A Thesis submitted for the degree of Doctor of Philosophy in the Faculty of Science

Stabilised finite element methods for fictitious domain problems

Cheherazada González Aguayo

23 January 2017

Department of Mathematics and Statistics
University of Strathclyde
Glasgow, UK

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

Contents

List of Figures

List of Tables

Acknowledgements

While it is difficult to write a formal document in English when it is not your mother tongue, it is even more difficult to use English to express your feelings. It is hard to believe that four years can pass so quickly and it has been a unique and unforgettable experience of which I have learned much in every way. “A jigsaw puzzle will be not composed if it is not for the union of many pieces” and I would like to highlight some of them.

This would not be possible without my supervisor, Dr. Gabriel Barrenechea. Thank Gabriel for trusting me from the beginning. Every meeting has been a great help. Muchas gracias Gabriel.

I also gratefully acknowledge to the University of Strathclyde, and in particular to the Department of Mathematics and Statistics, for giving me this opportunity to perform my studies in the best environment. And thank Andreas, Heather and Michal for being such good colleagues.

Thanks to Dr. Tomás Chacón, who advised me to take advantage of this PhD opportunity. His support was essential for me to start this adventure. Gracias Tomás.

The main “piece”, of this puzzle and of my life, has been my family. It was not easy for us when I started a new life in another country but they have always supported me with every decision I have made over my 30 years. “ I could not let you down and this has been my biggest motivation to build this puzzle which I started four years ago”. “No podía defraudaros y ha sido mi mayor motivación para no abandonar y conseguir mi objetivo que empecé hace cuatro años”. Gracias Freud, mamá y papá.

This paragraph maybe is the most difficult to express my largest acknowledgement. There is a person who has been with me every day, in my best and worst moments during my PhD. “He has illuminated the darkest days in Glasgow”. His love has made me stronger to finish this puzzle. “One step closer”. Gracias MTFVRT.

Last, but no least, another “thank you” to all my friends. There are too many to list by name, although I would like to mention one of them specially. She is not just my best friend, but my sister. Nobody knows me better than her and that is why her support has been very important during these four years telling me again and again “never give up”. Gracias Pi.

Abstract

This thesis deals with the solution of the Laplace and heat equations on complicated domains. The approach follows the idea of the fictitious domain method, in which a larger (simpler) domain is introduced with the idea of avoiding the use of meshes that resolve the geometry.

The first part of the thesis is dedicated to propose and analyse a new stabilised finite element method for the heat equation. The analysis, not available to date, is based on the introduction of a new projected initial condition that satisfies the boundary conditions of the original problem weakly. This allows us to prove inconditional stability and optimal convergence of the solution, thus avoiding the restriction linking the time discretisation and mesh width parameters present in previous references.

In the second part of this thesis the methodology has been adapted and extended to cover the case in which the problem at hand is posed in a domain containing several inclusions of small size. For this case, the usual fictitious domain approach is no longer applicable, and then a new method that compensates for the lack of stability of the original one is proposed, analysed and tested numerically. The numerical analysis has been carried out for the steady state case, but its applicability to time dependent problems is sketched and shown by means of numerical experiments.

Chapter 1

Introduction

This thesis aims at studying some stabilised Finite Element Methods (FEM) for problems on complicated domains. The approach used herein will allow the use of *uniform meshes*, that is, meshes containing cells with the same shape and size. Although the solution u of the problem is unknown, some of its physical properties may be known. Suppose, for example, that u is the temperature of a metal bar with constant length L. We provide heat at the ends of this bar. So, the magnitude of the temperature in these zones, $|u|$, is higher than in the rest of the bar. It will increase if we keep heating the metal bar. Otherwise, this magnitude will decrease and will become eventually constant throughout the bar. By using uniform meshes, we can exploit this fact and save computational effort when computing a discrete approximation u_h of u.

Let $\omega \subseteq \mathbb{R}^2$ be a bounded open domain with *Lipschitz continuous boundary* $\gamma := \partial\omega$, which means that the boundary ω can locally be parametrized by a Lipschitz continuous function. A more general definition can be found in [GR86].

Then for given data $\hat{f} : \omega \longrightarrow \mathbb{R}$ and $g : \gamma \longrightarrow \mathbb{R}$, we will consider the following *steady-state* problem for the Laplace equation:

find $\hat{u} : \omega \longrightarrow \mathbb{R}$ such that

$$\begin{cases} -\Delta\hat{u} = \hat{f} & \text{in} \quad \omega, \\ \hat{u} = g & \text{on} \quad \gamma. \end{cases} \tag{1.1}$$

The domain ω can be very complicated in shape (see Fig. 1.1).

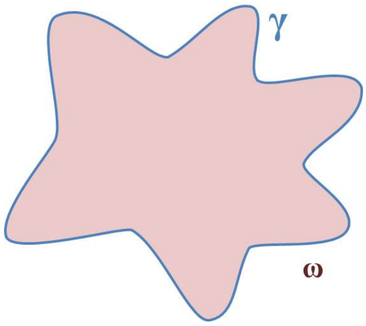


Fig. 1.1. The real domain ω.

We now present the weak form of the problem (1.1):
find $\hat{u} \in H^1(\omega)$ such that

$$\begin{cases} \int_\omega \nabla\hat{u}\nabla\hat{v}\,dx &= \int_\omega \hat{f}\hat{v}\,dx, \\ \hat{u} &= g \quad \text{on} \quad \gamma. \end{cases} \tag{1.2}$$

for all $v \in H_0^1(\omega)$. To obtain this formulation we integrated by parts over the domain ω, that is, we applied Green's formula (see [GR86]). This formula is proven to be valid on domains whose boundary is Lipschitz continuous. All domains considered in this thesis will have a Lipschitz continuous boundary.

Due to the complicated nature of ω, building a uniform mesh of it is not possible especially if its shape changes with time. Therefore, we need to solve the problem introducing a fictitious domain formulation. We first define Ω, an open bounded set such that $\omega \subseteq \Omega$ where Ω should be "simpler" than ω (see Fig. 1.2). A regular mesh can be defined in Ω (even a uniform mesh) and we avoid building a mesh for each time step. We introduce extensions $f : \Omega \longrightarrow \mathbb{R}, u_0 : \Omega \longrightarrow \mathbb{R}$ such that $f|_\omega = \hat{f}$.

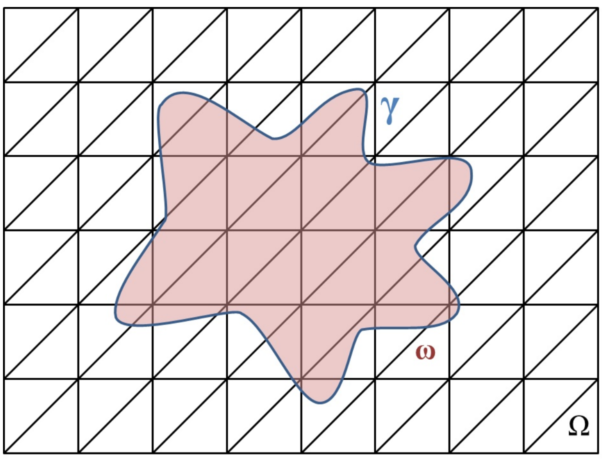


Fig. 1.2. The very complicated domain ω in a fictitious domain Ω.

Hereafter, we adopt the standard notation for Sobolev spaces (see, e.g., [EG04]). Let $D \subseteq \mathbb{R}^2$ and let $L^2(D)$ be the set of square integrable functions on D with the inner product denoted by $(u,v)_D := \int\limits_D uv\,dx$ and the corresponding norm $\|v\|_{0,D}^2 := (v,v)_D$.

For $l \in \mathbb{N}_+$, $H^l(D)$ denotes the usual Sobolev space containing functions whose weak (or distributional) derivatives up to order l belong to $L^2(D)$. The functions in $H_0^1(D)$ belong to $H^1(D)$ and vanish on the boundary ∂D. We denote the norm and seminorm in $H^l(D)$ by $\|.\|_{l,D}$ and $|.|_{l,D}$, respectively. Another ingredient that will be needed in what follows is the space of traces of functions in $H^1(D)$. First, we define the trace map

$$\begin{aligned}\gamma_0 : H^1(D) &\longrightarrow L^2(\partial D)\\ v &\longmapsto \gamma_0(v) = v|_{\partial D}.\end{aligned}$$

This mapping is well-defined and continuous (see [EG04]). Moreover, its range

$$H^{\frac{1}{2}}(\partial D) = \gamma_0(H^1(D))$$

is a Banach space when equipped with the trace norm

$$\|\mu\|_{\frac{1}{2},\partial D} = \inf_{\substack{v\in H^1(D)\\ \gamma_0(v)=\mu}} \|v\|_{1,D}.$$

Its dual with respect to the $L^2(\partial D)$ inner product is denoted by $H^{-\frac{1}{2}}(\partial D)$, and the duality pairing between them is denoted by $\langle .,. \rangle_{\partial D}$. Its norm is denoted by $\|.\|_{-\frac{1}{2},\partial D}$ (see [GR86] for more details). Then, using the following notation,

$$a(u,v) = (\nabla u, \nabla v)_\Omega \quad \text{and} \quad b(\lambda, v) = \langle \lambda, v\rangle_\gamma,$$

the problem (1.1) is written, in an equivalent form, as (see [GG95]):

find $(u,\lambda) \in H_0^1(\Omega) \times H^{-\frac{1}{2}}(\gamma)$ such that

$$\left\{\begin{array}{rcl} a(u,v) - b(\lambda,v) &=& (f,v)_\Omega\\ b(\mu,u) &=& b(\mu,g),\end{array}\right. \tag{1.3}$$

for all $v \in H_0^1(\Omega)$ and all $\mu \in H^{-\frac{1}{2}}(\gamma)$. More details about λ and the link between the solution of the problem (1.1) and the solution of (1.3) will be provided later on.

Throughout the manuscript, C will denote a positive constant independent of any discretization parameter, and whose value may vary whenever it is written in two different places.

The remainder of this chapter is structured as follows. We first set the neccesary notation and recall sufficient conditions to solve problems of type (1.3) followed by a literature review. Afterwards, an outline of the thesis is given.

1.1. Context for Literature review

We include this section to motivate some detailed statements in the later review of existing literature. We will assume basic knowledge of the FEM, see for example [Cia78, BS08] for thorough introductions to the topic.

We say a problem is *well-posed* if it has a unique solution which, in addition, depends continuously on the data. By applying the Babuška-Brezzi's theorem (cf. [Bab73] or [Bre74]), it is easy to prove that the problem (1.3) is well-posed. Indeed, the following result makes use of this theorem to prove this fact (see [GG95]).

Corollary 1.1. *If the bilinear form a is elliptic on $H_0^1(\Omega)$, that is, there exists a constant $k > 0$ such that*

$$a(v,v) \geq k\|v\|_{1,\Omega}^2 \quad \forall v \in H_0^1(\Omega), \tag{1.4}$$

and if b satisfies the inf-sup condition, so there exists a constant $\beta > 0$ such that

$$\sup_{v \in H_0^1(\Omega)} \frac{b(v,\mu)}{|v|_{1,\Omega}} \geq \beta\|\mu\|_{-\frac{1}{2},\gamma} \quad \forall \mu \in H^{-\frac{1}{2}}(\gamma), \tag{1.5}$$

then the problem (1.3) is well-posed.

Proof. The *Poincaré inequality*, $\|v\|_{0,\Omega} \leq C_\Omega |v|_{1,\Omega}$ for all $v \in H_0^1(\Omega)$, guarantees the validity of (1.4):

$$a(v,v) = |v|_{1,\Omega}^2 \geq k\|v\|_{1,\Omega}^2.$$

On the other hand, the inf-sup condition (1.5) is an easy consequence of the fact that, by definition of dual normal on $H^{-\frac{1}{2}}(\gamma)$,

$$\|\mu\|_{-\frac{1}{2},\gamma} = \sup_{\theta \in H^{\frac{1}{2}}(\gamma)} \frac{\langle v,\theta\rangle_\gamma}{\|\theta\|_{\frac{1}{2},\gamma}}.$$

Hence, the problem (1.3) has a unique solution. □

Condition (1.5) is called *inf-sup condition*, or *LBB condition*, where LBB are the initials of the authors Ladyzhenskaya, Babuška and Brezzi, who independently proposed the condition.

In the resulting mixed formulation (1.3), the primal variable and the multiplier are now members of linear spaces which are usually more amenable to approximation.

Provided the LBB condition holds, the mixed formulation is well-posed. Hence, the mixed problem (1.3) has a unique solution pair (u, λ) where u is the extension of the solution $\hat{u}$ of (1.1) to the fictitious domain Ω. Thus the restriction of u to the initial and complicated domain ω is the (unique) solution of the model problem. Futhermore, the *Lagrange multiplier* λ satisfies

$$\lambda = -[\![\partial_{\mathbf{n}} \hat{u}]\!]_{\gamma}, \tag{1.6}$$

where $[\![\partial_{\mathbf{n}} \hat{u}]\!]_{\gamma}$ denotes the jump of $\frac{\partial \hat{u}}{\partial \mathbf{n}}$ across γ, i.e.

$$[\![\partial_{\mathbf{n}} \hat{u}]\!]_{\gamma} = \nabla \hat{u}|_{\omega} \cdot \mathbf{n} - \nabla \hat{u}|_{\Omega \setminus \omega} \cdot \mathbf{n}, \tag{1.7}$$

and $\mathbf{n}$ denotes the unit normal vector to γ exterior to ω. A brief proof of (1.6) is the following. If $\varphi \in C_0^\infty(\Omega \setminus \omega)$:

$$\int\limits_{\Omega \setminus \omega} \nabla u \cdot \nabla \varphi \, dx = (f, \varphi)_{\Omega \setminus \omega} \quad \Rightarrow \quad -\Delta u = f \quad \text{in} \quad \Omega \setminus \omega.$$

Since $f \in L^2(\Omega)$, then the equality $-\Delta u = f$, a priori only valid in $D'(\Omega \setminus \omega)$, is also valid as function of $L^2(\Omega \setminus \omega)$. If $\varphi \in C_0^\infty(\omega)$:

$$\int\limits_{\omega} \nabla u \cdot \nabla \varphi \, dx = (f, \varphi)_{\omega} \quad \Rightarrow \quad -\Delta u = f \quad \text{in} \quad \omega.$$

Since $f \in L^2(\Omega)$, then the equality $-\Delta u = f$, a priori only valid in $D'(\omega)$, is also valid as function of $L^2(\omega)$. From the previous consideration, for $v \in H_0^1(\Omega)$ and applying Green's formula, we get

$$\int\limits_{\Omega \setminus \omega} (\Delta u - f) v \, dx + \int\limits_{\omega} (\Delta u - f) v \, dx - \langle (\partial_{\mathbf{n}}(u|_{\Omega \setminus \omega}) - \partial_{\mathbf{n}}(u|_{\omega})), v \rangle_{\gamma} - \langle \lambda, v \rangle_{\gamma} = 0,$$

for all $v \in H_0^1(\Omega)$. Therefore,

$$\langle \lambda, v \rangle_{\gamma} = -\langle [\![\partial_{\mathbf{n}} u]\!], v \rangle_{\gamma} \quad \forall v \in H_0^1(\Omega),$$

which implies

$$\langle \lambda, v \rangle_{\gamma} = -\langle [\![\partial_{\mathbf{n}} u]\!], v \rangle_{\gamma} \quad \forall v \in H^{\frac{1}{2}}(\gamma).$$

Hence $\lambda = -[\![\partial_{\mathbf{n}} u]\!] \in H^{-\frac{1}{2}}(\gamma)$ (see [GG95] and [GPP94] for more details).

Let us now state a fictitious domain formulation of (1.3). To this end, we define some terminology. A FEM seeks a solution in a trial space and tests the equality for every function in a test space. If trial and test spaces coincide, then the method is called the *Galerkin method*, otherwise it is called a *Petrov-Galerkin method* or a *non-standard Galerkin method* (see [EG04]).

Now, $\overline{\Omega}$ is covered by a family of triangulations $\{\mathcal{T}_h\}_{h>0}$ consisting of non-overlapping triangles where the subscript h refers to the level of refinement of the triangulation. Futhermore, let $V_h \subset V$ and $\Lambda_h \subset \Lambda$ be discrete, finite-dimensional, spaces associated with $\mathcal{T}_h$. We define the following finite element spaces:

$$\begin{aligned} V_h &= \{v_h \in C^0(\bar{\Omega}) \cap H_0^1(\Omega) : v_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}, \\ \Lambda_h &= \{\mu_h \in L^2(\gamma) : \mu_h|_e \in \mathbb{P}_0(e), \forall e \in \gamma_h\}, \end{aligned}$$

such that $\Lambda_{\tilde{h}} \subseteq \Lambda_h$. Then, a Galerkin method for (1.3) reads as follows:

find $(u_h, \lambda_h) \in V_h \times \Lambda_h$ such that

$$\begin{cases} a(u_h, v_h) - b(\lambda_h, v_h) &= (f, v_h)_\Omega \\ b(\mu_h, u_h) &= b(\mu_h, g), \end{cases} \tag{1.8}$$

for all $(v_h, \mu_h) \in V_h \times \Lambda_h$. Applying the classical results for this type of problem, (cf. [BF91]), we can state the following result.

Corollary 1.2. *Let us suppose that there exists a constant $k > 0$ independent of h such that*

$$a(v_h, v_h) \geq k\|v_h\|_{1,\Omega}^2 \quad \forall v_h \in V_h, \tag{1.9}$$

and there exists a constant $\tilde{\beta} > 0$ independent of h, such that

$$\sup_{v_h \in V_h} \frac{b(v_h, \mu_h)}{\|v_h\|_{1,\Omega}} \geq \tilde{\beta}\|\mu_h\|_{-\frac{1}{2},\gamma} \quad \forall \mu_h \in \Lambda_h. \tag{1.10}$$

Then the problem (1.8) has a unique solution and there exists a constant $C > 0$, independent of h and of u, v, u_h, v_h, such that

$$\|u - u_h\|_{1,\Omega} + \|\lambda - \lambda_h\|_{-\frac{1}{2},\gamma} \leq C\left\{\inf_{v_h \in V_h} \|u - v_h\|_{1,\Omega} + \inf_{\mu_h \in \Lambda_h} \|\lambda - \mu_h\|_{-\frac{1}{2},\gamma}\right\}. \tag{1.11}$$

Unlike an elliptic problem, (1.8) is not automatically well-posed by choosing subspaces $V_h \subset H_0^1(\Omega)$ and $\Lambda_h \subset H^{-\frac{1}{2}}(\lambda)$. This is one source of instability that may occur solving the problem (1.8).

A very popular approach to prove the discrete inf-sup condition (1.10) is using the Fortin Operator. If the continuous inf-sup condition holds, then the discrete inf-sup condition is equivalent to the existence of a restriction operator satisfying specific properties. A direct application of Fortin's lemma (see [For77]) gives the following result.

Lemma 1.3. *Assume that b satisfies the inf-sup condition* (1.5). *Then the discrete inf-sup condition* (1.10) *holds if and only if there exists a restriction operator $\Pi_h \in \mathcal{L}(V, V_h)$ with two properties:*

$$\|\Pi_h(v)\|_{1,\Omega} \leq C\|v\|_{1,\Omega} \quad \forall v \in H_0^1(\Omega), \tag{1.12}$$

where $C > 0$ is a constant independent of h, and

$$b(\Pi_h(v) - v, \mu_h) = 0 \quad \forall \mu_h \in \Lambda_h, \forall v \in V. \tag{1.13}$$

We now present the time-dependent problem to be considered in this thesis:
find $\hat{u} : \omega \times (0,T) \longrightarrow \mathbb{R}$ such that

$$\begin{cases} \partial_t \hat{u} - \Delta \hat{u} &= \hat{f} \quad \text{in} \quad \omega \times (0,T), \\ \hat{u} &= g \quad \text{on} \quad \partial\omega \times (0,T), \\ \hat{u}(x,0) &= \hat{u}_0(x) \quad \text{in} \quad \omega. \end{cases} \tag{1.14}$$

for given data $\hat{f} : \omega \times (0,T) \longrightarrow \mathbb{R}$, $g : \gamma \times (0,T) \longrightarrow \mathbb{R}$ and $\hat{u}_0 : \omega \longrightarrow \mathbb{R}$ with $T > 0$ a final time.

Using the fictitious domain formulation, and the extensions $f : \Omega \longrightarrow \mathbb{R}, u_0 : \Omega \longrightarrow \mathbb{R}$ such that $f|_\omega = \hat{f}, u_0|_\omega = \hat{u}_0$, the weak transient problem is given by:
find $(u, \lambda) \in L^2(0,T; H_0^1(\Omega)) \times L^\infty(0,T; H^{-\frac{1}{2}}(\gamma))$ such that $u(x,0) = u_0(x)$, and

$$\begin{cases} (\partial_t u, v)_\Omega + a(u,v) - b(\lambda, v) &= (f,v)_\Omega \\ b(\mu, u) &= b(\mu, g), \end{cases} \tag{1.15}$$

for all $v \in H_0^1(\Omega)$, all $\mu \in H^{-\frac{1}{2}}(\gamma)$, and almost all $t \in (0,T)$.

We discretise the problem (1.15) with respect to the space and time variables. Lagrange finite elements are used for the space discretisation, and first order backward difference formula are used for the time discretisation.

For space discretisation, let $\{\mathcal{T}_h\}_{h>0}$ denote a shape-regular family of triangulations of the domain Ω. For each triangulation $\mathcal{T}_h$, the subscript h refers to the level of refinement of the triangulation, which is defined by

$$h = \max_{K \in \mathcal{T}_h} h_K,$$

where $h_K = diam(K)$. Let γ_h and $\gamma_{\tilde{h}}$ be two partitions of γ such that the vertices of $\gamma_{\tilde{h}}$ are also vertices of γ_h, with edges $\tilde{e}$ satisfying the following (see [GG95]): there exists

$C > 0$ (independent of h) such that $3h \leqslant |\tilde{e}| \leqslant Ch$, for all $\tilde{e} \in \gamma_{\tilde{h}}$. We suppose that for all $\tilde{e} \in \gamma_{\tilde{h}}$, $\text{card}\{e \in \gamma_h : e \subset \tilde{e}\} \leqslant C$, where $C > 0$ is independent of $\tilde{e}$ and h. In particular, we can define γ_h as the partition of γ induced by $\mathcal{T}_h$. This is, the collection of edges e such that their end points are the intersections of γ with the edges of the triangulation $\mathcal{T}_h$, plus the angular points of γ (see Fig. 1.3).

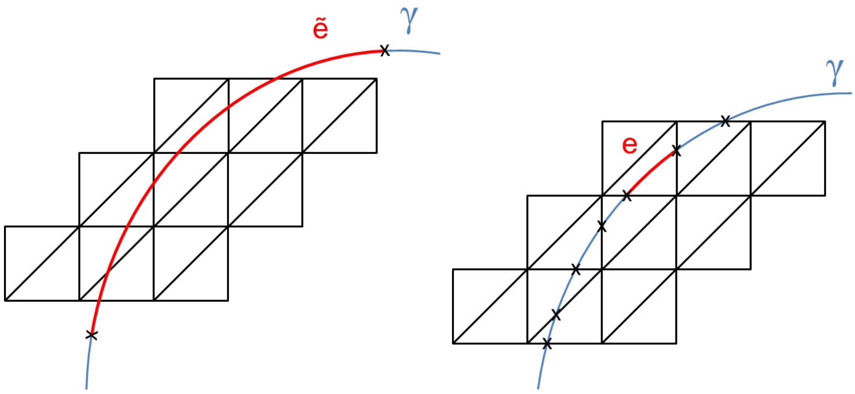


Fig. 1.3. Example of meshes on γ.

For the time discretisation, we use the implicit Euler method (although the same analysis can be extended to more involved schemes). Let $N \in \mathbb{N}_0$ be given. We consider a uniform partition $\{(t_n, t_{n+1}]\}_{0 \leqslant n \leqslant N-1}$, with $t_n = n\delta t$, of the time interval of interest $[0, T]$ with time-step size $\delta t = \frac{T}{N}$.

Then, given subspaces $V_h \subseteq H_0^1(\Omega)$ and $\Lambda_h \subseteq H^{\frac{1}{2}}(\gamma)$, the semi-discrete problem for (1.15) reads:

$$\left\{ \begin{array}{rcl} (\partial_t u_h, v_h)_\Omega + a(u_h, v_h) - b(\lambda_h, v_h) & = & (f, v_h)_\Omega \\ b(\mu_h, u_h) & = & b(\mu_h, g), \end{array} \right. \tag{1.16}$$

for all $v_h \in V_h$, all $\mu_h \in \Lambda_h$, and almost all $t \in (0, T)$.

For this time-dependent problem the situation is even more delicate. In fact, besides the discrete inf-sup condition (1.10), some compatibility in the discrete initial condition is needed (more on this later).

1.2. Literature review

In this section, we perform a literature review to solve the problem (1.8) through FEM with fictitious domain to obtain stability and optimal convergence. Fictitious domain method for partial differential equations have shown recently potential for solving complicated problems from Science and Engineering. The main reason for this popularity of fictitious domain methods (sometimes called *Domain Imbedding methods*) is that they allow the use of fairly structured meshes on a simple shape auxiliary domain containing the actual one, allowing therefore the use of fast solvers (see [GPP94] and [CJM96] for more details).

When we work with fictitious domains, a proper extension of the function has to be chosen, if we want to ensure optimal convergence. We need to choose an appropiate extension of $\hat{f}$ to Ω to ensure that (u, λ), the solution of the problem (1.3), is regular enough. We work on this direction in Chapter 2, although some preliminar results are given in [FGM13, Loz16]. Methods like the classical penalty method or the Lagrange multipliers method are known to produce a solution with a non-optimal order of the error, in particular for immersed bodies, as we can read in [FGM13]. The authors mention other methods that exist in the finite element framework in order to recover this optimal order, such as the Fat Boundary method, method based on Nitsche's formulation or control approach methods (see, e.g, [Mau01, CB09, ADG+91] for more details). The method presented in [FGM13] is in the spirit of these control approach methods and aims at recovering the optimal error order. The authors present a smooth extension method with a Poisson equation and introduce a method of the fictitious domain type to simulate the motion of immersed rigid bodies. Their idea is to find a smooth extension of the exact solution in the domain $\Omega \backslash \omega$ (where ω is a sphere included in Ω) to the whole and fictitious domain Ω by finding a suitable extension of the right-hand side in the inclusions. In [Loz16], an optimal convergence through fictitious domain method is proved without cut elements introducing stabilisation terms. A. Lozinski avoids the integration over the elements cut by the boundary of the problem domain as in Extended FEM (XFEM) and CutFEM for example. A Poisson problem is considered in a domain ω which is embedding into a simply shaped domain Ω and there is a quasi-uniform mesh $\mathcal{T}_h$ on Ω that can cut ω in an arbitrary manner. In this work, the author does not extend the solution u or the initial function f in ω to Ω as we have done in Chapter 3. He defines another fictitious domain ω_h such that $\omega \subset \omega_h \subset \Omega$ and extends u and f from ω to ω_h. The basic difference doing that is that he needs the extension only in a narrow strip which minimizes the effect of choosing a "wrong" extension and allows to avoid its explicit construction. The author mentions the difficulty of an extension of his work to higher-order finite element solving this one in $\mathbb{P}_1$ continuous finite element on a triangular mesh.

A fictitious domain formulation and its abstract discretisation was described in [GG95]. In this paper, a non-homogeneous elliptic Dirichlet problem is solved through FEM defining the model problem and the mixed problem. With regard to uniqueness and regularity of the solution of the model problem, one can use the following theorem given in [BH91].

Theorem 1.4. *Let* $\hat{f} \in H^p(\omega), p \geq 0$ *and* $g \in H^q(\gamma), q \geq \frac{1}{2}$*. Then there exists a unique weak solution* $\hat{u} \in H^1(\Omega)$ *of* (1.1) *and* $\frac{\partial \hat{u}}{\partial n} \in H^{-\frac{1}{2}}(\gamma)$.

The analysis of a low-order FEM for (1.8) was first presented in [GG95]. In there, the authors imposed a condition on the mesh $\mathcal{T}_h$ (which induces the space V_h), and on a partition of γ (which induces the space Λ_h) that allows them to prove a uniform inf-sup condition. More precisely, the authors prove the discrete inf-sup condition where the negative boundary norm (used until then) was replaced by the L^2 norm on the boundary through an inverse inequality and the discrete inf-sup condition was replaced by a sufficient condition that must be checked in the applications. This negative boundary norm was eliminated by constructing a Fortin operator.

To study the validity of (1.10), the use of linear triangular elements to approximate u and linear element for λ in the two-dimensional case is proposed in [Bab73]. The size of the elements for λ has to be greater than, or equal to, the diameter of the triangles along $\partial\Omega$ times a constant $C > 1$ depending on Ω. With this condition, stability and optimal-order convergence hold for the problem (1.8). The first works using the idea of introducing a Lagrange multiplier to impose a boundary condition weakly are [Bab73, AB72, BTOL77]. They established error estimates when the ratio between the boundary mesh size and the mesh size in the domain is greater than some constant depending on the domain. Unfortunately, the constant can be large and its dependence on the domain is not straightforward. Their results were refined by Pitkäranta in [Pit79] and Agouzal in [Ago93] where the boundary mesh points are directly related to the mesh points of the interior grid.

Later on, in [GG95], the authors propose a discretisation using linear continuous elements for V_h and piecewise constant elements for Λ_h, where the mesh size in given by γ is $\tilde{h} \neq h$. Under the condition $\tilde{h} \geq 3h$, they could prove the existence of a Fortin operator (and hence, the inf-sup condition). This condition has been shown, by performing numerical experiments, to be necessary. Unfortunately, it prevents the use of the most natural meshes, for example the mesh γ_n defined as the partition of γ induced by $\mathcal{T}_h$.

To overcome this fact, a local projection stabilisation method is proposed in [BC12] for solving a fictitious domain problem where this restriction between the meshes is not satisfied. The authors add a suitable fluctuation term to the formulation, thus yielding the natural space for the Lagrange multiplier stable. This stabilised problem is well-posed and stability and convergence are proved.

Other approaches have also been proposed in the recent years to avoid this restriction, such as using cut elements as in [BH10, BH12] or XFEM approaches as in [MDB99, MBT06, HR09] for instance.

In [BH10], the authors propose a stabilised Lagrange multiplier method introducing a fictitious domain formulation where the mesh is cut by the boundary. The main idea behind cut elements is to consider the restriction of the finite element space to the physical domain. This implies that the integration needs to be done in the intersection of the underlying mesh in the fictitious domain with the physical domain. One advantage of this approach is that it does not need the mesh on the physical boundary to be coarser than the underlying mesh. On the contrary, the intersection of the physical domain and the mesh needs to be tracked, and can be of very diverse nature giving very different elements in the matrices and leading to possibly very ill-conditioned matrices.

Similar ideas have been proposed in the framework of the Extended FEM. It was introduced in [MDB99] when dealing with cracked domains. The technique allows the entire crack to be presented independently of the mesh and so remeshing is not necessary as the crack grows. The method consists of a FEM enriching the solution space for solutions to differential equation with discontinuous functions. Based on this technique plus stabilisation like the one from Barbosa-Hughes (see [BH91]), an optimally convergent method was proposed in [HR09]. Since the problem is only considered on the physical domain, the Lagrange multiplier is given by the normal derivative of the primal variable (and not its jump as in the method of [GG95]).

The introduction of a Lagrange multiplier is also used for the approximation of interface problems in [ABG+15] and has some similarities with the fictitious domain approach with distributed Lagrange multiplier that is developed in [BGR14]. In [ABG+15], the authors consider the problem with two regions, Ω_1 and Ω_2, with different materials using two meshes which fit with the interface Γ. Both meshes have to share the nodes on Γ to impose continuity at the boundary but if Γ depends on time, one of these meshes has to change close to the interface. To avoid this inconvenience, in [BGR14] the authors propose a fictitious domain formulation with a distributed Lagrange multiplier where one mesh is for the whole domain Ω and one for the region Ω_2 where only one material is present. Both problems are equivalent if the solution u restricted to Ω_2 is the solution u_2. Then, introducing a Lagrange multiplier associated to this constraint, the authors prove the convergence of the method.

All the above approaches deal with steady-state problems for which the formulation and analysis of stabilised methods is fairly well understood. However, the design of robust and efficient stabilised methods for their transient counterparts is hardly a settled matter.

Some of the most effective algorithms for treating time-dependent problems can be defined through a process wherein the spatial and temporal discretisations are separated. Such algorithms are especially well adapted to the cylindrical nature of the time-space domain and usually posses excellent stability characteristics. Another reason for their popularity is that they reduce the partial differential equation to either a system of ordinary differential equations or, for problems with constraints, to a system of differential algebraic equations. Fully discrete formulations in which spatial and temporal discretisations are carried out separately are in much more common use than are coupled time-space formulations. An additional motivation for this choice is the desire to avoid the increase in the number of unknows required to achieve more accuracy in space-time formulations.

In the context of mixed problems, such as the Stokes problem or the problem (1.16), numerical experiments have shown that fully discrete algorithms are completely adequate for transient calculations carried out for moderate to relatively large time steps. However, in settings that require very fine time resolution, the behavior of such algorithms is not very well understood. The behavior just described is especially notorious if stabilisation is used in the space discretisation. As a matter of fact, if the Stokes problem is discretised in space using a residual stabilised finite element method, and the time is discretised using backward Euler, in [BGS04] it is shown that the fully discrete scheme is stable only if

$$\delta t \geq C h^2. \tag{1.17}$$

It is important to notice this is not limited to the backward Euler method. In fact, the same conclusion can be obtained if other time discretisation schemes are used (see [BB07, BF09] for discussions on this topic).

As the authors mention in [BB07], this condition relating the time step and the mesh size is required in the analysis of the stability of the methods described. This lower bound on the time step size was assumed in the stabilised methods for transient flows considered in [CVZ98, BC01, CB00, CPGB07]. It is interesting to observe that this condition may not be required if inf-sup stable discretisations are used in transient problems and some stabilised methods are also free of this restriction (see, e.g, [BGL07, He03]). Then this condition prevents the use of stabilised FEM in combination with very small time steps. Moreover, there is an even deeper reason for this. Loosely speaking, if the initial condition satisfies a restriction (for example, divergence-free in

Stokes, or the boundary condition $u = g$ in (1.16)), then the approximated initial condition should satisfy a similar restriction to prevent instabilities. This remark is the basis of the work [BF09] where an appropriate projection of the initial condition is used to avoid the instabilities. In Chapter 3 we use a similar approach to address problem (1.16).

The previous approaches deal with a physical domain whose size is comparable with the size of the fictitious domain. In this way the restriction imposed on the mesh (i.e. $\tilde{h} \geq 3h$) can be naturally satisfied. A different problem appears if the physical domain is perforated, i.e. domains with holes B_i (or inclusions) that are smaller than the characteristic mesh width. In our case, this amounts to stating that $diam(B_i) << diam(\Omega)$, and in turn, this will imply that, in most cases, $|\tilde{e}| < h$ where $\tilde{e}$ are the curved edges from the mesh on ∂B_i, which is precisely the case not allowed in [GG95]. Over the years many authors have proposed solutions to this problem. One alternative is the Composite FEM (see, e.g., [HS97]), where the geometrical features are included in the finite element space, thus proposing a method whose dimension does not necessarily depend on the number of geometrical inclusions. We can see the application of the same idea to the Stokes problem in [PS08], and an adaptive strategy associated to a discontinuous Galerkin version of this method in [GH14]. Alternatively, the geometrical features of the domain can be taken into account at the mesh generation step. This idea is at the basis of some recent developments on discontinuous Galerkin methods on general polyhedral meshes (see [CGH14], and [ACC+16] for a recent review). Finally, it is interesting to mention the approach described in [BLL14] (see also the references therein for an extensive review of this type of approach), where a multiscale problem on a domain with inclusions has been approximated using a multiscale finite element approach based on the enrichment of the Crouzeix-Raviart method with bubble functions.

1.3. Outline

We conclude this chapter with the outline of this thesis.

In Chapter 2, we have proved that optimal convergence is obtained if we choose a proper extension of the function $\hat{f}$ when we are working with a fictitious domain. We need to choose an appropiate extension of $\hat{f}$ to the fictitious domain Ω to ensure that (u, λ), solution of the problem (1.3), is regular enough. This is not an easy task, especially in dimensions higher than one. We do not know how to extend $\hat{f}$ such that $f \in H^{k-1}(\Omega)$ or $u \in H^{k+1}(\Omega)$ when we are working with n-dimensional problem.

Then we start to work with time-dependent problems in $\mathbb{P}_1$ where we need just $f \in L^2(\Omega)$. In Chapter 3, we develop transient problems studing full analysis for the backward difference formula for order one. The full discretisation of (1.16), using Lagrange elements for the space variables and the backward Euler formula for the time discretisation, is carried out.

In Chapter 4, the purpose is to approximate numerically an elliptic partial differential equation posed on domains with small perforations (or inclusions). The approach is based on the FEM, and since the method's interest lies in the case in which the geometrical features are not resolved by the mesh, we propose a stabilised FEM. Our objective was to propose a simple alternative to the approaches described in the previous sections. Then the stabilisation term is a simple, non-consistent penalization, that can be linked to the Barbosa-Hughes approach. Stability and optimal convergence are proved, and numerical results confirm the theory. We have already submitted a paper with these results [BG16a].

We conclude and present possible further extensions in Chapter 5. We show the conclusions obtained after work developed in this thesis and describe possible future work in two directions mainly.

Finally, we include two appendices. In Appendix A, we describe the calculation done to get the solution and errors of the problems worked in Chapter 2. We consider three different linear spaces to solve an especific problem through FEM with fictitious domain and without that. In Appendix B, we describe the proccess to define the matrix $\boldsymbol{C}$ in FreeFem++. This matrix is coming from introducing the established term j to our the stabilised method described in Chapter 3.

Chapter 2

Numerical experiments with fictitious domain and higher-order elements

2.1. Introduction

In this chapter, we focus on steady problems. We have solved elliptic problem using the FEM .

We have considered the following simple one-dimensional problem for $\mathbb{P}_1, \mathbb{P}_2$ and $\mathbb{P}_3$ finite dimensional spaces:

For $\hat{f}$ given, find $\hat{u}$ such that

$$\begin{cases} -\hat{u}'' &= \hat{f} \quad \text{in} \quad \omega, \\ \hat{u}(0) &= g_0, \\ \hat{u}(1) &= g_1, \end{cases} \tag{2.1}$$

where $g_0, g_1 \in \mathbb{R}$. We obtain the solution and calculate the error in the $L^2(\omega)$ and $H^1(\omega)$ norms between the approximate solution, $\hat{u}_h$, and the exact solution, $\hat{u}$.

Moreover, we consider the same problem (2.1) applying the fictitious domain method. As we have stated in the previous chapter, this approach relies on the introduction of a larger domain $\Omega \supset \omega$, an extension f of $\hat{f}$ to Ω, and the solution of the following mixed problem:

find $(u, \lambda) \in H_0^1(\Omega) \times H^{-\frac{1}{2}}(\partial\omega)$ such that

$$\begin{cases} a(u, v) - b(\lambda, v) &= (f, v)_\Omega \\ b(\mu, u) &= b(\mu, g), \end{cases} \tag{2.2}$$

for all $(v, \mu) \in H_0^1(\Omega) \times H^{-\frac{1}{2}}(\partial\omega)$ where $f \in L^2(\omega)$ and $g \in H^{\frac{1}{2}}(\partial\omega)$. We obtain the solution and calculate the errors in the $L^2(\omega)$ and $H^1(\omega)$ norms as well but restricted to the interval ω with different extensions f of $\hat{f}$ to Ω. The purpose of the calculations is to show that, if the extension f of $\hat{f}$ is not built carefully enough, then the extended solution u is not regular enough, which, in turn, affects the quality of the numerical results.

The remainder of this chapter is organised as follows. In Section 2.2, we consider the problem (2.1) and we show the numerical results implemented in *Matlab* [ACF99]. In Section 2.3, we apply the fictitious domain method to the problem solved in the previous section and we also show the numerical results.

2.2. Method without fictitious domain

In this section, we use u as the solution in the initial domain $(0,1)$ instead of $\hat{u}$ and, f instead of $\hat{f}$ to simplify the notation. Therefore, the weak problem is written as:

find $u \in H_0^1(0,1)$ such that

$$(u', v')_{(0,1)} = (f, v)_{(0,1)} \quad \forall v \in H_0^1(0,1),$$

where f is a given function.

