element method is introduced, applied to the Poisson problem. Sections 3 and 4 are concerned with the accuracy of the method and error estimates are proved in the  $H^1$  and  $L^2$ -norms. Finally in Section 5 we prove that the condition number of the matrix is independent of the way **the boundary** cuts the computational mesh.

## 2 The finite element formulation

Let  $\Omega$  be a bounded domain in  $\mathbb{R}^2$ , with smooth, or convex polygonal, boundary  $\Gamma$ . The Poisson equation that we propose as a model problem is given by

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

where  $\Gamma = \Gamma_D \cup \Gamma_N$  denotes the boundary of the domain  $\Omega$ , with outward pointing unit normal  $\mathbf{n}_\Gamma$ . Let

$$(u, v)_X := \int_X uv \, dx$$

denote the  $L^2$ -inner product on  $X \subset \mathbb{R}^d$  with associated norm  $\|u\|_X = (u, u)_X^{\frac{1}{2}}$  and

$$\langle u, v \rangle_Y := \int_Y uv \, ds$$

the  $L^2$ -inner product over  $Y \subset \mathbb{R}^{d-1}$ , where  $Y$  in practice will be either some subset of the boundary  $Y \subset \Gamma$ , or some element face in the computational

mesh. We will also use the notation  $\|u\|_{k,X}$  for the norm of a function  $u \in H^k(X)$ . If  $X$  is the computational mesh, or some subset thereof, the broken space is considered, i.e. functions such that  $v|_{K \cap X} \in H^k(K \cap X)$  for all  $K \cap X \neq \emptyset$ . Let  $V_g = \{v \in H^1(\Omega) : v|_{\Gamma_D} = g_D\}$ . We have the following weak formulation: find  $u \in V_g$  such that

$$a(u, v) = (f, v)_\Omega + \langle g_N, v \rangle_{\Gamma_N}, \quad \forall v \in V_0 \quad (2)$$

where

$$a(u, v) := (\nabla u, \nabla v)_\Omega.$$

We will assume that  $\Gamma_D, \Gamma_N$ , the shape of  $\Omega$ , and the regularity of the data are such that the following shift theorem holds

$$\|u\|_{2,\Omega} \leq C_S (\|f\|_{0,\Omega} + \|g_D\|_{\frac{3}{2},\Gamma_D} + \|g_N\|_{\frac{1}{2},\Gamma_N}) = C_{fg}. \quad (3)$$

Consider a family of quasi-uniform computational meshes  $\{\mathcal{T}_h\}_h$  such that  $\mathcal{T}_h = \{K\}$  where  $K$  denotes a shape regular simplex. The intersection of any two simplices is either the empty set, a vertex, or a face.

We do not assume that the mesh  $\mathcal{T}_h$  is fitted to the boundary of  $\Omega$ , only that  $\Omega \subset \mathcal{T}_h$  and that for all  $K \in \mathcal{T}_h$   $K \cap \Omega \neq \emptyset$ . Let  $\Omega_{\mathcal{T}}$  denote the domain covered by the mesh  $\mathcal{T}_h$ .

We will use the following notation for mesh related quantities. Let  $h_K$  be the diameter of  $K$  and  $h = \max_{K \in \mathcal{T}_h} h_K$ . By  $G_h := \{K \in \mathcal{T}_h : K \cap \Gamma \neq \emptyset\}$  we denote the set of elements that are intersected by the interface. For an element  $K \in G_h$ , let  $\Gamma_K := \Gamma \cap K$  be the part of  $\Gamma$  in  $K$ . We introduce the following assumptions regarding  $\Gamma$  and  $\mathcal{T}_h$ .

- ((A)) We assume that  $\Gamma$  intersects each element boundary  $\partial K$  exactly twice, and each (open) edge at most once for  $K \in G_h$ .
- ((B)) Let  $\Gamma_{K,h}$  be the straight line segment connecting the points of intersection between  $\Gamma$  and  $\partial K$ . We assume that  $\Gamma_K$  is a function of length on  $\Gamma_{K,h}$ ; in local coordinates

$$\Gamma_{K,h} = \{(\xi, \eta) : 0 < \xi < |\Gamma_{K,h}|, \eta = 0\}$$

and

$$\Gamma_K = \{(\xi, \eta) : 0 < \xi < |\Gamma_{K,h}|, \eta = \delta(\xi)\}.$$

- ((C)) We assume that for all  $K$  in  $G_h$  there exists  $K' \in \mathcal{T}_h$  such that  $K' \notin G_h$  and  $K \cap K' \neq \emptyset$  and such that that the measures of  $K$  and  $K'$  are comparable in the sense that there exists  $c_q > 1$  such that

$$c_q^{-1} \leq \frac{\text{meas}(K)}{\text{meas}(K')} \leq c_q \quad (4)$$

and that the faces  $F$  such that  $K \cap F \neq \emptyset$  and  $K' \cap F \neq \emptyset$  satisfy

$$|\text{meas}(K')| \leq c_q \text{meas}(F)^2. \quad (5)$$

Since the curvature of  $\Gamma$  is bounded except in a finite number of points, the first two assumptions are always fulfilled on sufficiently fine meshes.

We also introduce the set of element faces  $\mathcal{F}_G$  associated with  $G_h$ , defined as follows: for each face  $F \in \mathcal{F}_G$  there exists two simplices  $K$  and  $K'$  such that  $F = K \cap K'$  and at least one of the two is a member of  $G_h$ . This means in particular that the boundary faces of the mesh  $\mathcal{T}_h$  are excluded from  $\mathcal{F}_G$ . On a face  $F$  such that  $F = K \cap K'$ , define the jump of the grandient of  $v_h \in V^h$  by  $[\nabla v_h] = \mathbf{n}_F \cdot \nabla v_h|_K - \mathbf{n}_F \cdot \nabla v_h|_{K'}$ , where  $\mathbf{n}_F$  denotes a unit normal to the face  $F$  with fixed but arbitrary orientation. In the analysis we will frequently use the notation  $a \lesssim b$  meaning that  $a \leq Cb$  where  $C$  is a positive constant depending only on the shape regularity bound  $C_\rho$  or the particular form of the problem (1). In particular the constant is independent of the way  $\Gamma$  cuts the mesh.

We shall consider the following finite element space:

$$V^h = \{v \in C^0(\bar{\Omega}_T) : v|_K \in P^1(K), \forall K \in \mathcal{T}_h\}.$$

The finite element discretisation now takes the form: find  $u_h \in V^h$ , such that

$$A_h(u_h, v_h) = L(v_h) \quad \forall v_h \in V^h \quad (6)$$

where

$$\begin{aligned} L(v_h) := & (f, v_h)_\Omega + \langle g_D, \gamma_D h^{-1} v_h - \mathbf{n}_\Gamma \cdot \nabla v_h \rangle_{\Gamma_D} \\ & + \langle g_N, v_h + \gamma_N h \mathbf{n}_\Gamma \cdot \nabla v_h \rangle_{\Gamma_N} \end{aligned}$$

and

$$A_h(u_h, v_h) := a_h(u_h, v_h) + j(u_h, v_h)$$

with

$$\begin{aligned} a_h(u_h, v_h) := & a(u_h, v_h) - \langle \mathbf{n}_\Gamma \cdot \nabla u_h, v_h \rangle_{\Gamma_D} - \langle \mathbf{n}_\Gamma \cdot \nabla v_h, u_h \rangle_{\Gamma_D} \\ & + \langle \gamma_D h^{-1} u_h, v_h \rangle_{\Gamma_D} + \langle \gamma_N h \mathbf{n}_\Gamma \cdot \nabla u_h, \mathbf{n}_\Gamma \cdot \nabla v_h \rangle_{\Gamma_N} \end{aligned} \quad (7)$$

and

$$j(u_h, v_h) = \sum_{F \in \mathcal{F}_G} \langle \gamma_1 h [\nabla u_h], [\nabla v_h] \rangle_F.$$

Here  $\gamma_D, \gamma_N$  and  $\gamma_1$  are positive penalty parameters that will be discussed later. The formulation (6) satisfies the following consistency relation

**Lemma 1** (*Galerkin orthogonality*) Let  $u_h$  be the solution of the finite element formulation (6) and  $u \in H^1(\Omega)$  be the solution of (2). Then

$$a_h(u_h - u, v_h) = -j(u_h, v_h) \quad \forall v_h \in V^h. \quad (8)$$

**PROOF.** The proof is standard.

### 3 Approximation properties

We need to show that our approximating space  $V^h$  has optimal approximation properties. We introduce three norms for the analysis

$$|v|_{1,h}^2 := \|\nabla v\|_{0,\Omega}^2 + \|\gamma_N^{1/2} \mathbf{n}_\Gamma \cdot \nabla v\|_{-\frac{1}{2},h,\Gamma_N}^2 + \|\gamma_D^{1/2} v\|_{\frac{1}{2},h,\Gamma_D}^2,$$

$$|||v|||_*^2 := \|\nabla v\|_{0,\Omega}^2 + \|\mathbf{n}_\Gamma \cdot \nabla v\|_{-\frac{1}{2},h,\Gamma}^2 + \|\gamma_D^{1/2} v\|_{\frac{1}{2},h,\Gamma_D}^2$$

and

$$|||v|||_h^2 := \|\nabla v\|_{0,\Omega_T}^2 + \|\gamma_N^{1/2} \mathbf{n}_\Gamma \cdot \nabla v\|_{-\frac{1}{2},h,\Gamma_N}^2 + \|\gamma_D^{1/2} v\|_{\frac{1}{2},h,\Gamma_D}^2 + j(v, v).$$

The norms on the trace  $\Gamma$  are defined by  $\|v\|_{\alpha,h,\Gamma} = \|h^{-\alpha} v\|_{0,\Gamma}$ . Clearly  $\langle u, v \rangle_\Gamma \leq \|u\|_{-\frac{1}{2},h,\Gamma} \|v\|_{\frac{1}{2},h,\Gamma}$  by the Cauchy-Schwarz inequality. Note that  $|\cdot|_{1,h}$  and  $|||v|||_*$  require  $v$  to be in  $H^2(\Omega)$  to make sense, whereas  $|||v|||_h$  requires  $v$  to be in  $H^2(\mathcal{T}_h)$ . If  $\gamma_N = 0$ ,  $|\cdot|_{1,h}$  is well defined for functions in  $H^1(\Omega)$ . Below we will frequently use the following inequalities, the proofs of which may be found, e.g., in [4].

