A new fictitious domain method: optimal convergence without cut elements

Alexei Lozinski

Laboratoire de Mathématiques de Besançon - UMR CNRS 6623, Univ. Bourgogne Franche Comté, 16 route de Gray, 25030 Besançon Cedex, France.

Abstract
We present a method of the fictitious domain type for the Poisson problem with either Dirichlet or Neumann/Robin boundary conditions. The computational mesh is obtained from a background (typically uniform Cartesian) mesh by retaining only the elements intersecting the domain where the problem is posed. The resulting mesh does not thus fit the boundary of the problem domain. Several finite element methods (XFEM, CutFEM) adapted to such meshes have been recently proposed. The originality of the present article consists in avoiding integration over the elements cut by the boundary of the problem domain, while preserving the optimal convergence rates, as confirmed by both the theoretical estimates and the numerical results.

Keywords: Fictitious domain, finite elements.

1. Introduction.

In this article, we propose a new approach to the numerical solution of boundary value problems for partial differential equations using finite elements on non-matching meshes, circumventing the need to generate a mesh accurately fitting the physical boundaries or interfaces. Such approaches, classically known as the fictitious domain methods, have a long history dating back to [22] in the case of the Poisson problem with Dirichlet boundary conditions. They were later popularized, by Glowinski and co-workers, cf. [14] for example, and successfully applied in the context of particular flow simulations [13, 11]. The basic idea of the classical fictitious domain method is to embed the physical domain $\Omega$ into a bigger simply shaped domain $\Omega'$, to extend the physically meaningful solution on $\Omega$ by a fictitious solution on $\Omega' \setminus \Omega$ using the same governing equations as on $\Omega$, and to impose the boundary conditions on $\partial \Omega$ by Lagrange multipliers. At the numerical level, this means that one can work with simple meshes on $\Omega'$, but one also needs a mesh on the physical boundary $\partial \Omega$ for the Lagrange multiplier, which should be coarser than the first one in order to satisfy the inf-sup condition [10]. One is thus not completely free from meshing problems. Another unfortunate feature of such methods is their poor accuracy: one cannot expect the convergence order to be better than $1/2$ with the respect to the mesh size. Closely related penalty methods are well suited for both Dirichlet and Neumann boundary conditions [12,18], may be simpler to implement in practice than the fictitious domain methods with Lagrange multipliers, but share with them the poor convergence properties.

More recently, several optimally convergent finite element methods on non-matching meshes were proposed following the XFEM or CutFEM paradigms. XFEM (extended finite element method) was initially introduced in [20] for applications in the structural mechanics on cracked domains (Neumann boundary conditions on the crack). Its ability to impose Dirichlet boundary conditions was demonstrated in [19,23] and a properly stabilized version with proved optimal convergence was proposed in [15]. The CutFEM methods [3] were first introduced in a series of papers by Burman and Hansbo [4,5,6]. The common feature of all these methods is that the simple background mesh is only used to define the finite element space (only the mesh elements having non empty intersection with $\Omega$ are kept in the computational mesh), but the solution is no longer extended from the physical domain $\Omega$ to a larger fictitious domain. The integrals over $\Omega$ are thus maintained in the finite element formulation. The boundary conditions are imposed either through Lagrange multipliers living on the same mesh as the primary solution [15,4] or by the Nitsche method [5,6] stabilized by the ghost penalty [2]. The optimal convergence, i.e. the error estimates of the same order as those for finite element methods on a comparable matching mesh, are established for all the methods above.
As already mentioned, XFEM/CutFEM methods contain the integrals over \( \Omega \) in their formulations. This can be cumbersome in practice. Citing [4] “the only remaining difficulty of implementation is the actual integration on the boundary and on parts of elements cut by the boundary. This difficulty however is expected to arise in any optimal order fictitious domain method.” We attempt in the present article to prove that this statement is not entirely true. We propose in fact an optimal order fictitious domain method that does not involve the integrals on \( \Omega \) and thus does not require to perform the integration on parts of elements cut by the boundary. Our method can be regarded as a mix between the classical fictitious domain approach and CutFEM. As in CutFEM, we use the computational mesh constructed by keeping only the elements from the background mesh on the embedding domain \( O \) having non empty intersection with \( \Omega \). On the other hand, we extend the solution from \( \Omega \) to the computational domain, which is now only slightly larger than \( \Omega \). In fact, the extension is done on a narrow band of width of order of the mesh size, contrary to the extension to entire \( O \) as in the classical fictitious domain approach. This minimizes the effect of choosing a “wrong” extension and enables us, with the help of a proper stabilization mainly borrowed from CutFEM, to preserve the optimal convergence without integration on the cut mesh elements. Unfortunately, the integration on the cut mesh elements still needs to be performed, so that some non trivial quadrature techniques are still needed. We believe however that an integration over the surface (resp. the curve) cut by the mesh is simpler to implement than that on the cut mesh elements in 3D (resp. 2D).

Let us now give a first sketch of the methods that we propose in this article. We restrict ourselves to the simple model problem: the Poisson equation with either Dirichlet boundary conditions

\[- \Delta u = f \text{ in } \Omega, \quad u = g \text{ on } \Gamma \]

or Robin ones

\[- \Delta u = f \text{ in } \Omega, \quad \frac{\partial u}{\partial n} + \rho u = g \text{ on } \Gamma \]

with \( \rho \geq 0 \) (Neumann boundary conditions are included as a special case for \( \rho = 0 \)). Here \( \Omega \subset \mathbb{R}^d, d = 2, 3 \) is a domain with smooth boundary \( \Gamma = \partial \Omega \), \( f \) and \( g \) are given functions on \( \Omega \) and \( \Gamma \) respectively.

We start by embedding \( \Omega \) into a simply shaped domain \( O \) and introduce a quasi-uniform mesh \( T_h^O \) on \( O \) consisting of triangles/tetrahedrons of maximum diameter \( h \) that can be cut by the boundary \( \Gamma \) in an arbitrary manner. The computational mesh \( T_h \) is obtained from \( T_h^O \) by dropping out all the mesh element lying completely outside \( \Omega \):

\[
T_h = \{ T \in T_h^O : T \cap \Omega \neq \emptyset \}, \quad \Omega_h = (\cup_{T \in T_h})^c
\]

as illustrated in Fig. [1] \( \Omega_h \) is thus the domain occupied by the computational mesh, slightly larger than \( \Omega \). We denote its boundary by \( \Gamma_h := \partial \Omega_h \).

We outline first the derivation of our method in the case of Dirichlet boundary conditions, following the seminal article in [17]. We assume that the right-hand side \( f \) is extended smoothly from \( \Omega \) to \( \Omega_h \) and imagine (for the moment) that (1) can be solved on the extended domain \( \Omega_h \) while still imposing the boundary conditions on \( \Gamma \):

\[- \Delta u = f \text{ in } \Omega_h, \quad u = g \text{ on } \Gamma. \]

We keep here the same notations \( u \) and \( f \) for the functions on \( \Omega_h \) as for the originals on \( \Omega \). Integration by parts over \( \Omega_h \) and imposing the boundary conditions weakly on \( \Gamma \) as in the Nitsche method [21] yields

\[
\int_{\Omega_h} \nabla u \cdot \nabla v - \int_{\Omega_h} \frac{\partial u}{\partial n} v + \int_{\Gamma_h} \frac{\partial u}{\partial n} v + \frac{\gamma}{h} \int_{\Gamma} w v = \int_{\Omega_h} f v + \int_{\Gamma_h} \frac{\gamma}{h} \int_{\Gamma} g v, \quad \forall v \in H^1(\Omega_h)
\]

with \( \gamma > 0 \). Here, \( n \) on \( \Gamma \) (resp. \( \Gamma_h \)) denotes the unit normal looking outwards from \( \Omega \) (resp. \( \Omega_h \)). Our finite element method is then based on the weak formulation (5) adding to it a ghost penalty stabilization to assure the coerciveness of the bilinear form. The method will be fully presented in Subsection 2.1 and analyzed in Subsections 3.2 and 3.3.

\(^1\)The idea of constructing numerically a smooth extension to the whole \( O \) is explored in [8] resulting in an optimally convergent method. The price to pay is the necessity to solve an optimization problem by an iterative process, which can be expensive in practice.
Figure 1: Left: the “physical” domain \( \Omega \) represented by its boundary \( \Gamma \) embedded into a rectangle \( O \) with the “background” mesh \( T^O \) on it. Center: the computational mesh \( T_h \) obtained by dropping the unnecessary triangles from \( T^O \). Right: the band \( \Omega_{\Gamma}^h \) occupied by the cut elements \( T_{\Gamma}^h \) of the mesh \( T_h \).

We stress again that formulation (5) does not contain integrals on \( \Omega \). Otherwise, the resulting method is very close to the antisymmetric Nitsche CutFEM method from [2].

We turn now to the case of Robin boundary conditions (2). We start again by extending \( f \) from \( \Omega \) to \( \Omega_h \) and imagine (for the moment) that (2) can be solved on the extended domain \( \Omega_h \) while still imposing the boundary conditions on \( \Gamma \):

\[
-\Delta u = f \text{ in } \Omega_h, \quad \frac{\partial u}{\partial n} + \rho u = g \text{ on } \Gamma.
\]

Integration by parts over \( \Omega_h \) and imposing the boundary conditions weakly on \( \Gamma \) would yield

\[
\int_{\Omega_h} \nabla u \cdot \nabla v - \int_{\Gamma_h} \frac{\partial u}{\partial n} v + \int_{\Gamma} \rho u v = \int_{\Omega_h} f v + \int_{\Gamma} g v
\]

Such a formulation does not seem to lead to a reasonable finite element method: it is difficult to imagine a stabilization that would make the bilinear form on the left-hand side above coercive. Fortunately, we are able to devise an optimally convergent method by introducing a reconstruction \( y = -\nabla u \) of the gradient of \( u \) on the band of cut mesh elements \( \Omega_{\Gamma}^h \), cf. Fig. 1 on the right, and by replacing the weak formulation above with the following one

\[
\int_{\Omega_h} \nabla u \cdot \nabla v + \int_{\Gamma_h} v \cdot y + \int_{\Gamma} \rho u v + \gamma_1 \int_{\Omega_{\Gamma}^h} (y + \nabla u) \cdot (z + \nabla v) \quad \forall v \in H^1(\Omega_h), z \in L^2(\Omega_{\Gamma}^h)
\]

with \( \gamma_1 > 0 \). The terms multiplied by \( \gamma_1 \) serve to impose \( y = -\nabla u \) on \( \Omega_{\Gamma}^h \). Our finite element method (12), based on the weak formulation (7) with the addition of grad-div and ghost stabilizations, will be fully presented in Subsection 2.2 and analyzed in Subsections 3.4 and 3.5. It involves the additional vector-valued unknown \( y_h \), which results in some extra cost in practice, as compared with a simpler method for Dirichlet boundary conditions. Fortunately, this extra cost is negligible as \( h \to 0 \) since \( y_h \) lives only on the mesh elements cut by \( \Gamma \) which constitute smaller and smaller proportion of the total number of elements as the meshes are refined.

