



**Politecnico
di Torino**

Politecnico di Torino

Mathematical Engineering

A.a. 2025/2026

Graduation Session July 2026

The ALE-SBM Approach to Parabolic Problems in Moving Domains

Supervisors:

Claudio Canuto
Stefano Berrone

Candidate:

Stefano Iannarelli

Acknowledgements

Table of Contents

List of Figures	VI
1 Introduction	1
2 SBM for stationary problems	4
2.1 The Shifted Boundary Method	4
2.1.1 Definition of $\mathbf{M}_h$ for some smooth boundary	8
2.2 Coercivity and continuity analysis	11
2.3 Behavior of the Taylor remainder	17
2.4 Consistency and convergence analysis - Space	19
2.4.1 Enhanced error estimate in L^2	22
2.5 Numerical Results	26
2.5.1 Square Domain	27
2.5.2 Circular domain	30
2.5.3 Elliptic domain	33
2.5.4 Quarter-Circle domain	36
3 Time dependent problems	43
3.1 The Shifted Boundary Method for Time dependent problems	43
3.2 Stability, consistency and convergence analysis - Time	46
3.2.1 Stability	47
3.2.2 Consistency	50
3.2.3 Convergence	55
3.3 Numerical Results	57
3.3.1 Square domain	57
3.3.2 Quarter-circle domain	65
4 Moving Domain problems	73
4.1 ALE Formulation of the heat equation	75
4.2 The SBM formulation with ALE map	79
4.2.1 Stability of the ALE-SBM	81

4.2.2	ALE-free stability	83
4.3	Time semi-discretization	85
4.4	Full Discretization	92
4.5	Stability, consistency and convergence analysis - Moving domains .	99
4.5.1	Stability	99
4.5.2	Consistency	102
4.5.3	Convergence	107
4.6	Numerical Results	108
4.6.1	Extrapolation error	109
4.6.2	Square domain	112
4.6.3	L-shaped domain	116
4.6.4	Rotating ellipse	120
5	Conclusion	124
	Bibliography	127

List of Figures

2.1	The true domain Ω , the surrogate domain $\tilde{\Omega}_h \subset \Omega$ and their boundaries $\tilde{\Gamma}_h$ and Γ . Reproduced from [1].	5
2.2	Complications arising from the choice $\boldsymbol{\nu} = \tilde{\mathbf{n}}$. Reproduced from [1].	10
2.3	Projection comparison varying $\boldsymbol{\nu}$. $\tilde{\mathbf{n}}(\mathbf{x})$ is the normal vector of the surrogate boundary $\tilde{\Gamma}_h$. $\mathbf{n}(\mathbf{x})$ is the normal vector in the direction of the closest point projection (2.27 computed in $\mathbf{x}$). $\mathbf{n}(\tilde{\mathbf{x}})$ is the normal vector using $\mathbf{M}_h^z$ (2.27 computed in $\tilde{\mathbf{x}}$).	12
2.4	Square real and surrogate domains for $h = 0.137$	27
2.5	[Square] Comparison between the exact solution u and the numerical solution u_h for $h = 0.137$	28
2.6	[Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.	29
2.7	Circular real and surrogate domains for $h = 0.214$	30
2.8	[Circle] Comparison between the exact solution u and the numerical solution u_h for $h = 0.214$	31
2.9	[Circle] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.	32
2.10	Elliptic real and surrogate domains for $h = 0.201$	33
2.11	[Ellipse] Comparison between the exact solution u and the numerical solution u_h for $h = 0.201$	34
2.12	[Ellipse] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.	35
2.13	Quarter-Circle real and surrogate domains for $h = 0.107$	37
2.14	Angular and radial components of u	38
2.15	[Test 1 with $\alpha = 1, \text{Pe} = 5, k = 2$] Comparison between the exact solution u and the numerical solution u_h for $h = 0.107$	39
2.16	[Test 1 with $\alpha = 1, \text{Pe} = 5, k = 2$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.	40
2.17	[Test 2 with $\alpha = 5, \text{Pe} = 10, k = 5$] Comparison between the exact solution u and the numerical solution u_h for $h = 0.107$	41

2.18	[Test 2 with $\alpha = 5, \text{Pe} = 10, k = 5$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. . . .	42
3.1	[Square] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.137, \Delta t = 0.1$	59
3.2	[Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the right side ($\Delta t \sim \sqrt{h}$).	63
3.3	[Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h^2$) and Crank–Nicolson on the right side ($\Delta t \sim h$).	64
3.4	[Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.107, \Delta t = 0.05$ ($u(0) = 0$ is omitted). . . .	66
3.5	[Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the right side ($\Delta t \sim h$).	68
3.6	[Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.107, \Delta t = 0.05$ ($u(0) = 0$ is omitted). . . .	70
3.7	[Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the left side ($\Delta t \sim h$).	72
4.1	One-dimensional example of <i>Lagrangian</i> , <i>Eulerian</i> and ALE mesh and particle motion. Reproduced from [10].	74
4.2	Q^T comparison between stationary and moving domains.	77
4.3	Space-time representation of the ALE map. The diagram highlights how a generic function v maintains a constant value along the characteristics, becoming $\overleftarrow{v}$ at previous times.	86

4.4	Mapping from the reference space to the physical space. The domain features the real curved boundary Γ^n (blue) and the surrogate polygonal boundary $\tilde{\Gamma}_h^n$ (red). The physical element E lies inside the surrogate domain, touching $\tilde{\Gamma}_h^n$ strictly along the local edge e (green). The influence zone S_e bridges the gap between e and the true boundary along the projection directions, which fan out non-parallelly. The mappings B_E and B_{S_e} link the reference entities to their physical counterparts.	98
4.5	Surrogate domain (grey) and Extended strip S_h (yellow).	110
4.6	Comparison between the numerical solution u_h and his extension $\mathcal{E}_h u_h$ for $h = 0.0499$	110
4.7	Logarithmic L^2 -error regressions of $\mathcal{E}_h u_h - u$ and his corresponding theoretical convergence line (Rate 2.6285).	111
4.8	Square domain and its motion according to the velocity field $\mathbf{w}$. . .	112
4.9	[Moving Square] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137, \Delta t = 0.1$	113
4.10	[Moving Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with dG(0).	115
4.11	L-shaped domain: polar coordinates and its motion according to the velocity field $\mathbf{w}$	116
4.12	[Moving L-shaped domain] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137, \Delta t = 0.1$	117
4.13	[Moving L-shaped domain] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with dG(0). .	119
4.14	Rotating elliptic domain: local coordinate system tracked along the center's trajectory (left) and rigid rotation around the origin alongside its velocity field $\mathbf{w}$ (right).	120
4.15	[Rotating Ellipse] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137, \Delta t = 0.2$	121
4.16	[Rotating Ellipse] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with dG(0).	123

Chapter 1

Introduction

The shifted boundary method (SBM) is an approximate domain method for boundary value problems that falls within the broader class of unfitted, embedded, and immersed methods (see, e.g., [4, 6, 7, 8, 9, 12, 13, 14, 25, 26, 27, 30]).

In the SBM, the location where boundary conditions are applied is *shifted* from the true boundary to an approximate (surrogate) boundary. Simultaneously, the values of the boundary conditions are modified (also *shifted*) by means of Taylor expansions to prevent a reduction in the convergence rate of the overall formulation. Indeed, if the boundary conditions associated with the true domain are not appropriately modified on the surrogate domain, only first-order convergence can be expected. The appropriate (modified) boundary conditions are then applied weakly using a Nitsche-type strategy. This process yields a method that is simple, robust, accurate, and efficient.

The SBM was originally proposed for the Poisson and Stokes flow problems in [19]. It was subsequently generalized to the advection-diffusion and Navier-Stokes equations in [20], and to hyperbolic conservation laws in [31]. An extension of the SBM in conjunction with high-order gradient approximations was presented in [23], while the benefits of its application within the context of reduced-order modeling were analyzed in [15, 16, 17]. Further rigorous mathematical analysis for the Stokes and linear elasticity equations was carried out in [2, 3].

The primary objective of this thesis is to extend the SBM to parabolic equations, particularly within a moving domain setting, after first recalling its mathematical formulation for stationary problems such as the Poisson equation. To handle non-stationary domains, we adopt the Arbitrary Lagrangian-Eulerian (ALE) framework introduced in [5].

The analysis is developed for general two and three dimensional domains. Specifically, we present a comprehensive stability, consistency, and convergence analysis for a model Dirichlet problem governed by the Poisson and heat equations, containing both fixed and moving domains. The domain boundary may be smooth or

may contain corners and edges; notably, no convexity assumptions are made. For the sake of simplicity, only smooth domains are explicitly analyzed in this work; however, the extension to non-smooth domains is straightforward following the approach in [1].

Starting from an admissible, shape-regular triangulation that may not fit and may extend beyond the physical boundary, the surrogate domain is formed by the portion of such triangulation that is contained inside the physical domain. A natural condition on the distance between the surrogate and the exact boundaries, which should be bounded proportionally to the local meshsize, guarantees the coercivity of the discrete variational formulation. Classical conforming, piecewise-affine finite elements are used in the surrogate domain, although the extension to higher-order elements is possible. Due to the presence of corners and edges, the exact solution may not exhibit full elliptic regularity. This impacts the accuracy of the shifted boundary conditions to be applied on the surrogate boundary, since these are obtained by performing a truncated Taylor expansion therein. The consistency of the SBM discretization, and ultimately its convergence properties, depend on the behavior of the remainder in such expansion; this is thoroughly investigated in [1].

Together with the coercivity property discussed above, these consistency results allow us to prove convergence of the discretization in a suitable ‘energy’ norm with optimal rate. By a duality argument, we also establish an enhanced rate of decay of the error in the L^2 -norm, which however is optimal only if the shift direction is that of the normal vector to the surrogate boundary; since numerical results clearly indicate optimality irrespective of the orientation of the shift direction, this leaves an open problem for future research.

It is worth mentioning that while we use the simplest model of Poisson’s or heat equations, the mathematical framework and the analysis described here can easily be extended to handle more complex models, such as Darcy’s, or Stokes’, or elasticity equations.

This thesis is organized as follows: Chapter 2 introduces the Shifted Boundary Method for stationary problems, describing in particular how the exact boundary condition is mapped to the surrogate boundary. Coercivity and continuity are proved, establishing the well-posedness and stability of the discrete problem. Furthermore, the behavior of the consistency error is analyzed, demonstrating optimal convergence rates in the ‘energy’ norm and providing an enhanced L^2 -norm estimate. These theoretical findings are corroborated by numerical tests on square, circular, elliptical, and quarter-circle domains.

Chapter 3 investigates the heat equation, leveraging the results obtained for the stationary case. Time discretization is introduced via the method of lines; this approach involves first applying a semi-discretization in the spatial variables, and subsequently solving the resulting system of ordinary differential equations in time using established time-stepping schemes. Specifically, the Backward (Implicit)

Euler and Crank-Nicolson schemes are detailed. The stability, consistency, and convergence of the SBM-Backward Euler scheme are rigorously proved. To validate the theoretical findings, numerical tests, considering $\Delta t \sim h$ alongside other parameter combinations, are conducted on the same domains used for the stationary problems.

Finally, in Chapter 4 introduces the initial-boundary value problem for the heat equation on moving domains. To describe the domain's motion, the Arbitrary Lagrangian-Eulerian (ALE) framework is adopted. A stability estimate for the infinite-dimensional problem is derived; notably, under homogeneous boundary conditions, this estimate is considered ALE-free, as it is independent of the Eulerian velocity $\boldsymbol{w}$. In contrast to the previous chapter, the transverse method of lines (Rothe's method) is employed. Stability estimates for the ideal time, practical time, and fully discrete schemes are rigorously proved, highlighting that the ideal time discretization retains the ALE-free property. Furthermore, the consistency and convergence of the SBM-ALE scheme are analyzed. The theoretical framework is ultimately validated through numerical tests on square, L-shaped, and circular moving domains.

Chapter 2

SBM for stationary problems

In this chapter, the *Shifted Boundary Method* is introduced and analyzed, with a specific focus on its application to the Poisson equation. The fundamental principles of the method are presented, together with its main distinctive features which have been described in [1] over the past years.

It is important to remark that the main focus of this thesis is the solution of time dependent problems. For such problems, it should be noted that the spatial discretization coincides with that of the Poisson equation. For this reason, in this chapter we consider the stationary problem, since the results obtained here will also be employed in the time dependent analysis developed in Chapter 3.

2.1 The Shifted Boundary Method

Let Ω be a bounded and connected open region in $\mathbb{R}^n$ ($n = 2$ or 3), with Lipschitz boundary $\Gamma = \partial\Omega$; let Γ be formed by C^2 curves (in two dimensions) or surfaces (in three dimensions), which intersect at a finite number of vertices (in two dimensions) or edges (in three dimensions), and let $\mathbf{n}$ denote the outward-oriented unit normal vector to Γ . We assume that the physical problem of interest has been adimensionalized in such a way that $\text{diam}(\Omega) \simeq 1$.

We aim at numerically solving the Dirichlet boundary value problem for the Poisson equation

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= g && \text{on } \Gamma, \end{aligned} \tag{2.1}$$

where the data $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$ satisfy additional regularity assumptions

that will be made precise later on. In view of the application of the finite element method, we consider a closed domain $\mathcal{D}$ such that $\text{clos}(\Omega) \subseteq \mathcal{D}$ and we introduce a family $\mathcal{T}_h$ of admissible and shape-regular triangulations of $\mathcal{D}$. Then, we restrict each triangulation by selecting those elements that are contained in $\text{clos}(\Omega)$, i.e., we form

$$\tilde{\mathcal{T}}_h := \{T \in \mathcal{T}_h : T \subset \text{clos}(\Omega)\}. \quad (2.2)$$

This identifies the *surrogate domain*

$$\tilde{\Omega}_h := \text{int}\left(\bigcup_{T \in \tilde{\mathcal{T}}_h} T\right) \subseteq \Omega, \quad (2.3)$$

with *surrogate boundary* $\tilde{\Gamma}_h := \partial\tilde{\Omega}_h$ and outward-oriented unit normal vector $\tilde{\mathbf{n}}$

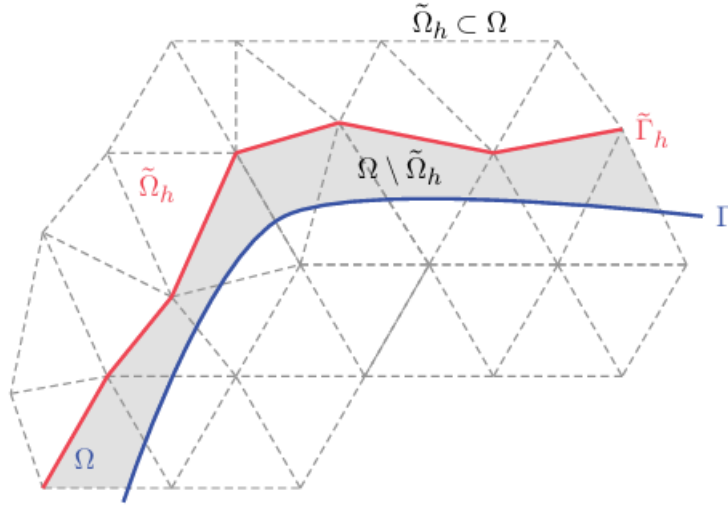


Figure 2.1: The true domain Ω , the surrogate domain $\tilde{\Omega}_h \subset \Omega$ and their boundaries $\tilde{\Gamma}_h$ and Γ . Reproduced from [1].

to $\tilde{\Gamma}_h$. Obviously, $\tilde{\Gamma}_h$ is an admissible and shape-regular triangulation of $\tilde{\Omega}_h$ (see Figure 2.1). As usual, we indicate by h_T the diameter of an element $T \in \tilde{\mathcal{T}}_h$ and by h the piecewise constant function in $\tilde{\Omega}_h$ such that $h|_T = h_T$ for all $T \in \tilde{\mathcal{T}}_h$. We introduce the following mesh parameters:

$$h_\Gamma := \max_{T \in \tilde{\mathcal{T}}_h : T \cap \tilde{\Gamma}_h \neq \emptyset} h_T, \quad h_\Omega := \max_{T \in \tilde{\mathcal{T}}_h} h_T. \quad (2.4)$$

We want to discretize Problem (2.1) in $\tilde{\Omega}_h$ rather than in Ω . Suppose that we knew the value of the exact solution u on $\tilde{\Gamma}_h$, say $u|_{\tilde{\Gamma}_h} = \tilde{g}$; then, we could consider the problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \tilde{\Omega}_h, \\ u &= \tilde{g} \quad \text{on } \tilde{\Gamma}_h, \end{aligned} \quad (2.5)$$

and discretize it by a Galerkin method using finite element trial and test functions on the triangulation $\tilde{\mathcal{T}}_h$. In particular, we could consider Nitsche's approach [22], which gains in flexibility by enforcing the boundary conditions in a weak manner by penalization. It is based on the following equivalent formulation of Problem (2.5)

$$\begin{aligned} (\nabla u, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} u, v)_{0, \tilde{\Gamma}_h} - (f, v)_{0, \tilde{\Omega}_h} \\ - (u - \tilde{g}, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1}(u - \tilde{g}), v)_{0, \tilde{\Gamma}_h} = 0, \end{aligned} \quad (2.6)$$

which hold for any sufficiently smooth test function v in $\tilde{\Omega}_h$; here, $\partial_{\tilde{\mathbf{n}}} = \tilde{\mathbf{n}} \cdot \nabla$ denotes a normal derivative to $\tilde{\Gamma}_h$, whereas $\gamma > 0$ is the (for the moment, arbitrary) penalization parameter. Introducing the finite dimensional subspace of $H^1(\tilde{\Omega}_h)$ made of continuous, piecewise affine functions on the triangulation $\tilde{\mathcal{T}}_h$,

$$V_h = V_h(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) := \{v \in H^1(\tilde{\Omega}_h) : v|_T \in \mathbb{P}_1(T), \forall T \in \tilde{\mathcal{T}}_h\}, \quad (2.7)$$

we could consider the following Galerkin discretization of Problem (2.5):

$$\begin{aligned} \text{Find } \tilde{u}_h \in V_h \text{ such that } \forall v_h \in V_h \\ (\nabla \tilde{u}_h, \nabla v_h)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} \tilde{u}_h, v_h)_{0, \tilde{\Gamma}_h} - (\tilde{u}_h, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \tilde{u}_h, v_h)_{0, \tilde{\Gamma}_h} \\ = (f, v_h)_{0, \tilde{\Omega}_h} - (\tilde{g}, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \tilde{g}, v_h)_{0, \tilde{\Gamma}_h}. \end{aligned} \quad (2.8)$$

To make this scheme practically feasible, we need a consistent approximation of $\tilde{g}$. To obtain it, let us select a mapping

$$\mathbf{M}_h : \tilde{\Gamma}_h \rightarrow \Gamma, \quad \tilde{\mathbf{x}} \mapsto \mathbf{x} \quad (2.9)$$

which associates to any point $\tilde{\mathbf{x}}$ on the surrogate boundary a point $\mathbf{x} = \mathbf{M}_h(\tilde{\mathbf{x}})$ on the true boundary. For example, $\mathbf{x}$ can be chosen as the closest point to $\tilde{\mathbf{x}}$ on Γ . We postpone to Section 2.1.1 a possible precise definition of a mapping $\mathbf{M}_h$ that works in practice with smooth boundary. The mapping (2.9) can be characterized through a distance vector function $\mathbf{d}_{\mathbf{M}_h}$ defined by

$$\mathbf{d}_{\mathbf{M}_h} := \mathbf{x} - \tilde{\mathbf{x}} = [\mathbf{M}_h - \mathbf{I}](\tilde{\mathbf{x}}). \quad (2.10)$$

For the sake of simplicity, we will actually set $\mathbf{d} = \mathbf{d}_{\mathbf{M}_h}$, so that we will write $\mathbf{x} = \tilde{\mathbf{x}} + \mathbf{d}(\tilde{\mathbf{x}})$. It will be useful to write $\mathbf{d} = \|\mathbf{d}\| \boldsymbol{\nu}$, where $\boldsymbol{\nu}$ is a unit vector defined on $\tilde{\Gamma}_h$. Note that $\boldsymbol{\nu}(\tilde{\mathbf{x}})$ may differ from both $\tilde{\mathbf{n}}(\tilde{\mathbf{x}})$ and $\mathbf{n}(\mathbf{M}_h(\tilde{\mathbf{x}}))$ (where $\mathbf{n}$ denotes the normal to the exact boundary Γ).

Assuming that u is sufficiently smooth in the strip between $\tilde{\Gamma}_h$ and Γ , so to admit a first-order Taylor expansion pointwise, we can write

$$g(\mathbf{x}) = u(\mathbf{x}) = u(\tilde{\mathbf{x}} + \mathbf{d}(\tilde{\mathbf{x}})) = u(\tilde{\mathbf{x}}) + (\nabla u \cdot \mathbf{d})(\tilde{\mathbf{x}}) + (R(u, \mathbf{d}))(\tilde{\mathbf{x}}) ,$$

where the remainder $R(u, \mathbf{d})$ satisfies $|R(u, \mathbf{d})| = o(\|\mathbf{d}\|)$ as $\|\mathbf{d}\| \rightarrow 0$. Equivalently, we can write

$$u(\tilde{\mathbf{x}}) = g(\mathbf{M}_h(\tilde{\mathbf{x}})) - (\nabla u \cdot \mathbf{d})(\tilde{\mathbf{x}}) - (R(u, \mathbf{d}))(\tilde{\mathbf{x}}) .$$

For the sake of simplicity, let us introduce the function on $\tilde{\Gamma}_h$

$$\bar{g}(\tilde{\mathbf{x}}) := g(\mathbf{M}_h(\tilde{\mathbf{x}})) ; \quad (2.11)$$

then, we see that the trace $\tilde{g}$ of u on $\tilde{\Gamma}_h$ satisfies

$$\tilde{g} = \bar{g} - \nabla u \cdot \mathbf{d} - R(u, \mathbf{d}) . \quad (2.12)$$

In conclusion, on $\tilde{\Gamma}_h$ it holds

$$0 = u - \tilde{g} = u + \nabla u \cdot \mathbf{d} - \bar{g} + R(u, \mathbf{d}) = S_h u - \bar{g} + R_h u , \quad (2.13)$$

where we have introduced the boundary operator on $\tilde{\Gamma}_h$

$$S_h v := v + \nabla v \cdot \mathbf{d} , \quad (2.14)$$

and $R_h u$ is a short-hand notation for the Taylor remainder $R(u, \mathbf{d})$. Neglecting the higher-order term (with respect to $\|\mathbf{d}\|$) in (2.13) and recalling (2.6), we deduce that the exact solution u satisfies the approximate equations

$$\begin{aligned} (\nabla u, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} u, v)_{0, \tilde{\Gamma}_h} - (f, v)_{0, \tilde{\Omega}_h} \\ - (S_h u - \bar{g}, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1}(S_h u - \bar{g}), S_h v)_{0, \tilde{\Gamma}_h} \approx 0 \end{aligned} \quad (2.15)$$

for any sufficiently smooth test function v in $\tilde{\Omega}_h$. Note that in the last inner product we have replaced v by $S_h v$ in order to obtain a symmetric stabilization term; on the other hand, unlike (2.6), the left-hand side of (2.15) is no longer symmetric in u and v .

Based on these approximate equations, we are led to introduce the following Galerkin discretization of Problem (2.5):

$$\begin{aligned} \text{Find } u_h \in V_h \text{ such that } \forall v_h \in V_h \\ (\nabla u_h, \nabla v_h)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} u_h, v_h)_{0, \tilde{\Gamma}_h} - (S_h u_h, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} S_h u_h, S_h v_h)_{0, \tilde{\Gamma}_h} \\ = (f, v_h)_{0, \tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \bar{g}, S_h v_h)_{0, \tilde{\Gamma}_h} . \end{aligned} \quad (2.16)$$

In view of the subsequent analysis, it is convenient to introduce the bilinear form

$$a_h(w, v) := (\nabla w, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} w, v)_{0, \tilde{\Gamma}_h} - (S_h w, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} S_h w, S_h v)_{0, \tilde{\Gamma}_h}, \quad (2.17)$$

and the linear form

$$F_h(v) := (f, v)_{0, \tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \bar{g}, S_h v)_{0, \tilde{\Gamma}_h}, \quad (2.18)$$

so that the proposed SBM discretization can be written in compact form as follows:

$$\text{Find } u_h \in V_h \text{ such that } a_h(u_h, v_h) = F_h(v_h) \quad \forall v_h \in V_h. \quad (2.19)$$

Note that the form a_h is well defined for all functions in $H^1(\tilde{\Omega}_h)$ whose gradient admits a trace in $L^2(\tilde{\Gamma}_h)$. This aspect will be further investigated in Sect. 2.2.

2.1.1 Definition of M_h for some smooth boundary

Let us presents a possible definition of the operator M_h for a two dimensional boundary Γ . As smooth boundary, we will consider circles or elliptic curves. In this case, the real boundary $\Gamma \subset \mathbb{R}^2$ can be parametrized as

$$\begin{cases} x = a \cos t \\ y = b \sin t \end{cases} \quad \text{with } \mathbf{x} = (x, y) \in \mathbb{R}^2, \quad t \in [0, 2\pi] \quad (2.20)$$

where a, b are strictly positive parameters ($r = a = b$ for the circle).

Given $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y}) \in \tilde{\Gamma}_h$, define

$$\phi(t) := \|\tilde{\mathbf{x}} - \mathbf{x}(t)\|^2 = (\tilde{x} - a \cos t)^2 + (\tilde{y} - b \sin t)^2, \quad (2.21)$$

where $\mathbf{x}(t) = (a \cos t, b \sin t)$ parametrizes the (smooth) ellipse Γ and $\|\cdot\|$ is the Euclidean norm.

The derivative of ϕ with respect to t is

$$\phi'(t) = -2a(a \cos t - \tilde{x}) \sin t + 2b(b \sin t - \tilde{y}) \cos t, \quad (2.22)$$

and the second derivative is

$$\phi''(t) = 4a^2 \sin^2 t - 2a^2 + 2a\tilde{x} \cos t - 4b^2 \sin^2 t + 2b^2 + 2b\tilde{y} \sin t. \quad (2.23)$$

Recalling (2.9), the closest-point projection M_h must be defined as the point on Γ which minimizes ϕ . A convenient definition is

$$t^* := \arg \min_{t \in [0, 2\pi]} \phi(t), \quad M_h(\tilde{\mathbf{x}}) := \mathbf{x}(t^*) = (a \cos t^*, b \sin t^*). \quad (2.24)$$

Equivalently, t^* is a solution of $\phi'(t) = 0$ which gives a *global* minimizer: if $\phi'(t^*) = 0$ and $\phi''(t^*) > 0$ (local minimum), then one must still check that $\phi(t^*)$ is the smallest among all critical values (over the compact interval $[0, 2\pi]$).

In order to find t^* , with (2.22) we have to solve the following trigonometric equation

$$2b\tilde{y}\cos t - 2a\tilde{x}\sin t + 2(a^2 - b^2)\cos t \sin t = 0. \quad (2.25)$$

This is a non trivial equation that could be numerically solved (e.g. with the Tangent-Newton method).

Remark 1. Solving (2.25) with Newton method needs a particular initial guess $t_0 = \arctan(\frac{ay_0}{bx_0})$ (see [24]). In [11] it is presented a similar definition of the distance function to (2.21). It is also presented the bisection method which it is the most robust method to solve this minimization problem.

Remark 2. A common alternative is to approximate the projection numerically by sampling the ellipse through a uniform parametrization

$$t_i = \frac{2\pi i}{N}, \quad i = 0, \dots, N,$$

and selecting the point that minimizes the discrete distance function

$$\phi(t_i) = \|(a \cos t_i, b \sin t_i) - \mathbf{x}_0\|^2.$$

However, this approach has computational cost $\mathcal{O}(N)$ and accuracy $\mathcal{O}(1/N)$, requiring very large values of N to achieve high precision. Consequently, when the projection must be computed for a large number of points (as in SBM), the total cost becomes prohibitive.

In contrast, solving numerically reduces the problem to the solution of a single scalar equation, with cost $\mathcal{O}(1)$ per projection. Newton's method converges quadratically and requires only a few iterations, independent of the desired accuracy. In practice, speedups between $50\times$ and $200\times$ are typical compared to a discretized search along the boundary, while maintaining full machine precision.

The presented approach provides an exact and extremely efficient method to compute the closest point projection on an ellipse. For this reason it is particularly suitable for FEM algorithms relying on surrogate or curved boundaries, where the projection must be evaluated repeatedly and with high accuracy.

Computing the closest point projection means to impose $\boldsymbol{\nu}(\tilde{\mathbf{x}}) = \mathbf{n}(\tilde{\mathbf{x}})$ for $\tilde{\mathbf{x}} \in \tilde{\Gamma}_h$ and finding $\mathbf{n}(\tilde{\mathbf{x}})$ is equivalent to what it has been discussed above. On the other hand, instead of computing the minimal distance maybe we could simplify just using $\boldsymbol{\nu}(\tilde{\mathbf{x}}) = \tilde{\mathbf{n}}(\tilde{\mathbf{x}})$. However, it could produce unphysical boundary conditions located far away from the edge, as shown in Figure 2.2. Let us introduce a different and simpler approach which will drastically reduce the computational cost of this task.

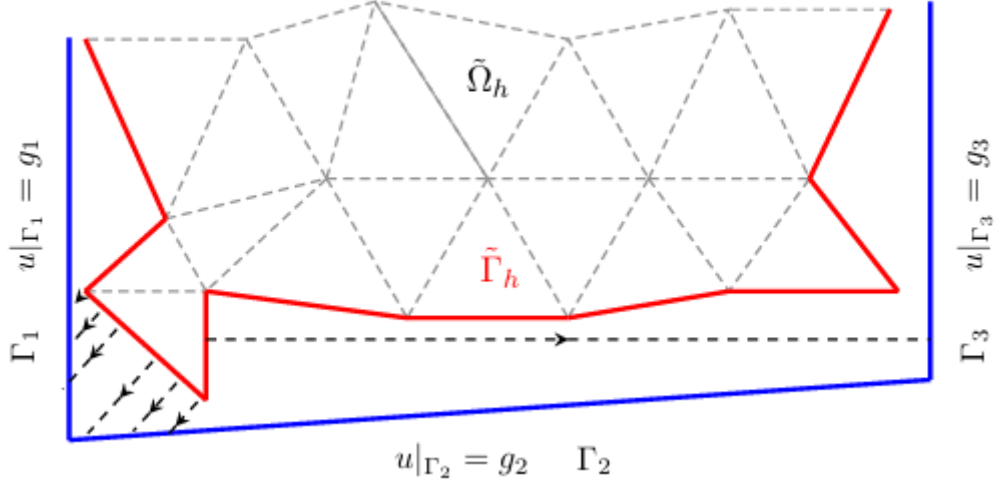


Figure 2.2: Complications arising from the choice $\boldsymbol{\nu} = \tilde{\boldsymbol{n}}$. Reproduced from [1].

Approximation of the mapping M_h in 2.24

First of all, we convert (2.20) in a two variable implicit equation using the classical ellipse formula. We define the contour line z , passing through a generic point $\hat{\boldsymbol{x}}$ in $\mathbb{R}^2$, as

$$z(\hat{\boldsymbol{x}}) := \frac{\hat{x}^2}{a^2} + \frac{\hat{y}^2}{b^2} - 1 \quad \forall \hat{\boldsymbol{x}} \in \mathbb{R}^2. \quad (2.26)$$

Note that for all $\boldsymbol{x}$ on the real boundary Γ we have $z(\boldsymbol{x}) = 0$. Moreover, for all of these points the normal vector $\boldsymbol{n}$ is computed as

$$\boldsymbol{n}(\boldsymbol{x}) = \frac{\nabla z(\boldsymbol{x})}{\|\nabla z(\boldsymbol{x})\|}, \quad \text{where} \quad \nabla z(\boldsymbol{x}) = \left(\frac{2x}{a^2}, \frac{2y}{b^2} \right)^T. \quad (2.27)$$

For each point $\tilde{\boldsymbol{x}}$ on the surrogate boundary $\tilde{\Gamma}_h$, we introduce the following projection on the real boundary Γ ,

$$\boldsymbol{M}_h^z(\tilde{\boldsymbol{x}}) := \tilde{\boldsymbol{x}} + \alpha \nabla z(\tilde{\boldsymbol{x}}), \quad (2.28)$$

with the step length $\alpha > 0$. In order to find α , we impose for all $\mathbf{x} \in \Gamma$:

$$\begin{aligned}
 z(\mathbf{x}) &= z(\mathbf{M}_h^z(\tilde{\mathbf{x}})) = 0 \\
 &= \frac{(\tilde{x} + \alpha \nabla z(\tilde{x}))^2}{a^2} + \frac{(\tilde{y} + \alpha \nabla z(\tilde{y}))^2}{b^2} - 1 \\
 &= \frac{(\tilde{x} + \alpha \frac{2\tilde{x}}{a^2})^2}{a^2} + \frac{(\tilde{y} + \alpha \frac{2\tilde{y}}{b^2})^2}{b^2} - 1 \\
 &= \underbrace{\alpha^2 \left(\frac{4\tilde{x}^2}{a^6} + \frac{4\tilde{y}^2}{b^6} \right)}_A + \underbrace{\alpha \left(\frac{4\tilde{x}^2}{a^4} + \frac{4\tilde{y}^2}{b^4} \right)}_B + \underbrace{\left(\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} - 1 \right)}_C = 0.
 \end{aligned}$$

Finding the step length α is equivalent to solve the second order equation above and we obtain

$$\alpha_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A},$$

where only $\alpha_2 = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$ is feasible since $-B$ is always negative. Hence, if we subtract the square root (always positive) we get a step length $\alpha < 0$ that means moving along the opposite direction.

In the end, this approach simplify a lot the computation that is reduced to simple algebraic operations. Although it is not a closest point projection, we are moving in a direction different from $\mathbf{n}$ and $\tilde{\mathbf{n}}$ (see Figure 2.3). Remember that for the SBM we do not have to use a closest point projection. For instance, this is a practical case where the $\mathbf{M}_h^z$ proposed is computationally cheap and easy to implement which still avoid the issue of Figure 2.2. In addition, when h goes to zero, the distance computed with this method will be close to the closest point projection.

2.2 Coercivity and continuity analysis

As a first result, we establish coercivity and continuity properties of the bilinear form a_h .

To this end, we first recall well-known scaled trace inequalities in shape-regular triangulations, which will be used in different forms throughout the paper.

Property 1. *There exists a constant $C_I > 0$ independent of the meshsize such that for any $T \in \tilde{\mathcal{T}}_h$ and any edge/face $E \subset \partial T$ it holds*

$$\|h_T^{1/2} \nabla w\|_{0,E}^2 \leq C_I \|\nabla w\|_{0,T}^2 \quad \forall w \in \mathbb{P}_1(T). \quad (2.29)$$

This immediately gives

$$\|h_T^{1/2} \nabla w\|_{0,\tilde{\Gamma}_h}^2 \leq C_I \|\nabla w\|_{0,\tilde{\Omega}_h}^2 \quad \forall w \in V_h. \quad (2.30)$$

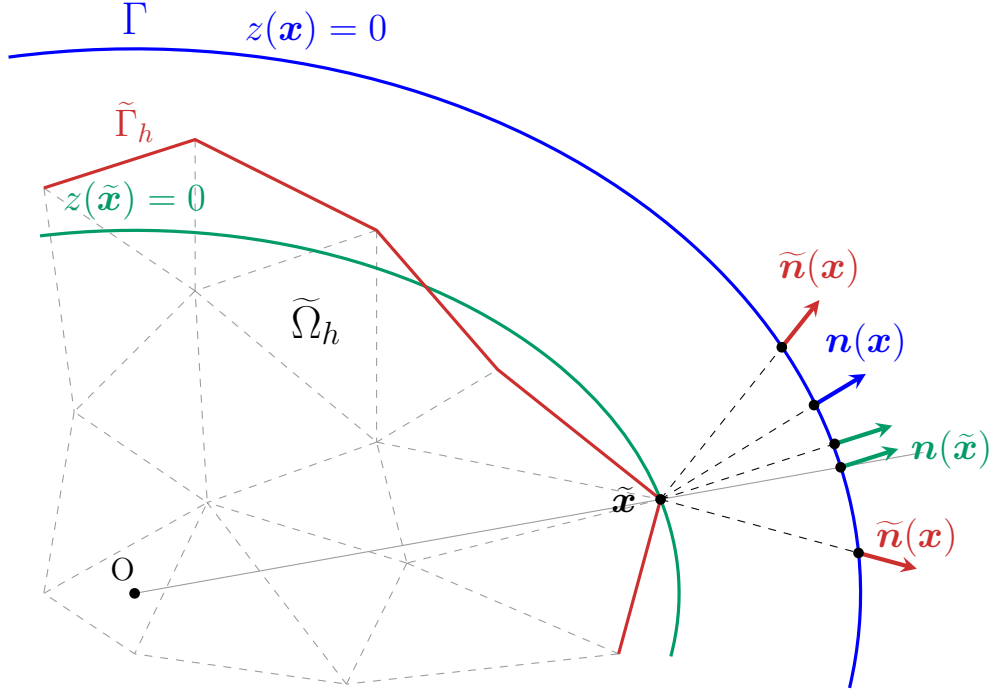


Figure 2.3: Projection comparison varying ν . $\tilde{n}(x)$ is the normal vector of the surrogate boundary $\tilde{\Gamma}_h$. $n(x)$ is the normal vector in the direction of the closest point projection (2.27 computed in x). $n(\tilde{x})$ is the normal vector using M_h^z (2.27 computed in $\tilde{x}$).

Moreover, we make the following assumption, on which all subsequent analysis will rely.

Assumption 1. The distance function d satisfies the condition

$$\|h^{-1}d\|_{L^\infty(\tilde{\Gamma}_h)} \leq \frac{\beta}{C_I} \quad \text{for some } \beta \in (0,1). \quad (2.31)$$

We exploit Assumption 1 to prove the uniform coercivity of the form a_h for sufficiently refined grids.

Proposition 1. Under Assumption 1, if the Nitsche penalization parameter γ satisfies $\gamma > \frac{2C_I}{1-\beta}$, there exists a constant $\alpha > 0$ independent of the mesh size, such that

$$a_h(v_h, v_h) \geq \alpha \|v_h\|_a^2, \quad \forall v_h \in V_h, \quad (2.32)$$

where

$$\|v\|_a^2 := \|\nabla v\|_{0,\tilde{\Omega}_h}^2 + \|h^{-1/2}S_h v\|_{0,\tilde{\Gamma}_h}^2 \quad (2.33)$$

Proof. One has

$$\begin{aligned}
 a_h(v_h, v_h) &= \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 - (\partial_{\tilde{\mathbf{n}}} v_h, v_h)_{0,\tilde{\Gamma}_h} - (S_h v_h, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} + \gamma \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2 \\
 &= \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 - (\partial_{\tilde{\mathbf{n}}} v_h, v_h)_{0,\tilde{\Gamma}_h} - (\nabla v_h \cdot \mathbf{d}, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} + (\nabla v_h \cdot \mathbf{d}, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} \\
 &\quad - (S_h v_h, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} + \gamma \|h^{-1} S_h v_h\|_{0,\tilde{\Gamma}_h}^2 \\
 &= \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 - 2(S_h v_h, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} \\
 &\quad + (\nabla v_h \cdot \mathbf{d}, \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} + \gamma \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2 \\
 &= \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 - 2(h^{-1/2} S_h v_h, h^{1/2} \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} \\
 &\quad + (\|h^{-1} \mathbf{d}\|_{L^\infty(\tilde{\Gamma}_h)} h^{1/2} \partial_{\nu} v_h, h^{1/2} \partial_{\tilde{\mathbf{n}}} v_h)_{0,\tilde{\Gamma}_h} + \gamma \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2,
 \end{aligned}$$

where we have used $\mathbf{d} = \|d\|\boldsymbol{\nu}$. Recalling (2.31) and applying (2.30), we get

$$\begin{aligned}
 a_h(v_h, v_h) &\geq \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 - 2C_I^{1/2} \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h} \|\nabla v_h\|_{0,\tilde{\Omega}_h} \\
 &\quad - \frac{\beta}{C_I} C_I \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 + \gamma \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2.
 \end{aligned}$$

By Young inequality the second term of the right-hand side can be written as $2C_I^{1/2} \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h} \|\nabla v_h\|_{0,\tilde{\Omega}_h} \leq \varepsilon C_I \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 + \frac{1}{\varepsilon} \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2$, whence

$$a_h(v_h, v_h) \geq (1 - \varepsilon C_I - \beta) \|\nabla v_h\|_{0,\tilde{\Omega}_h}^2 + (\gamma - \frac{1}{\varepsilon}) \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h}^2.$$

Choosing $\varepsilon = \frac{1-\beta}{2C_I}$, we obtain the result with $\alpha = \min\{\frac{1}{2}, \gamma - \frac{2C_I}{1-\beta}\}$. $\square$

A first consequence of Proposition 1 is the well-posedness and numerical stability of the SBM discretization (2.16) with respect to the norm $\|u_h\|_a$, which uniformly dominates the H^1 -norm of u_h (see Remark 4 below).

Proposition 2. *u_h solution of 2.19 satisfies the bound*

$$\|u_h\|_a \leq \frac{1}{\alpha} \|F_h\|_{V_h'} \quad (2.34)$$

Proof. First of all, considering $v_h = u_h$, we apply (2.32) on the left hand-side of (2.19)

$$\alpha \|u_h\|_a^2 \leq a_h(u_h, u_h) = F_h(u_h) \leq |F_h(u_h)|.$$

Recalling (2.18),

$$\begin{aligned}
 F_h(u_h) &= (f, u_h)_{0,\tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{\mathbf{n}}} u_h)_{0,\tilde{\Gamma}_h} + \gamma (h^{-1} \bar{g}, S_h u_h)_{0,\tilde{\Gamma}_h} \\
 &\leq \|f\|_{0,\tilde{\Omega}_h} \|u_h\|_{0,\tilde{\Omega}_h} + C_I^{1/2} \|h^{-1/2} \bar{g}\|_{0,\tilde{\Omega}_h} \|\nabla u_h\|_{0,\tilde{\Omega}_h} \\
 &\quad + \gamma \|h^{-1/2} \bar{g}\|_{0,\tilde{\Gamma}_h} \|h^{-1/2} S_h u_h\|_{0,\tilde{\Gamma}_h},
 \end{aligned}$$

where we have used Cauchy-Schwarz inequality for all the terms and (2.30) for the second term of the right-hand side. Moreover, considering the Poncaré's lemma we have the following bound

$$\|u_h\|_{0,\tilde{\Omega}_h} \leq C_P \left(\|\nabla u_h\|_{0,\tilde{\Omega}_h} + \|u_h\|_{0,\tilde{\Gamma}_h} \right).$$

Furthermore we get

$$\begin{aligned} \|u_h\|_{0,\tilde{\Omega}_h} &\leq C_P \left(\|\nabla u_h\|_{0,\tilde{\Omega}_h} + \|u_h \pm \nabla u_h \cdot \mathbf{d}\|_{0,\tilde{\Gamma}_h} \right) \\ &\leq C_P \left(\|\nabla u_h\|_{0,\tilde{\Omega}_h} + \|S_h u_h\|_{0,\tilde{\Gamma}_h} + \|\nabla u_h \cdot \mathbf{d}\|_{0,\tilde{\Gamma}_h} \right) \\ &\leq C_P \left(\|\nabla u_h\|_{0,\tilde{\Omega}_h} + \|h^{1/2}(h^{-1/2} S_h u_h)\|_{0,\tilde{\Gamma}_h} + \frac{\beta}{C_I} C_I^{1/2} \|h^{1/2} \nabla u_h\|_{0,\tilde{\Omega}_h} \right) \end{aligned}$$

using (2.30) again and (2.31). Finally we have

$$\begin{aligned} |F_h(u_h)| &\leq C_P \|f\|_{0,\tilde{\Omega}_h} \left(\|\nabla u_h\|_{0,\tilde{\Omega}_h} + C \|h^{1/2} u_h\|_a \right) + C_I^{1/2} \|h^{-1/2} \bar{g}\|_{0,\tilde{\Omega}_h} \|\nabla u_h\|_{0,\tilde{\Omega}_h} \\ &\quad + \gamma \|h^{-1/2} \bar{g}\|_{0,\tilde{\Gamma}_h} \|h^{-1/2} S_h u_h\|_{0,\tilde{\Gamma}_h} \\ &\leq C' \underbrace{(\|f\|_{0,\tilde{\Omega}_h} + \|h^{-1/2} \bar{g}\|_{0,\tilde{\Gamma}_h})}_{\|F_h\|_{V'_h}} \|u_h\|_a \end{aligned}$$

whence the result. $\square$

Note that this is not a classical stability results because the term $\|F_h\|_{V'_h}$ is not independent on the meshsize h .