**Lemma 2** (*Trace inequalities*) Let  $v \in H^1(K)$ . Then

$$\|v\|_{0,\partial K} \leq C_T (h_K^{-\frac{1}{2}} \|v\|_{0,K} + h_K^{\frac{1}{2}} \|\nabla v\|_{0,K}) \quad (9)$$

and

$$\|v\|_{0,\Gamma \cap K} \leq C_T (h_K^{-\frac{1}{2}} \|v\|_{0,K} + h_K^{\frac{1}{2}} \|\nabla v\|_{0,K}). \quad (10)$$

**Lemma 3** (*Inverse inequality*) Let  $v_h \in V^h$ . Then there holds

$$|||v_h|||_h \leq C_I h^{-1} \|v_h\|_{0,\Omega_T}.$$

**Lemma 4** (*Poincaré inequality*) For all  $v \in H^1(\Omega_T)$  there holds

$$\|v\|_{1,\Omega_T} \leq C_P (\|\nabla v\|_{0,\Omega_T} + \|v\|_{\frac{1}{2},h,\Gamma}).$$

Note that thanks to the Poincaré inequality all the above defined semi-norms are indeed norms. The constant in the Poincaré inequality depends on the boundary  $\Gamma$ , but not on its relation to the mesh, since it holds for all functions in  $H^1(\Omega_{\mathcal{T}})$ . A consequence of the second trace inequality is that, independently of  $\gamma_N$

$$|||u_h|||_* \lesssim |||u_h|||_h.$$

Following [5], we will use an  $H^2$ -extension on  $\Omega_{\mathcal{T}}$ ,  $\mathbb{E} : H^2(\Omega) \rightarrow H^2(\Omega_{\mathcal{T}})$ , such that  $|||\mathbb{E}u|||_{2,\Omega_{\mathcal{T}}} \lesssim \|u\|_{2,\Omega}$  (c.f. [3]) and the (Clément type) interpolation operator  $I^* : H^1(\Omega) \rightarrow V^h$  defined by  $I^*u := \mathcal{C}_h \mathbb{E}u$ , where  $\mathcal{C}_h : H^1(\Omega_{\mathcal{T}}) \rightarrow V^h$  denotes the standard Clément interpolant. We now state the approximation results needed for the convergence analysis

**Lemma 5** *For all  $u \in H^2(\Omega)$*

$$|||\mathbb{E}u - I^*u|||_h \lesssim h|u|_{2,\Omega} \quad (11)$$

and

$$|||u - I^*u|||_* + j(I^*u, I^*u)^{\frac{1}{2}} \lesssim h|u|_{2,\Omega}, \quad (12)$$

$$\|u - I^*u\|_{1,h} \lesssim h|u|_{2,\Omega}. \quad (13)$$

**PROOF.** First we note that

$$|u - I^*u|_{1,h} \leq |||\mathbb{E}u - I^*u|||_h$$

and  $j(I^*u, I^*u)^{\frac{1}{2}} = j(\mathbb{E}u - I^*u, \mathbb{E}u - I^*u)^{\frac{1}{2}}$ . Hence it is sufficient to prove (11) and the inequality

$$\|\mathbf{n}_{\Gamma} \cdot \nabla(u - I^*u)\|_{-\frac{1}{2},h,\Gamma} \lesssim h\|u\|_{2,\Omega}.$$

By the properties of the Clément interpolant we readily deduce that

$$\|\mathbb{E}u - I^*u\|_{0,\Omega_{\mathcal{T}}} \lesssim h^2\|u\|_{2,\Omega} \quad \text{and} \quad \|\nabla(\mathbb{E}u - I^*u)\|_{0,\Omega_{\mathcal{T}}} \lesssim h|u|_{2,\Omega}.$$

For the penalty terms on  $\Gamma$  we have, using the trace inequality (10),

$$\|u - I^*u\|_{\frac{1}{2},h,\Gamma}^2 \lesssim \sum_{K \in G_h} \left( h_K^{-2} \|\mathbb{E}u - I^*u\|_{0,K}^2 + \|\nabla(\mathbb{E}u - I^*u)\|_{0,K}^2 \right) \lesssim h^2\|u\|_{2,\Omega}^2$$

and

$$\|\mathbf{n}_{\Gamma} \cdot \nabla(u - I^*u)\|_{-\frac{1}{2},h,\Gamma} \lesssim \sum_{K \in G_h} \left( \|\nabla(\mathbb{E}u - I^*u)\|_{0,K}^2 + h^2\|D^2\mathbb{E}u\|_{0,K}^2 \right) \lesssim h^2\|u\|_{2,\Omega}^2.$$

Finally we consider the penalty term  $j(\cdot, \cdot)$  and we note that using a trace inequality we have, here for simplicity extending the sum to all elements of

the mesh,

$$j(\mathbb{E}u - I^*u, \mathbb{E}u - I^*u) \lesssim \sum_{K \in \Omega_{\mathcal{T}}} \left( \|\nabla(\mathbb{E}u - I^*u)\|_{0,K}^2 + h^2 \|D^2 \mathbb{E}u\|_{0,K}^2 \right) \lesssim h^2 \|u\|_{2,\Omega}^2.$$

#### 4 A priori error analysis

The key result of this note is the following Lemma showing that thanks to the gradient penalty in the boundary zone we have control of the  $L^2$ -norm of the gradient over the whole computational mesh  $\mathcal{T}_h$ .

**Lemma 6** (*Coercivity*)

Let  $\gamma_D$  be sufficiently large in a sense that will be made precise and let  $\gamma_1 = 1$  and  $\gamma_N \geq 0$ . Then

$$\|u_h\|_h^2 \lesssim a_h(u_h, u_h) + j(u_h, u_h), \quad \forall u_h \in V^h.$$

**PROOF.** First note that by the definition of  $a_h(\cdot, \cdot)$  there holds

$$a_h(u_h, u_h) + j(u_h, u_h) = |u_h|_{1,h}^2 - 2\langle \mathbf{n}_\Gamma \cdot \nabla u_h, u_h \rangle_{\Gamma_D} + j(u_h, u_h).$$

For future reference note that

$$2\langle \mathbf{n}_\Gamma \cdot \nabla u_h, u_h \rangle_{\Gamma_D} \leq \varepsilon \|\mathbf{n}_\Gamma \cdot \nabla u_h\|_{-\frac{1}{2},h,\Gamma_D}^2 + \frac{1}{\varepsilon} \|u_h\|_{\frac{1}{2},h,\Gamma_D}^2$$

for  $\varepsilon \in \mathbb{R}^+$ , and that

$$\|\mathbf{n}_\Gamma \cdot \nabla u_h\|_{-\frac{1}{2},h,\Gamma_D}^2 \leq C_T^2 \|\nabla u_h\|_{0,G_h}^2 \leq C_T^2 \|\nabla u_h\|_{0,\Omega_{\mathcal{T}}}^2.$$

The quantity  $C_T^2 \|\nabla u_h\|_{0,\Omega_{\mathcal{T}}}^2$  can not be controlled uniformly by  $|u_h|_{1,h}^2$  and therefore the standard Nitsche-method fails. We must show that the stabilization of the gradient jumps allows us to recover coercivity of the part of the triangles that are not in  $\Omega$ . For each triangle  $K \in G_h$ , let  $K'$  be a triangle such that  $K' \notin G_h$  and such that  $K' \cap K \neq \emptyset$  (the existence of  $K'$  is guaranteed by our assumptions). Note that the intersection of  $K$  and  $K'$  may consist of a face or only a vertex. Denote  $\mathcal{F}_{KK'}$  the set of faces that has to be crossed in order to pass from  $K$  to  $K'$ . By the shape regularity the cardinality of this set is uniformly upper bounded by some  $N_{\mathcal{F}}$ . It follows that, since the tangential component of the gradient is continuous we may write

$$\nabla u_h|_K = \nabla u_h|_{K'} + \sum_{F \in \mathcal{F}_{KK'}} \delta_F [\nabla u_h]|_F \mathbf{n}_F$$

where  $\delta_F = \pm 1$ , with sign chosen depending on the orientation of  $\mathbf{n}_F$  in such a way that the equality holds. Taking the square of both sides and applying Cauchy-Schwarz inequality and the geometric-arithmetic inequality repeatedly in the right hand side we obtain

$$|\nabla u_h|_K|^2 \leq (1 + N_{\mathcal{F}}) \left( |\nabla u_h|_{K'}|^2 + \sum_{F \in \mathcal{F}_{KK'}} |[\nabla u_h]|_F|^2 \right).$$

By the assumptions of quasi-uniformity in the interface zone (4) and (5) we have

$$\|\nabla u_h\|_K^2 \leq (1 + N_{\mathcal{F}}) c_q \left( \|\nabla u_h\|_{K'}^2 + \sum_{F \in \mathcal{F}_{KK'}} \|h_F [\nabla u_h]\|_{0,F}^2 \right).$$