Our finite element methods for both Dirichlet and Robin cases are presented in more detail in the next section. Their well-posedness and the optimal error estimates are proved in Section 3. We restrict ourselves here to \( P_1 \) continuous finite elements on a triangular/tetrahedral mesh. An extension to higher-order elements \( P_k \) is sketched in Section 4. We conclude by numerical experiments in Section 5, presenting only the results with \( P_1 \) elements.
2. Presentation of the methods

2.1. Dirichlet boundary conditions

We present first our discretization of Problem (1). Recall that $T_h$ is a quasi-uniform mesh obtained from a larger background mesh retaining only the elements lying inside $\Omega$ or cut by $\Gamma$, cf. [3] and Fig. 1, and $\Omega_h$ is the corresponding domain with boundary $\Gamma_h$. We inspire ourselves from the variational formulation (5), introduce the finite element space

$$V_h = \{v_h \in H^1(\Omega_h) : v_h|_T \in \mathbb{P}_1(T) \forall T \in T_h\}$$

with $\mathbb{P}_1$ denoting the set of polynomials of degree $\leq 1$, and introduce the following discrete problem:

Find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = \int_{\Omega_h} f v_h + \int_{\Gamma} g \frac{\partial v_h}{\partial n} + \gamma h \int_{\Gamma} g v_h \forall v_h \in V_h$$

where

$$a_h(u, v) = \int_{\Omega_h} \nabla u \cdot \nabla v - \int_{\Gamma} \frac{\partial u}{\partial n} v + \int_{\Gamma} u \frac{\partial v}{\partial n} + \gamma h \int_{\Gamma} uv + \sigma h \sum_{T \in T_h} \int_{\partial T} \left[ \frac{\partial u}{\partial n} \right] \left[ \frac{\partial v}{\partial n} \right]$$

and $\gamma, \sigma$ are some positive numbers properly chosen in a manner independent of $h$. The last term in (10) is the ghost penalty [2]. It is crucial to assure the coerciveness of $a_h$. The notations here are as follows: $[\cdot]$ stands for the jump over an internal facet of mesh $T_h$ and

$$T_\Gamma = \{E \text{ (an internal facet of } T_h) \text{ such that } \exists T \in T_h : T \cap \Gamma \neq \emptyset \text{ and } E \in \partial T\}$$

The ghost penalty is thus a properly scaled sum of the jumps of the normal derivatives over all the internal facets of the mesh either cut by $\Gamma$ themselves or owned by a mesh element cut by $\Gamma$.

2.2. Robin boundary conditions

We turn now to Problem (2). We inspire ourselves with the variational formulation (7) and introduce (along with $V_h$ as above) the auxiliary finite element space

$$Z_h = \{z_h \in H^1(\Omega_h)^d : z_h|_T \in \mathbb{P}_1(T)^d \forall T \in T_h\}$$

Here $T_h^\Gamma$ represents the cut elements of the mesh and $\Omega_h^\Gamma$ the corresponding subdomain of $\Omega_h$:

$$T_h^\Gamma = \{T \in T_h : T \cap \Gamma = \emptyset\} \text{ and } \Omega_h^\Gamma = (\bigcup_{T \in T_h^\Gamma} T)^\circ$$

We shall also denote by $\Gamma_h^\Gamma$ the internal boundary of $\Omega_h^\Gamma$, i.e. the ensemble of the facets separating $\Omega_h^\Gamma$ from the mesh elements inside $\Omega$, so that $\partial \Omega_h^\Gamma = \Gamma_h \cup \Gamma_h^\Gamma$.

Our finite element problem is: Find $u_h \in V_h, y_h \in Z_h$ such that

$$d_h^R(u_h, y_h; v_h, z_h) = \int_{\Omega_h} f v_h + \int_{\Gamma} g v_h + \gamma_{\text{div}} \int_{\Omega_h^\Gamma} \text{div } z_h \quad \forall (v_h, z_h) \in V_h \times Z_h$$

where

$$d_h^R(u, y; v, z) = \int_{\Omega_h} \nabla u \cdot \nabla v + \int_{\Gamma} y \cdot \nabla v - \int_{\Gamma} y \cdot \nabla u + \int_{\Gamma} \rho uv + \gamma_{\text{div}} \int_{\Omega_h^\Gamma} \text{div } y \text{ div } z$$

$$+ \gamma_1 \int_{\Gamma_h^\Gamma} (y + \nabla u) \cdot (z + \nabla v) + \sigma h \sum_{T \in T_h} \int_{\partial T} \left[ \frac{\partial u}{\partial n} \right] \left[ \frac{\partial v}{\partial n} \right]$$

with some positive numbers $\gamma_{\text{div}}, \gamma_1$ and $\sigma$ properly chosen in a manner independent of $h$. In addition to the variational formulation (7), we have introduced here

- a grad-div stabilization (the terms multiplied by $\gamma_{\text{div}}$) in the vein of [9], which is consistent with the governing equations since $y = -\nabla u$ and thus $\text{div } y = -\Delta u = f$ on $\Omega_h^\Gamma$;
Assumption 1. The subdomain \( \Omega_h \) is typically satisfied if the mesh is sufficiently refined with respect to \( \Gamma \).

Remark 2.1. The constant function \( u = 1 \) accompanied by \( \gamma = 0 \) is well in the kernel of \( d_\nu^3 \) if \( \rho = 0 \) reflecting the fact that the solution is not unique on both continuous and discrete levels in the case of Neumann boundary conditions. To treat properly the special case \( \rho = 0 \) we should change the definition of \( V_h \) by imposing \( \int_{\Gamma_h} v_h = 0 \). In order to keep the exposition brief, we shall mostly suppose from now on \( \rho > 0 \) and outline necessary modifications for \( \rho = 0 \) in the remarks.

3. Theoretical error analysis.

3.1. Geometrical assumptions and technical lemmas.

The theoretical analysis of the methods presented above will be done under the following assumptions on the boundary \( \Gamma \) and on the subdomain \( \Omega_h \) covered by the cut mesh elements, as defined in (11). Both these assumptions are typically satisfied if the mesh is sufficiently refined with respect to \( \Gamma \).

Assumption 1. The subdomain \( \Omega_h \) can be covered by open sets \( O_k \), \( k = 1, \ldots, K \) and one can introduce on every \( O_k \) local coordinates \( \xi_1, \xi_2 \) in the 2D case, \( d = 2 \) (resp. \( \xi_1, \xi_2, \xi_3 \) if \( d = 3 \)) such that

- \( \Gamma \cap O_k \) is given by \( \xi_d = 0 \) and \( \Omega \cap O_k \) by \( \xi_d < 0 \);
- \( \Omega_h \cap O_k \) is given by \(-a(\xi_1, \xi_2) < \xi_d < b(\xi_1, \xi_2)\) with some continuous non-negative functions \( a \) and \( b \);
- \( b(\xi_1, \xi_2) + a(\xi_1, \xi_2) \leq C_1 h \) with some \( C_1 > 0 \);
- all the partial derivatives \( \partial \xi_i / \partial \xi_j \) and \( \partial \xi_i / \partial x_j \) are bounded by some \( C_2 > 0 \).

Moreover, each point of \( \Omega_h \) is covered by at most \( N_{int} \) sets \( O_k \).

Assumption 2. The boundary \( \Gamma \) can be covered by element patches \( \{ \Pi_k \}_{k=1}^{N_{int}} \) having the following properties:

- Each \( \Pi_k \) is a connected set;
- Each \( \Pi_k \) is composed of a mesh element \( T_k \) lying inside \( \Omega \) and some mesh elements cut by \( \Gamma \), more precisely \( \Pi_k = T_k \cup \bigcup_{i=1}^{M} \Pi^i_k \) where \( T_k \in \mathcal{T}_h \), \( T_k \subset \Omega \), \( \Pi^i_k \subset \mathcal{T}_h \), and \( \Pi^i_k \) contains at most \( M \) mesh elements;
- \( \mathcal{T}_h = \bigcup_{k=1}^{N_{int}} \Pi^0_k \);
- \( \Pi_k \) and \( \Pi_l \) are disjoint if \( k \neq l \).

In what follows, we suppose both assumptions above to hold true and use the notation \( C \) for positive constants (which can change from one instance to another) that depend only on \( C_1 \), \( C_2 \), \( N_{int} \), \( M \) from the assumptions above and on the mesh regularity. We also recall that mesh \( \mathcal{T}_h \) is supposed quasi-uniform.

Assumption 1 is quite standard and essentially tells us that the boundary \( \Gamma \) is smooth and not too wiggly on the length scale \( h \), so that one can be sure that the band of cut mesh elements is of width \( \sim h \). Assumption 2 is slightly more technical. Let us explain the construction of element patches evoked there, cf. Fig. 2. Each patch \( \Pi_k \) is assigned to a facet \( E_k \subset \Gamma_h \) separating the cut elements from the interior ones. To form the patch \( \Pi_k \), one takes a mesh element \( T_k \) lying inside \( \Omega \) and attached to \( E_k \). One then picks up several cut elements touching \( E_k \) to form \( \Pi^i_k \) and set \( \Pi_k = T_k \cup \Pi^i_k \). If, again, the boundary \( \Gamma \) is not too wiggly with respect to the mesh \( \mathcal{T}_h \), one can partition the cut elements between the patches \( \Pi_k \) so that each patch contains a small number of elements (typically from 2 to 4 in 2D, and slightly more in 3D).

We begin with two technical lemmas. The first one is in the vein of Poincaré inequalities taking into account the assumption that \( \Omega_h^\Gamma \) is a strip of width \( \sim h \) around \( \Gamma \).