Remark 3. A different formulation of Nitsche's method, in which the penalty parameter γ is not used, is also presented in the literature. By neglecting the penalty term and changing the sign of the fourth term in the left-hand side of (2.6), we obtain

$$(\nabla u, \nabla v)_{0,\tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} u, v)_{0,\tilde{\Gamma}_h} - (f, v)_{0,\tilde{\Omega}_h} + (u - \tilde{g}, \partial_{\tilde{\mathbf{n}}} v)_{0,\tilde{\Gamma}_h} = 0.$$

With this formulation, the bilinear form a_h is automatically coercive, with a coercivity constant $\alpha^* > 0$, and no conditions on γ are required. From now on, we refer to this formulation whenever we consider the case $\gamma = 0$.

We now turn to discuss the continuity of the form a_h . In view of the convergence analysis, we consider a_h as defined on a product space $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \times V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$, where $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ is an infinite dimensional subspace of $H^1(\tilde{\Omega}_h)$, depending upon the triangulation $\tilde{\mathcal{T}}_h$ and containing V_h , made of functions for which the trace of the gradient on $\tilde{\Gamma}_h$ is well-defined and controlled in $L^2(\tilde{\Gamma}_h)$. For instance, $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ could be the space of functions with (broken) H^2 -regularity in $\tilde{\Omega}_h$. However,

since the solution u of our semi-discrete problem may exhibit singularities at the boundary Γ due to the presence of corners or edges, we choose a space of weaker regularity, that we are going to define.

Let us introduce a ‘reference’ element $\hat{T}$ for all our triangulations; as usual, we assume that $\hat{T}$ has unitary diameter. Let $\hat{E} \subset \partial\hat{T}$ any edge/face of $\hat{T}$. Let s, p be real numbers satisfying $0 < s \leq 1$ and $1 \leq p \leq \infty$. Then, if $t = s - \frac{1}{p}$ satisfies $t > 0$, the trace theorem (see e.g.) guarantees that the trace operator $\tau_{\hat{E}}v = v|_{\hat{E}}$ maps $W^{s,p}(\hat{T})$ on $W^{t,p}(\hat{E})$ continuously. On the other hand, the Sobolev imbedding theorem guarantees that $W^{t,p}(\hat{E}) \subseteq L^2(\hat{E})$ with continuous inclusion, provided $s \geq \frac{n}{p} - \frac{n-1}{2}$; this condition holds true if $s > \frac{1}{p}$ with $p \geq 2$. Thus, $v|_{\hat{E}} \in L^2(\hat{E})$ if $v \in W^{s,p}(\hat{T})$, and the bound

$$\|v\|_{0,\hat{E}}^2 \leq \hat{C}(\|v\|_{0,\hat{T}}^2 + |v|_{s,p,\hat{T}}^2)$$

holds with a constant $\hat{C} > 0$ independent of v . Here, $|v|_{s,p,\hat{T}}$ denotes the fractional seminorm $\left(\int_{\hat{T} \times \hat{T}} \frac{|v(\mathbf{x}) - v(\mathbf{y})|^p}{|\mathbf{x} - \mathbf{y}|^{n+sp}} d\mathbf{x} d\mathbf{y}\right)^{1/p}$ when $s < 1$, or the L^p -norm of ∇v when $s = 1$. Next, consider an element $T \in \tilde{\mathcal{T}}_h$ with edge/face E : a scaling argument yields after a simple computation the existence of a constant $C > 0$ independent of h such that

$$\|h_T^{1/2}v\|_{0,E}^2 \leq C \left(\|v\|_{0,T}^2 + h_T^{2s+(1-2/p)n} |v|_{s,p,T}^2 \right), \quad \forall v \in W^{s,p}(T). \quad (2.35)$$

We apply such bound to each component of the gradient of a function, yielding the existence of a constant, still denoted C_I as in (2.29), such that for all $w \in H^1(T)$ with $\nabla w \in W^{s,p}(T)$ it holds

$$\|h_T^{1/2}\nabla w\|_{0,E}^2 \leq C \left(\|\nabla w\|_{0,T}^2 + h_T^{2s+(1-2/p)n} |\nabla w|_{s,p,T}^2 \right). \quad (2.36)$$

(Note indeed that this inequality reduces to (2.29) when $w \in \mathbb{P}_1(T)$.) Thereby, considering the strip of elements in $\tilde{\mathcal{T}}_h$ with at least one edge/face on $\tilde{\Gamma}_h$, i.e., the subset

$$\tilde{\Omega}_h^b := \bigcup \{T \in \tilde{\mathcal{T}}_h^b\} \quad \text{where} \quad \tilde{\mathcal{T}}_h^b := \{T \in \tilde{\mathcal{T}}_h : \text{meas}(\partial T \cap \tilde{\Gamma}_h) > 0\}, \quad (2.37)$$

we are led to the following definition of the space $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$.

Definition 1. Let s, p be real numbers satisfying $0 < s \leq 1$ and $2 \leq p \leq \infty$ with $s > \frac{1}{p}$. Then, set

$$V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) = V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h, s, p) := \{w \in H^1(\tilde{\Omega}_h) : \nabla w|_T \in W^{s,p}(T) \quad \forall T \in \tilde{\mathcal{T}}_h^b\} \quad (2.38)$$

equipped with semi-norm

$$|h^{s_p} \nabla w|_{s,p,\tilde{\Omega}_h^b} := \left(\sum_{T \in \tilde{\mathcal{T}}_h^b} |h_T^{s_p} \nabla w|_{s,p,T}^p \right)^{1/p}, \quad s_p := s + (\tfrac{1}{2} - \tfrac{1}{p})n \quad (2.39)$$

and norm

$$\|w\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}^2 := \|w\|_a^2 + |h^{s_p} \nabla w|_{s,p,\tilde{\Omega}_h^b}^2. \quad (2.40)$$

With such definitions, we use Hölder's inequality in (2.36) and the property that the measure $|\tilde{\Gamma}_h|$ of $\tilde{\Gamma}_h$ is uniformly bounded by $|\Gamma|$, to get the existence of a constant $\bar{C}_I > 0$ such that

$$\|h_T^{1/2} \nabla w\|_{0,\tilde{\Gamma}_h}^2 \leq \bar{C}_I \left(\|\nabla w\|_{0,\tilde{\Omega}_h}^2 + |h_T^{s_p} \nabla w|_{s,p,\tilde{\Omega}_h^b}^2 \right), \quad \forall w \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \quad (2.41)$$

which extends (2.30); indeed, note that $V_h \subset V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$; and $\|v_h\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)} = \|v_h\|_a$ for all $v_h \in V_h$. Note as well that $W^{1+s,p}(\tilde{\Omega}_h) \subset V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$.

Remark 4. The norm $\|w\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}$ uniformly controls from above the standard norm $\|w\|_{1,\tilde{\Omega}_h}$. Indeed, the latter is equivalent to the norm $\left(\|\nabla w\|_{0,\tilde{\Omega}_h}^2 + \|w\|_{0,\tilde{\Gamma}_h}^2 \right)^{1/2}$ observing that

$$\begin{aligned} \|w\|_{0,\Gamma} &\leq C_\tau \|w\|_{1,\Omega} \quad (\text{trace inequality}), \\ \|w\|_{0,\Omega} &\leq C_P (\|\nabla w\|_{0,\Omega} + \|w\|_{0,\Gamma}) \quad (\text{Poincaré's lemma}), \\ \|\nabla w\|_{0,\Omega} + \|w\|_{0,\Omega} &\leq \|\nabla w\|_{0,\Omega} + C_P (\|\nabla w\|_{0,\Omega} + \|w\|_{0,\Gamma}) \leq C_1 (\|\nabla w\|_{0,\Omega} + \|w\|_{0,\Gamma}), \\ \|\nabla w\|_{0,\Omega} + \|w\|_{0,\Gamma} &\leq \|\nabla w\|_{0,\Omega} + C_\tau (\|\nabla w\|_{0,\Omega} + \|w\|_{0,\Omega}) \leq C_2 (\|\nabla w\|_{0,\Omega} + \|w\|_{0,\Omega}). \end{aligned}$$

Moreover, one has by (2.41) and (2.31) (assuming without loss of generality that $h \leq 1$ in $\tilde{\Omega}_h$)

$$\begin{aligned} \|w\|_{0,\tilde{\Gamma}_h} &\leq \|h^{-1/2} w\|_{0,\tilde{\Gamma}_h} \leq \|h^{-1/2} S_h w\|_{0,\tilde{\Gamma}_h} + \|h^{-1/2} \nabla w \cdot \mathbf{d}\|_{0,\tilde{\Gamma}_h} \\ &\leq \|h^{-1/2} S_h w\|_{0,\tilde{\Gamma}_h} + \frac{\beta}{C_I} \bar{C}_I^{1/2} \left(\|\nabla w\|_{0,\tilde{\Omega}_h}^2 + |h^{s_p} \nabla w|_{s,p,\tilde{\Omega}_h^b}^2 \right)^{1/2}, \end{aligned}$$

whence the result.

Proposition 3. *The bilinear form a_h is defined in $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \times V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ and uniformly continuous therein; precisely, there exists $\|a_h\| > 0$ independent of h such that*

$$|a_h(w, v)| \leq \|a_h\| \|w\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)} \|v\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}, \quad \forall w, v \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h). \quad (2.42)$$

Proof. One has

$$\begin{aligned}
 a_h(w, v) &= (\nabla w, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} w, S_h v)_{0, \tilde{\Gamma}_h} + (\partial_{\tilde{\mathbf{n}}} w, \nabla v \cdot \mathbf{d})_{0, \tilde{\Gamma}_h} \\
 &\quad - (S_h w, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} S_h w, S_h v)_{0, \tilde{\Gamma}_h} \\
 &= (\nabla w, \nabla v)_{0, \tilde{\Omega}_h} - (h^{1/2} \partial_{\tilde{\mathbf{n}}} w, h^{-1/2} S_h v)_{0, \tilde{\Gamma}_h} - (h^{-1/2} S_h w, h^{1/2} \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} \\
 &\quad + (\|h^{-1} \mathbf{d}\|_{L^\infty(\tilde{\Gamma}_h)} h^{1/2} \partial_{\nu} w, h^{1/2} \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1/2} S_h w, h^{-1/2} S_h v)_{0, \tilde{\Gamma}_h}.
 \end{aligned}$$

One concludes, using again (2.41) and (2.31), that each term in absolute value is bounded above by the product $\|a_h\| \|w\|_{V(\tilde{\Omega}_h; \tilde{\Gamma}_h)} \|v\|_{V(\tilde{\Omega}_h; \tilde{\Gamma}_h)}$. The constants $\|a_h\|$ is defined as

$$\|a_h\| := \sup_{\substack{w, v \in V(\tilde{\Omega}_h; \tilde{\Gamma}_h) \\ w, v \neq 0}} \frac{|a_h(w, v)|}{\|w\|_{V(\tilde{\Omega}_h; \tilde{\Gamma}_h)} \|v\|_{V(\tilde{\Omega}_h; \tilde{\Gamma}_h)}}.$$

□

Remark 5. Recalling the general non-homogeneous Dirichlet problem, in a classical Galerkin method the solution is usually decomposed as $u_h = u_h^o + u_h^g$, where u_h^o is the solution of the homogeneous problem and u_h^g is a lifting function such that $u_h^g|_{\Gamma} = g$ and zero elsewhere. Hence, for instance considering $V = H^1(\Omega)$ and $V_h^o \subset V^o = H_0^1(\Omega)$, we have

$$a(u_h^o, v_h) = F(v_h) - a(u_h^g, v_h) \quad \forall v_h \in V_h^o,$$

and by the Lax-Milgram theorem,

$$\|u_h^o\|_V \leq \frac{1}{\alpha} \left(\|F\|_{V_h'} + \|a\| \|u_h^g\|_V \right)$$

with $\|a\|, \|F\|_{V_h'}$ the classical norm related to the bilinear form a and the linear form F respectively. Considering the right-hand side, some assumption such as $\|u_h^g\|_V \lesssim \|g\|_{H^{1/2}(\Gamma)}$ is required to preserve stability. However, in the SBM method all the information on g is embedded in the linear form $F_h(v_h)$; therefore, assumptions on $\|F_h\|_{V_h'}$ alone are sufficient.

2.3 Behavior of the Taylor remainder

This section is devoted to the analysis of the behavior of the Taylor remainder $R_h u = \bar{g} - S_h u$ in the expansion of u on $\tilde{\Gamma}_h$ introduced in (2.13). In particular, we are interested in estimating a weighted L^2 -norm of $R_h u$ along $\tilde{\Gamma}_h$ in terms of the mesh parameter h_Γ . To this end, we make Assumption 2, which will be used in the subsequent analysis.

Assumption 2. There exists a constant $\delta > 0$ independent of h such that

$$|(\tilde{\mathbf{n}} \cdot \boldsymbol{\nu})(\tilde{\mathbf{x}})| \geq \delta, \quad \forall \tilde{\mathbf{x}} \in T \cap \tilde{\Gamma}_h, \quad \forall T \in \tilde{\mathcal{T}}_h, \quad (2.43)$$

where $\tilde{\mathbf{n}}$ is the unit normal vector to the surrogate boundary $\tilde{\Gamma}_h$, and $\boldsymbol{\nu}$ is the unit vector aligned with the distance $\mathbf{d}$.

The condition prevents the distance vector on any edge $\tilde{E} \subset \tilde{\Gamma}_h$ to be co-linear with the edge.

We will use the following *sufficient* condition, which guarantees that both Assumption 1 and Assumption 2 (when the mapping $\mathbf{M}_h$ is the closest-point projection).

Assumption 3. There exists constants $c_d > 0$ and $\zeta > 0$ such that

$$\|\mathbf{d}(\tilde{\mathbf{x}})\| \leq c_d h_T^{1+\zeta}, \quad \forall \tilde{\mathbf{x}} \in T \cap \tilde{\Gamma}_h, \quad \forall T \in \tilde{\mathcal{T}}_h \quad (2.44)$$

This condition requires the distance $\|\mathbf{d}\|$ between $\tilde{\Gamma}_h$ and Γ to go to zero slightly faster than the local meshsize. It can be realized in practice by slightly shifting the nodes on $\tilde{\Gamma}_h$ towards Γ while refining the grid.

One can prove the following result (see [1] for the proof).

Proposition 4. Let $\Gamma_S \subseteq \Gamma$ be a portion of the physical boundary which admits a C^2 parametrization, and let $\tilde{\Gamma}_{h,S}$ be $\{\tilde{E} \subset \tilde{\Gamma}_h : \mathbf{M}_h(\tilde{E}) \subset \Gamma_S\}$. Assume that in a suitable neighborhood Ω_S of Γ_S in Ω , which contains the region between Γ_S and $\tilde{\Gamma}_{h,S}$, the exact solution satisfies $u \in W^{2,p}(\Omega_S)$ for some $p \in [2, \infty]$. Then, there exists a constant $C_p > 0$ independent of h_Γ such that if h_Γ is sufficiently small one has

$$\|h^{-1/2} R_h u\|_{0, \tilde{\Gamma}_{h,S}} \leq C_p h_\Gamma^{3/2-1/p+\zeta'} \|\mathcal{H}u\|_{L^p(\Omega_S)},$$

where $\zeta' = (2 - 1/p)\zeta > 0$, and $\mathcal{H}u$ is the Hessian matrix of u .

In particular, for $p = 2$, neglecting the contribution from ζ' we obtain

$$\|h^{-1/2} R_h u\|_{0, \tilde{\Gamma}_{h,S}} \leq C_2 h_\Gamma |u|_{2, \Omega_S},$$

whereas if $u \in W^{2,p}(\Omega_S)$ for any arbitrarily large p , then we can take $p = 1/\zeta'$ and get

$$\|h^{-1/2} R_h u\|_{0, \tilde{\Gamma}_{h,S}} \leq C_\infty h_\Gamma^{3/2} |u|_{W^{2,\infty}(\Omega_S)}.$$

It is also possible to consider the neighborhood of a singularity in Γ (a corner point in a polygonal domain, an edge or a vertex in a polyhedral domain) which is proposed in [1].

2.4 Consistency and convergence analysis - Space

From now on, we pose the following regularity assumption on the solution of our Dirichlet problem.

Assumption 4. The solution u of Problem (2.1) belongs to $W^{1+s,p}(\Omega)$ for some s, p satisfying the conditions stated in Definition 1.

If the boundary Γ is C^2 and the data satisfy $f \in L^p(\Omega)$ and $g \in W^{2-1/p,p}(\Gamma)$, then $u \in W^{2,p}(\Omega)$, i.e., the assumption is satisfied with $s = 1$. On the other hand, if Γ contains a finite number of corners (in two dimensions) or edges or vertices (in three dimensions) other comments are necessary (see [1]). In all cases, Assumption 4 is fulfilled.

As a consequence of Assumption 4, u belongs to the related space $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h, s, p)$ introduced in (2.38), that we will simply indicate by $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ whenever no confusion may arise. In order to estimate the discretization error $u - u_h$ in the ‘energy’ norm, we invoke Strang’s Second Lemma, which reads

$$\|u - u_h\|_a \leq \left(1 + \frac{\|a_h\|}{\alpha}\right) E_{a,h}(u) + \frac{1}{\alpha} E_{c,h}(u), \quad (2.45)$$

where $E_{a,h}(u)$ is the approximation error

$$E_{a,h}(u) := \inf_{v_h \in V_h} \|u - v_h\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}, \quad (2.46)$$

whereas $E_{c,h}(u)$ is the consistency error

$$E_{c,h}(u) := \sup_{v_h \in V_h} \frac{a_h(u, v_h) - F_h(v_h)}{\|v_h\|_a} \quad (2.47)$$

We are going to estimate both errors in terms of the meshsize h_Ω introduced in (2.4).

Proposition 5. *There exists a constant $C_a > 0$ independent of u and the meshsize such that*

$$E_{a,h}(u) \leq C_a h_\Omega^s |\nabla u|_{s,p,\tilde{\Omega}_h}. \quad (2.48)$$

Proof. Thanks to the assumptions on s, p in Definition 1, $W^{1+s,p}(\Omega) \subset C^0(\Omega)$ with continuous injection. Hence, the piecewise linear interpolant $I_h u$ at the nodes of $\tilde{\mathcal{T}}_h$ is well-defined in V_h . So, we will estimate $\|u - I_h u\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}$. To this end, we observe that on the reference element $\hat{T}$ there exists a constant $\hat{c} > 0$ such that

$$\|\hat{u} - \hat{I}\hat{u}\|_{1,\hat{T}} \leq \hat{c} |\hat{\nabla} \hat{u}|_{s,p,\hat{T}} \quad \forall \hat{u} \in W^{1+s,p}(\hat{T}),$$

where $\hat{I}$ denotes linear interpolation at the vertices of $\hat{T}$. Using the standard transformations from an element $T \in \tilde{\mathcal{T}}_h$ to the reference element and back, one gets

$$\|h^{-1}(u - I_h u)\|_{0,\tilde{\Omega}_h} + \|\nabla(u - I_h u)\|_{0,\tilde{\Omega}_h} \leq c h_{\Omega}^s |\nabla u|_{s,p,\tilde{\Omega}_h}. \quad (2.49)$$

Here and in the sequel, c denotes a positive constant independent of the meshsize, which may be different in different formulas. On the other hand,

$$\|h^{-1/2} S_h(u - I_h u)\|_{0,\tilde{\Gamma}_h} \leq \|h^{-1/2}(u - I_h u)\|_{0,\tilde{\Gamma}_h} + \|h^{-1/2} \|\mathbf{d}\| \nabla(u - I_h u)\|_{0,\tilde{\Gamma}_h}$$

Using (2.35) with $p = 2$ and $s = \sigma = 1$ ($|\cdot|_1 = \|\nabla \cdot\|_0$) for bounding the first norm on the right-hand side, and (2.44) with (2.41) for bounding the second norm, one gets

$$\begin{aligned} \|h^{-1/2} S_h(u - I_h u)\|_{0,\tilde{\Gamma}_h} &\leq C \|h^{-1}(u - I_h u)\|_{0,\tilde{\Omega}_h^b} + |u - I_h u|_{s,p,\tilde{\Omega}_h^b} \\ &\quad + h_{\Gamma}^{\zeta} \|\nabla(u - I_h u)\|_{0,\tilde{\Omega}_h^b} + h_{\Gamma}^{s_p+\zeta} |\nabla(u - I_h u)|_{s,p,\tilde{\Omega}_h^b} \\ &\leq c \left(\|h^{-1}(u - I_h u)\|_{0,\tilde{\Omega}_h^b} + (1 + h_{\Gamma}^{\zeta}) \|\nabla(u - I_h u)\|_{0,\tilde{\Omega}_h^b} \right. \\ &\quad \left. + h_{\Gamma}^{s_p+\zeta} |\nabla(u - I_h u)|_{s,p,\tilde{\Omega}_h^b} \right). \end{aligned}$$

We conclude by (2.49) bounding the first two norms on the right-hand side, after observing that $s_p \geq s$ and $|\nabla(u - I_h u)|_{s,p,\tilde{\Omega}_h^b} = |\nabla u|_{s,p,\tilde{\Omega}_h^b}$ $\square$

Next, we estimate the consistency error $E_{c,h}(u)$. This will be accomplished in two steps.

Lemma 1. *There exists a constant $C_{c,1} > 0$ independent of u and the meshsize such that*

$$E_{c,h}(u) \leq C_{c,1} \|h^{-1/2} R_h u\|_{0,\tilde{\Gamma}_h}. \quad (2.50)$$

Proof. By integration by parts and application of (2.30), one gets for all $v_h \in V_h$

$$\begin{aligned} |a_h(u, v_h) - F_h(v_h)| &\leq |(S_h u - \bar{g}, \partial_{\mathbf{n}} v_h)_{0,\tilde{\Gamma}_h}| + \gamma |(h^{-1}(S_h u - \bar{g}), S_h v_h)_{0,\tilde{\Gamma}_h}| \\ &\leq \|h^{-1/2} \underbrace{(S_h u - \bar{g})}_{R_h u}\|_{0,\tilde{\Gamma}_h} \underbrace{\left(C_I^{1/2} \|\nabla v_h\|_{0,\tilde{\Omega}_h} + \gamma \|h^{-1/2} S_h v_h\|_{0,\tilde{\Gamma}_h} \right)}_{C_{c,1}}, \end{aligned}$$

whence the result. $\square$

The analysis of the Taylor remainder carried out in Sect. 2.3 provides an estimate for the scaled norm of $R_h u$ on $\tilde{\Gamma}_h$. We suppose that the assumptions made in Proposition 4 for all the smooth parts of Γ hold true. Then we obtain the following result.

Lemma 2. *Setting*

$$\sigma := 3/2 - 1/p + \zeta' = 3/2 - 1/p + (2 - 1/p)\zeta \quad (2.51)$$

there exists a constant $C_{c,2} > 0$ independent of u and the meshsize such that

$$\|h^{-1/2}R_h u\|_{0,\tilde{\Gamma}_h} \leq C_{c,2} h_{\Gamma}^{\sigma} K(u, f, g) , \quad (2.52)$$

where $K(u, f, g)$ is a sum of norms of u and the data f and g .

Concatenating (2.50) and (2.52), we arrive at the following estimate of the consistency error.

Proposition 6. *Under the previous assumptions, there exists a constant $C_c > 0$ independent of u and the meshsize such that*

$$E_{c,h}(u) \leq C_c h_{\Gamma}^{\sigma} K(u, f, g). \quad (2.53)$$

Summarizing, we obtain an estimate of the discretization error $u - u_h$ in the ‘energy’ norm by using (2.48) and (2.53) in (2.45), together with the inequality $h_{\Gamma} \leq h_{\Omega}$.

Theorem 1. *Under the assumptions made in this section, setting*

$$r := \min(s, \sigma) , \quad (2.54)$$

there exists a constant $C > 0$ independent of u and the meshsize such that

$$\|u - u_h\|_a \leq C h_{\Omega}^r (|\nabla u|_{s,p,\Omega} + K(u, f, g)) . \quad (2.55)$$

Remark 6. This estimate is optimal. Indeed, if data are sufficiently smooth and there is no boundary singularity, then $s = 1$ and $\sigma \geq 1$ (since $p \geq 2$), whence $r = 1$. On the other hand, if a boundary singularity exists, the rate is the largest possible as detailed in [1].

Finally, observing that $|h^{s_p} \nabla(u - u_h)|_{s,p,\tilde{\Omega}_h^b} = |h^{s_p} \nabla u|_{s,p,\tilde{\Omega}_h^b} \leq h_{\Omega}^s |\nabla u|_{s,p,\Omega}$, we immediately obtain a control of the error in the norm of $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$, too.

Corollary 1. *There exists a constant $C > 0$ independent of u and the meshsize such that*

$$\|u - u_h\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)} \leq C h_{\Omega}^r (|\nabla u|_{s,p,\Omega} + K(u, f, g)) . \quad (2.56)$$

2.4.1 Enhanced error estimate in L^2

In this subsection, we provide an estimate of the L^2 -norm of the discretization error $u - u_h$, by adapting to the present setting the classical Aubin-Nitsche duality argument. To this end, we need two technical results. The first one is a measure of the non-symmetry of the bilinear form a_h .

Lemma 3. *It holds $\forall v, w \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$*

$$a_h(w, v) - a_h(v, w) = (\partial_{\tilde{\mathbf{n}}} w, \nabla v \cdot \mathbf{d})_{0, \tilde{\Gamma}_h} - (\partial_{\tilde{\mathbf{n}}} v, \nabla w \cdot \mathbf{d})_{0, \tilde{\Gamma}_h} \quad (2.57)$$

Proof. By (2.17) one has

$$\begin{aligned} a_h(w, v) &= (\nabla w, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} w, S_h v)_{0, \tilde{\Gamma}_h} + (\partial_{\tilde{\mathbf{n}}} w, \nabla v \cdot \mathbf{d})_{0, \tilde{\Gamma}_h} \\ &\quad - (S_h w, \partial_{\tilde{\mathbf{n}}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} S_h w, S_h v)_{0, \tilde{\Gamma}_h}. \end{aligned}$$

Subtracting from this expression the analogous one in which w and v are exchanged, we obtain the result. $\square$

The second technical result replaces the Galerkin orthogonality property of standard interior discretizations.

Lemma 4. *It holds*

$$a_h(u - u_h, v_h) = (R_h u, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} - \gamma(h^{-1} R_h u, S_h v_h)_{0, \tilde{\Gamma}_h}, \quad \forall v_h \in V_h. \quad (2.58)$$

Proof. Consider (2.6) and replace therein the expression of $u - \tilde{g}$ given by (2.13), obtaining

$$\begin{aligned} &(\nabla u, \nabla v_h)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{\mathbf{n}}} u, v_h)_{0, \tilde{\Gamma}_h} - (S_h u + R_h u, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} \\ &\quad + \gamma(h^{-1} (S_h u + R_h u), S_h v_h)_{0, \tilde{\Gamma}_h} \\ &= (f, v_h)_{0, \tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{\mathbf{n}}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \bar{g}, S_h v_h)_{0, \tilde{\Gamma}_h}, \quad \forall v_h \in V_h. \end{aligned}$$

Subtracting from this identity the definition (2.16) of SBM solution, we obtain the result. $\square$

Note that these two results are strictly related to the non-conformity of the SBM method. In the Shifted Boundary Method, the solution u is mapped from the surrogate boundary to the true boundary through a Taylor expansion. The measure of the non-orthogonality property is therefore directly connected to the approximation error introduced by the truncation of the Taylor expansion, i.e., to its remainder term. Furthermore, the measure of the non-symmetry naturally arises from the only non-symmetric term present in the weak formulation.

We wish to estimate

$$\|u - u_h\|_{0,\tilde{\Omega}_h} = \sup_{\tilde{z} \in L^2(\tilde{\Omega}_h)} \frac{(u - u_h, \tilde{z})_{0,\tilde{\Omega}_h}}{\|\tilde{z}\|_{0,\tilde{\Omega}_h}}. \quad (2.59)$$

Given $\tilde{z} \in L^2(\tilde{\Omega}_h)$, let $z \in L^2(\Omega)$ be its extension by 0 outside $\tilde{\Omega}_h$, and let w be the solution of the auxiliary boundary-value problem

$$\begin{aligned} -\Delta w &= z \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \Gamma. \end{aligned} \quad (2.60)$$

Let $s_0 = 1, p = 2$ be such that $w \in W^{1+s_0,p}(\Omega) = H^2(\Omega)$, it holds $\|w\|_{1,\Omega} + |\nabla w|_{1,2,\Omega} = \|w\|_{2,\Omega} \leq c\|z\|_{0,\Omega} = c\|\tilde{z}\|_{0,\tilde{\Omega}_h}$ for some $c > 0$ independent of $\tilde{z}$. Thus, $w \in W(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) := V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h, 1, 2)$ as defined in (2.38).

The same argument that led to (2.12) shows that the trace $\tilde{q}$ of w on $\tilde{\Gamma}_h$ satisfies

$$\tilde{q} = 0 - \nabla w \cdot \mathbf{d} - R_h(w, \mathbf{d}), \quad (2.61)$$

whence on $\tilde{\Gamma}_h$ it holds

$$0 = w - \tilde{q} = S_h w + R_h w. \quad (2.62)$$

Thus, for any $v \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ we have

$$\begin{aligned} (\tilde{z}, v)_{0,\tilde{\Omega}_h} &= -(\Delta w, v)_{0,\tilde{\Omega}_h} = (\nabla w, \nabla v)_{0,\tilde{\Omega}_h} - (\partial_{\mathbf{n}} w, v)_{0,\tilde{\Gamma}_h} \\ &= (\nabla w, \nabla v)_{0,\tilde{\Omega}_h} - (\partial_{\mathbf{n}} w, v)_{0,\tilde{\Gamma}_h} - (S_h w + R_h w, \partial_{\mathbf{n}} v)_{0,\tilde{\Gamma}_h} \\ &\quad + \gamma(h^{-1}(S_h w + R_h w), S_h v)_{0,\tilde{\Gamma}_h} \\ &= a_h(w, v) - (R_h w, \partial_{\mathbf{n}} v)_{0,\tilde{\Gamma}_h} + \gamma(h^{-1} R_h w, S_h v)_{0,\tilde{\Gamma}_h} \\ &= a_h(v, w) + (\partial_{\mathbf{n}} w, \nabla v \cdot \mathbf{d})_{0,\tilde{\Gamma}_h} - (\partial_{\mathbf{n}} v, \nabla w \cdot \mathbf{d})_{0,\tilde{\Gamma}_h} \\ &\quad - (R_h w, \partial_{\mathbf{n}} v)_{0,\tilde{\Gamma}_h} + \gamma(h^{-1} R_h w, S_h v)_{0,\tilde{\Gamma}_h}, \end{aligned}$$

where in the last step we have used Lemma 3. We now pick $v = \psi := u - u_h$ and we use Lemma 4 with $v_h = w_I := I_h w$ to get

$$\begin{aligned} (\tilde{z}, \psi)_{0,\tilde{\Omega}_h} &= a_h(\psi, w) + E_1(\psi, w) + E_2(\psi, w) \\ &= a_h(\psi, w - w_I) + E_1(\psi, w) + E_2(\psi, w) + E_3(u, w_I), \end{aligned} \quad (2.63)$$

with the asymmetry error

$$E_1(\psi, w) := (\partial_{\mathbf{n}} w, \nabla \psi \cdot \mathbf{d})_{0,\tilde{\Gamma}_h} - (\partial_{\mathbf{n}} \psi, \nabla w \cdot \mathbf{d})_{0,\tilde{\Gamma}_h}, \quad (2.64)$$

the shifting error

$$E_2(\psi, w) := -(R_h w, \partial_{\mathbf{n}} \psi)_{0,\tilde{\Gamma}_h} + \gamma(h^{-1} R_h w, S_h \psi)_{0,\tilde{\Gamma}_h}, \quad (2.65)$$

the orthogonality error

$$E_3(u, w_I) := (R_h u, \partial_{\tilde{n}} w_I)_{0, \tilde{\Gamma}_h} - \gamma(h^{-1} R_h u, S_h w_I)_{0, \tilde{\Gamma}_h}. \quad (2.66)$$

We proceed to bound the four error terms on the right-hand side of (2.63). Concerning the first one, we observe that an argument similar to the one used in the proof of Proposition 5 yields the bound

$$\|w - w_I\|_{W(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)} \leq C h_{\Omega}^{s_0} |\nabla w|_{s_0, 2, \Omega}. \quad (2.67)$$

On the other hand, proceeding as in the proof of Proposition 3, one gets

$$|a_h(\psi, w - w_I)| \leq \|a_h\|_{V, W} \|\psi\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)} \|w - w_I\|_{W(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}$$

for some constant $\|a_h\|_{V, W} > 0$ independent of the meshsize. Hence, by (2.56) we deduce the existence of a constant $C_{0,0} > 0$ independent of $u, \tilde{z}$ and the meshsize, such that

$$|a_h(\psi, w - w_I)| \leq C_{0,0} h_{\Omega}^{r+s_0} (|\nabla u|_{s,p,\Omega} + K(u, f, g)) \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \quad (2.68)$$

Concerning the error term E_1 in (2.63), recalling Assumption 3 we have

$$|E_1(\psi, w)| \leq c_d h_{\Gamma}^{1/2+\zeta} \|\nabla w\|_{0, \tilde{\Gamma}_h} \|h^{1/2} \nabla \psi\|_{0, \tilde{\Gamma}_h}$$

with $\|\nabla w\|_{0, \tilde{\Gamma}_h} \leq c(\|w\|_{1, \Omega} + |\nabla w|_{s_0, \Omega})$ by the trace theorem, and $\|h^{1/2} \nabla \psi\|_{0, \tilde{\Gamma}_h} \leq \bar{C}_I \|\psi\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}$ by (2.41). Hence, we deduce the existence of a constant $C_{0,1} > 0$ independent of $u, \tilde{z}$ and the meshsize, such that

$$|E_1(\psi, w)| \leq C_{0,1} h_{\Omega}^{r+1/2+\zeta} (|\nabla u|_{s,p,\Omega} + K(u, f, g)) \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \quad (2.69)$$

Next, considering the error term E_2 in (2.63), we have

$$\begin{aligned} |E_2(\psi, w)| &\leq \|h^{-1/2} R_h w\|_{0, \tilde{\Gamma}_h} (\|h^{1/2} \nabla \psi\|_{0, \tilde{\Gamma}_h} + \|h^{-1/2} S_h \psi\|_{0, \tilde{\Gamma}_h}) \\ &\leq c \|h^{-1/2} R_h w\|_{0, \tilde{\Gamma}_h} \|\psi\|_{V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)}. \end{aligned}$$

To estimate the norm involving $R_h w$, we can apply the equivalent statement of Lemma 2 for w , with the summation index $p = 2$. Thus, denoting by σ_0 the constant defined in (2.51) with $p = 2$, we get

$$\|h^{-1/2} R_h w\|_{0, \tilde{\Gamma}_h} \leq c h_{\Gamma}^{\sigma_0} (|w|_{s,p,\Omega} + \|\tilde{z}\|_{0, \tilde{\Omega}_h}) \leq c' h_{\Gamma}^{\sigma_0} \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \quad (2.70)$$

Hence, we deduce the existence of a constant $C_{0,2} > 0$ independent of $u, \tilde{z}$ and the meshsize, such that

$$|E_2(\psi, w)| \leq C_{0,2} h_{\Omega}^{r+\sigma_0} (|\nabla u|_{s,p,\Omega} + K(u, f, g)) \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \quad (2.71)$$

Finally, the error term E_3 in (2.63) can be equivalently written as

$$E_3(u, w_I) = (R_h u, \partial_{\tilde{n}} w)_{0, \tilde{\Gamma}_h} - \gamma(h^{-1} R_h u, S_h w)_{0, \tilde{\Gamma}_h} - (R_h u, \partial_{\tilde{n}}(w - w_I))_{0, \tilde{\Gamma}_h} \\ + \gamma(h^{-1} R_h u, S_h(w - w_I))_{0, \tilde{\Gamma}_h}.$$

Recalling (2.62), one has

$$\begin{aligned} |E_3(u, w_I)| &\leq \|h^{-1/2} R_h u\|_{0, \tilde{\Gamma}_h} (h_{\Gamma}^{1/2} \|\nabla w\|_{0, \tilde{\Gamma}_h} + \gamma \|h^{-1/2} R_h w\|_{0, \tilde{\Gamma}_h} \\ &\quad + \underbrace{\|h^{1/2} \nabla(w - w_I)\|_{0, \tilde{\Gamma}_h}}_{\leq C_I^{1/2} \|\nabla(w - w_I)\|_{0, \tilde{\Omega}_h}} + \gamma \|h^{-1/2} S_h(w - w_I)\|_{0, \tilde{\Gamma}_h}) \\ &\leq \|h^{-1/2} R_h u\|_{0, \tilde{\Gamma}_h} (h_{\Gamma}^{1/2} \|\nabla w\|_{0, \tilde{\Gamma}_h} + \gamma \|h^{-1/2} R_h w\|_{0, \tilde{\Gamma}_h} \\ &\quad + \underbrace{c \|w - w_I\|_{W(\tilde{\Omega}_h; \tilde{\Gamma}_h)}}_{\leq c C h_{\Omega}^{s_0} |\nabla w|_{s_0, 2, \Omega}}) \\ &\leq C_{c,2} h_{\Gamma}^{\sigma} K(u, f, g) (c h_{\Gamma}^{1/2} + c' h_{\Gamma}^{\sigma_0} + c'' h_{\Omega}^{s_0}) \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \end{aligned}$$

We have used (2.52), (2.70) and (2.67). Observing that $\min(1/2, \sigma_0, s_0) = 1/2$, we deduce the existence of a constant $C_{0,3} > 0$ independent of $u, \tilde{z}$ and the meshsize, such that

$$|E_3(u, w_I)| \leq C_{0,3} h_{\Omega}^{\sigma+1/2} K(u, f, g) \|\tilde{z}\|_{0, \tilde{\Omega}_h}. \quad (2.72)$$

We are ready to state our estimate on the L^2 -norm of the discretization error $u - u_h$. Using the Riesz representation (2.59) together with (2.63) and the bounds (2.68), (2.69), (2.71), and (2.72), we arrive at the following result.

Theorem 2. *Under the same assumptions for the validity of Theorem 1, setting*

$$r_0 := \min(s_0, 1/2 + \zeta, \sigma_0, \sigma - r + 1/2), \quad (2.73)$$

there exists a constant $C > 0$ independent of u and the meshsize such that

$$\|u - u_h\|_{0, \tilde{\Omega}_h} \leq C h_{\Omega}^{r+r_0} (|\nabla u|_{s,p,\Omega} + K(u, f, g)). \quad (2.74)$$

Remark 7. It is worth analyzing the predicted decay rate in different situations, making the exponent $r + r_0$ explicit. In the most favorable case (smooth data and smooth boundary), one has $r = s_0 = 1$, and $\sigma = 3/2 + 2\zeta$, $\sigma_0 = 1 + 3\zeta/2$, whence $r + r_0 = 3/2 + \zeta$. Note that, the exponent $r + r_0$ is optimal ($= 2$) provided $\zeta \geq 1/2$. On the other hand, in the presence of boundary singularities the constraint is the same (see [1]).

Remark 8. Allowing $\zeta > 0$ to be arbitrarily small is of paramount importance in complex engineering applications, as it essentially avoids restrictions in the mesh generation near the boundary. In this case, the exponent in (2.74) is sub-optimal. Inspecting the behavior of the four error terms appearing in (2.63) reveals that in the most favorable situation (smooth data and smooth boundary), the terms $a_h(\psi, w - w_I)$, $E_2(\psi, w)$, and $E_3(u, w_I)$ exhibit the optimal decay h_Ω^2 (see (2.68), (2.71), and (2.72) with $p = 2$). Thus, the responsible for sub-optimality is the term $E_1(\psi, w)$, which reflects the discrepancy between the normal $\mathbf{n}$ to the exact boundary Γ and the normal $\tilde{\mathbf{n}}$ to the surrogate boundary $\tilde{\Gamma}_h$. To shed further light, assume that Γ is smooth and the mapping (2.9) is the closest-point projection, which implies $\mathbf{d} = \|\mathbf{d}\|\mathbf{n}$. Using the splitting $\partial_{\tilde{\mathbf{n}}} = (\tilde{\mathbf{n}} \cdot \mathbf{n})\partial_{\mathbf{n}} + (\tilde{\mathbf{n}} \cdot \mathbf{t})\partial_{\mathbf{t}}$, after cancellation of the common term involving the product $\partial_{\mathbf{n}}w \partial_{\mathbf{n}}\psi$, one gets

$$E_1(\psi, w) = \int_{\tilde{\Gamma}_h} \|\mathbf{d}\|(\tilde{\mathbf{n}} \cdot \mathbf{t})(\partial_{\mathbf{n}}\psi \partial_{\mathbf{t}}w - \partial_{\mathbf{n}}w \partial_{\mathbf{t}}\psi).$$

For sufficiently refined meshes, when $\tilde{\Gamma}_h$ tends to become parallel to Γ , the factor $\tilde{\mathbf{n}} \cdot \mathbf{t}$ may become small, and cancellations due to sign changes from one edge to the other along $\tilde{\Gamma}_h$ may occur. Indeed, the optimal rate h_Ω^2 for $\|u - u_h\|_{0, \tilde{\Omega}_h}$ is often observed in practical calculations with SBM.

Analogous considerations apply in the presence of geometric singularities.

2.5 Numerical Results

This section presents numerical experiments for the Poisson equation solved using the SBM. We consider four test cases: the unit square $[0,1]^2$, the circle centered at $(0,0)$ with radius $R = 1$, the ellipse centered at $(0,0)$ with semi-axes $a = 1$ and $b = 1/2$, and the quarter of a circle centered at $(0,0)$ with radius $R = 1$.

For all test cases, the exact solution u is known, allowing us to evaluate the L^2 -norm and the H^1 -seminorm (or H_0^1 -norm) of the error.