This problem has been solved using a standard FEM. For this, we introduce a uniform partition, $0 = x_0 \le x_1 \le ... \le x_{N+1} = 1$, of $(0,1)$ and denote $h = x_{i+1} - x_i$. Associated to this partition we introduce the finite element space $V_h = \{v_h \in C^0((0,1)) : v_h|_{[x_i,x_{i+1}]} \in \mathbb{P}_l[x_i, x_{i+1}] \quad \text{and} \quad v_h(0) = v_h(1) = 0 \quad \forall i\}$ for $l = 1, 2, 3$. Then, the finite element problem is:

find $u_h \in V_h$ such that

$$(u_h', v_h')_{(0,1)} = (f, v_h)_{(0,1)} \quad \forall v_h \in V_h. \tag{2.3}$$

2.2.1. Applications

We now present the numerical results of the errors solving the problem (2.1) for different values of N, number of subintervals in Ω of length $h = \frac{1}{N}$, and two different functions

$$f_1(x) = -12x^2 + 12x - 2,$$
$$f_2(x) = -e^x.$$

For finite elements of order k the following estimate holds (see [EG04]):

$$\| u - u_h \|_{0,(0,1)} \le Ch^{k+1} \mid u \mid_{k+1,(0,1)}, \tag{2.4}$$

$$\mid u - u_h \mid_{1,(0,1)} \le Ch^k \mid u \mid_{k+1,(0,1)} . \tag{2.5}$$

We have computed the errors $\|u - u_h\|_{0,(0,1)}$ and $|u - u_h|_{1,(0,1)}$. The results are given in Tables 2.1- 2.3 and they follow the optimal behavior predicted in (2.4) and (2.5).

		$\|u - u_h\|_{0,(0,1)}$	order	$\|u - u_h\|_{1,(0,1)}$	order
$f_1(x) = -12x^2 + 12x - 2$	$N = 4$	0.0103		0.0690	
	$N = 8$	0.0027	1.9316	0.0329	1.0685
	$N = 16$	6.8039e-04	1.9885	0.0162	1.0221
	$N = 32$	1.7055e-04	1.9962	0.0081	1
$f_2(x) = -e^x$	$N = 4$	0.0119		0.1288	
	$N = 8$	0.0030	1.9879	0.0645	0.9978
	$N = 16$	7.4906e-04	2.0018	0.0322	1.0022
	$N = 32$	1.8733e-04	1.9995	0.0161	1

Table 2.1. Finite element errors for $\mathbb{P}_1$ finite dimensional space.

		$\|u - u_h\|_{0,(0,1)}$	order	$\|u - u_h\|_{1,(0,1)}$	order
$f_1(x) = -12x^2 + 12x - 2$	$N = 4$	5.4256e-04		0.0162	
	$N = 8$	6.5794e-05	3.0437	0.0041	1.9823
	$N = 16$	8.1596e-06	3.0114	0.0010	2.0356
	$N = 32$	1.0179e-06	3.0029	2.5208e-04	1.9880
$f_2(x) = -e^x$	$N = 4$	1.3467e-04		0.0041	
	$N = 8$	1.6805e-05	3.0025	0.0010	2.0356
	$N = 16$	2.0997e-06	3.0006	2.6014e-04	1.9426
	$N = 32$	2.6244e-07	3.0001	6.5045e-05	1.9998

Table 2.2. Finite element errors for $\mathbb{P}_2$ finite dimensional space.

		$\|u - u_h\|_{0,(0,1)}$	order	$\|u - u_h\|_{1,(0,1)}$	order
$f_1(x) = -12x^2 + 12x - 2$	$N = 4$	2.0487e-05		0.0011	
	$N = 8$	1.2805e-06	3.9999	1.3905e-04	2.9838
	$N = 16$	8.0028e-08	4.0001	1.7381e-05	3.0000
	$N = 32$	5.0018e-09	4.0000	2.1726e-06	3.0000
$f_2(x) = -e^x$	$N = 4$	1.5344e-06		8.2485e-05	
	$N = 8$	9.6526e-08	3.9906	1.0344e-05	2.9953
	$N = 16$	6.0482e-09	3.9963	1.2941e-06	2.9988
	$N = 32$	3.7835e-10	3.9987	1.6179e-07	2.9997

Table 2.3. Finite element errors for $\mathbb{P}_3$ finite dimensional space.

2.3. Method with fictitious domain

The problem (2.1), with $g_0 = 0$ and $g_1 = 1$, is solved in this section with fictitious domain formulation through FEM.

The fictitious domain formulation is obtained by including $\omega = (0,1)$ in a larger domain $\Omega = (-1,2)$, an extension f of $\hat{f}$ to Ω and the solution of the following mixed problem:

find $(u,\lambda) \in H_0^1(-1,2) \times \mathbb{R}^2$ such that

$$\begin{cases} \int_{-1}^{2} u'v'\,dx - \lambda_1 v(1) - \lambda_0 v(0) &= \int_{-1}^{2} fv\,dx, \\ -\mu_1 u(1) - \mu_0 u(0) &= -\mu_1 \cdot 1 - \mu_0 \cdot 0, \end{cases} \tag{2.6}$$

for all $v \in H_0^1(-1,2)$ and for all $\mu_0, \mu_1 \in \mathbb{R}$. Both problems are linked by the fact that if (u,λ) satisfies (2.6), then $u\,|_\omega$ satisfies (2.1).

We use the same procedure as in Section 2.2 for $\mathbb{P}_1$, $\mathbb{P}_2$ and $\mathbb{P}_3$ but using these new data.

The errors are defined by Section 2.2 in the same way for each space but working with fictitious domain. It is worth emphasizing that we cannot obtain the errors in $L^2(-1,2)$ and $H^1(-1,2)$ norms because we do not always know the value of the solution outside the interval (0, 1). If f is defined using the same functional expression as $\hat{f}$, then it is possible to calculate the errors in the fictitious domain. Thus, we have calculated the errors, with different extensions of f, restricted to the interval $(0,1)$, more specifically to $[x_{ele+1}, x_k] \subseteq (0,1)$ in each one where the variables ele and k are defined as

$$\begin{aligned} ele &= max\{i : x_i \leq 0\}, \\ k &= max\{j : x_j \leq 1\}. \end{aligned}$$

The error committed using this approximation for $(0,1)$ is negligible.

2.3.1. Applications

We can prove that, if f is such that $u \in H^{k+1}(-1,2)$, then

$$|u - u_h|_{1,(-1,2)} \leq Ch^k |u|_{k+1,(-1,2)}.$$

This result relies on the fact that we are capable of building an extension of $\hat{f}$ such that the extended solution u is regular. This is not an inmediate task, and is even necessary condition for optimal convergence.

In the following numerical results we have used $\hat{f}(x) = \left(\frac{\pi}{2}\right)^2 \sin\left(\frac{\pi}{2}x\right)$ and three possible extensions,

$$f(x) = f_0(x) = \left(\frac{\pi}{2}\right)^2 \sin\left(\frac{\pi}{2}x\right), \tag{2.7}$$

$$f(x) = f_1(x) = \begin{cases} \hat{f}(x) & \text{if } x \in [0,1], \\ 0 & otherwise, \end{cases}$$

$$f(x) = f_2(x) = \begin{cases} \hat{f}(0) & \text{if } -1 < x < 0, \\ \hat{f}(x) & \text{if } x \in [0,1], \\ \hat{f}(1) & \text{if } 1 < x < 2. \end{cases}$$

For the first case, $f(x) = f_0(x)$ is regular in such a way that u (the extended solution to $(-1, 2)$) is regular as well. This will lead to the fictitious domain method providing optimal order estimates. This is not the case for $f(x) = f_1(x)$ and $f(x) = f_2(x)$. In both these cases u is not regular, which will show in the numerical results. This fact is apparent in Tables 2.4- 2.6 where the errors $\|\hat{u}-u_h\|_{0,(0,1)}$ and $|\hat{u}-u_h|_{1,(0,1)}$ are depicted. We see that the errors in these last two cases are suboptimal. As a matter of that, they show an erratic behavior which, never the less, hints a first order convergence, at best, which is clearly non optimal for $\mathbb{P}_2$ and $\mathbb{P}_3$ elements.

		$\|\hat{u} - u_h\|_{0,(0,1)}$	order	$\|\hat{u} - u_h\|_{1,(0,1)}$	order
$f(x) = f_0(x)$	$N = 16$	0.0040		0.1003	
	$N = 32$	9.9989e-04	2.0001	0.0501	1.0014
	$N = 64$	2.3641e-04	2.0805	0.0240	1.0618
	$N = 128$	5.8988e-05	2.0028	0.0120	1
	$N = 256$	1.4530e-05	2.0214	0.0059	1.0243
	$N = 512$	3.6309e-06	2.0006	0.0030	0.9758
$f(x) = f_1(x)$	$N = 16$	0.0179		0.1696	
	$N = 32$	0.0113	0.6637	0.1989	-0.2299
	$N = 64$	0.0050	1.1763	0.0750	1.4070
	$N = 128$	0.0027	0.8883	0.1013	-0.4337
	$N = 256$	0.0013	1.0544	0.0364	1.4766
	$N = 512$	0.000657	0.9846	0.0509	-0.4837
$f(x) = f_2(x)$	$N = 16$	0.0159		0.1105	
	$N = 32$	0.0087	0.8699	0.0708	0.6422
	$N = 64$	0.0042	1.0506	0.0300	1.2388
	$N = 128$	0.0022	0.9260	0.0269	0.1574
	$N = 256$	0.0011	1	0.0104	1.3710
	$N = 512$	5.3893e-04	1.0293	0.0123	-0.2421

Table 2.4. Finite element errors with fictitious domain for $\mathbb{P}_1$ finite dimensional space.

		$\|\hat{u}-u_h\|_{0,(0,1)}$	order	$\|\hat{u}-u_h\|_{1,(0,1)}$	order
$f(x)=f_0(x)$	$N=16$	1.0036e-04		0.0031	
	$N=32$	1.5640e-05	2.6819	8.6962e-04	1.8338
	$N=64$	1.8679e-06	3.0657	2.1732e-04	2.0006
	$N=128$	2.4544e-07	2.9280	5.5666e-05	1.9646
$f(x)=f_1(x)$	$N=16$	0.0093		0.2061	
	$N=32$	0.0074	0.3296	0.2159	-0.0670
	$N=64$	0.0019	1.9615	0.1085	0.9927
	$N=128$	0.0012	0.6629	0.1154	-0.0890
	$N=256$	0.000447	1.4247	0.0549	1.0718
	$N=512$	0.000243	0.8793	0.0586	-0.0940
$f(x)=f_2(x)$	$N=16$	0.0055		0.0528	
	$N=32$	0.0033	0.6781	0.0556	-0.0746
	$N=64$	0.0014	1.2370	0.0259	1.1021
	$N=128$	7.4326e-04	0.9135	0.0276	-0.0917
	$N=256$	3.5798e-04	1.0540	0.0129	1.0973
	$N=512$	1.8119e-04	0.9824	0.0138	-0.0973

Table 2.5. Finite element errors with fictitious domain for $\mathbb{P}_2$ finite dimensional space.

		$\|\hat{u}-u_h\|_{0,(0,1)}$	order	$\|\hat{u}-u_h\|_{1,(0,1)}$	order
$f(x)=f_0(x)$	$N=16$	1.2780e-06		8.9292e-05	
	$N=32$	8.1123e-08	3.9776	1.1167e-05	2.9993
	$N=64$	4.8945e-09	4.0509	1.3362e-06	3.0630
	$N=128$	3.0723e-10	3.9938	1.6699e-07	3.0003
	$N=256$	1.9623e-11	3.9687	2.0633e-08	3.0167
$f(x)=f_1(x)$	$N=16$	0.0034		0.0966	
	$N=32$	0.0049	-0.5272	0.1934	-1.0014
	$N=64$	9.5763e-04	2.3552	0.0528	1.8730
	$N=128$	7.5707e-04	0.3390	0.1004	-0.9271
	$N=256$	2.4670e-04	1.6177	0.0270	1.8947
$f(x)=f_2(x)$	$N=16$	0.0030		0.0261	
	$N=32$	0.0019	0.6589	0.0478	-0.8730
	$N=64$	8.0295e-04	1.2426	0.0128	1.9009
	$N=128$	4.3097e-04	0.8977	0.0238	-0.8949
	$N=256$	2.0439e-04	1.0763	0.0064	1.8948

Table 2.6. Finite element errors with fictitious domain for $\mathbb{P}_3$ finite dimensional space.

We observe the same fact graphically for the three possible extensions in the same order for each finite dimensional space. We check the optimal order in each first graph for each space.

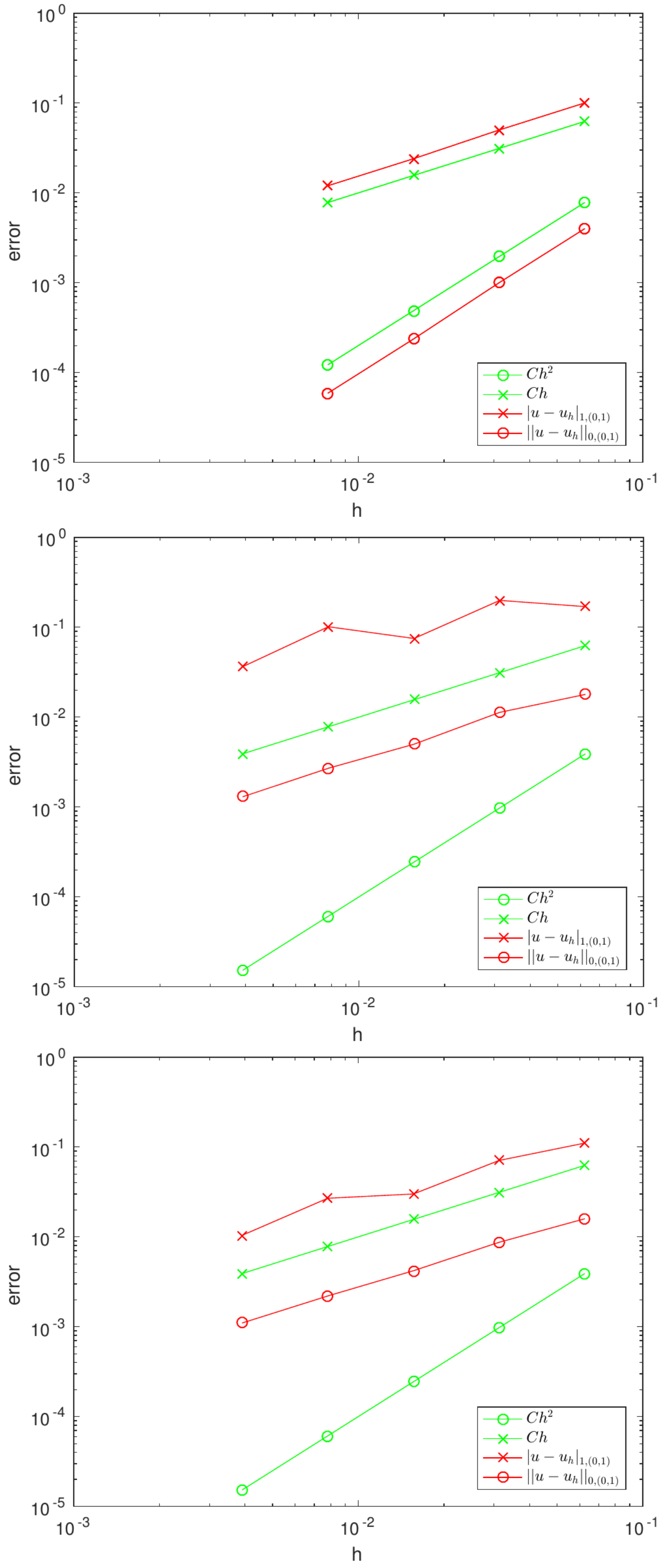


Fig. 2.1. Finite element errors with fictitious domain for $\mathbb{P}_1$ finite dimensional space depending on each extension: f_0 (top), f_1 (middle), f_2 (bottom).

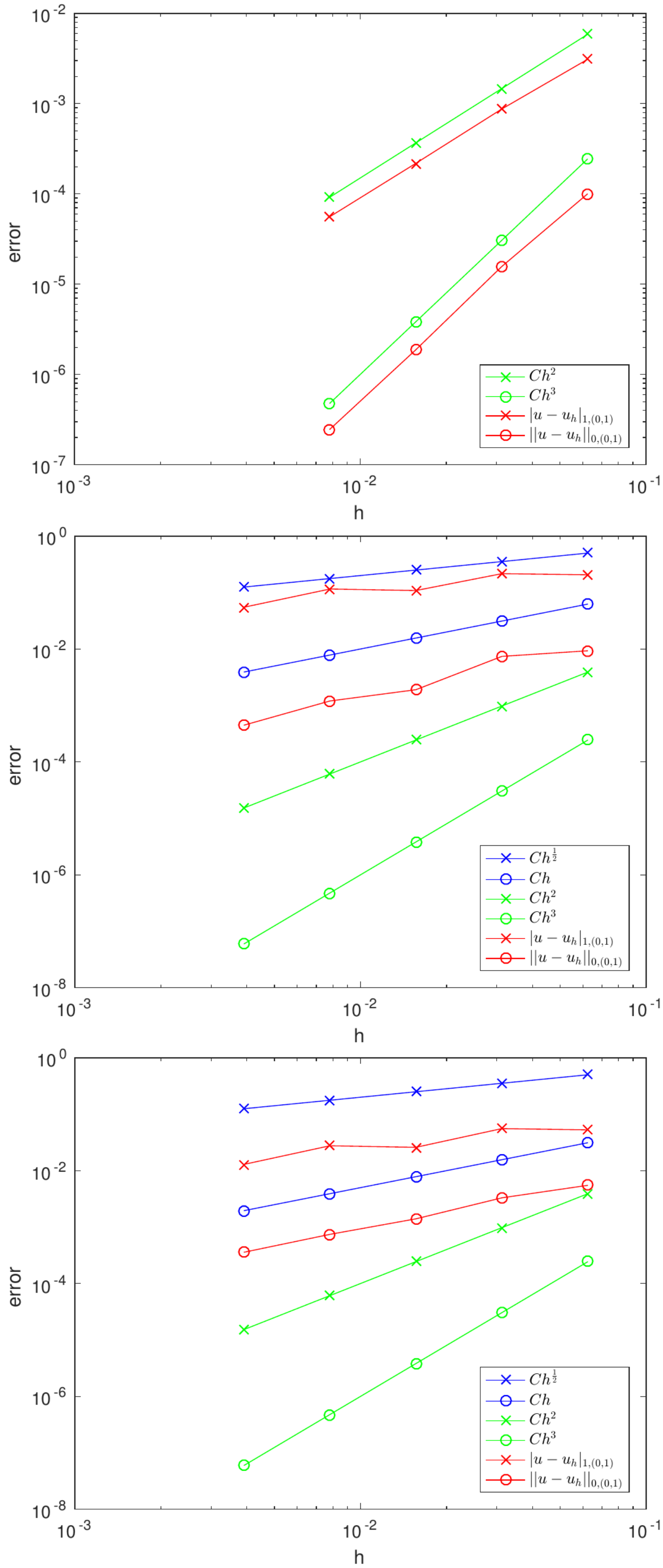


Fig. 2.2. Finite element errors with fictitious domain for $\mathbb{P}_2$ finite dimensional space depending on each extension: f_0 (top), f_1 (middle), f_2 (bottom).

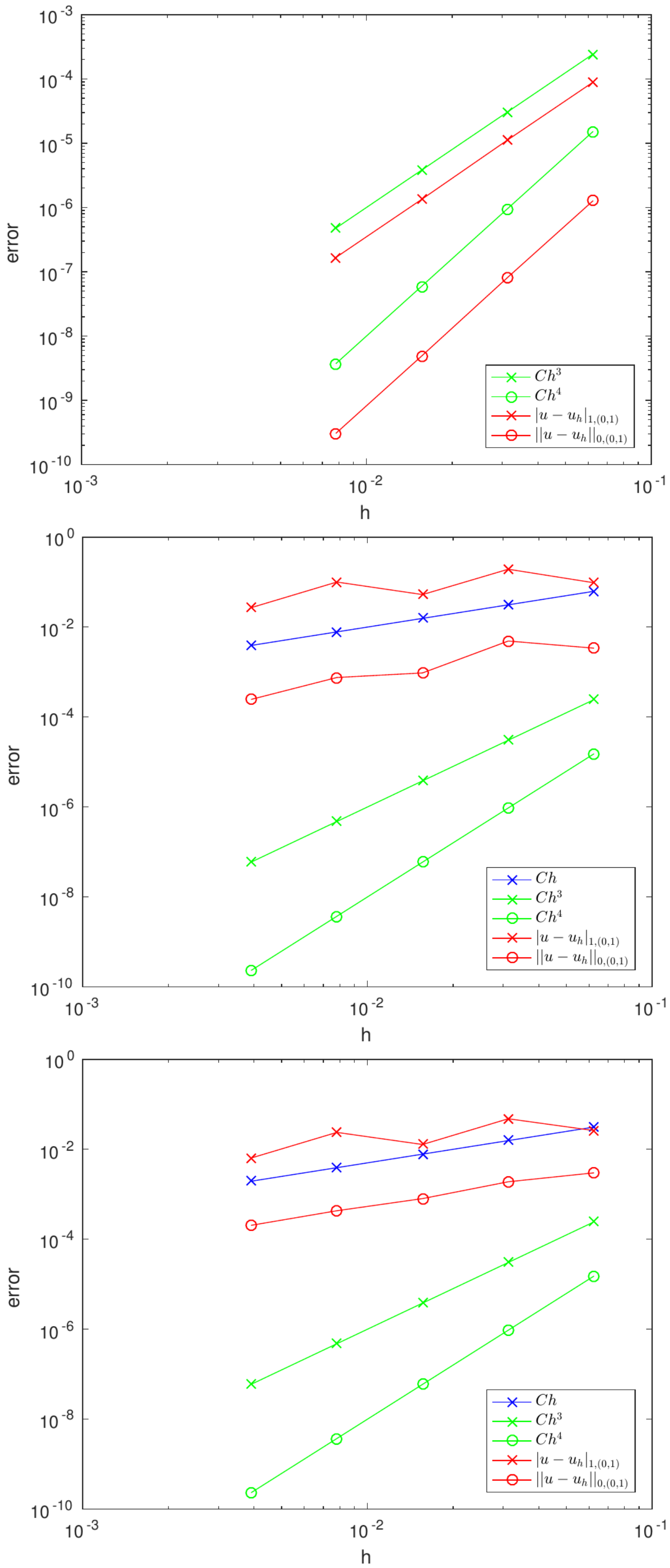


Fig. 2.3. Finite element errors with fictitious domain for $\mathbb{P}_3$ finite dimensional space depending on each extension: f_0 (top), f_1 (middle), f_2 (bottom).

Chapter 3

A fictitious domain method for the transient heat equation

3.1. Introduction

In this chapter, we prove unconditional stability of u and Lagrange multipliers, and optimal converge for the transient heat equation when the initial datum is chosen as a certain Ritz-type projection. In the case when a standard interpolation of the initial data is applied, an *inverse parabolic Courant-Friedrich-Lewy (CFL)-type* condition like the one presented in (1.17), must be respected in order to maintain stability for small time steps. Futhermore, we will distinguish two cases: inf-sup stable condition problem and not inf-sup stable where we have to add the stabilisation operator to obtain a stabilised method (see [BC12]). We base our approach on previous work done for the transient Stokes equations (see, e.g., [BGS04, BF09]).

The remainder of this chapter is organised as follows. In Section 3.2, we introduce the problem under consideration and some useful notation. We study the stability of our problem for the inf-sup stable case and the stabilised method with the space and time discretised formulations in Section 3.3. In Section 3.4, we have the convergence analysis for both cases too. Some numerical results are presented in Section 3.5.

3.2. Problem setting

Let ω be an open bounded domain in $\mathbb{R}^2$ with a Lipschitz continuous boundary $\gamma := \partial\omega$. We consider $\gamma = \bigcup_{i=1}^{M} \gamma_i$, where γ_i are the M smooth components of γ, that is, if γ has corners then $\gamma_i \cap \gamma_{i+1}$ will be its corner points (see Fig. 3.1).

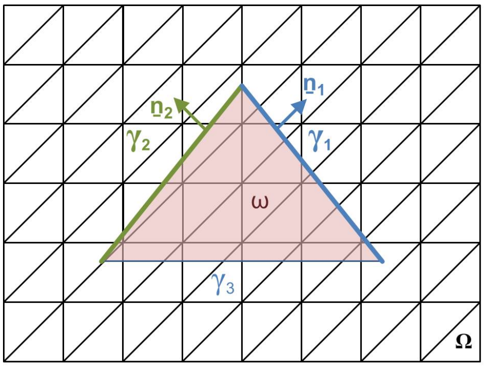


Fig. 3.1. Example of γ where $M = 3$.

In what follows, the following space is used:

$$\mathcal{H} := \prod_{i=1}^{M} H^{\frac{1}{2}}(\gamma_i) \quad \text{with} \quad \|\mu\|_{\mathcal{H}} := \sum_{i=1}^{M} \|\mu\|_{\frac{1}{2},\gamma_i}.$$

For $T > 0$ we consider the problem (recall problem (1.1)):

find $\hat{u} : \omega \times (0,T) \longrightarrow \mathbb{R}$ such that

$$\begin{cases} \partial_t \hat{u} - \Delta \hat{u} &= \hat{f} \quad \text{in} \quad \omega \times (0,T), \\ \hat{u} &= g \quad \text{on} \quad \gamma \times (0,T), \\ \hat{u}(x,0) &= \hat{u}_0(x) \quad \text{in} \quad \omega, \end{cases} \tag{3.1}$$

where $\hat{f} \in L^2(\omega)$ and $g \in H^{\frac{1}{2}}(\gamma)$ are given. Here, $\hat{f} : \omega \times [0,T] \longrightarrow \mathbb{R}$, $g : \gamma \times [0,T] \longrightarrow \mathbb{R}$, $\hat{u}_0 : \omega \longrightarrow \mathbb{R}$ stand for source term, and boundary and initial condition, respectively.

To introduce the fictitious domain formulation for this problem, we first define Ω as an open bounded set such that $\omega \subseteq \Omega$ (Ω should be "simpler" than ω) and extensions $f : \Omega \longrightarrow \mathbb{R}, u_0 : \Omega \longrightarrow \mathbb{R}$ such that $f|_\omega = \hat{f}, u_0|_\omega = \hat{u}_0$. We now rewrite this problem in an equivalent way following the approach presented in [GG95]. Defining

$$a(u,v) = (\nabla u, \nabla v)_\Omega \quad \text{and} \quad b(\lambda, v) = \langle \lambda, v \rangle_\gamma,$$

then the following equivalent weak form for (3.1) can be written:

find $(u,\lambda) \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega)) \times L^\infty(0,T;H^{-\frac{1}{2}}(\gamma))$ such that $u(x,0) = u_0(x)$, and

$$\begin{cases} (\partial_t u, v)_\Omega + a(u,v) - b(\lambda, v) &= (f,v)_\Omega \\ b(\mu, u) &= b(\mu, g), \end{cases} \tag{3.2}$$

for all $v \in H_0^1(\Omega)$, $\mu \in H^{-\frac{1}{2}}(\gamma)$ and almost all $t \in (0,T)$. Problems (3.1) and (3.2) are linked by the fact that if (u,λ) satisfies (3.2), then $u|_\omega$ satisfies (3.1) and λ coincides with the jump of the normal derivative of u on γ (see [GPP94, GG95] for details).

We discretise the problem (3.2) with respect to the space and time variables (for more details, see Chapter 1).

We define the following finite element spaces:

$$V_h = \{v_h \in C^0(\bar{\Omega}) \cap H_0^1(\Omega) : v_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\},$$
$$\Lambda_h = \{\mu_h \in L^2(\gamma) : \mu_h|_e \in \mathbb{P}_0(e), \forall e \in \gamma_h\},$$
$$\Lambda_{\tilde{h}} = \{\mu_{\tilde{h}} \in L^2(\gamma) : \mu_{\tilde{h}}|_{\tilde{e}} \in \mathbb{P}_0(\tilde{e}), \forall \tilde{e} \in \gamma_{\tilde{h}}\},$$

such that $\Lambda_{\tilde{h}} \subseteq \Lambda_h$. We denote $W_h := V_h \times \Lambda_h$. Thanks to the hypothesis on $\mathcal{T}_h$ and $\gamma_{\tilde{h}}$, the pair $W_{\tilde{h}} := V_h \times \Lambda_{\tilde{h}}$ satisfies the following discrete inf-sup condition (see [GG95]): there exists $\beta > 0$, independent of h and $\tilde{h}$, such that

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{\langle \mu_{\tilde{h}}, v_h \rangle_\gamma}{|v_h|_{1,\Omega}} \geq \beta \|\mu_{\tilde{h}}\|_{-\frac{1}{2},\gamma} \qquad \forall \mu_{\tilde{h}} \in \Lambda_{\tilde{h}}. \tag{3.3}$$

On the other hand, the pair $V_h \times \Lambda_h$ is not inf-sup stable. For this case, it is proven in [BC12] that there exist two constants $C, \beta > 0$, independent of h, such that

$$\sup_{v_h \in V_h \setminus \{0\}} \frac{\langle \mu_h, v_h \rangle_\gamma}{|v_h|_{1,\Omega}} + C\Big(\sum_{e \in \gamma_h} |e| \|\mu_h\|^2_{-\frac{1}{2},\gamma}\Big)^{\frac{1}{2}} \geqslant \beta \|\mu_h\|_{-\frac{1}{2},\gamma} \qquad \forall \mu_h \in \Lambda_h. \tag{3.4}$$

A part of this work is devoted to approximating (3.2) using the space W_h. Since we only have the weak inf-sup condition (3.4), the problem needs stabilisation. Then we introduce a bilinear form $j : L^2(\gamma) \times L^2(\gamma) \longrightarrow \mathbb{R}$ defined as

$$j(\mu, \xi) = \sum_{\tilde{e} \in \gamma_{\tilde{h}}} |\tilde{e}|(\mu - \Pi_{\tilde{h}}\mu, \xi - \Pi_{\tilde{h}}\xi)_{\tilde{e}},$$

where $\Pi_{\tilde{h}} : L^2(\gamma) \longrightarrow \Lambda_{\tilde{h}}$ is defined as $(\Pi_{\tilde{h}}\xi)|_{\tilde{e}} = |\tilde{e}|^{-1}(\xi, 1)_{\tilde{e}}$ for each $\tilde{e} \in \gamma_{\tilde{h}}$. This term satisfies the following properties:

Symmetry:

$$j(\mu, \xi) = j(\xi, \mu) \quad \forall \mu, \xi \in L^2(\gamma); \tag{3.5}$$

Continuity: There exists $C > 0$ such that

$$|j(\mu, \xi)| \leq Ch\|\mu - \Pi_{\tilde{h}}\mu\|_{0,\gamma}\|\xi - \Pi_{\tilde{h}}\xi\|_{0,\gamma} \quad \forall \mu, \xi \in L^2(\gamma); \tag{3.6}$$

Weak consistency:

$$j(\mu,\xi) \leq \sum_{\tilde{e}\in\gamma_{\tilde{h}}} |\tilde{e}| \|\mu - \Pi_{\tilde{h}}\mu\|_{0,\tilde{e}} \|\xi - \Pi_{\tilde{h}}\xi\|_{0,\tilde{e}} \leq Ch^2 \|\mu\|_{\frac{1}{2},\mathcal{H}} \|\xi\|_{\frac{1}{2},\mathcal{H}} \quad \forall \mu,\xi \in \mathcal{H}. \tag{3.7}$$

We end this section by presenting the fully discrete methods to be analyzed in this chapter. The fully discrete problems read as follows:

1) Given a suitable approximation of $u_h^0 \in V_h$ of u_0, then for all $(v_h, \mu_{\tilde{h}}) \in W_{\tilde{h}}$ and $0 \leq n \leq N-1$, find $(u_h^{n+1}, \lambda_{\tilde{h}}^{n+1}) \in W_{\tilde{h}}$ such that

$$\left\{ \begin{array}{rcl} \frac{1}{\delta t}(u_h^{n+1} - u_h^n, v_h)_\Omega + a(u_h^{n+1}, v_h) - b(\lambda_{\tilde{h}}^{n+1}, v_h) & = & (f(t_{n+1}), v_h)_\Omega \\ b(\mu_{\tilde{h}}, u_h^{n+1}) & = & b(\mu_{\tilde{h}}, g) \end{array} \right. \tag{3.8}$$

in the inf-sup stable context, and

2) Given a suitable approximation of $u_h^0 \in V_h$ of u_0 for all $(v_h, \mu_h) \in W_h$ and $0 \leq n \leq N-1$, find $(u_h^{n+1}, \lambda_h^{n+1}) \in W_h$ such that

$$\left\{ \begin{array}{rcl} \frac{1}{\delta t}(u_h^{n+1} - u_h^n, v_h)_\Omega + a(u_h^{n+1}, v_h) - b(\lambda_h^{n+1}, v_h) & = & (f(t_{n+1}), v_h)_\Omega \\ b(\mu_h, u_h^{n+1}) + j(\lambda_h^{n+1}, \mu_h) & = & b(\mu_h, g) \end{array} \right. \tag{3.9}$$

for the stabilised case.

3.3. Stability analysis

3.3.1. The inf-sup stable case

In this section, we analyze the stability of problem (3.8). For the purpose of the stability and convergence analysis below, we introduce the Ritz-projection operator

$$S_{\tilde{h}} : W \longrightarrow W_{\tilde{h}},$$

where $W := H_0^1(\Omega) \times H^{-\frac{1}{2}}(\gamma)$. For each $(w,\xi) \in W$, the projection $S_{\tilde{h}}(w,\xi) = (P_{\tilde{h}}(w,\xi), R_{\tilde{h}}(w,\xi)) \in W_{\tilde{h}}$ is defined as the unique solution of

$$\left\{ \begin{array}{rcl} (\nabla P_{\tilde{h}}(w,\xi), \nabla v_h)_\Omega - \langle R_{\tilde{h}}(w,\xi), v_h\rangle_\gamma & = & (\nabla w, \nabla v_h)_\Omega - \langle \xi, v_h\rangle_\gamma \\ \langle \mu_{\tilde{h}}, P_{\tilde{h}}(w,\xi)\rangle_\gamma & = & \langle \mu_{\tilde{h}}, w\rangle_\gamma \end{array} \right. \tag{3.10}$$

for all $(v_h, \mu_{\tilde{h}}) \in W_{\tilde{h}}$. Problem (3.10) is well-posed thanks to the inf-sup condition (3.3). Moreover, defining the norm

$$\|(v,\mu)\|_W^2 := |v|_{1,\Omega}^2 + \|\mu\|_{-\frac{1}{2},\gamma}^2,$$

then the following stability and approximation results hold (see [GG95]):

$$\|P_{\tilde{h}}(w,\xi)\|_W^2 \leqslant C\big(|w|_{1,\Omega}^2 + \|\xi\|_{-\frac{1}{2},\gamma}^2\big), \tag{3.11}$$

and if $(w,\xi) \in H^2(\Omega) \times \mathcal{H}$, then there exist $C > 0$ independent of h such that

$$|w - P_{\tilde{h}}(w,\xi)|_{1,\Omega} + \|\xi - R_{\tilde{h}}(w,\xi)\|_{-\frac{1}{2},\gamma} \leq Ch(|w|_{2,\Omega} + \|\xi\|_{\mathcal{H}}). \tag{3.12}$$

The next result is a first step towards proving the stability of (3.8).

Lemma 3.1. *Let* $\|u_h^0\|_{1,\Omega} \leq C|u_0|_{1,\Omega}$ *and let* $\Big\{(u_h^n, \lambda_{\tilde{h}}^n)\Big\}_{n=1}^N$ *be the solution of the fully discrete problem* (3.8). *Then, there exists* $C > 0$ *for* $1 \leq n \leq N$, *independent of* $h, \tilde{h}$ *and* δt, *such that*

$$\|u_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1}\Big(\delta t|u_h^{m+1}|_{1,\Omega}^2 + \|u_h^{m+1} - u_h^m\|_{0,\Omega}^2\Big)$$
$$\leq C|u_0|_{1,\Omega}^2 + C\sum_{m=0}^{n-1}\delta t\Big(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2\Big), \tag{3.13}$$

and

$$\sum_{m=0}^{n-1}\delta t\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2$$
$$\leq C|u_0|_{1,\Omega}^2 + \frac{C}{\beta^2}\sum_{m=0}^{n-1}\delta t\Big(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2\Big), \tag{3.14}$$

where β *is the constant from* (3.3), *and* $Du_h^{m+1} = \frac{u_h^{m+1} - u_h^m}{\delta t}$.