Lemma 3.1. For all \( v \in H^1(\Omega_h^\Gamma) \),

\[
\|v\|_{0,\Omega_h^\Gamma} \leq C \left( \sqrt{h\|v\|_{0,\Gamma}} + h\|v\|_{1,\Gamma} \right)
\]
Proof. Consider the 2D case \((d = 2)\). Recalling Assumption \(\text{[1]}\) we can pass to the local coordinates \(\xi_1, \xi_2\) on every set \(O_k\) assuming that \(\xi_1\) varies between 0 and \(L\), and to use the bounds on \(\xi_2\) and on the mapping \((x_1, x_2) \mapsto (\xi_1, \xi_2)\) to write

\[
\|v\|_{0, \Omega_k^{\xi_1}}^2 \leq C \int_0^L \int_{-a(\xi_1)}^{b(\xi_1)} v^2(\xi_1, \xi_2) d\xi_2 d\xi_1 = C \int_0^L \int_{-a(\xi_1)}^{b(\xi_1)} \left( v(\xi_1, 0) + \int_0^{\xi_2} \frac{\partial v}{\partial \xi_2}(\xi_1, t) dt \right)^2 d\xi_2 d\xi_1
\]

\[
\leq C \int_0^L \int_{-a(\xi_1)}^{b(\xi_1)} v^2(\xi_1, 0) + \left( \int_0^{\xi_2} \frac{\partial v}{\partial \xi_2}(\xi_1, t) dt \right)^2 d\xi_2 d\xi_1
\]

\[
\leq C \int_0^L \left( h v^2(\xi_1, 0) + h^2 \int_{-a(\xi_1)}^{b(\xi_1)} \left| \frac{\partial v}{\partial \xi_2}(\xi_1, t) \right|^2 dt \right) d\xi_1
\]

\[
\leq C h \|v\|_{0, \Gamma}^2 + C h^2 \|\nabla v\|_{0, \Omega_k^{\xi_1}}^2. 
\]

Summing over all neighborhoods \(O_k\) gives \(\|v\|_{0, \Omega_k} \leq C (\|v\|_{1, \Omega} + \|v\|_{0, \Gamma})\). The proof in the 3D case is the same up to the change of notations. \(\square\)

**Corollary 3.2.** For all \(v \in H^1(\Omega_k^\xi)\),

\[
\|v\|_{0, \Omega_k} \leq C (\|v\|_{1, \Omega_k} + \|v\|_{0, \Gamma}) 
\]

**Proof.** We start by a Poincaré-like inequality (easily proved via Petree-Tartar lemma)

\[
\|v\|_{0, \Omega} \leq C (\|v\|_{1, \Omega} + \|v\|_{0, \Gamma})
\]

We now combine it with Lemma \(\text{[3.1]}\)

\[
\|v\|_{0, \Omega_k} \leq \|v\|_{0, \Omega} + \|v\|_{0, \Omega_k^\xi} \leq C \left( \|v\|_{1, \Omega} + h \|v\|_{1, \Omega_k^\xi} + \left( 1 + \sqrt{h} \right) \|v\|_{0, \Gamma} \right)
\]

and conclude noting that \(h \leq \text{diam} (\Omega)\). \(\square\)

**Lemma 3.3.** For all \(v \in H^2(\Omega_h^\xi)\),

\[
\|v\|_{0, \Gamma} \leq C \left( \|v\|_{0, \Gamma} + h \|\nabla v\|_{0, \Gamma} + h^{3/2} \|v\|_{2, \Omega_h^\xi} \right)
\]

**Proof.** As in the proof of Lemma \(\text{[3.1]}\) we only consider the 2D case. Using again Assumption \(\text{[1]}\) we can represent \(v|_{\Gamma_h}\)

![Figure 2: Illustration of the construction of an element patch Πk: a portion of a triangular mesh Tk with the facets composing Πk represented by solid lines, the facets separating the cut triangles from the interior ones (Γk) represented by thick dashed lines, and the remaining facets in thin dashed lines. The patch Πk is composed of the interior element Tk ⊂ Ω attached to a facet Ek ⊂ Γk and several triangles cut by Γ (the remaining part of Πk denoted by Πk).](image-url)
inside every $O_k$ by a Taylor expansion with the integral remainder

$$\|v\|_{0, \Gamma; \partial \Omega_k}^2 \leq C \int_0^L v^2(\xi_1, b(\xi_1)) d\xi_1 = C \int_0^L \left( v(\xi_1, 0) + b(\xi_1) \frac{\partial v}{\partial \xi_2}(\xi_1, 0) + \int_0^{\xi_1} \frac{\partial^2 v}{\partial \xi_2^2}(\xi_1, t)(b(\xi_1) - t) dt \right)^2 d\xi_1$$

$$\leq C \int_0^L \left( v^2(\xi_1, 0) + h^2 \frac{\partial v}{\partial \xi_2}(\xi_1, 0)^2 + h^3 \left[ \int_0^{\xi_1} \frac{\partial^2 v}{\partial \xi_2^2}(\xi_1, t)^2 dt \right] \right) d\xi_1$$

$$\leq C(\|v\|_{0, \Gamma; \partial \Omega_k}^2 + h^2\|\nabla v\|_{0, \Gamma; \partial \Omega_k}^2 + h^3\|v\|_{2, \Gamma; \partial \Omega_k}^2)$$

Summing this over all $O_k$ gives (15).

**Lemma 3.4.** For all $v \in H^1(\Omega_h^\Gamma)$,

$$\sum_{E \in \mathcal{T}_h} \|v\|_{0,E}^2 \leq C(\|v\|_{0,T}^2 + h\|v\|_{1,T}^2).$$

and

$$\|v\|_{0,E}^2 \leq C(\|v\|_{0,T}^2 + h\|v\|_{1,T}^2).$$

**Proof.** Let $E$ be a mesh facet belonging to a mesh element $T \in \mathcal{T}_h$. Recall the well-known trace inequality

$$\|v\|_{0,E}^2 \leq C \left( \frac{1}{h} \|v\|_{0,T}^2 + h\|v\|_{1,T}^2 \right)$$

Summing this over all $E \in \mathcal{T}_h$ gives

$$\sum_{E \in \mathcal{T}_h} \|v\|_{0,E}^2 \leq C \left( \frac{1}{h} \|v\|_{0,T}^2 + h\|v\|_{1,T}^2 \right)$$

leading, in combination with (14), to (16). The proof of (17) is similar: it suffices to take the sum over the facets composing $\Gamma_h$.

3.2. **Dirichlet boundary conditions: coerciveness of $a_h$.**

We turn to the study of method (9)-(10) for the problem with Dirichlet boundary conditions. Our first goal is to prove the coerciveness of the bilinear form $a_h$ uniformly in $h$. The proof of this result, cf. Lemma 3.5, will be based on the following

**Lemma 3.5.** For any $\beta > 0$ one can choose $0 < \alpha < 1$ depending only on the mesh regularity and on parameter $M$ from Assumption 2 such that

$$|v|_{1, \Pi_h}^2 \leq \alpha |v|_{1, \Pi_h}^2 + \beta h \sum_{E \in \mathcal{T}_h} \left\| \frac{\partial v}{\partial n} \right\|_{0,E}^2$$

for all $v_h \in V_h$.

**Proof.** Choose any $\beta > 0$, consider the decomposition of $\Omega_h^\Gamma$ in element patches ($\Pi_k$) as in Assumption 2, cf. also Fig. 2. Introduce

$$\alpha := \max_{\Pi_k, \forall \pi \in \Pi_k} \frac{|v|_{1, \Pi_h}^2 - \beta h \sum_{E \in \mathcal{T}_h} \left\| \frac{\partial v}{\partial n} \right\|_{0,E}^2}{|v|_{1, \Pi_h}^2}$$

where the maximum is taken over all the possible configurations of a patch $\Pi_k$ allowed by the mesh regularity and over all the piecewise linear functions on $\Pi_k$. The subset $\mathcal{T}_h \subset \mathcal{T}_h$ gathers the facets internal to $\Pi_k$. Note that the quantity under the max sign in (3.2) is invariant under the scaling transformation $x \mapsto hx$ and is homogeneous with respect to $v_h$. Recall also that the patch $\Pi_k$ contains at most $M$ elements. Thus, the maximum is indeed attained since it is taken over a bounded set in a finite dimensional space (one can restrict the maximization to patches of diameter of order 1, since the quantity to maximize is invariant under rescaling).
Clearly, $\alpha \leq 1$. Suppose $\alpha = 1$ would lead to a contradiction. Indeed, if $\alpha = 1$ then we can take $\Pi_k$, $v_h$ yielding this maximum and suppose without loss of generality $|v_h|_{1,\Pi_k} = 1$. We observe then

$$|v_h|_{1, T_k}^2 + \beta h \sum_{E \in T_k} \left\| \frac{\partial v_h}{\partial n} \right\|^2_{0, E} = 0$$

since $|v_h|_{1, \Pi_k} = |v_h|_{1, T_k} + |v_h|_{1, \Pi_k}$. This implies $\nabla v_h = 0$ on $T_k$ and $\left\| \frac{\partial v_h}{\partial n} \right\|^2_{0, E} = 0$ on all $E \in T_k$, thus $\nabla v_h = 0$ on $\Pi_k$, which contradicts $|v_h|_{1, \Pi_k} = 1$.

This proves $\alpha < 1$. We have thus

$$|v_h|_{1, \Pi_k}^2 \leq \alpha |v_h|_{1, \Pi_k}^2 + \beta h \sum_{E \in \Pi_k} \left\| \frac{\partial v_h}{\partial n} \right\|^2_{0, E}$$

for all $v_h \in V_h$ and all the admissible patches $\Pi_k$. Summing this over $\Pi_k$, $k = 1, \ldots, N_{\Pi}$ yields (18).

**Lemma 3.6.** Provided $\sigma$ is sufficiently big, there exists an $h$-independent constant $c > 0$ such that

$$a(v_h, v_h) \geq c\|v_h\|_{h}^2$$

with

$$\|v_h\|_{h}^2 = |v_h|^2_{0, \Omega} + \frac{1}{h}|\nabla \Pi v_h|^2_{\Gamma}$$

for all $v_h \in V_h$.

**Proof.** Let $V_h \in V_h$ and $B_h$ be the strip between $\Gamma$ and $\Gamma_h$, i.e. $B_h = \Omega \setminus \Omega_h$ and $\partial B_h = \Gamma \cup \Gamma_h$. Recall that we assume the normal $n$ to look outward from $\Omega_h$ (resp. $\Omega$) on $\Gamma_h$ (resp. $\Gamma$). The outward looking normal on $\partial B_h$ coincides with $n$ on $\Gamma_h$ and with $-n$ on $\Gamma$. We have thus

$$\int_{\Gamma_h} \frac{\partial v_h}{\partial n} v_h - \int_{\Gamma} \frac{\partial v_h}{\partial n} v_h = \int_{\partial B_h} \sum_{T \in T^*} \frac{\partial v_h}{\partial n} v_h$$

(19)

(\text{using } \partial T = (\partial B_h \cap T) \cup (B_h \cap \partial T) \text{ and } \Delta v_h = 0 \text{ on } T)

Substituting this into the definition (10) of $a_h$ yields

$$a_h(v_h, v_h) = \int_{\Omega_h} |\nabla v_h|^2 - \int_{\Omega} |\nabla v_h|^2 + \sum_{T \in T^*} \int_{\partial B_h} \left[ \frac{\partial v_h}{\partial n} \right] + \frac{\gamma}{h} \int_{\Gamma} \left[ \frac{\partial \Pi v_h}{\partial n} \right]^2$$

Noting that $B_h \subset \Omega_h^T$ we can use (18) combined with the Young inequality (for any $\varepsilon > 0$) and (16) to write

$$a(v_h, v_h) \geq (1 - \varepsilon) |v_h|^2_{1, \Omega} + \left(\alpha - \beta - \frac{1}{2\varepsilon}\right) h \sum_{T \in T^*} \left\| \frac{\partial v_h}{\partial n} \right\|^2_{0, E} - \frac{\varepsilon}{2h} \sum_{T \in T^*} |v_h|^2_{0, E} + \frac{\gamma}{h} |\nabla \Pi v_h|^2_{\Gamma}$$

Taking $\varepsilon$ sufficiently small and $\sigma$ sufficiently big this bounds $a(v_h, v_h)$ from below by $c\|v_h\|_{h}^2$ as claimed.