Since the mesh generator employed requires as input the maximum area $|E|$ of the elements E_k , for $k = 1, \dots, N_{\text{elements}}$, rather than the mesh size h , we introduce the quantity h_{input} defined as

$$h_{\text{input}} := \sqrt{|E|}. \quad (2.75)$$

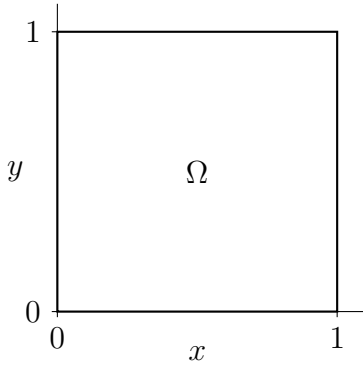
Note that h_{input} does not necessarily coincide with the effective maximum mesh size h_Ω (see (2.4)). Throughout the remainder of this work, whenever we write h , we will always refer to the effective maximum mesh size h_Ω in the surrogate domain.

2.5.1 Square Domain

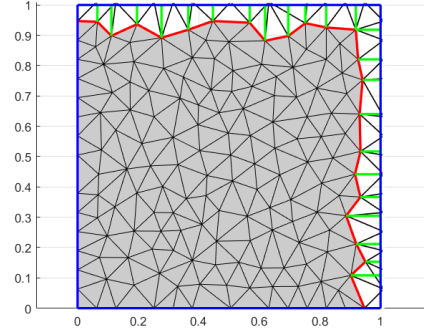
Let Ω be the square domain $[0,1]^2$. The surrogate domain $\tilde{\Omega}_h$ is constructed so that the left and top edges are unfitted (see Figure 2.4). The operator $\mathbf{M}_h$ denotes the closest point projection. The exact solution is chosen as

$$u = x + y + 16xy(1-x)(1-y).$$

The forcing term f is computed accordingly, and g is the trace of u on Γ , that is $g = x + y$. The Nitsche stabilization parameter γ is set to 10, following the choice in [1].



(a) Square domain $\Omega = [0,1]^2$.



(b) Surrogate domain $\tilde{\Omega}_h$ (grey) with corresponding mesh, true boundary Γ (blue) and surrogate boundary $\tilde{\Gamma}_h$ (red).

Figure 2.4: Square real and surrogate domains for $h = 0.137$.

In Figure 2.5 we compare the exact and the numerical solutions. Note that the numerical solution u_h is defined only on $\tilde{\Omega}_h$. Therefore, if one wishes to visualize u_h on the true domain Ω , a smooth extension of u_h is required on the elements belonging to $\Omega \setminus \tilde{\Omega}_h$. In general, the numerical solution slightly underestimates the exact one.

The error values $\|u - u_h\|_{L^2(\tilde{\Omega}_h)}$ and $|u - u_h|_{H^1(\tilde{\Omega}_h)}$ are reported in Table 2.1. The observed convergence rates agree with the theoretical prediction for the energy norm. As anticipated in Remark 8, the L^2 -error also exhibits the optimal convergence rate, showing the non-optimality of the estimate in (2.74). Finally, Figure 2.6 shows the logarithmic linear regressions of the errors, confirming the computed convergence rates.

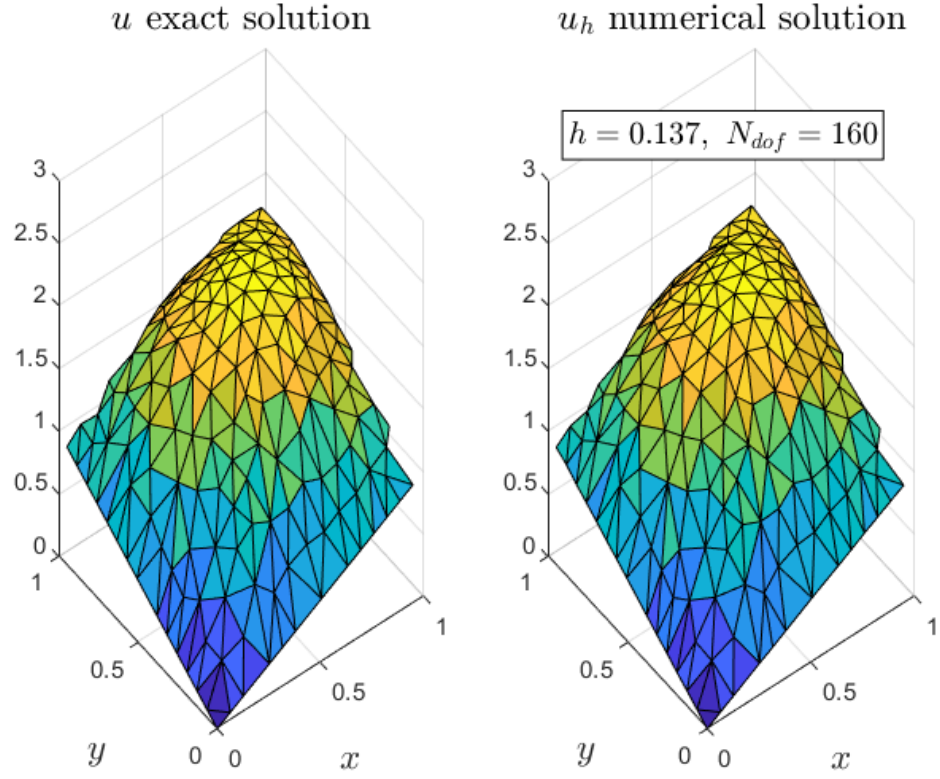
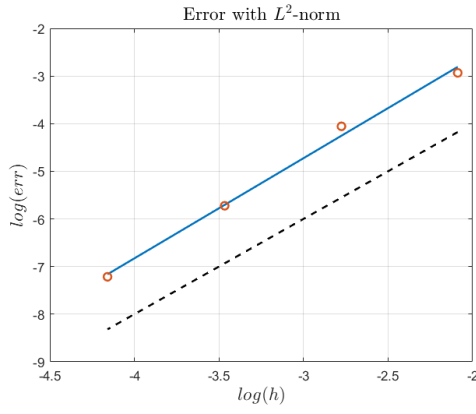


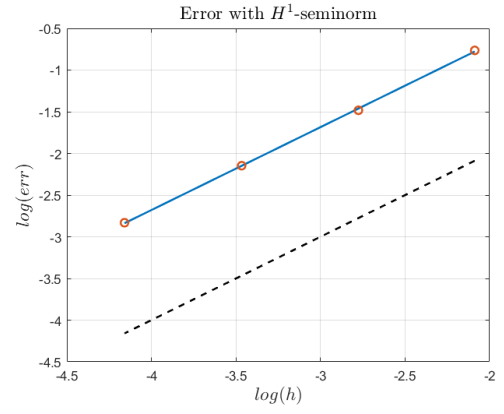
Figure 2.5: [Square] Comparison between the exact solution u and the numerical solution u_h for $h = 0.137$.

h_{input}	h	L^2 error	Rate (2.1025)	H_0^1 error	Rate (0.9946)
1/8	0.222	5.32×10^{-2}	—	4.657×10^{-1}	—
1/16	0.125	1.73×10^{-2}	1.96	2.271×10^{-1}	1.25
1/32	0.0663	3.3×10^{-3}	2.61	1.168×10^{-1}	1.05
1/64	0.0360	7×10^{-4}	2.54	5.89×10^{-2}	1.12

Table 2.1: [Square] Errors in L^2 and H_0^1 norms with corresponding convergence rates.



(a) L^2 logarithmic error regression
(Rate 2.1025).



(b) H_0^1 logarithmic error regression
(Rate 0.9946).

Figure 2.6: [Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.

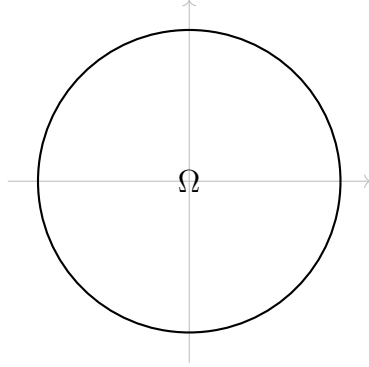
2.5.2 Circular domain

Let Ω be the disk of radius $R = 1$ centered at the origin. As shown in Figure 2.7 (b), the surrogate domain $\tilde{\Omega}_h$ is generated by selecting only the triangles that are strictly contained in the circular domain Ω . In this case, the projection operator is $\mathbf{M}_h^z$, introduced in Section 2.1.1.

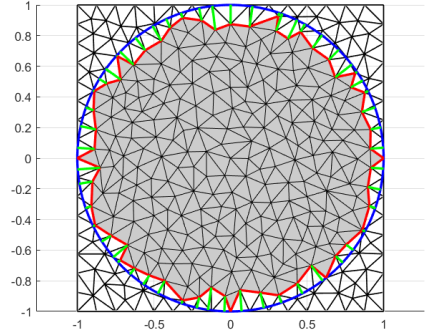
The exact solution is chosen as

$$u = 1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}, \quad (2.76)$$

which reduces to a paraboloid of rotation when $a = b = 1$. Homogeneous Dirichlet boundary conditions $u = 0$ on $\Gamma = \partial\Omega$ are imposed. The Nitsche penalization parameter γ is set to 60.



(a) Circular domain
 $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$.



(b) Surrogate domain $\tilde{\Omega}_h$ (grey) with corresponding mesh, true boundary Γ (blue) and surrogate boundary $\tilde{\Gamma}_h$ (red).

Figure 2.7: Circular real and surrogate domains for $h = 0.214$.

In Figure 2.8, the numerical and exact solutions are compared. The numerical solution satisfies the maximum principle, taking values that are always less than or equal to the exact solution. The convergence rates are optimal in both the L^2 and H_0^1 norms. The corresponding error values are reported in Table 2.2. Note that the initial meshsize is twice as large as in the square domain case. This is due to the fact that the external area of the circular domain is four times larger, since $|\mathcal{D}| \sim h^2$. In Figure 2.9, the logarithmic linear regressions of the errors are shown, confirming the expected convergence rates.

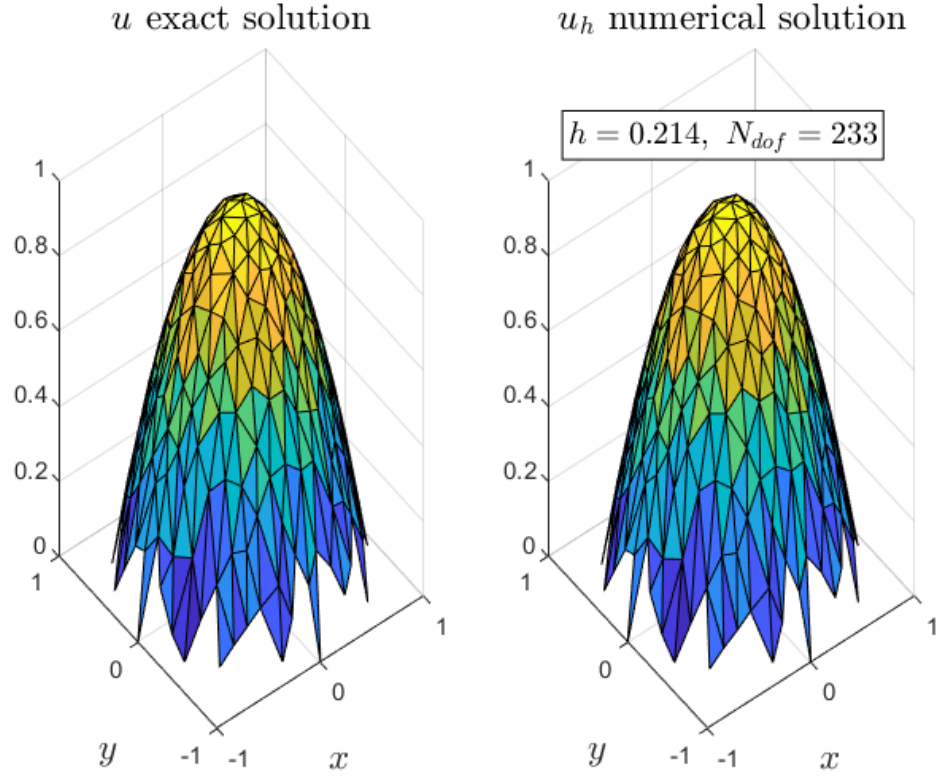
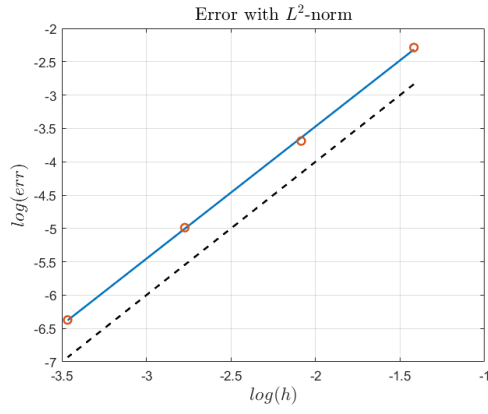


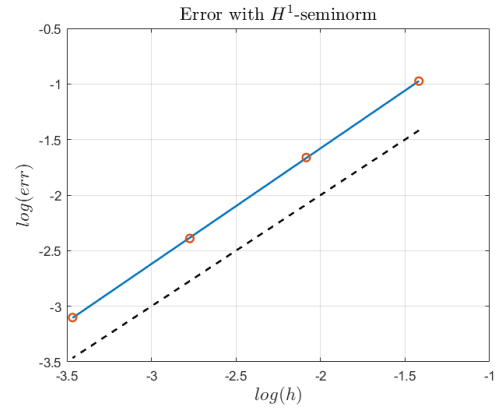
Figure 2.8: [Circle] Comparison between the exact solution u and the numerical solution u_h for $h = 0.214$.

h_{input}	h	L^2 error	Rate (1.9834)	H_0^1 error	Rate (1.0405)
1/4	0.4370	1.015×10^{-1}	—	3.774×10^{-1}	—
1/8	0.2329	2.50×10^{-2}	2.23	1.897×10^{-1}	1.09
1/16	0.1394	6.8×10^{-3}	2.54	9.16×10^{-2}	1.42
1/32	0.0705	1.7×10^{-3}	2.03	4.50×10^{-2}	1.04

Table 2.2: [Circle] Errors in L^2 and H_0^1 norms with corresponding convergence rates.



(a) L^2 logarithmic error regression
(Rate 1.9834).

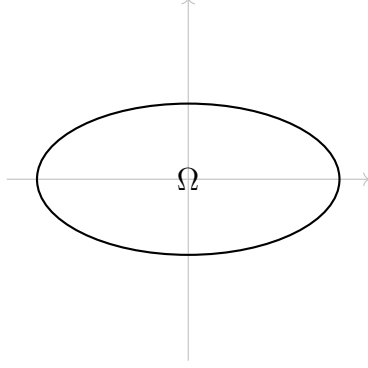


(b) H_0^1 logarithmic error regression
(Rate 1.0405).

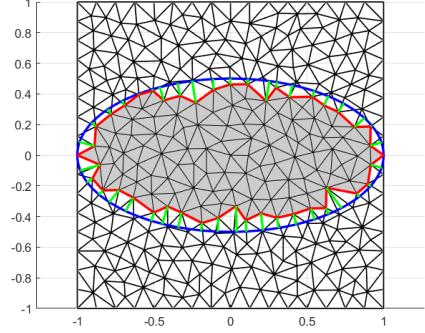
Figure 2.9: [Circle] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.

2.5.3 Elliptic domain

This section closely follows the previous one, as it represents a straightforward extension of the circular domain to the case where $a \neq b$. Let Ω be the ellipse centered at $(0,0)$ with semiaxes $a = 1$ and $b = 1/2$. The surrogate domain is generated in the same manner as for the circular case. Here as well, the operator $\mathbf{M}_h^z$ is employed. The exact solution is again given by (2.76), but now with the appropriate values of a and b . Homogeneous boundary conditions are imposed on $\Gamma = \partial\Omega$, and the Nitsche penalty parameter is set to $\gamma = 60$.



(a) Elliptic domain
 $\Omega = \{(x, y) : x^2 + 4y^2 \leq 1\}$.



(b) Surrogate domain $\tilde{\Omega}_h$ (grey) with corresponding mesh, true boundary Γ (blue) and surrogate boundary $\tilde{\Gamma}_h$ (red).

Figure 2.10: Elliptic real and surrogate domains for $h = 0.201$.

The qualitative interpretation of the results is analogous to the one discussed in Section 2.5.2. The numerical solution satisfies the maximum principle, as shown in Figure 2.11. The convergence rates obtained from the logarithmic regressions in Figure 2.12 are optimal. From Table 2.3, it can be observed that the error values are larger than those obtained in the circular test for both norms. Compared with the circular case, the L^2 convergence rate is slightly higher, whereas the H^1 convergence rate is slightly lower.

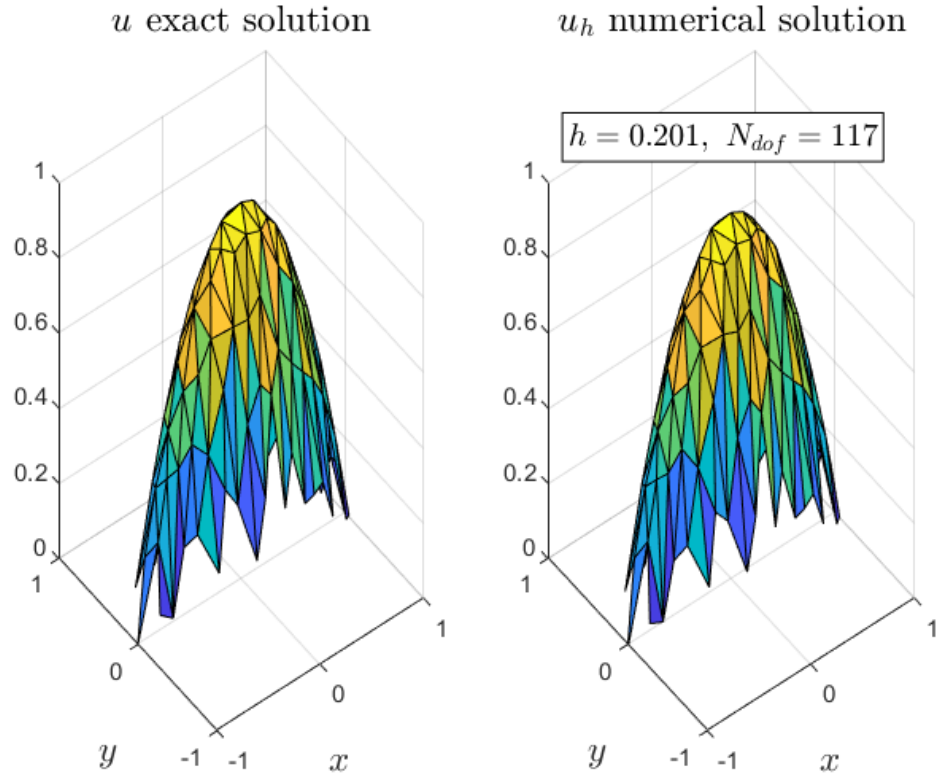
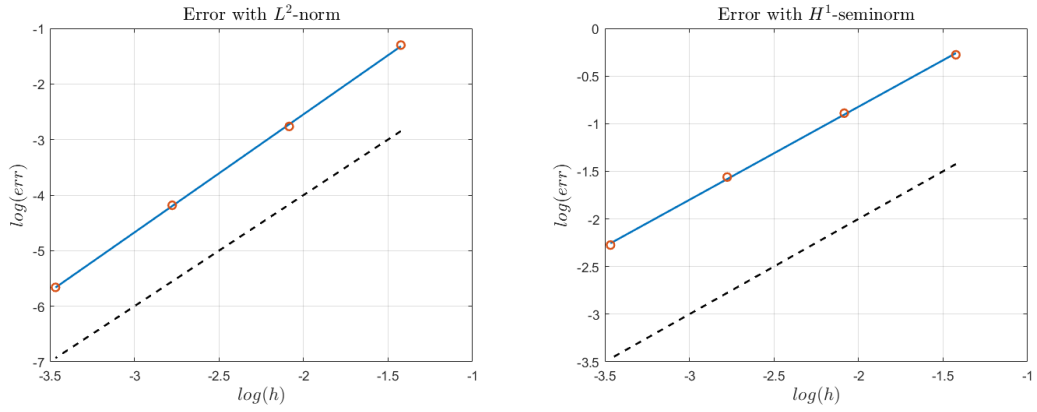


Figure 2.11: [Ellipse] Comparison between the exact solution u and the numerical solution u_h for $h = 0.201$.

h_{input}	h	L^2 error	Rate (2.1257)	H_0^1 error	Rate (0.9772)
1/4	0.4370	2.724×10^{-1}	—	7.580×10^{-1}	—
1/8	0.2329	6.30×10^{-2}	2.33	4.106×10^{-1}	0.97
1/16	0.1394	1.53×10^{-2}	2.76	2.100×10^{-1}	1.31
1/32	0.0705	3.5×10^{-3}	2.16	1.029×10^{-1}	1.05

Table 2.3: [Ellipse] Errors in L^2 and H_0^1 norms with corresponding convergence rates.



(a) L^2 logarithmic error regression
(Rate 2.1257).

(b) H_0^1 logarithmic error regression
(Rate 0.9772).

Figure 2.12: [Ellipse] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.

2.5.4 Quarter-Circle domain

In the previous sections, we performed numerical tests for simple domains. In this section, we consider a different setup. Let Ω be the domain shown in Figure 2.13 (a). Let u be a solution of the homogeneous Poisson equation (2.1) with $g = 0$, exhibiting a boundary layer near the curved portion of $\partial\Omega$. In polar coordinates, the exact solution can be expressed as

$$u(r, \theta) = \varphi(r) \psi(\theta), \quad 0 \leq r \leq 1, \quad 0 \leq \theta \leq \frac{\pi}{2}, \quad (2.77)$$

where

- $\psi(\theta) = \sin(2k\theta)$ for some integer $k \geq 1$, such that $\psi(0) = \psi(\frac{\pi}{2}) = 0$,
- $\varphi(r) = r^\alpha \eta(r)$ with $\alpha > 0$ and η solution of the one-dimensional convection-diffusion equation

$$\begin{cases} -\frac{1}{\text{Pe}} \eta'' + \eta' = 0 & \text{in } (0,1) \\ \eta(0) = 1, \eta(1) = 0 \end{cases} \quad (2.78)$$

with Péclet number $\text{Pe} > 1$. One easily checks that $\eta(r) = \frac{e^{\text{Pe}} - e^{\text{Pe}r}}{e^{\text{Pe}} - 1}$ and $\varphi(0) = \varphi(1) = 0$.

The Nitsche penalization parameter γ is set to 10 or 100, depending on the values of the parameters Pe , α , and k .

A few comments on u are in order. First, the parameter k of the angular component $\psi(\theta)$ is related to the number of sign changes (or equivalently, the number of maxima and minima) of ψ in the interval $[0, \pi/2]$ (see Figure 2.14 (a)). On the other hand, for the radial component $\varphi(r)$, the Péclet number must remain moderate, since it is responsible for the well-known numerical instabilities of the convection-diffusion equation in the convection-dominated regime. Indeed, if $\text{Pe} \gg 1$ in (2.78), the diffusion term becomes negligible (see Figure 2.14 (b)). Finally, if we are interested in a solution $u \in H^2(\tilde{\Omega}_h)$, this is equivalent to requiring that $\Delta u \in L^2(\tilde{\Omega}_h)$. By direct computation, we have

$$\begin{aligned} \Delta u &= \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \\ &= r^{\alpha-2} \left((\alpha^2 - 4k^2) \eta(r) + r(1 + 2\alpha + \text{Pe}r) \eta'(r) \right) \sin(2k\theta) \\ &= r^{\alpha-2} \chi(r, \theta), \end{aligned}$$

and, by the definition of square integrable functions,

$$\int_{\tilde{\Omega}_h} |\Delta u|^2 \leq \int_{\Omega} |\Delta u|^2 = \int_0^{\pi/2} d\theta \int_0^R r^{2(\alpha-2)} \chi(r, \theta)^2 r dr < +\infty,$$

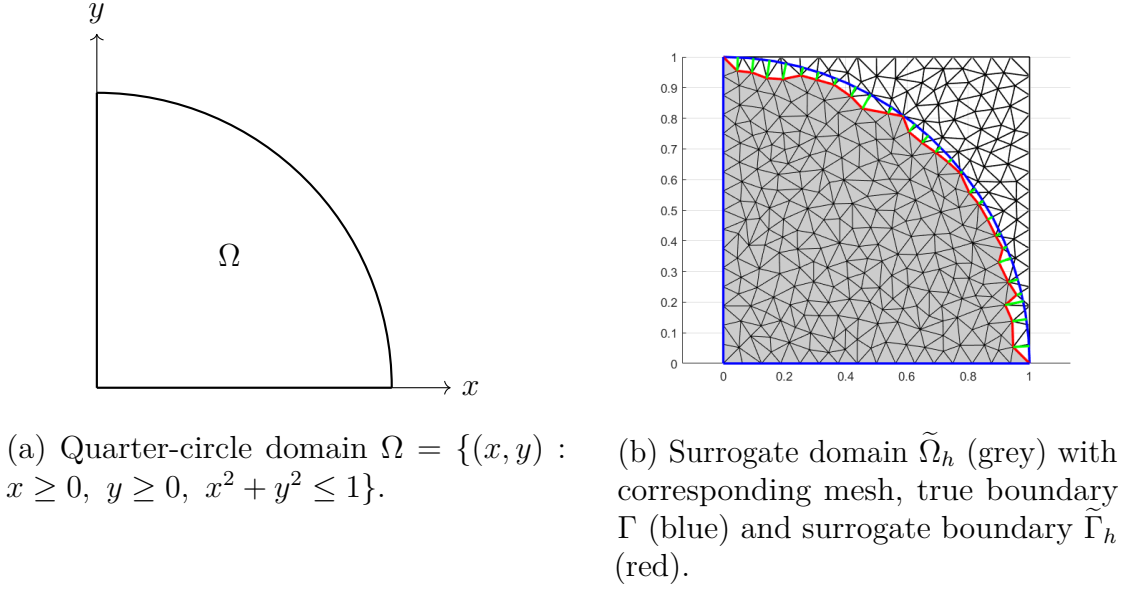


Figure 2.13: Quarter-Circle real and surrogate domains for $h = 0.107$.

which is satisfied if $2(\alpha - 2) + 1 > -1$, that is, $\alpha > 1$. The value $\alpha = 1$ is still admissible, recalling (2.38), but it implies that u has less regularity than $H^2(\tilde{\Omega}_h)$.

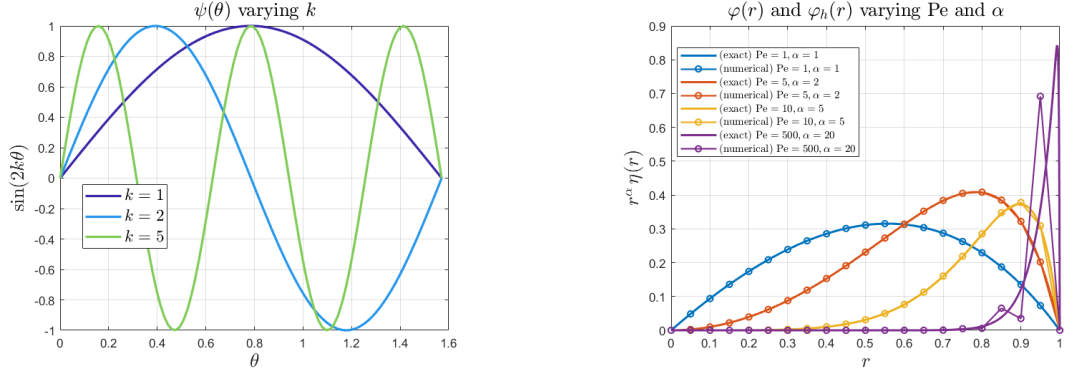
Based on this observation, we analyze the behavior of the solution as the parameters Pe , α , and k vary. In particular, we focus on the cases $\alpha = 1$, $\text{Pe} = 5$, $k = 2$ and $\alpha = 5$, $\text{Pe} = 10$, $k = 5$.

Test 1: $\alpha = 1, \text{Pe} = 5, k = 2$

The penalty parameter γ is set to 100. As shown in Figure 2.15, the numerical and exact solutions are nearly identical. The convergence rates are confirmed by the regression in Figure 2.16. However, from the values reported in Table 2.4, it appears that the error is still in a pre-asymptotic regime, particularly during the first refinement iterations. Nonetheless, the theoretical predictions are verified, and the order of magnitude of the errors is consistent with those observed in the curved domain tests.

Test 2: $\alpha = 5, \text{Pe} = 10, k = 5$

The penalty parameter γ is again set to 100. From Figure 2.17, it can be observed that the difference between the numerical and exact solutions is larger than in the previous scenarios. This observation is confirmed by the larger magnitude of the errors reported in Table 2.5 compared to the earlier tests. Nonetheless, the convergence rates remain correct, especially in the last iterations with finer meshes.


 (a) Angular component of u .

 (b) Radial component of u . $\varphi_h(r)$ is obtained with η_h numerical solution of (2.78) with $h = 0.05$. It is possible to see that the numerical solution starts to be instable with $Pe = 500$.

Figure 2.14: Angular and radial components of u .

h_{input}	h	L^2 error	Rate (2.2495)	H_0^1 error	Rate (0.9606)
1/16	0.1196	3.84×10^{-2}	—	6.027×10^{-1}	—
1/32	0.0697	1.05×10^{-2}	2.08	3.305×10^{-1}	0.89
1/64	0.0349	1.80×10^{-3}	2.54	1.566×10^{-1}	1.08
1/128	0.0170	4.00×10^{-4}	2.16	8.41×10^{-2}	0.90

Table 2.4: [Test 1 with $\alpha = 1, Pe = 5, k = 2$] Errors in L^2 and H_0^1 norms with corresponding convergence rates.

This behavior is highlighted by the logarithmic regression line shown in Figure 2.18, particularly for the L^2 error.

h_{input}	h	L^2 error	Rate (1.9049)	H_0^1 error	Rate (1.0469)
1/16	0.1196	6.69×10^{-2}	—	1.009,8	—
1/32	0.0697	3.07×10^{-2}	1.44	5.711×10^{-1}	1.06
1/64	0.0349	6.7×10^{-3}	2.20	2.473×10^{-1}	1.22
1/128	0.0170	1.4×10^{-3}	2.18	1.19×10^{-1}	1.01

Table 2.5: [Test 2 with $\alpha = 5, Pe = 10, k = 5$] Errors in L^2 and H_0^1 norms with corresponding convergence rates.

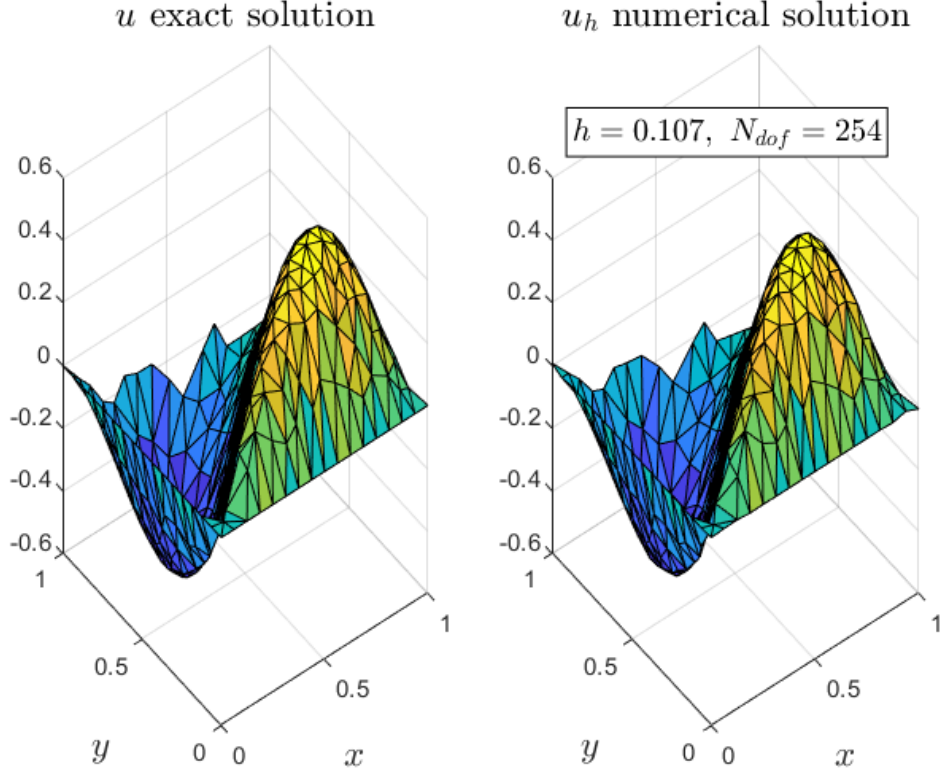
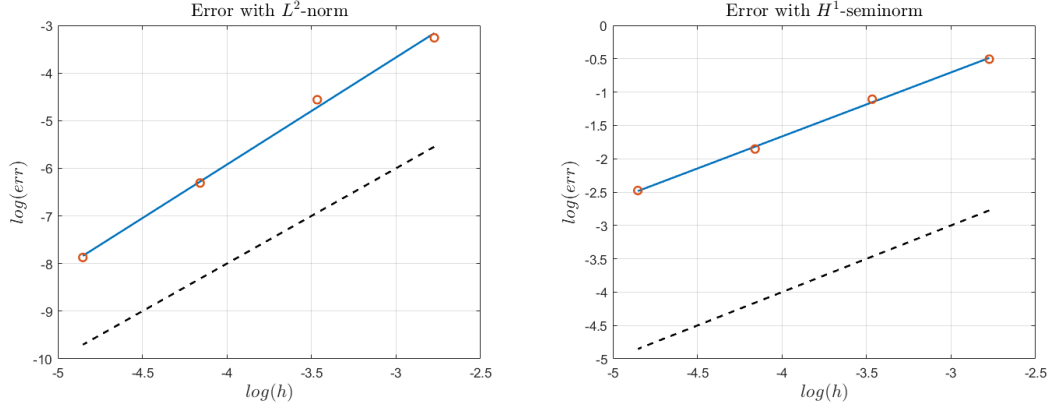


Figure 2.15: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 2$] Comparison between the exact solution u and the numerical solution u_h for $h = 0.107$.

In the end, several numerical tests are carried out by varying all the parameters, and the corresponding results are summarized in Table 2.6. As before, four mesh refinements are performed, each obtained by halving the mesh size, starting from $h_{\text{input}} = 1/16$ or $h_{\text{input}} = 1/20$.

It is worth noting that the most critical parameter is α : as α increases, the convergence rates tend to decrease. We believe this behaviour to be related to a pre-asymptotic regime, since the last error values exhibit the expected asymptotic rates. Moreover, in the case $\gamma = 0$ (see Remark 3), starting from a smaller initial mesh size improves the convergence rates. In general, choosing $\gamma = 100$ provides better stabilization. The parameter k also affects the convergence rates, although its influence is weaker compared to that of α .



(a) L^2 logarithmic error regression
(Rate 2.2495).

(b) H_0^1 logarithmic error regression
(Rate 0.9606).

Figure 2.16: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 2$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.

γ	Pe	α	k	h_{start}	L^2 rate	H^1 rate
100	10	2	1	1/16	1.9598	0.9432
10	10	2	1	1/16	1.8464	0.9079
0	10	2	1	1/16	1.6431	0.9931
0	10	2	1	1/20	1.7976	1.0145
100	10	3/2	1	1/16	1.9825	0.9571
10	10	3/2	5	1/16	1.9748	0.9761
100	10	3/2	10	1/16	1.8453	0.9201
10	10	3/2	10	1/16	1.8417	0.8989
100	10	5	1	1/16	1.8369	0.8553

Table 2.6: Summary of asymptotic convergence rates for various choices of $(\gamma, \text{Pe}, \alpha, k)$, with the corresponding starting mesh size h_{start} .

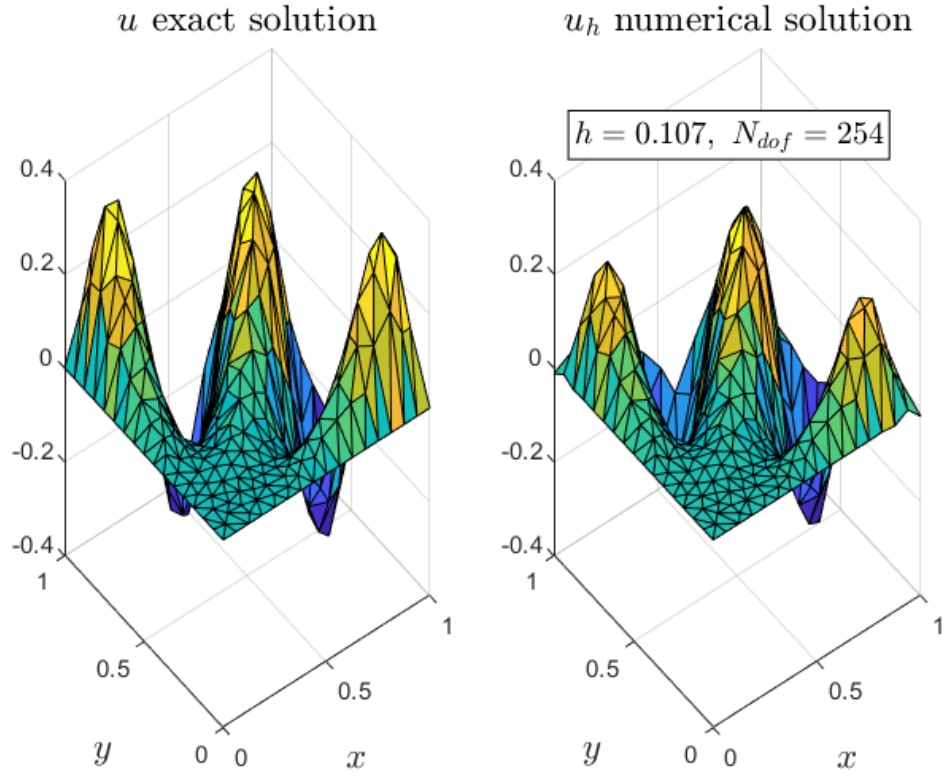
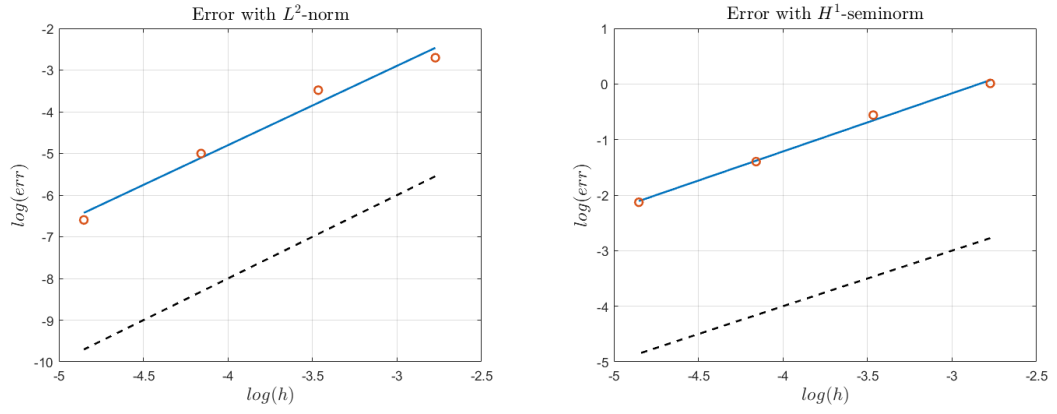


Figure 2.17: [Test 2 with $\alpha = 5, \text{Pe} = 10, k = 5$] Comparison between the exact solution u and the numerical solution u_h for $h = 0.107$.



(a) L^2 logarithmic error regression
(Rate 1.9049).

(b) H_0^1 logarithmic error regression
(Rate 1.0469).

Figure 2.18: [Test 2 with $\alpha = 5$, $Pe = 10$, $k = 5$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines.

Chapter 3

Time dependent problems

As announced in the previous chapter, we are now interested in studying the *Shifted Boundary Method* for time dependent problems. In particular, we will analyze the method applied to the heat equation. Since the space analysis is the same as for the Poisson equation, we will focus on the time analysis. The approach employed is the *lines method*, so solving this PDE implies the semi-discretization in the space variables at first and solving the remaining system in t later on, with the help of some well-known time advancing methods.

3.1 The Shifted Boundary Method for Time dependent problems

Let the assumption made in Section 2.1 be satisfied. We also consider $T > 0$ as the final time of our period of interest. We aim at numerically solving the Dirichlet boundary value problem for the heat equation

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f \quad \text{in } \Omega, \quad 0 < t \leq T, \\ u &= g \quad \text{on } \Gamma, \quad 0 < t \leq T, \\ u &= u_0 \quad \text{in } \Omega, \quad t = 0, \end{aligned} \tag{3.1}$$

where the data $f \in L^2(0, T; L^2(\Omega))$, $g \in L^2(0, T; H^{1/2}(\Gamma))$ and $u_0 \in L^2(\Omega)$ satisfy additional regularity assumptions that will be made precise later on. Recalling the definition of $\tilde{\mathcal{T}}_h$ and $\tilde{\Omega}_h$ given in (2.2) and (2.3) respectively, suppose that we knew the value of the exact solution u on $\tilde{\Gamma}_h$, say $u|_{\tilde{\Gamma}_h} = \tilde{g}$, we could consider the problem

$$\begin{aligned}
 \frac{\partial u}{\partial t} - \Delta u &= f \quad \text{in } \tilde{\Omega}_h, \quad 0 < t \leq T, \\
 u &= \tilde{g} \quad \text{on } \tilde{\Gamma}_h, \quad 0 < t \leq T, \\
 u &= u_0 \quad \text{in } \tilde{\Omega}_h, \quad t = 0,
 \end{aligned} \tag{3.2}$$

and discretize it by a Galerkin method using finite element trial and test functions on the triangulation $\tilde{\mathcal{T}}_h$. In particular, we could consider again Nitsche's approach [22]. It is based on the following equivalent formulation of Problem (3.2)

$$\begin{aligned}
 \frac{d}{dt}(u, v)_{0, \tilde{\Omega}_h} + (\nabla u, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{n}} u, v)_{0, \tilde{\Gamma}_h} - (f, v)_{0, \tilde{\Omega}_h} \\
 - (u - \tilde{g}, \partial_{\tilde{n}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1}(u - \tilde{g}), v)_{0, \tilde{\Gamma}_h} = 0,
 \end{aligned} \tag{3.3}$$

which hold for any sufficiently smooth test function v in $\tilde{\Omega}_h$, $0 < t \leq T$ (note that they are not time dependent so the time differential operator could go outside the L^2 -inner product in the first term on the left-hand side). Introducing V_h defined in (2.7), we could consider the following Galerkin discretization of Problem (3.2):

$$\begin{aligned}
 \text{Find } \tilde{u}_h \in L^2(0, T; V_h) \cap C^0([0, T]; L^2(\tilde{\Omega}_h)) \text{ such that } \forall v_h \in V_h \\
 \frac{d}{dt}(\tilde{u}_h, v_h)_{0, \tilde{\Omega}_h} + (\nabla \tilde{u}_h, \nabla v_h)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{n}} \tilde{u}_h, v_h)_{0, \tilde{\Gamma}_h} - (\tilde{u}_h, \partial_{\tilde{n}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \tilde{u}_h, v_h)_{0, \tilde{\Gamma}_h} \\
 = (f, v_h)_{0, \tilde{\Omega}_h} - (\tilde{g}, \partial_{\tilde{n}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \tilde{g}, v_h)_{0, \tilde{\Gamma}_h}, \quad 0 < t \leq T,
 \end{aligned} \tag{3.4}$$

To make this scheme practically feasible, we recall the consistent approximation of $\tilde{g}$ made in (2.12). As we have done before, neglecting the higher-order term (with respect to $\|\mathbf{d}\|$) in (2.13), we deduce that the exact solution u satisfies the approximate equations

$$\begin{aligned}
 \frac{d}{dt}(u, v)_{0, \tilde{\Omega}_h} + (\nabla u, \nabla v)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{n}} u, v)_{0, \tilde{\Gamma}_h} - (f, v)_{0, \tilde{\Omega}_h} \\
 - (S_h u - \bar{g}, \partial_{\tilde{n}} v)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1}(S_h u - \bar{g}), S_h v)_{0, \tilde{\Gamma}_h} \approx 0, \quad 0 < t \leq T,
 \end{aligned} \tag{3.5}$$

for any sufficiently smooth test function v in $\tilde{\Omega}_h$. Note that also in this case in the last inner product we have replaced v by $S_h v$ in order to obtain a symmetric stabilization term; on the other hand, unlike (3.3), the left-hand side of (3.5) is no longer symmetric in u and v .