Since any given  $K' \in \bar{\Omega}$  can appear in the upper bound of at most  $6N_{\mathcal{F}}$  triangles  $K \in G_h$  it follows that

$$\|\nabla u_h\|_{0,G_h}^2 \leq 6N_{\mathcal{F}}(1 + N_{\mathcal{F}}) c_q \left( \|\nabla u_h\|_{0,\Omega}^2 + \sum_{F \in \mathcal{F}_G} \|h_F [\nabla u_h]\|_{0,F}^2 \right).$$

Hence

$$\begin{aligned} \|\nabla u_h\|_{0,\Omega_T}^2 &\lesssim \|\nabla u_h\|_{0,G_h}^2 + \|\nabla u_h\|_{0,\Omega}^2 \\ &\leq (6N_{\mathcal{F}}(1 + N_{\mathcal{F}}) c_q + 1) \left( \|\nabla u_h\|_{0,\Omega}^2 + \sum_{F \in \mathcal{F}_G} \|h_F [\nabla u_h]\|_{0,F}^2 \right) \end{aligned}$$

showing that choosing  $\gamma_1 = 1$  and setting  $C_{\mathcal{F}_q} = (6N_{\mathcal{F}}(1 + N_{\mathcal{F}}) c_q + 1)^{-1}$  we have

$$C_{\mathcal{F}_q} \|\nabla u_h\|_{0,\Omega_T}^2 \leq \left( \|\nabla u_h\|_{0,\Omega}^2 + j(u_h, u_h) \right).$$

We now fix  $\varepsilon = \frac{C_{\mathcal{F}_q}}{2C_T^2}$  and  $\gamma_D = \frac{2}{\varepsilon} = \frac{2C_T^2}{C_{\mathcal{F}_q}}$  leading to

$$\frac{1}{2} \min\{C_{\mathcal{F}_q}, 1\} \|\nabla u_h\|_h^2 \leq a_h(u_h, u_h) + j(u_h, u_h)$$

which concludes the proof.

**Lemma 7 (Continuity)** *Let  $v = v_1 + v_2$  with  $v_1 \in H^2(\mathcal{T}_h)$ ,  $v_2 \in V^h$  and let  $y_h, w_h \in V^h$ . The bilinear forms  $A_h(\cdot, \cdot)$  and  $a_h(\cdot, \cdot)$  satisfy*

$$A_h(y_h, w_h) \leq C_a \|y_h\|_h \|w_h\|_h$$

and

$$a_h(v, w_h) \leq C_a \|v\|_* \|w_h\|_*$$

where  $C_a$  is a positive constant independent of how  $\Gamma$  cuts  $\mathcal{T}_h$ .



**PROOF.** Immediate by the Cauchy-Schwarz inequality.

**Proposition 8** (*Optimality with respect to interpolation*)

Let  $u \in H^2(\Omega)$  be the solution of (1) and  $u_h$  the solution of (6). Then there holds

$$\| \|u_h - I^*u\| \|_h \lesssim \left( \| \|u - I^*u\| \|_* + j(I^*u, I^*u)^{\frac{1}{2}} \right).$$

**PROOF.** Let  $e_h = u_h - I^*u \in V^h$ . By Lemma 6 we have

$$\| \|e_h\| \|_h^2 \lesssim a_h(e_h, e_h) + j(e_h, e_h).$$

Applying the Galerkin orthogonality of Lemma 1 with  $v_h = e_h$  we obtain

$$\| \|e_h\| \|_h^2 \lesssim a_h(u - I^*u, e_h) - j(I^*u, e_h).$$

We conclude by the continuity of Lemma 7, the inequality  $\| \|w_h\| \|_* \lesssim \| \|w_h\| \|_h$  and the Cauchy-Schwarz inequality  $j(I^*u, e_h) \leq j(I^*u, I^*u)^{\frac{1}{2}} j(e_h, e_h)^{\frac{1}{2}}$  leading to

$$\| \|e_h\| \|_h^2 \lesssim \left( \| \|u - I^*u\| \|_* + j(I^*u, I^*u)^{\frac{1}{2}} \right) \| \|e_h\| \|_h.$$

**Corollary 9** (*Optimal convergence*)

Let  $u \in H^2(\Omega)$  be the solution of (1) and  $u_h$  the solution of (6). Then there holds

$$\| \|u_h - u\| \|_* \lesssim C_{fg} h. \quad (14)$$

**PROOF.** First we note that

$$\| \|u_h - u\| \|_* \leq \| \|u - I^*u\| \|_* + \| \|u_h - I^*u\| \|_h.$$

By Lemma 5, equation (13) there holds  $\| \|I^*u - u\| \|_* \lesssim ch|u|_{2,\Omega}$ . The upper bound

$$\| \|u_h - I^*u\| \|_h \lesssim h|u|_{2,\Omega} \quad (15)$$

is immediate by Proposition 8 and by applying once again Lemma 5, equation (12). Finally apply equation (3).

**Proposition 10** ( *$L^2$ -norm convergence*)

Assume that  $u \in H^2(\Omega)$  is the solution of (1) and  $u_h \in V^h$  is the solution of (6). Then there holds

$$\| \|u - u_h\| \|_{0,\Omega} \lesssim C_{fg}^2 h^2.$$

**PROOF.** Consider the auxiliary problem

$$\begin{aligned} -\Delta z &= u - u_h && \text{in } \Omega, \\ z &= 0 && \text{on } \Gamma_D, \\ \nabla z \cdot n &= 0 && \text{on } \Gamma_N. \end{aligned} \tag{16}$$

By the inequality (3) there holds

$$\|z\|_{2,\Omega} \lesssim \|u - u_h\|_{0,\Omega}. \tag{17}$$

Multiplying the first equation of (16) with  $u - u_h$  and integrating by parts gives

$$\begin{aligned} \|u - u_h\|_{0,\Omega}^2 &= (\nabla(u - u_h), \nabla z) - \langle u - u_h, \nabla z \cdot n_\Gamma \rangle_\Gamma \\ &= (\nabla(u - u_h), \nabla z) - \langle u - u_h, \nabla z \cdot n_\Gamma \rangle_{\Gamma_D} - \langle \nabla(u - u_h) \cdot n, z \rangle_{\Gamma_D} \\ &\quad + \langle \frac{\gamma_D}{h}(u - u_h), z \rangle_{\Gamma_D} + \langle \gamma_N h \nabla(u - u_h) \cdot n_\Gamma, \nabla z \cdot n_\Gamma \rangle_{\Gamma_N} \\ &= a_h(u - u_h, z). \end{aligned}$$

Using now Galerkin orthogonality (1) and the continuity of Lemma 7 we may write

$$\begin{aligned} \|u - u_h\|_{0,\Omega}^2 &= a_h(u - u_h, z - I^*z) - j(u_h, I^*z) \\ &\lesssim \| |u - u_h| \|_* \|z - I^*z\|_* + j(u_h, u_h)^{\frac{1}{2}} j(I^*z, I^*z)^{\frac{1}{2}}. \end{aligned}$$

By Lemma 5 (equation (12)), the inequality

$$j(u_h, u_h)^{\frac{1}{2}} \lesssim \| |u_h - I^*u| \|_h + j(I^*u, I^*u)^{\frac{1}{2}}$$

and Corollary 9 (both equation (14) and equation (15)) we have

$$\begin{aligned} \|u - u_h\|_{0,\Omega}^2 &\lesssim \| |u - u_h| \|_* \|z - I^*z\|_* + \mathcal{C}_{fg} h^2 |z|_{2,\Omega} \\ &\lesssim \mathcal{C}_{fg} h^2 |z|_{2,\Omega}. \end{aligned}$$

The inequality (17) applied in the right hand side ends the proof.

## 5 Conditioning of the system matrix

In this section we will show that the condition number of the matrix satisfies a similar upper bound as the one resulting from a standard finite element method. The analysis follows the framework presented in [4]. For any vector  $U \in \mathbb{R}^N$ , where  $N = \dim V^h$  denote the standard Euclidean norm by  $|U|_N$  and the corresponding finite element function in  $V^h$  by  $u_h$ . Further let  $\mathcal{M}$  denote

the mass matrix defined by the bilinear form  $(u_h, v_h)_{\Omega_{\mathcal{T}}}$  and  $\mathcal{A}$  denote the system matrix defined by the bilinear form  $A_h(u_h, v_h)$ . Since  $\mathcal{T}_h$  is a **conforming, quasi-uniform**, mesh on the domain  $\Omega_{\mathcal{T}}$  we have the estimates

$$\mu_{min}^{\frac{1}{2}}|U|_N \leq \|u_h\|_{0,\Omega_{\mathcal{T}}} \leq \mu_{max}^{\frac{1}{2}}|U|_N \quad (18)$$

where  $\mu_{min}$  and  $\mu_{max}$  denotes the smallest and largest eigenvalues of  $\mathcal{M}$ . By definition the condition number  $\kappa(\mathcal{A})$  is given by

$$\kappa(\mathcal{A}) := \|\mathcal{A}\| \|\mathcal{A}^{-1}\|,$$

where

$$\|\mathcal{A}\| = \sup_{U \in \mathbb{R}^N} \frac{|\mathcal{A}U|_N}{|U|_N}.$$

With this notation we now prove the following:

**Lemma 11** *The condition number of the system matrix resulting from the formulation (6) satisfies the upper bound*

$$\kappa(\mathcal{A}) \leq \frac{\mu_{max}}{\mu_{min}} C_P^2 C_c^{-1} C_I^2 C_a h^{-2}.$$