### 3.3. Dirichlet boundary conditions: the error estimates.

We can now establish an optimal $H^1$ error estimate for the method (9)–(10) using the coerciveness of $a_h$ provided by Lemma 3.6. An $L^2$ error estimate will follow in Theorem 3.8.
Theorem 3.7. Suppose \( f \in H^1(\Omega_h), g \in H^{1/2}(\Gamma) \) and let \( u \in H^3(\Omega) \) be the solution to (1), \( u_h \in V_h \) be the solution to (9)–(10). Provided \( \sigma \) is sufficiently big, there exists an \( h \)-independent constant \( C > 0 \) such that

\[
|u - u_h|_{1,\Omega} + \frac{1}{\sqrt{h}} |u - u_h|_{0,\Gamma} \leq Ch(|f|_{1,\Omega} + ||g||_{L^2(\Gamma)}).
\]

Proof. Under the Theorem’s assumptions, the solution to (1) is indeed in \( H^3(\Omega) \) and it can be extended to a function \( \tilde{u} \in H^3(\Omega_h) \) such that \( \tilde{u} = u \) on \( \Omega \) and \( ||\tilde{u}||_{3,\Omega} \leq C(||f||_{1,\Omega} + ||g||_{L^2(\Gamma)}), \) cf. (1). Clearly, \( \tilde{u} \) satisfies

\[
a_h(\tilde{u}, v_h) = (\tilde{f}, v_h)_{L^2(\Omega_h)} + \int_{\Gamma} g \frac{\partial v_h}{\partial n} + \frac{\gamma}{h} \int_{\Gamma} v_h \quad \forall v_h \in V_h
\]

with \( \tilde{f} := -\Delta \tilde{u} \). It entails a Galerkin orthogonality relation

\[
a_h(\tilde{u} - u_h, v_h) = \int_{\Omega_h} (\tilde{f} - f)v_h, \quad \forall v_h \in V_h
\]

We have then using the standard nodal interpolation \( I_h : C(\bar{\Omega}_h) \to V_h \) and recalling Lemma 3.6

\[
\frac{1}{c} \|a_h - I_h \tilde{u}\| \leq \sup_{v_h \in V_h} \frac{a_h(\tilde{u} - u_h, v_h)}{\|v_h\|_h} = \sup_{v_h \in V_h} \frac{a_h(e, v_h) + (f - \tilde{f}, v_h)_{L^2(\Omega_h)}}{\|v_h\|_h}
\]

with \( e = \tilde{u} - I_h \tilde{u} \). Using the definition (10) of \( a_h \), we can now bound term by term

\[
a_h(e, v_h) \leq |e|_{1,\Omega} |v|_{1,\Omega} + \left| \frac{\partial e}{\partial n} \right|_{0,\Gamma} \left| v \right|_{0,\Gamma} + |e|_{1,\Omega} \left| \frac{\partial v}{\partial n} \right|_{0,\Gamma} + \frac{\gamma}{h} |e|_{0,\Gamma} \left| v \right|_{0,\Gamma}
\]

\[
+ \frac{\gamma}{h} \sum_{e \in \Gamma_h} \left| \frac{\partial e}{\partial n} \right|_{0,\Gamma} \left| v \right|_{0,\Gamma}^2
\]

\[
\leq C \left( |e|_{1,\Omega}^2 + \frac{h}{\sqrt{h}} \left| \frac{\partial e}{\partial n} \right|_{0,\Gamma}^2 + \frac{1}{h} \left| e \right|_{0,\Gamma}^2 + h \sum_{e \in \Gamma_h} \left| \frac{\partial e}{\partial n} \right|_{0,\Gamma}^2 \right)^{1/2} \left| v \right|_{0,\Gamma}
\]

We have used here Lemma 3.4 to bound \( h \sum_{e \in \Gamma_h} \left| \frac{\partial e}{\partial n} \right|_{0,\Gamma}^2 \) and \( |v|_{0,\Gamma} \). This entails thanks to the usual interpolation estimates

\[
a_h(e, v_h) \leq Ch ||e||_{2,\Omega_h} ||v||_h
\]

Moreover,

\[
(f - \tilde{f}, v_h)_{L^2(\Omega_h)} \leq ||f - \tilde{f}||_{0,\Omega} ||v||_{0,\Omega} \leq C ||f - \tilde{f}||_{0,\Omega_h} ||v||_{0,\Omega_h}
\]

thanks to Corollary 3.2. We conclude

\[
||u_h - I_h \tilde{u}||_h \leq C(h ||\tilde{u}||_{2,\Omega_h} + ||f - \tilde{f}||_{0,\Omega_h})
\]

We recall now that \( f = \tilde{f} \) on \( \Omega \) so that, thanks to Lemma 3.1

\[
||f - \tilde{f}||_{0,\Omega} = ||f - \tilde{f}||_{0,\Omega_h} \leq Ch ||f - \tilde{f}||_{0,\Omega_h} \leq Ch ||f||_{1,\Omega_h} + ||g||_{L^2(\Gamma)}.
\]

Combining the estimates above with the triangle inequality proves \( ||u_h - \tilde{u}||_h \leq Ch ||f||_{1,\Omega_h} + ||g||_{L^2(\Gamma)} \), as claimed. \( \square \)

Remark 3.1. The proof above does not rely on a solution to the non-standard boundary value problem (1) on \( \Omega_h \). We rather use the well defined solution \( u \) to problem (1) and extend it to \( \Omega_h \). The mismatch between \( \Omega_h \) and \( \Omega \) is then handled at the expense of a stronger than usual assumption on the right-hand side in (1): we need \( u \in H^3, f \in H^1 \) where as \( u \in H^2, f \in L^2 \) suffices for standard \( P_1 \) finite elements on a conforming mesh.
Lemma 3.9. For any $u_0$, provided $u_0$ reveals the optimal convergence rate $O(h^\alpha)$, similar to the state of the art in the study of the non-symmetric Nitsche method.

Theorem 3.8. Under the assumptions of Theorem 3.7, there exists an $h$-independent constant $C > 0$ such that

$$
\|u - u_0\|_{\Omega, f} \leq C h^{\alpha/3} (\|f\|_{L^2(\Omega)} + \|u_0\|_{L^2(\Omega)}).
$$

Proof. Let us introduce $w : \Omega \to \mathbb{R}$ such that

$$
-\Delta w = u - u_h \text{ in } \Omega, \quad w = 0 \text{ on } \Gamma.
$$

By elliptic regularity, $\|w\|_{L^2(\Omega)} \leq C \|u - u_h\|_{L^2(\Omega)}$. Let $\tilde{w}$ be an extension of $w$ from $\Omega$ to $\Omega_h$ preserving the $H^2$ norm estimate. Applying Lemma 3.3 to $\tilde{w}$ and to $\nabla \tilde{w}$ yields

$$
\|\tilde{w}\|_{L^2(\Omega)} \leq C h \|u - u_h\|_{L^2(\Omega)} \quad \text{and} \quad |\tilde{w}|_{L^2(\Omega)} \leq C \sqrt{h} |u - u_h|_{L^2(\Omega)}
$$

Similarly, applying Lemma 3.3 to $\tilde{w}$ yields

$$
\|\tilde{w}\|_{L^2(\Omega, f)} \leq C h |u - u_h|_{L^2(\Omega)}
$$

Taking $w_h = I_h \tilde{w}$, we can summarize the bounds above together with the interpolation estimates as

$$
\|w_w\|_{L^2(\Omega)} + \sqrt{h} \left\|\frac{\partial (w_w - w_h)}{\partial n}\right\|_{L^2(\Gamma_h)} + \frac{1}{\sqrt{h}} \left|\tilde{w} - w_h\right|_{L^2(\Omega, f)} + \sqrt{h} |w_h|_{L^2(\Gamma_h)} + \|w_h\|_{L^2(\Omega)} \leq C h |u - u_h|_{L^2(\Omega)}
$$

Using Galerkin orthogonality (21) and estimates (25) we arrive at (recall $B_h = \Omega \setminus \Omega_h$)

$$
\|u - u_h\|_{L^2(\Omega)} = \int_\Omega \nabla (u - u_h) : \nabla w - \int_\Gamma (u - u_h) \frac{\partial w}{\partial n} = a_h(\bar{u} - u_h, \tilde{w} - w_h) + \int_{\Gamma_h} \frac{\partial (\tilde{u} - u_h)}{\partial n} \tilde{w} - 2 \int_{\Gamma_h} (u - u_h) \frac{\partial \tilde{w}}{\partial n} - \int_{\Gamma_h} \nabla (\tilde{u} - u_h) : \nabla \tilde{w} + \int_{\Gamma_h} (\tilde{f} - f) w_h
$$

which gives the announced error estimate in $L^2(\Omega)$ norm thanks to the error estimates in the triple norm $\|\cdot\|$ and the bound on $\|\tilde{f} - f\|_{L^2(\Gamma_h)}$ already established in the proof of Theorem 3.7.

3.4. Robin boundary conditions: coerciveness of $a_h^R$.

We turn to the study of method (12)–(13) for the problem with Robin boundary conditions (assuming $\rho > 0$, cf. Remark 2.1). Our first goal is to prove the coerciveness of the bilinear form $a_h^R$ uniformly in $h$. To this end, we note that $a_h^R$ can be rewritten using the divergence Theorem as

$$
a_h^R(u, v) = \int_{\Omega} \nabla u : \nabla v + \int_{\Omega} (v \partial y + y \cdot \nabla v) + \int_{\Omega} \rho u v + \gamma_d \int_{\Omega} \partial y \partial z \partial y \partial z
$$

provided $y$ and $v$ are of regularity $H^1$. We have denoted here again $B_h = \Omega \setminus \Omega_h$. The following lemma will allow us to control $\int_{\Gamma_h} \partial y \cdot \nabla v$ while the other term $\int_{\Gamma_h} \partial y \nabla v$ can be controlled thanks to the grad-div stabilization.

Lemma 3.9. For any $\beta > 0$, there exist $0 < \alpha < 1$ and $\delta > 0$ depending only on the mesh regularity and on parameter $M$ in Assumption 2 such that

$$
\left\| \int_{\Gamma_h} \partial y \cdot \nabla v_h \right\| \leq \alpha \|v_h\|_{L^2(\Omega_h)} + \delta \|z_h + \nabla v_h\|_{L^2(\Gamma_h)}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{L^2(\Gamma_h)}^2
$$

for all $v_h \in \mathcal{V}_h, z_h \in \mathcal{Z}_h$.
\textbf{Proof.} The boundary $\Gamma$ can be covered by element patches $[\Pi_k]_{k=1,...,N_k}$ as in Assumption 2. Choose any $\beta > 0$ and consider

$$\alpha := \max_{I_k \ni \Omega_h \neq \emptyset} \frac{\|z_h\|_{0,I_k}^2 |v_h|_{1,I_k}^2 - \beta \|z_h + \nabla v_h\|_{0,I_k}^2 - \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2}{\frac{1}{2} \|z_h\|_{0,I_k}^2 + \frac{1}{2} |v_h|_{1,I_k}^2}$$ \hspace{1cm} (28)

where the maximum is taken over all the possible configurations of a patch $\Pi_k$ allowed by the mesh regularity and over all the piecewise linear functions $v_h$ and $z_h$ on $\Pi_k$. Note that the quantity under the max sign in (28) is invariant under the scaling transformation $x \mapsto hx, z_h \mapsto \frac{1}{h}z_h, v_h \mapsto v_h$, and is homogeneous with respect to $v_h, z_h$. Thus, the maximum is indeed attained since it is taken over a bounded set in a finite dimensional space.