Based on these approximate equations, we are led to introduce the following

Galerkin discretization of Problem (3.2):

$$\begin{aligned}
 & \text{Find } u_h \in L^2(0, T; V_h) \cap C^0([0, T]; L^2(\tilde{\Omega}_h)) \text{ such that } \forall v_h \in V_h \\
 & \frac{d}{dt}(u_h, v_h)_{0, \tilde{\Omega}_h} + (\nabla u_h, \nabla v_h)_{0, \tilde{\Omega}_h} - (\partial_{\tilde{n}} u_h, v_h)_{0, \tilde{\Gamma}_h} \\
 & \quad - (S_h u_h, \partial_{\tilde{n}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} S_h u_h, S_h v_h)_{0, \tilde{\Gamma}_h} \\
 & \quad = (f, v_h)_{0, \tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{n}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \bar{g}, S_h v_h)_{0, \tilde{\Gamma}_h}, \quad 0 < t \leq T,
 \end{aligned} \tag{3.6}$$

with $u_h(0) = u_{0,h} \in V_h$.

As we have done for the bilinear form a_h in (2.17) and the linear form F_h in (2.18), we introduce the bilinear form

$$b(w, v) := (w, v)_{0, \tilde{\Omega}_h}, \tag{3.7}$$

so that the proposed SBM discretization can be written in compact form as follows:

$$\begin{aligned}
 & \text{Find } u_h \in V_h \text{ such that} \\
 & \begin{cases} \frac{d}{dt} b(u_h, v_h) + a_h(u_h, v_h) = F_h(v_h) & \forall v_h \in V_h, \quad 0 < t \leq T, \\ u_h(0) = u_{0,h} & u_{0,h} \in V_h. \end{cases}
 \end{aligned} \tag{3.8}$$

Thanks to the rigorous definition of the infinite-dimensional space $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$ in (2.38), we recall and clarify some properties of the space $L^2(\tilde{\Omega}_h)$. $L^2(\tilde{\Omega}_h)$ is a separable Hilbert space with norm $\|\cdot\|_{0, \tilde{\Omega}_h}$, such that $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \subset L^2(\tilde{\Omega}_h)$. Note that this is still valid for the discrete problem because $V_h \subset V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \subset L^2(\tilde{\Omega}_h)$. We identify $L^2(\tilde{\Omega}_h)$ with its dual $L^2(\tilde{\Omega}_h)'$ by the Riesz representation theorem, so that the following continuous inclusions hold:

$$V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \subset L^2(\tilde{\Omega}_h) = L^2(\tilde{\Omega}_h)' \subset V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)'.$$

Also for the bilinear form b we can establish coercivity and continuity constants. So we can summarize that with the following trivial results.

Proposition 7. *The bilinear form b , defined in $L^2(\tilde{\Omega}_h) \times L^2(\tilde{\Omega}_h)$, is symmetric, continuous and coercive in $L^2(\tilde{\Omega}_h)$; precisely, there exists two constants $\|b\|, \beta > 0$ independent of the mesh size, such that*

$$b(v, v) \geq \beta \|v\|_{0, \tilde{\Omega}_h}^2, \quad \forall v \in L^2(\tilde{\Omega}_h), \tag{3.9}$$

$$|b(w, v)| \leq \|b\| \|w\|_{0, \tilde{\Omega}_h} \|v\|_{0, \tilde{\Omega}_h}, \quad \forall w, v \in L^2(\tilde{\Omega}_h). \tag{3.10}$$

Proof. Noting that b is defined as an $L^2(\tilde{\Omega}_h)$ -inner product, the proof is trivial and the constants are $\beta = \|b\| = 1$. $\square$

Property 2. Let us denote by $\|v\|_b := \sqrt{b(v, v)}$ the norm associated with b , which is uniformly equivalent to the norm of $L^2(\tilde{\Omega}_h)$:

$$\sqrt{\beta}\|v\|_{0,\tilde{\Omega}_h} \leq \|v\|_b \leq \sqrt{\|b\|}\|v\|_{0,\tilde{\Omega}_h}, \quad \forall v \in L^2(\tilde{\Omega}_h). \quad (3.11)$$

Considering the linear form F_h we can say the following.

Property 3. F_h is the time-dependent family of linear continuous forms

$$F_h(t) : V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h) \rightarrow \mathbb{R},$$

i.e., $F_h(t) \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)'$ for all t . We assume that $F_h \in L^2(0, T; V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)')$.

In the end, we formalize the condition on the initial data u_0 .

Property 4. The function $u_0 \in L^2(\tilde{\Omega}_h)$.

Once we have recalled these properties, we can definitely state the time-dependent variational problem on the correct infinite dimensional spaces.

$$[VP_{h,t}] \left\{ \begin{array}{l} \text{Find } u \in L^2(0, T; V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)) \cap C^0([0, T]; L^2(\tilde{\Omega}_h)) \text{ such that} \\ \frac{d}{dt}b(u, v) + a_h(u, v) = F_h(v) \quad \forall v \in V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h), \quad 0 < t \leq T, \\ u(0) = u_0. \end{array} \right. \quad (3.12)$$

3.2 Stability, consistency and convergence analysis - Time

In this section, we analyze the stability, consistency, and convergence of the SBM with respect to the time variable. After having discussed the semi-discretization in space in the previous section, we now introduce the Backward Euler scheme as a time-advancing method. Of course, this is only one of the possible choices; other suitable approaches such as the Crank-Nicolson or Runge-Kutta schemes could also be employed to solve this type of problem. However, we focus on the Backward Euler method because, unlike the Forward Euler scheme, it is *unconditionally stable*.

Before starting the analysis of the method, let us recall the definition of a Bochner space. Moreover, we introduce two useful norms corresponding to the Bochner spaces appearing in (3.12).

Definition 2. Given a normed space X and $1 \leq p \leq \infty$, the Bochner space $L^p(0, T; X)$ is the set of all function $w : (0, T) \rightarrow X$ such that the norm function $t \mapsto \|w(t)\|_X$ belongs to $L^p(0, T)$. It is a normed space with the norm

$$\|w\|_{L^p(0, T; X)} = \left(\int_0^T \|w(t)\|_X^p dt \right)^{1/p}. \quad (3.13)$$

The space $C^0([0, T]; X)$ is the set of all functions $w : [0, T] \rightarrow X$ which are continuous with respect to the norm of X , i.e., which satisfy

$$\lim_{t \rightarrow t_0} \|w(t) - w(t_0)\|_X = 0 \quad \forall t_0 \in [0, T].$$

It is a normed space with norm

$$\|w\|_{C^0([0, T]; X)} = \max_{0 \leq t \leq T} \|w(t)\|_X. \quad (3.14)$$

Once we have discretized the equation with respect to the spatial variable, let us introduce N_T time instants $t_n > 0$ satisfying $0 = t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_{N_T} = T$, and construct for each one the recursive approximation $u_h^n \approx u_h(t_n)$ of the solution; precisely, at time t_{n+1} we define u_h^{n+1} from the approximation $u_h^j, j \leq n$ computed at the previous steps.

The SBM fully discretized heat equation using the Backward Euler method is

$$\begin{cases} \frac{b(u_h^{n+1} - u_h^n, v_h)}{\Delta t} + a_h(u_h^{n+1}, v_h) = F_h(t_{n+1}; v_h), & \forall v_h \in V_h, \quad 0 \leq n \leq N_T - 1, \\ u_h^0 = u_{0,h}. \end{cases} \quad (3.15)$$

This results in a system of N_T equations, where at each time step t_n , as previously described, an SBM problem associated with the spatial variable must be solved. We emphasize the time dependence of the linear form F_h by writing it as $F_h(t; \cdot)$. The corresponding solutions form the sequence $\{u_h^n\}_{1 \leq n \leq N_T}$.

Remark 9. In the sequel, given a sequence $\{g^k : 1 \leq k \leq N_T\}$ contained in some vector space X , we define the piecewise-constant function $g_{\Delta t} : (0, T] \rightarrow X$ as follows:

$$g_{\Delta t}(t) = g^k \quad \text{if } t \in (t_{k-1}, t_k], \quad 1 \leq k \leq N_T. \quad (3.16)$$

3.2.1 Stability

We now derive bounds independent of meshsize h and time step Δt on certain norms of $\{u_h^n\}_{1 \leq n \leq N_T}$. Similar to the continuous case, we will prove the stability of the discrete method.

Proposition 8. *The fully discrete approximation $\{u_h^n\}$ solution of (3.15) satisfies the bounds*

$$\|u_{h,\Delta t}\|_{L^2(0,T;\|\cdot\|_a)} \leq \frac{1}{\alpha} \|F_{h,\Delta t}\|_{L^2(0,T;V_h')} + \frac{1}{\sqrt{\alpha}} \|u_{0,h}\|_b, \quad (3.17)$$

$$\|u_{h,\Delta t}\|_{C^0([0,T];L^2(\tilde{\Omega}_h))} \leq \frac{1}{\sqrt{\alpha}} \|F_{h,\Delta t}\|_{L^2(0,T;V_h')} + \|u_{0,h}\|_b. \quad (3.18)$$

Proof. Choosing $v_h = u_h^{n+1}$ in (3.15) and using the norm induced by the bilinear form b in Property 2, we get for all $0 \leq n \leq N_T - 1$

$$\|u_h^{n+1}\|_b^2 - (u_h^n, u_h^{n+1})_b + \Delta t a_h(u_h^{n+1}, u_h^{n+1}) = \Delta t F_h(t_{n+1}; u_h^{n+1}).$$

Noting that $\frac{1}{2}\|u_h^n - u_h^{n+1}\|_b^2 = \frac{1}{2}\|u_h^{n+1}\|_b^2 - (u_h^n, u_h^{n+1})_b + \frac{1}{2}\|u_h^n\|_b^2$ we can write

$$\frac{1}{2}\|u_h^{n+1}\|_b^2 - \frac{1}{2}\|u_h^n\|_b^2 + \frac{1}{2}\|u_h^n - u_h^{n+1}\|_b^2 + \Delta t a_h(u_h^{n+1}, u_h^{n+1}) = \Delta t F_h(t_{n+1}; u_h^{n+1}).$$

Applying (2.32) and continuity of the linear form F_h

$$\begin{aligned} \frac{1}{2}\|u_h^{n+1}\|_b^2 - \frac{1}{2}\|u_h^n\|_b^2 + \frac{1}{2}\|u_h^n - u_h^{n+1}\|_b^2 \\ + \Delta t \alpha \|u_h^{n+1}\|_a^2 \leq \Delta t \|F_h(t_{n+1})\|_{V_h'} \|u_h^{n+1}\|_a \\ \leq \frac{\Delta t}{2\alpha} \|F_h(t_{n+1})\|_{V_h'}^2 + \frac{\alpha \Delta t}{2} \|u_h^{n+1}\|_a^2, \end{aligned}$$

having used the Young's inequality with $\varepsilon = \sqrt{\alpha}$. Multiplying both sides by 2 and neglecting the term $\|u_h^{n+1} - u_h^n\|_b^2$ which is ≥ 0 , we obtain

$$\|u_h^{n+1}\|_b^2 - \|u_h^n\|_b^2 + \Delta t \alpha \|u_h^{n+1}\|_a^2 \leq \frac{\Delta t}{\alpha} \|F_h(t_{n+1})\|_{V_h'}^2.$$

Now, given any integer $m \in [1, N_T]$, we sum all these inequalities for n ranging from 0 to $m - 1$. Note that, on the left-hand sides, we have a *telescopic sum*, i.e., the second term in the n -th inequalities cancels with the first term in the $(n - 1)$ -th inequality, and so on. Furthermore, the initial condition gives $u_h^0 = u_{0,h}$. In this manner, we arrive at the inequalities

$$\|u_h^m\|_b^2 + \alpha \sum_{k=1}^m \|u_h^k\|_a^2 \Delta t \leq \frac{1}{\alpha} \sum_{k=1}^m \|F_h(t_k)\|_{V_h'}^2 \Delta t + \|u_{0,h}\|_b^2, \quad 1 \leq m \leq N_T. \quad (3.19)$$

Recalling Remark 9, we define the following piecewise-constant functions and their corresponding discrete norms:

- $u_{h,\Delta t} : (0, T] \rightarrow V_h$

$$u_{h,\Delta t}(t) = u_h^k \quad \text{if } t \in (t_{k-1}, t_k], \quad 1 \leq k \leq N_T, \quad (3.20)$$

with

$$\|u_{h,\Delta t}\|_{L^2(0,T;\|\cdot\|_a)} = \left(\sum_{k=1}^{N_T} \|u_h^k\|_a^2 \Delta t \right)^{1/2}$$

and also

$$\|u_{h,\Delta t}\|_{C^0([0,T];L^2(\tilde{\Omega}_h))} = \max_{1 \leq m \leq N_T} \|u_h^m\|_b.$$

- $F_{h,\Delta t} : (0, T] \rightarrow V'_h$

$$F_{h,\Delta t}(t) = F_h(t_k) \quad \text{if } t \in (t_{k-1}, t_k], \quad 1 \leq k \leq N_T, \quad (3.21)$$

with

$$\|F_{h,\Delta t}\|_{L^2(0,T;V'_h)} = \left(\sum_{k=1}^{N_T} \|F_h(t_k)\|_{V'_h}^2 \Delta t \right)^{1/2}.$$

Now, using (3.20) and (3.21), we can derive the two bounds with the two different norms.

At first, we choose $m = N_T$ in (3.19) and we neglect the first term on the left-hand side, which is ≥ 0 . This yields

$$\sum_{k=1}^{N_T} \|u_h^k\|_a^2 \Delta t \leq \frac{1}{\alpha^2} \sum_{k=1}^{N_T} \|F_h(t_k)\|_{V'_h}^2 \Delta t + \frac{1}{\alpha} \|u_{0,h}\|_b^2.$$

Taking the square root on both sides and applying the triangular inequality, we obtain (3.17).

Next, neglecting the sum on the left-hand side of (3.19) and taking the maximum value among the m , we get

$$\max_{1 \leq m \leq N_T} \|u_h^m\|_b^2 \leq \frac{1}{\alpha} \sum_{k=1}^{N_T} \|F_h(t_k)\|_{V'_h}^2 \Delta t + \|u_{0,h}\|_b^2.$$

(3.18) is obtained taking again the square root on both sides and using again the triangular inequality. $\square$

Few comments on the Proposition 8 are necessary. The bounds introduced prove the numerical stability of the SBM/Backward-Euler method. Indeed, if we assume F_h continuous (or just piecewise-continuous) in time, then

$$\|F_{h,\Delta t}\|_{L^2(0,T;V'_h)} \leq \|F_{h,\Delta t}\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)')} \rightarrow \|F_h\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)')} \quad \text{as } \Delta t \rightarrow 0,$$

since the integral of a continuous function is the limit of the integral sums of its piecewise- constant approximations; this implies

$\|F_{h,\Delta t}\|_{L^2(0,T;V_h')} \leq C_1 \|F_h\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)')}$ for some constant $C_1 > 0$ independent of Δt . On the other hand, it is not restrictive to assume that $u_{0,h}$ satisfies $\|u_{0,h}\|_{0,\tilde{\Omega}_h} \leq C_2 \|u_0\|_{0,\tilde{\Omega}_h}$ for some constant $C_2 > 0$ independent of h . We conclude that, under these assumptions, the right-hand side of (3.17) and (3.18) can be bounded by quantities depending on the data of the problem but independent of meshsize h and time step Δt .

Remark 10. Note that, consistently with the approach adopted in the previous chapter for Proposition 2, the stability results established in Proposition 8 do not require the stronger norm $\|\cdot\|_{V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)}$, which is instead reserved for the consistency and convergence analysis. Furthermore, it should be observed that in the discrete setting these two norms coincide; specifically, as stated in Section 2.2, by considering (2.40) it follows that

$$\|v_h\|_a = \|v_h\|_{V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)} \quad \forall v_h \in V_h.$$

From now on, in order to simplify the readability of the Bochner space, we will use $V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)$ instead of $\|\cdot\|_a$ even if the function does not belong to V_h .

3.2.2 Consistency

Considering the exact solution u of Problem (3.15), let us set $u^n = u(t_n)$ to simplify the notation. We are interested in estimating some norm of the errors

$$e^n := u^n - u_h^n, \quad n = 1, \dots, N_T. \quad (3.22)$$

To this end, let us introduce a projection operator $P_h : V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h) \rightarrow V_h$, which is strictly related to the Lemma 4 and satisfies the following.

Definition 3. Given $w \in V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)$, let $P_h w$ be the a_h -projection of w upon V_h , i.e.,

$$a_h(P_h w, v_h) = a_h(w, v_h) + e_\perp(v_h) \quad \forall v_h \in V_h, \quad (3.23)$$

where the term $e_\perp(v_h) := -(R_h w, \partial_{\tilde{n}} v_h)_{0,\tilde{\Gamma}_h} + \gamma(h^{-1} R_h w, S_h v_h)_{0,\tilde{\Gamma}_h} \neq 0$ since this is not an orthogonal projection.

Remark 11. Note that this choice is not unique. Other projection operator are also possible to be chosen, but they will modify the following computations for the proof of the consistency. For instance, one other possible choice should be the b -projector related to the bilinear form b .

Then, we write

$$e^n = (u^n - P_h u^n) + (P_h u^n - u_h^n) = \pi^n + \theta_h^n. \quad (3.24)$$

The quantities $\pi^n = (u^n - P_h u^n)$ are projection errors, which depend on the approximation properties of V_h . In particular, such errors can be studied by the technique presented in Section 2.4. We briefly summarize the previous chapter's results considering the following stationary boundary value problem. Given $w \in H^2(\Omega)$, set $\phi = -\Delta w$ and $g = w|_\Gamma$. Then, w is the solution of the following:

$$\begin{aligned} -\Delta w &= \phi \text{ in } \Omega, \\ w &= g \text{ on } \Gamma. \end{aligned} \quad (3.25)$$

If we approximate the problem (3.25) by the SBM, we find $w_{SBM} \in V_h$ such that:

$$a_h(w_{SBM}, v_h) = G_h(v_h) \quad \forall v_h \in V_h$$

with a_h defined in (2.17),

$$\begin{aligned} G_h(v_h) &= (\phi, v_h)_{0, \tilde{\Omega}_h} - (\bar{g}, \partial_{\tilde{n}} v_h)_{0, \tilde{\Gamma}_h} + \gamma(h^{-1} \bar{g}, S_h v_h)_{0, \tilde{\Gamma}_h}, \\ \bar{g} &= [S_h + R_h](w) \end{aligned}$$

where S_h is defined in (2.14) and R_h is the Taylor remainder. We define

$$P_h w := w_{SBM}$$

and, from Theorem 1 and 2, we have the bounds

$$\|w - P_h w\|_a \lesssim h |w|_{2, \tilde{\Omega}_h}, \quad (3.26)$$

$$\|w - P_h w\|_{0, \tilde{\Omega}_h} \lesssim h^{3/2} |w|_{2, \tilde{\Omega}_h} \quad (3.27)$$

for all $w \in H^2(\Omega)$. Of course other choices are possible for example $W^{1+s,p}(\Omega)$ but we prefer to have a lighter notation. This is a consistency assumption on V_h , which has the following consequence.

Proposition 9. *Assume that $u \in C^0([0, T]; H^2(\tilde{\Omega}_h))$. Thus, the error function $\pi_{\Delta t}$ (defined according to Remark 9) satisfies*

$$\|\pi_{\Delta t}\|_{C^0([0, T]; L^2(\tilde{\Omega}_h))} + \|\pi_{\Delta t}\|_{L^2(0, T; V(\tilde{\Omega}_h; \tilde{\Gamma}_h))} \leq C h \|u\|_{C^0([0, T]; H^2(\tilde{\Omega}_h))}. \quad (3.28)$$

Proof. For the first term of the left-hand side in (3.28) we just have to take the maximum value of the $\|\pi^n\|_{L^2(\tilde{\Omega}_h)}$ among the bound (3.27) for each time step,

$$\begin{aligned} \|\pi_{\Delta t}\|_{C^0([0, T]; L^2(\tilde{\Omega}_h))} &= \max_{1 \leq m \leq N_T} \|\pi^m\|_{L^2(\tilde{\Omega}_h)} \leq \max_{1 \leq m \leq N_T} c' h^{3/2} \|u^m\|_{H^2(\tilde{\Omega}_h)} \\ &\leq c'' h^{3/2} \|u\|_{C^0([0, T]; H^2(\tilde{\Omega}_h))}, \end{aligned} \quad (3.29)$$

noting that the discrete C^0 -norm is bounded above by the continuous one. On the other hand, for the second left-hand side in (3.28) we can write

$$\|\pi_{\Delta t}\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h))}^2 = \sum_{k=1}^{N_T} \|\pi^k\|_{V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h)}^2 \Delta t \leq c^2 h^2 \underbrace{\sum_{k=1}^{N_T} \|u^k\|_{H^2(\tilde{\Omega}_h)}^2 \Delta t}_{\|u_{\Delta t}\|_{L^2(0,T;H^2(\tilde{\Omega}_h))}^2},$$

using Corollary 1. Recalling that $u \in C^0([0, T]; H^2(\tilde{\Omega}_h))$, then

$$\|u_{\Delta t}\|_{L^2(0,T;H^2(\tilde{\Omega}_h))}^2 \leq \max_k \left(\sum_{k=1}^{N_T} \|u^k\|_{H^2(\tilde{\Omega}_h)}^2 \Delta t \right) \leq T \|u\|_{C^0([0,T];H^2(\tilde{\Omega}_h))}^2$$

and we have

$$\|\pi_{\Delta t}\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h))} \leq C h \|u\|_{C^0([0,T];H^2(\tilde{\Omega}_h))}. \quad (3.30)$$

Combining (3.29) and (3.30) we obtain the result. $\square$

Next, we focus on the errors $\theta_h^n = P_h u^n - u_h^n$, which are elements in V_h depending on the time discretization. The idea is to recognize that they are the solution of recurrence equations similar to (3.15); this will allow us to apply to them stability estimates similar to (3.17) or (3.18).

So let us write the left-hand side of (3.15) with θ_h^n instead of u_h^n .

$$\begin{aligned} \frac{b(\theta_h^{n+1} - \theta_h^n, v_h)}{\Delta t} + a_h(\theta_h^{n+1}, v_h) &= \frac{b(P_h u^{n+1} - u_h^{n+1}, v_h)}{\Delta t} - \frac{b(P_h u^n - u_h^n, v_h)}{\Delta t} \\ &\quad + a_h(P_h u^{n+1} - u_h^{n+1}, v_h) \\ &= \frac{b(P_h u^{n+1}, v_h)}{\Delta t} - \frac{b(P_h u^n, v_h)}{\Delta t} - \underbrace{\frac{b(u_h^{n+1}, v_h)}{\Delta t} + \frac{b(u_h^n, v_h)}{\Delta t}}_{-F_h(t_{n+1}; v_h)} - a_h(u_h^{n+1}, v_h) \\ &\quad + a_h(P_h u^{n+1}, v_h), \end{aligned}$$

where the labeled terms are the forcing term of (3.15) with opposite sign. Then, applying (3.23), we obtain

$$\frac{b(P_h(u^{n+1} - u^n), v_h)}{\Delta t} - F_h(t_{n+1}; v_h) + a_h(u^{n+1}, v_h) + e_{\perp}^n(v_h),$$

where $e_{\perp}^n(v_h)$ is the projection error at time t_n . This term is a spatial discretization error that could be added to the term π^n . Recalling (2.19), we can define the forcing term of the equation as

$$E(t_{n+1}; v_h) = \frac{b(P_h(u^{n+1} - u^n), v_h)}{\Delta t} - \left. \frac{d}{dt} b(u, v_h) \right|_{t=t_{n+1}} + e_{\perp}^n(v_h) \quad (3.31)$$

and finally we have

$$\frac{b(\theta_h^{n+1} - \theta_h^n, v_h)}{\Delta t} + a_h(\theta_h^{n+1}, v_h) = E(t_{n+1}, v_h), \quad \forall v_h \in V_h. \quad (3.32)$$

By Proposition 8 we have the stability of the method for the solutions $\{\theta_h^n\}$. Combining (3.17) and (3.18) we have the following.

Proposition 10. *Assume that $u \in C^0([0, T]; H^2(\tilde{\Omega}_h))$. Thus, the error functions $\theta_{h,\Delta t}, E_{\Delta t}$ (defined according to Remark 9) satisfies, with $C_\alpha > 0$,*

$$\|\theta_{h,\Delta t}\|_{C^0([0,T];L^2(\tilde{\Omega}_h))} + \|\theta_{h,\Delta t}\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\tau}_h))} \leq C_\alpha \left(\|\theta_{0,h}\|_b + \|E_{\Delta t}\|_{L^2(0,T;V'_h)} \right), \quad (3.33)$$

with $c > 0$

$$\|\theta_{0,h}\|_b \leq c h^{3/2} \|u_0\|_{H^2(\tilde{\Omega}_h)}, \quad (3.34)$$

and assuming $\frac{\partial u}{\partial t} \in L^2(0, T; H^2(\tilde{\Omega}_h))$, $\frac{\partial^2 u}{\partial t^2} \in L^2(0, T; L^2(\tilde{\Omega}_h))$, there exists constants $C_1, C_2 > 0$ independent of h and Δt such that

$$\begin{aligned} \|E_{\Delta t}\|_{L^2(0,T;V'_h)} &\leq C_1 h^{3/2} \left(\|u\|_{C^0([0,T];H^2(\tilde{\Omega}_h))} + \left\| \frac{\partial u}{\partial t} \right\|_{L^2(0,T;H^2(\tilde{\Omega}_h))} \right) \\ &\quad + C_2 \Delta t \left\| \frac{\partial^2 u}{\partial t^2} \right\|_{L^2(0,T;L^2(\tilde{\Omega}_h))}. \end{aligned} \quad (3.35)$$

Proof. The bound on $\|\theta_{0,h}\|_b$ is an easy application of (3.11) and (3.27), so

$$\begin{aligned} \|\theta_{0,h}\|_b &\leq \sqrt{\|b\|} \|\theta_{0,h}\|_{0,\tilde{\Omega}_h} = \sqrt{\|b\|} \|P_h u_0 - u_{0,h} - u_0 + u_0\|_{0,\tilde{\Omega}_h} \\ &= \sqrt{\|b\|} \|P_h u_0 - u_0\|_{0,\tilde{\Omega}_h} + \underbrace{\sqrt{\|b\|} \|u_0 - u_{0,h}\|_{0,\tilde{\Omega}_h}}_{=0 \text{ by CI}} \leq c h^{3/2} \|u_0\|_{H^2(\tilde{\Omega}_h)}. \end{aligned}$$

Recalling the definition of $\|E_{\Delta t}\|_{L^2(0,T;V'_h)}^2 = \sum_{k=1}^{N_T} \|E(t_k; v_h)\|_{V'_h}^2 \Delta t$ where

$$\|E(t_k; v_h)\|_{V'_h} = \sup_{v_h \in V_h} \frac{E(t_k; v_h)}{\|v_h\|_a},$$

we are intrested in estimating $|E(t_k; v_h)|$. Adding and subtracting $\frac{1}{\Delta t} b(u^{n+1} - u^n, v_h)$ in (3.31) we obtain

$$E(t_{n+1}; v_h) = \underbrace{\frac{b([P_h - I](u^{n+1} - u^n), v_h)}{\Delta t}}_{E_{P_h, \Delta t}} + \underbrace{\frac{b(u^{n+1} - u^n, v_h)}{\Delta t} - \frac{d}{dt} b(u, v_h) \Big|_{t_{n+1}}}_{E_{dt}} + e_\perp^n$$

The projection error is the same as (2.72) which is

$$|e_{\perp}^n(v_h)| \leq C_{0,3} h^{3/2} K(u^n, f^n, g^n) \|v_h\|_{0, \tilde{\Omega}_h}, \quad (3.36)$$

where $f^n = f(t_n)$, $g^n = g(t_n)$ and having observed that $r = 1 \leq \sigma$ by def in (2.54). Note that, recalling that $\|\cdot\|_a$ controls from above $\|\cdot\|_{1, \tilde{\Omega}_h}$ (see Remark 4) and $\|\cdot\|_{0, \tilde{\Omega}_h} \lesssim \|\cdot\|_{1, \tilde{\Omega}_h}$, we have

$$\|w_h\|_{0, \tilde{\Omega}_h} \lesssim \|w_h\|_a \quad \forall w_h \in V_h. \quad (3.37)$$

In order to find the bound for $E_{P_h, \Delta t}$, we apply the fundamental theorem of calculus and we exploit the commutativity between the projection operator P_h and the temporal derivative, so

$$\begin{aligned} |E_{P_h, \Delta t}(t_{n+1}; v_h)| &= \frac{1}{\Delta t} \left| \int_{t_n}^{t_{n+1}} b \left([P_h - I] \frac{\partial u}{\partial t}(s), v_h \right) ds \right| \\ &\leq \frac{\|b\|}{\Delta t} \int_{t_n}^{t_{n+1}} \left\| \left(P_h \frac{\partial u}{\partial t} - \frac{\partial u}{\partial t} \right)(s) \right\|_{L^2(\tilde{\Omega}_h)} ds \|v_h\|_{L^2(\tilde{\Omega}_h)}. \end{aligned}$$

Applying (3.27) we get

$$|E_{P_h, \Delta t}(t_{n+1}; v_h)| \leq \frac{1}{\Delta t} c h^{3/2} \int_{t_n}^{t_{n+1}} \left\| \frac{\partial u}{\partial t}(s) \right\|_{H^2(\tilde{\Omega}_h)} ds \|v_h\|_{0, \tilde{\Omega}_h}. \quad (3.38)$$

The final term E_{dt} is bounded using the Taylor formula with the reminder in integral form and the Cauchy-Schwarz inequality.

$$\begin{aligned} |E_{dt}(t_{n+1}; v_h)| &= \frac{1}{\Delta t} \left| \int_{t_n}^{t_{n+1}} (s - t_n) \frac{d^2}{dt^2} b(u, v_h) ds \right| \\ &\leq \frac{\|b\| \Delta t^{3/2}}{\sqrt{3} \Delta t} \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^2 u}{\partial t^2}(s) \right\|_{L^2(\tilde{\Omega}_h)}^2 ds \right)^{1/2} \|v_h\|_{L^2(\tilde{\Omega}_h)} \\ &\leq c''' \Delta t^{1/2} \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^2 u}{\partial t^2}(s) \right\|_{L^2(\tilde{\Omega}_h)}^2 ds \right)^{1/2} \|v_h\|_{0, \tilde{\Omega}_h}. \end{aligned} \quad (3.39)$$

Combining (3.37) with (3.36), (3.38) and (3.39) we obtain

$$\begin{aligned}
 \|E_{\Delta t}\|_{L^2(0,T;V_h')}^2 &= \sum_{k=1}^{N_T} \|E(t_k; v_h)\|_{V_h'}^2 \Delta t \\
 &\leq \sum_{k=1}^{N_T} \left(|E_{P_h, \Delta t}(t_k; v_h)|^2 + |E_{dt}(t_k; v_h)|^2 + |e_{\perp}^k(v_h)|^2 \right) \frac{\Delta t}{\|v_h\|_a^2} \\
 &\leq \sum_{k=1}^{N_T} \left[\Delta t \frac{c^2 h^3}{\Delta t^2} \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial u}{\partial t}(s) \right\|_{H^2(\tilde{\Omega}_h)}^2 ds \right)^2 \right. \\
 &\quad \left. + \Delta t (c''')^2 \Delta t \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^2 u}{\partial t^2}(s) \right\|_{L^2(\tilde{\Omega}_h)}^2 ds \right) \right. \\
 &\quad \left. + \Delta t C_{0,3}^2 h^3 K(u^k, f^k, g^k)^2 \right] \\
 &\leq \sum_{k=1}^{N_T} \left[\frac{c^2 h^3}{\Delta t} \Delta t \int_{t_n}^{t_{n+1}} \left\| \frac{\partial u}{\partial t}(s) \right\|_{H^2(\tilde{\Omega}_h)}^2 ds \right. \\
 &\quad \left. + \Delta t (c''')^2 \Delta t \left(\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^2 u}{\partial t^2}(s) \right\|_{L^2(\tilde{\Omega}_h)}^2 ds \right) \right. \\
 &\quad \left. + \Delta t C_{0,3}^2 h^3 K(u^k, f^k, g^k)^2 \right] \\
 &\leq \underbrace{c^2 h^3 \int_0^T \left\| \frac{\partial u}{\partial t}(s) \right\|_{H^2(\tilde{\Omega}_h)}^2 ds}_{\|\cdot\|_{L^2(0,T;H^2(\tilde{\Omega}_h))}^2} + \underbrace{(c''')^2 \Delta t^2 \int_0^T \left\| \frac{\partial^2 u}{\partial t^2}(s) \right\|_{L^2(\tilde{\Omega}_h)}^2 ds}_{\|\cdot\|_{L^2(0,T;L^2(\tilde{\Omega}_h))}^2} \\
 &\quad + T C_{0,3}^2 h^3 \|u\|_{C^0([0,T];H^2(\tilde{\Omega}_h))}^2,
 \end{aligned}$$

where we have applied again Cauchy-Schwartz inequality for the first term of the right-hand side. Taking the square root of the last equation, we get the result. $\square$

3.2.3 Convergence

Once we have provided the stability and the consistency of the fully discrete SBM/Backward-Euler method, we have all the ingredients to estimate the errors $\{e^n\} = \{u^n - u_h^n\}$ introduced in (3.22) and prove the convergence.

To this end, it is enough to apply triangle inequality of norms to the decomposition (3.24). The errors $\{\pi^n\}$ are estimated in Proposition 9, whereas the errors $\{\theta^n\}$ are estimated in Proposition 10. In this way, we obtain the following result.

Theorem 3. *Under previous assumptions in Definition 3, Proposition 9 and 10, let $\{u_h^n\}$ be defined by the SBM/Backward-Euler discretization in (3.15). Then, there exists a constant $C(u_0, u, u_t, u_{tt}, f, g)$ dependent on its arguments but independent of meshsize h and time step Δt such that*

$$\|e_{\Delta t}\|_{C^0([0,T];L^2(\tilde{\Omega}_h))} + \|e_{\Delta t}\|_{L^2(0,T;V(\tilde{\Omega}_h;\tilde{\mathcal{T}}_h))} \leq C(u_0, u, u_t, u_{tt}, f, g) (h + \Delta t) . \quad (3.40)$$

Proof. Recalling (3.24) also $e_{\Delta t} = \pi_{\Delta t} + \theta_{\Delta t}$, and combining (3.28) and (3.33 - 3.35) we get the result. $\square$

Looking at the proof of Theorem 3, as expected, the spatial error estimate is associated with the term $\pi_{\Delta t}$, while the temporal error estimate is associated with $\theta_{\Delta t}$. Although from (3.35) we have $\|\theta_{\Delta t}\|_{B(X)} \sim h^{3/2}$, this contribution is negligible, since $\|\pi_{\Delta t}\|_{B(X)} \sim h$ as $h \rightarrow 0$. Here, $B(X)$ denotes a generic Bochner space associated with the spatial norm X , which in the present setting corresponds to $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$.

Remark 12. The result confirms that the SBM/Backward-Euler is a first-order method in time and all the comments on the optimal rate decay of the spatial error, done in Section 2.4, are still valid. If one chooses $\Delta t \sim h$, the two parts of the error (space and time) decay with the same rate. On the other hand, if we are interested in the L^2 -norm (instead of $V(\tilde{\Omega}_h; \tilde{\mathcal{T}}_h)$) for the spatial error in (3.28), considering the suboptimality, the estimate of the error can be proven that satisfies

$$\|e_{\Delta t}\|_{C^0([0,T];L^2(\tilde{\Omega}_h))} + \|e_{\Delta t}\|_{L^2(0,T;L^2(\tilde{\Omega}_h))} \leq C(u_0, u, u_t, u_{tt}, f, g) (h^{r+r_0} + \Delta t) , \quad (3.41)$$

with r, r_0 following the definitions in (2.54) and (2.73). Also in this case, if one chooses $\Delta t \sim h^{r+r_0}$ we will have a balanced contribution for spatial and temporal errors ($r = 1, r_0 = 1/2$ if $s = 1, p = 2$).

Remark 13. Introducing a parameter $\theta_M \in [0,1]$, we can have a general formula of (3.15) as the following $\forall v_h \in V_h, 0 \leq n \leq N_T - 1$,

$$\begin{aligned} \frac{b(u_h^{n+1} - u_h^n, v_h)}{\Delta t} + a_h \left(\theta_M u_h^{n+1} + (1 - \theta_M) u_h^n, v_h \right) \\ = \theta_M F_h(t_{n+1}; v_h) + (1 - \theta_M) F_h(t_n; v_h) . \end{aligned} \quad (3.42)$$

The Backward Euler scheme is obtained setting $\theta_M = 1$. However, with $\theta_M = 1/2$ it can be proven that we have a second order method which is the well-known Crank–Nicolson scheme. Moreover, with a second order method the right-hand side term of (3.40) is $C^*(h + \Delta t^2)$ with $C^* > 0$ independent on h and Δt , so we need to choose $\Delta t \sim \sqrt{h}$ in order to have the same decay for spatial and temporal errors ($\Delta t \sim h$ for L^2 -norm estimate).

3.3 Numerical Results

This section presents numerical experiments for the heat equation solved using the SBM. The spatial semi-discretization is the same as that described in Chapter 2, while the time advancing schemes considered are Backward Euler and Crank–Nicolson. For convenience, the generic θ_M -method (3.42) is implemented; however, only the two classical choices $\theta_M = 1$ (Backward Euler) and $\theta_M = \frac{1}{2}$ (Crank–Nicolson) are considered.

First, the square domain introduced in Section 2.5.1 is used to verify the expected orders of convergence of both time advancing schemes. Subsequently, the methods are also tested on the quarter-circle domain described in Section 2.5.4, confirming the theoretical results stated in Theorem 3.

The errors are computed with respect to the norms (3.13) with $p = 2$ and (3.14), considering both $X = L^2(\tilde{\Omega}_h)$ and $X = H_0^1(\tilde{\Omega}_h)$. In the following, however, we focus primarily on the norms employed in the theoretical error estimates.

3.3.1 Square domain

In this section, the heat equation is solved on the same domain introduced in Section 2.5.1 for $t \in [0, \frac{1}{2}]$. The surrogate domain is also the same. Nitsche’s penalty parameter is set to $\gamma = 10$.

To verify the correct implementation of the two time advancing schemes, we consider exact solutions of the following form:

$$w(\mathbf{x}, t) = \sum_{i=0}^p v_i(\mathbf{x}) t^i, \quad p \in \mathbb{N}, \quad (3.43)$$

where each v_i is a generic function depending only on the spatial variable. Observe that this family of functions, which is polynomial in time, is exactly reproduced by a time advancing method of order p .

Consequently, if a solution linear in time ($p = 1$) is chosen, no temporal discretization error is introduced when using $\theta_M = 1$ (Backward Euler). Moreover, if the solution is at most quadratic in time ($p = 2$), the Crank–Nicolson scheme ($\theta_M = \frac{1}{2}$) yields an exact approximation with respect to the temporal variable. Since the discretization error depends solely on the spatial variables, we expect to observe the same convergence rates established in the previous chapter (see Section 2.5). Specifically, the error is characterized by a linear convergence rate in the H_0^1 -norm and a quadratic rate in the L^2 -norm.

Considering (3.43) the heat equation is solved with these test functions with $p = 1$:

- $w_1 = x + y + \left((1-x)(1-y)(e^x + \tan(xy)) + 1\right) t$

- $w_2 = \sin(y) + (e^x + \tan(xy)) t$

and with $p = 2$:

- $w_3 = y^4 + x^2 t + xy^2 t^2$
- $w_4 = \tan(xy) t^2$

Obviously, the corresponding right-hand side f and boundary data g are computed by substituting the exact solutions defined above into (3.1).

Recall that, if the temporal approximation of the solution is exact, the temporal discretization error is of the order of machine precision. Although it is not necessary to refine the time step Δt in this case, for all the numerical tests reported in Table 3.1 we set $\Delta t \sim h$.

p	Method	Exact solution	C^0 - L^2	C^0 - H^1	L^2 - L^2	L^2 - H^1	Total
1	BE	w_1	2.22	1.09	2.35	1.22	1.30
1	BE	w_2	2.30	1.05	2.37	1.17	1.23
2	CN	w_3	2.29	1.35	2.24	1.23	1.32
2	CN	w_4	2.30	0.92	2.50	1.13	1.18

Table 3.1: Summary of convergence rates for different exact solutions. Backward Euler is exact for linear time dependence, while Crank–Nicolson is exact for quadratic time dependence. C^0 - L^2 stands for $C^0([0, T]; L^2(\tilde{\Omega}_h))$ and so on.

Table 3.1 summarizes the observed convergence rates for different exact solutions and time dependencies starting from $h_{\text{input}} = 1/8$ with four refinements. As expected, Backward Euler exhibits optimal behavior for solutions linear in time, while Crank–Nicolson recovers the optimal temporal accuracy for quadratic time dependence. In all cases, the spatial convergence rates are consistent with the theoretical predictions.

Non-Polynomial in time

Once we have confirmed the order of the time advancing methods, we analyze the problem with an exact solution u that cannot be defined as in (3.43). In particular it is chosen

$$u = \sqrt{x^2 + y^2} \sin(t),$$

computing the corresponding f and g .