Proof. We start by proving the estimate for $\lambda_{\tilde{h}}$. Applying the inf-sup condition (3.3) and the definition of the method (3.8), we get

$$\delta t\beta\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma} \leq \sup_{v_h \in V_h \setminus \{0\}} \frac{\delta t b(v_h, \lambda_{\tilde{h}}^{m+1})}{|v_h|_{1,\Omega}}$$
$$= \sup_{v_h \in V_h \setminus \{0\}} \frac{\delta t(f(t_{m+1}), v_h)_\Omega - \delta t a(u_h^{m+1}, v_h) - (u_h^{m+1} - u_h^m, v_h)_\Omega}{|v_h|_{1,\Omega}}$$
$$\leq C \sup_{v_h \in V_h \setminus \{0\}} \frac{\Big(\delta t^2\|f(t_{m+1})\|_{0,\Omega}^2 + \delta t^2 \mid u_h^{m+1} \mid_{1,\Omega}^2 + \|u_h^{m+1} - u_h^m\|_{0,\Omega}^2\Big)^{\frac{1}{2}}|v_h|_{1,\Omega}}{|v_h|_{1,\Omega}}.$$

Thus,

$$\begin{aligned}
&\beta\|\lambda_{\tilde h}^{m+1}\|_{-\frac12,\gamma}\\
&\le \frac{C}{\delta t}\Big(\delta t^2\|f(t_{m+1})\|_{0,\Omega}^2+\delta t^2|u_h^{m+1}|_{1,\Omega}^2+\|u_h^{m+1}-u_h^m\|_{0,\Omega}^2\Big)^{\frac12}\\
&= C\Big(\|f(t_{m+1})\|_{0,\Omega}^2+|u_h^{m+1}|_{1,\Omega}^2+\delta t^{-2}\|u_h^{m+1}-u_h^m\|_{0,\Omega}^2\Big)^{\frac12}.
\end{aligned}$$

Then squaring, multiplying by δt and taking summation over $0\le m\le n-1$, we get

$$\sum_{m=0}^{n-1}\delta t\|\lambda_{\tilde h}^{m+1}\|_{-\frac12,\gamma}^2\le\frac{C}{\beta^2}\sum_{m=0}^{n-1}\delta t\Big(\|f(t_{m+1})\|_{0,\Omega}^2+|u_h^{m+1}|_{1,\Omega}^2+\|Du_h^{m+1}\|_{0,\Omega}^2\Big).\quad(3.15)$$

Then, to prove (3.13), we need to bound $\sum\limits_{m=0}^{n-1}\delta t|u_h^{m+1}|_{1,\Omega}^2$, which we now do. Taking $v_h=u_h^{m+1}$ and $\mu_{\tilde h}=\lambda_{\tilde h}^{m+1}$ in (3.8), we get

$$\begin{aligned}
&(u_h^{m+1}-u_h^m,u_h^{m+1})_\Omega+\delta t a(u_h^{m+1},u_h^{m+1})-\delta t b(\lambda_{\tilde h}^{m+1},u_h^{m+1})+\delta t b(\lambda_{\tilde h}^{m+1},u_h^{m+1})\\
&=\delta t(f(t_{m+1}),u_h^{m+1})_\Omega+\delta t\langle\lambda_{\tilde h}^{m+1},g(t_{m+1})\rangle_\gamma.
\end{aligned}$$

Using the equality $(a-b)a=\frac12(a^2-b^2)+\frac12(a-b)^2$ and the Cauchy Schwarz inequality, this gives

$$\begin{aligned}
&\frac12\Big(\|u_h^{m+1}\|_{0,\Omega}^2-\|u_h^m\|_{0,\Omega}^2+\|u_h^{m+1}-u_h^m\|_{0,\Omega}^2\Big)+\delta t|u_h^{m+1}|_{1,\Omega}^2\\
&=\delta t(f(t_{m+1}),u_h^{m+1})_\Omega+\delta t\langle\lambda_{\tilde h}^{m+1},g(t_{m+1})\rangle_\gamma\\
&\le C\delta t\|f(t_{m+1})\|_{0,\Omega}|u_h^{m+1}|_{1,\Omega}+\delta t\|\lambda_{\tilde h}^{m+1}\|_{-\frac12,\gamma}\|g(t_{m+1})\|_{\frac12,\gamma}\\
&\le C\frac{\delta t}{2}\|f(t_{m+1})\|_{0,\Omega}^2+\frac{\delta t}{2}|u_h^{m+1}|_{1,\Omega}^2+\delta t\|\lambda_{\tilde h}^{m+1}\|_{-\frac12,\gamma}\|g(t_{m+1})\|_{\frac12,\gamma}.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\frac12\Big(\|u_h^{m+1}\|_{0,\Omega}^2-\|u_h^m\|_{0,\Omega}^2+\|u_h^{m+1}-u_h^m\|_{0,\Omega}^2\Big)+\frac{\delta t}{2}|u_h^{m+1}|_{1,\Omega}^2\\
&\qquad\qquad\le C\frac{\delta t}{2}\|f(t_{m+1})\|_{0,\Omega}^2+\delta t\|\lambda_{\tilde h}^{m+1}\|_{-\frac12,\gamma}\|g(t_{m+1})\|_{\frac12,\gamma}.
\end{aligned}$$

After summation over $0 \leq m \leq n-1$,

$$\begin{aligned}
&\frac{1}{2}(\|u_h^n\|_{0,\Omega}^2 - \|u_h^0\|_{0,\Omega}^2) + \frac{1}{2}\sum_{m=0}^{n-1}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + \frac{\delta t}{2}\sum_{m=0}^{n-1}|u_h^{m+1}|_{1,\Omega}^2 \\
&\qquad \leq C\frac{\delta t}{2}\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \delta t\sum_{m=0}^{n-1}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}\|g(t_{m+1})\|_{\frac{1}{2},\gamma} \\
&\qquad \leq C\frac{\delta t}{2}\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\alpha}\delta t\sum_{m=0}^{n-1}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2 + \frac{\alpha}{2}\delta t\sum_{m=0}^{n-1}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2,
\end{aligned} \tag{3.16}$$

where $\alpha > 0$ will be chosen later. So, inserting (3.15) in (3.16), we get

$$\begin{aligned}
&\frac{1}{2}\Big(\|u_h^n\|_{0,\Omega}^2 - \|u_h^0\|_{0,\Omega}^2\Big) + \frac{1}{2}\sum_{m=0}^{n-1}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + \frac{\delta t}{2}\sum_{m=0}^{n-1}|u_h^{m+1}|_{1,\Omega}^2 \\
&\quad \leq C\Big(\frac{\delta t}{2}\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{\delta t}{2\alpha}\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{\delta t}{2\alpha}\sum_{m=0}^{n-1}\|Du_h^{m+1}\|_{0,\Omega}^2 \\
&\qquad + \frac{\delta t}{2\alpha}\sum_{m=0}^{n-1}|u_h^{m+1}|_{1,\Omega}^2 + \frac{\alpha}{2}\delta t\sum_{m=0}^{n-1}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2\Big).
\end{aligned}$$

Choosing $\alpha = 2C$,

$$\begin{aligned}
&\|u_h^n\|_{0,\Omega}^2 - \|u_h^0\|_{0,\Omega}^2 + \sum_{m=0}^{n-1}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + \delta t\sum_{m=0}^{n-1}|u_h^{m+1}|_{1,\Omega}^2 \\
&\leq C\sum_{m=0}^{n-1}\delta t\Big(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2\Big).
\end{aligned}$$

We obtain the estimate (3.14) inserting (3.13) into (3.15) and applying the inequality $\|u_h^0\|_{1,\Omega} \leq C\|u_0\|_{1,\Omega}$. □

The last result is a partial stability result. In fact, the term in the time derivative, namely Du_h^{m+1}, still needs to be estimated. To present a uniform estimate of this term, the initial condition has to be appropriately chosen. Moreover, to accommodate the possibility of a time-dependent boundary condition g, we need to suppose extra regularity. We do that in the next result.

Theorem 3.2. *Let* $\Big\{(u_h^n, \lambda_{\tilde{h}}^n)\Big\}_{n=1}^N$ *be the solution of the fully discrete problem* (3.8) *where we consider* $u_h^0 = P_{\tilde{h}}(u_0, 0)$. *Let us also assume* $\partial_t g \in L^\infty(0,T;H^{\frac{1}{2}}(\gamma))$. *Then* *the following estimates hold for all* $1 \le n \le N$

$$|u_h^n|_{1,\Omega}^2 + \sum_{m=0}^{n-1} \Big(\delta t|u_h^{m+1}|_{1,\Omega}^2 + \|u_h^{m+1} - u_h^m\|_{0,\Omega}^2\Big)$$
$$\le C\Big(|u_0|_{1,\Omega}^2 + \sum_{m=0}^{n-1} \delta t\big(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|\partial_t g\|_{L^\infty(t_m,t_{m+1};H^{\frac{1}{2}}(\gamma))}^2\big)\Big) \tag{3.17}$$

and

$$\sum_{m=0}^{n-1} \delta t\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2$$
$$\le C|u_0|_{1,\Omega}^2 + \frac{C}{\beta^2}\sum_{m=0}^{n-1} \delta t\Big(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|\partial_t g\|_{L^\infty(t_m,t_{m+1};H^{\frac{1}{2}}(\gamma))}^2\Big). \tag{3.18}$$

Proof. From Lemma 3.1, it is enough to prove that

$$\sum_{m=0}^{n-1} \delta t\|Du_h^{m+1}\|_{0,\Omega}^2 + |u_h^n|_{1,\Omega}^2$$
$$\le C\Big(|u_0|_{1,\Omega}^2 + \sum_{m=0}^{n-1} \delta t\big(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \epsilon\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2\big)\Big), \tag{3.19}$$

where $\epsilon > 0$ will be chosen later.

For $0 \le m \le N-1$, by taking $v_h = Du_h^{m+1}$ and $\mu_{\tilde{h}} = 0$ in (3.8) and using the Cauchy Schwarz inequality, we have

$$\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) - b(\lambda_{\tilde{h}}^{m+1}, Du_h^{m+1})$$
$$= (f(t_{m+1}), Du_h^{m+1})_\Omega$$
$$\le \|f(t_{m+1})\|_{0,\Omega}\|Du_h^{m+1}\|_{0,\Omega}$$
$$\le \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2.$$

So,

$$\frac{1}{2} \parallel Du_h^{m+1} \parallel_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) - b(\lambda_{\tilde{h}}^{m+1}, Du_h^{m+1}) \le \frac{1}{2} \parallel f(t_{m+1}) \parallel_{0,\Omega}^2 . \tag{3.20}$$

On the other hand, for $1 \le m \le N-1$, testing (3.8) at the time levels $m+1$ and m with $v_h = 0$ and $\mu_{\tilde{h}} = \lambda_{\tilde{h}}^{m+1}$, we have

$$b(\lambda_{\tilde{h}}^{m+1}, u_h^{m+1}) = b(\lambda_{\tilde{h}}^{m+1}, g(t_{m+1})) \quad \text{and} \quad b(\lambda_{\tilde{h}}^{m+1}, u_h^m) = b(\lambda_{\tilde{h}}^{m+1}, g(t_m)). \tag{3.21}$$

Therefore, by subtracting these equalities and applying a Taylor expansion, we get

$$
\begin{aligned}
&b(\lambda_{\tilde{h}}^{m+1}, u_h^{m+1} - u_h^m) = b(\lambda_{\tilde{h}}^{m+1}, g(t_{m+1}) - g(t_m)) \\
&= \langle \lambda_{\tilde{h}}^{m+1}, g(t_{m+1}) - g(t_m) \rangle_\gamma \\
&\leq \|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma} \|g(t_{m+1}) - g(t_m)\|_{\frac{1}{2},\gamma} = \|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma} \|\delta t \partial_t g(\xi_m)\|_{\frac{1}{2},\gamma} \\
&\leq \frac{1}{2\epsilon} \delta t \|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2} \delta t \|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2,
\end{aligned}
$$

where $\xi_m \in (t_m, t_{m+1})$ and then

$$
b(\lambda_{\tilde{h}}^{m+1}, Du_h^{m+1}) \leq \frac{1}{2\epsilon} \|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2} \|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2 \quad \text{for} \quad 1 \leq m \leq N-1. \tag{3.22}
$$

Then it follows from (3.20) and (3.22) that

$$
\frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) \leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2. \tag{3.23}
$$

Next, using the symmetry and bilinearity of $a(.,.)$, we have

$$
Da(u_h^{m+1}, u_h^{m+1}) = \frac{a(u_h^{m+1}, u_h^{m+1}) - a(u_h^m, u_h^m)}{\delta t},
$$

and

$$
\begin{aligned}
&a(Du_h^{m+1}, Du_h^{m+1}) \\
&= a\left(\frac{u_h^{m+1} - u_h^m}{\delta t}, \frac{u_h^{m+1} - u_h^m}{\delta t} \right) = \frac{1}{\delta t} a(u_h^{m+1}, Du_h^{m+1}) - \frac{1}{\delta t} a(u_h^m, Du_h^{m+1}).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&Da(u_h^{m+1}, u_h^{m+1}) + \delta t a(Du_h^{m+1}, Du_h^{m+1}) \\
&= \frac{a(u_h^{m+1}, u_h^{m+1}) - a(u_h^m, u_h^{m+1}) + a(u_h^{m+1}, u_h^m) - a(u_h^m, u_h^m)}{\delta t} \\
&+ a(u_h^{m+1}, Du_h^{m+1}) - a(u_h^m, Du_h^{m+1}) \\
&= a(Du_h^{m+1}, u_h^{m+1}) + a(Du_h^{m+1}, u_h^m) + a(u_h^{m+1}, Du_h^{m+1}) - a(u_h^m, Du_h^{m+1}) \\
&= 2a(u_h^{m+1}, Du_h^{m+1}).
\end{aligned}
$$

Therefore,

$$
a(u_h^{m+1}, Du_h^{m+1}) = \frac{1}{2} Da(u_h^{m+1}, u_h^{m+1}) + \frac{\delta t}{2} a(Du_h^{m+1}, Du_h^{m+1}) \geq \frac{1}{2} Da(u_h^{m+1}, u_h^{m+1}).
$$

Hence, (3.23) becomes

$$\frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2 + \frac{1}{2}Da(u_h^{m+1}, u_h^{m+1}) \leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2,$$

for $1 \leq m \leq N-1$. After multiplication by δt and summation over $1 \leq m \leq n-1$, it follows that

$$\sum_{m=1}^{n-1} \delta t\|Du_h^{m+1}\|_{0,\Omega}^2 + \sum_{m=1}^{n-1} \delta t Da(u_h^{m+1}, u_h^{m+1})$$
$$\leq \sum_{m=1}^{n-1} \delta t\big(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2\big),$$

and since

$$\sum_{m=1}^{n-1} \delta t Da(u_h^{m+1}, u_h^{m+1}) = a(u_h^n, u_h^n) - a(u_h^1, u_h^1),$$

we get

$$\sum_{m=1}^{n-1} \delta t\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^n, u_h^n)$$
$$\leq a(u_h^1, u_h^1) + \sum_{m=1}^{n-1} \delta t\big(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2\big).$$

Thus,

$$\sum_{m=1}^{n-1} \delta t\|Du_h^{m+1}\|_{0,\Omega}^2 + |u_h^n|_{1,\Omega}^2$$
$$\leq |u_h^1|_{1,\Omega}^2 + \sum_{m=1}^{n-1} \delta t\big(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2\epsilon}\|\partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2\big). \quad (3.24)$$

Since the initial approximation of u is given in terms of the Ritz-projection, $u_h^0 = P_{\tilde{h}}(u_0, 0)$ with $u_0 \in H_0^1(\Omega)$, by setting $\lambda_{\tilde{h}}^0 = R_{\tilde{h}}(u_0, 0)$ it follows that (3.22) also holds for $m = 0$. This is

$$\begin{aligned}
\delta t b(\lambda_{\tilde{h}}^1, Du_h^1) &= \delta t \left\langle \lambda_{\tilde{h}}^1, \frac{u_h^1 - u_h^0}{\delta t} \right\rangle_\gamma \\
&= \langle \lambda_{\tilde{h}}^1, u_h^1 \rangle_\gamma - \langle \lambda_{\tilde{h}}^1, u_h^0 \rangle_\gamma = \langle \lambda_{\tilde{h}}^1, g(t_1) \rangle_\gamma - \langle \lambda_{\tilde{h}}^1, g(t_0) \rangle_\gamma \\
&= \langle \lambda_{\tilde{h}}^1, g(t_1) - g(t_0) \rangle_\gamma \\
&\leq \|\lambda_{\tilde{h}}^1\|_{-\frac{1}{2},\gamma} \|g(t_1) - g(t_0)\|_{\frac{1}{2},\gamma} \\
&\leq \delta t \|\lambda_{\tilde{h}}^1\|_{-\frac{1}{2},\gamma} \|\partial_t g(\xi_0)\|_{\frac{1}{2},\gamma} \\
&\leq \frac{1}{2\epsilon} \delta t \|\partial_t g(\xi_0)\|^2_{\frac{1}{2},\gamma} + \frac{\epsilon}{2} \delta t \|\lambda_{\tilde{h}}^1\|^2_{-\frac{1}{2},\gamma},
\end{aligned} \tag{3.25}$$

where $\xi_0 \in (t_0, t_1)$. Now, testing for $m = 0$ in (3.8) with $v_h = Du_h^1$ and $\mu_{\tilde{h}} = 0$, gives

$$(Du_h^1, Du_h^1)_\Omega + a(u_h^1, Du_h^1) - b(\lambda_{\tilde{h}}^1, Du_h^1) = (f(t_1), Du_h^1)_\Omega,$$

and then

$$\begin{aligned}
&2\delta t \|Du_h^1\|^2_{0,\Omega} + 2\frac{\delta t}{2} Da(u_h^1, u_h^1) + 2\frac{\delta t^2}{2} a(Du_h^1, Du_h^1) - 2\delta t b(\lambda_h^1, Du_h^1) \\
&= 2\delta t (f(t_1), Du_h^1)_\Omega \\
&\leq 2\delta t \|f(t_1)\|_{0,\Omega} \|Du_h^1\|_{0,\Omega} \\
&\leq \delta t \|f(t_1)\|^2_{0,\Omega} + \delta t \|Du_h^1\|^2_{0,\Omega}.
\end{aligned}$$

Hence, applying (3.25) to this last inequality we arrive at

$$\delta t \|Du_h^1\|^2_{0,\Omega} + a(u_h^1, u_h^1) - a(u_h^0, u_h^0) \leq \delta t \|f(t_1)\|^2_{0,\Omega} + \frac{1}{\epsilon} \delta t \|\partial_t g(\xi_0)\|^2_{\frac{1}{2},\gamma} + \epsilon \delta t \|\lambda_{\tilde{h}}^1\|^2_{-\frac{1}{2},\gamma}, \tag{3.26}$$

which, applying the stability of the Ritz-projection (3.11), gives

$$\begin{aligned}
\delta t \|Du_h^1\|^2_{0,\Omega} + |u_h^1|^2_{1,\Omega} &\leq |u_h^0|^2_{1,\Omega} + \delta t \big(\|f(t_1)\|^2_{0,\Omega} + \frac{1}{\epsilon} \|\partial_t g(\xi_0)\|^2_{\frac{1}{2},\gamma} + \epsilon \|\lambda_{\tilde{h}}^1\|^2_{-\frac{1}{2},\gamma} \big) \\
&\leq C|u_0|^2_{1,\Omega} + \delta t \big(\|f(t_1)\|^2_{0,\Omega} + \frac{1}{\epsilon} \|\partial_t g(\xi_0)\|^2_{\frac{1}{2},\gamma} + \epsilon \|\lambda_{\tilde{h}}^1\|^2_{-\frac{1}{2},\gamma} \big).
\end{aligned} \tag{3.27}$$

Then, using (3.24) and (3.27),

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|Du_h^{m+1}\|^2_{0,\Omega} + |u_h^n|^2_{1,\Omega} \\
&\leq C|u_0|^2_{1,\Omega} + \sum_{m=0}^{n-1} \delta t \big(\|f(t_{m+1})\|^2_{0,\Omega} + \frac{1}{\epsilon} \|\partial_t g(\xi_m)\|^2_{\frac{1}{2},\gamma} + \epsilon \|\lambda_{\tilde{h}}^{m+1}\|^2_{-\frac{1}{2},\gamma} \big),
\end{aligned}$$

which proves (3.19). Taking $\epsilon = \frac{1}{2C}$, estimate (3.17) is obtained by inserting (3.19) into (3.13) and the estimate (3.18) is obtained by inserting (3.19) into (3.14). □

Remark 3.3. *If we do not start with the Ritz-projection of* u_0*, then a weaker stability result can be obtained. In fact, if* u_0 *is given in terms of a general interpolant,* $u_h^0 = i_h(u_0)$*, using the approximation properties of* i_h *and supposing that* $\delta t \geq Ch^2$*, we get, instead of* (3.25)*,*

$$\begin{aligned}
\delta t b(\lambda_{\tilde{h}}^1, Du_h^1) &= \delta t \langle \lambda_{\tilde{h}}^1, \frac{u_h^1 - u_h^0}{\delta t} \rangle_\gamma \\
&= \langle \lambda_{\tilde{h}}^1, u_h^1 \rangle_\gamma - \langle \lambda_{\tilde{h}}^1, u_h^0 \rangle_\gamma \\
&= \langle \lambda_h^1, g(t_1) - g(t_0) \rangle_\gamma + \langle \lambda_h^1, g(t_0) - i_h(u_0) \rangle_\gamma \\
&\leq \frac{1}{2\epsilon} \delta t \|\partial_t g\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2} \delta t \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2 + Ch|u_0|_{2,\Omega}^2 \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2 \\
&\leq \frac{1}{2\epsilon} \delta t \|\partial_t g\|_{\frac{1}{2},\gamma}^2 + \frac{\epsilon}{2} \delta t \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2 + C\frac{1}{2\epsilon}|u_0|_{2,\Omega}^2 + \frac{\epsilon}{2} h^2 \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2 \\
&\leq \frac{1}{2\epsilon} \delta t \|\partial_t g\|_{\frac{1}{2},\gamma}^2 + C|u_0|_{2,\Omega}^2 + C\epsilon \delta t \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2.
\end{aligned}$$

Then, the same results as in Theorem 3.2 can be obtained supposing $\delta t \geq Ch^2$*.*

3.3.2. The stabilised method

In this section, we analize the stability of problem (3.9). The first step is to modify the definition of the Ritz-projection $S_h : W \longrightarrow W_h$ to accommodate it to the stabilised method. For each $(w, \xi) \in W$, the projection $S_h(w, \xi) = (P_h(w, \xi), R_h(w, \xi)) \in W_h$ is now defined as the unique solution of

$$\left\{ \begin{array}{rcl} (\nabla P_h(w, \xi), \nabla v_h)_\Omega - \langle R_h(w, \xi), v_h \rangle_\gamma &=& (\nabla w, \nabla v_h)_\Omega - \langle \xi, v_h \rangle_\gamma \\ \langle \mu_h, P_h(w, \xi) \rangle_\gamma + j(R_h(w, \xi), \mu_h) &=& \langle \mu_h, w \rangle_\gamma, \end{array} \right. \tag{3.28}$$

for all $(v_h, \mu_h) \in W_h$. The well-posedness of (3.28) has been studied in [BC12]. Moreover, defining the norm

$$|||(v_h, \mu_h)|||_h^2 = |v_h|_{1,\Omega}^2 + j(\mu_h, \mu_h),$$

then the following stability and approximation results hold (see [BC12]):

$$|||(P_h(w, \xi), R_h(w, \xi))|||_h^2 \leq C \left(|w|_{1,\Omega}^2 + \|\xi\|_{-\frac{1}{2},\gamma}^2 \right), \tag{3.29}$$

and if $(w, \xi) \in H^2(\Omega) \times \mathcal{H}$, then there exists $C > 0$ independent of h such that

$$\|\xi - R_h(w, \xi)\|_{-\frac{1}{2},\gamma} + |||(w - P_h(w, \xi), \xi - R_h(w, \xi))|||_h \leq Ch \left(|w|_{2,\Omega} + \|\xi\|_{\mathcal{H}} \right).$$

As it was done in the last section, we give the following analogues of Lemma 3.1 and Theorema 3.2.

Lemma 3.4. *Let $\|u_h^0\|_{1,\Omega} \le C|u_0|_{1,\Omega}$ and let $\{(u_h^n, \lambda_h^n)\}_{n=1}^N$ be the solution of the fully discrete problem* (3.9)*. Then there exists $C > 0$ independent of h and δt such that the following estimate holds for $1 \le n \le N$:*

$$\begin{aligned} &\|u_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2 \\ &\le C|u_0|_{1,\Omega}^2 + C \sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2 \right), \end{aligned} \tag{3.30}$$

and

$$\begin{aligned} &\sum_{m=0}^{n-1} \delta t \|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2 \\ &\le \frac{C}{\beta^2} \left(|u_0|_{1,\Omega}^2 + \sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2 \right) \right), \end{aligned} \tag{3.31}$$

where β is the constant from (3.4)*.*

Proof. Using (3.4) and (3.9) with $\mu_h = 0$, the Cauchy Schwarz and Poincaré inequalities, we get

$$\begin{aligned} \beta\|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma} &\le \sup_{v_h \in V_h \setminus \{0\}} \frac{|b(\lambda_h^{m+1}, v_h)|}{|v_h|_{1,\Omega}} + Cj(\lambda_h^{m+1}, \lambda_h^{m+1})^{\frac{1}{2}} \\ &= \sup_{v_h \in V_h \setminus \{0\}} \frac{(f(t_{m+1}), v_h)_\Omega - a(u_h^{m+1}, v_h) - (Du_h^{m+1}, v_h)_\Omega}{|v_h|_{1,\Omega}} + Cj(\lambda_h^{m+1}, \lambda_h^{m+1})^{\frac{1}{2}} \\ &\le C\|f(t_{m+1})\|_{0,\Omega} + |u_h^{m+1}|_{1,\Omega} + \frac{C}{\delta t}\|u_h^{m+1} - u_h^m\|_{0,\Omega} + Cj(\lambda_h^{m+1}, \lambda_h^{m+1})^{\frac{1}{2}}. \end{aligned}$$

Squaring, multiplying by δt and adding over $0 \le m \le n-1$,

$$\begin{aligned} &\delta t \sum_{m=0}^{n-1} \|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2 \\ &\le \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + |u_h^{m+1}|_{1,\Omega}^2 + \left\| \frac{u_h^{m+1} - u_h^m}{\delta t} \right\|_{0,\Omega}^2 + j(\lambda_h^{m+1}, \lambda_h^{m+1}) \right) \\ &= \frac{C}{\beta^2} \sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2 + |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2 \right). \end{aligned} \tag{3.32}$$

To prove (3.30), we take $v_h = u_h^{m+1}$ and $\mu_h = \lambda_h^{m+1}$ in (3.9),

$$\left(\frac{u_h^{m+1} - u_h^m}{\delta t}, u_h^{m+1}\right)_\Omega + a(u_h^{m+1}, u_h^{m+1}) - b(\lambda_h^{m+1}, u_h^{m+1}) + b(\lambda_h^{m+1}, u_h^{m+1})$$
$$+ j(\lambda_h^{m+1}, \lambda_h^{m+1}) = (f(t_{m+1}), u_h^{m+1})_\Omega + b(\lambda_h^{m+1}, g(t_{m+1})).$$

Using the equality $(a-b)a = \frac{1}{2}(a^2 - b^2) + \frac{1}{2}(a-b)^2$ and the Cauchy Schwarz inequality in the above expression, we get

$$\frac{1}{2\delta t}\Big(\|u_h^{m+1}\|_{0,\Omega}^2 - \|u_h^m\|_{0,\Omega}^2\Big) + \frac{1}{2\delta t}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$$
$$= (f(t_{m+1}), u_h^{m+1}) + b(\lambda_h^{m+1}, g(t_{m+1}))$$
$$\leq C\|f(t_{m+1})\|_{0,\Omega}|u_h^{m+1}|_{1,\Omega} + \|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}.$$

Now, after multiplying by $2\delta t$ and adding over $0 \leq m \leq n-1$, we obtain

$$\|u_h^n\|_{0,\Omega}^2 - \|u_h^0\|_{0,\Omega}^2 + \sum_{m=0}^{n-1}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + 2\delta t\sum_{m=0}^{n-1}|||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$$
$$\leq C\delta t\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}|u_h^{m+1}|_{1,\Omega} + 2\delta t\sum_{m=0}^{n-1}\|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}$$
$$\leq C\delta t\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \delta t\sum_{m=0}^{n-1}|u_h^{m+1}|_{1,\Omega}^2 + \epsilon\delta t\sum_{m=0}^{n-1}\|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2$$
$$+ \frac{1}{\epsilon}\delta t\sum_{m=0}^{n-1}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2,$$

where $\epsilon > 0$ will be chosen later. Then as $|u_h^{m+1}|_{1,\Omega}^2 \leq |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$, using (3.32), we get

$$\|u_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + \delta t\sum_{m=0}^{n-1}|||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$$
$$\leq \|u_h^0\|_{0,\Omega}^2 + C\delta t\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \epsilon\delta t\sum_{m=0}^{n-1}\|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2 + \frac{1}{\epsilon}\delta t\sum_{m=0}^{n-1}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2$$
$$\leq \|u_h^0\|_{0,\Omega}^2 + C\delta t\sum_{m=0}^{n-1}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\epsilon}\delta t\sum_{m=0}^{n-1}\|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2$$
$$+ \epsilon\frac{C}{\beta^2}\delta t\sum_{m=0}^{n-1}\Big(|||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2 + \|Du_h^{m+1}\|_{0,\Omega}^2 + \|f(t_{m+1})\|_{0,\Omega}^2\Big).$$

Then, taking $\epsilon = \frac{\beta^2}{2C}$, we obtain

$$\|u_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \|u_h^{m+1} - u_h^m\|_{0,\Omega}^2 + \delta t \sum_{m=0}^{n-1} |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$$
$$\leq C\|u_h^0\|_{0,\Omega}^2 + C\sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|Du_h^{m+1}\|_{0,\Omega}^2 \right),$$

which proves (3.30) applying the inequality $\|u_h^0\|_{1,\Omega} \leq C\|u_0\|_{1,\Omega}$. We obtain the estimate (3.31) inserting (3.30) into (3.32). □

The last result is a partial stability result. In fact, the term in the time derivative, namely Du_h^{m+1}, still needs to be estimated. We do that in the next result.

Theorem 3.5. *Let $\{(u_h^n, \lambda_h^n)\}_{n=1}^N$ be the solution of the fully discrete problem* (3.9) *where we consider $u_h^0 = P_h(u_0, 0)$. Let us also assume $\partial_t g \in L^\infty(0,T; H^{\frac{1}{2}}(\gamma))$. Then for all $1 \leq n \leq N$ the following estimate holds*

$$\|u_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |||(u_h^{m+1}, \lambda_h^{m+1})|||_h^2$$
$$\leq C|u_0|_{1,\Omega}^2 + C\sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|\partial_t g\|_{L^\infty(t_m,t_{m+1};H^{\frac{1}{2}}(\gamma))}^2 \right), \tag{3.33}$$

and

$$\sum_{m=0}^{n-1} \delta t \|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2$$
$$\leq \; C|u_0|_{1,\Omega}^2 + C\sum_{m=0}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \|g(t_{m+1})\|_{\frac{1}{2},\gamma}^2 + \|\partial_t g\|_{L^\infty(t_n,t_{n+1};H^{\frac{1}{2}}(\gamma))}^2 \right). \tag{3.34}$$

Proof. Based on the previous lemma, it is enough to prove that

$$\sum_{m=0}^{n-1} \delta t \|Du_h^{m+1}\|_{0,\Omega}^2 + |||(u_h^n, \lambda_h^n)|||_h^2$$
$$\leq C\Big(|u_0|_{1,\Omega}^2 + \sum_{m=0}^{n-1} \delta t \big(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{\delta t}{\epsilon} \|\partial_t g\|_{\frac{1}{2},\gamma}^2 + \epsilon \|\lambda_h^{m+1}\|_{-\frac{1}{2},\gamma}^2 \big)\Big). \tag{3.35}$$

where $\epsilon > 0$ will be chosen later.

For $0 \leq m \leq N-1$, taking $v_h = Du_h^{m+1}$ and $\mu_h = 0$ in (3.9) and using the Cauchy

Schwarz inequality, we have

$$\begin{aligned}
&\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) - b(\lambda_h^{m+1}, Du_h^{m+1}) \\
&= (f(t_{m+1}), Du_h^{m+1})_\Omega \\
&\leq \|f(t_{m+1})\|_{0,\Omega}\|Du_h^{m+1}\|_{0,\Omega} \\
&\leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2,
\end{aligned}$$

which implies

$$\frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) - b(\lambda_h^{m+1}, Du_h^{m+1}) \leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2. \tag{3.36}$$

On the other hand, for $1 \leq m \leq N-1$, testing (3.9) at the time levels $m+1$ and m with $v_h = 0$ and $\mu_h = \lambda_h^{m+1}$, we have

$$\begin{aligned}
b(\lambda_h^{m+1}, g(t_{m+1})) - b(\lambda_h^{m+1}, u_h^{m+1}) &= j(\lambda_h^{m+1}, \lambda_h^{m+1}), \\
b(\lambda_h^{m+1}, g(t_m)) - b(\lambda_h^{m+1}, u_h^{m}) &= j(\lambda_h^{m}, \lambda_h^{m+1}).
\end{aligned} \tag{3.37}$$

Therefore, by subtracting these equalities and using the bilinearity of $j(.,.)$, we obtain

$$-b(\lambda_h^{m+1}, Du_h^{m+1}) = j(D\lambda_h^{m+1}, \lambda_h^{m+1}) - \frac{1}{\delta t}b(\lambda_h^{m+1}, g(t_{m+1}) - g(t_m)), \tag{3.38}$$

for $1 \leq m \leq N-1$. It then follows from (3.36) and a Taylor expansion that

$$\begin{aligned}
&\frac{1}{2}\|Du_h^{m+1}\|_{0,\Omega}^2 + a(u_h^{m+1}, Du_h^{m+1}) + j(D\lambda_h^{m+1}, \lambda_h^{m+1}) \\
&\leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\delta t}b(\lambda_h^{m+1}, g(t_{m+1}) - g(t_m)) \\
&\leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\delta t}\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}\|\delta t \partial_t g(\xi_m)\|_{\frac{1}{2},\gamma} \\
&\leq \frac{1}{2}\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\epsilon}\|\delta t \partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \epsilon\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2,
\end{aligned} \tag{3.39}$$

where $\xi_m \in (t_m, t_{m+1})$.

On the other hand, using the symmetry and bilinearity of $a(.,.)$ and $j(.,.)$, we have

$$\begin{aligned}
a(u_h^{m+1}, Du_h^{m+1}) &= \frac{1}{2}Da(u_h^{m+1}, u_h^{m+1}) + \frac{\delta t}{2}a(Du_h^{m+1}, Du_h^{m+1}), \\
j(\lambda_h^{m+1}, D\lambda_h^{m+1}) &= \frac{1}{2}Dj(\lambda_h^{m+1}, \lambda_h^{m+1}) + \frac{\delta t}{2}j(D\lambda_h^{m+1}, D\lambda_h^{m+1}).
\end{aligned}$$

Hence, (3.39) becomes

$$\begin{aligned}
&\|Du_h^{m+1}\|_{0,\Omega}^2 + D(a(u_h^{m+1}, u_h^{m+1}) + j(\lambda_h^{m+1}, \lambda_h^{m+1})) \\
&\leq \|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\epsilon}\|\delta t \partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \epsilon\|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2,
\end{aligned}$$

for $1 \leq m \leq N-1$. After multiplication by δt and summation over $1 \leq m \leq n-1$, it follows that

$$\begin{aligned}
&\sum_{m=1}^{n-1} \delta t \|Du_h^{m+1}\|_{0,\Omega}^2 + \|(u_h^n, \lambda_h^n)\|_h^2 \\
&\leq \left|\left|\left|(u_h^1, \lambda_h^1)\right|\right|\right|_h^2 + \sum_{m=1}^{n-1} \delta t \left(\|f(t_{m+1})\|_{0,\Omega}^2 + \frac{1}{\epsilon} \|\delta t \partial_t g(\xi_m)\|_{\frac{1}{2},\gamma}^2 + \epsilon \|\lambda_{\tilde{h}}^{m+1}\|_{-\frac{1}{2},\gamma}^2 \right).
\end{aligned} \tag{3.40}$$

Since the initial approximation of u is given in terms of the Ritz-projection, $u_h^0 = P_h(u_0, 0)$, by setting $\lambda_h^0 = R_h(u_0, 0)$ it follows that (3.38) also holds for $m = 0$. This is

$$b(\lambda_h^1, Du_h^1) = \frac{1}{\delta t} b(\lambda_h^1, g(t_1) - g(t_0)) - j(D\lambda_h^1, \lambda_h^1). \tag{3.41}$$

Taking $v_h = Du_h^1$, $\mu_h = 0$ in (3.9) and multiplying by $2\delta t$, we get

$$\delta t \|Du_h^1\|_{0,\Omega}^2 + a(u_h^1, u_h^1) - a(u_h^0, u_h^0) \leq \delta t \|f(t_1)\|_{0,\Omega}^2 + b(\lambda_h^1, Du_h^1),$$

which implies

$$\delta t \|Du_h^1\|_{0,\Omega}^2 + \left|\left|\left|u_h^1, \lambda_h^1\right|\right|\right|_h^2 \leq \left|\left|\left|u_h^0, \lambda_h^0\right|\right|\right|_h^2 + \delta t \|f(t_1)\|_{0,\Omega}^2 + b(\lambda_h^1, Du_h^1).$$

Applying (3.41) to this last inequality, we obtain

$$\begin{aligned}
&\delta t \|Du_h^1\|_{0,\Omega}^2 + \left|\left|\left|u_h^1, \lambda_h^1\right|\right|\right|_h^2 \\
&\leq \left|\left|\left|(u_h^0, \lambda_h^0)\right|\right|\right|_h^2 + \delta t \|f(t_1)\|_{0,\Omega}^2 + b(\lambda_h^1, g(t_1) - g(t_0)) \\
&\leq \left|\left|\left|(u_h^0, \lambda_h^0)\right|\right|\right|_h^2 + \delta t \|f(t_1)\|_{0,\Omega}^2 + \|\lambda_h^1\|_{-\frac{1}{2},\gamma} \|\delta t \partial_t g(\xi_0)\|_{\frac{1}{2},\gamma} \\
&\leq \left|\left|\left|(u_h^0, \lambda_h^0)\right|\right|\right|_h^2 + \delta t \|f(t_1)\|_{0,\Omega}^2 + \frac{\epsilon}{2} \delta t \|\lambda_h^1\|_{-\frac{1}{2},\gamma}^2 + \frac{\delta t}{2\epsilon} \|\partial_t g(\xi_0)\|_{\frac{1}{2},\gamma}^2
\end{aligned} \tag{3.42}$$

where $\xi_0 \in (t_0, t_1)$. Then we get the estimate (3.35) adding (3.42) to (3.40) and using the stability of the Ritz-projection (3.29), $\left|\left|\left|(u_h^0, \lambda_h^0)\right|\right|\right|_h^2 \leq C|u_0|_{1,\Omega}^2$.

Finally, we obtain the estimate (3.33) taking $\epsilon = \frac{1}{2C}$ in (3.35) and the estimate (3.34) is obtained in an analogous way. □

3.4. Convergence analysis

In this section we prove optimal order error estimates for the fully discrete methods (3.8) and (3.9).

3.4.1. The inf-sup stable case

We start by presenting the following result on consistency. Its proof is direct verification.

Lemma 3.6. *Let* (u, λ) *be the solution of* (3.2) *and let* $\{(u_h^n, \lambda_{\tilde{h}}^n)\}_{0 \leq n \leq N}$ *be the solution of* (3.8). *Assume that* $u \in H^1(0,T; L^2(\Omega)) \cap L^2(0,T; H_0^1(\Omega))$ *and let* $\lambda \in L^\infty(0,T; H^{-\frac{1}{2}}(\gamma))$. *Then, for* $0 \leq n \leq N-1$, *there holds*

$$
\begin{aligned}
&(Du(t_{n+1}) - Du_h^{n+1}, v_h)_\Omega + a(u(t_{n+1}) - u_h^{n+1}, v_h) - b(\lambda(t_{n+1}) - \lambda_{\tilde{h}}^{n+1}, v_h) \\
&+ b(\mu_{\tilde{h}}, u(t_{n+1}) - u_h^{n+1}) = (Du(t_{n+1}) - \partial_t u(t_{n+1}), v_h)_\Omega,
\end{aligned}
$$

for all $(v_h, \mu_{\tilde{h}}) \in W_{\tilde{h}}$.

We prove the following result which states an optimal error estimate for u, and an estimate for λ which is optimal, up to a term that will be treated later.

Lemma 3.7. *Let us assume that* $u \in H^1(0,T; H^2(\Omega)) \cap H^2(0,T; L^2(\Omega)) \cap C^0(0,T; H^2(\Omega))$, $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$ *and* $\lambda \in H^1(0,T; \mathcal{H})$, *and set* $u_h^0 \in V_h$ *as* $u_h^0 = C_h(u_0)$ *where* $C_h : H_0^1(\Omega) \longrightarrow V_h$ *stands for the Clément interpolation operator. Then, the following estimates hold for* $1 \leq n \leq N$:

$$
\begin{aligned}
&\|u_h^n - u(t_n)\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |u_h^{m+1} - u(t_{m+1})|_{1,\Omega}^2 \\
&\leq Ch^2\Big(\|u_0\|_{2,\Omega}^2 + \|\lambda(0)\|_{\mathcal{H}}^2\Big) + Ch^2\delta t^2\Big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 + \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2\Big) \\
&+ C\Big(\delta t^2 \|\partial_{tt} u\|_{L^2(0,t_n;L^2(\Omega))}^2 + h^2\|\lambda\|_{C^0(t_1,t_n;\mathcal{H})}^2 + h^2\|u\|_{C^0(t_1,t_n;H^2(\Omega))}^2\Big),
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\lambda_{\tilde{h}}^{m+1} - \lambda(t_{m+1})\|_{-\frac{1}{2},\gamma}^2 \\
&\leq \sum_{m=0}^{n-1} \delta t \bigg(Ch^2\Big(\|u_0\|_{2,\Omega}^2 + \|\lambda(0)\|_{\mathcal{H}}^2\Big) + Ch^2\delta t^2\Big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 + \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2\Big) \\
&+ C\Big(\delta t^2 \|\partial_{tt} u\|_{L^2(0,t_n;L^2(\Omega))}^2 + h^2\|\lambda\|_{C^0(t_1,t_n;\mathcal{H})}^2 + h^2\|u\|_{C^0(t_1,t_n;H^2(\Omega))}^2\Big) \\
&+ \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|_{0,\Omega}^2 \bigg),
\end{aligned}
$$

with $C > 0$ *a positive constant independent of* h *and* δt.