Clearly, $\alpha \leq 1$. Supposing $\alpha = 1$ would lead to a contradiction. Indeed, if $\alpha = 1$, we can then take $\Pi_k, v_h, z_h$ yielding this maximum (in particular, $|v_h|_{1,I_k}^2 + \|z_h\|_{0,I_k}^2 > 0$). We observe then

$$\frac{1}{2} |v_h|_{1,I_k}^2 + \|z_h\|_{0,I_k}^2 |v_h|_{1,I_k}^2 + \frac{1}{2} \|z_h\|_{0,I_k}^2 + \beta |z_h + \nabla v_h\|_{0,I_k}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2 = 0$$

and consequently (recall $|v_h|_{1,I_k}^2 = |v_h|_{1,T_k}^2 + |v_h|_{1,I_k}^2$)

$$\frac{1}{2} |v_h|_{1,T_k}^2 + \beta |z_h + \nabla v_h\|_{0,I_k}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2 = 0$$ \hspace{1cm} (29)

This implies $\|z_h + \nabla v_h\|_{0,I_k}^2 = 0$, i.e. $\nabla v_h = -z_h$ on $\Pi_k$. Since $\nabla v_h$ is piecewise constant and $z_h$ is continuous, it means that $\nabla v_h = -z_h =$ const on $\Pi_k$. We also have $\left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2 = 0$, which implies $\nabla v_h =$ const on the whole $\Pi_k$. Since, by (29), $\nabla v_h = 0$ on $T_k$, we have finally $\nabla v_h = 0$ on $\Pi_k$ and $z_h = 0$ on $\Pi_k$, which is in contradiction with $|v_h|_{1,I_k}^2 + \|z_h\|_{0,I_k}^2 > 0$.

Thus $\alpha < 1$ and

$$\|z_h\|_{0,I_k}^2 |v_h|_{1,I_k}^2 \leq \frac{\alpha}{2} \|z_h\|_{0,I_k}^2 + \frac{\alpha}{2} |v_h|_{1,I_k}^2 + \beta |z_h + \nabla v_h\|_{0,I_k}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2$$

for all $v_h, z_h$ and all admissible patches $\Pi_k$. We now observe

$$\left| \int_{\partial I_k} z_h \cdot \nabla v_h \right| \leq \sum_k \left| \int_{\partial \Omega_h \cap \partial I_k} z_h \cdot \nabla v_h \right| \leq \sum_k \|z_h\|_{0,I_k}^2 |v_h|_{1,I_k}^2$$

$$\leq \frac{\alpha}{2} \|z_h\|_{0,I_k}^2 + \frac{\alpha}{2} |v_h|_{1,I_k}^2 + \beta |z_h + \nabla v_h\|_{0,I_k}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2$$

$$\leq \alpha \left( 1 + \frac{e}{2} \right) |v_h|_{1,I_k}^2 + \beta \left( \frac{\alpha}{2} + \frac{\alpha}{2e} \right) \|z_h + \nabla v_h\|_{0,I_k}^2 + \beta h \left\| \frac{\partial v_h}{\partial n} \right\|_{0,\partial \Omega_h \cap \partial I_k}^2$$

for any $\varepsilon > 0$. We have used here the following estimate valid by Young inequality

$$\|z_h\|_{0,I_k}^2 \leq \|z_h + \nabla v_h\|_{0,I_k}^2 + \|\nabla v_h\|_{0,I_k}^2 - 2(z_h + \nabla v_h, \nabla v_h)_{0,I_k}$$

$$\leq \left( 1 + \frac{1}{\varepsilon} \right) \|z_h + \nabla v_h\|_{0,I_k}^2 + (1 + \varepsilon) |v_h|_{1,I_k}^2$$

Taking $\varepsilon$ sufficiently small, redefining $\alpha$ as $\alpha \left( 1 + \frac{e}{2} \right)$ and putting $\delta = \left( \beta + \frac{\alpha}{2} + \frac{\alpha}{2e} \right)$ we obtain [27]. \hfill \square
Lemma 3.10. Provided \( \rho > 0 \) and \( \gamma_{\text{div}}, \gamma_1 \) are sufficiently big, there exists an \( h \)-independent constant \( c > 0 \) such that for all \( v_h \in V_h, z_h \in Z_h \),

\[
da_h^R(v_h, z_h; v_h, z_h) \geq c \|v_h, z_h\|_h^2
\]

with

\[
\|v, z\|_h^2 = \|v\|_{1, \Omega_h}^2 + \rho \|v\|_{0, \Gamma_h}^2 + \|\text{div } z\|_{0, \Gamma_h}^2 + \|z + \nabla v\|_{0, \Gamma_h}^2 + h \left\| \frac{\partial v}{\partial n} \right\|_{0, \Gamma_h}^2
\]

Proof. Expression (26) for \( a_h^R \) implies for all \( v_h \in V_h, z_h \in Z_h \)

\[
da_h^R(v_h, z_h; v_h, z_h) = |v_h|_{1, \Omega_h}^2 + \rho |v|_{0, \Gamma_h}^2 + \int_{\Omega_h} (v_h \text{div } z_h + z_h \cdot \nabla v_h)
+ \gamma_{\text{div}} \|\text{div } z_h\|_{0, \Gamma_h}^2 + \gamma_1 |z_h + \nabla v_h|_{0, \Gamma_h}^2 + \alpha h \left\| \frac{\partial v_h}{\partial n} \right\|_{0, \Gamma_h}^2
\]

A combination of (27) with the Young inequality (for any \( \varepsilon > 0 \)) yields

\[
da_h^R(v_h, z_h; v_h, z_h) \geq (1 - \alpha) |v_h|_{1, \Omega_h}^2 + \rho |v|_{0, \Gamma_h}^2 - \frac{\varepsilon}{2} |v|_{0, \Gamma_h}^2
+ \left( \gamma_{\text{div}} - \frac{1}{2\varepsilon} \right) \|\text{div } z_h\|_{0, \Gamma_h}^2 + (\gamma_1 - \delta) |z_h + \nabla v_h|_{0, \Gamma_h}^2 + (\sigma - \beta) h \left\| \frac{\partial v_h}{\partial n} \right\|_{0, \Gamma_h}^2
\]

Recall a Poincaré-type inequality (valid thanks to Petree-Tartar lemma)

\[
|v_h|_{1, \Omega_h}^2 \leq C_h^2 (|v_h|_{0, \Omega_h}^2 + |v|_{0, \Gamma_h}^2)
\]

(30)

This entails

\[
da_h^R(v_h, z_h; v_h, z_h) \geq (1 - \alpha - \frac{\varepsilon}{2} C_h^2) |v_h|_{1, \Omega_h}^2 + \left( \rho - \frac{\varepsilon}{2} C_h^2 \right) |v|_{0, \Gamma_h}^2 + \left( \gamma_{\text{div}} - \frac{1}{2\varepsilon} \right) \|\text{div } z_h\|_{0, \Gamma_h}^2
+ (\gamma_1 - \delta) |z_h + \nabla v_h|_{0, \Gamma_h}^2 + (\sigma - \beta) h \left\| \frac{\partial v_h}{\partial n} \right\|_{0, \Gamma_h}^2
\]

Taking \( \varepsilon, \beta \) sufficiently small and \( \gamma_1, \gamma_{\text{div}} \) sufficiently big this bounds \( a_h^R(v_h, z_h; v_h, z_h) \) from below by \( c \|v_h, z_h\|_h^2 \) as claimed.

\( \square \)

Remark 3.2. The coercivity constant \( c \) in the lemma above vanishes in the limit \( \rho \to 0 \). This lemma is thus not directly suitable for the study of Neumann boundary conditions, \( \rho = 0 \). We can easily prove though that the coercivity constant of \( a_h^R \) is robust with respect to \( \rho \) on the subspace \( V_h^0 \times Z_h \) with

\[
V_h^0 = \{ v_h \in V_h : \int_{\Omega} v_h = 0 \}
\]

(31)

Indeed, (30) can be replaced by \( |v_h|_{0, \Omega_h} \leq C_p |v_h|_{1, \Omega_h} \) for \( v_h \in V_h^0 \) so that the term \( \rho |v_h|_{0, \Gamma_h}^2 \) is no longer perturbed in the final estimate of the proof.

3.5. Robin boundary conditions: error estimates.

We can now establish an optimal \( H^1 \) error estimate for the method (12)-(13) using the coerciveness of \( a_h^R \) provided by Lemma 3.10. An \( L^2 \) error estimate will follow in Theorem 3.12.
Theorem 3.11. Suppose \( f \in H^1(\Omega_0) \), \( g \in H^{3/2}(\Gamma) \) and let \( u \in H^3(\Omega) \) be the solution to (2), \((u_h, y_h) \in V_h \times Z_h\) be the solution to (12). Provided \( \rho > 0 \) and \( \gamma_{\text{div}}, \gamma_1 \) are sufficiently big, there exists an \( h \)-independent constant \( C > 0 \) such that
\[
|u - u_h|_{1, \Omega} + \sqrt{h} |u - u_h|_{0, \Gamma} \leq C h(\|f\|_{1, \Omega_0} + \|g\|_{3/2, \Gamma}).
\] (32)