As shown in Figure 3.1, the numerical solution is compared with the exact solution. The time step is chosen here solely to simplify the visualization of the

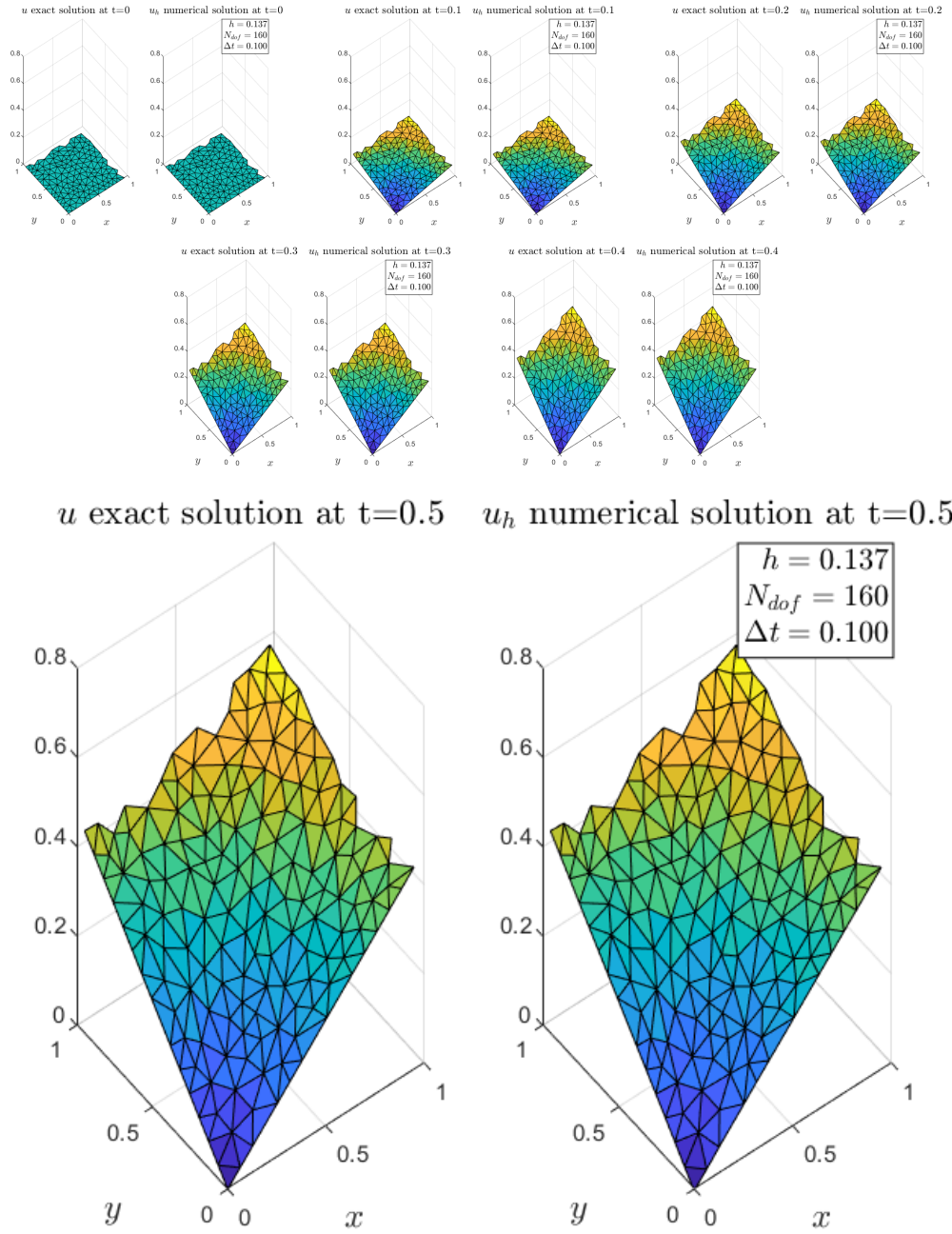


Figure 3.1: [Square] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.137, \Delta t = 0.1$.

solutions. However, when mesh refinements are performed, it will be selected appropriately in accordance with Remark 12 and Remark 13.

In Tables 3.2 and 3.4, the errors in all considered Bochner norms are reported. As discussed previously, the spatial energy norms are optimal, whereas the L^2 norm is suboptimal. The total error remains of first order, which is consistent with the theoretical expectation.

The expected quadratic decay rate is correctly observed in Tables 3.3 and 3.5. Comparing Backward Euler to Crank–Nicolson, the latter, due to its higher order, achieves smaller error magnitudes for the same convergence rate. In Figure 3.2, the convergence rates for the norms $C^0([0, T]; L^2(\tilde{\Omega}_h))$ and $L^2(0, T; H_0^1(\tilde{\Omega}_h))$ are shown along with their theoretical slopes. Since the $L^2(0, T; H_0^1(\tilde{\Omega}_h))$ norm was already optimal, Figure 3.3 displays only the $C^0([0, T]; L^2(\tilde{\Omega}_h))$ norm, which is finally observed to be optimal as well.

Backward Euler $\Delta t \sim h$				
h_{input}	1/8	1/16	1/32	1/64
h	0.222	0.125	0.0663	0.0360
$\ e_{\Delta t}\ _{C^0([0, T]; L^2(\tilde{\Omega}_h))}$	1.27×10^{-3} —	6.34×10^{-4} 1.21	3.54×10^{-4} 0.92	2.11×10^{-4} 0.85
$\ e_{\Delta t}\ _{C^0([0, T]; H_0^1(\tilde{\Omega}_h))}$	4.64×10^{-2} —	2.80×10^{-2} 0.88	1.31×10^{-2} 1.20	7.51×10^{-3} 0.91
$\ e_{\Delta t}\ _{L^2(0, T; L^2(\tilde{\Omega}_h))}$	6.71×10^{-4} —	2.98×10^{-4} 1.42	1.47×10^{-4} 1.11	8.39×10^{-5} 0.92
$\ e_{\Delta t}\ _{L^2(0, T; H_0^1(\tilde{\Omega}_h))}$	2.45×10^{-2} —	1.37×10^{-2} 1.02	5.77×10^{-3} 1.36	3.18×10^{-3} 0.98
Total error	2.58×10^{-2} —	1.44×10^{-2} 1.02	6.13×10^{-3} 1.35	3.39×10^{-3} 0.97

Table 3.2: [Square] Errors for the Backward Euler scheme and successive mesh refinements.

Backward Euler $\Delta t \sim h^2$				
h_{input}	1/8	1/16	1/32	1/64
h	0.222	0.125	0.0663	0.0360
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	1.52×10^{-3}	4.71×10^{-4}	1.11×10^{-4}	3.41×10^{-5}
	–	1.91	2.27	1.93
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	5.09×10^{-2}	2.79×10^{-2}	1.39×10^{-2}	7.92×10^{-3}
	–	0.87	1.02	0.91
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	6.64×10^{-4}	1.98×10^{-4}	4.58×10^{-5}	1.41×10^{-5}
	–	1.98	2.33	1.93
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	2.24×10^{-2}	1.19×10^{-2}	5.77×10^{-3}	3.29×10^{-3}
	–	0.96	1.09	0.92
Total error	2.40×10^{-2}	1.23×10^{-2}	5.88×10^{-3}	3.33×10^{-3}
	–	0.98	1.11	0.93

Table 3.3: [Square] Errors for the Backward Euler scheme and successive mesh refinements.

Crank–Nicolson $\Delta t \sim \sqrt{h}$				
h_{input}	1/8	1/16	1/32	1/64
h	0.222	0.125	0.0663	0.0360
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	1.33×10^{-3}	4.34×10^{-4}	2.27×10^{-4}	1.19×10^{-4}
	–	1.96	1.02	1.06
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	4.90×10^{-2}	2.02×10^{-2}	7.48×10^{-3}	6.13×10^{-3}
	–	1.55	1.56	0.33
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	9.12×10^{-4}	2.58×10^{-4}	1.15×10^{-4}	6.10×10^{-5}
	–	2.21	1.27	1.04
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	3.36×10^{-2}	1.20×10^{-2}	3.80×10^{-3}	3.01×10^{-3}
	–	1.80	1.81	0.38
Total error	3.50×10^{-2}	1.25×10^{-2}	4.02×10^{-3}	3.12×10^{-3}
	–	1.80	1.79	0.42

Table 3.4: [Square] Errors for the Crank–Nicolson scheme and successive mesh refinements.

Crank–Nicolson $\Delta t \sim h$				
h_{input}	1/8	1/16	1/32	1/64
h	0.222	0.125	0.0663	0.0360
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	1.46×10^{-3} –	4.99×10^{-4} 1.88	1.08×10^{-4} 2.41	3.28×10^{-5} 1.95
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	4.64×10^{-2} –	2.80×10^{-2} 0.885	1.30×10^{-2} 1.20	7.46×10^{-3} 0.92
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	7.53×10^{-4} –	2.39×10^{-4} 2.01	4.69×10^{-5} 2.57	1.37×10^{-5} 2.01
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	2.45×10^{-2} –	1.37×10^{-2} 1.02	5.74×10^{-3} 1.37	3.15×10^{-3} 0.98
Total error	2.60×10^{-2} –	1.42×10^{-2} 1.06	5.85×10^{-3} 1.40	3.19×10^{-3} 0.99

Table 3.5: [Square] Errors for the Crank–Nicolson scheme and successive mesh refinements.

First Order Convergence Errors

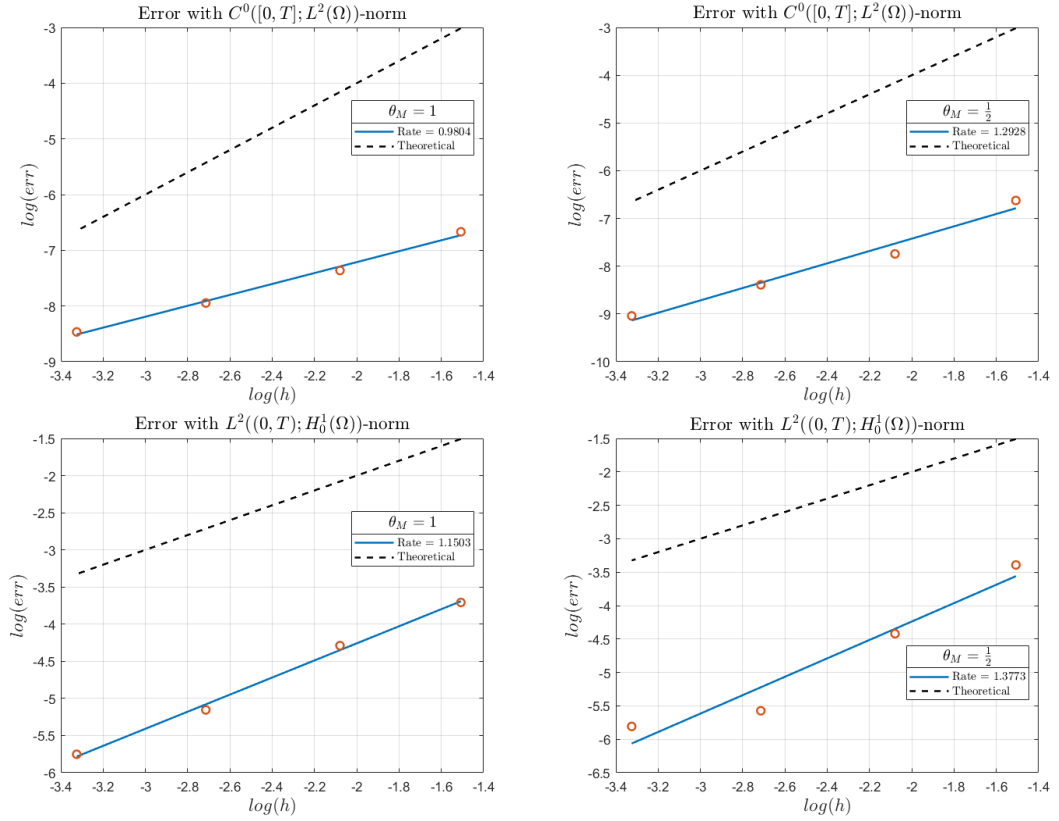


Figure 3.2: [Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the right side ($\Delta t \sim \sqrt{h}$).

Second Order Convergence Errors

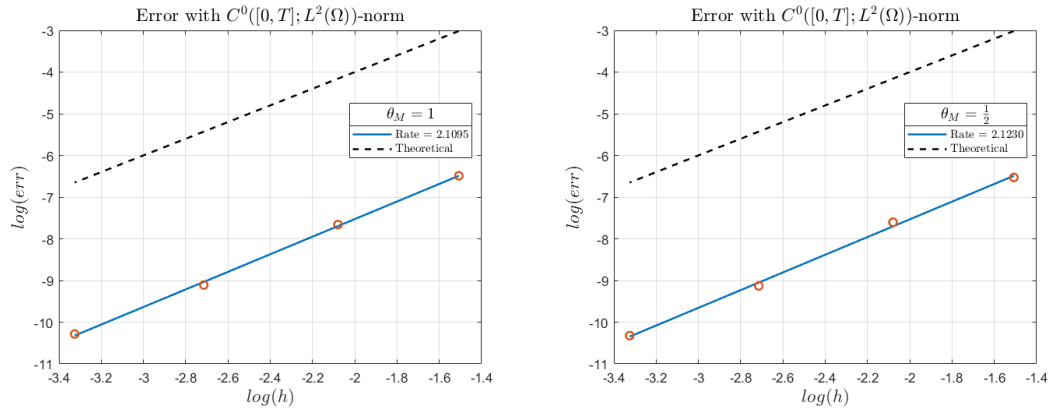


Figure 3.3: [Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h^2$) and Crank–Nicolson on the right side ($\Delta t \sim h$).

3.3.2 Quarter-circle domain

In this section, numerical experiments are performed in the quarter-circle domain defined in Section 2.5.4 for $t \in [0, 1/2]$. The surrogate domain remains the same, and the Nitsche penalization parameter γ is set to 100.

As the exact solution, we consider the same function u defined in (2.77), but with a time-dependent radial component. Specifically, it takes the form

$$\varphi(r, t) = \frac{3}{10} \sin(4\pi r t) r^\alpha \eta(r).$$

The parameters are still α , Pe , and k . The test cases are similar to those in the stationary problem, but here we choose $k = 1$ and $k = 3$ to produce simpler solution shapes. Note that k does not introduce any critical behavior in the numerical solution, unlike α and Pe .

Test 1: $\alpha = 1, \text{Pe} = 5, k = 1$

Numerical and exact solutions are compared in Figure 3.4 for all time steps, showing no significant differences.

The convergence rates are displayed in Figure 3.5, comparing Backward Euler and Crank–Nicolson, both with $\Delta t \sim h$. The total error is optimal for both schemes.

Tables 3.6 and 3.7 report the error values along with their experimental orders of convergence (EOC), which are consistent with the logarithmic linear regression. As expected, the Crank–Nicolson method results in a smaller magnitude of discretization error.

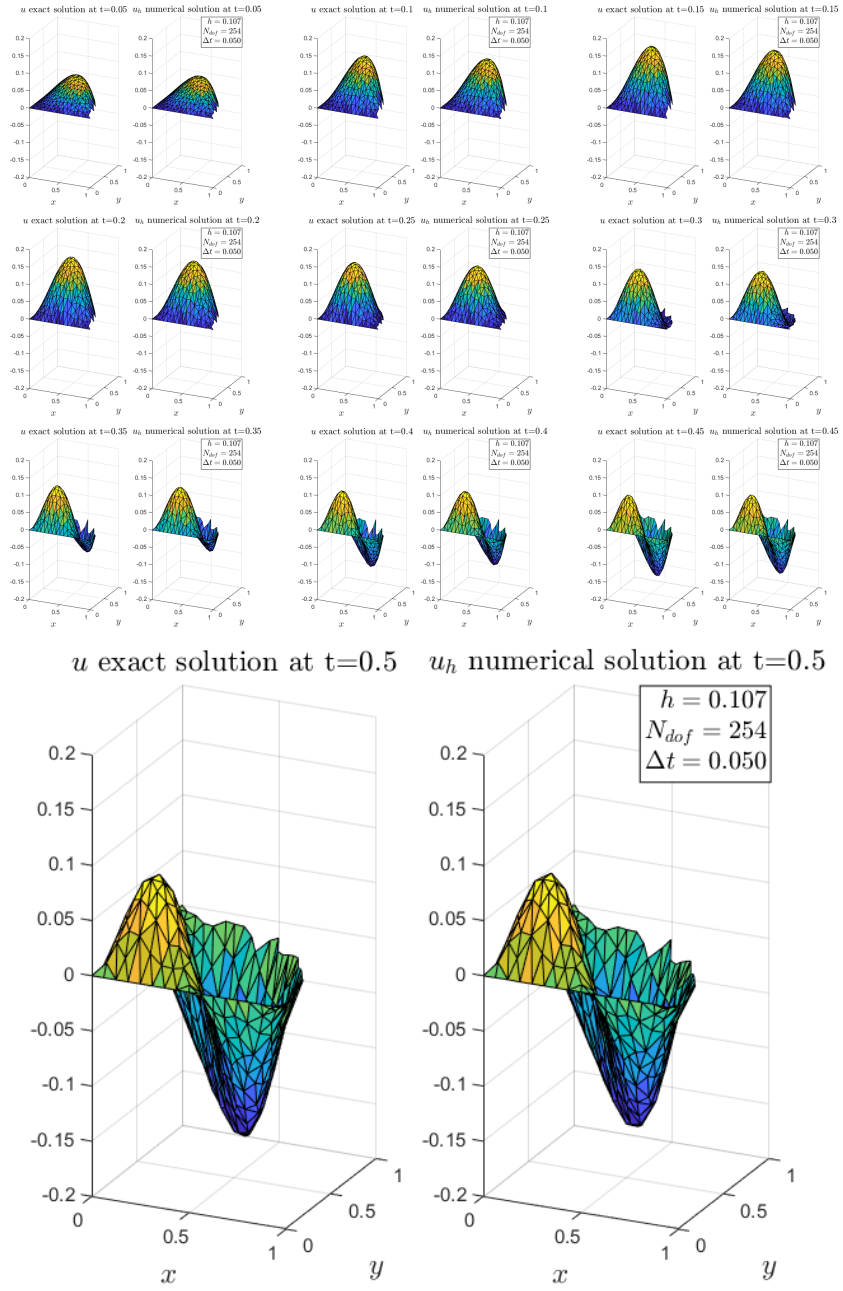


Figure 3.4: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.107, \Delta t = 0.05$ ($u(0) = 0$ is omitted).

Backward Euler $\Delta t \sim h$				
h_{input}	1/8	1/16	1/32	1/64
h	0.2185	0.1196	0.0697	0.0349
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	2.63×10^{-2}	1.15×10^{-2}	6.74×10^{-3}	3.31×10^{-3}
	–	1.28	0.76	1.06
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	1.90×10^{-1}	1.15×10^{-1}	5.33×10^{-2}	2.77×10^{-2}
	–	0.74	1.11	0.96
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	1.59×10^{-2}	6.66×10^{-3}	3.18×10^{-3}	1.52×10^{-3}
	–	1.32	1.06	1.06
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	1.12×10^{-1}	6.43×10^{-2}	3.02×10^{-2}	1.52×10^{-2}
	–	0.81	1.10	1.00
Total error	1.38×10^{-1}	7.58×10^{-2}	3.69×10^{-2}	1.85×10^{-2}
	–	0.87	1.05	1.00

Table 3.6: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Errors for the Backward Euler scheme and successive mesh refinements.

Crank–Nicolson $\Delta t \sim h$				
h_{input}	1/8	1/16	1/32	1/64
h	0.2185	0.1196	0.0697	0.0349
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	2.39×10^{-2}	8.93×10^{-3}	2.35×10^{-3}	5.42×10^{-4}
	–	1.49	2.01	2.02
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	1.91×10^{-1}	1.10×10^{-1}	4.63×10^{-2}	2.45×10^{-2}
	–	0.83	1.28	0.91
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	1.37×10^{-2}	4.46×10^{-3}	1.13×10^{-3}	2.51×10^{-4}
	–	1.74	2.00	2.05
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	1.04×10^{-1}	5.80×10^{-2}	2.54×10^{-2}	1.28×10^{-2}
	–	0.88	1.17	0.99
Total error	1.28×10^{-1}	6.70×10^{-2}	2.77×10^{-2}	1.33×10^{-2}
	–	0.97	1.28	1.06

Table 3.7: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Errors for the Crank–Nicolson scheme and successive mesh refinements.

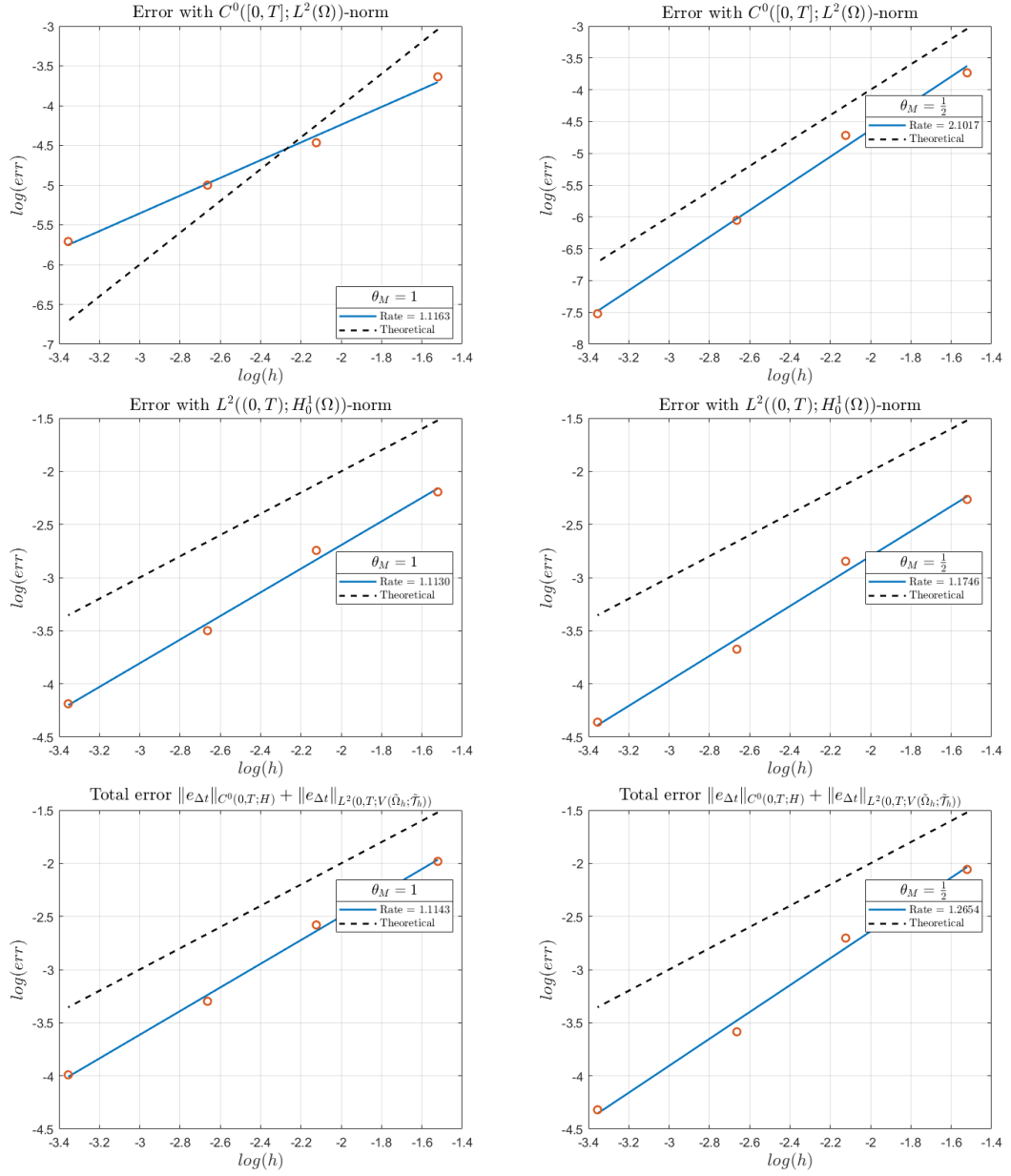


Figure 3.5: [Test 1 with $\alpha = 1, \text{Pe} = 5, k = 1$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the right side ($\Delta t \sim h$).

Test 2: $\alpha = 5, \text{Pe} = 10, k = 3$

As in the previous test, numerical and exact solutions are compared in Figure 3.6 for all time steps, showing no significant differences.

To ensure we are in the asymptotic regime, the mesh refinements start from $h_{\text{input}} = 1/16$. Discretization errors decrease in the last refinements, which is also reflected in a superlinear convergence rate for $\|\cdot\|_{C^0([0,T];L^2(\tilde{\Omega}_h))}$ using Backward Euler, as shown in Figure 3.7.

All error values are summarized in Tables 3.8 and 3.9, computed using Backward Euler and Crank–Nicolson, respectively. The observed convergence rates are consistent with the theoretical results discussed previously.

Backward Euler $\Delta t \sim h$				
h_{input}	1/16	1/32	1/64	1/128
h	0.1196	0.0697	0.0349	0.0170
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	1.86×10^{-2}	7.64×10^{-3}	1.93×10^{-3}	5.90×10^{-4}
	–	1.65	1.99	1.65
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	1.99×10^{-1}	1.09×10^{-1}	5.79×10^{-2}	2.94×10^{-2}
	–	1.12	0.91	0.94
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	9.88×10^{-3}	3.84×10^{-3}	9.59×10^{-4}	2.90×10^{-4}
	–	1.75	2.01	1.67
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	1.12×10^{-1}	5.91×10^{-2}	2.99×10^{-2}	1.49×10^{-2}
	–	1.18	0.99	0.97
Total error	1.31×10^{-1}	6.67×10^{-2}	3.18×10^{-2}	1.55×10^{-2}
	–	1.25	1.07	1.00

Table 3.8: [Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Errors for the Backward Euler scheme and successive mesh refinements.

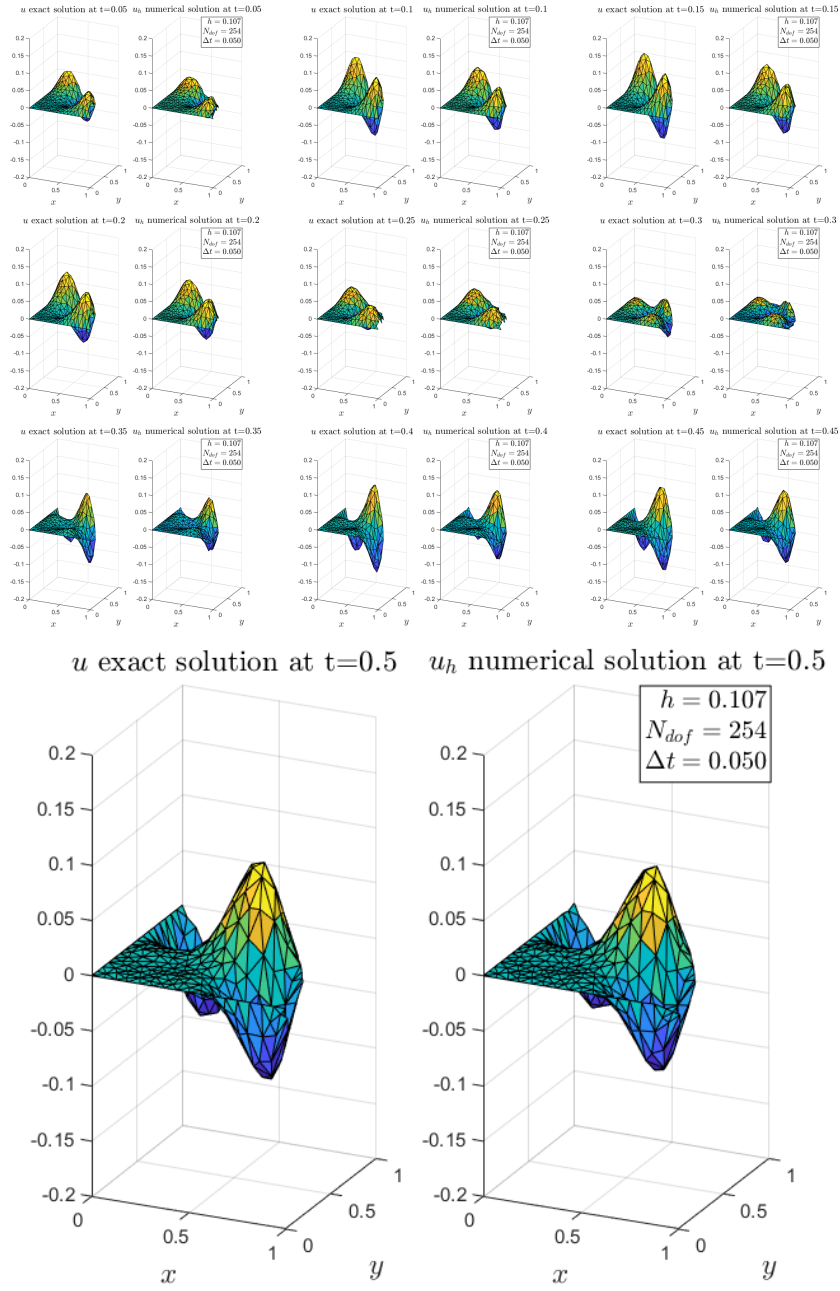


Figure 3.6: [Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Comparison at each time step between the exact solution u and the numerical solution u_h using Backward Euler for $h = 0.107, \Delta t = 0.05$ ($u(0) = 0$ is omitted).

Crank–Nicolson $\Delta t \sim h$				
h_{input}	1/16	1/32	1/64	1/128
h	0.1196	0.0697	0.0349	0.0170
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h))}$	1.76×10^{-2} –	7.13×10^{-3} 1.67	1.52×10^{-3} 2.23	3.70×10^{-4} 1.96
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h))}$	1.96×10^{-1} –	1.08×10^{-1} 1.10	5.77×10^{-2} 0.91	2.94×10^{-2} 0.94
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h))}$	9.64×10^{-3} –	3.56×10^{-3} 1.84	7.60×10^{-4} 2.23	1.79×10^{-4} 2.01
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h))}$	1.11×10^{-1} –	5.83×10^{-2} 1.19	2.96×10^{-2} 0.98	1.48×10^{-2} 0.96
Total error	1.29×10^{-1} –	6.55×10^{-2} 1.26	3.12×10^{-2} 1.07	1.52×10^{-2} 1.00

Table 3.9: [Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Errors for the Crank–Nicolson scheme and successive mesh refinements.

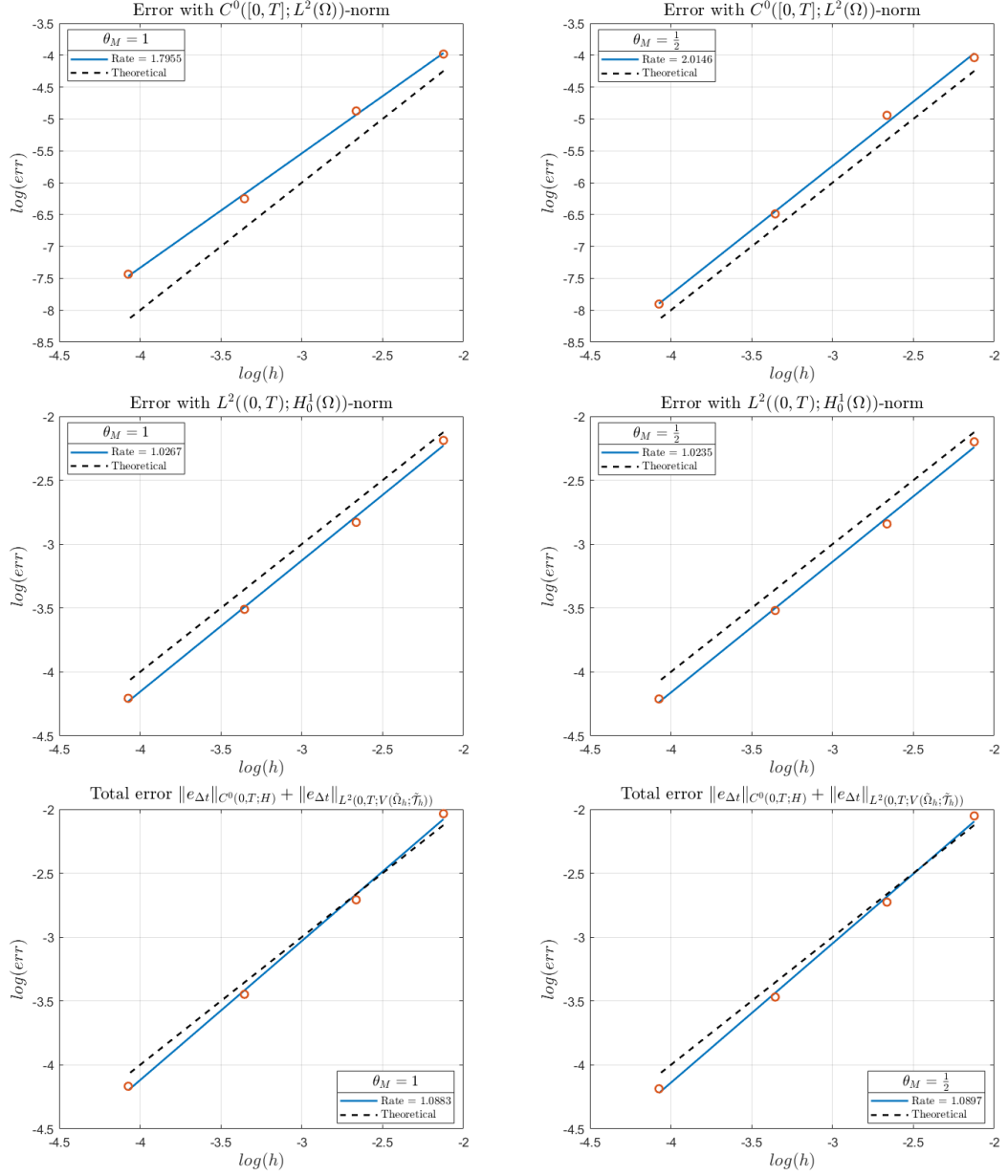


Figure 3.7: [Test 2 with $\alpha = 5, \text{Pe} = 10, k = 3$] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines. Comparison between Backward Euler on the left side ($\Delta t \sim h$) and Crank–Nicolson on the left side ($\Delta t \sim h$).

Chapter 4

Moving Domain problems

The analysis of the *Shifted Boundary Method* is extended to the solution of partial differential equations posed on moving domains. In classical fitted methods, the computational mesh is required to follow the evolution of the physical domain. As a consequence, during the time-stepping procedure the mesh may progressively distort and deform, leading to a deterioration of its quality. To overcome this issue, frequent remeshing of the domain is required, which significantly increases the computational cost.

By contrast, the *Shifted Boundary Method* allows the use of a single background mesh that covers the entire trajectory of the moving domain. At each time step, the surrogate domain is defined by suitably labeling and selecting a subset of the elements of this fixed mesh. In this way, mesh motion and remeshing procedures are completely avoided, resulting in a substantial reduction of the overall computational cost.

To solve the PDE on a moving domain, we adopt the *Arbitrary Lagrangian Eulerian* (ALE) [10] approach. In *Lagrangian* algorithms, each individual node of the computational mesh follows the associated material particle during motion (see Figure 4.1). These algorithms are mainly used in structural mechanics. The *Lagrangian* description allows an easy tracking of free surfaces and interfaces between different materials. It also facilitates the treatment of materials with history-dependent constitutive relations. Its weakness is its inability to follow large distortions of the computational domain without recourse to frequent remeshing operations. *Eulerian* algorithms are widely used in fluid dynamics. Here, as shown in Figure 4.1, the computational mesh is fixed and the continuum moves with respect to the grid. In the *Eulerian* description, large distortions in the continuum motion can be handled with relative ease, but generally at the expense of precise interface definition and the resolution of flow details. Because of the shortcomings of purely *Lagrangian* and purely *Eulerian* descriptions, the ALE technique has been developed that succeeds, to a certain extent, in combining the best features

of both the *Lagrangian* and the *Eulerian* approaches. As announced before, with the ALE approach remeshing is still needed when the quality is decreased under a certain tolerance, nevertheless the SBM solve this issue. In the end, the adjective *Arbitrary* stands for the freedom in moving the computational mesh.

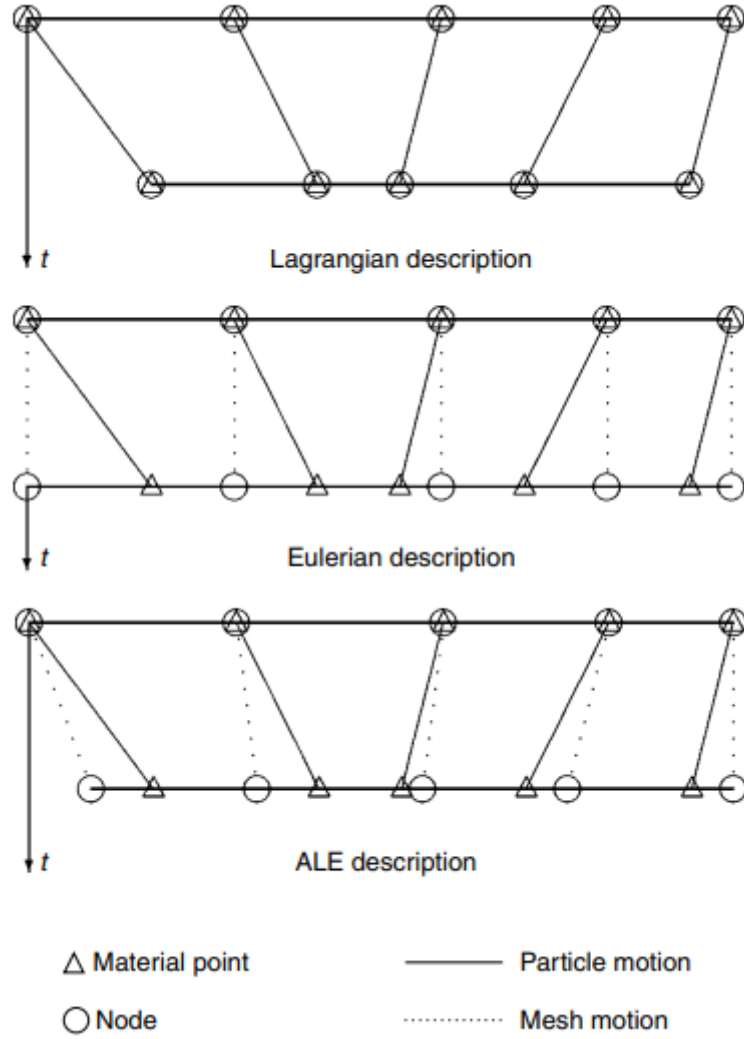


Figure 4.1: One-dimensional example of *Lagrangian*, *Eulerian* and ALE mesh and particle motion. Reproduced from [10].

4.1 ALE Formulation of the heat equation

Let $\Omega^t \subset \mathbb{R}^d$ with $d = 2, 3$ be a moving or deformable domain (open, bounded and connected) with uniformly Lipschitz boundary $\partial\Omega^t$ that evolves in time $t \in [0, T]$, with $0 < T < \infty$. The motion is prescribed or passive and is dictated by a (*Lagrangian*) velocity field $\hat{\mathbf{w}} : \Omega^0 \times [0, T] \rightarrow \mathbb{R}^d$ of class at least $W^{1,\infty}$ defined in the reference or initial domain Ω^0 (to simplify we write only Ω^t , but we should write $\bar{\Omega}^t$ considering also the boundary). This induces a trajectory $\mathbb{x}(\mathbf{x}_0, t)$ emanating from $\mathbf{x}_0 \in \Omega^0$ (*Lagrangian* coordinate) given by

$$\mathbb{x}(\mathbf{x}_0, t) := \mathbf{x}_0 + \int_0^t \hat{\mathbf{w}}(\mathbf{x}_0, s) \, ds$$

also called *forward characteristics*, which in turn solves the initial value problem for an ODE

$$\begin{cases} \frac{d\mathbb{x}}{dt}(\mathbf{x}_0, t) = \hat{\mathbf{w}}(\mathbf{x}_0, t) & t \in (0, T], \\ \mathbb{x}(\mathbf{x}_0, 0) = \mathbf{x}_0, \end{cases} \quad (4.1)$$

and yields $\Omega^t := \{\mathbb{x}(\mathbf{x}_0, t) : \mathbf{x}_0 \in \Omega^0\} = \mathbb{x}(\Omega^0, t)$. We point out that $\mathbb{x}(\mathbf{x}_0, t)$ inherits the space regularity of $\hat{\mathbf{w}}(\mathbf{x}_0, t)$, whence Ω^t is Lipschitz for all $t \in [0, T]$. Hence this guarantees uniqueness of (4.1) so that trajectories do not cross; for every point $\mathbf{x} \in \Omega^t$ (*Eulerian* coordinate) there exists a unique $\mathbf{x}_0 \in \Omega^0$ such that $\mathbf{x} = \mathbb{x}(\mathbf{x}_0, t)$. Moreover the trajectories are invertible such that $\mathbf{x}_0 = \mathbb{x}^{-1}(\mathbf{x}, t)$. We can thus define the *Eulerian* counterpart $\mathbf{w}$ of $\hat{\mathbf{w}}$ as follows

$$\mathbf{w}(\mathbf{x}, t) := \hat{\mathbf{w}}(\mathbb{x}^{-1}(\mathbf{x}, t), t) = \hat{\mathbf{w}}(\mathbf{x}_0, t) \quad \text{if } \mathbf{x} = \mathbb{x}(\mathbf{x}_0, t) \in \Omega^t. \quad (4.2)$$

Moreover, we could start the flow (4.1) at any time $s \in [0, T]$ with initial condition $\mathbf{z}_s = \mathbb{x}(\mathbf{x}_0, s)$ and produce the same outcome $\mathbb{x}(\mathbf{x}_0, t)$ for $t \in [s, T]$ (i.e. (4.1) satisfies a semi-group property). We abbreviate this flow as $\mathcal{A}(\cdot, s, t) : \Omega^s \rightarrow \Omega^t$ and call it the ALE map, namely

$$\mathcal{A}(\mathbf{z}_s, s, t) = \mathbb{x}(\mathbf{x}_0, t) \quad \text{if } \mathbf{z}_s = \mathbb{x}(\mathbf{x}_0, s), \quad 0 \leq s \leq t \leq T.$$

In view of (4.1) and (4.2), we readily deduce

$$\begin{cases} \frac{d\mathcal{A}}{dt}(\mathbf{z}_s, s, t) = \mathbf{w}(\mathcal{A}(\mathbf{z}_s, s, t), t) \\ \mathcal{A}(\mathbf{z}_s, s, s) = \mathbf{z}_s, \end{cases} \quad (4.3)$$

and observe that $\frac{d\mathcal{A}}{dt}(\mathbf{z}_s, s, t)|_{s=t} = \mathbf{w}(\mathbf{z}, s)$ if $\mathbf{z} = \mathcal{A}(\mathbf{z}_s, s, t)$. If $t > s$ then (4.3) describes the flow *forward characteristics*, whereas if $t < s$ then (4.3) corresponds to the flow *backward characteristics*.