Proof. As usual, for $s = 0, ..., N$, we decompose the error into interpolation and discrete errors as follows

$$
\begin{aligned}
u(t_s) - u_h^s &= \underbrace{u(t_s) - P_{\tilde{h}}(u(t_s), \lambda(t_s))}_{\theta_\pi^s} + \underbrace{P_{\tilde{h}}(u(t_s), \lambda(t_s)) - u_h^s}_{\theta_h^s} \\
&=: \theta_\pi^s + \theta_h^s, \qquad (3.43)
\end{aligned}
$$

$$
\begin{aligned}
\lambda(t_s) - \lambda_{\tilde{h}}^s &= \underbrace{\lambda(t_s) - R_{\tilde{h}}(u(t_s), \lambda(t_s))}_{y_\pi^s} + \underbrace{R_{\tilde{h}}(u(t_s), \lambda(t_s)) - \lambda_{\tilde{h}}^s}_{y_h^s} \\
&=: y_\pi^s + y_h^s, \qquad (3.44)
\end{aligned}
$$

where $(P_{\tilde{h}}, R_{\tilde{h}})$ is defined in (3.10). To simplify the notation, we write

$$
\begin{aligned}
\tilde{u}^s &= P_{\tilde{h}}(u(t_s), \lambda(t_s)), \\
\tilde{\lambda}^s &= R_{\tilde{h}}(u(t_s), \lambda(t_s)).
\end{aligned}
$$

Let $n \in 1, ..., N$, and $m \in 0, ..., n-1$. First, using (3.12), it follows that

$$
|\theta_\pi^{m+1}|_{1,\Omega} + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma} \le Ch \left(\|u(t_{m+1})\|_{2,\Omega} + \|\lambda(t_{m+1})\|_{\mathcal{H}} \right).
$$

In order to estimate θ_h^{m+1}, we first note that a quick calculation gives

$$
(Du_h^{m+1}, u_h^{m+1})_\Omega = \frac{1}{2} D\|u_h^{m+1}\|_{0,\Omega}^2 + \frac{1}{2\delta t}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2.
$$

Then, using the definition of the bilinear form $a(.,.)$, we get

$$
\begin{aligned}
&\frac{1}{2} D\|\theta_h^{m+1}\|_{0,\Omega}^2 + |\theta_h^{m+1}|_{1,\Omega}^2 \le (D\theta_h^{m+1}, \theta_h^{m+1})_\Omega + |\theta_h^{m+1}|_{1,\Omega}^2 \\
&= \underbrace{(D\theta_h^{m+1}, \theta_h^{m+1})_\Omega + a(\theta_h^{m+1}, \theta_h^{m+1}) + b(y_h^{m+1}, \theta_h^{m+1}) - b(y_h^{m+1}, \theta_h^{m+1})}_{T_1^{m+1}}. \qquad (3.45)
\end{aligned}
$$

In addition, using (3.43)-(3.44), we have

$$
\begin{aligned}
T_1^{m+1} &= -(D\theta_\pi^{m+1}, \theta_h^{m+1})_\Omega - a(\theta_\pi^{m+1}, \theta_h^{m+1}) + a(u(t_{m+1}) - u_h^{m+1}, \theta_h^{m+1}) \\
&+ (Du(t_{m+1}) - Du_h^{m+1}, \theta_h^{m+1})_\Omega - b(\tilde{\lambda}^{m+1} - \lambda(t_{m+1}), \theta_h^{m+1}) + b(y_h^{m+1}, \tilde{u}^{m+1} - u(t_{m+1})) \\
&- b(\lambda(t_{m+1}) - \lambda_{\tilde{h}}^{m+1}, \theta_h^{m+1}) + b(y_h^{m+1}, u(t_{m+1}) - u_h^{m+1}).
\end{aligned}
$$

We know that $(\tilde{u}^{m+1}, \tilde{\lambda}^{m+1})$ is the finite element approximation of $(u(t_{m+1}), \lambda(t_{m+1}))$, so

$$
b(\mu_{\tilde{h}}, u(t_{m+1}) - \tilde{u}^{m+1}) = 0 \quad \forall \mu_{\tilde{h}} \in \Lambda_{\tilde{h}}.
$$

Futhermore, using Lemma 3.6, T_1^{m+1} can be written as

$$\begin{aligned}
&T_1^{m+1} \\
&= -(D\theta_\pi^{m+1}, \theta_h^{m+1})_\Omega + (Du(t_{m+1}) - \partial_t u(t_{m+1}), \theta_h^{m+1})_\Omega - a(\theta_\pi^{m+1}, \theta_h^{m+1}) \\
&+ b(y_\pi^{m+1}, \theta_h^{m+1}).
\end{aligned}$$

Now, using Schwarz and Poincaré inequalities and approximation properties of the operator, we get

$$\begin{aligned}
&T_1^{m+1} \\
&\leq C\Big(\underbrace{\|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega} + \|D\theta_\pi^{m+1}\|_{0,\Omega}}_{T_2^{m+1}} + |\theta_\pi^{m+1}|_{1,\Omega} + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma}\Big)|\theta_h^{m+1}|_{1,\Omega} \\
&\leq \frac{C}{2}\left(T_2^{m+1} + |\theta_\pi^{m+1}|_{1,\Omega} + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma}\right)^2 + \frac{1}{2}|\theta_h^{m+1}|_{1,\Omega}^2. \qquad (3.46)
\end{aligned}$$

The term T_2^{m+1} can be treated as two terms. For the first one, a Taylor expansion gives

$$u(t_m) = u(t_{m+1}) - \delta t \partial_t u(t_{m+1}) - \frac{1}{2}\int_{t_m}^{t_{m+1}} \partial_{tt} u(s)(t_m - s)ds,$$

which implies

$$\begin{aligned}
&\left|\frac{u(t_{m+1}) - u(t_m)}{\delta t} - \partial_t u(t_{m+1})\right| \leq \frac{1}{2\delta t}\int_{t_m}^{t_{m+1}} |\partial_{tt} u(s)||(t_m - s)|ds \\
&\leq \frac{1}{2}\int_{t_m}^{t_{m+1}} |\partial_{tt} u(s)|ds \leq \frac{1}{2}\left(\int_{t_m}^{t_{m+1}} 1^2 ds\right)^{\frac{1}{2}}\left(\int_{t_m}^{t_{m+1}} |\partial_{tt} u(s)|^2 ds\right)^{\frac{1}{2}} \\
&= \frac{1}{2}\delta t^{\frac{1}{2}}\|\partial_{tt} u\|_{L^2(t_m, t_{m+1})}.
\end{aligned}$$

Then,

$$\begin{aligned}
&\left\|\frac{u(t_{m+1}) - u(t_m)}{\delta t} - \partial_t u(t_{m+1})\right\|_{0,\Omega}^2 = \int_\Omega \left|\frac{u(t_{m+1}) - u(t_m)}{\delta t} - \partial_t u(t_{m+1})\right|^2 dx \\
&\leq \frac{1}{4}\delta t\int_\Omega \|\partial_{tt} u\|_{L^2(t_m, t_{m+1})}^2 dx = \frac{\delta t}{4}\|\partial_{tt} u\|_{L^2(t_m, t_{m+1}; L^2(\Omega))}^2. \qquad (3.47)
\end{aligned}$$

For the second term we have, using (3.12) once again and a Taylor expansion,

$$
\begin{aligned}
&\|D\theta_\pi^{m+1}\|_{0,\Omega}^2 = \left\| \frac{\theta_\pi^{m+1} - \theta_\pi^m}{\delta t} \right\|_{0,\Omega}^2 \\
&= \left\| \frac{u(t_{m+1}) - P_{\tilde{h}}(u(t_{m+1}), \lambda(t_{m+1})) - u(t_m) + P_{\tilde{h}}(u(t_m), \lambda(t_m))}{\delta t} \right\|_{0,\Omega}^2 \\
&= \left\| \frac{u(t_{m+1}) - u(t_m)}{\delta t} - P_{\tilde{h}} \left(\frac{u(t_{m+1}) - u(t_m)}{\delta t}, \frac{\lambda(t_{m+1}) - \lambda(t_m)}{\delta t} \right) \right\|_{0,\Omega}^2 \\
&\leq Ch^2 \Big(\left| \frac{u(t_{m+1}) - u(t_m)}{\delta t} \right|_{2,\Omega}^2 + \left\| \frac{\lambda(t_{m+1}) - \lambda(t_m)}{\delta t} \right\|_{\mathcal{H}}^2 \Big) \\
&\leq Ch^2 \delta t \left(\|\partial_t u\|_{L^2(t_m,t_{m+1};H^2(\Omega))}^2 + \|\partial_t \lambda\|_{L^2(t_m,t_{m+1};\mathcal{H})}^2 \right). \qquad (3.48)
\end{aligned}
$$

Gathering (3.45)-(3.48), we arrive at

$$
\begin{aligned}
&\frac{1}{2} D\|\theta_h^{m+1}\|_{0,\Omega}^2 + |\theta_h^{m+1}|_{1,\Omega}^2 \\
&\leq C\Big(\delta t \|\partial_{tt} u\|_{L^2(t_m,t_{m+1};L^2(\Omega))}^2 + Ch^2 \delta t \Big(\|\partial_t u\|_{L^2(t_m,t_{m+1};H^2(\Omega))}^2 + \|\partial_t \lambda\|_{L^2(t_m,t_{m+1};\mathcal{H})}^2 \Big) \\
&+ |\theta_\pi^{m+1}|_{1,\Omega}^2 + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma}^2 \Big) + \frac{1}{2} |\theta_h^{m+1}|_{1,\Omega}^2.
\end{aligned}
$$

Multiplying the resulting expression by $2\delta t$ and summing over $0 \leq m \leq n-1$, we obtain

$$
\begin{aligned}
&\|\theta_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |\theta_h^{m+1}|_{1,\Omega}^2 \\
&\leq \|\theta_h^0\|_{0,\Omega}^2 + C\delta t \sum_{m=0}^{n-1} \Bigg(\delta t \|\partial_{tt} u\|_{L^2(t_m,t_{m+1};L^2(\Omega))}^2 \\
&+ Ch^2 \delta t \big(\|\partial_t u\|_{L^2(t_m,t_{m+1};H^2(\Omega))}^2 + \|\partial_t \lambda\|_{L^2(t_m,t_{m+1};\mathcal{H})}^2 \big) + |\theta_\pi^{m+1}|_{1,\Omega}^2 + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma}^2 \Bigg) \\
&= \|\theta_h^0\|_{0,\Omega}^2 + C\Bigg(\delta t^2 \|\partial_{tt} u\|_{L^2(0,t_n;L^2(\Omega))}^2 + Ch^2 \delta t^2 \big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 \\
&+ \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2 \big) + \sum_{m=0}^{n-1} \delta t \Big(|\theta_\pi^{m+1}|_{1,\Omega}^2 + \|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma}^2 \Big) \Bigg). \qquad (3.49)
\end{aligned}
$$

The only term that remains to be bounded is $\|\theta_h^0\|_{0,\Omega}^2$. For this, we write

$$
\begin{aligned}
\theta_h^0 &= P_{\tilde{h}}(u(0), \lambda(0)) - i_h(u_0) \\
&= u_0 - i_h(u_0) + P_{\tilde{h}}(u(0), 0) - u_0 + P_{\tilde{h}}(0, \lambda(0)).
\end{aligned}
$$

Due to the approximation properties of i_h and $P_{\tilde{h}}$ we have

$$
\|u_0 - i_h(u_0)\|_{0,\Omega} \leq Ch^2 \|u_0\|_{2,\Omega}, \qquad (3.50)
$$

and

$$\|P_{\tilde{h}}(u(0),0)-u_0\|_{0,\Omega}\le Ch\|u_0\|_{2,\Omega}. \tag{3.51}$$

In addition, $P_{\tilde{h}}(0,\lambda(0))$ is the finite element approximation of 0, and then

$$\|P_{\tilde{h}}(0,\lambda(0))\|_{0,\Omega}=\|P_{\tilde{h}}(0,\lambda(0))-0\|_{0,\Omega}\le Ch\|\lambda(0)\|_{\mathcal{H}}. \tag{3.52}$$

Hence, inserting (3.50)-(3.52) in (3.49), we arrive at

$$\begin{aligned}
&\|\theta_h^n\|_{0,\Omega}^2+\sum_{m=0}^{n-1}\delta t|\theta_h^{m+1}|_{1,\Omega}^2\\
&\le C\Big(h^2(\|u_0\|_{2,\Omega}^2+\|\lambda(0)\|_{\mathcal{H}}^2)+\delta t^2\|\partial_{tt}u\|_{L^2(0,t_n;L^2(\Omega))}^2+h^2\delta t^2\Big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2\\
&+\|\partial_t\lambda\|_{L^2(0,t_n;\mathcal{H})}^2\Big)+h^2\sum_{m=0}^{n-1}\delta t\Big(\|u(t_{m+1})\|_{2,\Omega}^2+\|\lambda(t_{m+1})\|_{\mathcal{H}}^2\Big)\Big).
\end{aligned} \tag{3.53}$$

For the Lagrange multiplier error estimate we first note that, from (3.44), it suffices to control $\|y_h^{n+1}\|_{-\frac{1}{2},\gamma}$. To this end, we use the inf-sup condition (3.3):

$$\beta\|y_h^{n+1}\|_{-\frac{1}{2},\gamma}\le\sup_{v_h\in V_h}\frac{b(y_h^{n+1},v_h)}{|v_h|_{1,\Omega}}.$$

From (3.44) we have

$$b(y_h^{n+1},v_h)=-b(y_\pi^{n+1},v_h)+b(\lambda(t_{n+1})-\lambda_{\tilde{h}}^{n+1},v_h).$$

Using the definition of b and the trace inequality, we get

$$b(y_\pi^{n+1},v_h)\le\|y_\pi^{n+1}\|_{-\frac{1}{2},\gamma}|v_h|_{1,\Omega}.$$

On the other hand, using the modified Galerkin orthogonality (see Lemma 3.6) with $\mu_{\tilde{h}}=0$, we have

$$\begin{aligned}
&b(\lambda(t_{n+1})-\lambda_{\tilde{h}}^{n+1},v_h)=a(u(t_{n+1})-u_h^{n+1},v_h)+(\partial_t u(t_{n+1})-Du_h^{n+1},v_h)_\Omega\\
&\le|u(t_{n+1})-u_h^{n+1}|_{1,\Omega}|v_h|_{1,\Omega}+\|\partial_t u(t_{n+1})-Du_h^{n+1}\|_{0,\Omega}\|v_h\|_{0,\Omega}.
\end{aligned}$$

As a result, from the above estimations, we have

$$\beta\|y_h^{n+1}\|_{-\frac{1}{2},\gamma}\le\Big(\|y_\pi^{n+1}\|_{-\frac{1}{2},\gamma}+|u(t_{n+1})-u_h^{n+1}|_{1,\Omega}\Big)+C\|\partial_t u(t_{n+1})-Du_h^{n+1}\|_{0,\Omega}.$$

Therefore,

$$
\begin{aligned}
&\beta^2 \sum_{m=0}^{n-1} \delta t \|y_h^{m+1}\|^2_{-\frac{1}{2},\gamma} \\
&\leq C \sum_{m=0}^{n-1} \delta t \left(\|y_\pi^{m+1}\|^2_{-\frac{1}{2},\gamma} + |u(t_{m+1}) - u_h^{m+1}|^2_{1,\Omega} + \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega} \right).
\end{aligned}
$$

Finally, standard interpolation estimate 3.12 gives

$$
\|y_\pi^{m+1}\|_{-\frac{1}{2},\gamma} \leq Ch\|\lambda(t_{n+1})\|_{\mathcal{H}}.
$$

This together with the estimate for u finish the proof. □

The only problematic term is in the second estimate in Lemma 3.7, $\|\partial_t u - Du_h^{n+1}\|_{0,\Omega}$, which we bound in the following theorem using the hypothesis $u_h^0 = P_h(u_0, 0)$.

Theorem 3.8. *Under the assumptions of Lemma 3.7, assuming that $\lambda \in H^1(0,T;\mathcal{H})$, $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, and $u_h^0 := P_{\tilde{h}}(u_0, 0)$, for $1 \leq n \leq N$ we have*

$$
\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\lambda_{\tilde{h}}^{m+1} - \lambda(t_{m+1})\|^2_{-\frac{1}{2},\gamma} \\
&\leq C\Big(h^2\|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} + h^2\|u_0\|^2_{2,\Omega} + h^2\|u\|^2_{C^0(t_1,t_n;H^2(\Omega))} + \delta t^2 \|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} \\
&+ h^2 \delta t^2 \big(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(0,t_n;\mathcal{H})}\big) + h^2\|\lambda(0)\|^2_{\mathcal{H}}\Big).
\end{aligned}
$$

Proof. From Lemma 3.7, we only need to prove the following estimate:

$$
\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega} + |P_{\tilde{h}}(u(t_n), \lambda(t_n)) - u_h^n|^2_{1,\Omega} \\
&\leq C\Big(\delta t^2 \|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} + h^2 \delta t^2 \big(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(0,t_n;\mathcal{H})}\big)\Big) \\
&+ Ch^2 \|\lambda(0)\|^2_{\mathcal{H}}.
\end{aligned} \tag{3.54}
$$

We use the descomposition of the error given in (3.43)-(3.44). Using the triangle inequality we get

$$
\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega} \\
&\leq C \sum_{m=0}^{n-1} \delta t \Big(\|\partial_t u(t_{m+1}) - Du(t_{m+1})\|^2_{0,\Omega} + \|D\theta_\pi^{m+1}\|^2_{0,\Omega} + \|D\theta_h^{m+1}\|^2_{0,\Omega} \Big).
\end{aligned}
$$

The first and second terms are bounded in Lemma 3.7. For the third term, we use the definition of the Ritz-projection (3.10) which implies

$$a(u(t_{n+1}) - \tilde{u}^{n+1}, v_h) - b(\lambda(t_{n+1}) - \tilde{\lambda}^{n+1}, v_h) = a(\theta_\pi^{n+1}, v_h) - b(y_\pi^{n+1}, v_h) = 0.$$

The modified Galerkin orthogonality (see Lemma 3.6) with $\mu_{\tilde{h}} = 0$ gives

$$\begin{aligned}
&\|D\theta_h^{n+1}\|_{0,\Omega}^2 + a(\theta_h^{n+1}, D\theta_h^{n+1}) - b(y_h^{n+1}, D\theta_h^{n+1}) \\
&= \Big(D(P_{\tilde{h}}(u(t_{n+1}), \lambda(t_{n+1})) - u_h^{n+1} \pm u(t_{n+1})), D\theta_h^{n+1}\Big)_\Omega \\
&+ a\Big(P_{\tilde{h}}(u(t_{n+1}), \lambda(t_{n+1})) - u_h^{n+1} \pm u(t_{n+1}), D\theta_h^{n+1}\Big) \\
&- b\Big(R_{\tilde{h}}(u(t_{n+1}), \lambda(t_{n+1})) - \lambda_{\tilde{h}}^{n+1} \pm \lambda(t_{n+1}), D\theta_h^{n+1}\Big) \\
&= -(D\theta_\pi^{n+1}, D\theta_h^{n+1})_\Omega - a(\theta_\pi^{n+1}, D\theta_h^{n+1}) \\
&+ b(y_\pi^{n+1}, D\theta_h^{n+1}) + (Du(t_{n+1}) - \partial_t u(t_{n+1}), D\theta_h^{n+1})_\Omega \\
&= -(D\theta_\pi^{n+1}, D\theta_h^{n+1})_\Omega + (Du(t_{n+1}) - \partial_t u(t_{n+1}), D\theta_h^{n+1})_\Omega.
\end{aligned}$$

and, using Young's inequality, we arrive at

$$\begin{aligned}
&\frac{1}{2}\|D\theta_h^{n+1}\|_{0,\Omega}^2 + a(\theta_h^{n+1}, D\theta_h^{n+1}) - b(y_h^{n+1}, D\theta_h^{n+1}) \\
&\leq C\left(\|D\theta_\pi^{n+1}\|_{0,\Omega}^2 + \|Du(t_{n+1}) - \partial_t u(t_{n+1})\|_{0,\Omega}^2\right).
\end{aligned} \tag{3.55}$$

Moreover, since, for every $0 \leq m \leq N$ the following Galerkin orthogonality holds

$$b(\mu_{\tilde{h}}, \theta_h^m) = 0 \quad \forall \mu_{\tilde{h}} \in \Lambda_{\tilde{h}},$$

then

$$b(y_h^{n+1}, D\theta_h^{n+1}) = \frac{1}{\delta t}\left(b(y_h^{n+1}, \theta_h^{n+1}) - b(y_h^{n+1}, \theta_h^n)\right) = 0.$$

So, inserting this into (3.55), we get

$$\frac{1}{2}\|D\theta_h^{n+1}\|_{0,\Omega}^2 + a(\theta_h^{n+1}, D\theta_h^{n+1}) \leq C\Big(\|D\theta_\pi^{n+1}\|_{0,\Omega}^2 + \|Du(t_{n+1}) - \partial_t u(t_{n+1})\|_{0,\Omega}^2\Big). \tag{3.56}$$

Using the symmetry of $a(.,.)$, we get

$$a(\theta_h^{n+1}, D\theta_h^{n+1}) = \frac{1}{2}Da(\theta_h^{n+1}, \theta_h^{n+1}) + \frac{\delta t}{2}a(D\theta_h^{n+1}, D\theta_h^{n+1}),$$

which, after (3.56), gives

$$\frac{1}{2}\|D\theta_h^{n+1}\|_{0,\Omega}^2 + \frac{1}{2}Da(\theta_h^{n+1}, \theta_h^{n+1}) \leq \|D\theta_\pi^{n+1}\|_{0,\Omega}^2 + \|Du(t_{n+1}) - \partial_t u(t_{n+1})\|_{0,\Omega}^2.$$

Thus, after multiplication by $2\delta t$ and adding over $0 \leq m \leq n-1$, we have

$$\sum_{m=0}^{n-1} \delta t \|D\theta_h^{m+1}\|_{0,\Omega}^2 + |\theta_h^n|_{1,\Omega}^2$$
$$\leq |\theta_h^0|_{1,\Omega}^2 + C \sum_{m=0}^{n-1} \delta t \left(\|D\theta_\pi^{m+1}\|_{0,\Omega}^2 + \|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega}^2 \right),$$

where we just have to bound the initial term because the other terms have already bounded in (3.47) and (3.48). Therefore,

$$\theta_h^0 = P_{\tilde{h}}(u(0), \lambda(0)) - u_h^0 = \underbrace{P_{\tilde{h}}(0, \lambda(0))}_{:=w_h \in V_h} + \underbrace{P_{\tilde{h}}(u(0), 0) - u_h^0}_{=0}.$$

We realise that $P_{\tilde{h}}(0, \lambda(0))$ is the finite element approximation of $(0, \lambda(0))$, and then

$$|\theta_h^0|_{1,\Omega} = |P_h(0, \lambda(0)) - 0|_{1,\Omega} \leqslant Ch(|0|_{2,\Omega} + \|\lambda(0)\|_{\mathcal{H}}).$$

This implies

$$|\theta_h^0|_{1,\Omega}^2 \leq Ch^2 \|\lambda(0)\|_{\mathcal{H}}^2.$$

This gives rise to the following estimate:

$$\sum_{m=0}^{n-1} \delta t \|D\theta_h^{m+1}\|_{0,\Omega}^2 + |\theta_h^n|_{1,\Omega}^2$$
$$\leq C\Big(h^2 \|\lambda(0)\|_{\mathcal{H}}^2 + \delta t^2 \|\partial_{tt} u\|_{L^2(0,t_n;L^2(\Omega))}^2 + h^2 \delta t^2 \big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2$$
$$+ \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2 \big) \Big)$$

for all $1 \leq n \leq N$, which proves (3.54). □

3.4.2. The stabilised method

We start by presenting the following result on consistency. Its proof is direct verification.

Lemma 3.9. *Let (u, λ) be the solution of (3.1) and let $\{(u_h^n, \lambda_h^n)\}_{0 \leq n \leq N}$ be the solution of (3.9). Assume that $u \in \mathcal{C}^0(0, T; H_0^1(\Omega))$ and let $\lambda \in \mathcal{C}^0(0, T; H^{-\frac{1}{2}}(\gamma))$. Then, for $0 \leq n \leq N-1$, there holds*

$$(Du(t_{n+1}) - Du_h^{n+1}, v_h)_\Omega + a(u(t_{n+1}) - u_h^{n+1}, v_h) - b(\lambda(t_{n+1}) - \lambda_h^{n+1}, v_h)$$
$$+ b(\mu_h, u(t_{n+1}) - u_h^{n+1}) = j(\lambda_h^{n+1}, \mu_h) + (Du(t_{n+1}) - \partial_t u(t_{n+1}), v_h)_\Omega,$$

for all $(v_h, \mu_h) \in W_h$.

We prove the following result which states an optimal error estimate for u, and an estimate for λ which is optimal, up to a term that will be treated later.

Lemma 3.10. *Let us assume that* $u \in H^1(0,T;H^2(\Omega)) \cap H^2(0,T;L^2(\Omega)) \cap C^0(0,T;H^2(\Omega))$, $u_0 \in H^2(\Omega) \cap H^1_0(\Omega)$ *and* $\lambda \in H^1(0,T;\mathcal{H})$, *and set* $u_h^0 \in V_h$ *as* $u_h^0 = i_h(u)$ *where* $i_h : H^1_0(\Omega) \longrightarrow V_h$ *stands for the Clément interpolation operator. Then the following estimates hold for* $1 \le n \le N$*:*

$$\begin{aligned}
&\|u_h^n - u(t_n)\|^2_{0,\Omega} + \sum_{m=0}^{n-1} \delta t \big|\!\big|\!\big|(u_h^{m+1} - u(t_{m+1}), \lambda(t_{m+1}) - \lambda_h^{m+1})\big|\!\big|\!\big|_h^2 \\
&\le Ch^2\Big(\|u_0\|^2_{2,\Omega} + \|\lambda(0)\|^2_{\mathcal{H}}\Big) + Ch^2\delta t^2\left(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(0,t_n;\mathcal{H})}\right) \\
&+ C\left(\delta t^2 \|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} + h^2\|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} + h^2\|u\|^2_{C^0(t_1,t_n;H^2(\Omega))}\right),
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\lambda_h^{m+1} - \lambda(t_{m+1})\|^2_{-\frac{1}{2},\gamma} \le C\Big(1 + \frac{1}{\beta^2}\Big) h^2 \|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} \\
&+ \frac{C}{\beta^2}\Big(Ch^2\Big(\|u_0\|^2_{2,\Omega} + \|\lambda(0)\|^2_{\mathcal{H}}\Big) + Ch^2\delta t^2\Big(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t\lambda\|^2_{L^2(0,t_n;\mathcal{H})}\Big) \\
&+ C\Big(\delta t^2\|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} + h^2\|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} + h^2\|u\|^2_{C^0(t_1,t_n;H^2(\Omega))}\Big) \\
&+ \sum_{m=0}^{n-1} \delta t \|\partial t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega}\Big),
\end{aligned}$$

with $C > 0$ *a positive constant independent of* h *and* δt.

Proof. As usual, for $s = 0, ..., N$, we decompose the error in to interpolation and discrete errors as follows

$$\begin{aligned}
u(t_s) - u_h^s &= \underbrace{u(t_s) - P_h(u(t_s), \lambda(t_s))}_{\theta_\pi^s} + \underbrace{P_h(u(t_s), \lambda(t_s)) - u_h^s}_{\theta_h^s} \\
&:= \theta_\pi^s + \theta_h^s, \qquad (3.57)
\end{aligned}$$

$$\begin{aligned}
\lambda(t_s) - \lambda_h^s &= \underbrace{\lambda(t_s) - R_h(u(t_s), \lambda(t_s))}_{y_\pi^s} + \underbrace{R_h(u(t_s), \lambda(t_s)) - \lambda_h^s}_{y_h^s} \\
&:= y_\pi^s + y_h^s, \qquad (3.58)
\end{aligned}$$

where (P_h, R_h) is defined (3.28). To facilite the notation, let

$$\hat{u}^s = P_h(u(t_s), \lambda(t_s)),$$
$$\hat{\lambda}^s = R_h(u(t_s), \lambda(t_s)).$$

The first term θ_π^{m+1} can be bounded using (3.12). In order to estimate θ_h^{m+1} using the definition of the bilinear form $a(.,.) + j(.,.)$ and

$$(Du_h^{m+1}, u_h^{m+1}) = \frac{1}{2}D\|u_h^{m+1}\|_{0,\Omega}^2 + \frac{1}{2\delta t}\|u_h^{m+1} - u_h^m\|_{0,\Omega}^2,$$

we get

$$\begin{aligned}
&\frac{1}{2}D\|\theta_h^{m+1}\|_{0,\Omega}^2 + |||(\theta_h^{m+1}, y_h^{m+1})|||_h^2 \\
&\leq (D\theta_h^{m+1}, \theta_h^{m+1})_\Omega + |||(\theta_h^{m+1}, y_h^{m+1})|||_h^2 \\
&= \underbrace{(D\theta_h^{m+1}, \theta_h^{m+1})_\Omega + a(\theta_h^{m+1}, \theta_h^{m+1}) + b(y_h^{m+1}, \theta_h^{m+1}) - b(y_h^{m+1}, \theta_h^{m+1}) + j(y_h^{m+1}, y_h^{m+1})}_{T_1^{m+1}}.
\end{aligned} \tag{3.59}$$

In addition, using (3.57)-(3.58), we have

$$\begin{aligned}
T_1^{m+1} &= -(D\theta_\pi^{m+1}, \theta_h^{m+1})_\Omega - a(\theta_\pi^{m+1}, \theta_h^{m+1}) + a(u(t_{m+1}) - u_h^{m+1}, \theta_h^{m+1}) \\
&+ j(\hat{\lambda}^{m+1}, y_h^{m+1}) - j(\lambda_h^{m+1}, y_h^{m+1}) + (Du(t_{m+1}) - Du_h^{m+1}, \theta_h^{m+1})_\Omega \\
&- b(\hat{\lambda}^{m+1} - \lambda(t_{m+1}), \theta_h^{m+1}) + b(y_h^{m+1}, \hat{u}^{m+1} - u(t_{m+1})) \\
&- b(\lambda(t_{m+1}) - \lambda_h^{m+1}, \theta_h^{m+1}) + b(y_h^{m+1}, u(t_{m+1}) - u_h^{m+1}).
\end{aligned}$$

By the modified Galerkin orthogonality (see Lemma 3.9), this expression reduces to

$$\begin{aligned}
T_1^{m+1} &= -(D\theta_\pi^{m+1}, \theta_h^{m+1})_\Omega + (Du(t_{m+1}) - \partial_t u(t_{m+1}), \theta_h^{m+1})_\Omega - a(\theta_\pi^{m+1}, \theta_h^{m+1}) \\
&+ j(\hat{\lambda}^{m+1}, y_h^{m+1}) + b(y_\pi^{m+1}, \theta_h^{m+1}) - b(y_h^{m+1}, \theta_\pi^{m+1}) \\
&= I + II + III + IV + V.
\end{aligned} \tag{3.60}$$

We now bound the above right-hand side term by term using the Cauchy Schwarz and the Poincaré inequalities:

$$\begin{aligned}
I &\leq \Big(\|D\theta_\pi^{m+1}\|_{0,\Omega} + \|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega}\Big)\|\theta_h^{m+1}\|_{0,\Omega} \\
&\leq \Big(\|D\theta_\pi^{m+1}\|_{0,\Omega} + \|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega}\Big) C_p|\theta_h^{m+1}|_{1,\Omega} \\
&\leq \Big(\|D\theta_\pi^{m+1}\|_{0,\Omega} + \|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega}\Big) C_p|||(\theta_h^{m+1}, y_h^{m+1})|||_h,
\end{aligned} \tag{3.61}$$

$$II \leq |a(\theta_\pi^{m+1}, \theta_h^{m+1})| \leq |\theta_\pi^{m+1}|_{1,\Omega} \|\theta_h^{m+1}\|_{0,\Omega} \leq |\theta_\pi^{m+1}|_{1,\Omega} \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h, \tag{3.62}$$

$$III \leq j(\hat{\lambda}^{m+1}, \hat{\lambda}^{m+1})^{\frac{1}{2}} j(y_h^{m+1}, y_h^{m+1})^{\frac{1}{2}} \leq j(\hat{\lambda}^{m+1}, \hat{\lambda}^{m+1})^{\frac{1}{2}} \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h, \tag{3.63}$$

$$\begin{aligned} IV &= \langle y_\pi^{m+1}, \theta_h^{m+1} \rangle_\gamma \\ &\leq \|\lambda(t_{m+1}) - \hat{\lambda}^{m+1}\|_{-\frac{1}{2},\gamma} \|\theta_h^{m+1}\|_{\frac{1}{2},\gamma} \\ &\leq Ch \|\lambda(t_{m+1})\|_{\mathcal{H}} \|\theta_h^{m+1}\|_{1,\Omega} \\ &\leq Ch \|\lambda(t_{m+1})\|_{\mathcal{H}} \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h, \end{aligned} \tag{3.64}$$

$$\begin{aligned} V &\leq |b(y_h^{m+1}, \theta_\pi^{m+1})| = \langle y_h^{m+1}, \theta_\pi^{m+1} \rangle_\gamma = j(\hat{\lambda}^{m+1}, y_h^{m+1}) \\ &= j(\hat{\lambda}^{m+1} - \lambda(t_{m+1}), y_h^{m+1}) + j(\lambda(t_{m+1}), y_h^{m+1}) \\ &\leq \Big(j(\hat{\lambda}^{m+1} - \lambda(t_{m+1}), \hat{\lambda}^{m+1} - \lambda(t_{m+1})) + j(\lambda(t_{m+1}), \lambda(t_{m+1})) \Big)^{\frac{1}{2}} j(y_h^{m+1}, y_h^{m+1})^{\frac{1}{2}} \\ &\leq Ch \|\lambda(t_{m+1})\|_{\mathcal{H}} \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h. \end{aligned} \tag{3.65}$$

Then collecting (3.60)-(3.65), we get

$$\begin{aligned} T_1^{m+1} \leq\ & \underbrace{\Big(\|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega} + \|D\theta_\pi^{m+1}\|_{0,\Omega} \Big)}_{T_2^{m+1}} C_p \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h \\ &+ \Big(|\theta_\pi^{m+1}|_{1,\Omega} + j(\hat{\lambda}^{m+1}, \hat{\lambda}^{m+1})^{\frac{1}{2}} + Ch \|\lambda(t_{m+1})\|_{\mathcal{H}} \Big) \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h. \end{aligned} \tag{3.66}$$

The term T_2^{m+1} has already been analysed (see (3.47) and (3.48)). Thus, from (3.66) and using Young's inequality, it follows that

$$\begin{aligned} T_1^{m+1} \leq\ & \frac{1}{2} \big|\big|\big|(\theta_h^{m+1}, y_h^{m+1})\big|\big|\big|_h^2 + C\Big(\frac{\delta t}{4} \|\partial_{tt} u\|^2_{L^2(t_m, t_{m+1}; L^2(\Omega))} \\ &+ Ch^2 \delta t \big(\|\partial_t u\|^2_{L^2(t_m, t_{m+1}; H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(t_m, t_{m+1}; \mathcal{H})} \big) \\ &+ |\theta_\pi^{m+1}|^2_{1,\Omega} + j(\hat{\lambda}^{m+1}, \hat{\lambda}^{m+1}) + h^2 \|\lambda(t_{m+1})\|^2_{\mathcal{H}} \Big). \end{aligned}$$

By inserting this expression into (3.59), multiplying the resulting expression by $2\delta t$, and adding over $0 \leq m \leq n-1$, we obtain

$$\begin{aligned}
&\|\theta_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |||(\theta_h^{m+1}, y_h^{m+1})|||_h^2 \\
&\leq \|\theta_h^0\|_{0,\Omega}^2 + C\bigg(\frac{\delta t^2}{2}\|\partial_{tt}u\|_{L^2(0,t_n;L^2(\Omega))}^2 + Ch^2\delta t^2\Big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 \\
&+ \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2\Big) + \sum_{m=0}^{n-1} \delta t \left(|\theta_\pi^{m+1}|_{1,\Omega}^2 + j(\hat{\lambda}^{n+1}, \hat{\lambda}^{n+1}) + h^2\|\lambda(t_{n+1})\|_{\mathcal{H}}^2\right)\bigg).
\end{aligned}$$

where $|\theta_\pi^{m+1}|_{1,\Omega}^2$ has already bounded and

$$j(\hat{\lambda}^{n+1}, \hat{\lambda}^{n+1}) = j\left(\hat{\lambda}^{n+1} - \lambda(t_{n+1}), \hat{\lambda}^{n+1} - \lambda(t_{n+1})\right) + j\left(\lambda(t_{n+1}), \lambda(t_{n+1})\right),$$

$$\begin{aligned}
&j\left(\hat{\lambda}^{n+1} - \lambda(t_{n+1}), \hat{\lambda}^{n+1} - \lambda(t_{n+1})\right) \leq \left|\left|\left|(\hat{u}^{n+1} - u(t_{n+1}), \hat{\lambda}^{n+1} - \lambda(t_{n+1}))\right|\right|\right|_h^2 \\
&\leq Ch^2(|u(t_{n+1})|_{2,\Omega}^2 + \|\lambda(t_{n+1})\|_{\mathcal{H}}),
\end{aligned}$$

and

$$j(\lambda(t_{n+1}), \lambda(t_{n+1})) = \sum_{e\in\gamma} |e| \|\lambda(t_{n+1}) - \Pi_{\tilde{h}}\lambda(t_{n+1})\|_{0,e}^2 \leq C\sum_{e\in\gamma} |e|^2\|\lambda(t_{n+1})\|_{\mathcal{H}}^2.$$

Therefore,

$$j(\hat{\lambda}^{n+1}, \hat{\lambda}^{n+1}) \leq Ch^2(|u(t_{n+1})|_{2,\Omega}^2 + \|\lambda(t_{n+1})\|_{\mathcal{H}}^2).$$

The error estimate for u is obtained using approximation properties of a projection operator and the consistency of the Lagrange multiplier stabilisation, which yields

$$\begin{aligned}
&\|\theta_h^n\|_{0,\Omega}^2 + \sum_{m=0}^{n-1} \delta t |||(\theta_h^{m+1}, y_h^{m+1})|||_h^2 \\
&\leq \|\theta_h^0\|_{0,\Omega}^2 + C\Big(\delta t^2\|\partial_{tt}u\|_{L^2(0,t_n;L^2(\Omega))}^2 + h^2\delta t^2\left(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 + \|\partial_t\lambda\|_{L^2(0,t_n;\mathcal{H})}^2\right) \\
&+ h^2\sum_{m=0}^{n-1}\delta t\|u(t_{m+1})\|_{2,\Omega}^2 + h^2\sum_{m=0}^{n-1}\delta t\|\lambda(t_{m+1})\|_{\mathcal{H}}^2\Big).
\end{aligned}$$

Finally,

$$\begin{aligned}
\theta_h^0 &= P_h(u_0, \lambda(0)) - u_h^0 \\
&= P_h(u_0, 0) - u_h^0 + P_h(0, \lambda(0)),
\end{aligned}$$

and, procceding as before,

$$\|\theta_h^0\|_{0,\Omega}^2 \leq Ch(\|u_0\|_{2,\Omega} + \|\lambda(0)\|_{\mathcal{H}}).$$

For the Lagrange multiplier error estimate we first note that, from (3.58), it suffices to control $\|y_h^{n+1}\|_{0,\Omega}$. To this end, we use the inf-sup condition (3.4):

$$\beta\|y_h^{n+1}\|_{-\frac{1}{2},\gamma} \leq \sup_{v_h\in V_h\setminus\{0\}} \frac{b(y_h^{n+1},v_h)}{|v_h|_{1,\Omega}} + Cj(y_h^{n+1},y_h^{n+1})^{\frac{1}{2}}.$$

From (3.58), we have

$$b(y_h^{n+1},v_h) = -b(y_\pi^{n+1},v_h) + b(\lambda(t_{n+1})-\lambda_h^{n+1},v_h).$$

The first term can be bounded using the continuity of $b(.,.)$ and the trace inequality, which yields

$$b(y_\pi^{n+1},v_h) \leq \|y_\pi^{n+1}\|_{-\frac{1}{2},\gamma}|v_h|_{1,\Omega}.$$

On the other hand, using the modified Galerkin orthogonality (see Lemma 3.9) with $\mu_h = 0$, we have

$$\begin{aligned}
&b(\lambda(t_{n+1})-\lambda_h^{n+1},v_h) = a(u(t_{n+1})-u_h^{n+1},v_h) + (\partial_t u(t_{n+1}) - Du_h^{n+1},v_h)_\Omega \\
&\leq C|||(u(t_{n+1})-u_h^{n+1},0)|||_h|v_h|_{1,\Omega} + \|\partial_t u(t_{n+1}) - Du_h^{n+1}\|_{0,\Omega}|v_h|_{1,\Omega}.
\end{aligned}$$

As a result, from the above estimations, we have

$$\begin{aligned}
&\beta\|y_h^{n+1}\|_{-\frac{1}{2},\gamma} \leq \sup_{v_h\in V_h\setminus\{0\}} \frac{|-b(y_\pi^{n+1},v_h)+b(\lambda(t_{n+1})-\lambda_h^{n+1},v_h)|}{|v_h|_{1,\Omega}} + Cj(y_h^{n+1},y_h^{n+1})^{\frac{1}{2}} \\
&\leq C\Bigg(\frac{\|y_\pi^{n+1}\|_{0,\Omega}|v_h|_{1,\Omega}}{|v_h|_{1,\Omega}} + \frac{|||(u(t_{n+1})-u_h^{n+1},y_h^{n+1})|||_h|v_h|_{1,\Omega}}{|v_h|_{1,\Omega}} \\
&+ \frac{\|\partial_t u(t_{n+1}) - Du_h^{n+1}\|_{0,\Omega}|v_h|_{1,\Omega}}{|v_h|_{1,\Omega}}\Bigg) \\
&\leq C\Big(\|y_\pi^{n+1}\|_{-\frac{1}{2},\gamma} + |||(u(t_{n+1})-u_h^{n+1},y_h^{n+1})|||_h + C_p\|\partial_t u(t_{n+1}) - Du_h^{n+1}\|_{0,\Omega}\Big).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\beta^2\sum_{m=0}^{n-1}\delta t\|y_h^{m+1}\|^2_{-\frac{1}{2},\gamma} \\
&\leq C\sum_{m=0}^{n-1}\delta t\Big(\|y_\pi^{m+1}\|^2_{-\frac{1}{2},\gamma} + |||(u(t_{m+1})-u_h^{m+1},y_h^{m+1})|||_h^2 + \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega}\Big)
\end{aligned}$$

and we conclude using the error estimate for u. □

We solve the problem of the Lagrange multiplier convergence by providing an error estimate for the term $\|\partial_t u - Du_h^{n+1}\|_{0,\Omega}$.