Proof. Under the Theorem’s assumptions, the solution to (2) is indeed in \( H^3(\Omega) \) and it can be extended to a function \( \bar{u} \in H^3(\Omega_0) \) such that \( \bar{u} = u \) on \( \Omega \) and \( \|\bar{u}\|_{3, \Omega_0} \leq C(\|f\|_{1, \Omega_0} + \|g\|_{3/2, \Gamma}) \). Introduce \( y = -\nabla \bar{u} \) on \( \Omega_0^\Gamma \). Clearly, \( \bar{u}, y \) satisfy
\[
a_h^R(\bar{u}; \gamma; v_h, z_h) = \int_{\Omega_h} f v_h + \int_{\Gamma} g v_h + \gamma_{\text{div}} \int_{\Omega_h} \bar{f} \text{div } z_h \quad \forall (v_h, z_h) \in V_h \times Z_h
\]
with \( \bar{f} := -\Delta \bar{u} \). It entails a Galerkin orthogonality relation
\[
a_h^R(\bar{u} - u_h, y - y_h; v_h, z_h) = \int_{\Omega_h} (\bar{f} - f) v_h + \gamma_{\text{div}} \int_{\Omega_h} (\bar{f} - f) \text{div } z_h, \quad \forall (v_h, z_h) \in V_h \times Z_h
\] (33)
We have then using the standard nodal interpolation \( I_h : C(\Omega_0) \to V_h \) and recalling Lemma 3.10
\[
\frac{1}{c} \|u_h - I_h \bar{u}, y_h - I_h y\|_{0, \Gamma} \leq \sup_{(v_h, z_h) \in V_h \times Z_h} \frac{a_h^R(u_h - I_h \bar{u}, y_h - I_h y; v_h, z_h)}{\|v_h, z_h\|_h}
\]
with \( e_u = \bar{u} - I_h \bar{u}, e_y = y - I_h y \). Recalling (26), we can bound
\[
a_h^R(e_u, e_y; v_h, z_h) \leq |e_u|_{1, \Omega_0} \|v_h|_{1, \Omega_0} + \|\text{div } e_y\|_{0, \Gamma} \|v_h\|_{0, \Gamma} + \|e_y\|_{0, \Gamma} \|v_h\|_{0, \Gamma} + \rho \|e_u\|_{0, \Gamma} \|v_h\|_{0, \Gamma} + \gamma_{\text{div}} \|\text{div } z_h\|_{0, \Gamma} + \gamma_1 |e_y| + \Delta e_u \|v_h\|_{0, \Gamma} + \sigma h \left( \frac{\partial e_u}{\partial n} \right|_{\Gamma_0^\gamma} \right) \frac{\partial v_h}{\partial n} \|v_h, z_h\|_h
\]
By the usual interpolation estimates this entails
\[
a_h^R(e_u, e_y; v_h, z_h) \leq C h(\|\bar{u}\|_{2, \Omega_0} + |y|_{2, \Omega_0^\Gamma}) \|v_h, z_h\|_h \leq C h(\|\bar{u}\|_{3, \Omega_0} \|v_h, z_h\|_h
\]
since \( y = -\nabla \bar{u} \) on \( \Omega_0^\Gamma \). We now conclude as in the proof of Theorem 3.7. 

Remark 3.3. The error estimate of the preceding theorem remains valid in the case \( \rho = 0 \) (Neumann boundary conditions) if one replaces \( V_h \) in (12) by its subspace \( V_h^\rho \) given by (37). This follows from a modification of Lemma 3.10 suggested in Remark 3.2. The \( L^2 \) error estimate of the following theorem remains also true assuming that the exact solution satisfies \( \int_{\Gamma} u = 0 \).

Theorem 3.12. Under the assumptions of Theorem 3.11 there exists an \( h \)-independent constant \( C > 0 \) such that
\[
|u - u_h|_{1, \Omega} \leq C h^{3/2}(\|f\|_{1, \Omega_0} + \|g\|_{3/2, \Gamma}).
\] (34)

Proof. Let us introduce \( w : \Omega \to \mathbb{R} \) such that
\[
-\Delta w = u - u_h \text{ in } \Omega, \quad \frac{\partial w}{\partial n} + \rho w = 0 \text{ on } \Gamma.
\]
By elliptic regularity, \( \|w\|_{2,\Omega} \leq C\|u-u_h\|_{0,\Omega} \). Let \( \tilde{w} \) be an extension of \( w \) from \( \Omega \) to \( \Omega_h \) preserving the \( H^2 \)-norm estimate and set \( w_h = I_h \tilde{w} \). Integration by parts and interpolation estimates yield
\[
\|u-u_h\|^2_{0,\Omega} = \int_{\Omega} \nabla(u-u_h) \cdot \nabla(w-w_h) + \rho \int_{\Gamma}(u-u_h)(w-w_h) \\
+ \int_{\Omega} \nabla(u-u_h) \cdot \nabla w_h + \rho \int_{\Gamma}(u-u_h)w_h \\
\leq Ch \|u-u_h\|^2_{0,\Omega} + \rho \|u-u_h\|^2_{0,\Gamma}^{1/2} \|w_h\|_{2,\Omega_h} + \int_{\Omega} \nabla(u-u_h) \cdot \nabla w_h + \rho \int_{\Gamma}(u-u_h)w_h
\]
We now rewrite the bilinear form (36) as
\[
a_h^*(u, v; z) = \int_{\Omega} \nabla u \cdot \nabla v + \rho \int_{\Gamma} uv + \int_{\Omega} (v \div y + (y + \nabla u) \cdot \nabla v) + \gamma \div \int_{\Gamma} \frac{\partial u}{\partial n} \cdot \frac{\partial v}{\partial n}
\]
so that the Galerkin orthogonality relation (33) with \( v_h = w_h, z_h = 0 \) becomes
\[
\int_{\Omega} \nabla(\tilde{u} - u_h) \cdot \nabla w_h + \rho \int_{\Gamma}(\tilde{u} - u_h)w_h \\
+ \gamma \int_{\Omega} (y + \nabla \tilde{u} - y - \nabla u_h) \cdot \nabla w_h + \rho \int_{\Gamma} \frac{\partial(\tilde{u} - u_h)}{\partial n} \cdot \frac{\partial w_h}{\partial n}
\]
\[
= \int_{\Omega} (\tilde{f} - f)w_h
\]
Recalling the definition of the triple norm from Lemma 3.10 this leads to
\[
\left| \int_{\Omega} \nabla(\tilde{u} - u_h) \cdot \nabla w_h + \rho \int_{\Gamma}(\tilde{u} - u_h)w_h \right| \leq C\|\tilde{u} - I_h \tilde{u}, y - I_h y\|_{0,\Gamma} \left( \|w_h\|_{1,\Omega_h} + \sqrt{h} \left\| \frac{\partial w_h}{\partial n} \right\|_{0,\Gamma} \right)
\]
\[
+ \|\tilde{f} - f\|_{0,\Omega_h} \|w_h\|_{0,\Omega_h}
\]
By Lemma 3.1 and interpolation estimates
\[
\|w_h\|_{0,\Omega_h} \leq \|\tilde{w} - I_h \tilde{w}\|_{0,\Omega_h} + \|\tilde{w}\|_{0,\Omega_h} \leq Ch \|\tilde{w}\|_{2,\Omega} + C \left( \sqrt{h}\|\tilde{w}\|_{0,\Gamma} + h\|\tilde{w}\|_{1,\Omega_h} \right) \leq C \sqrt{h}\|\tilde{w}\|_{2,\Omega_h}
\]
and similarly
\[
\|\nabla w_h\|_{0,\Omega_h} \leq \|\nabla(\tilde{w} - I_h \tilde{w})\|_{0,\Omega_h} + \|\nabla \tilde{w}\|_{0,\Omega_h} \leq Ch \|\tilde{w}\|_{2,\Omega} + C \left( \sqrt{h}\|\nabla \tilde{w}\|_{0,\Gamma} + h\|\nabla \tilde{w}\|_{1,\Omega_h} \right) \leq C \sqrt{h}\|\tilde{w}\|_{2,\Omega_h}
\]
Combining all the estimates above we arrive at
\[
\|u-u_h\|^2_{0,\Omega} \leq C \sqrt{h}\|\tilde{u} - I_h \tilde{u}, y - I_h y\|_{0,\Gamma} + \|\tilde{f} - f\|_{0,\Omega_h} \|\tilde{w}\|_{2,\Omega_h}
\]
We conclude recalling the bound on the triple norm of the error from the proof of Theorem 3.11 the estimate \( \|f - f\|_{0,\Omega_h} \leq Ch(\|f\|_{1,\Omega} + \|\tilde{w}\|_{1,\Omega_h}) \) from the proof of Theorem 3.7 and the regularity estimate \( \|w_h\|_{2,\Omega_h} \leq C\|u-u_h\|_{0,\Omega} \) .

4. Extension to \( P_k \) finite elements

The methods presented above can be easily extended to \( P_k \) finite elements giving optimal convergence of order \( h^k \) in the \( H^1 \)-norm, \( k \geq 2 \). We thus consider in this section the finite element space
\[
V_h^{(k)} = \{ v_h \in H^1(\Omega_h) : v_h|_T \in P_k(T), \ \forall T \in \mathcal{T}_h \}
\]
with \( P_k(T) \) representing the polynomials of degree \( \leq k \), propose some modifications to be introduced to the methods above, first for Dirichlet boundary conditions and then for Neumann-Robin ones, and outline the convergence proofs in both these cases.
4.1. Dirichlet boundary conditions

The method [7–10] is modified as follows:

Find $u_h \in V_h^{(k)}$ s.t.

$$a_h(u_h, v_h) = \int_{\Omega_h} f v_h + \int_T g \frac{\partial v_h}{\partial n} + \mu \int_T g v_h - \sigma h^2 \sum_{T \subset T_h} \int_T f \Delta v_h, \quad \forall v_h \in V_h^{(k)}$$

with

$$a_h(u, v) = \int_{\Omega_h} \nabla u \cdot \nabla v - \int_{\Gamma_h} \frac{\partial u}{\partial n} v + \int_T u \frac{\partial v}{\partial n} + \mu \int_T u v + \sigma h^2 \sum_{T \subset T_h} \int_T (\Delta u)(\Delta v) + \sigma \sum_{E \in \mathcal{F}_h} \sum_{j=1}^k h^{2j-1} \int_E \left[ \frac{\partial u}{\partial n_j} \right] \left[ \frac{\partial v}{\partial n_j} \right]$$

Note that the additional stabilization term with the product $(\Delta u)(\Delta v)$ is strongly consistent since $-\Delta u = f$. Also note that the ghost penalty term is extended to control the normal derivatives of all orders up to $k$, cf. [6].

Revisiting the theoretical analysis of Section 3 reveals the following:

- Lemma 3.5 remains valid thanks to the extended ghost penalty. Indeed, the only thing to recheck in its proof is the following implication: if $\nabla v_h = 0$ on $T$ and all the norms of the jumps contained in ghost penalty vanish, then $\nabla v_h = 0$ on $T$. This is true since the extended ghost penalty controls all the derivatives present in our finite element space.

- The first relation (19) in the proof of Lemma 3.6 now becomes

$$\int_{\Gamma_h} \frac{\partial v_h}{\partial n} - \int_T \frac{\partial v_h}{\partial n} \cdot \cdots = \int_{\Gamma_h} |\nabla v_h|^2 + \sum_{T \subset T_h} \int_{T \cap \Gamma_h} (\Delta v_h)v_h - \sum_{E \in \mathcal{F}_h} \int_{E \cap \Gamma_h} v_h \left[ \frac{\partial v_h}{\partial n} \right]$$

since one can no longer assume $\nabla v_h = 0$ on $T$. The new term with $\Delta v_h$ can be then controlled thanks to the additional $\Delta \cdot \Delta$ stabilization term in the bilinear form $a_h$ so that Lemma 3.6 remains valid.