Summarizing, $\mathcal{A}(\cdot, s, t)$ establishes an isomorphism between Ω^s and Ω^t such that observing

$$\begin{aligned}\mathcal{A}(\cdot, 0, t) : \Omega^0 &\rightarrow \Omega^t \\ \mathcal{A}(\mathbf{x}_0, 0, t) &= \mathbb{x}(\mathbf{x}_0, t)\end{aligned}\tag{4.4}$$

and

$$\begin{aligned}\mathcal{A}^{-1}(\cdot, 0, t) : \Omega^t &\rightarrow \Omega^0 \\ \mathcal{A}^{-1}(\mathbf{x}, 0, t) &= \mathbb{x}^{-1}(\mathbf{x}, t)\end{aligned}$$

we have

$$\mathcal{A}(\cdot, s, t) = \mathcal{A}(\cdot, 0, t) \circ \mathcal{A}^{-1}(\cdot, 0, s).\tag{4.5}$$

We can also recall the *Eulerian* $\mathbf{w}$ as

$$\mathbf{w}(\mathbf{x}, t) = \underbrace{\hat{\mathbf{w}}(\mathcal{A}^{-1}(\mathbf{x}, 0, t), t)}_{\mathbf{x}_0 \in \Omega^0}.\tag{4.6}$$

Remark 14. For practical applications, as described in [18], the velocity $\mathbf{w}$ at $t = 0$, denoted by $\mathbf{w}_b$, is prescribed only on the boundary $\partial\Omega^0$. The term *Arbitrary* also indicates that the extension of $\mathbf{w}$ to the interior of Ω^0 can be constructed using any suitable method. In general, the velocity field is obtained as the solution of the following harmonic partial differential equation.

$$\begin{cases} \Delta \mathbf{w} = \mathbf{0} & \text{in } \Omega^0, \\ \mathbf{w} = \mathbf{w}_b & \text{on } \partial\Omega^0. \end{cases}\tag{4.7}$$

Moreover, denoting by $D\mathcal{A}$ the Jacobian matrix of $\mathcal{A}$ with respect to the spatial variables, we define $J(\mathbf{z}_s, s, t) := \det(D\mathcal{A}(\mathbf{z}_s, s, t))$ to be the Jacobian of the ALE map starting from $\mathbf{z}_s$ at $t = s$ either forward or backward, and note that it satisfies the following initial value problem:

$$\begin{cases} \frac{dJ}{dt}(\mathbf{z}_s, s, t) = \nabla \cdot \mathbf{w}(\mathcal{A}(\mathbf{z}_s, s, t), t) J(\mathbf{z}_s, s, t) \\ J(\mathbf{z}_s, s, s) = 1 \end{cases}\tag{4.8}$$

and corresponding classical Liouville's formula

$$J(\mathbf{z}_s, s, t) = \exp\left(\int_s^t \nabla \cdot \mathbf{w}(\mathcal{A}(\mathbf{z}_s, s, s'), s') ds'\right).\tag{4.9}$$

We next denote by Q^T the non-cylindrical space-time domain

$$Q^T := \left\{(\mathbf{x}, t) \in \mathbb{R}^d \times [0, T] : \mathbf{x} = \mathcal{A}(\mathbf{y}_0, 0, t), \mathbf{y}_0 \in \Omega^0\right\} = \bigcup_{t \in [0, T]} \Omega^t \times \{t\},\tag{4.10}$$

and define the associated Lebesgue and Sobolev spaces to be

$$L^2(Q^T) := \left\{ v : Q^T \rightarrow \mathbb{R}^d : \int_0^T \|v(\cdot, t)\|_{L^2(\Omega^t)}^2 dt < \infty \right\}, \quad (4.11)$$

$$H^1(Q^T) := \left\{ v \in L^2(Q^T) : \int_0^T \left(\left\| \frac{\partial v}{\partial t}(\cdot, t) \right\|_{L^2(\Omega^t)}^2 + \|\nabla v(\cdot, t)\|_{L^2(\Omega^t)}^2 \right) dt < \infty \right\}. \quad (4.12)$$

Note that, for a stationary domain defined as $\Omega^t = \Omega$, $\forall t \in [0, T]$, as considered in Section 3.1, Q^T is a cylindrical space-time domain such that, according to (4.10), it reduces to $\Omega \times [0, T]$ (see Figure 4.2). In this case, recalling the definition of Bochner spaces given in Definition 2, (4.11) can be equivalently denoted as $L^2(0, T; L^2(\Omega))$ and (4.12) as $L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$.

For the sake of simplicity, we extend Definition 2 to the moving domain setting. Given a normed space X , for any $w(\cdot, t) \in X(\Omega^t)$, we define

$$\|w\|_{L^p(0, T; X(\Omega^t))} = \left(\int_0^T \|w(t)\|_{X(\Omega^t)}^p dt \right)^{1/p}, \quad (4.13)$$

and

$$\|w\|_{C^0([0, T]; X(\Omega^t))} = \max_{0 \leq t \leq T} \|w(t)\|_{X(\Omega^t)}. \quad (4.14)$$

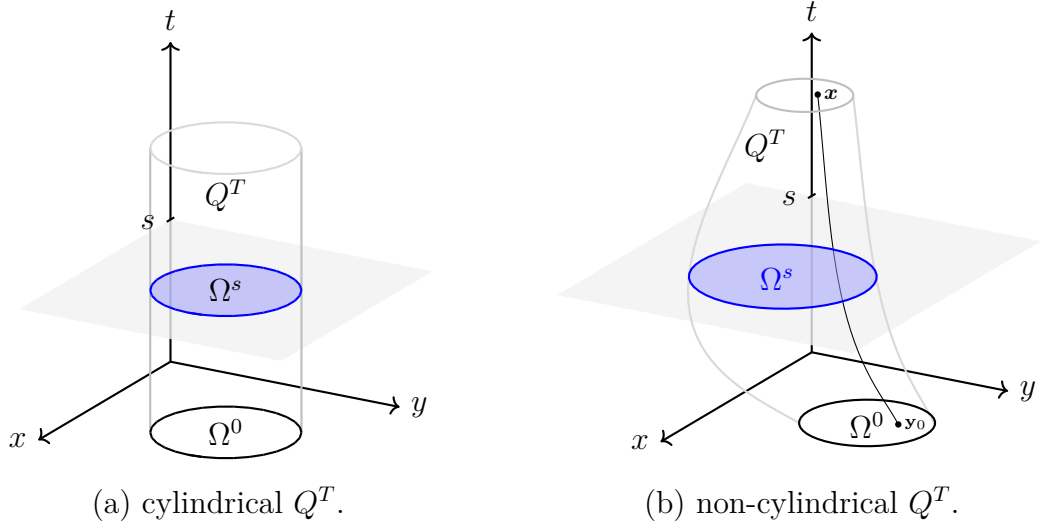


Figure 4.2: Q^T comparison between stationary and moving domains.

We finally introduce the *material derivative* of $v \in H^1(Q^T)$ by setting for any $(\mathbf{x}, t) \in Q^T$

$$D_t v(\mathbf{x}, t) := \frac{d}{dt} v(\mathcal{A}(\mathbf{y}_0, 0, t), t) \quad \text{if } \mathbf{x} = \mathcal{A}(\mathbf{y}_0, 0, t).$$

The chain rule yields

$$\begin{aligned} \frac{dv}{dt}(\mathcal{A}(\mathbf{y}_0, 0, t), t) &= \nabla v(\mathcal{A}(\mathbf{y}_0, 0, t), t) \cdot \frac{d\mathcal{A}}{dt}(\mathbf{y}_0, 0, t) + \frac{\partial v}{\partial t}(\mathcal{A}(\mathbf{y}_0, 0, t), t) \\ &= \nabla v(\mathbf{x}, t) \cdot \mathbf{w}(\mathbf{x}, t) + \frac{\partial v}{\partial t}(\mathbf{x}, t) \end{aligned}$$

whence

$$D_t v = \frac{\partial v}{\partial t} + \mathbf{w} \cdot \nabla v.$$

or, equivalently,

$$\frac{\partial v}{\partial t} = D_t v - \mathbf{w} \cdot \nabla v. \quad (4.15)$$

This yields the following inequality which holds for $0 \leq s < t \leq T$ and any $v \in H^1(Q^T)$, if $\mathbf{y} = \mathcal{A}(\mathbf{y}_0, 0, \tau)$ and $\mathbf{y}_0 \in \Omega^0$.

$$\begin{aligned} &\|v(\mathcal{A}(\mathbf{y}_0, 0, s), s) - v(\mathcal{A}(\mathbf{y}_0, 0, t), t)\|_{L^2(\Omega^0)} \\ &= \left\| \int_s^t \frac{dv}{dt}(\mathcal{A}(\mathbf{y}_0, 0, \tau), \tau) d\tau \right\|_{L^2(\Omega^0)} = \left\| \int_s^t D_\tau v(\mathbf{y}, \tau) d\tau \right\|_{L^2(\Omega^0)} \\ &\leq \int_s^t \|D_\tau v(\mathbf{y}, \tau)\|_{L^2(\Omega^0)} d\tau \\ &\leq \|J^{-1}(\mathbf{y}_0, 0, \tau)\|_{L^\infty(\Omega^0)} \int_s^t \|D_\tau v(\mathbf{y}, \tau)\|_{L^2(\Omega^\tau)} d\tau \\ &\leq C_J \int_s^t \left\| \frac{\partial v}{\partial \tau}(\mathbf{y}, \tau) \right\|_{L^2(\Omega^\tau)} + \|\nabla v(\mathbf{y}, \tau)\|_{L^2(\Omega^\tau)} \|\mathbf{w}(\mathbf{y}, \tau)\|_{L^\infty(\Omega^\tau)} d\tau \end{aligned} \quad (4.16)$$

where $C_J > 0$ is a bound for $\|J^{-1}(\mathbf{y}_0, 0, \tau)\|_{L^\infty(\Omega^0)}$, which naturally arises from the inverse change of volume measure $d\mathbf{y}_0 = J^{-1}(\mathbf{y}_0, 0, \tau) d\mathbf{y}$ associated with the ALE mapping $\mathbf{y} = \mathcal{A}(\mathbf{y}_0, 0, \tau)$. This in turn implies that $v(\mathcal{A}(\mathbf{y}_0, 0, t), t)$ is uniformly continuous in $[0, T]$ with values in $L^2(\Omega^0)$, for all $v \in H^1(Q^T)$.

Furthermore, we recall the well-known *Reynolds' identity* that will be useful during the analysis.

Property 5. *Considering previous assumptions, the function $v \in H^1(Q^T)$ satisfies*

$$\frac{d}{dt} \int_{\Omega^t} v = \int_{\Omega^t} \frac{\partial v}{\partial t} + \nabla \cdot (\mathbf{w} v) = \int_{\Omega^t} D_t v + (\nabla \cdot \mathbf{w}) v. \quad (4.17)$$

4.2 The SBM formulation with ALE map

Considering $\Gamma^t := \partial\Omega^t$, we are interested in solving the heat equation for $t \in [0, T]$:

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f \quad \text{in } \Omega^t, \quad 0 < t \leq T, \\ u &= g \quad \text{on } \Gamma^t, \quad 0 < t \leq T, \\ u &= u_0 \quad \text{in } \Omega^0, \quad t = 0, \end{aligned} \quad (4.18)$$

with data $f \in L^2(Q^T)$, $g \in L^2(0, T; H^{1/2}(\Gamma^t))$, $u_0 \in L^2(\Omega^0)$. Consequently, using (4.15), the first equation in (4.18) can be written equivalently in non-conservative form as follows:

$$D_t u - \Delta u - \mathbf{w} \cdot \nabla u = f \quad \text{in } \Omega^t, \quad 0 < t \leq T. \quad (4.19)$$

In view of the application of the finite element method, we consider a closed domain $\mathcal{D}$ such that

$$\text{clos}(\Omega^t) \subseteq \mathcal{D} \quad \forall t \in [0, T] \quad (4.20)$$

and we introduce a family $\mathcal{T}_h$ of admissible and shape-regular triangulations of $\mathcal{D}$. Then, for $t \in [0, T]$, we restrict each triangulation by selecting those elements that are contained in $\text{clos}(\Omega^t)$, i.e., we form

$$\tilde{\mathcal{T}}_h^t := \{E \in \mathcal{T}_h : E \subset \text{clos}(\Omega^t)\} \quad \forall t \in [0, T]. \quad (4.21)$$

This identifies the *surrogate moving domain*

$$\tilde{\Omega}_h^t := \text{int}\left(\bigcup_{E \in \tilde{\mathcal{T}}_h^t} E\right) \subseteq \Omega^t, \quad (4.22)$$

with *surrogate moving boundary* $\tilde{\Gamma}_h^t := \partial\tilde{\Omega}_h^t$ and outward-oriented unit normal vector $\tilde{\mathbf{n}}_t$ to $\tilde{\Gamma}_h^t$. Obviously, $\tilde{\mathcal{T}}_h^t$ is an admissible and shape-regular triangulation of $\tilde{\Omega}_h^t$. As we have done in previous chapters, we want to discretize problem (4.18) in $\tilde{\Omega}_h^t$ rather than in Ω^t . If we knew the value of the exact solution u on $\tilde{\Gamma}_h^t$, say $u|_{\tilde{\Gamma}_h^t} = \tilde{g}$, we could consider the problem

$$\begin{aligned} D_t u - \Delta u - \mathbf{w} \cdot \nabla u &= f \quad \text{in } \tilde{\Omega}_h^t, \quad 0 < t \leq T, \\ u &= \tilde{g} \quad \text{on } \tilde{\Gamma}_h^t, \quad 0 < t \leq T, \\ u &= u_0 \quad \text{in } \tilde{\Omega}_h^0, \quad t = 0, \end{aligned} \quad (4.23)$$

and discretize it by a Galerkin method using finite element trial and test functions on the triangulation $\tilde{\mathcal{T}}_h^t$ for all $t \in [0, T]$. In particular, exploiting Nitsche's approach in problem (4.19), we get

$$\begin{aligned} (D_t u, v)_{0, \tilde{\Omega}_h^t} - (\mathbf{w} \cdot \nabla u, v)_{0, \tilde{\Omega}_h^t} + (\nabla u, \nabla v)_{0, \tilde{\Omega}_h^t} - (\partial_{\tilde{\mathbf{n}}_t} u, v)_{0, \tilde{\Gamma}_h^t} - (f, v)_{0, \tilde{\Omega}_h^t} \\ - (u - \tilde{g}, \partial_{\tilde{\mathbf{n}}_t} v)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1}(u - \tilde{g}), v)_{0, \tilde{\Gamma}_h^t} = 0, \end{aligned} \quad (4.24)$$

which hold for any sufficiently smooth test function v in $\tilde{\Omega}_h^t$, $0 < t \leq T$; here, $\partial_{\tilde{\mathbf{n}}_t} = \tilde{\mathbf{n}}_t \cdot \nabla$ denotes a normal derivative to $\tilde{\Gamma}_h^t$, whereas $\gamma > 0$ is the penalization parameter.

Introducing the finite dimensional subspace of $H^1(\tilde{\Omega}_h^t)$ made of continuous, piecewise affine functions on the triangulation $\tilde{\mathcal{T}}_h^t$,

$$V_h^t = V_h(\tilde{\Omega}_h^t; \tilde{\mathcal{T}}_h^t) := \{v \in H^1(\tilde{\Omega}_h^t) : v|_E \in \mathbb{P}_1(E), \forall E \in \tilde{\mathcal{T}}_h^t\}, \quad (4.25)$$

we could consider the following SBM semi-discretization in space of Problem (4.24):

$$\begin{aligned} \text{Find } u_h(\cdot, t) \in V_h^t \text{ such that } \forall v_h \in V_h^t \\ (D_t u_h, v_h)_{0, \tilde{\Omega}_h^t} - (\mathbf{w} \cdot \nabla u_h, v_h)_{0, \tilde{\Omega}_h^t} + (\nabla u_h, \nabla v_h)_{0, \tilde{\Omega}_h^t} - (\partial_{\tilde{\mathbf{n}}_t} u_h, v_h)_{0, \tilde{\Gamma}_h^t} \\ - (S_h^t u_h, \partial_{\tilde{\mathbf{n}}_t} v_h)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1} S_h^t u_h, S_h^t v_h)_{0, \tilde{\Gamma}_h^t} \\ = (f, v_h)_{0, \tilde{\Omega}_h^t} - (\bar{g}, \partial_{\tilde{\mathbf{n}}_t} v_h)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1} \bar{g}, S_h^t v_h)_{0, \tilde{\Gamma}_h^t}, \quad 0 < t \leq T, \\ u_h(0) = u_{0,h}, \end{aligned} \quad (4.26)$$

where S_h^t is the boundary operator defined in (2.14) for all $t \in [0, T]$ and $\bar{g}$ is defined in (2.11).

We can rewrite in compact form:

$$\begin{aligned} \text{Find } u_h(\cdot, t) \in V_h^t \text{ such that:} \\ d(t; u_h, v_h) - c(t; u_h, v_h) + a_h(t; u_h, v_h) = F_h(t; v_h) \quad \forall v_h \in V_h^t, \quad 0 < t \leq T, \\ u_h(0) = u_{0,h}, \end{aligned} \quad (4.27)$$

where

$$d(t; u, v) := (D_t u, v)_{0, \tilde{\Omega}_h^t}, \quad (4.28)$$

$$c(t; u, v) := (\mathbf{w} \cdot \nabla u, v)_{0, \tilde{\Omega}_h^t}, \quad (4.29)$$

$$a_h(t; u, v) := (\nabla u, \nabla v)_{0, \tilde{\Omega}_h^t} - (\partial_{\tilde{\mathbf{n}}_t} u, v)_{0, \tilde{\Gamma}_h^t} - (S_h^t u, \partial_{\tilde{\mathbf{n}}_t} v)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1} S_h^t u, S_h^t v)_{0, \tilde{\Gamma}_h^t}, \quad (4.30)$$

$$F_h(t; v) := (f, v)_{0, \tilde{\Omega}_h^t} - (\bar{g}, \partial_{\tilde{\mathbf{n}}_t} v)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1} \bar{g}, S_h^t v)_{0, \tilde{\Gamma}_h^t}. \quad (4.31)$$

Note that, for all $t \in [0, T]$, the properties of the time-dependent forms $a_h(t; \cdot, \cdot)$ and $F_h(t; \cdot)$, such as coercivity and continuity, are inherited from those of a_h and F_h defined in (2.17) and (2.18), respectively. Hence, given $u, v \in V_h^t$, there exist constants $\alpha > 0$ and $\|a\| > 0$, $\forall t \in [0, T]$

$$a_h(t; u, u) \geq \alpha \|u\|_a^2, \quad (4.32)$$

$$a_h(t; u, v) \leq \|a\| \|u\|_{V(\tilde{\Omega}_h^t; \tilde{\mathcal{T}}_h^t)} \|v\|_{V(\tilde{\Omega}_h^t; \tilde{\mathcal{T}}_h^t)}, \quad (4.33)$$

$$F_h(t; v) \leq \|F_h\|_{(V_h^t)'} \|v\|_a. \quad (4.34)$$

With the relation $\mathfrak{x}(\mathbf{x}_0, t) = \mathcal{A}(\mathbf{x}_0, 0, t)$ from (4.4) at hand, we can reformulate problem (4.27) as a time-dependent diffusion-advection-reaction PDE with variable coefficients on the reference domain $\tilde{\Omega}_h^0 \subset \Omega^0$. In view of (4.16), this rigorously justifies the meaning of the initial condition $u(\cdot, 0) = u_0$. Furthermore, the parabolic nature of (4.19) yields the additional regularity $\Delta u, D_t u \in L^2(Q^T)$ for any solution $u \in H^1(Q^T)$, which is essential for the subsequent error analysis.

4.2.1 Stability of the ALE-SBM

Once we have introduced the discretization of the spacial variables, we are ready to solve the Problem (4.26). In this section the stability estimate of the SBM semi-discrete formulation is presented.

Lemma 5. *Considering $\frac{C_{\tau,a}}{2} \max_{0 \leq t \leq T} \|\mathbf{w}(t) \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \leq \frac{\alpha}{p}$ with p integer and $C_{\tau,a} > 0$ independent on meshsize, Problem (4.26) satisfies*

$$\|u_h(t)\|_{0,\tilde{\Omega}_h^t}^2 + \alpha \int_0^t \|u_h(s)\|_a^2 \leq \Lambda \left(\|u_0\|_{0,\tilde{\Omega}_h^0}^2 + \frac{2}{\alpha} \int_0^t \|F_h\|_{(V_h^s)'}^2 \right) \quad \forall t \in [0, T], \quad (4.35)$$

with $\Lambda = \Lambda(\mathbf{w}, \tilde{\mathbf{n}}_t, t) := \exp \left(\frac{2}{C_{\tau,a}} \int_0^t \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \right) < \infty$.

Proof. Take $v_h = u_h$, observe that

$$c(t; u_h, u_h) = \frac{1}{2} \int_{\tilde{\Omega}_h^t} \mathbf{w} \cdot \nabla(u_h^2) = \frac{1}{2} \int_{\tilde{\Gamma}_h^t} (\mathbf{w} \cdot \tilde{\mathbf{n}}_t) u_h^2 - \frac{1}{2} \int_{\tilde{\Omega}_h^t} (\nabla \cdot \mathbf{w}) u_h^2$$

and rewrite (4.27) as follows

$$\frac{1}{2} \int_{\tilde{\Omega}_h^t} D_t(u_h^2) + \frac{1}{2} \int_{\tilde{\Omega}_h^t} (\nabla \cdot \mathbf{w}) u_h^2 - \frac{1}{2} \int_{\tilde{\Gamma}_h^t} (\mathbf{w} \cdot \tilde{\mathbf{n}}_t) u_h^2 + a_h(u_h, u_h) = F_h(u_h).$$

Applying Reynolds' identity in (4.17) we get

$$\frac{1}{2} \frac{d}{dt} \int_{\tilde{\Omega}_h^t} u_h^2 + a_h(t; u_h, u_h) = F_h(t; u_h) + \frac{1}{2} \int_{\tilde{\Gamma}_h^t} (\mathbf{w} \cdot \tilde{\mathbf{n}}_t) u_h^2.$$

Moreover, invoking the coercivity of a_h in (4.32) and continuity of F_h in (4.34) we have

$$\frac{1}{2} \frac{d}{dt} \|u_h\|_{0,\tilde{\Omega}_h^t}^2 + \alpha \|u_h\|_a^2 \leq \|F_h\|_{(V_h^t)'} \|u_h\|_a + \frac{1}{2} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \|u_h\|_{0,\tilde{\Gamma}_h^t}^2,$$

where the second term of the right-hand side can be bounded using the multiplicative trace inequality and Young inequality

$$\begin{aligned}
 \frac{1}{2} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \|u_h\|_{0,\tilde{\Gamma}_h^t}^2 &\leq \frac{1}{2} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} C_\tau \|u_h\|_{0,\tilde{\Omega}_h^t} \|u_h\|_{1,\tilde{\Omega}_h^t} \\
 &\leq \frac{C_{\tau,a}}{2} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \|u_h\|_{0,\tilde{\Omega}_h^t} \|u_h\|_a \\
 &\leq \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \left(\frac{1}{2\varepsilon} \|u_h\|_{0,\tilde{\Omega}_h^t}^2 + \frac{\varepsilon}{2} \|u_h\|_a^2 \right).
 \end{aligned} \tag{4.36}$$

where $\varepsilon := \frac{C_{\tau,a}}{2}$. Applying Young inequality again, and considering

$$\frac{C_{\tau,a}}{2} \max_{0 \leq t \leq T} \|\mathbf{w}(t) \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \leq \frac{\alpha}{p} \tag{4.37}$$

where p is an integer depending on how many inequalities of the a -norm are involved (here $p = 2$), we get

$$\begin{aligned}
 \frac{1}{2} \frac{d}{dt} \|u_h\|_{0,\tilde{\Omega}_h^t}^2 + \alpha \|u_h\|_a^2 &\leq \frac{2}{2\alpha} \|F_h\|_{(V_h^t)'}^2 + \frac{\alpha}{4} \|u_h\|_a^2 + \frac{\alpha}{4} \|u_h\|_a^2 \\
 &\quad + \frac{1}{C_{\tau,a}} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \|u_h\|_{0,\tilde{\Omega}_h^t}^2, \\
 \frac{d}{dt} \|u_h\|_{0,\tilde{\Omega}_h^t}^2 + \alpha \|u_h\|_a^2 &\leq \frac{2}{\alpha} \|F_h\|_{(V_h^t)'}^2 + \frac{2}{C_{\tau,a}} \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)} \|u_h\|_{0,\tilde{\Omega}_h^t}^2.
 \end{aligned}$$

We now integrate in time from 0 to $t \in [0, T]$ and, applying the Gronwall Lemma with $\Lambda(\mathbf{w}, \tilde{\mathbf{n}}_t, t) := \exp\left(\frac{2}{C_{\tau,a}} \int_0^t \|\mathbf{w} \cdot \tilde{\mathbf{n}}_t\|_{L^\infty(\tilde{\Gamma}_h^t)}\right) < \infty$, we get the result. $\square$

Furthermore, we analyze the term $-\frac{1}{2} \int_{\tilde{\Gamma}_h^t} (\mathbf{w} \cdot \tilde{\mathbf{n}}_t) u_h^2$. Note that if $u_h|_{\tilde{\Gamma}_h^t} = 0$, that is very unlikely, or $\mathbf{w} \cdot \tilde{\mathbf{n}}_t = 0 \implies \mathbf{w} \perp \tilde{\mathbf{n}}_t$ we do not need to apply the Gronwall Lemma and the stability result in Lemma 5 will be ALE-free because it is not sensitive to the ALE velocity $\mathbf{w}$.

- If $\mathbf{w} \cdot \tilde{\mathbf{n}}_t > 0$, the velocity flux points outwards, representing a domain expansion. In this case, the integral contribution is negative, acting as a dissipative term that reduces the overall energy of the solution u .
- If $\mathbf{w} \cdot \tilde{\mathbf{n}}_t < 0$, the velocity flux points inwards, representing a domain contraction. The integral contribution becomes positive, effectively acting as a source term that increases the energy of the solution u .

4.2.2 ALE-free stability

In this section, we will focus on the original problem (4.18) analyzing the basic stability estimate that, in case of homogeneous boundary conditions, is considered ALE-free because it does not depend on the velocity $\mathbf{w}$ of the ALE map. The variational formulation of (4.19), using the ALE map, needs the decomposition $u = u^o + u^g$ for all $t \in [0, T]$, where $u^o \in H_0^1(\Omega^t)$ and the lifting function u^g satisfies $u^g|_{\Gamma^t} = g$. In order to have a lighter notation let us set $u := u^o$.

We are interested to solve the following problem

$$\begin{aligned} \text{Find } u(\cdot, t) \in H_0^1(\Omega^t) \text{ such that } \forall v \in H_0^1(\Omega^t), 0 < t \leq T : \\ (D_t u, v)_{0, \Omega^t} + (\nabla u, \nabla v)_{0, \Omega^t} - (\mathbf{w} \cdot \nabla u, v)_{0, \Omega^t} = (f, v)_{0, \Omega^t} - R_g(u^g, v), \quad (4.38) \\ u(0) = u_0 \end{aligned}$$

where $R_g(u^g, v) := (D_t u^g, v)_{0, \Omega^t} - (\nabla u^g, \nabla v)_{0, \Omega^t} - (\mathbf{w} \cdot \nabla u^g, v)_{0, \Omega^t}$. In order to formulate a rigorous stability estimate we need the following assumptions.

Assumption 5. The lifting $u^g \in H^1(\Omega^t)$ satisfies for all $t \in [0, T]$:

$$\|u^g\|_{1, \Omega^t} \leq C_\nabla \|g\|_{H^{1/2}(\Gamma^t)}, \quad (4.39)$$

$$\|D_t u^g\|_{0, \Omega^t} \leq C_{D_t} \|D_t g\|_{H^{1/2}(\Gamma^t)}. \quad (4.40)$$

We now proceed to exhibit the basic stability estimate of Problem (4.38).

Lemma 6. Problem (4.38) satisfies

$$\|u(t)\|_{0, \Omega^t}^2 + \int_0^t \|\nabla u\|_{0, \Omega^s}^2 \leq \|u_0\|_{0, \Omega^0}^2 + C \int_0^t \|\mathcal{F}(s)\|^2 \quad \forall t \in [0, T], \quad (4.41)$$

where $\|\mathcal{F}(t)\|^2 := \|f\|_{0, \Omega^t}^2 + \|g\|_{H^{1/2}(\Gamma^t)}^2 + \|D_t g\|_{H^{1/2}(\Gamma^t)}^2$.

Proof. Take $v = u$, observe that

$$\int_{\Omega^t} (\mathbf{w} \cdot \nabla u) u = \frac{1}{2} \int_{\Omega^t} \mathbf{w} \cdot \nabla (u^2) = \frac{1}{2} \underbrace{\int_{\Gamma^t} (\mathbf{w} \cdot \tilde{\mathbf{n}}_t) u^2}_{=0} - \frac{1}{2} \int_{\Omega^t} (\nabla \cdot \mathbf{w}) u^2$$

and rewrite (4.27) as follows

$$\frac{1}{2} \int_{\Omega^t} D_t (u^2) + \frac{1}{2} \int_{\Omega^t} (\nabla \cdot \mathbf{w}) u^2 + (\nabla u, \nabla u)_{0, \Omega^t} = (f, u)_{0, \Omega^t} - R_g(u^g, u).$$

Applying Reynolds' identity in (4.17) we get

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega^t} u^2 + \|\nabla u\|_{0, \Omega^t}^2 = (f, u)_{0, \Omega^t} - R_g(u^g, u).$$

Moreover, applying Cauchy-Schwarz and Poincaré inequalities to the first term of the right-hand side, we have

$$\frac{1}{2} \frac{d}{dt} \|u\|_{0,\Omega^t}^2 + \|\nabla u\|_{0,\Omega^t}^2 \leq C_P(\Omega^t) \|f\|_{0,\Omega^t} \|\nabla u\|_{0,\Omega^t} + |R_g(u^g, u)|.$$

Let us focus on the second term of the right-hand side

$$|R_g(u^g, u)| = |(D_t u^g, u)_{0,\Omega^t}| + |(\nabla u^g, \nabla u)_{0,\Omega^t}| + |(\mathbf{w} \cdot \nabla u^g, u)_{0,\Omega^t}|,$$

where each term can be bounded using Cauchy-Schwarz and Poincaré inequalities with (4.39) and (4.40):

- $|(D_t u^g, u)_{0,\Omega^t}| \leq \|D_t u^g\|_{0,\Omega^t} \|u\|_{0,\Omega^t} \leq C_{D_t} \|D_t g\|_{H^{1/2}(\Gamma^t)} C_P(\Omega^t) \|\nabla u\|_{0,\Omega^t},$
- $|(\nabla u^g, \nabla u)_{0,\Omega^t}| \leq \|\nabla u^g\|_{0,\Omega^t} \|\nabla u\|_{0,\Omega^t} \leq C_\nabla \|g\|_{H^{1/2}(\Gamma^t)} \|\nabla u\|_{0,\Omega^t},$
- $|(\mathbf{w} \cdot \nabla u^g, u)_{0,\Omega^t}| \leq \|\mathbf{w}\|_{L^\infty(\Omega^t)} C_\nabla \|g\|_{H^{1/2}(\Gamma^t)} C_P(\Omega^t) \|\nabla u\|_{0,\Omega^t}.$

Summarizing, we have

$$|R_g(u^g, u)| \leq (C_g \|g\|_{H^{1/2}(\Gamma^t)} + C_{\dot{g}} \|D_t g\|_{H^{1/2}(\Gamma^t)}) \|\nabla u\|_{0,\Omega^t}.$$

Applying Young inequality we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u\|_{0,\Omega^t}^2 + \|\nabla u\|_{0,\Omega^t}^2 &\leq \frac{3C_P^2(\Omega^t)}{2} \|f\|_{0,\Omega^t}^2 + \frac{1}{6} \|\nabla u\|_{0,\Omega^t}^2 + \frac{3C_g^2}{2} \|g\|_{H^{1/2}(\Omega^t)}^2 \\ &\quad + \frac{1}{6} \|\nabla u\|_{0,\Omega^t}^2 + \frac{3C_{\dot{g}}^2}{2} \|D_t g\|_{H^{1/2}(\Gamma^t)}^2 + \frac{1}{6} \|\nabla u\|_{0,\Omega^t}^2, \\ \frac{d}{dt} \|u\|_{0,\Omega^t}^2 + \|\nabla u\|_{0,\Omega^t}^2 &\leq C \left(\|f\|_{0,\Omega^t}^2 + \|g\|_{H^{1/2}(\Gamma^t)}^2 + \|D_t g\|_{H^{1/2}(\Gamma^t)}^2 \right). \end{aligned}$$

We now integrate in time from 0 to $t \in [0, T]$ getting the result. $\square$

Remark 15. If in (4.41) we integrate in time from $t = 0$ to $t = \tau$, with arbitrary $\tau \in (0, T]$, we obtain

$$\|u(\tau)\|_{0,\Omega_\tau}^2 + \int_0^\tau \|\nabla u\|_{0,\Omega^t}^2 \leq \|u_0\|_{0,\Omega^0}^2 + C \int_0^\tau \|\mathcal{F}(s)\|^2.$$

Taking $\tau = T$ and neglecting the first term on the left-hand side, we obtain a stability estimate similar to the one featuring the $L^2(0, T; \cdot)$ -norm in (3.17), but tailored to moving domains. Conversely, by taking the supremum over $(0, T]$ and omitting the integral term, we obtain the analogue of the $C^0([0, T]; \cdot)$ -norm estimate in (3.18).

As demonstrated in the proof, the constant C depends on several parameters, including the mesh velocity $\mathbf{w}$. However, in the case of homogeneous boundary conditions, we have $R_g(u^g, v) = 0$. Consequently, the constant C reduces to C_P^2 , where $C_P := \sup_{t \in [0, T]} C_P(\Omega^t)$ denotes the uniform Poincaré constant over the time interval. Under these specific conditions, the stability of (4.38) is shown to be ALE-free, as it no longer explicitly depends on the mesh motion:

$$\|u(t)\|_{0, \Omega^t}^2 + \int_0^t \|\nabla u\|_{0, \Omega^s}^2 \leq \|u_0\|_{0, \Omega^0}^2 + C_P^2 \int_0^t \|f\|_{0, \Omega^s}^2 \quad \forall t \in [0, T].$$

The ideal time discretization discussed next preserves this property, but this is no longer possible with practical time and space discretizations. This is the topic of next section.

4.3 Time semi-discretization

In this section we present and discuss the semi-discretization in time of the ALE formulation of the heat equation with SBM. Conversely to the approach of Chapter 3, we will follow the *traverse line* or *Rothe* method. Indeed, the formulation relies on a time discretization based on the degree-zero discontinuous Galerkin scheme, also denoted as dG(0), coupled with the spatial SBM discretization discussed in Chapter 2. It is worth noting that the dG(0) method is first-order accurate in time, a property that perfectly matches the first-order spatial accuracy derived from the Taylor expansion within the SBM framework. Since we begin with the temporal discretization, we first analyze Problem (4.38) in the time-dependent domain Ω^t . Subsequently, once the spatial discretization is introduced, we shall investigate Problem (4.26) in $\tilde{\Omega}_h^t$.

As announced before, let us introduce the following time discretization. We subdivide the time interval $[0, T]$ into N_T subintervals $I_n := [t_n, t_{n+1}]$ defined by the nodes

$$0 = t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_{N_T} = T.$$

We assume for simplicity that this partition is uniform and denote by $\Delta t = T/N_T$ the time step. It will be convenient later to consider $s = t_n$ and $t \in I_n$ and simplify the notation of $\mathcal{A}$ to read

$$\mathcal{A}_n(\mathbf{y}_n, t) := \mathcal{A}(\mathbf{y}_n, s = t_n, t) \quad \mathbf{y}_n \in \Omega^n, \quad t \in I_n, \quad (4.42)$$

where $\Omega^n := \Omega^{t_n}$. We may regard the trajectory $\mathbf{y} = \mathcal{A}_n(\mathbf{y}_n, t) \in \Omega^t$ backwards, namely emanating from $\mathbf{y}_{n+1} = \mathcal{A}_n(\mathbf{y}_n, t_{n+1}) \in \Omega^{n+1}$ at time t_{n+1} and going along the backward characteristics in time for $t \in I_n$: $\mathbf{y} = \mathcal{A}(\mathbf{y}_{n+1}, t_{n+1}, t)$. This map is well defined because $\mathcal{A}_n(\cdot, t) : \Omega^n \rightarrow \Omega^t$ is bijective for all $t \in I_n$. We set $\mathcal{A}_n(\cdot) := \mathcal{A}_n(\cdot, t_{n+1}) : \Omega^n \rightarrow \Omega^{n+1}$ and denote its inverse by $\mathcal{A}_n^{-1}(\cdot) : \Omega^{n+1} \rightarrow \Omega^n$.

Given a function $v \in H^1(\Omega^{n+1})$, we extend it constant along the backward characteristics and denote it by

$$\overleftarrow{v}(\mathbf{y}, t) := v(\mathbf{y}_{n+1}) = v \circ \mathcal{A}_n(\mathbf{y}_n) \quad \forall \mathbf{y} \in \Omega^t, \quad t \in I_n, \quad (4.43)$$

where $\mathbf{y}_{n+1} = \mathcal{A}^{-1}(\mathbf{y}, t_{n+1}, t) = \mathcal{A}(\mathbf{y}, t, t_{n+1})$ and $\mathbf{y}_n = \mathcal{A}(\mathbf{y}, t, t_n) = \mathcal{A}_n^{-1}(\mathbf{y}_{n+1})$; note that $\overleftarrow{v}(\mathbf{y}_{n+1}, t_{n+1}) = v(\mathbf{y}_{n+1})$ (see Figure 4.3). This particular constant extension will be useful for the following ALE time discrete formulation.

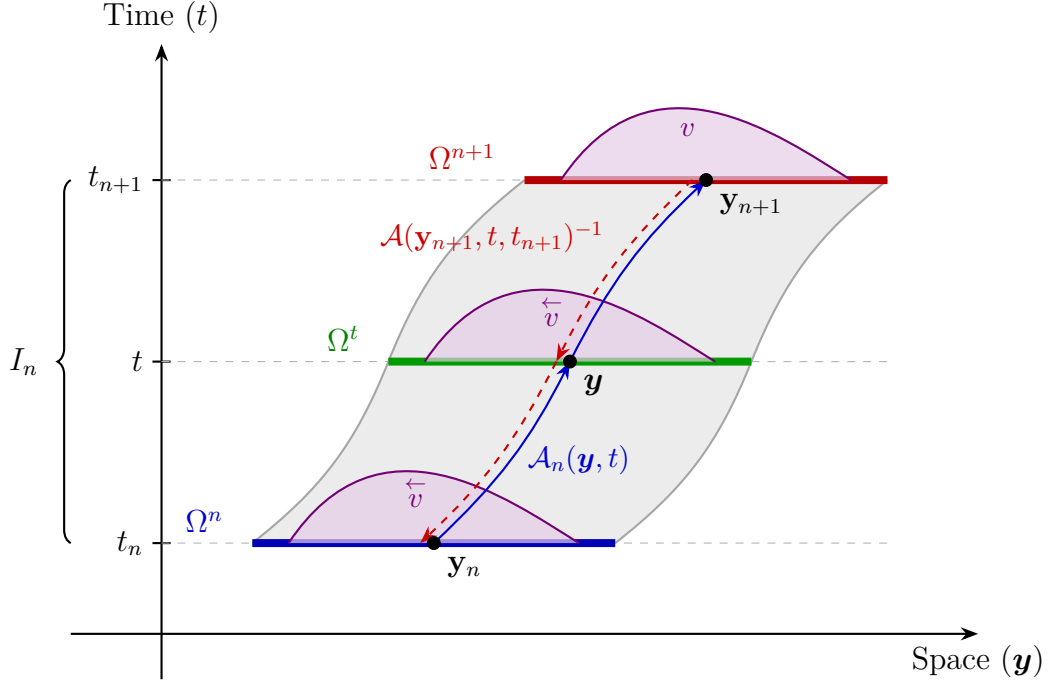


Figure 4.3: Space-time representation of the ALE map. The diagram highlights how a generic function v maintains a constant value along the characteristics, becoming $\overleftarrow{v}$ at previous times.

We now introduce an *ideal* time discretization, in the sense that it preserves the stability properties described in Section 4.2.2. To this end, we integrate (4.38) over I_n using the discontinuous Galerkin (dG) method to approximate the bilinear form $(D_t u, v)_{0, \Omega^t}$. Using the integration by parts formula, we obtain

$$\begin{aligned} \int_{I_n} (D_s u, v)_{0, \Omega^s} ds &= [(u, v)_{\Omega^s}]_{t_n}^{t_{n+1}} - \int_{I_n} (u, D_s v)_{\Omega^s} ds \\ &= (u_-^{n+1}, v_-^{n+1})_{\Omega^{n+1}} - (u_+^n, v_+^n)_{\Omega^n} - \int_{I_n} (u, D_s v)_{\Omega^s} ds \end{aligned}$$

where $\phi^n := \phi(t_n)$, and ϕ_+^n, ϕ_-^n denote the right-hand and left-hand limits of the function ϕ at time t_n , respectively.

Note that in a continuous (classical) formulation, the solution must satisfy $u^n(t_n) = u^{n+1}(t_n)$ for all $0 \leq n < N_T$. Conversely, within the dG framework, discontinuous solutions are allowed. To account for this, we replace the value u_+^n with its left-hand limit u_-^n , which yields

$$(u_-^{n+1}, v_-^{n+1})_{\Omega^{n+1}} - (u_-^n, v_+^n)_{\Omega^n} - \int_{I_n} (u, D_s v)_{\Omega^s} ds.$$

Applying the integration by parts formula once more, we have

$$- \int_{I_n} (u, D_s v)_{\Omega^s} ds = -[(u, v)_{\Omega^s}]_{t_n}^{t_{n+1}} + \int_{I_n} (D_s u, v)_{0, \Omega^s} ds,$$

hence the approximation of $\int_{I_n} (D_s u, v)_{\Omega^s}$ becomes

$$\int_{I_n} (D_s u, v)_{0, \Omega^s} ds + (u_+^n - u_-^n, v_+^n)_{\Omega^n}. \quad (4.44)$$

By considering dG functions that are piecewise constant in time along the characteristics, the first term in (4.44) vanishes. This is equivalent to setting $\overleftarrow{u}^{n+1} := u_+^n$, $u^n := u_-^n$, and $\overleftarrow{v} := v_+^n$.

Ideal time discretization: given $u^n \in H_0^1(\Omega^n)$, find $u^{n+1} \in H_0^1(\Omega^{n+1})$ such that, setting $\overleftarrow{u} = \overleftarrow{u}^{n+1} = u^{n+1} \circ \mathcal{A}_n$, we have for all $v \in H_0^1(\Omega^{n+1})$

$$\begin{aligned} (u^{n+1} \circ \mathcal{A}_n - u^n, v \circ \mathcal{A}_n)_{0, \Omega^n} &= \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{v})_{0, \Omega^s} + \int_{I_n} (\nabla \overleftarrow{u}, \nabla \overleftarrow{v})_{0, \Omega^s} \\ &= \int_{I_n} (f, \overleftarrow{v})_{0, \Omega^s} - \int_{I_n} R_g(u^g, \overleftarrow{v}), \quad 0 < n \leq N_T. \end{aligned} \quad (4.45)$$

Note that all terms are computed along backward characteristics and that the first term, corresponding to the material derivative in (4.38), is evaluated at t_n rather than t_{n+1} . This hinges on the dG approach developed in [5] and is justified in the following lemma. But before we do so, we observe that such a term can be equivalently written at time t_{n+1} via a change of variables that involves the Jacobian $J_n^{-1} := \det(D\mathcal{A}_n)^{-1} : \Omega^n \rightarrow \mathbb{R}$

$$(u^{n+1} \circ \mathcal{A}_n - u^n, v \circ \mathcal{A}_n)_{0, \Omega^n} = (u^{n+1} - u^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v)_{0, \Omega^{n+1}} \quad \forall v \in H^1(\Omega^{n+1}). \quad (4.46)$$

This will be useful later for the actual implementation of the fully discrete method.

Remark 16. It is important to comment on the choice of the test function $v \in H^1(\Omega^{n+1})$, which is assumed to be constant along the characteristics. In the previous chapter, when dealing with fixed (stationary) domains, the test function

was chosen to be constant in time. In the current setting, since we are working with moving domains, the variational formulation is stated along the characteristics. Therefore, to maintain the same approach as the classical parabolic formulation for stationary domains, the test function should be considered constant along the characteristics. Indeed, if we consider a stationary domain as a particular case of a moving domain with straight characteristics, these two choices for the test functions are perfectly consistent.