Theorem 3.11. *Under the assumptions of Lemma 3.10, assuming that* $\lambda \in H^1(0,T;\mathcal{H})$, $u_0 \in H^2(\Omega) \cap H^1_0(\Omega)$, *and* $u_h^0 := P_h(u_0, 0)$, *for* $1 \leq n \leq N$ *we have*

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\lambda_h^{m+1} - \lambda(t_{m+1})\|^2_{-\frac{1}{2},\gamma} \\
&\leq C\left(1 + \frac{1}{\beta^2}\right) h^2 \|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} + Ch^2\|u_0\|^2_{2,\Omega} \\
&+ Ch^2\delta t^2 \left(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(0,t_n;\mathcal{H})}\right) \\
&+ C\Big(\delta t^2 \|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} + h^2\|\lambda\|^2_{C^0(t_1,t_n;\mathcal{H})} + h^2\|u\|^2_{C^0(t_1,t_n;H^2(\Omega))}) + Ch^2\|\lambda(0)\|^2_{\mathcal{H}}\Big).
\end{aligned}$$

Proof. Based on the previous lemma, we just need to prove

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|Du_h^{m+1} - \partial_t u(t_{m+1})\|^2_{0,\Omega} + \left|\left|\left|(\hat{u}^n - u_h^n, \hat{\lambda}^n - \lambda_h^n)\right|\right|\right|^2_h \\
&\leq C\Big(\delta t^2 \|\partial_{tt} u\|^2_{L^2(0,t_n;L^2(\Omega))} + h^2\delta t^2 \left(\|\partial_t u\|^2_{L^2(0,t_n;H^2(\Omega))} + \|\partial_t \lambda\|^2_{L^2(0,t_n;\mathcal{H})}\right) \\
&+ Ch^2\|\lambda(0)\|^2_{\mathcal{H}}.
\end{aligned} \tag{3.67}$$

As usual, in order to provide an optimal error estimate, we decompose the error as in (3.57)-(3.58). Using the triangle inequality, we have

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t \|\partial_t u(t_{m+1}) - Du_h^{m+1}\|^2_{0,\Omega} \\
&= \sum_{m=0}^{n-1} \delta t \|\partial_t u(t_{m+1}) - Du(t_{m+1}) + D\theta_\pi^{m+1} + D\theta_h^{m+1}\|^2_{0,\Omega} \\
&\leq C \sum_{m=0}^{n-1} \delta t \Big(\|\partial_t u(t_{m+1}) - Du(t_{m+1})\|^2_{0,\Omega} + \|D\theta_\pi^{m+1}\|^2_{0,\Omega} + \|D\theta_h^{m+1}\|^2_{0,\Omega}\Big).
\end{aligned}$$

For the first and second terms, we proceed as in the inf-sup case (see Lemma 3.7). For the third term, we use the modified Galerkin orthogonality (see Lemma 3.9) with $\mu_h = 0$ and the definition of the Ritz-projection (3.28) to obtain

$$\begin{aligned}
&\|D\theta_h^{n+1}\|^2_{0,\Omega} + a(\theta_h^{n+1}, D\theta_h^{n+1}) - b(y_h^{n+1}, D\theta_h^{n+1}) \\
&= -(D\theta_\pi^{n+1}, D\theta_h^{n+1})_\Omega - a(\theta_\pi^{n+1}, D\theta_h^{n+1}) + b(y_\pi^{n+1}, D\theta_h^{n+1}) \\
&+ (Du(t_{n+1}) - \partial_t u(t_{n+1}), D\theta_h^{n+1})_\Omega \\
&= -(D\theta_\pi^{n+1}, D\theta_h^{n+1})_\Omega + (Du(t_{n+1}) - \partial_t u(t_{n+1}), D\theta_h^{n+1})_\Omega.
\end{aligned}$$

Next, Young's inequality yields

$$\begin{aligned}&\frac{1}{2}\|D\theta_h^{n+1}\|_{0,\Omega}^2 + a(\theta_h^{n+1}, D\theta_h^{n+1}) - b(y_h^{n+1}, D\theta_h^{n+1}) \\ &\leq C\Big(\|D\theta_\pi^{n+1}\|_{0,\Omega}^2 + \|Du(t_{n+1}) - \partial_t u(t_{n+1})\|_{0,\Omega}^2\Big).\end{aligned} \tag{3.68}$$

In addition, for $0 \leq n \leq N$, testing (3.28) at the time level n with $v_h = 0$, we have

$$b\big(\mu_h, \hat{u}^n\big) = -j(\hat{\lambda}^n, \mu_h) + b(\mu_h, g(t_n)). \tag{3.69}$$

For $1 \leq n \leq N$, testing (3.9) at the time level n with $v_h = 0$ and since by definition, $u_h^0 = P_h(u_0, 0)$, we have

$$b(\mu_h, u_h^n) = -j(\lambda_h^n, \mu_h) + b(\mu_h, g(t_n)), \tag{3.70}$$

for all $\mu_h \in \Lambda_h$ and $0 \leq n \leq N$ and where we have defined $\lambda_h^0 = R_h(u_0, 0)$. As a result, from (3.69) - (3.70), we have

$$b(\mu_h, \theta_h^n) = -j(y_h^n, \mu_h),$$

for all $\mu_h \in \Lambda_h$ and $0 \leq n \leq N$. We therefore have, for $0 \leq n \leq N-1$,

$$b(y_h^{n+1}, D\theta_h^{n+1}) = -j(Dy_h^{n+1}, y_h^{n+1}). \tag{3.71}$$

On the other hand, using the symmetry of a and j, we get

$$\begin{aligned}a(\theta_h^{n+1}, D\theta_h^{n+1}) &= \frac{1}{2}Da(\theta_h^{n+1}, \theta_h^{n+1}) + \frac{\delta t}{2}a(D\theta_h^{n+1}, D\theta_h^{n+1}),\\ j(y_h^{n+1}, Dy_h^{n+1}) &= \frac{1}{2}Dj(y_h^{n+1}, y_h^{n+1}) + \frac{\delta t}{2}j(Dy_h^{n+1}, Dy_h^{n+1}).\end{aligned}$$

Then, replacing (3.71) and the last set of equalities in (3.68), we get

$$\begin{aligned}&\frac{1}{2}\|D\theta_h^{n+1}\|_{0,\Omega}^2 + \frac{1}{2}D(a(\theta_h^{n+1}, \theta_h^{n+1}) + j(y_h^{n+1}, y_h^{n+1}))\\ &\leq \|D\theta_\pi^{n+1}\|_{0,\Omega}^2 + \|Du(t_{n+1}) - \partial_t u(t_{n+1})\|_{0,\Omega}^2.\end{aligned}$$

Thus, after multiplication by $2\delta t$ and summation over $0 \leq m \leq n-1$, we have

$$\begin{aligned}&\sum_{m=0}^{n-1}\delta t\|D\theta_h^{m+1}\|_{0,\Omega}^2 + |||(\theta_h^n, y_h^n)|||_h^2\\ &\leq \Big|\Big|\Big|(\theta_h^0, y_h^0)\Big|\Big|\Big|_h^2 + C\sum_{m=0}^{n-1}\delta t\Big(\|D\theta_\pi^{m+1}\|_{0,\Omega}^2 + \|Du(t_{m+1}) - \partial_t u(t_{m+1})\|_{0,\Omega}^2\Big).\end{aligned}$$

For the initial term, we use the linearity of the Ritz-projection and its approximation properties

$$\theta_h^0 = u_h^0 - P_h(u(0), \lambda(0)) = P_h(u(0), 0) - P_h(u(0), \lambda(0))$$
$$= P_h(u(0), 0) - P_h(u(0), 0) - P_h(0, \lambda(0)) = -P_h(0, \lambda(0)),$$

$$y_h^0 = \lambda_h^0 - R_h(u(0), \lambda(0)) = R_h(u(0), 0) - R_h(u(0), 0) - R_h(0, \lambda(0)) = -R_h(0, \lambda(0)).$$

Hence

$$\left|\!\left|\!\left|(\theta_h^0, y_h^0)\right|\!\right|\!\right|_h = |||(P_h(0, \lambda(0)), R_h(0, \lambda(0)))|||_h.$$

We have

$$|P_h(0, \lambda(0))|_{1,\Omega} = |P_h(0, \lambda(0)) - 0|_{1,\Omega} \leq Ch\Big(|0|_{2,\Omega} + \|\lambda(0)\|_{\mathcal{H}}\Big) = Ch\|\lambda(0)\|_{\mathcal{H}},$$

and

$$\begin{aligned}
&j(R_h(0, \lambda(0)), R_h(0, \lambda(0))) \\
&= j(R_h(0, \lambda(0)) - \lambda(0), R_h(0, \lambda(0))) + j(\lambda(0), R_h(0, \lambda(0))) \\
&\leq \Big(j(R_h(0, \lambda(0)) - \lambda(0)), R_h(0, \lambda(0)) - \lambda(0)) \\
&+ j(\lambda(0), \lambda(0))\Big)^{\frac{1}{2}} \Big(j(R_h(0, \lambda(0)), R_h(0, \lambda(0))\Big)^{\frac{1}{2}} \\
&\leq Ch\|\lambda(0)\|_{\mathcal{H}}.
\end{aligned}$$

Thus, we obtain

$$\left|\!\left|\!\left|(\theta_h^0, y_h^0)\right|\!\right|\!\right|_h^2 = |||(P_h(0, \lambda(0)), R_h(0, \lambda(0)))|||_h^2 \leq Ch^2\|\lambda(0)\|_{\mathcal{H}}^2.$$

Therefore, using (3.47) and (3.48), we have

$$\begin{aligned}
&\sum_{m=0}^{n-1} \delta t\|D\theta_h^{m+1}\|_{0,\Omega}^2 + |||(\theta_h^n, y_h^n)|||_h^2 \\
&\leq C\Big(h^2\|\lambda(0)\|_{\mathcal{H}}^2 + \delta t^2\|\partial_{tt} u\|_{L^2(0,t_n;L^2(\Omega))}^2 + h^2\delta t^2\big(\|\partial_t u\|_{L^2(0,t_n;H^2(\Omega))}^2 \\
&+ \|\partial_t \lambda\|_{L^2(0,t_n;\mathcal{H})}^2\big)\Big)
\end{aligned}$$

for $1 \leq n \leq N$. This proves (3.67), and then the proof is finished. □

3.5. Numerical studies

In this section, we will report the results of numerical experiments that support the analytical results of Section 3.3 and Section 3.4. We present computations demonstrating the optimal convergence using the space discretisation and the time discretisation defined in Section 3.2. We also verify numerically that the choice of the initial condition $u_h^0 = P_h(u_0, 0)$ guarantees a uniform approximation of λ, thus confirming the results in Theorem 3.8 and Theorem 3.11. It is also discussed that for other choices of discrete initial condition. The results for small time steps are not stable in the sense that the error explodes. All computations have been performed using $FreeFem++$ [Hec12].

We consider problem (3.1) with $\omega = \{(x, y) \in \mathbb{R}^2 : x^2+y^2 < 1\}$, and its reformulation (3.2) using $\Omega = (-2.4, 4) \times (-2, 2)$. We have taken $T = 1$.

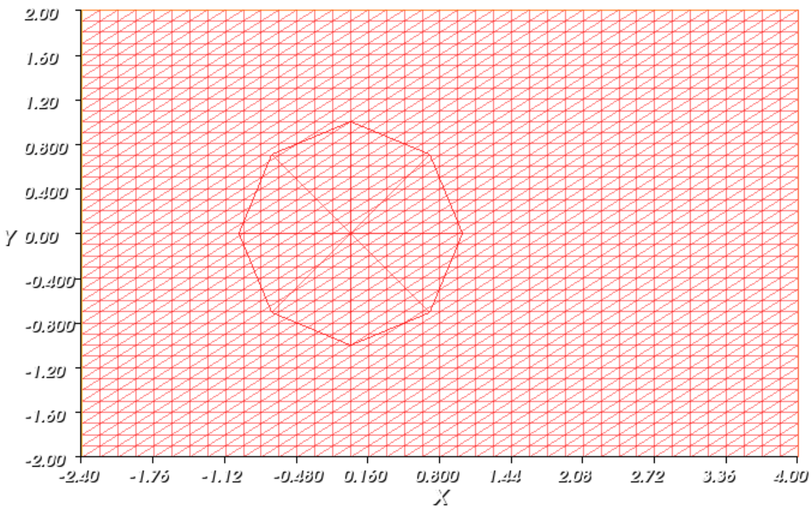


Fig. 3.2. Meshes for inf-sup stable case when $n = 1$.

We have tested two examples with known analytical solution. We first choose f in $\Omega \times (0, T)$ such that the exact solution of (3.2) is given by

$$u_1(x, t) = e^t(x^2 + y^2 - 1)$$

thus giving $g = 0$ on $\gamma \times (0, T)$. Also, we have chosen another case where $g \neq 0$ on γ given by

$$u_2(x, t) = e^t(x^2 + xy).$$

We have discretised this problem using a sequence of uniform meshes. The starting point is the mesh depicted in Fig. 3.2. To build the meshes, a parameter n is given (n denotes an integer parameter which refines the meshes). Then Ω is divided horizontally into $20 \cdot 2^n$ segments and vertically into $20 \cdot 2^n$ segments. The resulting quadrilateral mesh is then divided into triangles to form the mesh in Fig. 3.2 (where $n = 1$ is depicted). To build the mesh on γ, (i.e., $\gamma_{\tilde{h}}$), we divide γ into $4 \cdot 2^n$ curved segments. The pair of meshes $\mathcal{T}_h \times \gamma_{\tilde{h}}$ is used to implement the inf-sup stable method (3.8). For the stabilised method (3.9) we use $\mathcal{T}_h \times \gamma_h$, where the mesh γ_h is obtained after dividing each segment $\tilde{e}$ of $\gamma_{\tilde{h}}$ into 4 equally spaced curved segments.

For both test problems we have $\lambda = 0$. To measure convergence, we have computed the norms

$$\|u - u_{\tilde{h},\delta t}\|_*^2 := \|u_h^N - u(t_N)\|_{0,\Omega}^2 + \sum_{m=0}^{N-1} \delta t |u_h^{m+1} - u(t_{m+1})|_{1,\Omega}^2,$$

$$\|\lambda - \lambda_{\tilde{h},\delta t}\|_{**}^2 := \sum_{m=0}^{N-1} \delta t \|\lambda_{\tilde{h}}^{m+1} - \lambda(t_{m+1})\|_{0,\gamma}^2,$$

in the case of (3.8), and

$$\|(u - u_{h,\delta t}, \lambda - \lambda_{h,\delta t})\|_+^2 := \|u_h^N - u(t_N)\|_{0,\Omega}^2 + \sum_{m=0}^{N-1} \delta t |||(u_h^{m+1} - u(t_{m+1}), \lambda_h^{m+1} - \lambda(t_{m+1}))|||_h^2,$$

$$\|\lambda - \lambda_{h,\delta t}\|_{++}^2 := \sum_{m=0}^{N-1} \delta t \|\lambda_h^{m+1} - \lambda(t_{m+1})\|_{0,\gamma}^2,$$

for the stabilised method (3.9). Recall the error estimates $\|error\| \leq C(h + \delta t)$. So, to balance both terms, we have chosen $\delta t = h$ in the experiments. The numerical results are depicted in Figs. 3.3- 3.6 and measured the behavior of the corresponding norm with respect to h. We see that all norms go to zero as predicted by the theory.

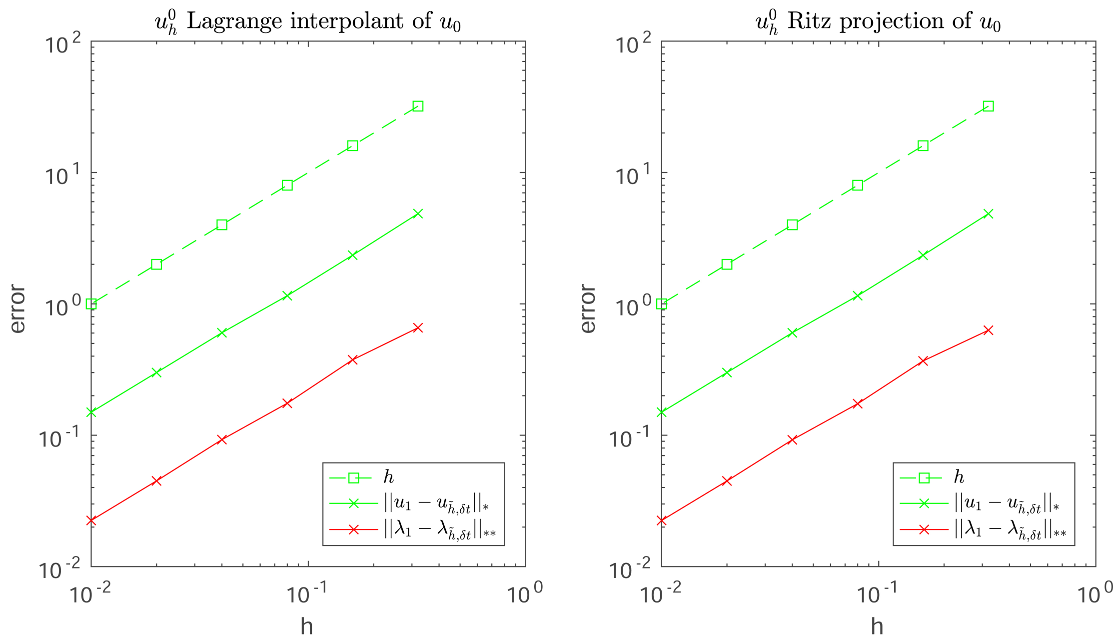


Fig. 3.3. Error of u and λ for inf-sup stable case (Theorem 3.8) for homogeneous BCs.

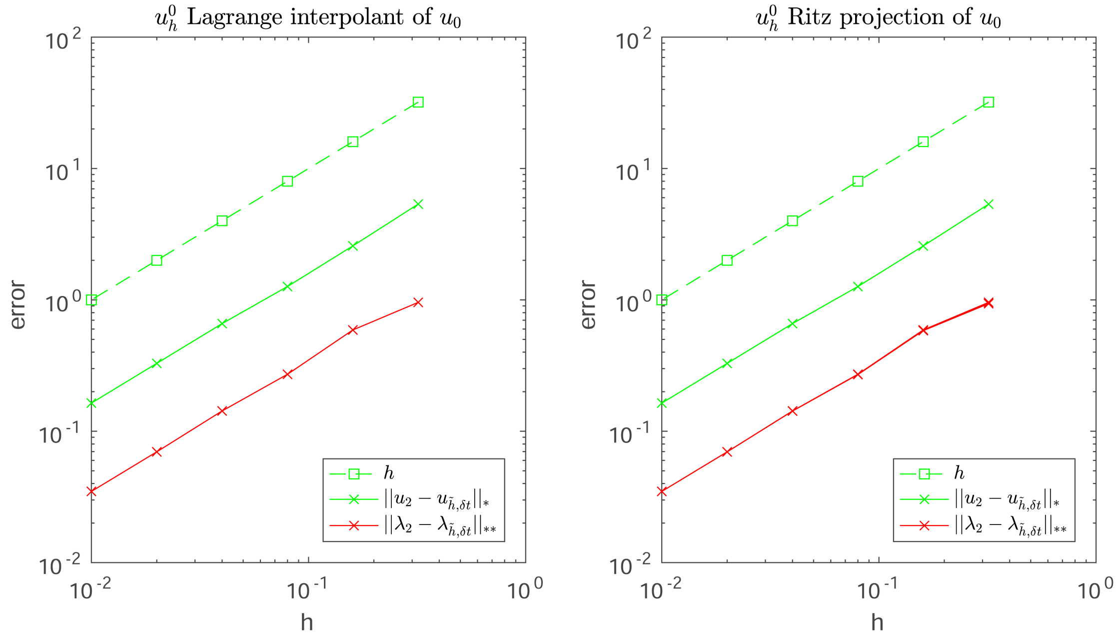


Fig. 3.4. Error of u and λ for inf-sup stable case (Theorem 3.8) for nonhomogeneous BCs.

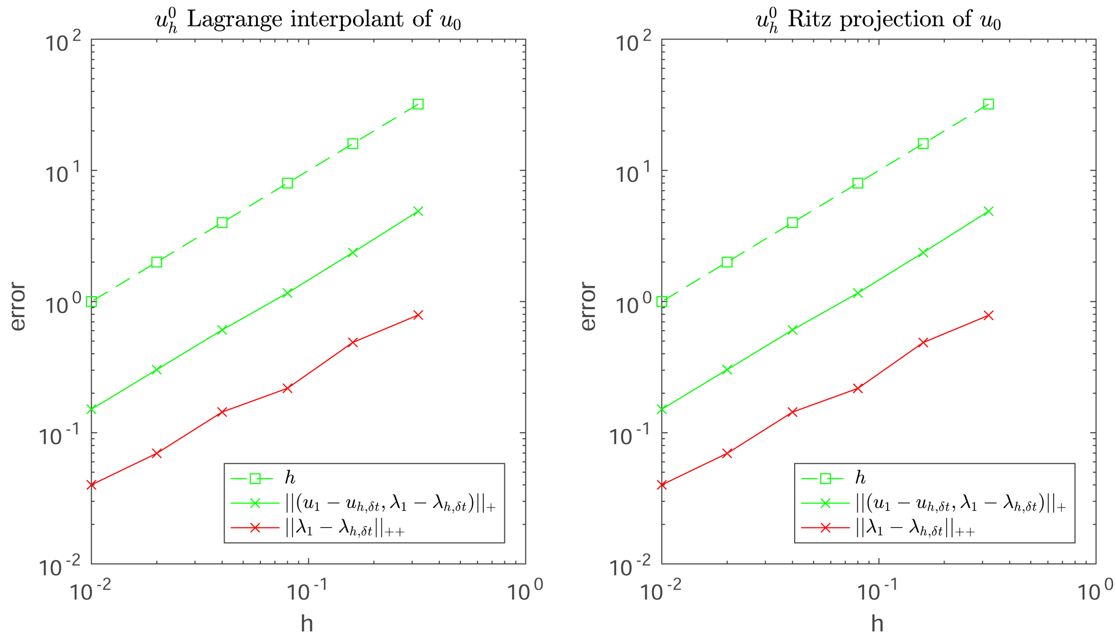


Fig. 3.5. Error of u and λ for stabilised method (Theorem 3.11) for homogeneous BCs.

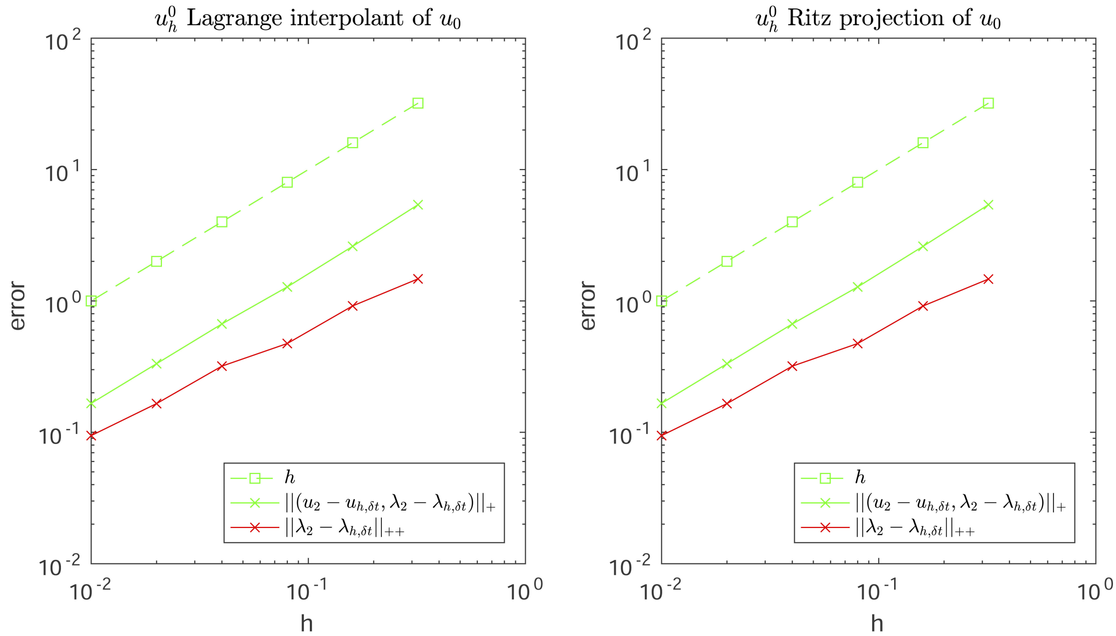


Fig. 3.6. Error of u and λ for stabilised method (Theorem 3.11) for nonhomogeneous BCs.

The results depicted on the right-hand side of the above figures do not contradict the theory. To stress this fact, we have performed a numerical experiment reminiscent of the one from [JN15]. This is, we have fixed one level $n = 1$ and have taken $\delta t \longrightarrow 0$. We have then measured $\|\lambda(\delta t) - \lambda_h^1\|_{0,\gamma}$ both considering $u_h^0 = i_h(u_0)$ and $u_h^0 = P_h(u_0, 0)$ defined before. The results, depicted in Figs. 3.7-3.10 show that, unless the initial condition is chosen appropriately, the error in λ grows as $\delta t \longrightarrow 0$, i. e. the approximation of λ cannot be guaranteed. These results confirm the sharpness of the stability results presented in Section 3.3.

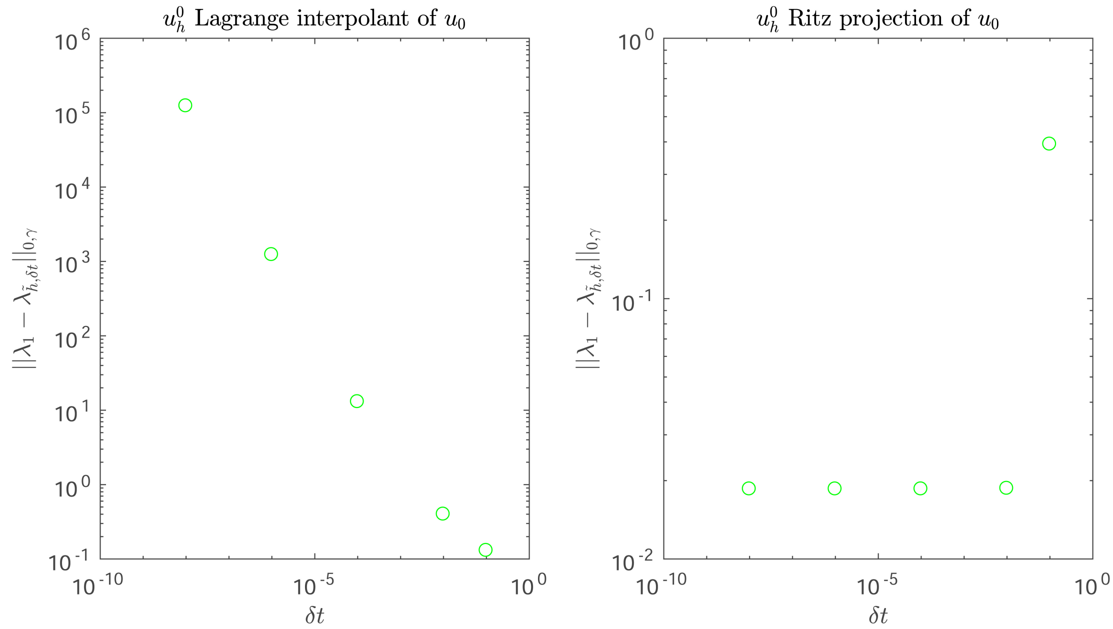


Fig. 3.7. Lambda behavior for inf-sup stable case for homogeneous BCs.

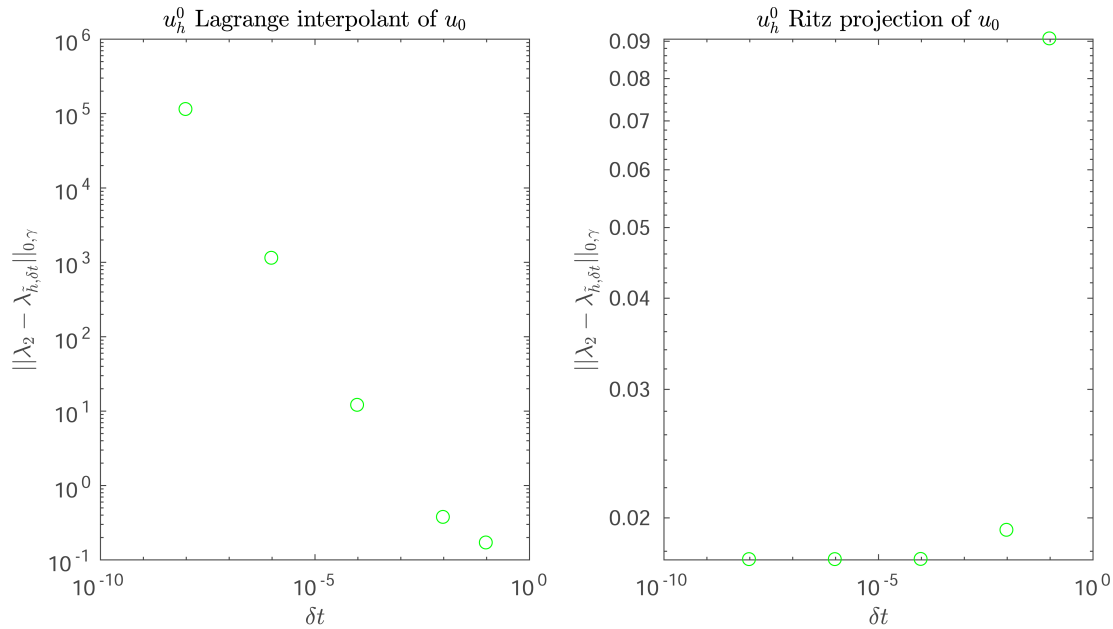


Fig. 3.8. Lambda behavior for inf-sup stable case for nonhomogeneous BCs.

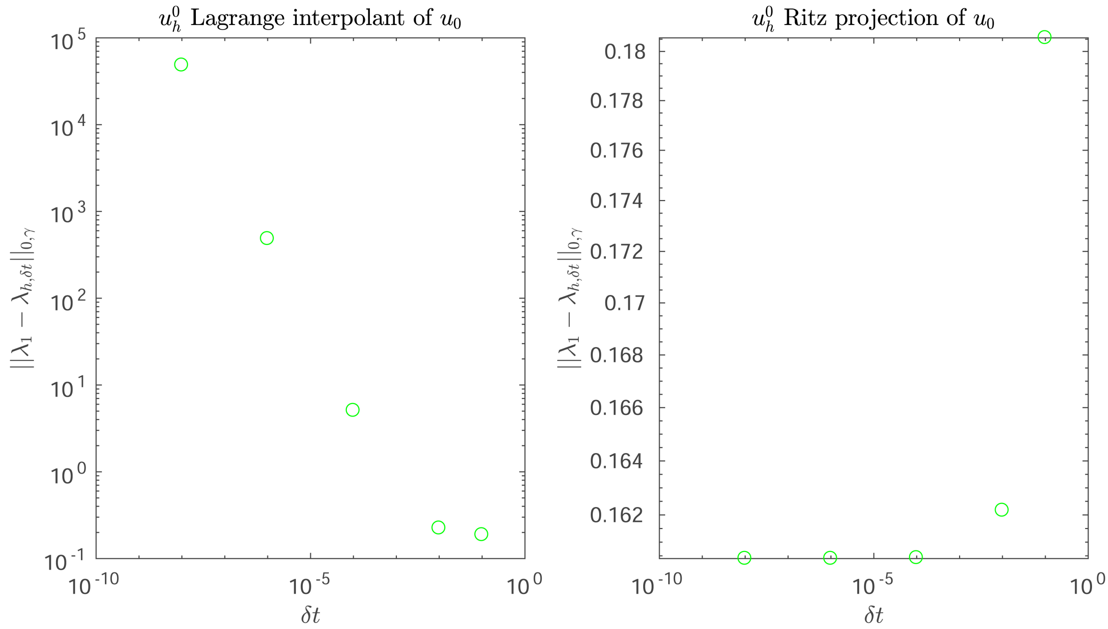


Fig. 3.9. Lambda behavior for stabilised method for homogeneous BCs.

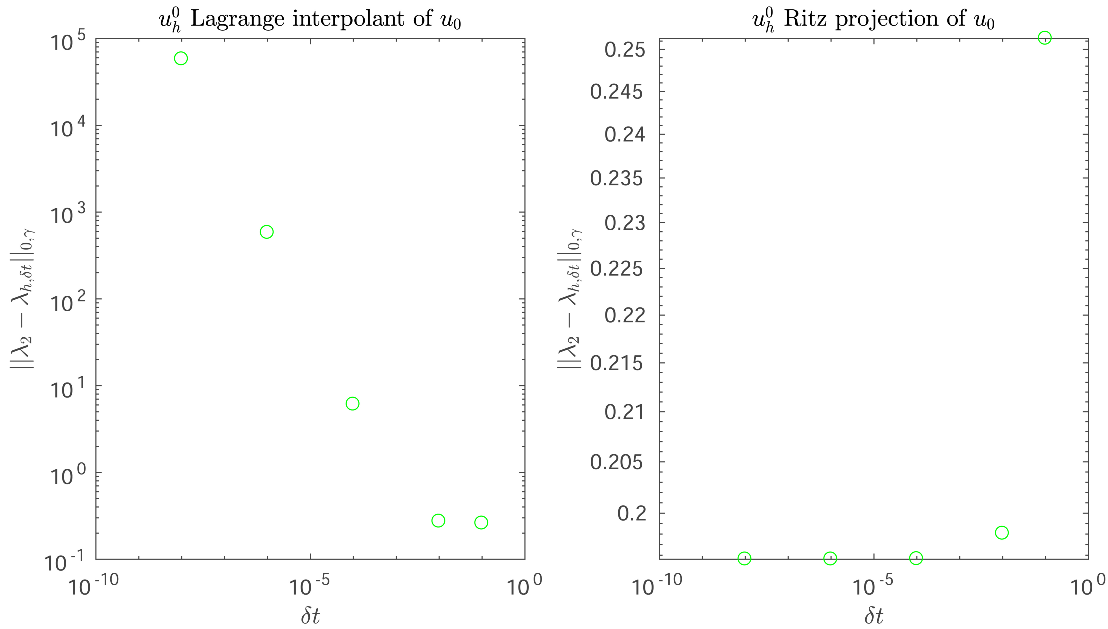


Fig. 3.10. Lambda behavior for stabilised method for nonhomogeneous BCs.

Chapter 4

A stabilised finite element method for a fictitious domain problem allowing small inclusions

4.1. Introduction

In this chapter, we extend the fictitious domain approach to the situation in which a domain contains small holes, or inclusions, that cannot be resolved by the finite element mesh. We consider a problem in such domain restricted to Dirichlet boundary conditions which is of importance. Another type of boundary conditions, e.g., Neumann boundary conditions, would lead to a different approach. Nevertheless, our motivation is to use a method like the one proposed in this chapter to approximate a problem like (4.1), but in incompressible fluid mechanics, i.e., solving Stokes, or even Navier-Stokes, equation. In such a case, Dirichlet conditions posed on each one of the perforations are the typical ones.

The situation in which the geometrical details of the domain are of a smaller size than the finite element mesh is not considered in the classical theory of fictitious domain method. In fact, in [GG95] the condition $|\tilde{e}| \geq 3h$ was imposed in order to prove stability and error estimates where $\tilde{e}$ denote the mesh on γ and h is the mesh size of the mesh in Ω. This last condition was relaxed in [BC12] by using a stabilised FEM, but the need for an auxiliary space satisfying $|\tilde{e}| \geq 3h$ was still needed for convergence. Then, this work presents an alternative to both these approaches by introducing a different stabilisation term, motivated by the method of Barbosa-Hughes (cf. [BH91], see also [HR09] for an application of this idea in the context of fictitious domain using cut elements).

The remainder of this chapter is organised as follows. In Section 4.2, we present the problem under consideration and useful notation. In Section 4.3, we study the stability of the stabilised problem, its convergence in Section 4.4 and we show the numerical results in Section 4.5. In Section 4.6, we extend the work done in the previous sections to the problem depending on time where the numerical results are also shown.

4.2. Problem setting

Let $\Omega \subseteq \mathbb{R}^2$ be an open bounded fictitious domain which contains the initial physical domain $\omega := \Omega \setminus \bigcup_{i=1}^{M} B_i$ where $M \in \mathbb{N}$. Each domain B_i is a closed, simply connected domain that can be, a priori, of any shape and size with $\gamma_i := \partial B_i$ (see Fig. 4.1 for a typical case).

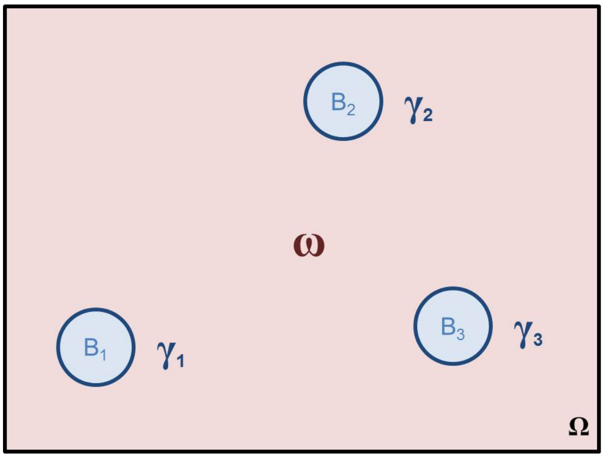


Fig. 4.1. Physical domain ω, fictitious domain Ω and inclusions B_i.

The problem of interest reads as follows:
find $\hat{u} : \omega \longrightarrow \mathbb{R}$ such that

$$\begin{cases} -\Delta \hat{u} &= \hat{f} \quad \text{in} \quad \omega, \\ \hat{u} &= g_i \quad \text{on} \quad \gamma_i, \quad i = 1, ..., M, \\ u &= 0 \quad \text{on} \quad \partial\Omega, \end{cases} \tag{4.1}$$

where $\hat{f} \in L^2(\omega)$ and $g_i \in H^{\frac{1}{2}}(\gamma_i)$ for all $i = 1, ..., M$. To present the fictitious domain method we introduce an extension f of $\hat{f}$ to Ω, and the solution of the following mixed problem:

find $(u, \boldsymbol{\lambda}) \in W := H_0^1(\Omega) \times \prod_{i=1}^{M} H^{-\frac{1}{2}}(\gamma_i)$, where $\boldsymbol{\lambda} = (\lambda_i)_{i=1}^{M}$, such that

$$\begin{cases} (\nabla u, \nabla v)_\Omega - \sum_{i=1}^{M} \langle \lambda_i, v \rangle_{\gamma_i} &= (f, v)_\Omega, \\ \sum_{i=1}^{M} \langle \mu_i, u \rangle_{\gamma_i} &= \sum_{i=1}^{M} \langle \mu_i, g_i \rangle_{\gamma_i}, \end{cases} \tag{4.2}$$

for all $(v, \boldsymbol{\mu}) \in W, \boldsymbol{\mu} = (\mu_i)_{i=1}^M$. Problems (4.1) and (4.2) are linked by the fact that if $(u, \boldsymbol{\lambda})$ satisfies (4.2), then $u|_\omega$ satisfies (4.1) and the Lagrange multipliers λ_i satisfy $\lambda_i = [\![\partial_\mathbf{n} u]\!]_{\gamma_i}$, for $i = 1, ..., M$, where $[\![v]\!]_{\gamma_i}$ stands for the jump of a function v across γ_i (see [GG95] for details).