- In Theorem 3.7, we should now suppose $f \in H^k(\Omega_h)$, $g \in H^{k+3/2}(\Gamma)$ so that the solution $u$ to (1) is in $H^{k+2}(\Omega)$. The proof of the Theorem can be then followed using appropriate Sobolev spaces and standard interpolation estimates to $P_k$ finite elements. The only non-trivial change in the proof concerns the estimate (22), which should be changed to

$$\|f - \bar{f}\|_{0, \Omega_h} \leq C h^k |f|_{k, \Omega_h} \leq C h^k (|f|_{k, \Omega_h} + \|g\|_{k+2, \Omega_h}). \quad (36)$$

This can be proved by refining the argument of Lemma 3.1. We recall that $f = \bar{f}$ on $\Omega$ and both $f$ and $\bar{f}$ are in $H^k(\Omega)$ so that all the derivatives up to order $k - 1$ of $f$ and $\bar{f}$ coincide on $\Gamma$. The bound (36) then follows as in Lemma 3.1 employing a Taylor expansion with the integral remainder of order $k$.

Keeping in mind the modifications above, we can easily establish the convergence of the $P_k$ version of our method: given $f \in H^k(\Omega_h)$ and $g \in H^{k+3/2}(\Gamma)$, and supposing $\sigma$ sufficiently big, one has

$$|u - u_h|_{1, \Omega} + \frac{1}{\sqrt{h}} \|u - u_h\|_{0, \Gamma} \leq C h^k (|f|_{k, \Omega_h} + \|g\|_{k+3/2, \Gamma}).$$

The $L^2$ error estimate of order $h^{k+1/2}$ can be also proved as in Theorem 3.8.
4.2. Neumann-Robin boundary conditions

We turn now to Problem (2). The goal is to extend the method (12)–(13) to $P_k$ finite elements. We thus recall the space $V_h^{(k)}$ as in (35) and introduce the auxiliary finite element space

$$Z_h^{(k)} = \{ z_h \in H^1(\Omega_h^{(k)}; z_h|_{\Gamma_h^\ast} \in P_k(\Gamma) ; \forall \Gamma \in \mathcal{T}_h^\ast \}$$

Our finite element problem is: Find $u_h \in V_h^{(k)}$, $y_h \in Z_h^{(k)}$ solving (12) with the bilinear form $a_h^k$ modified as follows

$$a_h^k(u_h; v_h; z_h) = \int_{\Omega_h} \nabla u_h \cdot \nabla v_h + \int_{\Gamma_h} y_h \cdot n v_h - \int_{\Gamma_h} y_h \cdot n v_h + \int_{\Gamma_h} \rho u v + \gamma_{\text{div}} \int_{\Omega_h} \text{div} y \text{div} z_h$$

$$+ \gamma_1 \int_{\Omega_h} (y_h + \nabla u) \cdot (z_h + \nabla v) + \sigma \sum_{j=1}^k h_\ast^{2j-1} \int_{\Gamma_h} \left| \frac{\partial^j u}{\partial n^j} \right| + \sigma \sum_{E \in \mathcal{T}_h^\ast} \sum_{j=1}^{k-1} h_\ast^{2j+1} \left| \frac{\partial^j u}{\partial n^j} \right|$$

We have added here extra penalization terms controlling the jumps of the higher normal derivatives of $u$ on $\Gamma_h^\ast$ and an additional ghost stabilization term controlling the jumps of derivatives of $y$ on the cut facets, denoted by $\mathcal{T}_h^\ast \setminus \mathcal{T}^\Gamma_h$. The theoretical analysis: $\Gamma$ is in fact approximated by a sequence of straight segments with the endpoints obtained by approximate intersections of $\Gamma$ with the edges of $\mathcal{T}_h$. This error is of order $h^2$ and thus can be assumed negligible in the case of $P_1$ finite elements (which is the only case studied numerically in this paper). We presume that a subtler approximation should be employed when dealing with higher order finite elements, as in [7].

5. Numerical experiments.

We shall illustrate our methods (9)–(10) and (12)–(13) by numerical experiments in a 2D domain $\Omega$ defined by a level-set function $\varphi$:

$$\Omega = \{ (x, y) \mid \varphi(x, y) < 0 \} \text{ with } \varphi := r^4(5 + 3 \sin(7 \theta + 7 \pi / 36)) / 2 - R^4$$

where $(r, \theta)$ are the polar coordinates $r = \sqrt{x^2 + y^2}$, $\theta = \arctan \frac{y}{x}$, and $R = 0.47$ (this example is taken from [15]). To construct the computational mesh, we embed $\Omega$ into the square $\hat{\Omega} = (-0.5, 0.5) \times (-0.5, 0.5)$, introduce a regular $N \times N$ criss-cross mesh on $\hat{\Omega}$, and drop the triangles outside $\Omega$ to produce $\mathcal{T}_h$ and $\mathcal{T}_h^\Gamma$, as illustrated at Fig. 1 in the case $N = 16$. In some of our experiments, the domain $\Omega$ will be rotated by an angle $\theta_0$ counter-clockwise around the origin. This is achieved by redefining $\varphi$ as $r = \arctan \frac{y}{x} - \theta_0$.

All the computations are done in FreeFem++ [16] taking advantage of its level-set capabilities (the key word levelset in numerical integration commands int1d and int2d to deal with the integrals over $\Gamma$ and over $\Omega$, cf. Subsection 5.2). Note that the numerical integration on $\Gamma$ introduces an additional error which is not covered by our theoretical analysis: $\Gamma$ is in fact approximated by a sequence of straight segments with the endpoints obtained by approximate intersections of $\Gamma$ with the edges of $\mathcal{T}_h$. This error is of order $h^2$ and thus can be assumed negligible in the case of $P_1$ finite elements (which is the only case studied numerically in this paper). We presume that a subtler approximation should be employed when dealing with higher order finite elements, as in [7].

5.1. Dirichlet boundary conditions

We start by solving numerically the Dirichlet problem (1) in domain (37) with zero right-hand side $f = 0$ and a non-homogeneous boundary condition $g$ set so that the exact solution is given by

$$u = \sin(x)e^y$$

We employ the method (9)–(10) taking the following parameter values: $\gamma = 1$, $\sigma = 0.01$. Fig. 3 represents the solution obtained on a $16 \times 16$ mesh. We observe that the numerical method captures well the exact solution, the perturbation being essentially concentrated in the narrow fictitious domain $B_h = \Omega_h \setminus \Omega$. Fig. 4 reports the evolution of the error.
Figure 3: Poisson-Dirichlet problem \( (1) \) with the exact solution \( (38) \) on domain \( (37) \). Top-Left: the numerical solution \( u_h \) on the \( 16 \times 16 \) mesh as in Fig. 1 produced by method \( (9) \)–\( (10) \). Top-Right: the exact solution \( u \). Bottom: the error \( u - u_h \).

Figure 4: Poisson-Dirichlet problem as in Fig. 3 the relative errors in \( L^2(\Omega) \) and \( H^1(\Omega) \) norms as functions of \( h \) under the mesh refinement.
Figure 5: Poisson-Dirichlet problem (1) on the domain as in Figs. 1 and 3 rotated by an angle $\theta_0$ around the origin: the relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of the rotation angle $\theta_0$.

Figure 6: Poisson-Dirichlet problem (1) as in Fig. 3 but on the $32 \times 32$ mesh. The relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of parameters $\gamma$ and $\sigma$. 
under the mesh refinement (always using the regular criss-cross meshes as mentioned above) and confirms the optimal convergence order of the method in both $H^1$ and $L^2$ norms.

In order to explore the robustness of the method with respect to the placement of the physical domain on the computational mesh, we now redo the calculations above, rotating $\Omega$ by a series of angles $\theta_0$ ranging from 0 to $\frac{\pi}{2}$ as described in the preamble of this Section. For each rotation angle $\theta_0$, the boundary $\Gamma$ cuts the triangles of the background mesh in a different manner, creating sometimes the “dangerous” situations when certain mesh triangles of $\mathcal{T}_h$ have only a tiny portion inside the physical domain $\Omega$. Fig. 5 presents the errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of $\theta_0$ at different discretization levels. We observe that the errors do not vary much from one position to another, especially when measured in the $H^1(\Omega)$ norm. Moreover, this variability decreases as the meshes are refined.

Finally, we explore the influence of the parameters $\gamma$ and $\sigma$ in (10) on the precision of the method. The error norms on a fixed $32 \times 32$ mesh for various choices of $\gamma$ and $\sigma$ are presented at Fig. 6. We observe that the method is not very sensitive to the parameters, especially when the error is measured in the $H^1$ norm. The accuracy does not seem to deteriorate catastrophically even in the limit $\gamma, \sigma \to 0$. This is somewhat surprising in view of the theoretical analysis of Section 3 and may indicate that a subtler theory could reveal the optimal convergence properties of the method (9)–(10) without stabilization. In our numerical experiments above, we have preferred however to remain on a safer side and have chosen a rather large value of the Nitsche parameter $\gamma = 1$ while keeping the ghost stabilization parameter $\sigma = 10^{-2}$ small. Note that larger values of $\gamma$ seem to make the method more sensible and unpredictable with respect to the choice of $\sigma$. On the other hand, the error is clearly monotonically increasing with $\sigma$ in the regime $\gamma \leq 1$.

5.2. Comparisons with CutFEM

As already mentioned in Introduction, our method (9)–(10) is very close to CutFEM methods, the essential difference being that we avoid the integration over $\Omega$, i.e. the numerical integration over the cut mesh elements. Although such an integration is in principle difficult to implement (in particular, it is more difficult than the integration over the segments of the boundary $\Gamma$), it is already available in FreeFEM++, so that we can easily compare numerically the performance of CutFEM with that of our method.