**PROOF.** By definition

$$|\mathcal{A}U|_N = \sup_{W \in \mathbb{R}^N} \frac{(\mathcal{A}U, W)_N}{|W|_N} = \sup_{W \in \mathbb{R}^N} \frac{A_h(u_h, w_h)}{|W|_N}.$$

Using now the continuity of  $A_h(\cdot, \cdot)$ , Lemma 3 and the bound (18) we may write

$$A_h(u_h, w_h) \leq C_a C_I \| \|u_h\| \|_h h^{-1} \mu_{max}^{\frac{1}{2}} |W|_N$$

and hence

$$|\mathcal{A}U|_N \leq C_a C_I \| \|u_h\| \|_h h^{-1} \mu_{max}^{\frac{1}{2}}.$$

Applying once again Lemma 3 and the bound (18) we have

$$|\mathcal{A}U|_N \leq C_a C_I^2 h^{-2} \mu_{max} |U|_N$$

resulting in

$$\|\mathcal{A}\| = \sup_{U \in \mathbb{R}^N} \frac{|\mathcal{A}U|_N}{|U|_N} \leq C_a C_I^2 \mu_{max} h^{-2}. \quad (19)$$

Similarly for the norm of the inverse we have

$$\begin{aligned} C_P^{-1} C_c \mu_{min}^{\frac{1}{2}} |U|_N &\leq C_P^{-1} C_c \|u_h\|_{0,\Omega_{\mathcal{T}}} \\ &\leq C_c \| \|u_h\| \|_h \leq \frac{A_h(u_h, u_h)}{\| \|u_h\| \|_h} = \frac{(\mathcal{A}U, U)_N}{\| \|u_h\| \|_h} \\ &\leq \frac{|\mathcal{A}U|_N \mu_{min}^{-\frac{1}{2}} C_P}{\| \|u_h\| \|_h} = \mu_{min}^{-\frac{1}{2}} C_P |\mathcal{A}U|_N. \end{aligned}$$

Since  $U$  is arbitrary, it follows, by setting  $V = \mathcal{A}U$  and  $U = \mathcal{A}^{-1}V$  that

$$\|\mathcal{A}^{-1}\| = \sup_{V \in \mathbb{R}^N} \frac{|\mathcal{A}^{-1}V|_N}{|V|_N} \leq \mu_{min}^{-1} C_P^2 C_c^{-1}. \quad (20)$$

We conclude by inserting the bounds (19) and (20) into the definition of  $\kappa(\mathcal{A})$ .

**Remark 12** *Since the constants  $C_c, C_a, C_I, \mu_{min}, \mu_{max}, C_P$  all are independent of how the boundary cuts the mesh we conclude that the condition number of the system matrix  $\mathcal{A}$  is robust with respect to the position of the boundary.*

## 6 Numerical examples

In the numerical examples below, we have not integrated on the exact boundary, but on a polygonal approximation of the boundary given by a discrete representation of a level set function (on the same mesh used for the actual computation). This introduces an additional error which is not analysed here but is expected not to destroy the order of the method, as also indicated by our examples.

### 6.1 A problem with smooth solution

We consider a problem on a circular domain, from [2], with exact solution  $u = (r^3 - R^3)/9$ , where  $r$  is the distance from the center of the circle and  $R$  is the radius of the circle. We chose  $\gamma_D = 5$  and  $\gamma_1 = 0.1$ .

In Fig. 1 we show the mesh with the boundary of the domain indicated (for the third mesh in the convergence study), in Fig. 2 the corresponding meshed domain  $\Omega$ . In Figure 3 we show the obtained optimal convergence in  $L_2(\Omega)$  (second order) and in  $H^1(\Omega)$  (first order). An elevation of the solution is shown in Fig. 4, and of the interpolated error on the same mesh in Fig. 5. We note that the error is largest at the boundary, which is to be expected due to the lack of mesh symmetry there.

### 6.2 The effect of $\gamma_1$ on the condition number of the stiffness matrix

We consider two cases: one where the mesh is cut with small triangles, and one where the mesh is cut with slivers. The second case is more difficult. In

the inverse inequality

$$\|\mathbf{n}_\Gamma \cdot \nabla u_h\|_{-\frac{1}{2},h,\Gamma_D \cap T}^2 \leq C_T^2 \|\nabla u_h\|_{0,T \cap \Omega}^2$$

$C_T$  grows faster for a sliver cut than for a triangular cut. If we were to redefine  $h$  to denote the distance to the closest interior node, then the corresponding  $C_T$  would be bounded for a triangular cut but not for a sliver cut, cf. [6]. However, conditioning is also adversely affected by integrating the energy product on very small element pieces. These conditioning problems are alleviated by the introduction of the jump terms, as we shall see.

We first consider the fixed mesh with small triangular cuts in Fig. 6. The domain  $(-1, 1) \times (-1, 1)$  is cut by the line  $x + y + 0.49 = 0$ . In Fig. 7 we show the condition number as a function of  $\gamma_1$ . We note that too small  $\gamma_1$  leads to bad conditioning. The dotted line corresponds to the condition number using  $\gamma_1 = 0.01$  on a mesh where the triangles are less severely cut ( $x + y + 0.3 = 0$ , cf. Fig. 8). The quotient between the best condition number in the first case and the regular case is a factor of three. Finally, in Fig. 9 we show the condition number as a function of  $h$  corresponding to the cut  $x + y + 0.49 = 0$ ,  $\gamma_1 = 0.1$ , which shows  $O(h^{-2})$  as in the regular case.

Next, we consider the sliver cut case, Fig. 10 corresponding to  $x + 0.24 = 0$ . The condition number as a function of  $\gamma_1$  is given in Fig. 11. Note that  $\gamma_1$  has to be chosen larger in this case for moderate condition numbers. The dotted line corresponds to the line  $x + 0.125 = 0$  using  $\gamma_1 = 0.1$ , cf. Fig. 12. Again, we observe a factor of about 3 increase in condition number. In Fig. 13 we observe the expected dependence of the condition number upon  $h$  ( $\gamma_1 = 0.4$ ), and, finally, in Fig. 14 we show how the condition number depends on the distance from the vertical cut in Fig. 10 to the closest node. We note that the condition number stabilizes on a reasonable level as the distance approaches zero.

### 6.3 Sensitivity with respect to $\gamma_1$

The stabilizing parameter  $\gamma_1$  has the effect of preventing spurious jumps in the normal derivative. As  $\gamma_1$  grows, the gradient tends toward a constant on  $\Gamma$  which reduces the accuracy of the method. Thus it is important to know quantitatively how large we can set  $\gamma_1$ . To illustrate this, we again consider the domain in Section 6.1, but now with exact solution  $u = e^{10x}$ , which leads to strongly varying normal derivative on the boundary. In Figure 15 we show the effect upon  $\|u - u_h\|_{L_2(\Omega)}$ , on three consecutively refined meshes, of increasing  $\gamma_1$ . We conclude that  $\gamma_1$  should be chosen small compared to  $\gamma_D$ . Combined with the conditioning aspect it seems  $\gamma_1 = 1/2$  is a reasonable choice. However,

convergence of the method is clearly present also for larger  $\gamma_1$ , the effect is only present on a fixed mesh.

## References

- [1] R. Becker, E. Burman and P. Hansbo. A Nitsche extended finite element method for incompressible elasticity with discontinuous modulus of elasticity. *Comput. Methods Appl. Mech. Engrg.*, 198(41-44):3352–3360, 2009.
- [2] E. Burman and P. Hansbo. Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method. *Comput. Methods Appl. Mech. Engrg.*, 199(41-44):2680–2686, 2010.
- [3] R. Dautray and J.-L. Lions. Mathematical Analysis and Numerical Methods for Science and Technology. Volume 2: Functional and Variational Methods. *Springer-Verlag, Berlin*, 1988.
- [4] A. Ern and J.-L. Guermond. Theory and Practice of Finite Elements, volume 159 of *Applied Mathematical Sciences*. *Springer-Verlag, New York, NY*, 2004.
- [5] A. Hansbo and P. Hansbo. An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems. *Comput. Methods Appl. Mech. Engrg.*, 191(47-48):5537–5552, 2002.
- [6] P. Hansbo. Nitsche’s method for interface problems in computational mechanics. *GAMM-Mitt.*, 28(2):183–206, 2005.
- [7] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971.

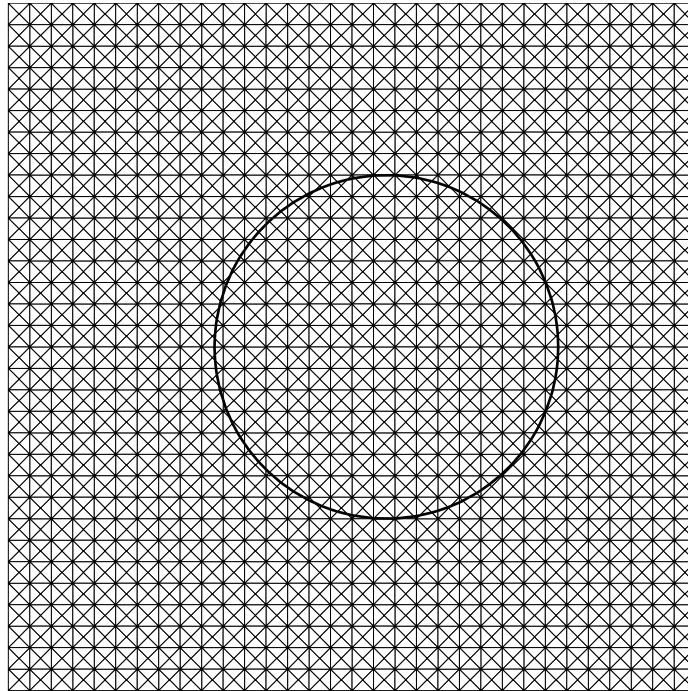


Fig. 1. The domain  $\mathcal{T}_h$  with the boundary of  $\Omega$  indicated. In the computation triangles not intersected by  $\Omega$  are disregarded.