To solve this weak problem, we introduce $\mathcal{T}_h$, a regular triangulation of $\overline{\Omega}$ built using triangles K with diameter h_K, and $h := \max_{K \in \mathcal{T}_h} h_K$. Additionally, each γ_i is partitioned into a different mesh $\gamma_{i,\tilde{h}}$ with curved edges $\tilde{e}$, where $\tilde{h} := \max |\tilde{e}|$. We are interested in the case in which $diam(B_i) << diam(\Omega)$ (and then $|\tilde{e}| < h$), which is precisely the case not allowed in [GG95]. Associated to these partitions, we define the following finite element spaces:

$$V_h = \{v_h \in C^0(\bar{\Omega}) : v_h|_k \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\},$$
$$\Lambda_{i,\tilde{h}} = \{\mu_{\tilde{h}} \in L^2(\gamma_i) : \mu_{\tilde{h}}|_{\tilde{e}} \in \mathbb{P}_0(\tilde{e}), \forall \tilde{e} \in \gamma_{i,\tilde{h}}\} \quad \text{for} \quad i = 1, ..., M,$$
$$\boldsymbol{\Lambda}_{\tilde{h}} = \prod_{i=1}^M \Lambda_{i,\tilde{h}}.$$

We now give a little bit more insight on the assumption on the geometry of γ_i. In fact, the case we are interested in is depicted in Fig. 4.2, where K and K' are in the triangulation $\mathcal{T}_h$ of the domain Ω defined above and γ_i is a circle.

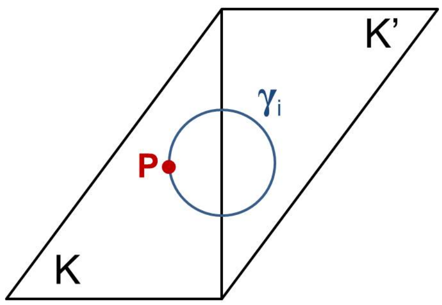


Fig. 4.2. A typical situation in which the inclusion B_i is a circle.

This implies that, for almost every point on γ_i, say P, we can be sure the jump of the normal derivative of u is zero. Namely, $v_h \in V_h$ so $v_h|_K \in \mathbb{P}_1(K)$. Hence $\nabla v_h|_K \in \mathbb{R}^2$ which implies $[\![\nabla v_h \cdot \mathbf{n}]\!]_{\gamma \cap K} = 0$.

4.3. The stabilised formulation and its stability

As it has been shown in [GG95], the inf-sup condition would only be valid if $|\tilde{e}| \geq 3h$. Due to our assumptions on B_i (i.e. $h > diam(B_i)$ for all $i = 1, ..., M$), this situation cannot occur. Then, we now propose the following stabilised FEM to approximate (4.2) using $V_h \times \mathbf{\Lambda}_{\tilde{h}}$ as finite element space:

find $(u_h, \boldsymbol{\lambda}_{\tilde{h}}) \in W_h := V_h \times \mathbf{\Lambda}_{\tilde{h}}$ such that

$$\mathbf{B}\left((u_h, \boldsymbol{\lambda}_{\tilde{h}}), (v_h, \boldsymbol{\mu}_{\tilde{h}})\right) = (f, v_h)_\Omega - \sum_{i=1}^{M} \langle \mu_{i,\tilde{h}}, g_i \rangle_{\gamma_i} \quad \forall (v_h, \boldsymbol{\mu}_{\tilde{h}}) \in W_h, \tag{4.3}$$

where

$$\mathbf{B}\left((u_h, \boldsymbol{\lambda}_{\tilde{h}}), (v_h, \boldsymbol{\mu}_{\tilde{h}})\right) = (\nabla u_h, \nabla v_h)_\Omega - \sum_{i=1}^{M} \langle \lambda_{i,\tilde{h}}, v_h \rangle_{\gamma_i} - \sum_{i=1}^{M} \langle \mu_{i,\tilde{h}}, u_h \rangle_{\gamma_i}$$
$$- \sum_{i=1}^{M} h \langle \lambda_{i,\tilde{h}}, \mu_{i,\tilde{h}} \rangle_{\gamma_i}. \tag{4.4}$$

We define the following norm

$$\|(v_h, \boldsymbol{\mu}_{\tilde{h}})\|_{W_h}^2 := |v_h|_{1,\Omega}^2 + \sum_{i=1}^{M} h \|\mu_{i,\tilde{h}}\|_{0,\gamma_i}^2,$$

and prove next the stability of the method.

Theorem 4.1. *For all $(v_h, \boldsymbol{\mu}_{\tilde{h}}) \in W_h$, the following holds*

$$\boldsymbol{B}\left((v_h, \boldsymbol{\mu}_{\tilde{h}}), (v_h, -\boldsymbol{\mu}_{\tilde{h}})\right) = \|(v_h, \boldsymbol{\mu}_{\tilde{h}})\|_{W_h}^2.$$

Hence, problem (4.3) *is well-posed.*

Proof. It follows directly from the definition of the bilinear form $\mathbf{B}$. □

The next result states the consistency of the method for smooth solutions.

Lemma 4.2. *Let $(u, \boldsymbol{\lambda})$ be the solution of* (4.2) *and $(u_h, \boldsymbol{\lambda}_{\tilde{h}}) \in W_h$ be the solution of* (4.3)*. Then*

$$\boldsymbol{B}\left((u - u_h, \boldsymbol{\lambda} - \boldsymbol{\lambda}_{\tilde{h}}), (v_h, \boldsymbol{\mu}_{\tilde{h}})\right) = -\sum_{i=1}^{M} h \langle \lambda_i, \mu_{i,\tilde{h}} \rangle_{\gamma_i}, \tag{4.5}$$

for all $(v_h, \boldsymbol{\mu}_{\tilde{h}}) \in W_h$. Moreover, if $u \in H^{\frac{3}{2}+\epsilon}(\Omega)$ for some $\epsilon > 0$, then

$$\boldsymbol{B}\left((u - u_h, \boldsymbol{\lambda} - \boldsymbol{\lambda}_{\tilde{h}}), (v_h, \boldsymbol{\mu}_{\tilde{h}})\right) = 0, \tag{4.6}$$

for all $(v_h, \boldsymbol{\mu}_{\tilde{h}}) \in W_h$.

Proof. From the definition of $\mathbf{B}$, and the fact that $(u, \boldsymbol{\lambda})$ solves (4.2) and $(u_h, \boldsymbol{\lambda}_{\tilde{h}})$ solves (4.3), it follows that

$$\begin{aligned}\mathbf{B}\left((u,\boldsymbol{\lambda}),(v_h,\boldsymbol{\mu}_{\tilde{h}})\right) &= (\nabla u, \nabla v_h)_\Omega - \sum_{i=1}^{M}\langle \lambda_i, v_h\rangle_{\gamma_i} - \sum_{i=1}^{M}\langle \mu_{i,\tilde{h}}, u\rangle_{\gamma_i} - \sum_{i=1}^{M} h\langle \lambda_i, \mu_{i,\tilde{h}}\rangle_{\gamma_i} \\ &= (f, v_h)_\Omega - \sum_{i=1}^{M}\langle \mu_{i,\tilde{h}}, g_i\rangle_{\gamma_i} - \sum_{i=1}^{M} h\langle \lambda_i, \mu_{i,\tilde{h}}\rangle_{\gamma_i} \\ &= \mathbf{B}\left((u_h,\boldsymbol{\lambda}_{\tilde{h}}),(v_h,\boldsymbol{\mu}_{\tilde{h}})\right) - \sum_{i=1}^{M} h\langle \lambda_i, \mu_{i,\tilde{h}}\rangle_{\gamma_i},\end{aligned}$$

which proves (4.5). Now, if $u \in H^{\frac{3}{2}+\epsilon}(\Omega)$, then $\lambda_i = [\![\partial_{\mathbf{n}} u]\!]_{\gamma_i} = 0$, for $i = 1, ..., M$ and we get (4.6) directly from (4.5). $\square$

4.4. Error analysis

We start this section by making two assumptions. First, we will suppose that the inclusions B_i satisfy $dist(B_i, \partial\Omega) \geq \frac{h}{2}$. This assumption will be needed on what follows, and, although it may seem restrictive, the factor $\frac{1}{2}$ can be relaxed to any positive constant, as long as it is fixed. We will also assume, just to simplify the presentation, that the inclusions B_i are convex. This latter hypothesis is made only for simplicity, the same results are valid, up to minor modifications, if this does not hold.

The next result is a first technical tool, the local trace inequality, that will be used in the error analysis. The proof of this result is similar to the one given in [CGS16] for the case of curved elements.

Lemma 4.3. *Let $\tilde{B}_i$ be the local annular neighborhood defined as $\tilde{B}_i := \left\{\boldsymbol{x} \in \omega : dist(\boldsymbol{x}, \gamma_i) \leq \frac{h}{2}\right\}$, for every $i = 1, ..., M$. Then, the following local trace inequality holds*

$$\|v\|_{0,\gamma_i}^2 \leq 8h^{-1}\|v\|_{0,\tilde{B}_i}^2 + 8\|v\|_{0,\tilde{B}_i}\|\nabla v\|_{0,\tilde{B}_i}, \tag{4.7}$$

for every $v \in H^1(\Omega)$.

Proof. Let $v \in H^1(\Omega)$. We split $\gamma_i = \bigcup_{j=1}^{R}\overline{\gamma_i^j}$, where the γ_i^j are disjoint, and introduce a collection of points $\mathbf{x}_i^j \in \Omega \setminus B_i$ (see Fig. 4.3) such that

$$dist(\mathbf{x}_i^j, \gamma_i^j) = \frac{1}{2}h \quad \text{and} \quad (\mathbf{x} - \mathbf{x}_i^j)\cdot \mathbf{n}|_{\gamma_i} \geq \frac{1}{4}h, \tag{4.8}$$

for all $\mathbf{x} \in \gamma_i^j$. Using these points, we build the curved triangle T_{ij} as depicted in Fig. 4.3.

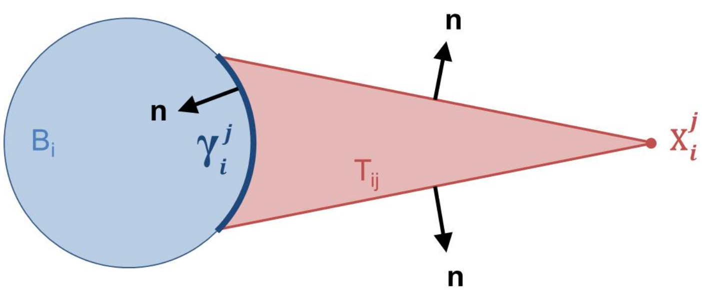


Fig. 4.3. The curved triangle T_{ij} used in the proof of Lemma 4.3.

Let $\mathbf{m}(\mathbf{x}) = \mathbf{x} - \mathbf{x}_i^j$. Then $\nabla \cdot \mathbf{m} = 2$, and $\mathbf{m} \cdot \mathbf{n} = 0$ on $\partial T_{ij} \setminus \gamma_i^j$. Applying Green's theorem we get,

$$(\mathbf{m} \cdot \mathbf{n}, v^2)_{\gamma_i^j} = (\mathbf{m} \cdot \mathbf{n}, v^2)_{\partial T_{ij}} = (\nabla(v^2 \cdot \mathbf{m}), 1)_{T_{ij}} = 2\|v\|_{0,T_{ij}}^2 + 2(v, \nabla v \cdot \mathbf{m})_{T_{ij}}.$$

On the other hand, from (4.8) we get $\mathbf{m} \cdot \mathbf{n}|_{\gamma_i^j} \geq \frac{1}{4}h$. Then, applying Hölder's inequality to the last expresion, and supposing there R is large enough to have $\|\mathbf{m}\|_{\infty,T_{ij}} \leq h$, we arrive at

$$\begin{aligned}\frac{1}{4}h\|v\|_{0,\gamma_i^j}^2 \leq (\mathbf{m} \cdot \mathbf{n}, v^2)_{\gamma_i^j} &= 2\|v\|_{0,T_{ij}}^2 + 2(v, \nabla v \cdot \mathbf{m})_{T_{ij}} \\ &\leq 2\|v\|_{0,T_{ij}}^2 + 2\|v\|_{0,T_{ij}}\|\nabla v\|_{0,T_{ij}}\|\mathbf{m}\|_{\infty,T_{ij}} \\ &\leq 2\|v\|_{0,T_{ij}}^2 + 2h\|v\|_{0,T_{ij}}\|\nabla v\|_{0,T_{ij}}.\end{aligned}$$

Then, the following holds

$$\|v\|_{0,\gamma_i^j}^2 \leq 8h^{-1}\|v\|_{0,T_{ij}}^2 + 8\|v\|_{0,T_{ij}}\|\nabla v\|_{0,T_{ij}},$$

for each γ_i^j. Hence, adding over $j = 1, ..., R$, using that the curved triangles T_{ij} are disjoint, $\tilde{B}_i \supset \bigcup_{j=1}^{R} T_{ij}$ and applying Cauchy-Schwarz's inequality, we obtain

$$\begin{aligned}\|v\|_{0,\gamma_i}^2 &= \sum_{j=1}^{R} \|v\|_{0,\gamma_i^j}^2 \leq 8h^{-1}\sum_{j=1}^{R} \|v\|_{0,T_{ij}}^2 + 8\sum_{j=1}^{R} \|v\|_{0,T_{ij}}\|\nabla v\|_{0,T_{ij}} \\ &\leq 8h^{-1}\|v\|_{0,\tilde{B}_i}^2 + 8\|v\|_{0,\tilde{B}_i}\|\nabla v\|_{0,\tilde{B}_i}.\end{aligned}$$

This finishes the proof. □

Remark 4.4. *More precisely, if the inclusion is not convex, then the set of points x_i^j defined in the proof of Lemma 4.3 would need to be chosen in such a way that the curved triangles T_{ij} are convex. Moreover, the convexity assumption ensures that the curved* *triangles T_{ij} are disjoint. If B_i is not convex, then this is no longer time, and the constant in* (4.7) *will change to include maximun number of non-empty intersections.*

In order to prove the error estimate, we split the error into interpolation and discrete errors as follows: $(e^u, e^{\boldsymbol{\lambda}}) := (u-u_h, \boldsymbol{\lambda}-\boldsymbol{\lambda}_{\tilde{h}}) = (u-i_h u, \boldsymbol{\lambda}-\Pi_{\tilde{h}}\boldsymbol{\lambda})+(i_h u-u_h, \Pi_{\tilde{h}}\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\tilde{h}}) =: (\eta^u, \eta^{\boldsymbol{\lambda}}) - (e^u_h, e^{\boldsymbol{\lambda}}_{\tilde{h}})$. Here $i_h : C^0(\overline{\Omega}) \longrightarrow V_h$ stands for the Lagrange interpolation operator, and $\Pi_{\tilde{h}}\boldsymbol{\lambda} \in \Lambda_{i,\tilde{h}}$ is defined by $\Pi_{\tilde{h}}\boldsymbol{\lambda} = (\Pi_{\tilde{h}}\lambda_i)_{i=1}^N$ where $\Pi_{\tilde{h}}\lambda_i|_{\tilde{e}} := |\tilde{e}|^{-1}(\lambda_i, 1)_{\tilde{e}}$ for all $\tilde{e} \in \gamma_{i,\tilde{h}}$, and all $i = 1, ..., M$. We now state the main error estimate for the method (4.3).

Theorem 4.5. *Let us suppose there exits $\epsilon > 0$ such that $u \in H^{1+s}(\Omega)$, for $s \in [\frac{1}{2} - \epsilon, 2]$, and $\boldsymbol{\lambda} \in \prod_{i=1}^{N} H^{\delta}(\gamma_i)$, for $\delta \in [0, \frac{1}{2}]$. Then, there exits a constant $C > 0$, independent of h and $\tilde{h}$, such that*

$$\|(e^u, e^{\boldsymbol{\lambda}})\|_{W_h} \le C \left(h^s |u|_{1+s,\Omega} + h^{\frac{1}{2}+\delta} \Big(\sum_{i=1}^{M} \|\lambda_i\|^2_{\delta,\gamma_i} \Big)^{\frac{1}{2}} \right). \tag{4.9}$$

Proof. First, using standard interpolation estimates (see [EG04]) and $\tilde{h} \le h$, we obtain

$$\|(\eta^u, \eta^{\boldsymbol{\lambda}})\|_{W_h} \le C \left(h^s |u|_{1+s,\Omega} + h^{\frac{1}{2}+\delta} \Big(\sum_{i=1}^{M} \|\lambda_i\|^2_{\delta,\gamma_i} \Big)^{\frac{1}{2}} \right). \tag{4.10}$$

To bound the discrete error, we suppose $s > \frac{1}{2}$. Then, using Theorem 4.1 and Lemma 4.2, we arrive at

$$\begin{aligned} \|(e^u_h, e^{\boldsymbol{\lambda}}_{\tilde{h}})\|^2_{W_h} &= B\big((e^u_h, e^{\boldsymbol{\lambda}}_{\tilde{h}}), (e^u_h, -e^{\boldsymbol{\lambda}}_{\tilde{h}})\big) = B\big((\eta^u, \eta^{\boldsymbol{\lambda}}), (e^u_h, -e^{\boldsymbol{\lambda}}_{\tilde{h}})\big) \\ &= (\nabla \eta^u, \nabla e^u_h)_\Omega - \sum_{i=1}^{M} \langle \eta^{\lambda_i}, e^u_h \rangle_{\gamma_i} + \sum_{i=1}^{M} \langle e^{\lambda_i}_{\tilde{h}}, \eta^u \rangle_{\gamma_i} + \sum_{i=1}^{M} h \langle -\eta^{\lambda_i}, -e^{\lambda_i}_{\tilde{h}} \rangle_{\gamma_i} \\ &= I + II + III + IV. \end{aligned} \tag{4.11}$$

We now bound the above right-hand side term by term. The main arguments are the approximation properties of i_h and $\Pi_{\tilde{h}}$ (see [EG04]), the duality between $H^{-\frac{1}{2}}(\gamma_i)$ and $H^{\frac{1}{2}}(\gamma_i)$, the Trace theorem in each B_i, the local trace result from Lemma 4.3, and the Cauchy-Schwarz and Poincaré inequalities:

$$\begin{aligned} I &\le |\eta^u|_{1,\Omega} |e^u_h|_{1,\Omega} \\ &\le C h^{2s} |u|^2_{1+s,\Omega} + \frac{1}{4} |e^u_h|^2_{1,\Omega}, \end{aligned} \tag{4.12}$$

$$
\begin{aligned}
II &\leq \sum_{i=1}^{M} \|\eta^{\lambda_i}\|_{-\frac{1}{2},\gamma_i} \|e_h^u\|_{\frac{1}{2},\gamma_i} \\
&\leq \sum_{i=1}^{M} \|\eta^{\lambda_i}\|_{-\frac{1}{2},\gamma_i} \|e_h^u\|_{1,B_i} \\
&\leq C \sum_{i=1}^{M} h^{1+2\delta} \|\lambda_i\|_{\delta,\gamma_i}^2 + \frac{1}{4} |e_h^u|_{1,\Omega}^2, \qquad (4.13)
\end{aligned}
$$

$$
\begin{aligned}
III &\leq \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2 + \sum_{i=1}^{M} h^{-1} \|\eta^u\|_{0,\gamma_i}^2 \\
&\leq \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2 + \sum_{i=1}^{M} \left(8h^{-2} \|\eta^u\|_{0,\tilde{B}_i}^2 + 4h^{-1} \|\eta^u\|_{0,\tilde{B}_i} \|\nabla \eta^u\|_{0,\tilde{B}_i} \right) \\
&\leq \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2 + C h^{2s} |u|_{1+s,\Omega}^2, \qquad (4.14)
\end{aligned}
$$

$$
\begin{aligned}
IV &\leq \sum_{i=1}^{M} h \|\eta^{\lambda_i}\|_{0,\gamma_i}^2 + \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2 \\
&\leq C \sum_{i=1}^{M} h^{1+2\delta} \|\lambda_i\|_{\delta,\gamma_i}^2 + \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2. \qquad (4.15)
\end{aligned}
$$

Collecting then (4.11)-(4.15), we get

$$
\|(e_h^u, e_h^{\boldsymbol{\lambda}})\|_{W_h}^2 \leq C \left(h^{2s} |u|_{1+s,\Omega}^2 + h^{1+2\delta} \sum_{i=1}^{M} \|\lambda_i\|_{\delta,\gamma_i}^2 \right) + \frac{1}{2} \|(e_h^u, e_{\tilde{h}}^{\boldsymbol{\lambda}})\|_{W_h}^2,
$$

and (4.9) follows by rearranging terms and applying the triangle inequality and (4.10). Next, if $s \leq \frac{1}{2}$, proceeding as above, we get

$$
\|(e_h^u, e_{\tilde{h}}^{\boldsymbol{\lambda}})\|_{W_h}^2 \leq I + II + III + IV + \sum_{i=1}^{M} h \langle \lambda_i, e_{\tilde{h}}^{\lambda_i} \rangle_{\gamma_i}.
$$

The first four terms have already been bounded. The fifth is bounded as follows:

$$
\sum_{i=1}^{M} h \langle \lambda_i, e_{\tilde{h}}^{\lambda_i} \rangle_{\gamma_i} \leq \sum_{i=1}^{M} h \|\lambda_i\|_{0,\gamma_i} \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i} \leq 8 \sum_{i=1}^{M} h \|\lambda_i\|_{0,\gamma_i}^2 + \frac{1}{4} \sum_{i=1}^{M} h \|e_{\tilde{h}}^{\lambda_i}\|_{0,\gamma_i}^2.
$$

Then (4.9) follows rearranging terms, and applying the triangle inequality and (4.10).□

Remark 4.6. *Even if the proof of this result does not depend on M, the number of* *perforations, it is important to remark that, for a given mesh, the constant on the error estimate may degenerate with M. The reason for this lies in the bound for the third term in the above proof. In fact, the constant in the estimate of III depends on the overlap between the annular regions $\tilde{B}_i$. If the number of inclusions increases, then a larger number of these regions are expected to have non-empty intersections, thus increasing that constant. This problem is of no importance once the mesh is refined enough, since the annular regions $\tilde{B}_i$ become more and more separated, but this might be an issue when the mesh is coarse.*

Remark 4.7. *It is worth remarking that the method can be written in a completely consistent way, at least in the case in which all the γ_i are curved boundaries. As a matter of fact, in such a case we have $[\![\partial_{\boldsymbol{n}} v_h]\!]_{\gamma_i} = 0$ for all $v_h \in V_h$ and all $i = 1, ..., M$ (see Fig. 4.2 for a typical situation, as can be seen there, $\nabla v_h$is the same constant both sides of the curve γ_i, at almost every point of it). Then the bilinear form $\boldsymbol{B}$ can be rewritten as*

$$\boldsymbol{B}\big((u_h, \boldsymbol{\lambda}_{\tilde{h}}), (v_h, \boldsymbol{\mu}_{\tilde{h}})\big) = (\nabla u_h, \nabla v_h)_\Omega - \sum_{i=1}^{M} \langle \lambda_{i,\tilde{h}}, v_h \rangle_{\gamma_i} - \sum_{i=1}^{M} \langle \mu_{i,\tilde{h}}, u_h \rangle_{\gamma_i} + \sum_{i=1}^{M} h \langle [\![\partial_{\boldsymbol{n}} u_h]\!] - \lambda_{i,\tilde{h}}, [\![\partial_{\boldsymbol{n}} v_h]\!] + \mu_{i,\tilde{h}} \rangle_{\gamma_i}, \tag{4.16}$$

and then (4.6) follows using $\lambda_i - [\![\partial_{\boldsymbol{n}} u]\!]_{\gamma_i} = 0$ for all $i = 1, ..., M$. Looking at this last definition, the link to the Barbosa-Hughes method (and the method from [HR09]) is apparent. Nevertheless, we have preferred to keep the non-consistent presentation here since the defintion (4.16) leads to technical difficulties. More precisely, for this alternative writing of $\boldsymbol{B}$, Lemma 4.3 would have to be applied with respect to the interior of B_i, which would lead, ultimately, to error estimates that depend on diam$(B_i)^{-1}$.

4.5. Numerical studies

In this section, we report the results of numerical experiments that support the analysis carried out in previous sections. All computations have been performed using a code written in $FreeFem++$ [Hec12].

We consider the stabilised problem (4.3) with $\omega = \Omega \setminus \bigcup_{i=1}^{3} B_i$, where $\Omega = (0, 10)^2$, and $B_i = [\mathbf{c}_i, r]$, where $r > 0$. The centers of the balls B_i are $\mathbf{c}_1 = (1.7, 7.4)$, $\mathbf{c}_2 = (5.7, 8.4)$ and $\mathbf{c}_3 = (8.7, 3.4)$, respectively. To build the meshes, a parameter n is given. Then Ω is divided horizontally and vertically into $10n$ segments. The mesh on each γ_i is built by dividing each one of them into $8n$ curved segments. The resulting meshes, where

$n = 1$ and $r = 0.1$, are depicted in Fig. 4.4 and Fig. 4.5 is a zoom around B_1 of the resulting mesh where we can observe that the finite element mesh does not resolve the geometrical feature that is B_1. We can observe that $diam(B_i) \leq h$. More importantly, for all values of n, the discretisation parameter $|\tilde{e}|$ on γ_i is such that $|\tilde{e}| < h$, and then stabilisation is indeed needed. We remark that we have also implemented the unstabilised version of the method, and it has lead to singular matrices (in all the cases). Thus, there is a real need for stabilisation in a situation such as this one.

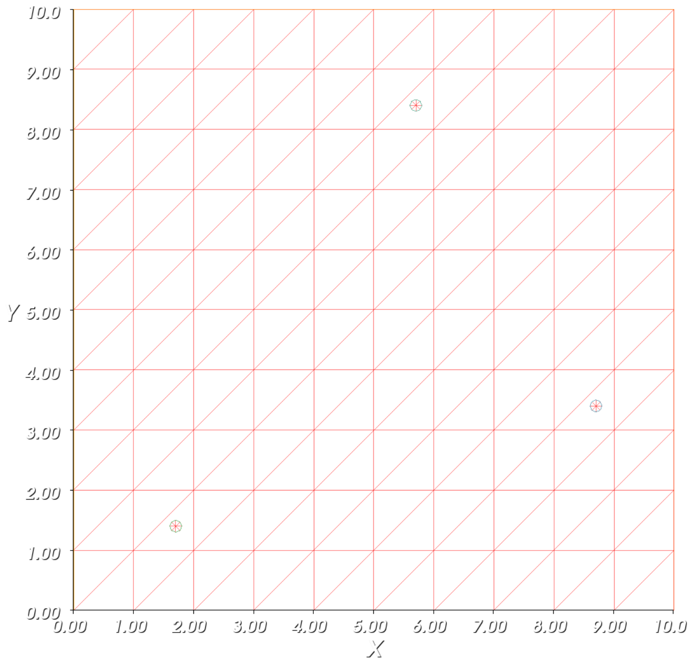


Fig. 4.4. Meshes when $n = 1$.

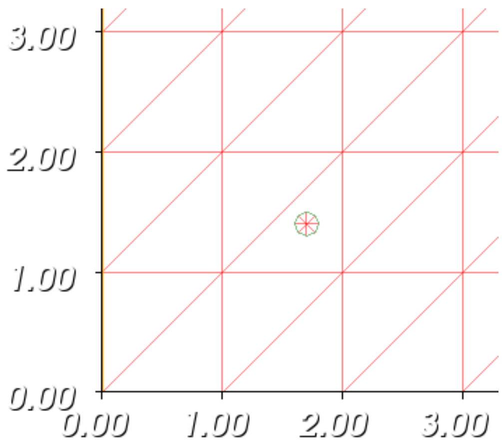


Fig. 4.5. A zoom of the computational mesh in a neighborhood of B_1 for $r = 0.1$ and $n = 1$. We can observe that the mesh does not resolve the inclusion.

We first build two examples in which we know the exact solution. We set the right-hand side f and the boundary conditions in such a way that the exact solutions are given by two different examples

$$u_1(x,y) = \sin(x)\sin(y) \quad \text{and} \quad u_2(x,y) = x^2(10-x)^2 y^2 (10-y)^2 \quad \text{in} \quad \Omega.$$

These two examples are built just to show that if an extension of f that makes the extended solution regular is available, then the method converges in an optimal way. In this sense, we can interpret this method as a way of imposing Dirichlet conditions on the interior of the domain weakly.

The results of the simulation for these first two examples are given in Tables 4.1, 4.2 and 4.3 (where we have taken $r = 0.2, r = 0.1$ and $r = 0.025$, respectively). We observe that, since $u \in H^2(\Omega)$ and $\boldsymbol{\lambda} = \mathbf{0} \in \prod_{i=1}^{M} H^1(\gamma_i)$, then the optimal order of convengerce $o(h)$ is obtained with an even better than optimal converge for $\boldsymbol{\lambda}$. This is most likely due to the fact $\boldsymbol{\lambda} = \mathbf{0}$, and is compatible with the results obtained in [BC12].

		$\|u-u_h\|_{1,\Omega}$	order	$\left(\sum_{i=1}^{3} h\|\lambda-\lambda_{\tilde{h}}\|^2_{0,\gamma_i}\right)^{\frac{1}{2}}$	order
$u=u_1(x,y)$	$n=1$	3.5808		0.2486	
	$n=2$	1.7996	0.9926	0.0691	1.8471
	$n=3$	1.1988	1.0019	0.0336	1.7783
	$n=4$	0.8989	1.0008	0.0180	2.1696
	$n=5$	0.7187	1.0026	0.0121	1.7799
	$n=6$	0.5988	1.0011	0.0086	1.8728
	$n=7$	0.5131	1.0020	0.0061	2.2282
	$n=8$	0.4489	1.0010	0.0047	1.9525
	$n=9$	0.3995	0.9898	0.0037	2.0311
	$n=10$	0.3592	1.0092	0.0030	1.9905
	$n=11$	0.3266	0.9982	0.0028	0.7239
	$n=12$	0.3001	0.9725	0.0023	2.2607
	$n=13$	0.2762	1.0368	0.0020	1.7461
	$n=14$	0.2563	1.0090	0.0018	1.4217
	$n=15$	0.2395	0.9826	0.001490	2.6426
	$n=16$	0.2244	1.0091	0.001466	0.2516
$u=u_2(x,y)$	$n=1$	189836		8013.27	
	$n=2$	92234.4	1.0414	2082.72	1.9439
	$n=3$	61106.9	1.0154	1047.71	1.6945
	$n=4$	45743.9	1.0065	587.443	2.0112
	$n=5$	36542.8	1.0064	433.107	1.3659
	$n=6$	30435.2	1.0031	282.196	2.3496
	$n=7$	26070.1	1.0043	213.894	1.7977
	$n=8$	22805	1.0021	159.669	2.1896
	$n=9$	20282.5	0.9952	130.767	1.6954
	$n=10$	18245.4	1.0046	112.195	1.4539
	$n=11$	16584	1.0017	101.369	1.0646
	$n=12$	15252.4	0.9620	84.1687	2.1370
	$n=13$	14030.7	1.0431	78.6999	0.8393
	$n=14$	13015.1	1.0139	69.2829	1.7197
	$n=15$	12160.4	0.9845	61.8237	1.6511
	$n=16$	11390.6	1.0133	54.8456	1.8557

Table 4.1. Finite element errors for the smooth examples and $r=0.2$.

		$\|u-u_h\|_{1,\Omega}$	order	$\left(\sum_{i=1}^{3} h\|\lambda-\lambda_{\tilde{h}}\|^2_{0,\gamma_i}\right)^{\frac{1}{2}}$	order
$u=u_1(x,y)$	$n=1$	3.5800		0.1986	
	$n=2$	1.7993	0.9925	0.0595	1.7389
	$n=3$	1.1986	1.0019	0.0256	2.0800
	$n=4$	0.8989	1.0002	0.0156	1.7218
	$n=5$	0.7187	1.0026	0.0113	1.4451
	$n=6$	0.5988	1.0011	0.0069	2.7056
	$n=7$	0.5131	1.0020	0.0057	1.2394
	$n=8$	0.4489	1.0010	0.0038	3.0365
	$n=9$	0.3994	0.9920	0.0032	1.4590
	$n=10$	0.3592	1.0069	0.0025	2.3430
	$n=11$	0.3265	1.0015	0.0022	1.3412
	$n=12$	0.3001	0.9690	0.0020	1.0954
	$n=13$	0.2763	1.0323	0.0016	2.7878
	$n=14$	0.2563	1.0139	0.0014	1.8018
	$n=15$	0.2395	0.9826	0.0012	2.2343
	$n=16$	0.2244	1.0091	0.0011	1.3482
$u=u_2(x,y)$	$n=1$	189830		5920.23	
	$n=2$	92230.1	1.0414	1538.89	1.9438
	$n=3$	61106.1	1.0153	731.663	1.8337
	$n=4$	45742.8	1.0066	439.74	1.7698
	$n=5$	36542.9	1.0063	337.51	1.1857
	$n=6$	30434.6	1.0032	226.654	2.1839
	$n=7$	26070.2	1.0041	187.245	1.2388
	$n=8$	22804.9	1.0021	127.731	2.8648
	$n=9$	20282.6	0.9952	103.571	1.7801
	$n=10$	18245.4	1.0046	76.8754	2.8291
	$n=11$	16584.1	1.0017	81.1063	-0.5621
	$n=12$	15252.4	0.9620	77.5004	0.5227
	$n=13$	14030.7	1.0431	56.0019	4.0591
	$n=14$	13015	1.0140	50.6272	1.3615
	$n=15$	12160.4	0.9844	43.6408	2.1523
	$n=16$	11390.6	1.0133	41.8489	0.6496

Table 4.2. Finite element errors for the smooth examples and $r=0.1$.

		$\|u-u_h\|_{1,\Omega}$	order	$\left(\sum_{i=1}^{3} h\|\lambda-\lambda_{\tilde{h}}\|^2_{0,\gamma_i}\right)^{\frac{1}{2}}$	order
$u=u_1(x,y)$	$n=1$	3.5804		0.1088	
	$n=2$	1.7993	0.9927	0.0384	1.5025
	$n=3$	1.1986	1.0019	0.0152	2.2857
	$n=4$	0.8988	1.0006	0.0124	0.7077
	$n=5$	0.7186	1.0027	0.0067	2.7587
	$n=6$	0.5988	1.0003	0.0060	0.6052
	$n=7$	0.5131	1.0020	0.0034	3.6846
	$n=8$	0.4489	1.0010	0.0033	0.2236
	$n=9$	0.3994	0.9920	0.0029	1.0970
	$n=10$	0.3592	1.0069	0.0015	6.2570
	$n=11$	0.3265	1.0015	0.0019	-2.4802
	$n=12$	0.3001	0.9690	0.0015	2.7168
	$n=13$	0.2762	1.0368	0.0012	2.7878
	$n=14$	0.2563	1.0090	0.0011	1.1741
	$n=15$	0.2395	0.9826	0.0010	1.3814
	$n=16$	0.2243	1.0160	0.0008	3.4575
$u=u_2(x,y)$	$n=1$	189822		3171.73	
	$n=2$	92227.5	1.0414	798.105	1.9906
	$n=3$	61104.7	1.0153	468.024	1.3163
	$n=4$	45742.1	1.0066	364.426	0.8697
	$n=5$	36542.6	1.0063	121.598	4.9188
	$n=6$	30434.6	1.0032	178.351	-2.1009
	$n=7$	26069.7	1.0043	98.6184	3.8436
	$n=8$	22804.8	1.0020	91.6176	0.5514
	$n=9$	20282.7	0.9951	86.4029	0.4975
	$n=10$	18245.4	1.0047	44.2459	6.3521
	$n=11$	16584	1.0017	58.2477	-2.8847
	$n=12$	15252.3	0.9620	49.7358	1.8156
	$n=13$	14030.6	1.0431	35.2758	4.2918
	$n=14$	13015	1.0139	30.919	1.7788
	$n=15$	12160.4	0.9844	32.2092	-0.5925
	$n=16$	11390.6	1.0133	25.1221	3.8505

Table 4.3. Finite element errors for the smooth examples and $r=0.025$.

Next, we move onto a case in which the exact solution is not known. More precisely, we consider the approximation of the following boundary value problem:

$$\begin{cases} -\Delta \hat{u} &= 0 \quad \text{in} \quad \omega, \\ \hat{u} &= 10 \quad \text{on} \quad \gamma_i, \quad i=1,2,3, \\ u &= 0 \quad \text{on} \quad \partial\Omega. \end{cases} \tag{4.17}$$

Then, we have computed a reference solution using a standard FEM in ω using a mesh containing 71200 elements. We denote this solution by u_{ref}. In Tables 4.4, 4.5 and 4.6 we report the errors $\|u_{ref} - u_h\|_{0,\omega}$ and $\|u_{ref} - u_h\|_{1,\omega}$. We see that these tend to zero, but with a slower rate. The slower rate of convergence is due to two factors. For the coarser meshes the results show an erratic behavior typical of preasymptotic regimes (since $diam(B_i) < h$). Later, for fine meshes the convergence orders tend to stabilise on a suboptimal order, which is linked to the fact that u does not belong to $H^2(\Omega)$.

	$\|u_{ref} - u_h\|_{0,\omega}$	order	$\|u_{ref} - u_h\|_{1,\omega}$	order
$n = 1$	17.9261		28.069	
$n = 2$	12.4341	0.5278	20.0713	0.4838
$n = 3$	9.5128	0.6605	15.7798	0.5933
$n = 4$	7.9418	0.6274	13.0286	0.6660
$n = 5$	6.6404	0.8020	11.2396	0.6619
$n = 6$	5.6830	0.8539	9.8531	0.7221
$n = 7$	5.0598	0.7535	8.7309	0.7844
$n = 8$	4.5119	0.8583	7.9581	0.6941
$n = 12$	3.1452	0.8899	5.9118	0.7331
$n = 16$	2.3856	0.9609	4.7040	0.7944
$n = 20$	1.9860	0.8216	4.002	0.7243
$n = 24$	1.6406	1.0479	3.3968	0.8993
$n = 28$	1.4105	0.9803	3.0996	0.5940
$n = 32$	1.2245	1.0590	2.8213	0.7045

Table 4.4. The errors $\|u_{ref} - u_h\|_{0,\omega}$ and $\|u_{ref} - u_h\|_{1,\Omega}$ for $r = 0.2$.

	$\|u_{ref} - u_h\|_{0,\omega}$	order	$\|u_{ref} - u_h\|_{1,\omega}$	order
$n = 1$	17.1514		27.2304	
$n = 2$	12.7182	0.4314	21.4572	0.3438
$n = 3$	10.591	0.4514	17.3827	0.5194
$n = 4$	8.6825	0.6907	15.0703	0.4962
$n = 5$	7.5684	0.6154	13.1351	0.6159
$n = 6$	6.6633	0.6986	11.7117	0.6291
$n = 7$	5.8129	0.8857	10.6428	0.6209
$n = 8$	5.3928	0.5618	9.5588	0.8045
$n = 12$	3.8631	0.8227	7.2007	0.6987
$n = 16$	3.0487	0.8230	5.7899	0.7580
$n = 20$	2.5217	0.8505	5.0093	0.6490
$n = 24$	2.0896	1.0309	4.2809	0.8618
$n = 28$	1.8109	0.9286	3.7392	0.8776
$n = 32$	1.5872	0.9874	3.5047	0.4850

Table 4.5. The errors $\|u_{ref} - u_h\|_{0,\omega}$ and $\|u_{ref} - u_h\|_{1,\Omega}$ for $r = 0.1$.

	$\|u_{ref} - u_h\|_{0,\omega}$	order	$\|u_{ref} - u_h\|_{1,\omega}$	order
$n = 1$	14.8711		24.4094	
$n = 2$	13.1243	0.1803	22.3855	0.1249
$n = 3$	11.8203	0.2581	20.2568	0.2464
$n = 4$	10.5745	0.3871	18.9668	0.2287
$n = 5$	9.6422	0.4136	17.5545	0.3468
$n = 6$	8.8191	0.4894	16.4094	0.3700
$n = 7$	8.3809	0.3306	15.0603	0.5565
$n = 8$	7.4988	0.8329	14.536	0.2654
$n = 9$	6.9397	0.6579	13.8106	0.4346
$n = 10$	6.9632	-0.0321	12.5167	0.9337
$n = 11$	6.1493	1.3042	12.4309	0.0722
$n = 12$	5.8831	0.5086	11.7656	0.6322
$n = 13$	5.8012	0.1751	10.9178	0.9343
$n = 14$	5.3615	1.0637	10.6914	0.2828
$n = 15$	5.1095	0.6978	10.266	0.5885
$n = 16$	4.9088	0.6208	9.8383	0.6594

Table 4.6. The errors $\|u_{ref} - u_h\|_{0,\omega}$ and $\|u_{ref} - u_h\|_{1,\Omega}$ for $r = 0.025$.