We have considered the following variants of CutFEM with $V_h$ denoting everywhere the $P_1$ finite elements space on the mesh $\mathcal{T}_h$, as in (8).

a) A version with Lagrange multipliers approximated by $P_0$ finite elements on the cut triangles $\mathcal{T}_h^\Gamma$ (4):

Find $u_h \in V_h$, $\lambda_h \in W_h := \{\mu_h \in L^2(\Omega_h^\Gamma) : \mu_h|_T \in P_0(T) \forall T \in \mathcal{T}_h^\Gamma\}$ s.t.

$$
\int_\Omega \nabla u_h \cdot \nabla v_h + \int_\Gamma \lambda_h v_h = \int_\Omega f v_h \quad \forall v_h \in V_h
$$

$$
\int_\Gamma \mu_h v_h - \sigma h \sum_{E \in \mathcal{E}^\Gamma} \int_E [\lambda_h]|_{E} = \int_\Gamma g v_h \quad \forall \mu_h \in W_h
$$

(39)

b) A version based on the symmetric Nitsche method (5):

Find $u_h \in V_h$ s.t.

$$
\int_\Omega \nabla u_h \cdot \nabla v_h - \int_\Gamma \frac{\partial u_h}{\partial n} v_h - \int_\Gamma u_h \frac{\partial v_h}{\partial n} + \gamma h \int_\Gamma u_h v_h + \sigma h \sum_{E \in \mathcal{E}^\Gamma} \int_E \left[ \frac{\partial u_h}{\partial n} \right] \left[ \frac{\partial v_h}{\partial n} \right] = \int_\Omega f v_h - \int_\Gamma g \frac{\partial v_h}{\partial n} + \gamma h \int_\Gamma g v_h \quad \forall v_h \in V_h
$$

(40)

c) A version based on the antisymmetric Nitsche method (similar to the above, only the sign in front of the third term changes) (2):

Find $u_h \in V_h$ s.t.

$$
\int_\Omega \nabla u_h \cdot \nabla v_h - \int_\Gamma \frac{\partial u_h}{\partial n} v_h - \int_\Gamma u_h \frac{\partial v_h}{\partial n} + \gamma h \int_\Gamma u_h v_h + \sigma h \sum_{E \in \mathcal{E}^\Gamma} \int_E \left[ \frac{\partial u_h}{\partial n} \right] \left[ \frac{\partial v_h}{\partial n} \right] = \int_\Omega f v_h + \int_\Gamma g \frac{\partial v_h}{\partial n} + \gamma h \int_\Gamma g v_h \quad \forall v_h \in V_h
$$

(41)
Figure 7: Poisson-Dirichlet problem with CutFEM. The relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of the rotation angle $\theta_0$, cf. Fig. 5.
The results are presented at Fig. 7. We consider there the same setup as in Fig. 5 i.e. domain $\Omega$ given by (37) rotated around the origin at a series of angles $\theta_0$, the exact solution given by (38). The stabilization parameters $\gamma$ and $\sigma$ are given in the captions (we have taken the same parameters for the antisymmetric version as for our method in Fig. 5, while a larger value for $\gamma$ was necessary for the symmetric version). Comparing the results in Figs. 5 and 7, we observe that all the methods have overall almost the same accuracy in the $H^1(\Omega)$ norm, but CutFEM produces better results (gaining by a factor around 2) when the error is measured in the $L^2(\Omega)$ norm. However, the performance of CutFEM (with the exception of the symmetric Nitsche variant) can drastically degrade at certain position of the domain $\Omega$ with respect to the mesh (the spikes of the error on the graphs as functions of the rotation angle $\theta_0$). This is certainly an implementation issue, which can be conceivably explained by inaccuracies in the numerical integration over the cut triangles and/or by an incorrect determination of such triangles due to the round-off errors. We do not attempt here to investigate this issue further, and merely note that the absence of integration over the cut triangles in our method (9)–(10) permits us to avoid some delicate implementation issues.

5.3. Neumann boundary conditions

We now turn to the Poisson-Neumann problem (2) with $\rho = 0$. Our test case is similar to that used before: the domain is given by (37) and the exact solution by (38). The right-hand side in (2) is thus $f = 0$ and the boundary condition $g$ is set up as $g = n \cdot \nabla u$ with the normal $n$ defined via the levelset function $\varphi$ in (37) as $n = \frac{\nabla \varphi}{|\nabla \varphi|}$. We employ the method (12)–(13) taking the following parameter values: $\gamma_{\text{div}} = 1$, $\gamma_1 = 10$, $\sigma = 0.01$. We enforce the uniqueness of the solution by imposing $\int_{\Omega} u_h = 0$ with the help of a Lagrange multiplier, thus increasing the size of the system matrix by 1. A convergence study under the mesh refinement is reported at Fig. 8. It confirms the optimal convergence order of the method in both $H^1$ and $L^2$ norms.

Figure 8: Poisson-Neumann problem on domain and meshes as in Figs. 1, 3: the relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of $h$ under the mesh refinement.

In order to explore the robustness of the method with respect to the placement of the physical domain on the computational mesh, we now redo the calculations above, rotating $\Omega$ by a series of angles $\theta_0$ ranging from 0 to $2\pi$, same as in the Dirichlet case. Fig. 9 presents the errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of $\theta_0$ at different discretization levels. Overall, the errors are of the same order as in the Dirichlet case, but they are now much more sensitive to the position of the domain with respect to the mesh. This is especially true for the $L^2$ error whose variability does not fade out when the meshes are refined. We postpone a more detailed study of this unfortunate phenomenon to future endeavors.

Finally, the influence of the stabilization parameters $\gamma_{\text{div}}$, $\gamma_1$ and $\sigma$ in (13) is presented at Fig. 10. The error norms on a fixed $32 \times 32$ mesh are given there for various choices of $\gamma_{\text{div}}$ and $\sigma$, with $\gamma_1$ set to 1 and 10. The behavior of the error with respect to all these 3 parameters is difficult to summarize, other than saying that the method is not too much sensitive to the parameter in a wide range. The accuracy does not seem to deteriorate catastrophically even in the limit $\gamma_{\text{div}}, \sigma \to 0$, at least with $\gamma_1 = 1$. This is somewhat surprising in view of the theoretical analysis of Section 5 and
Figure 9: Poisson-Neumann problem in the domain as in Figs. 1 and 2 rotated by an angle $\theta_0$ around the origin: the relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of the rotation angle $\theta_0$.

Figure 10: Poisson-Neumann problem. The relative errors in $L^2(\Omega)$ and $H^1(\Omega)$ norms as functions of parameters $\gamma_1$, $\sigma$ and $\gamma_1$ for two values of $\gamma_1$. 

(a) $\gamma_1 = 1$

(b) $\gamma_1 = 10$
may indicate that a subtler theory could reveal the optimal convergence properties of the method \cite{12--13} without grad-div and ghost stabilizations.

6. Conclusions

We have presented an optimally convergent method of the fictitious domain/XFEM/CutFEM type avoiding the numerical integration on cut mesh elements. The numerical experiments confirm the optimal convergence and the robustness of the method (with a possible slight deficiency on the level of $L^2$ errors in the case of Neumann boundary conditions). We have restricted ourselves to simple model problem (Poisson) in this first publication, but the methods should be applicable in more realistic settings as, for example, elasticity problem on a cracked domain (the method would then be similar to XFEM of \cite{20} but avoiding the integration on the mesh elements cut by the crack). Further research could be aimed, apart from extensions to more complex problems, at a finer understanding of the influence of stabilization parameters.

Acknowledgements

I am grateful to Grégoire Allaire for stimulating discussions and suggestions concerning the implementation. The implementation in FreFEM++ would not be possible without advices and help generously provided by Frédéric Hecht.

References

\begin{thebibliography}{99}
  \bibitem{1} Adams, R.A., Fournier, J.J.: Sobolev spaces, vol. 140. Elsevier (2003)
  \bibitem{2} Burman, E.: Ghost penalty. Comptes Rendus Mathematique \textbf{348}(21), 1217–1220 (2010)
  \bibitem{3} Burman, E., Claus, S., Hansbo, P., Larson, M.G., Massing, A.: Cutfem: discretizing geometry and partial differential equations. International Journal for Numerical Methods in Engineering \textbf{104}(7), 472–501 (2015)
  \bibitem{4} Burman, E., Hansbo, P.: Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method. Computer Methods in Applied Mechanics and Engineering \textbf{199}(41), 2680–2686 (2010)
  \bibitem{5} Burman, E., Hansbo, P.: Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method. Applied Numerical Mathematics \textbf{62}(4), 326–341 (2012)
  \bibitem{6} Burman, E., Hansbo, P.: Fictitious domain methods using cut elements: Iii. a stabilized nitsche method for stokes problem. ESAIM: Mathematical Modelling and Numerical Analysis \textbf{48}(3), 859–874 (2014)
  \bibitem{7} Burman, E., Hansbo, P., Larson, M.: A cut finite element method with boundary value correction. Mathematics of Computation \textbf{87}(310), 633–657 (2018)
  \bibitem{8} Fabrèges, B., Gouarin, L., Maury, B.: A smooth extension method. Comptes Rendus Mathematique \textbf{351}(9), 361–366 (2013)
  \bibitem{9} Franca, L.P., Hughes, T.J.: Two classes of mixed finite element methods. Computer Methods in Applied Mechanics and Engineering \textbf{69}(1), 89–129 (1988)
  \bibitem{10} Girault, V., Glowinski, R.: Error analysis of a fictitious domain method applied to a dirichlet problem. Japan Journal of Industrial and Applied Mathematics \textbf{12}(3), 487 (1995)
  \bibitem{11} Glowinski, R., Kuznetsov, Y.: Distributed lagrange multipliers based on fictitious domain method for second order elliptic problems. Computer Methods in Applied Mechanics and Engineering \textbf{196}(8), 1498–1506 (2007)
  \bibitem{12} Glowinski, R., Pan, T.W.: Error estimates for fictitious domain/penalty/finte element methods. Calcolo \textbf{29}(1), 125–141 (1992)
  \bibitem{13} Glowinski, R., Pan, T.W., Hesla, T.I., Joseph, D.D.: A distributed lagrange multiplier/fictitious domain method for particulate flows. International Journal of Multiphase Flow \textbf{25}(5), 755–794 (1999)
  \bibitem{14} Glowinski, R., Pan, T.W., Periaux, J.: A fictitious domain method for dirichlet problem and applications. Computer Methods in Applied Mechanics and Engineering \textbf{111}(3-4), 283–303 (1994)
  \bibitem{15} Haslinger, J., Renard, Y.: A new fictitious domain approach inspired by the extended finite element method. SIAM Journal on Numerical Analysis \textbf{47}(2), 1474–1499 (2009)
  \bibitem{16} Hecht, F.: New development in FreeFem++. J. Numer. Math. \textbf{20}(3-4), 251–265 (2012)
  \bibitem{17} Lozinski, A.: A new fictitious domain method: Optimal convergence without cut elements. Comptes Rendus Mathematique \textbf{354}(7), 741–746 (2016)
  \bibitem{18} Maury, B.: Numerical analysis of a finite element/volume penalty method. SIAM Journal on Numerical Analysis \textbf{47}(2), 1126–1148 (2009)
  \bibitem{19} Moës, N., Béchet, E., Tourbier, M.: Imposing dirichlet boundary conditions in the extended finite element method. International Journal for Numerical Methods in Engineering \textbf{67}(12), 1641–1669 (2006)
  \bibitem{20} Moës, N., Dolbow, J., Belytschko, T.: A finite element method for crack growth without remeshing. International journal for numerical methods in engineering \textbf{46}(1), 131–150 (1999)
  \bibitem{21} Nitsche, J.: 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 \textbf{36}(1), 9–15 (1971)
  \bibitem{22} Saal’ev, V.K.: On the solution of some boundary value problems on high performance computers by fictitious domain method. Siberian Math. J \textbf{4}(4), 912–925 (1963)
\end{thebibliography}
[23] Sukumar, N., Chopp, D.L., Moës, N., Belytschko, T.: Modeling holes and inclusions by level sets in the extended finite-element method. Computer methods in applied mechanics and engineering 190(46-47), 6183–6200 (2001)