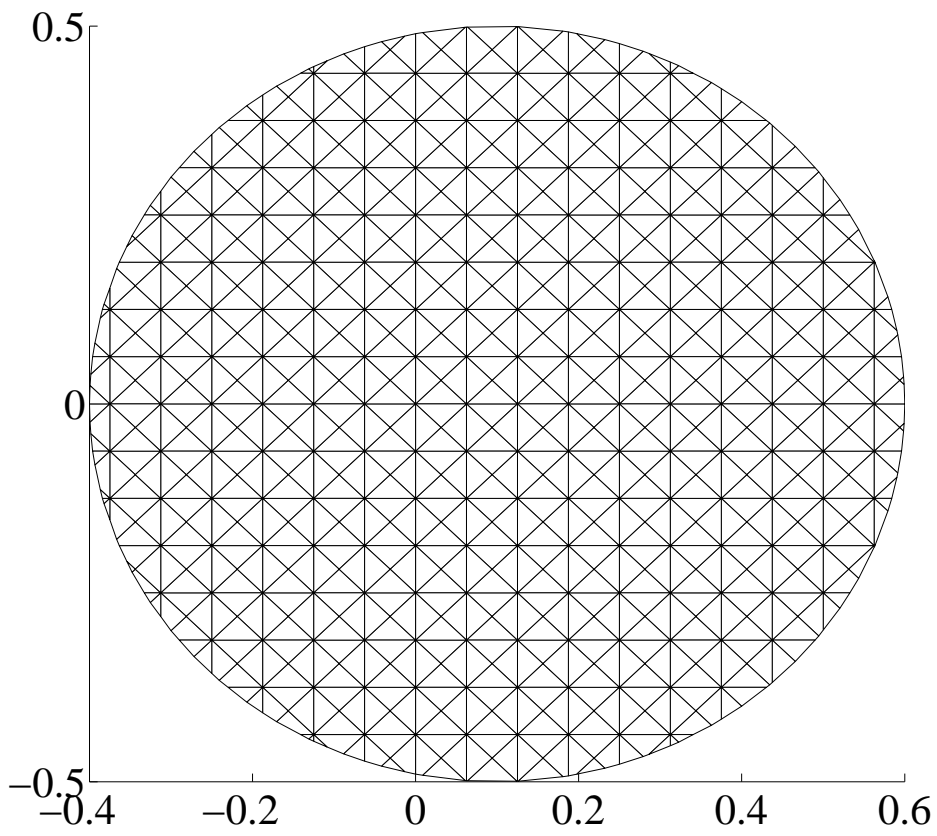


Fig. 2. The domain  $\Omega$ .



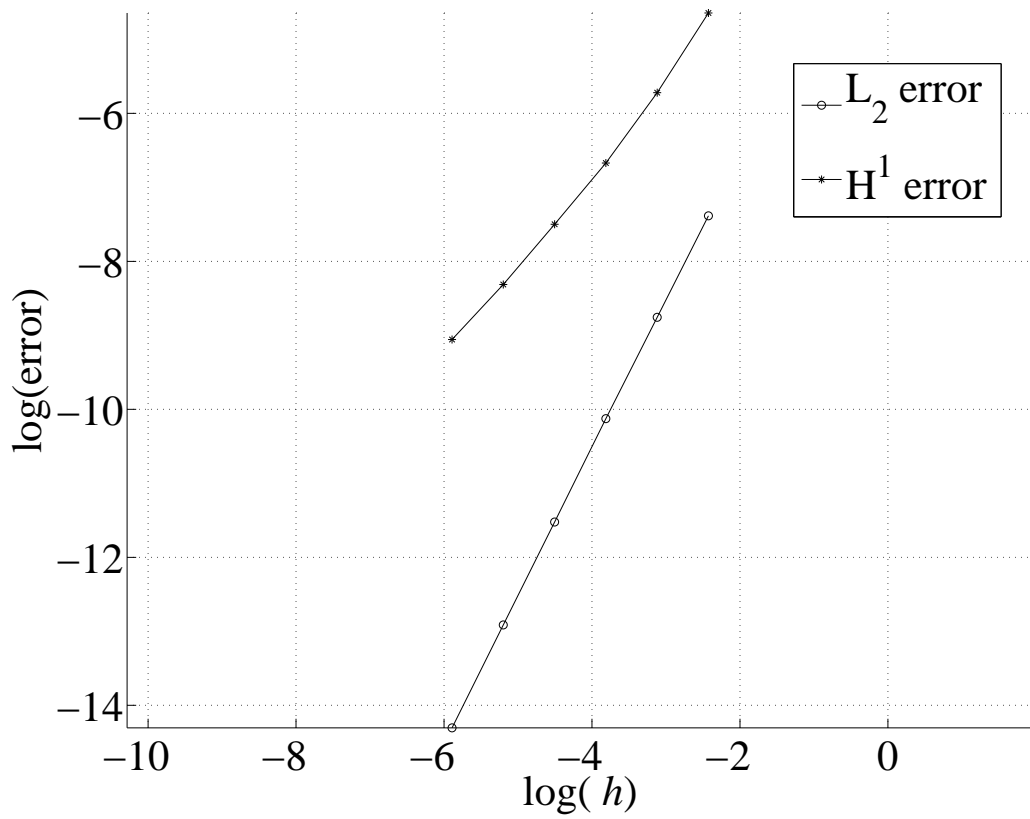


Fig. 3. Convergence in  $L_2$  and  $H^1$ .

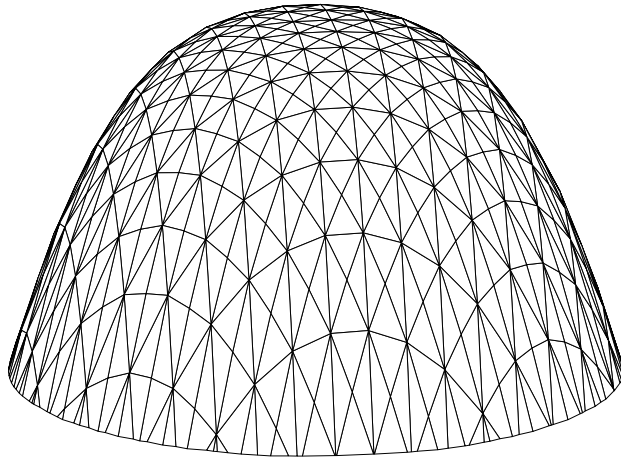


Fig. 4. Elevation of the solution.

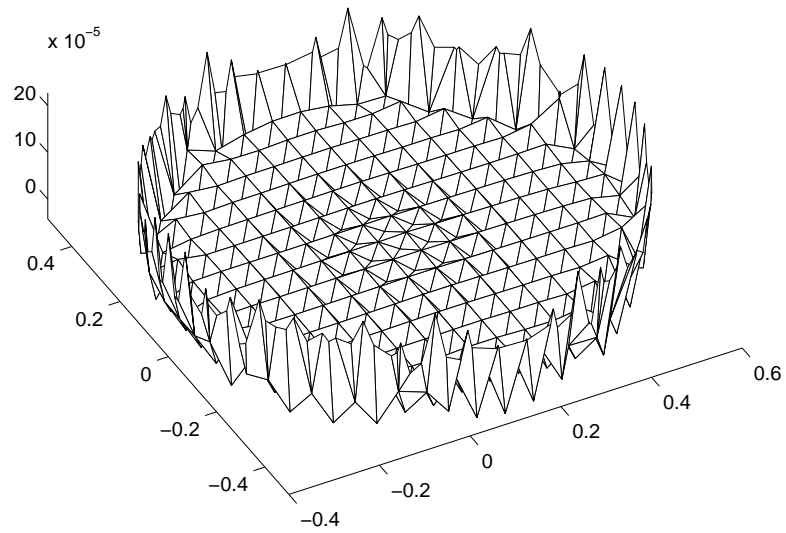


Fig. 5. Elevation of the (interpolated) error.