We observe the behavior of the reference and approximate solutions for $n = 10$ when $r = 0.2$ in Fig. 4.6- 4.9.

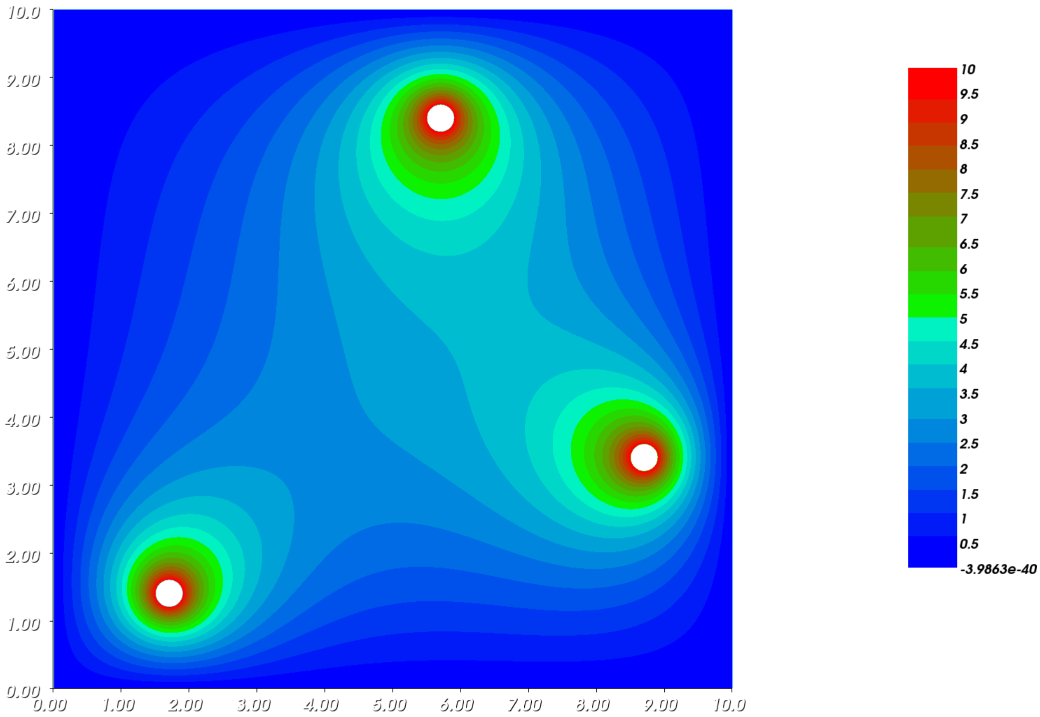


Fig. 4.6. Reference solution, u_{ref}, of the problem (4.17) for $r = 0.2$.

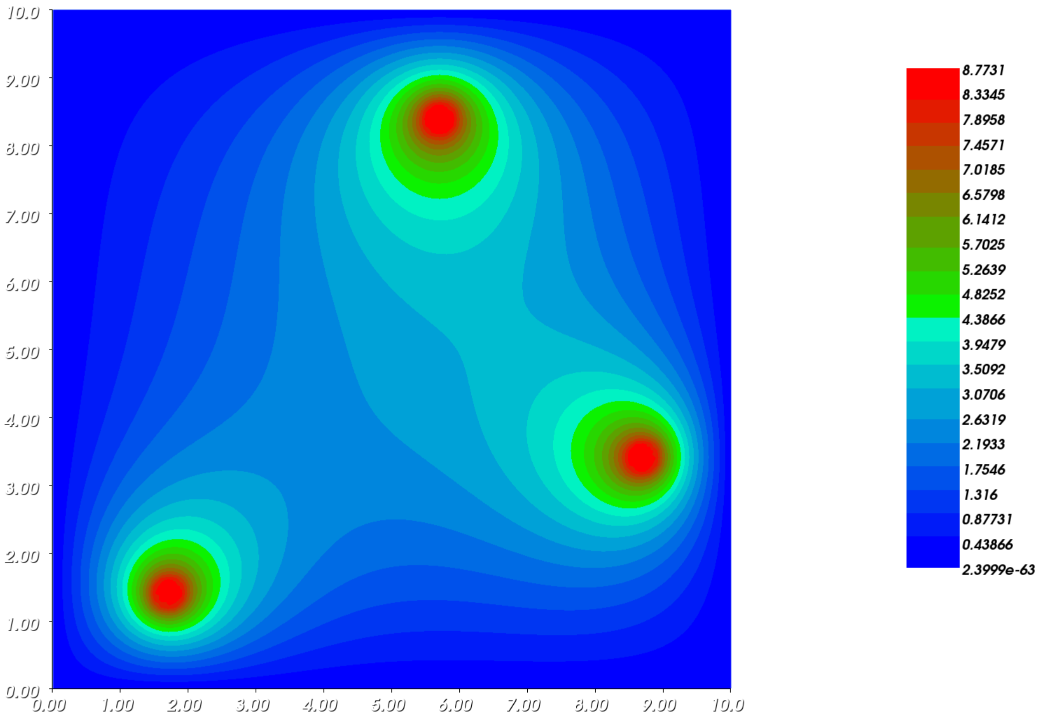


Fig. 4.7. Approximate solution, u_h, of the problem (4.17) when $n = 10$ for $r = 0.2$.

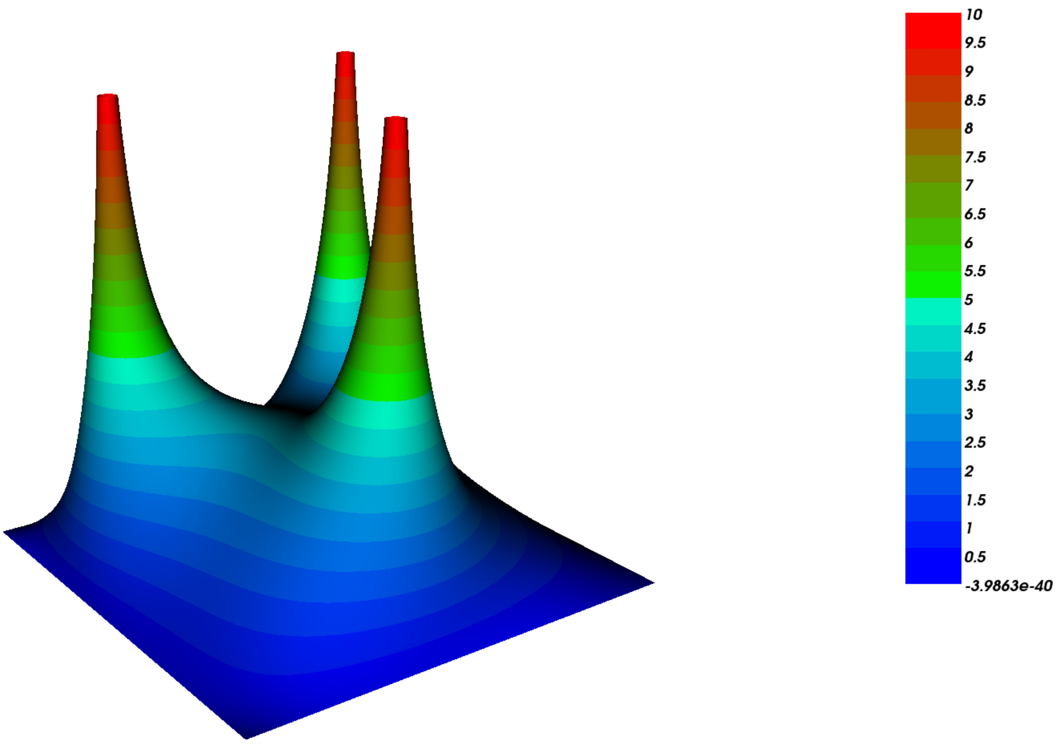


Fig. 4.8. Reference solution, u_{ref}, of the problem (4.17) for $r = 0.2$.

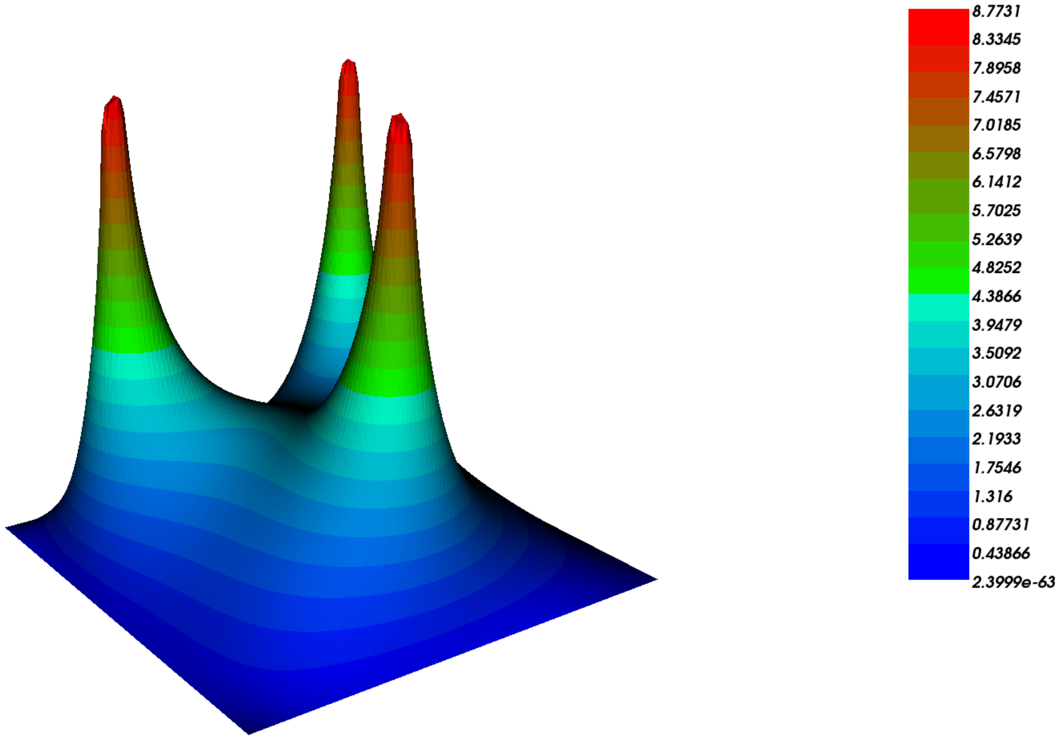


Fig. 4.9. Approximate solution, u_h, of the problem (4.17) when $n = 10$ for $r = 0.2$.

For both solutions, the cross section is shown in Fig. 4.10, when $x \in (0, 10)$, $r = 0.2$ and we fix $y = 8.4$, with different mesh refinements .

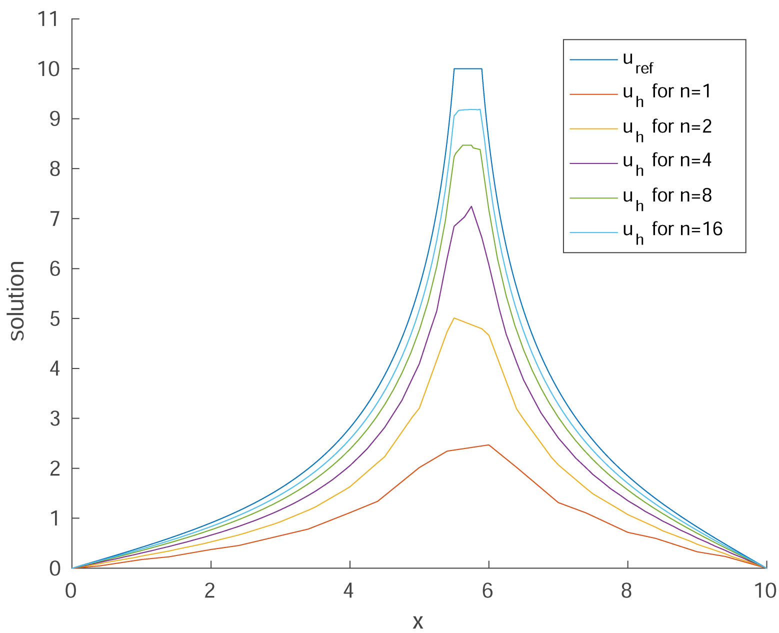


Fig. 4.10. Cross section of the approximate solution for some values of n and the approximate solution along the line $y = 8.4$ when $r = 0.2$.

4.6. Extension to time-dependent problems

We extend the work done in this chapter to the problem depending on time. The problem which we have consider is the following:

find $\hat{u} : \omega \times (0,T) \longrightarrow \mathbb{R}$ such that

$$\begin{cases} \partial_t \hat{u} - \Delta \hat{u} &= \hat{f} \quad \text{in} \quad \omega \times (0,T), \\ u &= g_i \quad \text{on} \quad \gamma_i \times (0,T), \quad i = 1, ..., M. \\ u &= 0 \quad \text{on} \quad \partial\Omega \times (0,T), \\ \hat{u}(x,0) &= \hat{u}_0(x) \quad \text{in} \quad \omega, \end{cases} \tag{4.18}$$

where $\hat{f} \in L^2(\omega)$, $g_i \in H^{\frac{1}{2}}(\gamma_i)$, and we keep the same notation as on the previous section.

We can apply the theory from Chapter 3 with fictitious domain to this stabilised problem:

find $(u, \boldsymbol{\lambda}) \in W := H^1_0(\Omega) \times \prod_{i=1}^{M} H^{-\frac{1}{2}}(\gamma_i)$, where $\boldsymbol{\lambda} = (\lambda_i)_{i=1}^M$, such that

$$\begin{cases} (\partial_t u, v)_\Omega + (\nabla u, \nabla v)_\Omega - \sum_{i=1}^{M} \langle \lambda_i, v \rangle_{\gamma_i} &= (f, v)_\Omega, \\ \sum_{i=1}^{M} \langle \mu_i, u \rangle_{\gamma_i} &= \sum_{i=1}^{M} \langle \mu_i, g_i \rangle_{\gamma_i}, \end{cases} \tag{4.19}$$

for all $(v, \boldsymbol{\mu}) \in W, \boldsymbol{\mu} = (\mu_i)_{i=1}^M$. Then, the fully discrete problem is as follows:

Given a suitable approximation of $u_h^0 \in V_h$ of u_0 for all $(v_h, \boldsymbol{\mu}_{\tilde{h}}) \in W_h$ and $0 \leq n \leq N-1$, find $(u_h^{n+1}, \boldsymbol{\lambda}_{\tilde{h}}^{n+1}) \in W_h$ such that

$$\begin{cases} \frac{1}{\delta t}(u_h^{n+1} - u_h^n, v_h)_\Omega + a(u_h^{n+1}, v_h) - \sum_{i=1}^{M} \langle \lambda_{i,\tilde{h}}^{n+1}, v \rangle_{\gamma_i} &= (f(t_{n+1}), v_h)_\Omega \\ \sum_{i=1}^{M} \langle \mu_{i,\tilde{h}}, u_h^{n+1} \rangle_{\gamma_i} + j(\boldsymbol{\lambda}_{\tilde{h}}^{n+1}, \boldsymbol{\mu}_{\tilde{h}}) &= \sum_{i=1}^{M} \langle \mu_{i,\tilde{h}}, g_i \rangle_{\gamma_i}. \end{cases} \tag{4.20}$$

The stabilised term added in this case to circumvent the inf-sup condition,

$$j(\boldsymbol{\lambda}_{\tilde{h}}^{n+1}, \boldsymbol{\mu}_{\tilde{h}}) = \sum_{i=1}^{M} h \langle \lambda_{i,\tilde{h}}^{n+1}, \mu_{i,\tilde{h}} \rangle_{\gamma_i},$$

satisfies the requirements of the analysis in Chapter 3. More precisely, (3.5) is satisfied automatically. Conditions (3.6) and (3.7) are used primarily in the error analysis, and even if the current definition of j does not fulfill them, we notice that if $(u, \boldsymbol{\lambda})$ is the exact solution of (4.19), and $u \in H^{\frac{3}{2}+\epsilon}(\Omega)$, then $\boldsymbol{\lambda} = 0$ and $j(\boldsymbol{\lambda}_{\tilde{h}}^{n+1}, \boldsymbol{\mu}_{\tilde{h}}) = 0$ for all $\boldsymbol{\mu}_{\tilde{h}} \in \Lambda_{\tilde{h}}$, which is enough to carry the analysis over.

4.6.1. Numerical studies

We consider the problem (4.18) with the domains and meshes defined in Section 4.5 with $r = 0.2$.

We first build two examples in which we know the exact solution. We set the right-hand side f and the boundary conditions in such a way that the solutions are given by two different examples:

$$u_1(x,t) = e^t(x^2 + xy) \quad \text{and} \quad u_2(x,y) = \cos(t)(x^2 + xy) \quad \text{in} \quad \Omega,$$

taking the time step $\delta t = h$. We solve the problem (4.18) defining the initial value for the solution (when $t = 0$) by the Ritz-projection as we have seen in Chapter 3. In Table 4.7 we report the values of the errors $\|u(.,T) - u_h(.,T)\|_{1,\Omega}$ and $j\big(\boldsymbol{\lambda}(.,T) - \boldsymbol{\lambda}_{\tilde{h}}(.,T), \boldsymbol{\lambda}(.,T) - \boldsymbol{\lambda}_{\tilde{h}}(.,T)\big)^{\frac{1}{2}}$, where we have used $T = 1$. We can observe that these quantities tend to zero with optimal rates.

		$\|u-u_h\|_{1,\Omega}$	order	$\left(\sum_{i=1}^{3} h\|\lambda-\lambda_{\tilde{h}}\|^2_{0,\gamma_i}\right)^{\frac{1}{2}}$	order
$u=u_1(x,y)$	$n=1$	522.107		86.9896	
	$n=2$	270.092	0.9509	50.9639	0.7714
	$n=3$	181.752	0.9770	35.0871	0.9206
	$n=4$	137.091	0.9802	25.7068	1.0813
	$n=5$	109.991	0.9870	20.2107	1.0780
	$n=6$	91.8498	0.9886	16.454	1.1279
	$n=7$	78.8835	0.9872	13.6429	1.2154
	$n=8$	69.1239	0.9891	11.5956	1.2176
	$n=9$	61.5557	0.9845	9.98372	1.2707
	$n=10$	55.4355	0.9939	8.72377	1.2804
	$n=11$	50.411	0.9969	7.79123	1.1862
	$n=12$	46.2607	0.9874	6.92771	1.3500
	$n=13$	42.7205	0.9946	6.2629	1.2604
	$n=14$	39.7034	0.9883	5.64608	1.3991
	$n=15$	37.0754	0.9926	5.15584	1.3165
	$n=16$	34.7705	0.9945	4.73283	1.3264
$u=u_2(x,y)$	$n=1$	197.257		32.6503	
	$n=2$	109.657	0.8471	20.323	0.6840
	$n=3$	75.7344	0.9128	14.1366	0.8952
	$n=4$	57.8828	0.9344	10.3649	1.0788
	$n=5$	46.8138	0.9511	8.12328	1.0921
	$n=6$	39.2982	0.9598	6.58655	1.1502
	$n=7$	33.8793	0.9625	5.4457	1.2339
	$n=8$	29.7689	0.9686	4.61124	1.2456
	$n=9$	26.5629	0.9674	3.96006	1.2925
	$n=10$	23.9661	0.9764	3.45152	1.3045
	$n=11$	21.8236	0.9826	3.0706	1.2270
	$n=12$	20.0498	0.9743	2.72619	1.3673
	$n=13$	18.5328	0.9829	2.45772	1.2952
	$n=14$	17.2381	0.9772	2.21345	1.4126
	$n=15$	16.1077	0.9831	2.01815	1.3389
	$n=16$	15.1148	0.9858	1.84933	1.3536

Table 4.7. Finite element errors for the smooth examples with $\delta t = h$ and $r = 0.2$.

We next move onto a case in which the exact solution is not known. In this case, we define the initial value for the solution by zero. More precisely, we consider the approximation of the following boundary value problem:

$$\begin{cases} \partial_t \hat{u} - \Delta \hat{u} &= 0 \quad \text{in} \quad \omega \times (0,1], \\ \hat{u} &= 10t \quad \text{on} \quad \gamma_i \times (0,1], \quad i = 1,2,3. \\ u &= 0 \quad \text{on} \quad \partial\Omega \times (0,1], \\ \hat{u}(x,0) &= 0 \quad \text{in} \quad \omega. \end{cases} \tag{4.21}$$

Then, we have computed the reference and approximate solutions as in Section 4.5, taking $\delta t = 10^{-4}$ and $\delta t = h$ respectively. As expected, the errors decrease, but not with an optimal rate.

We observe the behavior of the reference solution for $\delta t = 10^{-4}$ and approximate solution for $\delta t = h$ and $n = 10$, when $r = 0.2$, in Fig. 4.11- 4.14.

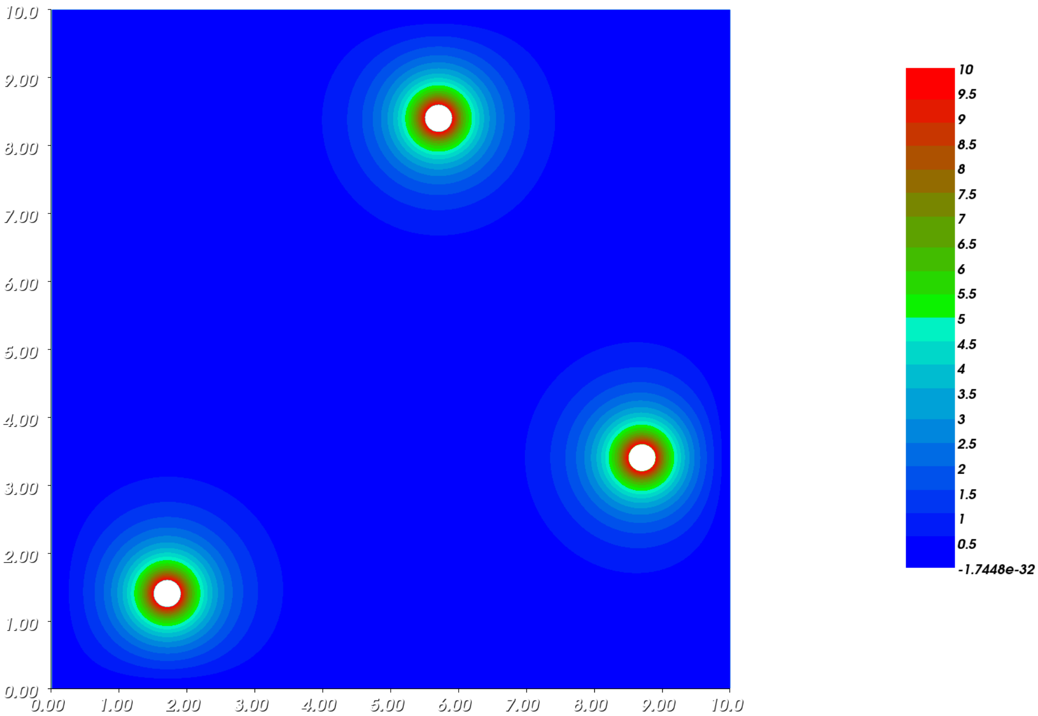


Fig. 4.11. Reference solution, u_{ref}, of the problem (4.21) for $r = 0.2$.

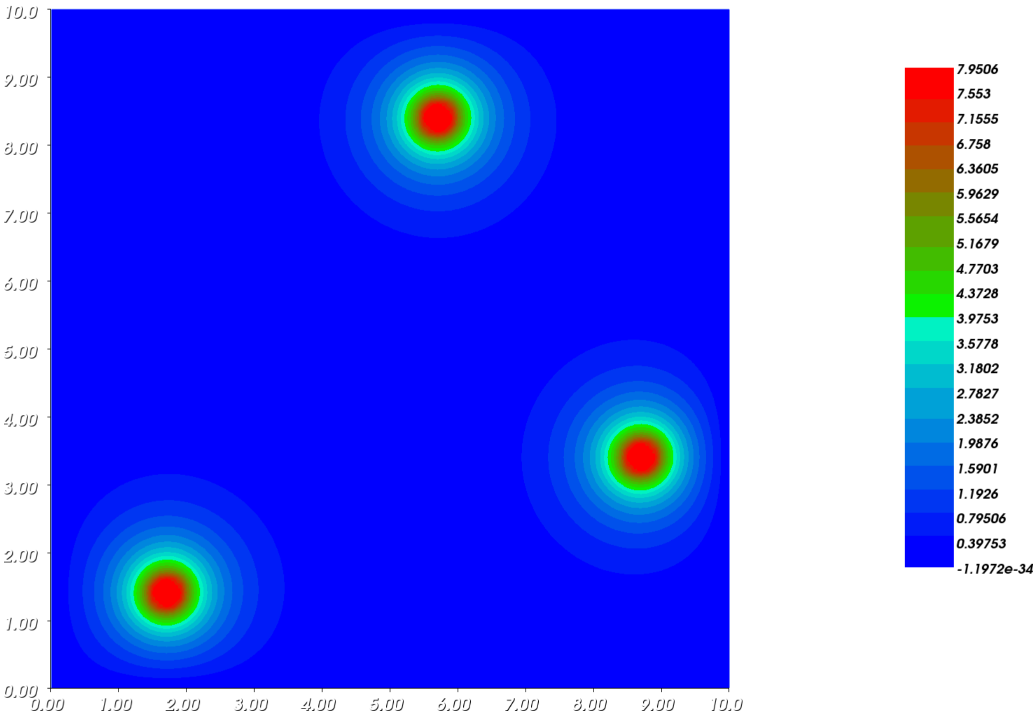


Fig. 4.12. Approximate solution, u_h, of the problem (4.21) when $n = 10$ for $r = 0.2$.

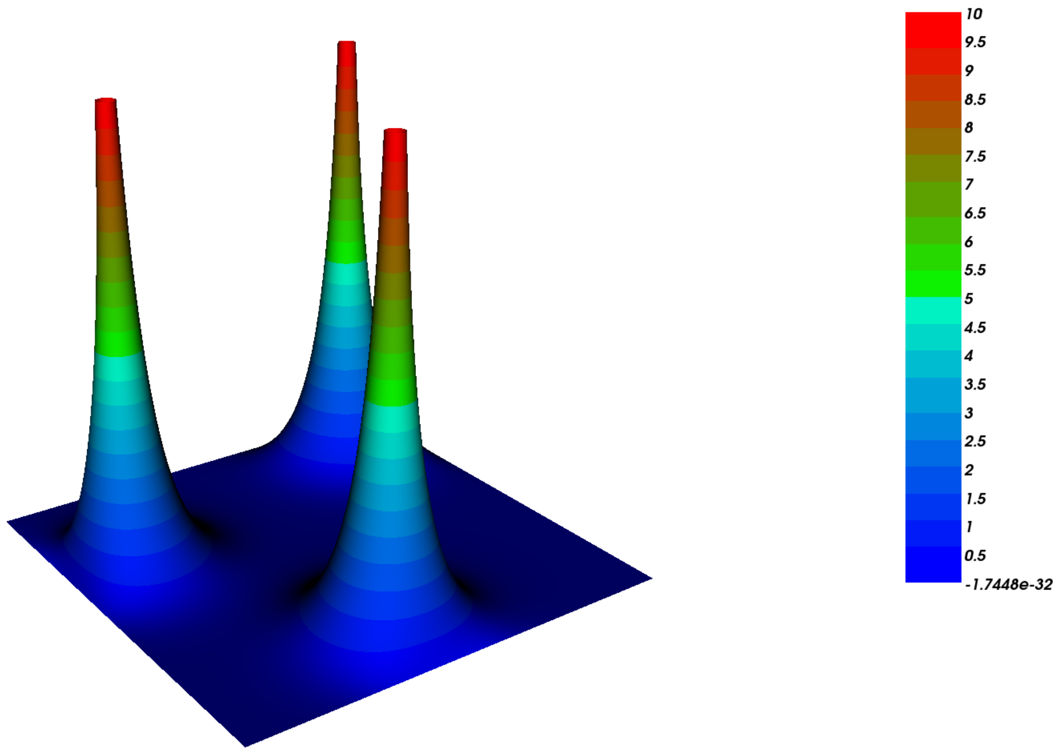


Fig. 4.13. Reference solution, u_{ref}, of the problem (4.21) for $r = 0.2$.

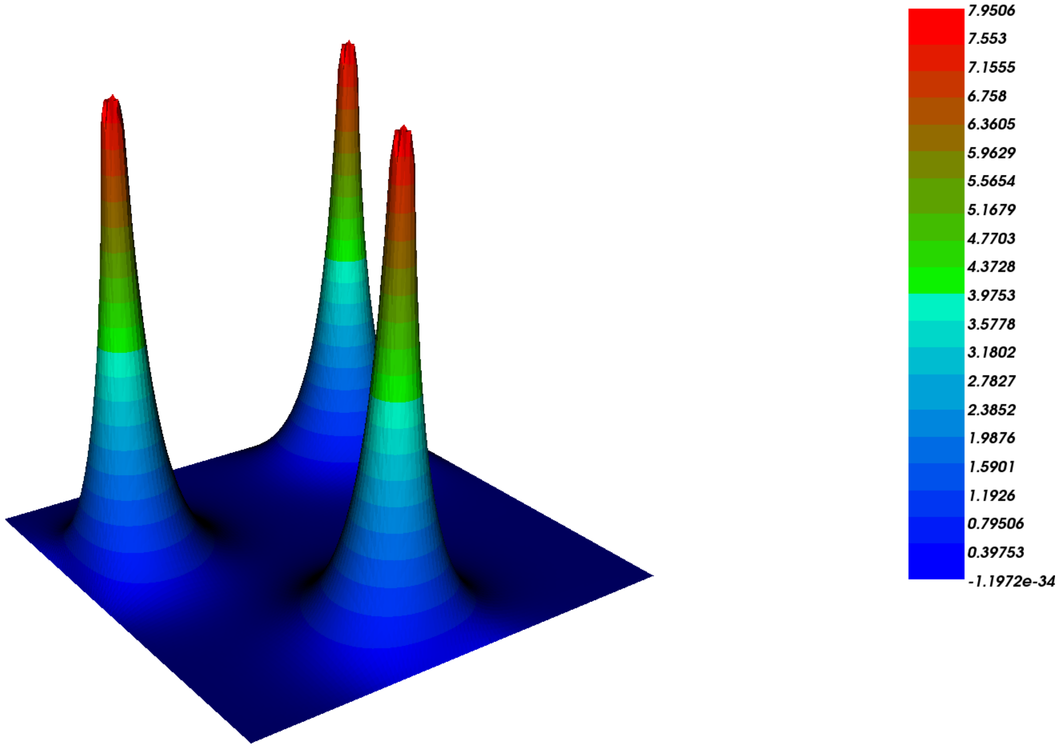


Fig. 4.14. Approximate solution, u_h, of the problem (4.21) when $n = 10$ for $r = 0.2$.

In Fig. 4.15 we depict cross-section of the reference and approximate solution at $T = 1$, obtained using different mesh refinement levels, and $\delta t = h$.

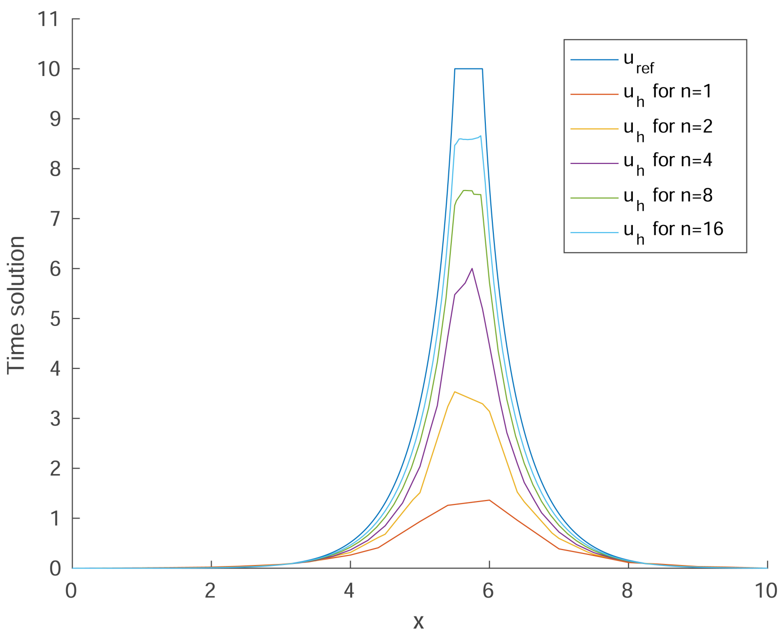


Fig. 4.15. Cross section of the approximate solution of the problem (4.21) for some values of n and the reference solution along the line $y = 8.4$ when $r = 0.2$.

In addition, in Fig. 4.16 we depict cross-section of the reference solutions and approximate solution at different stages using $n = 8$. Here then $h = \frac{1}{8} = \delta t$, which implies the approximation is quite coarse. We observe then a rough approximation of the boundary condition. This is due, mostly, to the space discretisation error. As a matter of fact, we have repeated the same test case using $h = \frac{1}{8}$ as before but $\delta t = 10^{-2}$ for the fictitious domain method. The results are depicted in Fig. 4.17, where we observe that the solution given by the fictitious domain method is very close to the one using $\delta t = \frac{1}{8}$.

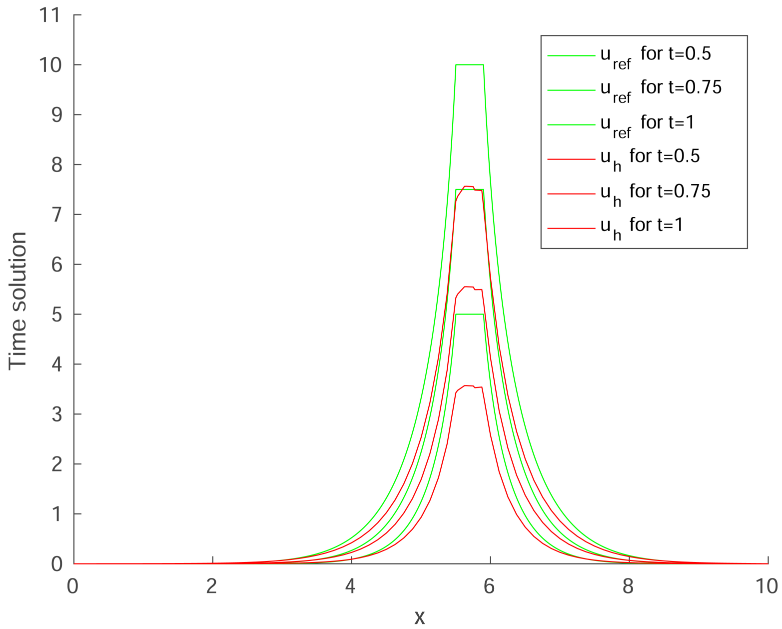


Fig. 4.16. Cross section of the approximate solution for $n = 8$ and the reference solution of the problem (4.21) along the line $y = 8.4$, $r = 0.2$ and different times when $h = \delta t$.

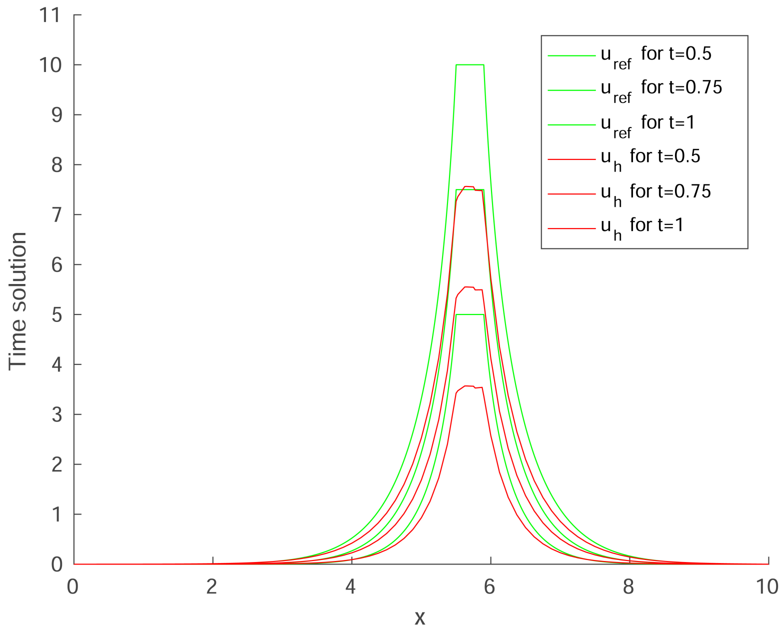


Fig. 4.17. Cross section of the approximate solution for $n = 8$ and the reference solution of the problem (4.21) along the line $y = 8.4$, $r = 0.2$ and different times when $\delta t = 10^{-2}$ and $h = \frac{1}{8}$.

Chapter 5

Conclusions and future work

In this chapter, we first present conclusions, and then we show our ideas for future projects.

5.1. Conclusions

Our goal in this thesis has been to solve partial differential equations through FEM. The domain, where our problem is defined, can be very complicated in shape and we have used fictitious domain methods.

Our first motivation was to rewrite the problem defined in the initial domain ω as a problem in the fictitious domain Ω where it is easier to solve it. In Chapter 2, we have observed that if the choice of extension f makes the solution u less regular, then the method will not converge with an optimal rate. In fact, the numerical experiments shown prove that. Otherwise, optimal convergence is obtained if we choose an appropriate extension of the function $\hat{f}$. Pursuing this purpose is not an easy task, especially in dimensions higher than one.

We have implemented the time-dependent problem (1.14), described in Chapter 1, with and without fictitious domain as well. We have checked the same fact for both cases. The regularity of the solution depends on the extension of $\hat{f}$ and it not easy to define f when we are working with an n-dimensional problem. We avoid to work with higher-order methods for the time being owing to the fact that we do not know how to extend $\hat{f}$ in such a way that $u \in H^{k+1}(\Omega)$ when we are working with an n-dimensional problem. So, we focus on transient problems with fictitious domain for $\mathbb{P}_1/\mathbb{P}_0$ finite element spaces where we need just $f \in L^2(\Omega)$ which was our second motivation described in the Chapter 3 after reading papers as [BF09, BC12].

In Chapter 3, we have proved unconditional stability and optimal convergence for the transient heat problem. We have shown that for small time steps the use of a Ritz-projection for the initial data is essential to avoid instabilities, unless the condition between time and space discretisation parameters ($\delta t \geq Ch^2$) is satisfied. We have carried out the full analysis for the backward difference formula for order one when the inf-sup stable condition is satisfied or not. To circumvent the inf-sup stable condition, a stabilisation operator has been added to obtain a stabilised method.

In Chapter 4, we have proposed a simple stabilised FEM for approximating the solution of partial differential equations posed in domains containing a moderate amount of small perforations. The method is a fictitious domain method, enhanced with a stabilisation term that, in some cases, is reminiscent of the Barbosa-Hughes stabilised method. The numerical results show that, at least for the cases presented in this chapter, this method can be used to approximate the solution (especially the far field) with a good accuracy, without the need to modify the finite element space, or to consider elements with special geometries. We do not expect this method to give accurate results if we consider a domain with a very large number of perforations but, as long as this number remains moderate, we believe this method presents a simple alternative to previously existing references. This analysis has also been extended to the time-dependent problem building on the analysis performed in Chapter 3.

5.2. Future projects

We have new concerns related to what we have mentioned in this document. We would like to extend our future work in two directions mainly.

On the one hand, we have started the study of the described problem here with a new alternative: when the domain ω is moving over time. This idea is applied by *Fluid-Structure Interaction problems* (FSI). Fluid-structure interactions can be stable or oscillatory. In the oscillatory case, the strain induced in the solid structure causes it to move such that the source of strain is reduced, and the structure returns to its former state only for the process to repeat. The study of fluid-structure interactions are essential in the field of engineering or medicine for example. One very famous example is the spectacular collapse of the Tacoma Narrows Bridge which occurred in 1940 (see, e.g., [AVKW41] for more details), where the excessive oscilations caused by wind action were the main problem. Nowadays, this type of problems are really important in the development of medicine, for example, for the simulation of biological problems related to the blood flow in the heart.