Lemma 7 (material derivative). *If $\overleftarrow{u}^{n+1} := u^{n+1} \circ \mathcal{A}^n, 0 \leq n < N_T$, the following identity holds*

$$\begin{aligned} (\overleftarrow{u}^{n+1} - u^n, \overleftarrow{u}^{n+1})_{0, \Omega^n} - \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{0, \Omega^s} &= \frac{1}{2} \|u^{n+1}\|_{0, \Omega^{n+1}}^2 - \frac{1}{2} \|u^n\|_{0, \Omega^n}^2 \\ &\quad + \frac{1}{2} \|\overleftarrow{u}^{n+1} - u^n\|_{0, \Omega^n}^2. \end{aligned} \quad (4.47)$$

Proof. The assertion follows from the classical identity

$$(\overleftarrow{u}^{n+1} - u^n, \overleftarrow{u}^{n+1})_{\Omega^n} = \frac{1}{2} \|\overleftarrow{u}^{n+1}\|_{0, \Omega^n}^2 - \frac{1}{2} \|u^n\|_{0, \Omega^n}^2 + \frac{1}{2} \|\overleftarrow{u}^{n+1} - u^n\|_{0, \Omega^n}^2.$$

and the equality

$$- \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{0, \Omega^s} = \frac{1}{2} \|u^{n+1}\|_{0, \Omega^{n+1}}^2 - \frac{1}{2} \|\overleftarrow{u}^{n+1}\|_{0, \Omega^n}^2$$

that leads to cancellation of the problematic term $\frac{1}{2} \|\overleftarrow{u}^{n+1}\|_{0, \Omega^n}^2$. To prove the latter, we observe that

$$\int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{\Omega^s} = \frac{1}{2} \int_{I_n} (\mathbf{w}, \nabla \overleftarrow{u}^2)_{\Omega^s} = -\frac{1}{2} \int_{I_n} (\nabla \cdot \mathbf{w}, \overleftarrow{u}^2)_{\Omega^s}$$

and we use Reynolds' identity (4.17) for $v = \overleftarrow{u}^2$, in conjunction with (4.43) which implies $D_t \overleftarrow{u}^2 = 0$, to obtain

$$\begin{aligned} - \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{\Omega^s} &= \frac{1}{2} \int_{I_n} (\nabla \cdot \mathbf{w}, \overleftarrow{u}^2)_{\Omega^s} = \frac{1}{2} \int_{I_n} \frac{d}{dt} \|\overleftarrow{u}(t)\|_{0, \Omega^s}^2 \\ &= \frac{1}{2} \|u^{n+1}\|_{0, \Omega^{n+1}}^2 - \frac{1}{2} \|\overleftarrow{u}^{n+1}\|_{0, \Omega^n}^2 \end{aligned}$$

whence the result. $\square$

Since the remaining terms in (4.45) are similar to those in (4.38), proceeding exactly as in Lemma 6 we can derive the following result, whose proof is trivial.

Lemma 8 (stability of ideal time discretization). *There exists a constant $C > 0$ independent of the semidiscrete solution $\overleftarrow{u} \in H_0^1(\Omega^t)$ such that*

$$\|u^n\|_{0,\Omega^n}^2 + \int_0^{t_n} \|\nabla \overleftarrow{u}\|_{0,\Omega^s}^2 \leq \|u_0\|_{0,\Omega^0}^2 + C \int_0^{t_n} \|\mathcal{F}(s)\|^2, \quad (4.48)$$

for all $0 < n \leq N_T$. The constant C is the same of (4.41) that, as we explained before, could still be ALE-free.

Since the ideal scheme (4.45) is not practical, we now consider a *practical* variant of it that incorporates quadrature to approximate the time integral $\int_{I_n}$. Given a data $\phi = \phi(\mathbf{x}, t)$, we denote by $\phi^n = \phi^n(\mathbf{x}) := \phi(\mathbf{x}, t_n)$ its restriction to Ω^n .

Practical time discretization: given $u^n \in H_0^1(\Omega^n)$, find $u^{n+1} \in H_0^1(\Omega^{n+1})$ such that for all $v \in H_0^1(\Omega^{n+1})$

$$\begin{aligned} (u^{n+1} \circ \mathcal{A}_n - u^n, v \circ \mathcal{A}_n)_{0,\Omega^n} - \Delta t (\mathbf{w}^{n+1} \cdot \nabla u^{n+1}, v)_{0,\Omega^{n+1}} + \Delta t (\nabla u^{n+1}, \nabla v)_{0,\Omega^{n+1}} \\ = \Delta t (f^{n+1}, v)_{0,\Omega^{n+1}} - \Delta t R_g^{n+1}(u^g, v), \quad 0 < n \leq N_T, \end{aligned} \quad (4.49)$$

with

$$\begin{aligned} R_g^{n+1}(u^g, v) &:= (D_t u^g(t_{n+1}), v)_{0,\Omega^{n+1}} - (\nabla u^g(t_{n+1}), \nabla v)_{0,\Omega^{n+1}} \\ &\quad - (\mathbf{w}^{n+1} \cdot \nabla u^g(t_{n+1}), v)_{0,\Omega^{n+1}}. \end{aligned}$$

We can rewrite this scheme as follows to exploit the results of Lemma 10:

$$\begin{aligned} (u^{n+1} \circ \mathcal{A}_n - u^n, v \circ \mathcal{A}_n)_{0,\Omega^n} - \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{0,\Omega^s} + \Delta t (\nabla u^{n+1}, \nabla v)_{0,\Omega^{n+1}} \\ = \Delta t (f^{n+1}, v)_{0,\Omega^{n+1}} - \Delta t R_g^{n+1}(u^g, v) \\ + Q^{n+1}(\overleftarrow{u}, \overleftarrow{v}), \quad 0 < n \leq N_T. \end{aligned} \quad (4.50)$$

where $Q^{n+1}(\overleftarrow{u}, \overleftarrow{v})$ stands for the quadrature rule error

$$Q^{n+1}(\overleftarrow{u}, \overleftarrow{v}) := - \int_{I_n} (\mathbf{w} \cdot \nabla \overleftarrow{u}, \overleftarrow{u})_{0,\Omega^s} + \Delta t (\mathbf{w}^{n+1} \cdot \nabla u^{n+1}, v)_{0,\Omega^{n+1}}. \quad (4.51)$$

The stability properties of (4.50) are similar to those of (4.45) upon taking $v = u^{n+1} \in H_0^1(\Omega^{n+1})$, except for the new term Q^{n+1} in (4.51). Since estimating Q^{n+1} requires dealing with $t \in I_n$, we simplify the notation slightly upon setting for all $\mathbf{y}_{n+1} \in \Omega^{n+1}$.

$$\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t) := \mathcal{A}^{-1}(\mathbf{y}_{n+1}, t, t_{n+1}) = \mathcal{A}(\mathbf{y}_{n+1}, t_{n+1}, t), \quad (4.52)$$

$$J_{n+1}^-(\mathbf{y}_{n+1}, t) := J(\mathbf{y}_{n+1}, t_{n+1}, t) = \exp \left(\int_{t_{n+1}}^t \nabla \cdot \mathbf{w}(\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, s), s) \right), \quad (4.53)$$

and for all $\mathbf{y} = \mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t)$

$$\omega(\mathbf{y}, t) := \nabla \cdot \mathbf{w}(\mathbf{y}, t).$$

Lemma 9 (temporal quadrature error). *For $0 \leq n < N_T$, the following inequality holds*

$$|Q^{n+1}(\overset{\leftarrow}{u}, \overset{\leftarrow}{u})| \leq \frac{\Delta t}{2} C_n \|u^{n+1}\|_{0, \Omega^{n+1}}^2 \quad (4.54)$$

provided $\exp\left(\int_{I_n} \|\omega\|_{L^\infty(\Omega^s)}\right) \leq M_J$, where

$$C_n := M_J \int_{I_n} \left(\left\| \frac{\partial \omega}{\partial s} \right\|_{L^\infty(\Omega^s)} + \|\nabla \omega\|_{L^\infty(\Omega^s)} \|\mathbf{w}\|_{L^\infty(\Omega^s)} + \|\omega\|_{L^\infty(\Omega^s)}^2 \right) \quad (4.55)$$

and $M_J > 1$.

Proof. If we define $\varphi(t) := (\mathbf{w} \cdot \nabla \overset{\leftarrow}{u}, \overset{\leftarrow}{u})_{0, \Omega^t}$, then $Q^{n+1}(\overset{\leftarrow}{u}, \overset{\leftarrow}{u}) = -\int_{I_n} \varphi + \Delta t \varphi^{n+1}$. Exploiting the Taylor-Lagrange formula we have

$$\varphi(t) = \varphi^{n+1} + \varphi'(s)(t - t_{n+1}) \quad s \in (t, t_{n+1})$$

and integrating over I_n we get

$$\begin{aligned} \int_{I_n} \varphi(t) &= \Delta t \varphi^{n+1} + \int_{I_n} \varphi'(s)(t - t_{n+1}) dt, \\ Q^{n+1}(\overset{\leftarrow}{u}, \overset{\leftarrow}{u}) &= -\int_{I_n} \varphi'(s)(t - t_{n+1}) dt. \end{aligned}$$

Observing that $\Delta t = \max_{t \in I_n} |t - t_{n+1}|$, we obtain

$$|Q^{n+1}(\overset{\leftarrow}{u}, \overset{\leftarrow}{u})| \leq \Delta t \int_{I_n} |\varphi'|.$$

Since changing variables $\Omega^{n+1} \ni \mathbf{y}_{n+1} \mapsto \mathbf{y} = \mathcal{A}_{n+1}^-(y_{n+1}, t) \in \Omega^t$ gives

$$\begin{aligned} \varphi(t) &= -\frac{1}{2} \int_{\Omega^t} \overset{\leftarrow}{u}(\mathbf{y}, t)^2 \nabla \cdot \mathbf{w}(\mathbf{y}, t) d\mathbf{y} \\ &= -\frac{1}{2} \int_{\Omega^{n+1}} u^{n+1}(\mathbf{y}_{n+1})^2 \nabla \cdot \mathbf{w}(\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t), t) J_{n+1}^-(\mathbf{y}_{n+1}, t) d\mathbf{y}_{n+1}, \end{aligned}$$

setting

$$\psi(\mathbf{y}_{n+1}, t) := \nabla \cdot \mathbf{w}(\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t), t) J_{n+1}^-(\mathbf{y}_{n+1}, t)$$

we realize that

$$|Q^{n+1}(\overset{\leftarrow}{u}, \overset{\leftarrow}{u})| \leq \frac{\Delta t}{2} \|u^{n+1}\|_{0, \Omega^{n+1}}^2 \int_{I_n} \left\| \frac{\partial \psi}{\partial t} \right\|_{L^\infty(\Omega^{n+1})}$$

because u^{n+1} does not depend on t . The next step is to note the dependence on t on both arguments of $\omega = \nabla \cdot \mathbf{w}$ in order to compute $\frac{\partial \psi}{\partial t}(\mathbf{y}_{n+1}, t)$. In view of (4.3),

(4.8) and the definition of $\omega(\mathbf{y}, t)$ with $\mathbf{y} = \mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t)$, this yields

$$\begin{aligned} \frac{\partial \psi}{\partial t}(\mathbf{y}_{n+1}, t) &= \frac{d}{dt} \left(\omega(\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t), t) J_{n+1}^-(\mathbf{y}_{n+1}, t) \right) \\ &= \left(\nabla \omega(\mathbf{y}, t) \cdot \frac{d}{dt} \mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t) + \frac{\partial \omega}{\partial t}(\mathbf{y}, t) \right) J_{n+1}^-(\mathbf{y}_{n+1}, t) \\ &\quad + \omega(\mathbf{y}, t) \frac{d}{dt} J_{n+1}^-(\mathbf{y}_{n+1}, t) \\ &= \left(\frac{\partial \omega}{\partial t}(\mathbf{y}, t) + \nabla \omega(\mathbf{y}, t) \cdot \mathbf{w}(\mathbf{y}, t) + \omega(\mathbf{y}, t)^2 \right) J_{n+1}^-(\mathbf{y}_{n+1}, t). \end{aligned}$$

Since $J_{n+1}^-(\mathbf{y}_{n+1}, t) \leq M_J$, we readily obtain $\int_{I_n} \|\frac{\partial \psi}{\partial t}\|_{L^\infty(\Omega^{n+1})} \leq C_n$ whence (4.54) follows. $\square$

We are now in a position to prove the stability of the practical semidiscrete scheme (4.50) under the following condition.

Assumption 6 (regularity of ALE velocity). The Eulerian ALE velocity $\mathbf{w}$ of (4.2) satisfies $\Lambda(\mathbf{w}, T) < \infty$, where

$$\begin{aligned} \Lambda(\mathbf{w}, T) &:= \exp \left(M_J \Delta t \int_0^T \left\| \frac{\partial \omega}{\partial s} \right\|_{L^\infty(\Omega^s)} + \|\nabla \omega\|_{L^\infty(\Omega^s)} \|\mathbf{w}\|_{L^\infty(\Omega^s)} + \|\omega\|_{L^\infty(\Omega^s)}^2 \right) \\ &= \exp \left(\Delta t \sum_{n=0}^{N_T-1} C_n \right). \end{aligned} \tag{4.56}$$

Proposition 11 (stability of practical time discretization). *Let the Eulerian ALE velocity $\mathbf{w}$ satisfy Assumption 6. Then, assuming that all time-steps satisfy*

$$\Delta t \max_{0 \leq n < N_T} C_n \leq 1, \tag{4.57}$$

where C_n is the constant given in (4.55), there holds

$$\begin{aligned} \|u^{N_T}\|_{0, \Omega_{N_T}}^2 + 2 \sum_{n=0}^{N_T-1} \|\bar{u}^{n+1} - u^n\|_{0, \Omega^n}^2 + 2 \sum_{n=0}^{N_T-1} \Delta t \|\nabla u^{n+1}\|_{0, \Omega^{n+1}}^2 \\ \leq \Lambda(\mathbf{w}, T) \left(2\|u_0\|_{0, \Omega^0}^2 + C \sum_{n=0}^{N_T-1} \Delta t \|\mathcal{F}^{n+1}\|^2 \right), \end{aligned} \tag{4.58}$$

provided $\exp \left(\Delta t \sup_{t \in [0, T]} \|\nabla \cdot \mathbf{w}\|_{L^\infty(\Omega^t)} \right) \leq M_J$.

Proof. Take $v = u^{n+1} \in H_0^1(\Omega^{n+1})$ in (4.50). The only difference w.r.t. Lemma 8 (stability of ideal time discretization) is the presence of $Q^{n+1}(\overleftarrow{u}, \overleftarrow{u})$ in (4.50). Since $\exp\left(\Delta t \sup_{t \in I_n} \|\nabla \cdot \mathbf{w}\|_{L^\infty(\Omega^t)}\right) \leq M_J$ applying Lemma 9 (temporal quadrature error), we end up with the inequality

$$\begin{aligned} \|u^{n+1}\|_{0,\Omega^{n+1}}^2 + \sum_{k=0}^n \|\overleftarrow{u}^{k+1} - u^k\|_{0,\Omega^k}^2 + \sum_{k=0}^n \Delta t \|\nabla u^{k+1}\|_{0,\Omega^{k+1}}^2 \\ \leq \|u_0\|_{0,\Omega^0}^2 + C \Delta t \sum_{k=0}^n \|\mathcal{F}^{n+1}\|^2 + \frac{\Delta t}{2} \sum_{k=0}^n C_k \|u^{k+1}\|_{0,\Omega^{k+1}}^2, \end{aligned}$$

for $0 \leq n \leq N_T - 1$, where C is as in Lemma 8 (stability of ideal time discretization) and $C_k := \frac{C_n}{2}$. We now recall the *discrete Gronwall inequality* for a sequence of non negative numbers $\alpha_n, \beta_n, \gamma_n, \delta_n$ with α_n non-decreasing

$$\gamma_n + \delta_n \leq \alpha_n + \sum_{k=0}^{n-1} \beta_k \gamma_k \implies \gamma_n + \delta_n \leq \alpha_n \exp\left(\sum_{k=0}^{n-1} \beta_k\right) \quad 0 \leq n \leq N_T - 1. \quad (4.59)$$

Note that, by defining $\beta_k := \Delta t C_k$ and $\gamma_k := \|u^{k+1}\|_{0,\Omega^{k+1}}^2$, an additional term appears at $k = n$ which prevents the direct application of the Gronwall inequality. However, provided that $\beta_n \leq 1$, we observe that:

$$\frac{1}{2} \sum_{k=0}^{n-1} \beta_k \|u^{k+1}\|_{0,\Omega^{k+1}}^2 + \frac{\beta_n}{2} \|u^{n+1}\|_{0,\Omega^{n+1}}^2 \leq \frac{1}{2} \sum_{k=0}^{n-1} \beta_k \|u^{k+1}\|_{0,\Omega^{k+1}}^2 + \frac{1}{2} \|u^{n+1}\|_{0,\Omega^{n+1}}^2.$$

Consequently, the last term on the right-hand side of the stability estimate can be absorbed into the corresponding term on the left-hand side. This leads to (4.58) with a stability constant $\exp\left(\Delta t \sum_{k=0}^{N_T-2} C_k\right) \leq \Lambda(\mathbf{w}, T)$, consistent with (4.56), which concludes the proof. $\square$

It is worth realizing that the stability constraint (4.57) is not very stringent because the constant C_n scales like Δt for a $W^{2,\infty}$ Eulerian velocity field $\mathbf{w}$. Consequently, assuming a realistic value of $M_J = 2$, the condition (4.57) reduces to

$$2\Delta t^2 \sup_{t \in [0,T]} \left(\left\| \frac{\partial(\nabla \cdot \mathbf{w})}{\partial t} \right\|_{L^\infty(\Omega^t)} + \|\nabla(\nabla \cdot \mathbf{w})\|_{L^\infty(\Omega^t)} \|\mathbf{w}\|_{L^\infty(\Omega^t)} + \|\nabla \cdot \mathbf{w}\|_{L^\infty(\Omega^t)}^2 \right) \leq 1.$$

4.4 Full Discretization

Once the time discretization has been introduced, a sequence of PDEs must be solved in Ω^n for each $0 \leq n < N_T$. Since we aim to solve Problem (4.38) with the SBM, we must properly define the domain $\tilde{\Omega}_h^n := \tilde{\Omega}_h^{t_n}$ (and $\tilde{\Gamma}_h^n := \tilde{\Gamma}_h^{t_n}$ with normal vector $\tilde{\mathbf{n}}_n$). Recalling Section 4.2, at each time step we define:

- i) $\tilde{\mathcal{T}}_h^n := \tilde{\mathcal{T}}_h^{t_n}$, the triangulation defined in (4.21),
- ii) $\tilde{\Omega}_h^n := \tilde{\Omega}_h^{t_n}$, the surrogate domain defined in (4.22),
- iii) $V_h^n := V_h^{t_n} = V_h(\tilde{\Omega}_h^{t_n}; \tilde{\mathcal{T}}_h^{t_n})$, the finite dimensional subspace of $H^1(\tilde{\Omega}_h^n)$ corresponding to the triangulation and the domain defined in (4.25).

We also recall that, as seen in (4.20), the domain $\mathcal{D}$ is stationary by definition. Thus $\mathcal{T}_h$, the admissible and shape-regular triangulation of $\mathcal{D}$, it does not move under the ALE map $\mathcal{A}$. Therefore, to construct $\tilde{\mathcal{T}}_h^n$, we simply select the elements $E \in \mathcal{T}_h$ that satisfy (4.21) at each time step.

Moreover, it is necessary to say that, while the ALE map $\mathcal{A}_n$ is a bijection between Ω^n and Ω^{n+1} , in general does not map $\tilde{\Omega}_h^n$ onto $\tilde{\Omega}_h^{n+1}$, but is still well-defined since $\tilde{\Omega}_h^n \subset \Omega^n$ for all $0 \leq n \leq N_T$ by definition of surrogate domain in (4.22). Therefore, we are led to introduce the domain

$$\tilde{\Omega}_h^n := \mathcal{A}_n^{-1}(\tilde{\Omega}_h^{n+1}), \quad (4.60)$$

so that $\mathcal{A}_n$ is a bijection between $\tilde{\Omega}_h^n$ and $\tilde{\Omega}_h^{n+1}$. The Jacobian of the inverse of such map will be still denoted by J_n^{-1} .

Given a function u_h^n defined in $\tilde{\Omega}_h^n$, the term $u_h^n \circ \mathcal{A}_n^{-1}$ may not be defined in $\tilde{\Omega}_h^{n+1}$ because, as we said before, $\mathcal{A}_n$ needs a point in $\tilde{\Omega}_h^n$ as input. To overcome this drawback, it will be convenient to define the following discrete extension operator.

Definition 4. Let us call $\mathcal{E}_h^n : V_h^n(\tilde{\Omega}_h^n) \rightarrow L^2(\tilde{\Omega}_h^n)$ the discrete extension operator such that:

$$\mathcal{E}_h^n u_h^n(\tilde{\mathbf{x}}) = p(\tilde{\mathbf{x}}), \quad (4.61)$$

if $p \in \mathbb{P}_1(\mathbb{R}^2)$ such that

$$p(\tilde{\mathbf{x}})|_{E(\tilde{\mathbf{x}})} = u_h^n|_{E(\tilde{\mathbf{x}})} \quad (4.62)$$

where $E(\tilde{\mathbf{x}})$ is the closest element $E \in \tilde{\mathcal{T}}_h$ to $\tilde{\mathbf{x}} \in \tilde{\Omega}_h^n$.

Note that we can define the operator pointwise because we are using the $\mathbb{P}_1$ -Lagrangian basis with piecewise continuous polynomial functions. This operator satisfies two important properties.

Property 6. The discrete extension operator $\mathcal{E}_h^n$ satisfies, for all v_h in V_h^n ,

$$\|\mathcal{E}_h^n v_h\|_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}^2 \leq C_{\mathcal{E}_h} \|v_h\|_{0, \tilde{\Omega}_h^n}^2 \quad (4.63)$$

$$\|\mathcal{E}_h^n v_h\|_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}^2 \leq C'_{\mathcal{E}_h} h^2 \|v_h\|_a^2 \quad (4.64)$$

with $C_{\mathcal{E}_h}, C'_{\mathcal{E}_h} > 0$ independent of the meshsize.

Proof. Let us consider the strip S_h between the real and the surrogate domain defined as

$$S_h := \Omega^n \setminus \tilde{\Omega}_h^n = \bigcup_{\substack{e \subset \partial E \cap \tilde{\Gamma}_h^n, \\ E \in \tilde{\mathcal{T}}_h}} S_e, \quad (4.65)$$

where e is the edge of the nearest element E that is also edge on the surrogate boundary $\tilde{\Gamma}_h^n$. The patches S_e , with e edge of E , are the ‘influence zones’ of E outside the surrogate domain $\tilde{\Omega}_h^n$, hence the parts where E is the nearest element. Considering the reference system (x, y) along the edge, we introduce the following change of variable $(x, y) \mapsto (s, t)$

$$\begin{aligned} x(s, t) &= s, \\ y(s, t) &= t \cos \alpha(s), \end{aligned}$$

where s is the parametrization of the edge with length a such that $0 \leq s \leq a$. On the other hand, y parametrizes the distance from the surrogate boundary $\tilde{\Gamma}_h^n$ to the real boundary Γ^n and satisfies $0 \leq t \leq \delta(s)$. Moreover, let us consider $\alpha(s)$ as the angle between t and y at $x = s$.

Setting $\tilde{v}_h = \mathcal{E}_h^n v_h$, we have

$$\int_{S_e} \tilde{v}_h(x, y) \, dx dy = \int_0^a \left(\int_0^{\delta(s)} \tilde{v}_h(s, t) \cos \alpha(s) \, dt \right) ds,$$

where $\cos \alpha(s)$ is the determinant of the Jacobian related to the change of variable. Using the Taylor expansion we have

$$\tilde{v}_h(s, t) = \tilde{v}_h(s, 0) + \int_0^t \frac{\partial \tilde{v}_h}{\partial t}(s, \xi) d\xi \quad (4.66)$$

where

$$\frac{\partial \tilde{v}_h}{\partial t} = \frac{\partial \tilde{v}_h}{\partial x} \underbrace{\frac{\partial x}{\partial t}}_{=0} + \frac{\partial \tilde{v}_h}{\partial y} \underbrace{\frac{\partial y}{\partial t}}_{=\cos \alpha(s)} = \nabla \tilde{v}_h \cdot \boldsymbol{\nu}(s)$$

is constant along t . So we obtain

$$\tilde{v}_h^2(s, t) \lesssim \tilde{v}_h^2(s, 0) + |\nabla \tilde{v}_h|^2 t^2.$$

The L^2 -norm of $\tilde{v}_h$ results in

$$\begin{aligned} \|\tilde{v}_h\|_{0, S_e}^2 &= \int_{S_e} \tilde{v}_h^2(s, t) \cos^2 \alpha(s) \, ds dt \\ &\lesssim \int_0^a \left(\int_0^{\delta(s)} \tilde{v}_h^2(s, 0) \, dt \right) ds + |\nabla \tilde{v}_h|^2 \int_0^a \left(\int_0^{\delta(s)} t^2 \, dt \right) ds \\ &= \int_0^a \tilde{v}_h^2(s, 0) \delta(s) \, ds + |\nabla \tilde{v}_h|^2 \int_0^a \frac{1}{3} \delta^3(s) \, ds \\ &\leq h \int_0^a \tilde{v}_h^2(s, 0) \, ds + |\nabla \tilde{v}_h|^2 h^4, \end{aligned}$$

where $|\delta(s)| \leq h$. Recall that by definition of extension operator, in $s = 0$ we have $\tilde{v}_{h|e} = v_{h|e}$ and $\nabla \tilde{v}_{h|S_e} = \nabla v_{h|E}$ with E nearest element and e one of his edges that lies on $\tilde{\Gamma}_h^n$. Note that

$$\|\nabla v_h\|_{0,E}^2 = \int_E |\nabla v_h|^2 = |\nabla v_h|^2 |E| \simeq |\nabla v_h|^2 h^2 \quad \text{and} \quad \|\nabla v_h\|_{0,e}^2 \simeq |\nabla v_h|^2 h,$$

hence, exploiting the definition of $S_h v_h$ on e and using (2.31), we obtain

$$\begin{aligned} \|\tilde{v}_h\|_{0,S_e}^2 &\lesssim h \|v_h\|_{0,e}^2 + h^2 \|\nabla v_h\|_{0,E}^2 \lesssim h \|S_h v_h\|_{0,e}^2 + h \|\nabla v_h \cdot \mathbf{d}\|_{0,e}^2 + h^2 \|\nabla v_h\|_{0,E}^2 \\ &\lesssim h^2 \|h^{-1/2} S_h v_h\|_{0,e}^2 + h^3 \|\nabla v_h\|_{0,e}^2 + h^2 \|\nabla v_h\|_{0,E}^2 \\ &\lesssim h^2 \left(\|h^{-1/2} S_h v_h\|_{0,e}^2 + \|\nabla v_h\|_{0,E}^2 \right). \end{aligned}$$

Summing all over the elements in $\tilde{\mathcal{T}}_h^n$ and considering that $\tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n \subset S_h$, we get the desired result in (4.64). Using the inverse inequality we can easily obtain (4.63). $\square$

Property 7. $\mathcal{E}_h^n$ satisfies, $\forall v \in H^2(\Omega^n)$ solution of Problem (4.18) and $v_h \in V_h^n(\tilde{\Omega}_h^n; \tilde{\mathcal{T}}_h^n)$ solution of Problem (4.27) both at $t = t_n$,

$$\|\mathcal{E}_h^n v_h - v\|_{0,\tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n} \leq C''_{\mathcal{E}_h} h^2 |v|_{2,\Omega^n} \quad (4.67)$$

with $C''_{\mathcal{E}_h} > 0$ independent of the meshsize.

Proof. Let us set $w_h := \mathcal{E}_h^n v_h - v$ and note that, by the definition of the extension operator, w_h vanishes on the boundary $\tilde{\Gamma}_h^n$ of $\tilde{\Omega}_h^n$. Similarly to (4.66), after applying the change of variables in a patch S_e , the Taylor expansion of w_h reads

$$w_h(s, t) = \underbrace{w_h(s, 0)}_{=0} + \int_0^t \frac{\partial w_h}{\partial t}(s, \xi) d\xi = \frac{\partial w_h}{\partial t} t,$$

yielding

$$\|w_h\|_{0,S_e}^2 \lesssim |\nabla w_h|_{S_e}|^2 \int_0^a \int_0^{\delta(s)} t^2 dt ds \lesssim |\nabla w_h|_{S_e}|^2 h^4.$$

We can obtain the same result by applying the Poincaré inequality, observing that $w_h \in H_0^1(\tilde{\Omega}_h^n)$ since it vanishes on the boundary.

Furthermore, recalling that E is the nearest element to S_e , we observe that

$$\begin{aligned} h^2 |\nabla w_h|_{S_e}|^2 &= h^2 |\nabla(\mathcal{E}_h^n v_h - v)|_{S_e}|^2 = h^2 |\nabla(v_{h|E} - v_{|S_e})|^2 \\ &\lesssim \|\nabla v_{h|E} - \mu_E(\nabla v)\|_{0,E}^2 + h^2 |\mu_E(\nabla v) - \mu_{S_e}(\nabla v)|^2 \\ &\quad + \|\mu_{S_e}(\nabla v) - \nabla v_{|S_e}\|_{0,S_e}^2, \end{aligned}$$

where $\mu_E(\varphi) = \frac{1}{|E|} \int_E \varphi$ denotes the integral mean. The first term on the right-hand side can be bounded using the Poincaré-Wirtinger inequality; recalling that the size of the domain is $O(h)$, we obtain

$$\begin{aligned} \|\nabla v_h|_E - \mu_E(\nabla v)\|_{0,E}^2 &\leq \|\nabla v_h|_E - \nabla v\|_{0,E}^2 + \|\nabla v - \mu_E(\nabla v)\|_{0,E}^2 \\ &\lesssim \|v_h|_E - v\|_{1,E}^2 + h^2 |v|_{2,E}^2. \end{aligned} \quad (4.68)$$

Similarly, we obtain the bound

$$\|\mu_{S_e}(\nabla v) - \nabla v|_{S_e}\|_{0,S_e}^2 \lesssim h^2 |v|_{0,S_e}^2. \quad (4.69)$$

The difference between the two means requires further analysis. In particular, let z denote a single component of the vector ∇v . Note that if $v \in H^2(\Omega^n)$, it follows that $z \in H^1(\Omega^n)$. Let $\hat{E}$ and $\hat{S}_e$ be the reference element and the reference patch, respectively, associated with the affine transformations B_E and B_{S_e} . Furthermore, we define $\hat{z} := z \circ B_i$ for $i \in \{E, S_e\}$ (see Figure 4.4).

It is important to observe that, since B_E is an affine mapping, its Jacobian determinant $\det B_E$ is constant. Therefore, we have

$$\int_E d\mathbf{x} = \int_{\hat{E}} \det B_E d\hat{\mathbf{x}} \quad \implies \quad \det B_E = \frac{\int_E d\mathbf{x}}{\int_{\hat{E}} d\hat{\mathbf{x}}} = \frac{|E|}{|\hat{E}|},$$

which yields

$$\mu_E(z) = \frac{1}{|E|} \int_E z d\mathbf{x} = \frac{\det B_E}{|E|} \int_{\hat{E}} \hat{z} d\hat{\mathbf{x}} = \frac{1}{|\hat{E}|} \int_{\hat{E}} \hat{z} d\hat{\mathbf{x}} = \mu_{\hat{E}}(\hat{z}).$$

This result is analogously valid for S_e . In addition, for any constant polynomial $p_0 \in \mathbb{P}_0(\Omega^n)$, it trivially holds that

$$\mu_E(p_0) = \frac{1}{|E|} \int_E p_0 d\mathbf{x} = \frac{p_0}{|E|} \int_E d\mathbf{x} = p_0 = \mu_{S_e}(p_0).$$

Thus, we can write

$$\begin{aligned} |\mu_E(z) - \mu_{S_e}(z)| &= |\mu_{\hat{E}}(\hat{z}) - \mu_{\hat{S}_e}(\hat{z})| = |\mu_{\hat{E}}(\hat{z} + p_0) - \mu_{\hat{S}_e}(\hat{z} + p_0)| \\ &\lesssim \|\hat{z} + p_0\|_{0,\hat{E} \cup \hat{S}_e} \leq \|\hat{z} + p_0\|_{1,\hat{E} \cup \hat{S}_e} \quad \forall p_0 \in \mathbb{P}_0(\hat{E} \cup \hat{S}_e), \end{aligned}$$

and applying the Deny-Lions lemma we finally obtain

$$|\mu_E(z) - \mu_{S_e}(z)| \leq \inf_{p_0 \in \mathbb{P}_0(\hat{E} \cup \hat{S}_e)} \|\hat{z} + p_0\|_{1,\hat{E} \cup \hat{S}_e} \lesssim |\hat{z}|_{1,\hat{E} \cup \hat{S}_e} \lesssim |z|_{1,E \cup S_e}, \quad (4.70)$$

where, in the last step, we used a standard scaling argument which, for the H^1 -seminorm in dimension $d = 2$, yields a scaling factor of $h^{1-d/2} = h^0 = 1$. Note

that the shape regularity and the $O(h)$ size of S_e make this analysis via reference elements and patches consistent.

Summing (4.68) over all elements $E \in \tilde{\mathcal{T}}_h^n$ and recalling that the H^1 -norm is bounded by the a -norm, we can control the first term using (2.55). This yields

$$\|v - v_h\|_{1, \tilde{\Omega}_h^{n, \Gamma}}^2 \lesssim \|v - v_h\|_{a, \tilde{\Omega}_h^{n, \Gamma}}^2 \lesssim h^2 |v|_{2, \tilde{\Omega}_h^{n, \Gamma}}, \quad (4.71)$$

where we have defined the boundary layer of elements as

$$\tilde{\Omega}_h^{n, \Gamma} := \bigcup_{\substack{\partial E \cap \tilde{\Gamma}_h^n, \\ E \in \tilde{\mathcal{T}}_h}} E,$$

and $\|\cdot\|_{a, \tilde{\Omega}_h^{n, \Gamma}}$ denotes the a -norm introduced in (2.33), restricted to $\tilde{\Omega}_h^{n, \Gamma}$.

Combining the bounds (4.68), (4.69), (4.70), and (4.71) over all elements and patches, we obtain the desired result. It is important to observe that, recalling (4.65), the seminorm on the right-hand side of (4.67) should rigorously be computed over $S_h \cup \tilde{\Omega}_h^{n, \Gamma} \subset \Omega^n$; however, we have chosen to denote it simply over Ω^n for the sake of notational simplicity. $\square$

We are now ready to introduce the full discrete scheme solving with the SBM all the boundary value problems at each time step. Applying the definitions in (4.30) and (4.31) to the practical scheme in (4.50) and recalling (4.46), thereby obtaining the following

Space-time discretization scheme: find $u_h^{n+1} \in V_h^{n+1}$

$$\begin{aligned} (u_h^{n+1} - \mathcal{E}_h^n u_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} - \Delta t c^{n+1}(u_h^{n+1}, v_h) + \Delta t a_h^{n+1}(u_h^{n+1}, v_h) \\ = \Delta t F_h^{n+1}(v_h), \quad \forall v_h \in V_h^{n+1}, \quad 0 \leq n < N_T \end{aligned} \quad (4.72)$$

where

$$\begin{aligned} c^{n+1}(u_h^{n+1}, v_h) &:= c(t_{n+1}; u_h^{n+1}, v_h), \\ a_h^{n+1}(u_h^{n+1}, v_h) &:= a_h(t_{n+1}; u_h^{n+1}, v_h), \\ F_h^{n+1}(v_h) &:= F_h(t_{n+1}; v_h). \end{aligned}$$

Note that, since u_h^n is defined in $\tilde{\Omega}_h^n$, we need to extrapolate in $\tilde{\Omega}_h^n$ with the help of $\mathcal{E}_h^n$ that we have defined in (4.61). Next section will be devoted to the stability analysis of this scheme, whereas the error analysis will be detailed in Sect 4.5.2.

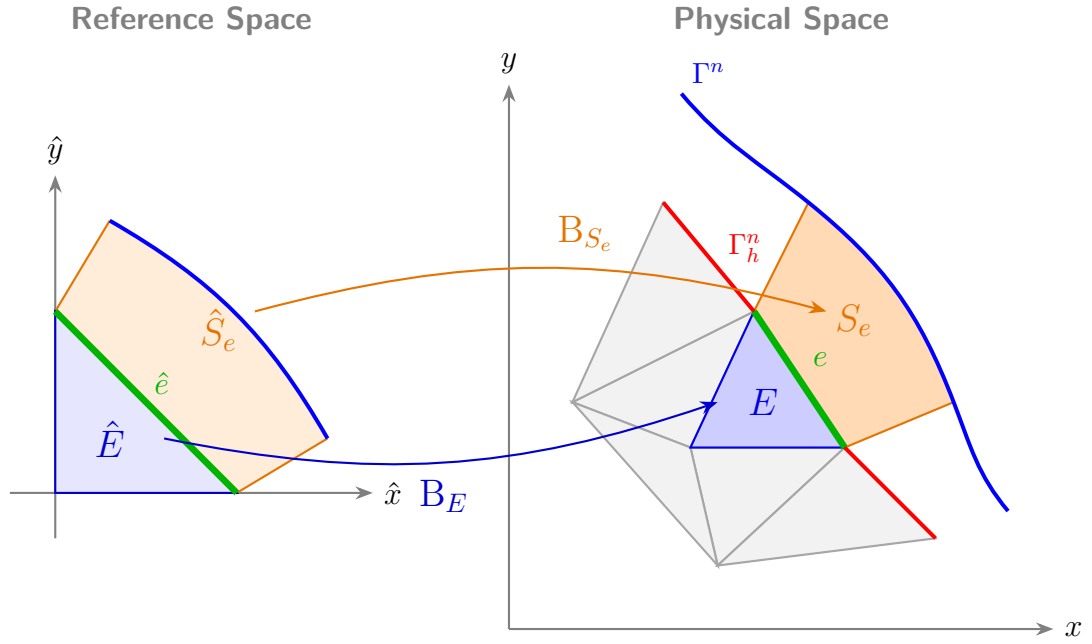


Figure 4.4: Mapping from the reference space to the physical space. The domain features the real curved boundary Γ^n (blue) and the surrogate polygonal boundary $\tilde{\Gamma}_h^n$ (red). The physical element E lies inside the surrogate domain, touching $\tilde{\Gamma}_h^n$ strictly along the local edge e (green). The influence zone S_e bridges the gap between e and the true boundary along the projection directions, which fan out non-parallelly. The mappings B_E and B_{S_e} link the reference entities to their physical counterparts.

4.5 Stability, consistency and convergence analysis - Moving domains

4.5.1 Stability

First of all, take $v_h = u_h^{n+1}$ in (4.72) to get

$$\begin{aligned} (u_h^{n+1} - \mathcal{E}_h^n u_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} u_h^{n+1})_{0, \tilde{\Omega}_h^{n+1}} - \Delta t c^{n+1}(u_h^{n+1}, u_h^{n+1}) + \Delta t a_h^{n+1}(u_h^{n+1}, u_h^{n+1}) \\ = \Delta t F_h^{n+1}(u_h^{n+1}), \quad \forall v_h \in V_h^{n+1}, \quad 0 \leq n < N_T \end{aligned}$$

which we can write as

$$\begin{aligned} (u_h^{n+1} - \mathcal{E}_h^n u_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} u_h^{n+1})_{0, \tilde{\Omega}_h^{n+1}} - \frac{\Delta t}{2} (\nabla \cdot \mathbf{w}^{n+1}, (u_h^{n+1})^2)_{0, \tilde{\Omega}_h^{n+1}} \\ + \Delta t a_h^{n+1}(u_h^{n+1}, u_h^{n+1}) = \Delta t F_h^{n+1}(u_h^{n+1}) + \frac{\Delta t}{2} (\mathbf{w}^{n+1} \cdot \tilde{\mathbf{n}}_{n+1}, u_h^2)_{0, \tilde{\Gamma}_h^{n+1}}. \end{aligned} \quad (4.73)$$

To proceed further, it is convenient to recall from (4.52) the definition of the backward map $\mathbf{y} = \mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t) = \mathcal{A}^{-1}(\mathbf{y}_{n+1}, t, t_{n+1}) = \mathcal{A}(\mathbf{y}_{n+1}, t_{n+1}, t)$ for all $\mathbf{y}_{n+1} \in \tilde{\Omega}_h^{n+1}$ and $t \in I_n$. Note that $\mathcal{A}_{n+1}^-(\mathbf{y}_{n+1}, t) = \mathcal{A}_n^-(\mathbf{y}_{n+1})$. Extending the definition (4.60), the images of $\tilde{\Omega}_h^{n+1}$ through the mappings $\mathcal{A}_{n+1}^-$ define the domains $\tilde{\Omega}_h^t := \mathcal{A}_{n+1}^-(\tilde{\Omega}_h^{n+1}, t)$ for any $t \in I_n$. Recalling (4.43), by backward motion, the function $u_h^{n+1} \in V_h^{n+1}$ defines the function $\overleftarrow{u}_h$ by setting

$$\overleftarrow{u}_h(\mathbf{y}, t) := u_h^{n+1}(\mathbf{y}_{n+1}) \quad \forall \mathbf{y} \in \tilde{\Omega}_h^t, \quad t \in I_n, \quad (4.74)$$

if $\mathbf{y}_{n+1} = \mathcal{A}(\mathbf{y}, t, t_{n+1})$.

Thus, applying a change of variable between $\tilde{\Omega}_h^{n+1}$ and $\tilde{\Omega}_h^t$ similar to (4.46), the first inner product in (4.73) satisfies

$$(u_h^{n+1} - \mathcal{E}_h^n u_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} u_h^{n+1})_{0, \tilde{\Omega}_h^{n+1}} = (\overleftarrow{u}_h^{n+1} - \mathcal{E}_h^n u_h^n, \overleftarrow{u}_h^{n+1})_{0, \tilde{\Omega}_h^t}, \quad (4.75)$$

whereas the second inner product can be written as

$$\Delta t (\nabla \cdot \mathbf{w}^{n+1}, (\overleftarrow{u}_h^{n+1})^2)_{0, \tilde{\Omega}_h^{n+1}} = \int_{I_n} (\nabla \cdot \mathbf{w}^{n+1}, (\overleftarrow{u}_h^{n+1})^2)_{0, \tilde{\Omega}_h^t} + Q_h^{n+1}(\overleftarrow{u}_h, \overleftarrow{u}_h) \quad (4.76)$$

where the quadrature error $Q_h^{n+1}(\overleftarrow{u}_h, \overleftarrow{u}_h)$ is defined as in (4.51) with Ω^t replaced by $\tilde{\Omega}_h^t$. In particular,

$$Q_h^{n+1}(\overleftarrow{u}_h, \overleftarrow{u}_h) := - \int_{I_n} (\nabla \cdot \mathbf{w}^{n+1}, (\overleftarrow{u}_h^{n+1})^2)_{0, \tilde{\Omega}_h^t} + \Delta t (\nabla \cdot \mathbf{w}^{n+1}, (\overleftarrow{u}_h^{n+1})^2)_{0, \tilde{\Omega}_h^{n+1}}. \quad (4.77)$$

Equations (4.75) and (4.76) suggest to use the following discrete version of Lemma 7 (material derivative).