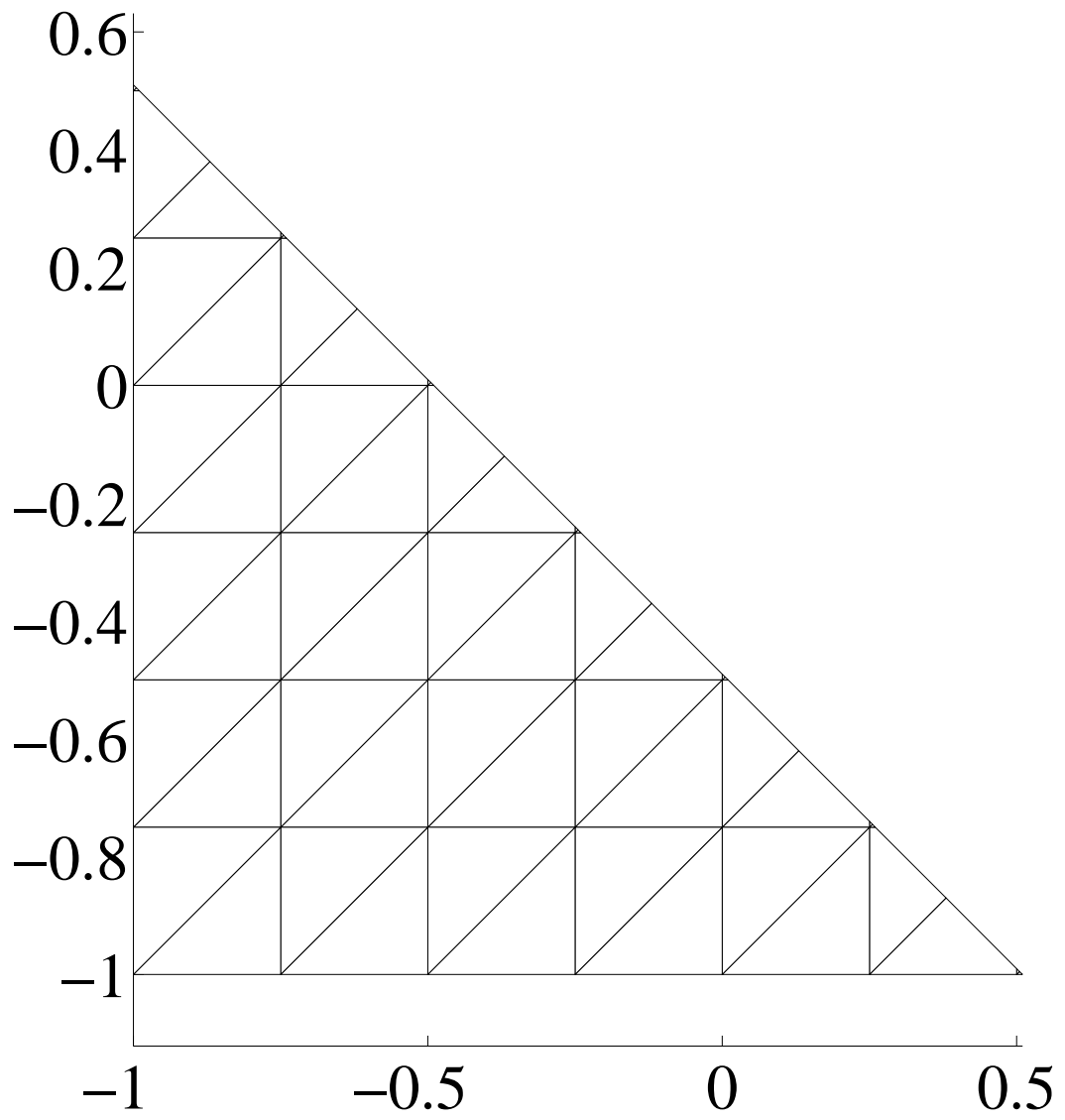


Fig. 6. Cut with small triangular pieces.

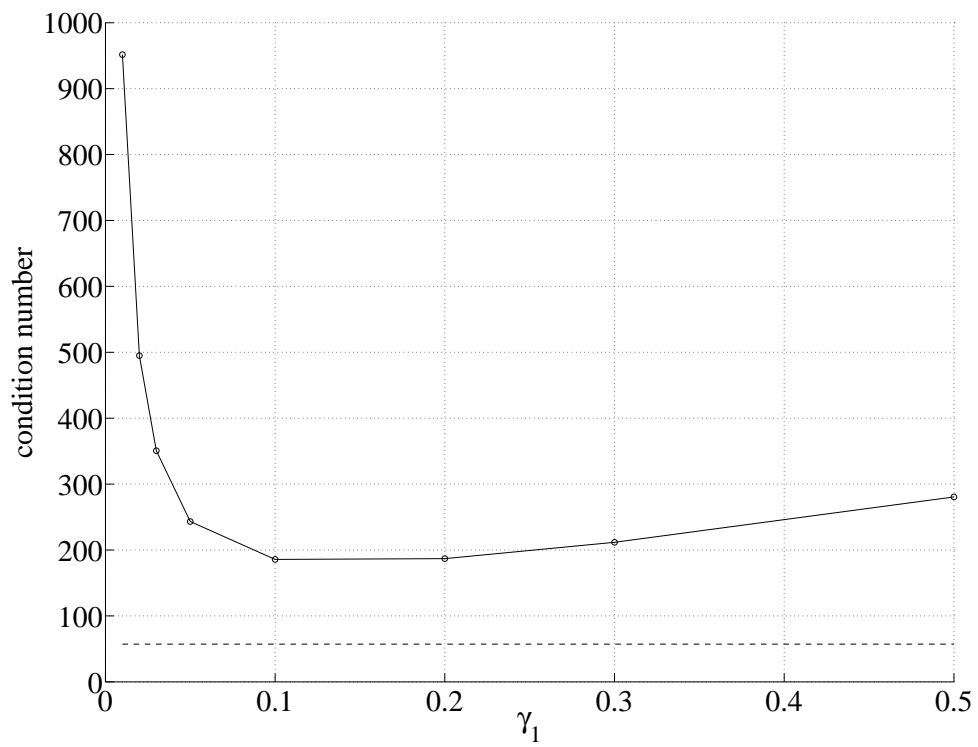


Fig. 7. Condition number as a function of  $\gamma_1$  for the triangle cut.

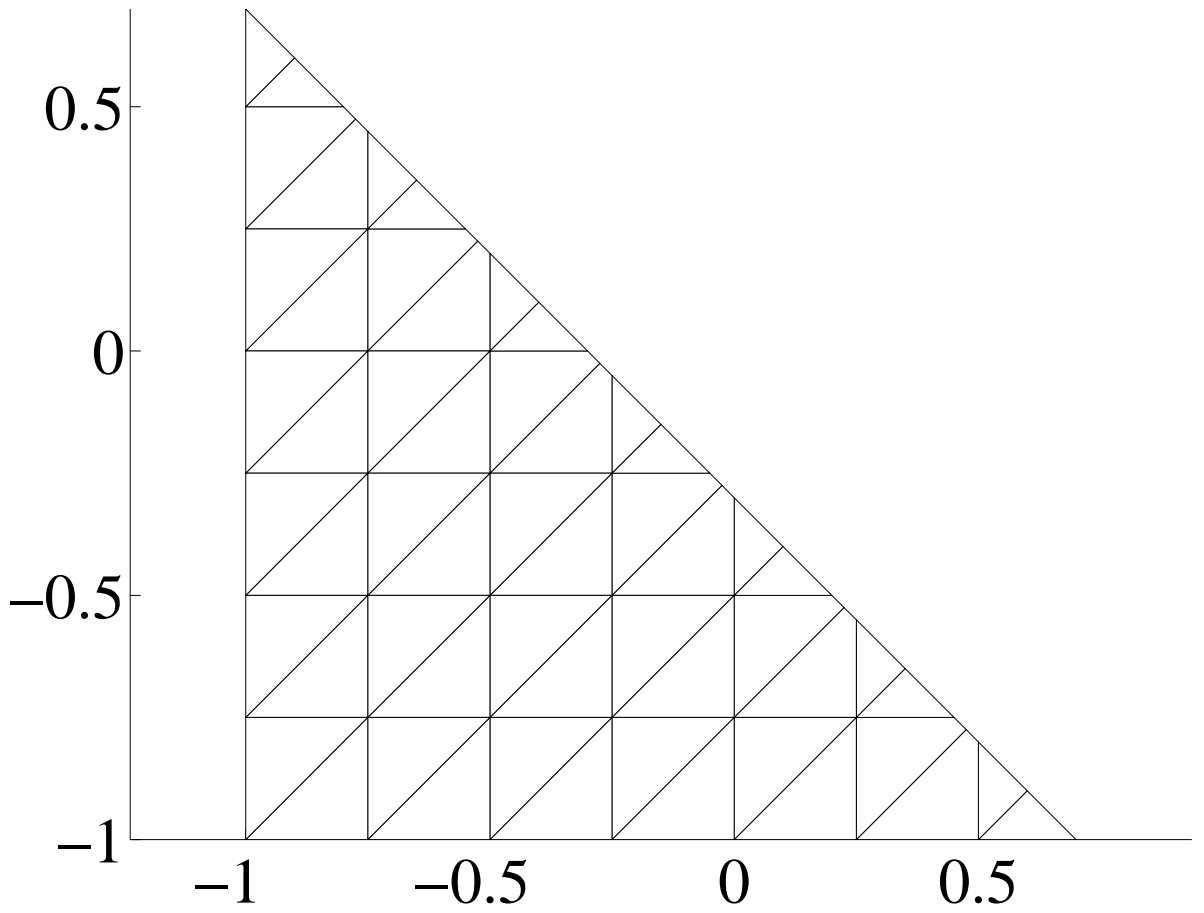


Fig. 8. Cut with large triangular pieces.

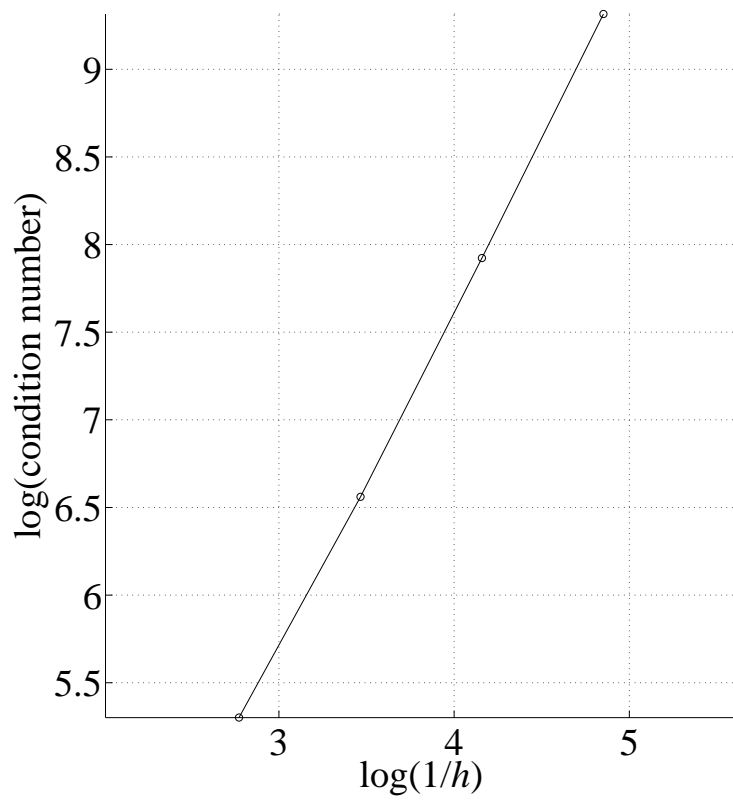


Fig. 9. Condition number as a function of  $1/h$ , triangular case.