Navier-Stokes equations are applied to FSI, in general, because they describe the motion of a fluid. We would like to apply our results to FSI considering that the velocity of the fluid is the gradient of a potential and then, it is a heat equation. Thus, we think that we can prove the unconditional stability and optimal convergence for our problem when the domain is moving following the FSI structure. In papers such as [BRS11, BG16b, JDP10], the stability results are provided and optimal convergence estimates are proved for coupled problems, such as FSI. In [JDP10], it is shown that the overall problem is likely to be less regular stable and less accurate than the individual sub-problems consisting of two or more domains unless special measures are taken. An appropiate time integration scheme is applied to each domain to obtain better results. In [BRS11] the authors present the full analysis of a problem for the interaction between the wind and a sail (applying it to nautical sports). They propose a fictitious domain formulation of the problem, involving the wind velocity stream function and a Lagrange multiplier considering two domains, the fictitious domain and the moving domain (the sail). As in our case, the authors in this paper work with a FEM based on piecewise linear continuous elements to approximate the stream function and piecewise constant ones for the Lagrange multiplier and the analysis follows essentially the paper [GPP94] as well. The work done in [BRS11] is very similar to our study described in Chapter 3. The method is used to determine the sail shape under the action of the wind to obtain, for example, a greater speed. We will consider the paper [JGP02] as well, where the authors look at the motion of rigid bodies settling in an incompressible viscous fluid based on fictitious domain formulation.

On the other hand, with respect to the work done in Chapter 4, more intensive numerical testing and the extension to problems in incompressible fluid and solid mechanics will be the subject of future research as well. And also we have in mind to continue with the idea proposed in Chapter 2 and try to find a good extension of $\hat{f}$ in dimensions higher than one.

Chapter A

Appendix A

We shall specify the calculus used to get the solution and errors of the problems worked on Chapter 2.

A.1. Method without fictitious domain

In this section, u is the solution in the initial domain $(0,1)$ instead of $\hat{u}$ and f is used instead of $\hat{f}$ to simplify notation.

The problem has been solved through FEM, so the discrete form is given by (2.3). We consider three different linear spaces and the process to solve it is depending on each one in details shown. All computations have been performed using $Matlab$ [ACF99].

A.1.1. $\mathbb{P}_1$ finite dimensional space

Let us consider problem (2.1). We introduce the linear space

$$V = \{v \in C^0(\omega) : v' \text{ is piecewise continuous and bounded } \omega, v = 0 \text{ on } \partial\omega\}$$

where $V \supset V_h = \{v_h \in C^0(\omega) : v_h|_{[x_i,x_{i+1}]} \in \mathbb{P}_1[x_i, x_{i+1}]\}$.

We can see the problem as a linear system of equations with N+1 equations in N+1 unknowns $u_1, u_2, ..., u_{N+1}$ where $\boldsymbol{u} = [u_1, ..., u_{N+1}]^\top$ is the solution. The linear system to be solved is

$$\sum_{j=1}^{N+1} u_i(\varphi_i', \varphi_j') = (f, \varphi_j),$$

which can be written in matrix form as $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{F}$, where $\boldsymbol{A} = (\varphi_i', \varphi_j'), u = [u_1, ..., u_{N+1}]^T$ and $\boldsymbol{F} = (f, \varphi_j)$.

Let N be the number of subintervals in $\omega = (0,1)$ of length $h = \frac{1}{N}$. Let us introduce the basis functions $\varphi_j \in V_h$ for all $j = 1, ..., N+1$ defined by

$$\varphi_j(x_i) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

The elements $a_{ij} = (\varphi_i', \varphi_j')$ of the stiffness matrix $\boldsymbol{A}$ can easily be computed. First, we observe that $(\varphi_i', \varphi_j') = 0$ if $\mid i - j \mid > 1$ since, in this case, for all $x \in (0,1)$ either $\varphi_i(x)$ or $\varphi_j(x)$ is equal to zero. So, the matrix $\boldsymbol{A}$ is tri-diagonal where

$$(\varphi_i', \varphi_i') = \int_{x_{i-1}}^{x_i} \frac{1}{h^2} dx + \int_{x_i}^{x_{i+1}} \frac{1}{h^2} dx = \frac{1}{h} + \frac{1}{h} = \frac{2}{h}$$
$$(\varphi_i', \varphi_{i-1}') = (\varphi_{i-1}', \varphi_i') = -\int_{x_{i-1}}^{x_i} \frac{1}{h^2} dx = -\frac{1}{h}$$

Note also that the matrix $\boldsymbol{A}$ is symmetric and positive definite. Specially,

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{h} & -\frac{1}{h} & 0 & 0 & 0 & \cdots & 0 \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & 0 & 0 & & \\ 0 & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & 0 & & \vdots \\ \vdots & & & & \ddots & & \\ 0 & & \cdots & & 0 & -\frac{1}{h} & \frac{1}{h} \end{bmatrix}$$

If we impose the boundary conditions, we will change the first and the last row. Hence, in this problem, the stiffness matrix is

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & & 0 \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & 0 & 0 & & & \\ 0 & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & 0 & & & \vdots \\ \vdots & & & & \ddots & & & \\ 0 & & \cdots & & & 0 & 0 & 1 \end{bmatrix}$$

The load vector, $\boldsymbol{F} = [F_1, ..., F_{N+1}]^\top$, is defined as

$$\boldsymbol{F} = \int_0^1 f(x)\varphi(x)dx,$$

and where restricted to each subinterval is

$$F_i = \int_{x_{i-1}}^{x_{i+1}} f(x)\varphi_i(x)dx = hf(x_i).$$

Moreover, the boundary conditions are applied, $F_1 = F_{n+1} = 0$.

Once we have developed this work, we can obtain the solution by solving the linear system $\boldsymbol{Au} = \boldsymbol{F}$.

We shall now study the error between the solution of (2.1) and the solution of the finite element problem (2.3).

The error in $L^2(0,1)$ norm is given by

$$\| u - u_h \|^2_{0,(0,1)} = \int_0^1 (u - u_h)^2(x)dx = \sum_{i=1}^{N} \int_{x_i}^{x_{i+1}} (u - u_h)^2(x)dx,$$

which is approximated by the Simpson's rule.

The error in $H^1(0,1)$ norm is defined by:

$$\begin{aligned} | u - u_h |^2_{1,(0,1)} &= \int_0^1 (u' - u'_h)^2 = \sum_{i=1}^{N} \int_{x_i}^{x_{i+1}} (u' - u'_h)^2(x)dx \\ &= \sum_{i=1}^{N} \int_{x_i}^{x_{i+1}} [u'(x) - u'_h(x)]^2 dx \\ &= \sum_{i=1}^{N} \int_{x_i}^{x_{i+1}} [u'(x) - \frac{u_h(i+1) - u_h(i)}{h}]^2 dx. \end{aligned}$$

A.1.2. $\mathbb{P}_2$ finite dimensional space

Let us consider the problem (2.1) but, in this case, the linear space is $V_h = \{v_h \in C^0(\omega) : v_h|_{[x_{2i-1}, x_{2i+1}]} \in \mathbb{P}_2[x_{2i-1}, x_{2i+1}]\}$.

We follow the same procedure as in previous section.

We will have $2N+1$ nodes in $\mathbb{P}_2$, so the stiffness matrix $\boldsymbol{B} \in \boldsymbol{M}(2N+1, 2N+1)$. In the reference interval $(0,1)$, the bases functions are

$$\begin{aligned} \hat{\varphi}_1(x) &= (1-x)(1-2x) && \Longrightarrow && \hat{\varphi}'_1(x) = -3 + 4x, \\ \hat{\varphi}_2(x) &= 4x(1-x) && \Longrightarrow && \hat{\varphi}'_2(x) = 4 - 8x, \\ \hat{\varphi}_3(x) &= x(2x-1) && \Longrightarrow && \hat{\varphi}'_3(x) = 4x - 1. \end{aligned}$$

Then, we define the matrix $\boldsymbol{BG}$ as:

$$\boldsymbol{BG} := [(\hat{\varphi}'_j \hat{\varphi}'_i)] \qquad \text{for} \quad i, j = 1, 2, 3.$$

A simple calculation in our particular case gives:

$$
\begin{aligned}
&(\hat{\varphi}_1'\hat{\varphi}_1') = \frac{7}{3},\\
&(\hat{\varphi}_2'\hat{\varphi}_2') = \frac{16}{3},\\
&(\hat{\varphi}_3'\hat{\varphi}_3') = \frac{7}{3},\\
&(\hat{\varphi}_1'\hat{\varphi}_2') = (\hat{\varphi}_2'\hat{\varphi}_1') = -\frac{8}{3},\\
&(\hat{\varphi}_1'\hat{\varphi}_3') = (\hat{\varphi}_3'\hat{\varphi}_1') = \frac{1}{3},\\
&(\hat{\varphi}_2'\hat{\varphi}_3') = (\hat{\varphi}_3'\hat{\varphi}_2') = -\frac{8}{3}.
\end{aligned}
$$

If we work with the matrix $\boldsymbol{BG}$ and we perform the change of variables

$$
\begin{aligned}
T : (0,1) &\longrightarrow (x_{2i-1}, x_{2i+1})\\
\hat{x} &\longrightarrow T\hat{x} = x_{2i-1} + \hat{x}h,
\end{aligned}
$$

leads to

$$
\begin{aligned}
\hat{\varphi}(\hat{x}) &= \varphi(x), \quad x = x_{2i-1} + \hat{x}h,\\
\hat{\varphi}'(\hat{x}) &= \varphi'(x)\frac{dx}{d\hat{x}} = \varphi'(x)h.
\end{aligned}
$$

This will help us to calculate the stiffness matrix $\boldsymbol{B}$. Moreover, we impose the boundary conditions where the first and the last row are zeros except $\boldsymbol{B}(1,1) = \boldsymbol{B}(2N+1, 2N+1) = 1$.

Otherwise, we calculate the load vector, $\boldsymbol{F2} = [F2_1, ..., F2_{2N+1}]^\top$. In this case we have to differenciate the basis functions of the node in odd or ever position, i.e.,

$$
\begin{aligned}
\int_0^1 f\varphi_{2i} &= \int_{x_{2i-1}}^{x_{2i+1}} f\varphi_{2i} = \frac{2h}{3} f(x_{2i}),\\
\int_0^1 f\varphi_{2i+1} &= \int_{x_{2i-1}}^{x_{2i+1}} f\varphi_{2i+1} + \int_{x_{2i+1}}^{x_{2i+3}} f\varphi_{2i+1} = \frac{h}{3} f(x_{2i+1}),
\end{aligned}
$$

applying the Simpson's rule.

In this case, to calculate the errors, we have to use quadrature formula of higher order because the error due to the quadrature rule has to be smaller than the finite element error. For that, we use the quadrature formula

$$\int_a^b f(x)dx = w_1 f(y_1) + w_2 f(y_2) + w_3 f(y_3),$$

where y_1, y_2, and y_3 are the quadrature points and w_1, w_2 and w_3 are the weights (see [EG04]). In our problem

$$\begin{aligned}
y_1 &= \frac{1}{2}(a+b) - \frac{(b-a)}{2}\sqrt{\frac{3}{5}},\\
y_2 &= \frac{1}{2}(a+b),\\
y_3 &= \frac{1}{2}(a+b) + \frac{(b-a)}{2}\sqrt{\frac{3}{5}},\\
w_1 &= w_3 = \frac{5}{18} \quad \text{and} \quad w_2 = \frac{8}{18}.
\end{aligned}$$

Thus, with $h = (x_{2i+1} - x_{2i-1})$,

$$\begin{aligned}
\int_0^1 (u-u_h)^2(x)dx = \sum_{i=1}^{N} \int_{x_{2i-1}}^{x_{2i+1}} (u-u_h)^2(x)dx \\
\simeq \sum_{i=1}^{N} \Big\{ \frac{5}{18} h(u(y_1) - u_h(y_1))^2 \\
+ \frac{8}{18} h(u(y_2) - u_h(y_2))^2 \\
+ \frac{5}{18} h(u(y_3) - u_h(y_3))^2 \Big\}.
\end{aligned}$$

In this expression, $u_h(y_1)$ and $u_h(y_3)$ are unknowns. To find these values, we shall calculate $\varphi_{2i-1}(y_i)$ and $\varphi_{2i}(y_i), \varphi_{2i+1}(y_i)$, for $i = 1, 3$. It will be easier if we work in the reference interval and we obtain the following matrix

$$\boldsymbol{Q} = \begin{bmatrix} \hat{\varphi}_1(\hat{y}_1) & \hat{\varphi}_2(\hat{y}_1) & \hat{\varphi}_3(\hat{y}_1) \\ \hat{\varphi}_1(\hat{y}_2) & \hat{\varphi}_2(\hat{y}_2) & \hat{\varphi}_3(\hat{y}_2) \\ \hat{\varphi}_1(\hat{y}_3) & \hat{\varphi}_2(\hat{y}_3) & \hat{\varphi}_3(\hat{y}_3) \end{bmatrix},$$

where

$$\begin{aligned}
\hat{y}_1 &= \frac{1}{2} - \frac{1}{2}\sqrt{\frac{3}{5}},\\
\hat{y}_2 &= \frac{1}{2},\\
\hat{y}_3 &= \frac{1}{2} + \frac{1}{2}\sqrt{\frac{3}{5}}
\end{aligned}$$

and $\hat{\varphi}_i$ are the $\mathbb{P}_2$ basis functions in the reference interval for $i = 1, 2, 3$.

For the error in $H^1(0,1)$ norm,

$$\int_0^1 (u' - u'_h)^2(x)dx = \sum_{i=1}^{N} \int_{x_{2i-1}}^{x_{2i+1}} (u'(x) - u'_h(x))^2 dx$$
$$\simeq \sum_{i=1}^{N} \Big\{ \frac{5}{18} h(u'(y_1) - u'_h(y_1))^2 + \frac{8}{18} h(u'(y_2) - u'_h(y_2))^2 + \frac{5}{18} h(u'(y_3) - u'_h(y_3))^2 \Big\},$$

where

$$u'_h(y_i) = u_h(x_{2i-1})\varphi'_{2i-1}(y_i) + u_h(x_{2i})\varphi'_{2i}(y_i) + u_h(x_{2i+1})\varphi'_{2i+1}(y_i) \quad \text{for} \quad i = 1,2,3,$$

and $\varphi'_{2i-1}(y_i), \varphi'_{2i}(yi), \varphi'_{2i+1}(y_i)$ are unknown. To facilitate calculations, we will calculate the derivative of the basis functions in the reference interval again, such that

$$\boldsymbol{P} = \begin{bmatrix} \hat{\varphi}'_1(\hat{y}_1) & \hat{\varphi}'_2(\hat{y}_1) & \hat{\varphi}'_3(\hat{y}_1) \\ \hat{\varphi}'_1(\hat{y}_2) & \hat{\varphi}'_2(\hat{y}_2) & \hat{\varphi}'_3(\hat{y}_2) \\ \hat{\varphi}'_1(\hat{y}_3) & \hat{\varphi}'_2(\hat{y}_3) & \hat{\varphi}'_3(\hat{y}_3) \end{bmatrix}.$$

A.1.3. $\mathbb{P}_3$ finite dimensional space

Let us consider the problem (2.1) with $V_h = \{v_h \in C^0(\omega) : v_h|_{[x_{3i-2}, x_{3i+1}]} \in \mathbb{P}_3[x_{3i-2}, x_{3i+1}]\}$ and the same steps are followed as before.

We define with the stiffness matrix $\boldsymbol{C}$. Let N be the number of subintervals in ω of lenght $h = \frac{1}{N}$. We will have $3N+1$ nodes in $\mathbb{P}_3$, so $\boldsymbol{C} \in \boldsymbol{M}(3N+1, 3N+1)$. In the reference interval $(0,1)$, the basis functions are

$$\hat{\varphi}_1(x) = -\frac{9}{2}\left(\frac{1}{3} - x\right)\left(x - \frac{2}{3}\right)(1-x) \implies \hat{\varphi}'_1(x) = -\frac{27}{2}x^2 + 18x - \frac{11}{2},$$
$$\hat{\varphi}_2(x) = x\left(\frac{2}{3} - x\right)(x-1)\left(-\frac{27}{2}\right) \implies \hat{\varphi}'_2(x) = \frac{81}{2}x^2 - 45x + 9,$$
$$\hat{\varphi}_3(x) = \frac{27}{2}x\left(x - \frac{1}{3}\right)(1-x) \implies \hat{\varphi}'_3(x) = -\frac{81}{2}x^2 + 36x - \frac{9}{2},$$
$$\hat{\varphi}_4(x) = -\frac{9}{2}x\left(\frac{1}{3} - x\right)\left(x - \frac{2}{3}\right) \implies \hat{\varphi}'_4(x) = \frac{27}{2}x^2 - 9x + 1.$$

Then we define the matrix $\boldsymbol{CG}$ as

$$\boldsymbol{CG} := [(\hat{\varphi}'_j, \hat{\varphi}'_i)] \qquad \text{for} \quad i,j = 1,2,3,4.$$

We calculate it using a quadrature formula. If we work with the matrix $\boldsymbol{CG}$ and we perform the change of variables, we will be able to calculate the stiffness matrix $\boldsymbol{C}$. Moreover, we impose the boundary conditions where the first and the last row are zeros except $\boldsymbol{C}(1,1) = \boldsymbol{C}(3N+1, 3N+1) = 1$. Moreover, we calculate the load vector, $\boldsymbol{F3} = [F3_1, ..., F3_{3N+1}]^\top$, as

$$\boldsymbol{F} = \int_0^1 f(x)\varphi(x)dx \simeq \sum_{i=1}^{3N+1} f(x_k) \int_0^1 \varphi_k(x)\varphi_i(x)dx = \boldsymbol{M}[f(x_1), f(x_2), ..., f(x_{3N+1})]^\top,$$

where

$$M_{ki} = \int_0^1 \varphi_k(x)\varphi_i(x)dx.$$

To calculate the mass matrix $\boldsymbol{M}$,

$$\int_{x_k}^{x_{k+1}} \varphi_k(x)\varphi_i(x)dx = h\int_0^1 \hat{\varphi}_k(x)\hat{\varphi}_i(x)d\hat{x},$$

where the matrix MG in the reference interval is defined by

$$\boldsymbol{MG} = \left[\int_0^1 \hat{\varphi}_k(x)\hat{\varphi}_i(x)dx\right].$$

We obtain $\boldsymbol{MG}$ through quadrature formula and thus, the matrix $\boldsymbol{M}$. The error is calculated using the same procedure as in $\mathbb{P}_2$. We use the quadrature formula

$$\int_a^b f(x)dx = w_1 f(y_1) + w_2 f(y_2) + w_3 f(y_3) + w_4 f(y_4),$$

where

$$y_1 = \frac{(a+b)}{2} - \frac{(b-a)}{2}\sqrt{\frac{15+2\sqrt{30}}{35}},$$

$$y_2 = \frac{(a+b)}{2} - \frac{(b-a)}{2}\sqrt{\frac{15-2\sqrt{30}}{35}},$$

$$y_3 = \frac{(a+b)}{2} + \frac{(b-a)}{2}\sqrt{\frac{15-2\sqrt{30}}{35}},$$

$$y_4 = \frac{(a+b)}{2} + \frac{(b-a)}{2}\sqrt{\frac{15+2\sqrt{30}}{35}},$$

$$w_1 = w_4 = \left(\frac{1}{4} - \frac{1}{12}\sqrt{\frac{5}{6}}\right)(b-a),$$

$$w_2 = w_3 = \left(\frac{1}{4} + \frac{1}{12}\sqrt{\frac{5}{6}}\right)(b-a).$$

A.2. Method with fictitious domain

The problem (2.1) is solved in this section with fictitous domain formulation through FEM. In particular,we consider the mixed problem (2.6).

We use the same procedure as in Section A.1 for $\mathbb{P}_1$, $\mathbb{P}_2$ and $\mathbb{P}_3$ but using these new data. Thus, to implement the problem with $Matlab$, we need to change the domain, boundary conditions, the function $\hat{f}(x)$, exact solution $uex(x)$ and define the extension function $f(x)$. Futhermore, we introduce the variables ele and k define as

$$\begin{aligned} ele &= max\{i : x_i \leq 0\}, \\ k &= max\{j : x_j \leq 1\}, \end{aligned}$$

in such a way that $[x_{ele+1}, x_k] \subseteq (0, 1)$. The code to implement the problem in $Matlab$ is the same to get the solution as in the method without fictitous domain just changing the data for each finite dimensional space and following the details shown in Chapter 2 for this case.

Chapter B

Appendix B

When we are working on the stabilised method, we defined a bilinear form $j : \Lambda_h \times \Lambda_h \longrightarrow \mathbb{R}$ as

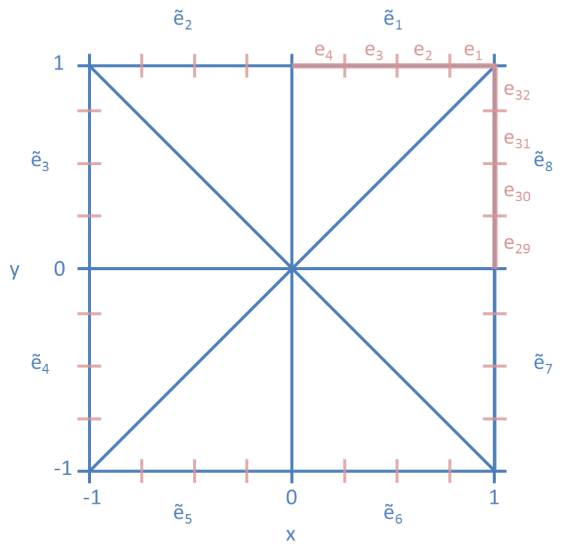
$$j(\lambda, \mu) = \sum_{\tilde{e} \in \gamma_{\tilde{h}}} |\tilde{e}|(\lambda - \Pi_{\tilde{h}}\lambda, \mu - \Pi_{\tilde{h}}\mu)_{\tilde{e}},$$

where $\Pi_{\tilde{h}} : L^2(\gamma) \longrightarrow \Lambda_{\tilde{h}}$ is defined as $(\Pi_{\tilde{h}}\xi)|_{\tilde{e}} = |\tilde{e}|^{-1}(\xi, 1)_{\tilde{e}}$. This stabilised term, j, will be introduced in the problem as a matrix $\boldsymbol{C}$ in our case. Thus, in this section, we want to describe the proccess to define this matrix in $FreeFem++$.

B.1. Matrix C from the stabilised term

We work with two uniform meshes on the boundary of ω, γ. Let γ_h be the partition of γ with edges e. Let also $\gamma_{\tilde{h}}$ be a partition of γ, whose vertices are also vertices of γ_h, with edges $\tilde{e}$ satisfying the following: there exists $C > 0$ (independent of h) such that $3h \leqslant| \tilde{e} |\leqslant Ch$, for all $\tilde{e} \in \gamma_{\tilde{h}}$. Using the mesh regularity of $\mathcal{T}_h$, a shape-regular family of triangulations of the fictitious domain Ω, it is easy to see that for all $\tilde{e} \in \gamma_{\tilde{h}}$, $\text{card}\{e \in \gamma_h : e \subset \tilde{e}\} \leqslant C$, where $C > 0$ is independent of $\tilde{e}$ and h. The angular points of γ also belong to $\gamma_{\tilde{h}}$.

We show this proccess for the case where ω is a square of side two. The partition of both meshes on γ change depending of n, given in the full code of the problem, which gives us the refinement of the meshes. For example, when $n = 0$,



First, we need to load "lapack" in the full code to define the matrix $\boldsymbol{C}$ by blocks which changes for every value of n.

After that, we start implementing the code for the matrix $\boldsymbol{C}$. We define the square block matrix $\boldsymbol{B}_1$ which arises from the term

$$|\tilde{e}|(\lambda - \Pi_{\tilde{h}}\lambda, \mu - \Pi_{\tilde{h}}\mu)_{\tilde{e}}.$$

In our example,

$$\begin{pmatrix} (\lambda_1 - \Pi\lambda_1, \lambda_1 - \Pi\lambda_1) & \cdots & (\lambda_1 - \Pi\lambda_1, \lambda_4) \\ \vdots & \ddots & \vdots \\ (\lambda_1, \lambda_4 - \Pi\lambda_4) & \cdots & (\lambda_4 - \Pi\lambda_4, \lambda_4 - \Pi\lambda_4) \end{pmatrix} = \begin{pmatrix} \frac{3}{16} & \cdots & \frac{3}{16} \\ \vdots & \ddots & \vdots \\ \frac{3}{16} & \cdots & \frac{3}{16} \end{pmatrix} * |\tilde{e}|$$

which multiplied by $|\tilde{e}|$ again, from the definition of j, gives us the square block matrix $\boldsymbol{B}_1$ whose dimension is equal to four because we split each $\tilde{e}$ of the mesh $\tilde{\gamma}$ in four e of γ. So, in $FreeFem++$, we define the real number $a = \frac{3}{16}$ (line 5) and knowing that the lenght of one of the elements $\tilde{e}$ is equal to $\frac{1}{n+1}$, we can build the matrix $\boldsymbol{B}_1$ (lines 6-15).

One you have defined $\boldsymbol{B}_1$, the diagonal matrix $\boldsymbol{C}$ is formed with this block $8*(n+1)$ times,

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{B}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{B}_1 \end{bmatrix}.$$

To implement the matrix $\boldsymbol{C}$ in $FreeFem++$, we will do it in two steps: we define array $\boldsymbol{D}$ and after that, we copy array $\boldsymbol{D}$ to matrix $\boldsymbol{C}$ just to simplify the code. We need to convert to matrix in order to work in our full code where we find the solution to our problem of matrix form.

To define $\boldsymbol{D}$, we need to introduce new variables. Number of times that $\boldsymbol{B}_1$ is repeated in $\boldsymbol{D}$ which depend of the value of n, $numberOfBlocks$. A variable, $startIdx$, whose value is the index of the element of $\boldsymbol{D}$ in the first row and column in each block $\boldsymbol{B}_1$. When we are working in C++, the element of a matrix which is in the first row and first column has $index = 0$. And $blockwidth$, which says the dimension of the square block matrix $\boldsymbol{B}_1$ (line 25).

We start defining the array $\boldsymbol{D}$ equal to zero whose dimension is the product between $numberOfBlocks$ and $blockwidth$ (lines 27-30) and we do some changes explained later on (lines 31-39). We are transforming the array $\boldsymbol{D}$ to one which is not sparse and it is formed with $\boldsymbol{B}_1$ in the diagonal. To this end, we begin with a loop where the variable iB takes integer values from 1 to 8, both included, for the case $n = 0$ (line 31). For each value of this variable, $startIdx$ is changing as well (line 32), and give us the number of row and column of $\boldsymbol{D}$ where we will write the element $b1_{00}$ of the matrix $\boldsymbol{B}_1$. The following loops, which depend on j and ir, change the elements of $\boldsymbol{D}$ equal to zeros to the rest of the matrix $\boldsymbol{B}_1$ (lines 34-38). For example, to insert the second block matrix in $\boldsymbol{D}$, $iB = 2$, $startIdx = 4$, j takes the value 4 until 7. For $j = 4$, integer number ir can be 4 or less than 8. Thus, we get $d_{44} = b1_{00}$, $d_{54} = b1_{10}$, $d_{64} = b1_{20}$ and $d_{74} = b1_{30}$ when $ir = 4$. We need to continue with this proccess with every value in the loops to get the second block of $\boldsymbol{D}$ and the final array $\boldsymbol{D}$.

Therefore, as we said before, last step is convert the array $\boldsymbol{D}$ to the matrix $\boldsymbol{C}$ (line 42).

In such a way that we have implemented the matrix $\boldsymbol{C}$, stabilised term, in $FreeFem+$ +. The code reads as follows

```
1  load "lapack"
2  // Define matrix C, stabilised term.
3  
4  real a=3/16.;
5  matrix B1=[
6              [-a/((n+1)*(n+1)),-a/((n+1)*(n+1)),
7               -a/((n+1)*(n+1)), -a/((n+1)*(n+1))],
8              [-a/((n+1)*(n+1)), -a/((n+1)*(n+1)),
9               -a/((n+1)*(n+1)), -a/((n+1)*(n+1))],
10             [-a/((n+1)*(n+1)), -a/((n+1)*(n+1)),
11              -a/((n+1)*(n+1)), -a/((n+1)*(n+1))],
12             [-a/((n+1)*(n+1)), -a/((n+1)*(n+1)),
```

```
            -a/((n+1)*(n+1)), -a/((n+1)*(n+1))]
];


// define the block matrix which dimension depends on "n":

// 1. define array D to use loop to assemble "array"
// 2. copy array D to matrix C


// Step 1
int numberOfBlocks=8*(n+1), startIdx, blockwidth=4;


real[int,int]
       D(numberOfBlocks*blockwidth,numberOfBlocks*blockwidth);

D=0;
for(int iB=1;iB<numberOfBlocks+1;iB++){
        startIdx = (iB-1)*blockwidth;
        // (since C++ starts with index zero)
        for(int j=startIdx;j<startIdx+4;j++){
                for(int ir=startIdx;ir<startIdx+4;ir++){
                        D(ir,j) = B1(ir%4, j%4);
                };
        };
};

// Step 2
matrix C = D;
```

Bibliography

[AB72] A. K. Aziz and Ivo Babuška, *Survey lectures on the mathematical foundations of the finite element method. The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations* , Academic Press, 1972.

[ABG+15] Ferdinando Auricchio, Daniele Boffi, Lucia Gastaldi, Adrien Lefieux, and Alessandro Reali, *On a fictitious domain method with distributed Lagrange multiplier for interface problems* , Applied Numerical Mathematics **95** (2015), 36 – 50.

[ACC+16] Paola Antonietti, Andrea Cangiani, Joe Collis, Zhaonan Dong, Emmanuil H. Georgoulis, Stefano Giani, and Paul Houston, *Review of discontinuous Galerkin finite element methods for partial differential equations on complicated domains. In building bridges, connections and challenges in modern approaches to numerical partial differential equations. Lecture notes in Computational Science and Engineering, Vol.114*, Springer, Heidelberg, 2016, to appear.

[ACF99] Jochen Alberty, Carsten Carstensen, and Stefan A. Funken, *Remarks around 50 lines of Matlab: short finite element implementation*, Numer. Algorithms **20** (1999), no. 2-3, 117–137.

[ADG+91] C. Atamian, Q.V. Dinh, Roland Glowinski, Jiwen He, and Jacques Périaux, *Control approach to fictitious-domain methods. Application to fluid dynamics and electro-magnetics*, Proceedings of the Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, R. Glowinski, Y. Kuznetsov, G. Meurant, J. Périaux, and OB Widlund, eds., SIAM, Philadelphia, 1991, pp. 275–309.

[Ago93] Abdellatif Agouzal, *Analyse numérique de méthodes de décomposition de domaines: Méthodes de domaines fictifs avec multiplicateurs de Lagrange*, Ph.D. thesis, 1993.

[AVKW41] Othmar H Amman, Theodore Von Kármán, and Glenn B Woodruff, *The failure of the Tacoma Narrows bridge*, CaltechAUTHORS (1941).

[Bab73] Ivo Babuška, *The finite element method with lagrangian multipliers*, Numer. Math. **20** (1973), no. 3, 179–192.

[BB07] Gabriel R. Barrenechea and Jordi Blasco, *Pressure stabilization of finite element approximations of time-dependent incompressible flow problems*, Computer Methods in Applied Mechanics and Engineering **197** (2007), no. 14, 219 – 231.

[BC01] Jordi Blasco and Ramón Codina, *Space and time error estimates for a first order, pressure stabilized finite element method for the incompressible NavierStokes equations*, Applied Numerical Mathematics **38** (2001), no. 4, 475 – 497.

[BC12] Gabriel R. Barrenechea and Franz Chouly, *A local projection stabilized method for fictitious domains*, Appl. Math. Lett. **25** (2012), no. 12, 2071–2076.

[BF09] Erik Burman and Miguel A. Fernández, *Galerkin finite element methods with symmetric pressure stabilization for the transient Stokes equations: stability and convergence analysis*, SIAM J. Numer. Anal. **47** (2008/09), no. 1, 409–439.

[BF91] Franco Brezzi and Michel Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991.

[BG16a] Gabriel R. Barrenechea and Cheherazada González, *A stabilised finite element method for a fictitious domain problem allowing small inclusions*, Numerical methods for partial differential equaitions (2016), submitted.

[BG16b] Daniele Boffi and Lucia Gastaldi, *A fictitious domain approach with Lagrange multiplier for fluid-structure interactions*, Numerische Mathematik (2016), to appear.

[BGL07] Pavel B. Bochev, Max D. Gunzburger, and Richard B. Lehoucq, *On stabilized finite element methods for the stokes problem in the small time step limit*, International Journal for Numerical Methods in Fluids **53** (2007), no. 4, 573–597.

[BGR14] Daniele Boffi, Lucia Gastaldi, and Michele Ruggeri, *Mixed formulation for interface problems with distributed Lagrange multiplier*, Computers Mathematics with Applications **68** (2014), no. 12, Part B, 2151 – 2166.

[BGS04] Pavel B. Bochev, Max D. Gunzburger, and John N. Shadid, *On infsup stabilized finite element methods for transient problems* , Computer Methods in Applied Mechanics and Engineering **193** (2004), no. 1516, 1471 – 1489.

[BH91] Helio J. C. Barbosa and Thomas J. R. Hughes, *The finite element method with Lagrange multipliers on the boundary: circumventing the Babuška-Brezzi condition*, Comput. Methods Appl. Mech. Engrg. **85** (1991), no. 1, 109–128. MR 1088318

[BH10] Erik Burman and Peter Hansbo, *Fictitious domain finite element methods using cut elements: I. A stabilized Lagrange multiplier method* , Computer Methods in Applied Mechanics and Engineering (2010).

[BH12] Erik Burman and Peter Hansbo, *Fictitious domain finite element methods using cut elements: II. A stabilized Nitsche method*, Applied Numerical Mathematics (2012).

[BLL14] Claude Le Bris, Frédéric Legoll, and Alexei Lozinski, *An MsFEM Type Approach for Perforated Domains*, Multiscale Modeling & Simulation **12** (2014), no. 3, 1046–1077.

[Bre74] Franco Brezzi, *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique **8** (1974), no. 2, 129–151.

[BRS11] Alfredo Bermúdez, Rodolfo Rodríguez, and María Luisa Seoane, *A fictitious domain method for the numerical two-dimensional simulation of potential flows past sails*, ESAIM: M2AN **45** (2011), no. 6, 1033–1058.

[BS08] Susanne Brenner and Ridgway Scott, *The mathematical theory of finite element methods*, third ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008.

[BTOL77] Ivo Babuška, J. Tinsley Oden, and J.K. Lee, *Mixed-hybrid finite element approximations of second-order elliptic boundary-value problems*, Computer Methods in Applied Mechanics and Engineering **11** (1977), no. 2, 175 – 206.

[CB00] Ramón Codina and Jordi Blasco, *Analysis of a pressure-stabilized finite element approximation of the stationary navier-stokes equations*, Numerische Mathematik **87** (2000), no. 1, 59–81.

[CB09] Ramon Codina and Joan Baiges, *Approximate imposition of boundary conditions in immersed boundary methods*, International Journal for Numerical Methods in Engineering **80** (2009), no. 11, 1379–1405.

[CGH14] Andrea Cangiani, Emmanuil H. Georgoulis, and Paul Houston, *Hp-version discontinuous Galerkin methods on polygonal and polyhedral meshes*, Math. Models Methods Appl. Sci. **24** (2014), no. 10, 2009–2041.

[CGS16] Andrea Cangiani, Emmanuil H. Georgoulis, and Y. Sabawi, *Adaptive discontinuous Galerkin methods for elliptic interface problems*, submitted.

[Cia78] Philippe G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978.

[CJM96] Francis Collino, Patrick Joly, and Florence Millot, *Fictitious domain method for unsteady problems: Application to electromagnetic scattering*, Research Report RR-2963, INRIA, 1996.

[CPGB07] Ramón Codina, Javier Príncipe, Oriol Guasch, and Santiago Badia, *Time dependent subscales in the stabilized finite element approximation of incompressible flow problems*, Computer Methods in Applied Mechanics and Engineering **196** (2007), no. 2124, 2413 – 2430.

[CVZ98] Ramón Codina, M. Vázquez, and O. C. Zienkiewicz, *A general algorithm for compressible and incompressible flows. part iii: The semi-implicit form*, International Journal for Numerical Methods in Fluids **27** (1998), no. 1-4, 13–32.

[EG04] Alexandre Ern and Jean-Luc Guermond, *Theory and practice of finite elements*, Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004.

[FGM13] Benoit Fabréges, Loc Gouarin, and Bertrand Maury, *A smooth extension method*, Comptes Rendus Mathematique **351** (2013), no. 910, 361 – 366.

[For77] Michel Fortin, *An analysis of the convergence of mixed finite element methods*, RAIRO Anal. Numér. **11** (1977), no. 4, 341–354, iii.

[GG95] Vivette Girault and Roland Glowinski, *Error analysis of a fictitious domain method applied to a Dirichlet problem*, Japan J. Indust. Appl. Math. **12** (1995), no. 3, 487–514.

[GH14] Stefano Giani and Paul Houston, *hp-adaptive composite discontinuous Galerkin methods for elliptic problems on complicated domains*, Numerical Methods for Partial Differential Equations **30** (2014), no. 4, 1342–1367.

[GPP94] Roland Glowinski, Tsorng-Whay Pan, and Jacques Periaux, *A fictitious domain method for Dirichlet problem and applications*, Computer Methods in Applied Mechanics and Engineering **111** (1994), no. 34, 283 – 303.

[GR86] Vivette Girault and Pierre-Arnaud Raviart, *Finite element methods for Navier Stokes equations*, Springer, Berlin, 1986.

[He03] Yinnian He, *A fully discrete stabilized finite-element method for the time-dependent Navier-Stokes problem*, IMA J. Numer. Anal. (2003).

[Hec12] Frédéric Hecht, *New development in FreeFem++*, J. Numer. Math. **20** (2012), no. 3-4, 251–265. MR 3043640

[HR09] Jaroslav Haslinger and Yves Renard, *A new fictitious domain approach inspired by the extended finite element method*, SIAM Journal on Numerical Analysis **47** (2009), no. 2, 1474–1499.

[HS97] Wolfgang Hackbusch and Stephen A. Sauter, *Composite finite elements for the approximation of pdes on domains with complicated micro-structures*, Numerische Mathematik **75** (1997), no. 4, 447–472.

[JDP10] Michel M. Joosten, Wulf G. Dettmer, and Djordje Perić, *On the temporal stability and accuracy of coupled problems with reference to fluidstructure interaction*, International Journal for Numerical Methods in Fluids **64** (2010), no. 10-12, 1363–1378.

[JGP02] L. H. Juárez, Roland Glowinski, and Tsorng-Whay Pan, *Numerical simulation of the sedimentation of rigid bodies in an incompressible viscous fluid by Lagrange multiplier/fictitious domain methods combined with the Taylor-Hood finite element approximation*, Journal of Scientific Computing **17** (2002), no. 1, 683–694.

[JN15] Volker John and Julia Novo, *Analysis of the pressure stabilized Petrov–Galerkin method for the evolutionary Stokes equations avoiding time step restrictions*, SIAM Journal on Numerical Analysis **53** (2015), no. 2, 1005–1031.

[Loz16] Alexei Lozinski, *A new fictitious domain method: Optimal convergence without cut elements*, Comptes Rendus Mathematique (2016), to appear.

[Mau01] Bertrand Maury, *A fat boundary method for the Poisson problem in a domain with holes*, Journal of Scientific Computing **16** (2001), no. 3, 319–339.

[MBT06] Nicolas Mos, Eric Béchet, and Matthieu Tourbier, *Imposing Dirichlet boundary conditions in the extended finite element method*, International Journal for Numerical Methods in Engineering **67** (2006), no. 12, 1641–1669.

[MDB99] Nicolas Moës, John Dolbow, and Ted Belytschko, *A finite element method for crack growth without remeshing*, International Journal for Numerical Methods in Engineering **46** (1999), no. 1, 131–150.

[Pit79] Juhani Pitkranta, *Boundary subspaces for the finite element method with Lagrange multipliers*, Numerische Mathematik **33** (1979), 273–290.

[PS08] Daniel Peterseim and Stefan A. Sauter, *The composite mini element-coarse mesh computation of Stokes flows on complicated domains*, SIAM J. Numer. Anal. **46** (2008), no. 6, 3181–3206.