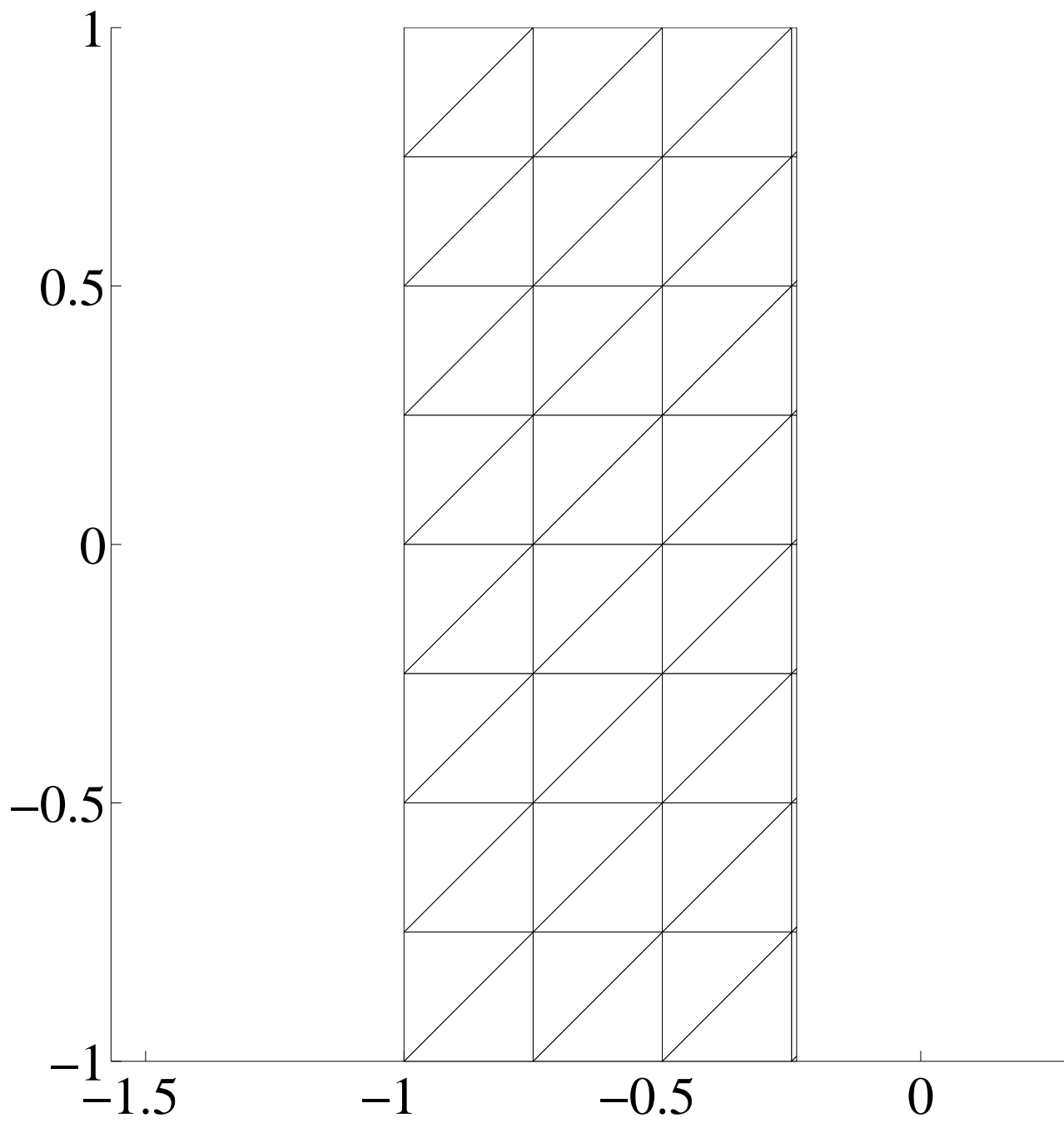


Fig. 10. Cut with small sliver pieces.

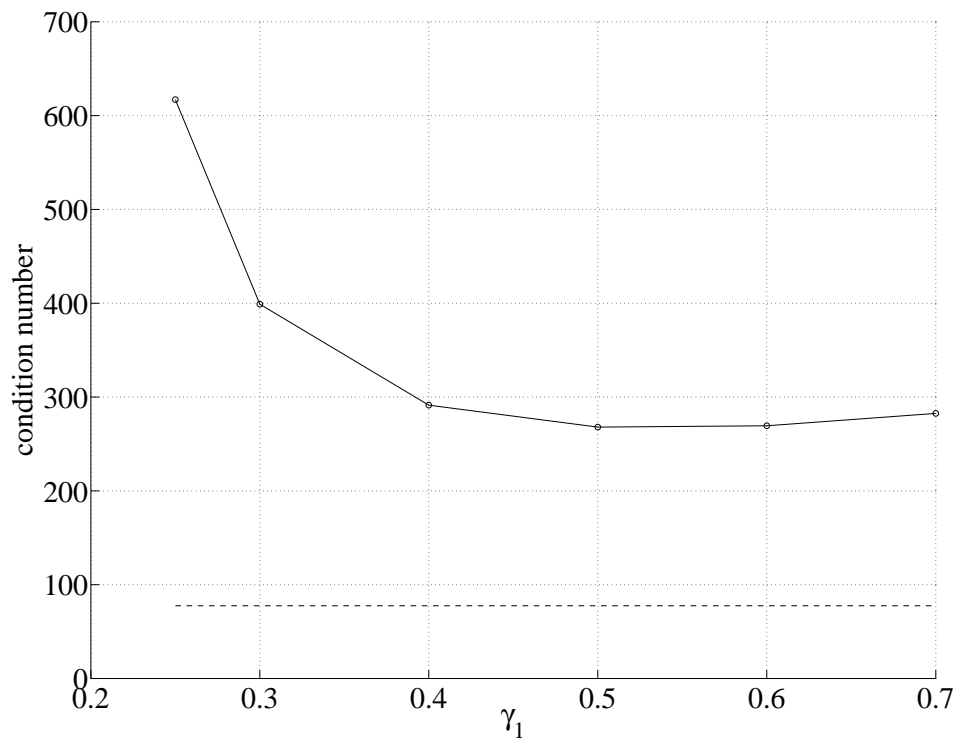


Fig. 11. Condition number as a function of  $\gamma_1$  for the sliver cut.



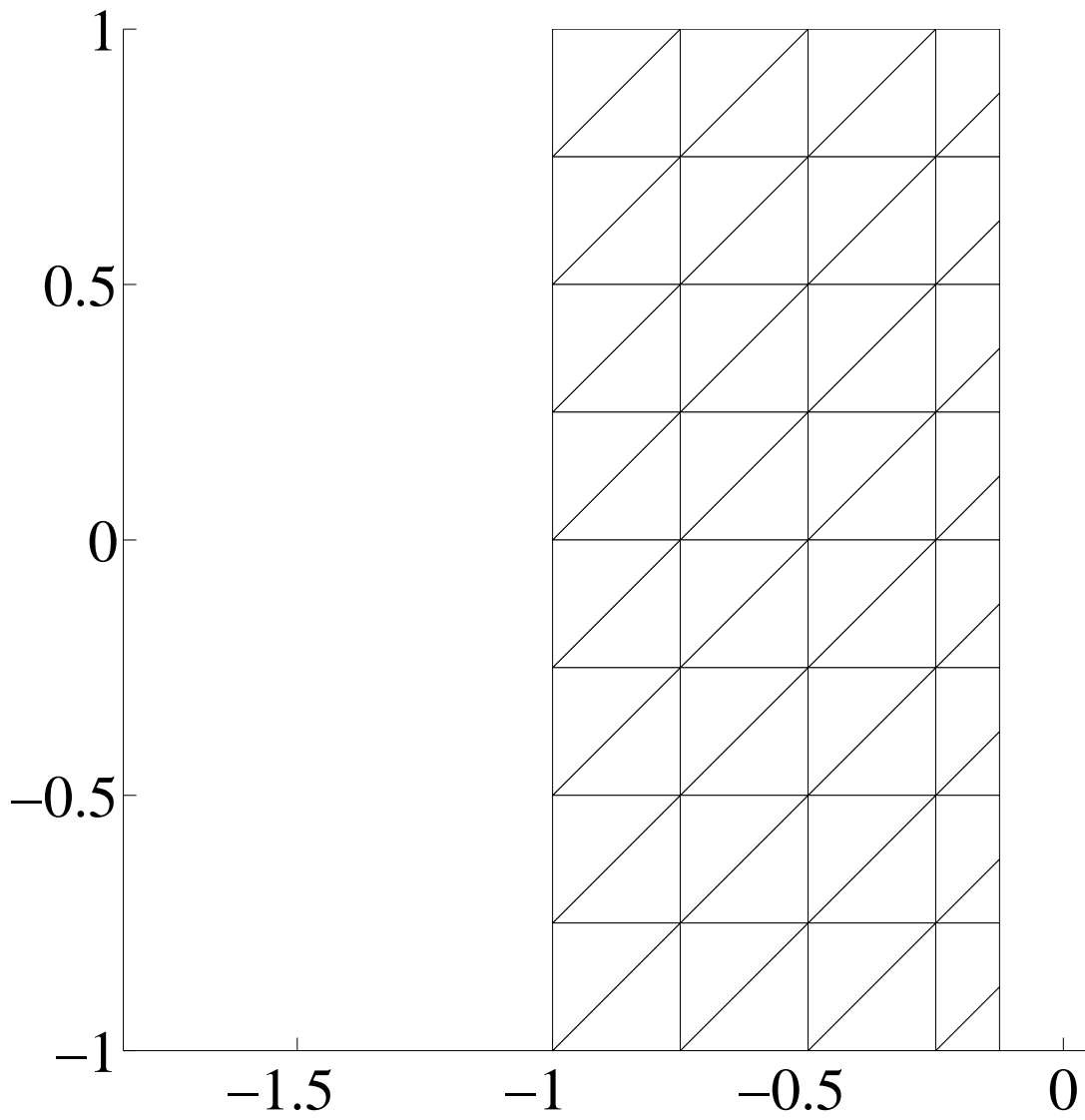


Fig. 12. Cut with large sliver pieces.

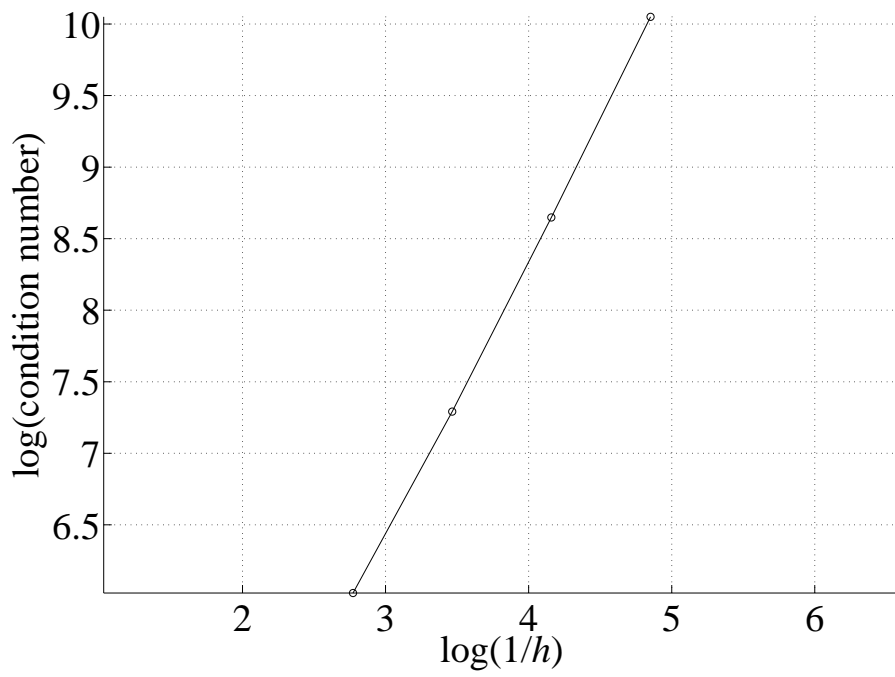


Fig. 13. Condition number as a function of  $1/h$ , sliver case.

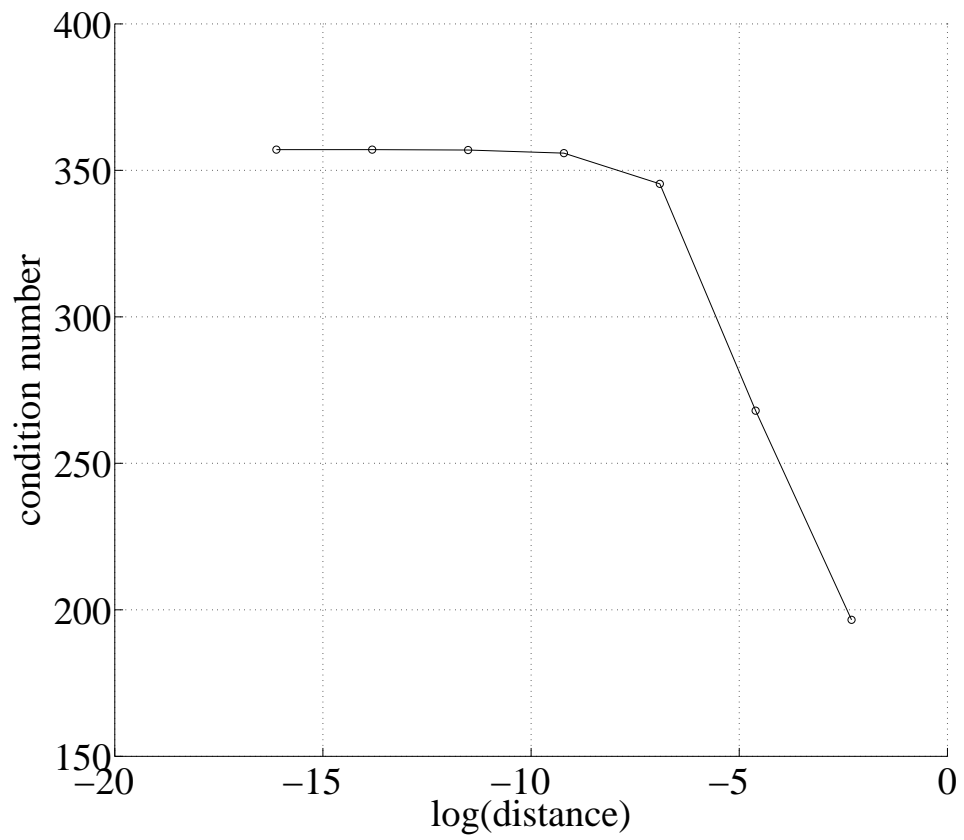


Fig. 14. Condition number as a function of the distance from the vertical cut to the closest node, sliver case.

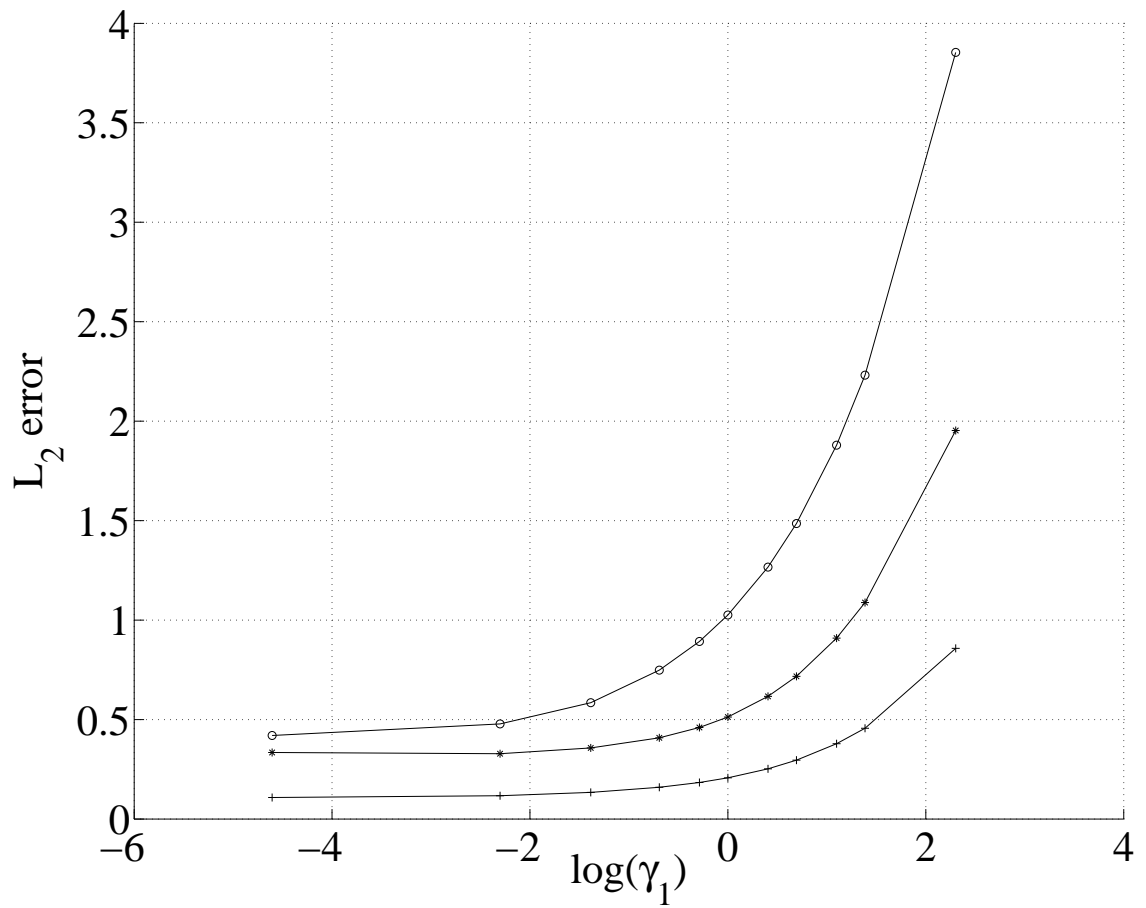


Fig. 15. Sensitivity of the error with respect to  $\gamma_1$ .