Lemma 10 (discrete material derivative). *The following inequality holds*

$$\begin{aligned} & (\bar{u}_h^{n+1} - \mathcal{E}_h^n u_h^n, \bar{u}_h^{n+1})_{0, \bar{\Omega}_h^n} - \frac{1}{2} \int_{I_n} (\nabla \cdot \mathbf{w}^{n+1}, (\bar{u}_h^{n+1})^2)_{0, \bar{\Omega}_h^s} \\ & \geq \frac{1}{2} \|u_h^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}}^2 - \frac{1}{2} \|u_h^n\|_{0, \tilde{\Omega}_h^n}^2 + \frac{1}{2} \|\bar{u}_h^{n+1} - \mathcal{E}_h^n u_h^n\|_{0, \bar{\Omega}_h^n}^2 - \frac{C_{\mathcal{E}_h}' h^2}{2} \|u_h^n\|_a^2. \end{aligned} \quad (4.78)$$

Proof. Lemma 7 applied to $\bar{u}_h^{n+1}$ and $\mathcal{E}_h^n u_h^n$ in the family of discrete domains $\bar{\Omega}_h^t$ yields

$$\begin{aligned} & (\bar{u}_h^{n+1} - \mathcal{E}_h^n u_h^n, \bar{u}_h^{n+1})_{0, \bar{\Omega}_h^n} - \frac{1}{2} \int_{I_n} (\nabla \cdot \mathbf{w}^{n+1}, (\bar{u}_h^{n+1})^2)_{0, \bar{\Omega}_h^s} = \frac{1}{2} \|u_h^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}}^2 \\ & \quad - \frac{1}{2} \|\mathcal{E}_h^n u_h^n\|_{0, \bar{\Omega}_h^n}^2 + \frac{1}{2} \|\bar{u}_h^{n+1} - \mathcal{E}_h^n u_h^n\|_{0, \bar{\Omega}_h^n}^2 \end{aligned}$$

where, looking at the proof of Lemma 7, the integral term needs to be like in (4.77) in order to apply Reynolds' identity (4.17). We can split the second term of the right-hand side as

$$\int_{\bar{\Omega}_h^n} (\mathcal{E}_h^n u_h^n)^2 = \int_{\bar{\Omega}_h^n \cap \tilde{\Omega}_h^n} (u_h^n)^2 + \int_{\bar{\Omega}_h^n \setminus \tilde{\Omega}_h^n} (\mathcal{E}_h^n u_h^n)^2$$

because by (4.61) we have $\mathcal{E}_h^n u_h^n = u_h^n$ in $\tilde{\Omega}_h^n$. Finally, applying (4.64) we get the result. $\square$

Next, we combine equations (4.73), (4.75), (4.76), and Lemma 10. In addition, we invoke the coercivity of a_h in (4.32), and we use (4.36) and (4.37) with $p = 3$. This yields the **first stability inequality**:

$$\begin{aligned} & \frac{1}{2} \|u_h^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}}^2 - \frac{1}{2} \|u_h^n\|_{0, \tilde{\Omega}_h^n}^2 + \frac{1}{2} \|\bar{u}_h^{n+1} - \mathcal{E}_h^n u_h^n\|_{0, \bar{\Omega}_h^n}^2 + \alpha \Delta t \|u_h^{n+1}\|_a^2 \\ & \leq \Delta t \|F_h\|_{(V_h^{n+1})'} \|u_h^{n+1}\|_a + \left| Q_h^{n+1}(\bar{u}_h, \bar{u}_h) \right| + \frac{C_{\mathcal{E}_h}' h^2}{2} \|u_h^n\|_a^2 \\ & \quad + \frac{2\Delta t}{C_{\tau, a}} \|\mathbf{w}^{n+1} \cdot \tilde{\mathbf{n}}^{n+1}\|_{L^\infty(\tilde{\Gamma}_h^{n+1})} \|u_h^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}}^2 + \frac{\alpha}{6} \|u_h^{n+1}\|_a^2. \end{aligned} \quad (4.79)$$

Observe that the quadrature error $\left| Q_h^{n+1} \right|$ could be bounded as in Lemma 9, obtaining

$$\left| Q_h^{n+1}(\bar{u}_h, \bar{u}_h) \right| \leq \frac{\Delta t}{2} \bar{C}_{0, n} \|u_h^{n+1}\|_{\tilde{\Omega}_h^{n+1}}^2 \quad (4.80)$$

where the constant $\bar{C}_{0, n}$ is defined as in (4.55), but with Ω^s replaced by the extended domain $\bar{\Omega}_h^s$. The result holds provided

$$\exp \left(\int_{I_n} \|\omega\|_{L^\infty(\bar{\Omega}_h^s)} \right) \leq M_J. \quad (4.81)$$

For the sake of simplicity, let us set

$$\overleftarrow{C}_n := 2\overleftarrow{C}_{0,n} + \frac{8}{C_{\tau,a}} \|\mathbf{w}^{n+1} \cdot \tilde{\mathbf{n}}^{n+1}\|_{L^\infty(\tilde{\Omega}_h^{n+1})}. \quad (4.82)$$

Before establishing the stability estimate, let us state the following condition that links the time step with the meshsize in similar manner that the recalls the famous CFL condition.

Assumption 7 (relation between h and Δt). The meshsize and the time step satisfy

$$Ch^2 \leq \Delta t \frac{\alpha}{p}, \quad (4.83)$$

with p integer in according to the one selected for (4.37), $\alpha > 0$ coercivity constant defined in Proposition 1, and $C > 0$ independent of h and Δt .

Proposition 12 (stability of the fully discrete scheme). *Let Assumptions 6 (regularity of ALE velocity) and 7 hold and the condition (4.81) be valid for all $0 \leq n < N_T$. Let the time-step Δt satisfy the condition*

$$\Delta t \max_{0 \leq n < N_T} \overleftarrow{C}_n \leq 1, \quad (4.84)$$

where $\overleftarrow{C}_n$ is defined in (4.82). Then the solution of the fully discrete scheme (4.72) fulfills the stability estimate

$$\begin{aligned} & \|u_h^{N_T}\|_{0,\tilde{\Omega}_h^{N_T}}^2 + \sum_{n=0}^{N_T-1} \|\overleftarrow{u}_h^{n+1} - \mathcal{E}_h^n u_h^n\|_{0,\tilde{\Omega}_h^n}^2 + \alpha \sum_{n=0}^{N_T-1} \Delta t \|u_h^{n+1}\|_a^2 \\ & \leq \Lambda(\mathbf{w}, T) \left(2\|u_{0,h}\|_{0,\tilde{\Omega}_h^0}^2 + \frac{6}{\alpha} \sum_{n=0}^{N_T-1} \Delta t \|F_h^{n+1}\|_{(V_h^{n+1})'}^2 \right), \end{aligned} \quad (4.85)$$

where $\Lambda = \Lambda(\mathbf{w}, T) = \exp\left(\Delta t \sum_{n=0}^{N_T-2} \overleftarrow{C}_n\right)$ is an exponential constant depending on the ALE map.

Proof. Summing (4.79) over $0 \leq n \leq \nu$ for any $\nu \leq N_T - 1$ and applying Young's inequality to the first term of the right-hand side we obtain

$$\begin{aligned} & \|u_h^{n+1}\|_{0,\tilde{\Omega}_h^{n+1}}^2 + \sum_{n=0}^{\nu} \|\overleftarrow{u}_h^{n+1} - \mathcal{E}_h^n u_h^n\|_{0,\tilde{\Omega}_h^n}^2 + \alpha \sum_{n=0}^{\nu} \Delta t \|u_h^{n+1}\|_a^2 \\ & \leq \|u_{0,h}\|_{0,\tilde{\Omega}_h^0}^2 + \frac{3\Delta t}{\alpha} \sum_{n=0}^{\nu} \|F_h\|_{(V_h^{n+1})'}^2 + \frac{\Delta t}{2} \sum_{n=0}^{\nu} \overleftarrow{C}_n \|u_h^{n+1}\|_{0,\tilde{\Omega}_h^{n+1}}^2. \end{aligned}$$

We next apply Gronwall's inequality (4.59), as in the proof of Proposition 11, to get the asserted estimate. $\square$

4.5.2 Consistency

Before starting with the error analysis, let us extend the definition of projector operator stated in Section 3.2.2.

Definition 5. Given $w \in V(\tilde{\Omega}_h^t; \tilde{\mathcal{T}}_h^t)$, let $P_h^t w$ be the a_h -projection of w upon V_h^t , i.e., for all $0 < t \leq T$

$$a_h(t; P_h^t w, v_h) = a_h(t; w, v_h) + e_\perp^t(v_h) \quad \forall v_h \in V_h, \quad (4.86)$$

where the term $e_\perp^t(v_h) := -(R_h^t w, \partial_{\tilde{n}_t} v_h)_{0, \tilde{\Gamma}_h^t} + \gamma(h^{-1} R_h^t w, S_h^t v_h)_{0, \tilde{\Gamma}_h^t} \neq 0$ since this is not an orthogonal projection (R_h^t is the Taylor remainder at time t). Therefore, assuming $w \in H^2(\tilde{\Omega}_h^t)$, we have

$$\|w - P_h^t w\|_a \lesssim h |w|_{2, \tilde{\Omega}_h^t}, \quad (4.87)$$

$$\|w - P_h^t w\|_{0, \tilde{\Omega}_h^t} \lesssim h^{3/2} |w|_{2, \tilde{\Omega}_h^t}. \quad (4.88)$$

Let us set

$$\mathcal{L}_h(t; u, v) := -c(t; u, v) + a_h(t; u, v)$$

as the SBM operator. Note that the analysis in Chapter 2 could be extended to a generic diffusion-convection-reaction PDE since with SBM we only need to handle the boundary term deriving from partial integration of the diffusive term $-\Delta u$. Convection and reaction terms could be added without any problem, they will simply change the value of coercivity and continuity constants that we will set as $\alpha_{\mathcal{L}}$ and $\|a\|_{\mathcal{L}}$ both greater than zero and independent of the meshsize. For this reason, let us consider $P_h^{t_n}$ as the $\mathcal{L}_h$ -projector with the same properties of Definition 5. To simplify, let us set $P_h^n := P_h^{t_n}$.

In order to estimate the error

$$e^n = u(t_n) - u_h^n, \quad (4.89)$$

let us introduce the projection $p_h^n := P_h^n u(t_n) \in V_h^n$ of the exact solution at time t_n , and let us split

$$e^n = (u(t_n) - p_h^n) + (p_h^n - u_h^n) =: \pi^n + \theta_h^n. \quad (4.90)$$

The projection error π^n can be estimated using (4.87) or (4.88), whereas the discrete error θ_h^n can be controlled similarly to the stability estimate in Proposition 12 (stability of the fully discrete scheme), provided we show it satisfies a recursive equation similar to (4.72). To this end, let us write

$$\begin{aligned} & (\theta_h^{n+1} - \mathcal{E}_h^n \theta_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} + \Delta t \mathcal{L}_h^{n+1}(\theta_h^{n+1}, v_h) \\ &= (p_h^{n+1} - \mathcal{E}_h^n p_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} + \Delta t \mathcal{L}_h^{n+1}(p_h^{n+1}, v_h) \\ & \quad - \underbrace{(u_h^{n+1} - \mathcal{E}_h^n u_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} - \Delta t \mathcal{L}_h^{n+1}(u_h^{n+1}, v_h)}_{-\Delta t F_h^{n+1}(v_h)} \end{aligned}$$

and using (4.86) for the $\mathcal{L}_h$ operator we get

$$\begin{aligned}
 & (p_h^{n+1} - \mathcal{E}_h^n p_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} + \Delta t \mathcal{L}_h^{n+1}(p_h^{n+1}, v_h) - \Delta t F_h^{n+1}(v_h) \\
 &= (p_h^{n+1} - \mathcal{E}_h^n p_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} + \Delta t \mathcal{L}_h^{n+1}(u^{n+1}, v_h) \\
 &\quad + \Delta t e_\perp^{n+1}(v_h) - \Delta t F_h^{n+1}(v_h) \\
 &= (p_h^{n+1} - \mathcal{E}_h^n p_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} - \Delta t d(t_{n+1}; u, v_h) + \Delta t e_\perp^{n+1}(v_h) \\
 &=: \Delta t E_c^{n+1}(v_h)
 \end{aligned} \tag{4.91}$$

recalling the definition of $d(t; \cdot, \cdot)$ in (4.28) and the Problem (4.27). We have obtained the recursive equation satisfied by the discrete error.

Discrete error equation: For $n = 0, \dots, N_T - 1$, find $\theta_h^{n+1} \in V_h^{n+1}$ such that

$$\begin{aligned}
 & (\theta_h^{n+1} - \mathcal{E}_h^n \theta_h^n \circ \mathcal{A}_n^{-1}, J_n^{-1} v_h)_{0, \tilde{\Omega}_h^{n+1}} + \Delta t \mathcal{L}_h^{n+1}(\theta_h^{n+1}, v_h) \\
 &= \Delta t E_c^{n+1}(v_h), \quad \forall v_h \in V_h^{n+1},
 \end{aligned} \tag{4.92}$$

with consistency error

$$\begin{aligned}
 E_c^{n+1}(v_h) &= \underbrace{\frac{1}{\Delta t} (\overset{\leftarrow}{p}_h^{n+1} - p_h^n, \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \cap \tilde{\Omega}_h^n} - \frac{1}{\Delta t} (\overset{\leftarrow}{u}(t_{n+1}) - u(t_n), \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \cap \tilde{\Omega}_h^n}}_{E_{P_h,1}^{n+1}} \\
 &\quad + \underbrace{\frac{1}{\Delta t} (\overset{\leftarrow}{p}_h^{n+1} - \mathcal{E}_h^n p_h^n, \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n} - \frac{1}{\Delta t} (\overset{\leftarrow}{u}(t_{n+1}) - \mathcal{E}_h^n \bar{u}(t_n), \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}}_{E_{P_h,2}^{n+1}} \\
 &\quad + \underbrace{\frac{1}{\Delta t} (\mathcal{E}_h^n \bar{u}(t_n) - u(t_n), \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}}_{E_{\mathcal{E}_h}^{n+1}} \\
 &\quad + \underbrace{\frac{1}{\Delta t} (\overset{\leftarrow}{u}(t_{n+1}) - u(t_n), J_n \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n} - d(t_{n+1}; u, v_h)}_{E_{D_t,1}^{n+1}} \\
 &\quad + \underbrace{\frac{1}{\Delta t} (\overset{\leftarrow}{u}(t_{n+1}) - u(t_n), \overset{\leftarrow}{v}_h - J_n \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n}}_{E_{D_t,2}^{n+1}} \\
 &\quad + e_\perp^{n+1}(v_h)
 \end{aligned} \tag{4.93}$$

obtained by applying the change of variables in (4.75), adding and subtracting the terms $\frac{1}{\Delta t} (\overset{\leftarrow}{u}(t_{n+1}) - u(t_n), \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n}$, $\frac{1}{\Delta t} (\mathcal{E}_h^n \bar{u}(t_n), \overset{\leftarrow}{v}_h)_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}$ from (4.91), where $\bar{u} := u|_{\tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n}$.

Let us define the discrete non-cylindrical space-time domain, extending the definition of Q^T to the pulled-back domain $\bar{\Omega}_h^n$:

$$Q_h^T := \bigcup_{0 \leq n \leq N_T} \left(\bigcup_{s \in I_n} \bar{\Omega}_h^s \times \{s\} \right). \quad (4.94)$$

We are now ready to state the consistency result for the discrete error θ_h^n .

Proposition 13. *For all $t \in [0, T]$, assume that $u(\cdot, t) \in H^2(\Omega^t)$. Thus, the error functions θ_h^n, E_c^n satisfy,*

$$\begin{aligned} \|\theta_h^{N_T}\|_{0, \bar{\Omega}_h^{N_T}}^2 + \sum_{n=0}^{N_T-1} \|\bar{\theta}_h^{\leftarrow n+1} - \mathcal{E}_h^n \theta_h^n\|_{0, \bar{\Omega}_h^n}^2 + \alpha \sum_{n=0}^{N_T-1} \Delta t \|\theta_h^{n+1}\|_a^2 \\ \leq \Lambda(\mathbf{w}, T) \left(2\|\theta_h^0\|_{0, \bar{\Omega}_h^0}^2 + \frac{6}{\alpha} \sum_{n=0}^{N_T-1} \Delta t \|E_c^{n+1}\|_{(V_h^{n+1})'}^2 \right), \end{aligned} \quad (4.95)$$

with $c > 0$

$$\|\theta_h^0\|_{0, \bar{\Omega}_h^0} \leq c h^{3/2} |u^0|_{2, \bar{\Omega}_h^0}, \quad (4.96)$$

and assuming $D_t^2 u, D_t u \in L^2(Q_h^T)$, $D_t u \in L^2(0, T; H^2(\bar{\Omega}_h^t))$, and also $D_t P_h^t u \in L^1(0, T; L^2(\Omega^t))$, there exists positive constants C_1, C_2, C_3 independent of h and Δt such that

$$\begin{aligned} \sum_{n=0}^{N_T-1} \Delta t \|E_c^{n+1}\|_{(V_h^{n+1})'}^2 \leq C_1 h^3 \left(\|D_t u(s)\|_{L^2(0, T; H^2(\bar{\Omega}_h^s))}^2 + \|u\|_{C^0([0, T]; H^2(\bar{\Omega}_h^t))}^2 \right) \\ + C_2 h \|u\|_{C^0([0, T]; H^2(\Omega^t))}^2 \\ + C_3 \Delta t^2 \left(\|D_t u\|_{L^2(Q_h^T)}^2 + \|D_t^2 u\|_{L^2(Q_h^T)}^2 \right), \end{aligned} \quad (4.97)$$

considering that the meshsize and the time step satisfies

$$C h \leq \Delta t \quad (4.98)$$

with $C > 0$ independent on the meshsize.

Proof. Note that the discrete error equation in (4.92) satisfies Proposition 12 (stability of the fully discrete scheme). The error at the initial time step is bounded by applying (4.88):

$$\begin{aligned} \|\theta_h^0\|_{0, \bar{\Omega}_h^0} &= \|p_h^0 - u_h^0\|_{0, \bar{\Omega}_h^0} \leq \|p_h^0 - u^0\|_{0, \bar{\Omega}_h^0} + \underbrace{\|u^0 - u_h^0\|_{0, \bar{\Omega}_h^0}}_{=0 \text{ by CI}} \\ &\leq c h^{3/2} |u^0|_{2, \bar{\Omega}_h^0}. \end{aligned}$$

Furthermore, the second term is bounded by recalling the definition of the dual norm:

$$\|E_c^{n+1}\|_{(V_h^{n+1})'} = \sup_{v_h \in V_h^{n+1}} \frac{|E_c^{n+1}(v_h)|}{\|v_h\|_a},$$

and we now proceed to bound each term of (4.93). To simplify the notation, let us set

$$J_{max} := \max_{s \in I_n} \|J(t_n, s)\|_{L^\infty(\tilde{\Omega}_h^s)}.$$

Assuming that $D_t P_h^t u \in L^1(0, T; L^2(\Omega^t))$, the first term is a projection error at time $n+1$ and is bounded using the fundamental theorem of calculus, exploiting the commutativity between the projection operator P_h^t and the time derivative. Hence, we obtain

$$\begin{aligned} |E_{P_h,1}^{n+1}| &= \frac{J_{max}}{\Delta t} \left| \int_{I_n} ([P_h^t - I] D_t \tilde{u}(s), \tilde{v}_h)_{0, \tilde{\Omega}_h^s \cap \tilde{\Omega}_h^s} ds \right| \\ &\lesssim \frac{1}{\Delta t} \int_{I_n} \|P_h^t D_t \tilde{u}(s) - D_t \tilde{u}(s)\|_{0, \tilde{\Omega}_h^s \cap \tilde{\Omega}_h^s} ds \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}, \\ &\lesssim \frac{1}{\Delta t} h^{3/2} \int_{I_n} |D_t \tilde{u}(s)|_{2, \tilde{\Omega}_h^s \cap \tilde{\Omega}_h^s} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}, \end{aligned}$$

where we have applied (4.88) in the last step. The second term of the consistency error is bounded in a similar manner, noting that $\mathcal{E}_h^{n+1} \phi = \phi$ in $\tilde{\Omega}_h^{n+1}$. Thus, we have

$$\begin{aligned} |E_{P_h,2}^{n+1}| &= \frac{J_{max}}{\Delta t} \left| \int_{I_n} (\mathcal{E}_h^t \{[P_h^t - I]\} D_t \tilde{u}(s), \tilde{v}_h)_{0, \tilde{\Omega}_h^s \setminus \tilde{\Omega}_h^s} ds \right| \\ &\lesssim \frac{1}{\Delta t} \int_{I_n} \|\mathcal{E}_h^t (P_h^t D_t \tilde{u}(s) - D_t \tilde{u}(s))\|_{0, \tilde{\Omega}_h^s \setminus \tilde{\Omega}_h^s} ds \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}, \\ &\leq \frac{C_{\mathcal{E}_h}^*}{\Delta t} h^{3/2} \int_{I_n} |D_t \tilde{u}(s)|_{2, \tilde{\Omega}_h^s} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}, \end{aligned}$$

where we have used (4.63) with $C_{\mathcal{E}_h}^* := \max_{I_n} C_{\mathcal{E}_h}$. The term $E_{\mathcal{E}_h}^{n+1}$ is bounded using (4.63), (4.67), and (4.88) yielding

$$\begin{aligned} |E_{\mathcal{E}_h}^{n+1}| &\leq \frac{J_{max}}{\Delta t} \left(\|\mathcal{E}_h^n [\bar{u}(t_n) - p_h^n]\|_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n} + \|\mathcal{E}_h^n p_h^n - u(t_n)\|_{0, \tilde{\Omega}_h^n \setminus \tilde{\Omega}_h^n} \right) \|v_h\|_{0, \tilde{\Omega}_h^{n+1}} \\ &\lesssim \frac{1}{\Delta t} \left(C_{\mathcal{E}_h} h^{3/2} |u(t_n)|_{2, \tilde{\Omega}_h^n} + h^2 |u(t_n)|_{2, \Omega^n} \right) \|v_h\|_{0, \tilde{\Omega}_h^{n+1}} \\ &\lesssim h^{1/2} |u(t_n)|_{2, \Omega^n} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}, \end{aligned}$$

where, in the last step, we have used the assumption $h \lesssim \Delta t$. By applying Taylor's

formula with the Lagrange remainder, we can bound the fourth term, obtaining

$$\begin{aligned}
 |E_{D_t,1}^{n+1}| &= \frac{J_{max}}{\Delta t} \left| \int_{I_n} (s - t_n) (D_s^{2\leftarrow} u(s), \leftarrow v_h)_{0, \tilde{\Omega}_h^s} ds \right| \\
 &\lesssim \frac{\Delta t^{3/2}}{\sqrt{3}\Delta t} \left(\int_{I_n} \|D_t^{2\leftarrow} u(s)\|_{0, \tilde{\Omega}_h^s}^2 ds \right)^{1/2} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}} \\
 &\lesssim \Delta t^{1/2} \left(\int_{I_n} \|D_t^{2\leftarrow} u(s)\|_{0, \tilde{\Omega}_h^s}^2 ds \right)^{1/2} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}}.
 \end{aligned}$$

The bound on $E_{D_t,2}^{n+1}$ follows from another application of the fundamental theorem of calculus. Indeed,

$$\begin{aligned}
 |E_{D_t,2}^{n+1}| &= \frac{J_{max}}{\Delta t} \left| \int_{I_n} (D_t^{\leftarrow} u(s), \leftarrow v_h(1 - J_n))_{0, \tilde{\Omega}_h^s} ds \right| \\
 &\lesssim \frac{\Delta t^{1/2}}{\Delta t} \|1 - J_n\|_{L^\infty(\tilde{\Omega}_h^s)} \left(\int_{I_n} \|D_t^{\leftarrow} u(s)\|_{0, \tilde{\Omega}_h^s}^2 ds \right)^{1/2} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}} \\
 &\lesssim \Delta t^{1/2} \left(\int_{I_n} \|D_t^{\leftarrow} u(s)\|_{0, \tilde{\Omega}_h^s}^2 ds \right)^{1/2} \|v_h\|_{0, \tilde{\Omega}_h^{n+1}},
 \end{aligned}$$

where, recalling (4.9) and the Maclaurin expansion of the exponential function,

$$\|1 - J_n\|_{L^\infty(\tilde{\Omega}_h^s)} = \left\| 1 - \exp \left(\int_{t_n}^s \nabla \cdot \mathbf{w} \right) \right\|_{L^\infty(\tilde{\Omega}_h^s)} \leq \left\| - \int_{t_n}^s \nabla \cdot \mathbf{w} \right\|_{L^\infty(\tilde{\Omega}_h^s)} \lesssim \Delta t.$$

Finally, the orthogonality error $e_\perp^{n+1}(v_h)$ is bounded as in (3.36). Combining these

bounds, and recalling that $\|v_h\|_{0,\tilde{\Omega}_h^{n+1}} \lesssim \|v_h\|_a$, we obtain

$$\begin{aligned}
 & \sum_{n=0}^{N_T-1} \Delta t \|E_c^{n+1}\|_{(V_h^{n+1})'}^2 \\
 & \leq \sum_{n=0}^{N_T-1} \frac{\Delta t}{\|v_h\|_a^2} \left(|E_c^{n+1}(v_h)|^2 \right) \\
 & \lesssim \sum_{n=0}^{N_T-1} \left[\frac{h^3}{\Delta t} \left(\int_{I_n} |D_t u(s)|_{2,\tilde{\Omega}_h^s \cap \tilde{\Omega}_h^s} \right)^2 + \frac{h^3}{\Delta t} \left(\int_{I_n} |D_t u(s)|_{2,\tilde{\Omega}_h^s} \right)^2 + \Delta t h |u(t_n)|_{2,\Omega^n}^2 \right. \\
 & \quad \left. + \Delta t^2 \left(\int_{I_n} \|D_t^2 u(s)\|_{0,\tilde{\Omega}_h^s}^2 + \int_{I_n} \|D_t u(s)\|_{0,\tilde{\Omega}_h^s}^2 \right) + \Delta t h^3 |u(t_{n+1})|_{2,\tilde{\Omega}_h^{n+1}}^2 \right] \\
 & \lesssim \sum_{n=0}^{N_T-1} \left[h^3 \int_{I_n} |D_t u(s)|_{2,\tilde{\Omega}_h^s}^2 + \Delta t h |u(t_n)|_{2,\Omega^n}^2 \right. \\
 & \quad \left. + \Delta t^2 \left(\int_{I_n} \|D_t^2 u(s)\|_{0,\tilde{\Omega}_h^s}^2 + \int_{I_n} \|D_t u(s)\|_{0,\tilde{\Omega}_h^s}^2 \right) + \Delta t h^3 |u(t_{n+1})|_{2,\tilde{\Omega}_h^{n+1}}^2 \right] \\
 & \lesssim h^3 \int_0^T |D_t u(s)|_{2,\tilde{\Omega}_h^s}^2 + T h \max_{0 \leq n < N_T} |u(t_n)|_{2,\Omega^n}^2 \\
 & \quad + \Delta t^2 \left(\int_0^T \|D_t^2 u(s)\|_{0,\tilde{\Omega}_h^s}^2 + \int_0^T \|D_t u(s)\|_{0,\tilde{\Omega}_h^s}^2 \right) + T h^3 \max_{0 \leq n < N_T} |u(t_{n+1})|_{2,\tilde{\Omega}_h^{n+1}}^2,
 \end{aligned}$$

yielding the desired result. $\square$

Remark 17. Note that, to prove the stability estimate in (4.85), we introduced Assumption 7, resulting in a time scheme that is not *unconditionally stable*, even though it has a structure very similar to the Backward Euler method. Moreover, the consistency proof required a stronger condition, given in (4.98). Nevertheless, although the numerical results presented in Section 4.6 were obtained with $h \sim \Delta t$, we have tested several different configurations which suggest that the fully discrete SBM-ALE scheme is actually *unconditionally stable* in practice. Naturally, this aspect requires further investigation.

4.5.3 Convergence

Once we have provided the stability and the consistency of the fully discrete SBM-ALE method with dG(0) scheme, we have all the ingredients to estimate the errors $\{e^n\} = \{u(t_n) - u_h^n\}$ introduced in (4.89) and prove the convergence.

To this end, it is enough to apply triangle inequality of norms to the decomposition (4.90). As we said before, the errors $\{\pi^n\}$ are estimated using (4.87) or (4.88), whereas the errors $\{\theta_h^n\}$ are estimated in Proposition 13. In this way, we obtain the following result.

Theorem 4. *Under previous assumptions in Definition 5 with (4.87-4.88), Proposition 13, let $\{u_h^n\}$ be defined by the ALE-SBM fully discrete scheme in (4.72). Then, there exists a constant $C(u_0, u, D_t u, D_t^2 u, f, g)$ dependent on its arguments but independent of meshsize h and time step Δt such that*

$$\max_{0 \leq n < N_T} \|e^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}} + \alpha \sum_{n=0}^{N_T-1} \Delta t \|e^{n+1}\|_a \leq C(u_0, u, D_t u, D_t^2 u, f, g) (h^{1/2+\zeta} + \Delta t) . \quad (4.99)$$

Proof. Combining Stability, Consistency, and taking the maximum of the term $\|e^{n+1}\|_{0, \tilde{\Omega}_h^{n+1}}$ we get the desired result. We write $h^{1/2+\zeta}$ recalling the suboptimality we have already discussed in Remark 7 and 8. $\square$

It is important to emphasize that the suboptimality is strictly related to the SBM and is independent of the time discretization or Rothe's approach. Nevertheless, as observed in the stationary case, the numerical experiments demonstrate optimal rates of convergence.

Therefore, since we are approximating the material derivative with the dG(0) scheme, assuming $\Delta t \sim h$, the ALE-SBM method results in a first-order method. Neglecting the error term computed with the a -norm, even if the spatial L^2 error has the enhanced order of convergence presented in (4.88), the total error accounting for the time discretization is still first-order. It can be proven that, in order to obtain the optimal or enhanced error in the L^2 -norm, a higher-order time scheme such as Crank-Nicolson should be employed.

4.6 Numerical Results

In this section, we present numerical results to validate the theoretical bounds established in this chapter. In particular, we analyze the total error measured in the Bochner norms introduced in Chapter 3. As in Section 3.3, all possible combinations are evaluated: specifically, we present both the C^0 -in-time error (taking the maximum over time instants) and the discrete L^2 -in-time error (computed as the sum of the errors at each time step weighted by Δt). The spatial errors are measured in the L^2 and H_0^1 norms. Hence we compute the discrete version of (4.14) and (4.13).

First of all, we present numerical tests for a stationary problem to evaluate the extrapolation error $\|\mathcal{E}_h u_h - u\|_{0, S_h}$ within a strip S_h , defined as the set of elements within a distance h of the surrogate boundary $\tilde{\Gamma}_h$.

Furthermore, The numerical simulations are performed on a moving square, an L-shaped moving domain, and a rotating ellipse. For all these tests, we prescribe the Eulerian velocity $\mathbf{w}$ a priori, although we recall that in practical applications,

$\mathbf{w}$ is typically computed as discussed in Remark 14. The meshsize and the time step are set with the same order of magnitude, that is to say $\Delta t \sim h$.

4.6.1 Extrapolation error

In this section, the extrapolation error within a strip of width $O(h)$ around the surrogate domain is investigated. Specifically, the numerical test is performed for a stationary boundary value problem, as described in (2.5), in the square domain $\Omega = (0,1)^2$. The strip $S_h = \Omega_\delta \setminus \tilde{\Omega}_h$ is constructed as the set of elements contained in an auxiliary domain $\Omega_\delta = (0,1+\delta)^2$, where δ is chosen as a small fraction of the meshsize h , for instance, $\delta = \frac{h}{100}$.

A comparison between the extended numerical solution $\mathcal{E}_h u_h$ and the exact solution u is presented in Figure 4.6, where it can be observed that the profile of the numerical solution is perfectly preserved.

Within the strip S_h , the L^2 -error between the extended numerical solution and the exact solution u is computed, confirming the theoretical results in (4.67).

In Table 4.1, we observe a super-quadratic rate of convergence. This behavior depends on the fact that the measure of S_h is $O(h)$, which confirms the theoretical predictions. Indeed, one has

$$\|\mathcal{E}_h u_h - u\|_{0,S_h}^2 = \int_{S_h} |\mathcal{E}_h u_h - u|^2 dx \leq |S_h| \max_{x \in S_h} |\mathcal{E}_h u_h - u|^2 \leq Ch^5 |u|_{2,\Omega}^2.$$

Analogous considerations hold for the logarithmic error regression lines displayed in Figure 4.7.

h_{input}	1/20	1/40	1/80	1/160
h	0.0499	0.0250	0.0125	0.0062
$\ \mathcal{E}_h u_h - u\ _{0,S_h}$	1.18×10^{-2}	1.60×10^{-3}	2.81×10^{-4}	4.87×10^{-5}
	—	2.88	2.52	2.53

Table 4.1: L^2 Extension errors and convergence rates.

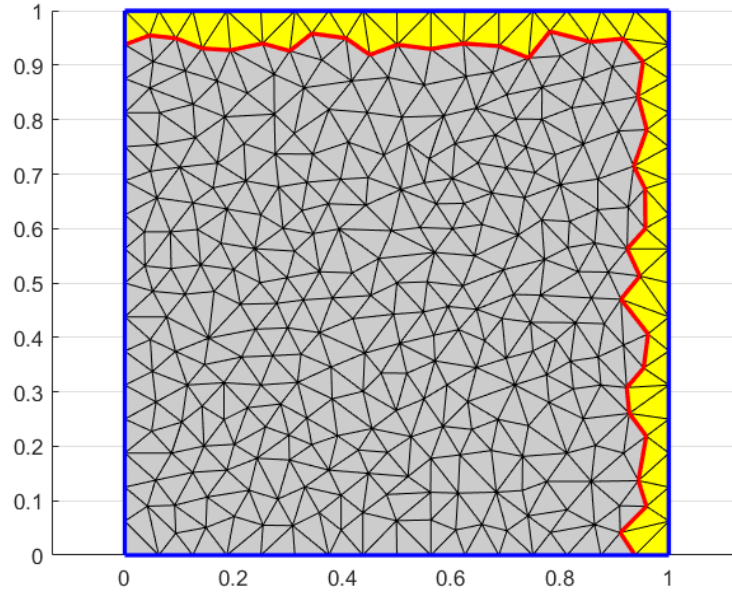


Figure 4.5: Surrogate domain (grey) and Extended strip S_h (yellow).

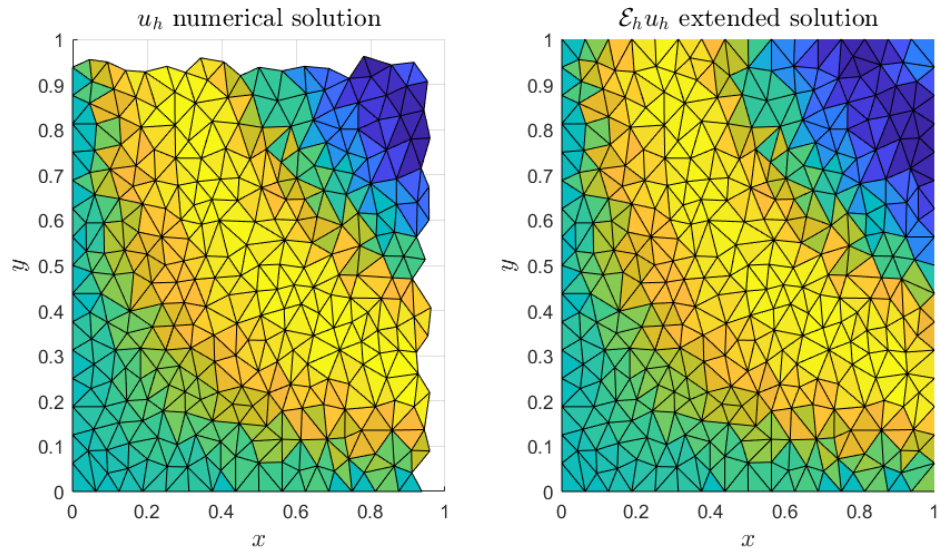


Figure 4.6: Comparison between the numerical solution u_h and his extension $\mathcal{E}_h u_h$ for $h = 0.0499$.

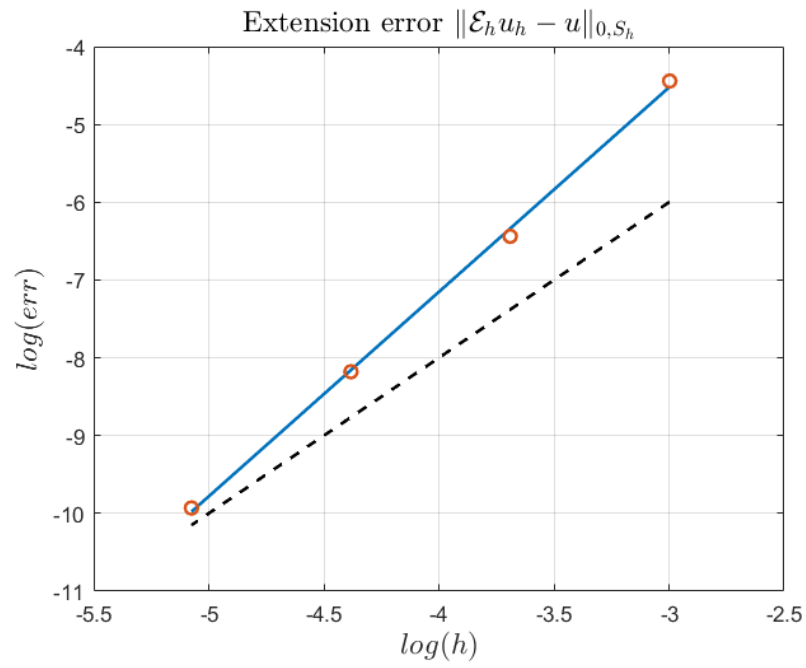


Figure 4.7: Logarithmic L^2 -error regressions of $\mathcal{E}_h u_h - u$ and his corresponding theoretical convergence line (Rate 2.6285).

4.6.2 Square domain

The first numerical test is performed on a square domain expanding with the Eulerian velocity

$$\mathbf{w} = \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty}, \quad (4.100)$$

for $t \in [0, \frac{1}{2}]$, starting from the initial domain $\Omega^0 = (0, \frac{1}{2})^2$ and reaching the final domain $\Omega^T = (0, 1)^2$. The manufactured exact solution is given by

$$u = \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{\pi y}{L}\right), \quad \text{with } L = \frac{1}{2} + t,$$

and the forcing term f and boundary data g are computed accordingly. Note that u is chosen to vanish on the boundary at each time step; thus, homogeneous Dirichlet boundary conditions are imposed. The Nitsche penalty parameter is set to 10.

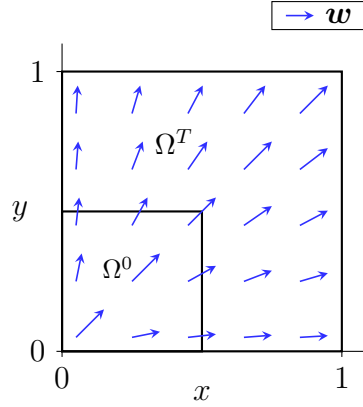


Figure 4.8: Square domain and its motion according to the velocity field $\mathbf{w}$.

Figure 4.9 presents a comparison between the numerical and exact solutions at each time step. The time step is chosen here solely to facilitate the visualization of the solutions. However, during the mesh refinement study, it will be appropriately scaled such that $\Delta t \sim h$.

In Table 4.2, the errors and their corresponding experimental orders of convergence (EOC) are reported for all possible combinations of Bochner norms. As expected, a first-order convergence rate is achieved in the final refinement steps across all norms. It can be observed that the EOC for the Bochner norms involving the spatial L^2 -norm approaches 1 from above, starting from initial values close to 2. This behavior is likely due to the fact that the spatial error in the L^2 -norm converges quadratically; however, when coupled with the first-order dG(0) scheme, the overall asymptotic convergence rate is ultimately dominated by the time discretization, thus reducing to first order. Conversely, the EOC for the Bochner norms involving

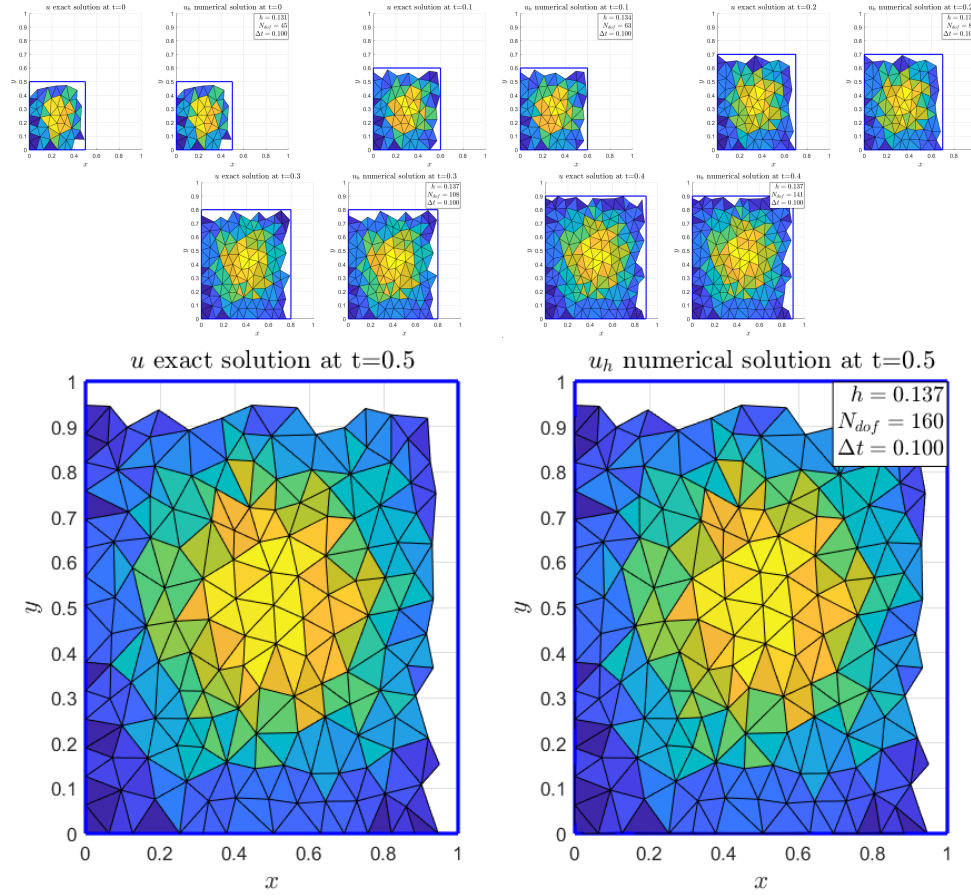


Figure 4.9: [Moving Square] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137, \Delta t = 0.1$.

the spatial H_0^1 -norm approaches 1 from below, as the spatial error in the H_0^1 -norm intrinsically converges linearly.

The logarithmic regression of the errors, shown in Figure 4.10, illustrates nearly first-order convergence rates overall, since the linear fit is influenced by the pre-asymptotic behavior of the initial refinements. Nevertheless, the values at the finest refinement levels are consistent with the preceding discussion and perfectly match the theoretical convergence rates for the H^1 Bochner norms.

h_{input}	1/8	1/16	1/32	1/64
h	0.2479	0.1207	0.0645	0.0345
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h^t))}$	4.41×10^{-2}	1.21×10^{-2}	3.88×10^{-3}	1.84×10^{-3}
	–	1.80	1.81	1.19
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h^t))}$	5.52×10^{-1}	3.33×10^{-1}	1.84×10^{-1}	9.80×10^{-2}
	–	0.70	0.94	1.01
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h^t))}$	2.40×10^{-2}	7.15×10^{-3}	2.49×10^{-3}	1.09×10^{-3}
	–	1.68	1.69	1.31
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h^t))}$	3.22×10^{-1}	1.82×10^{-1}	9.71×10^{-2}	5.13×10^{-2}
	–	0.80	1.00	1.02
Total error	3.67×10^{-1}	1.94×10^{-1}	1.01×10^{-1}	5.31×10^{-2}
	–	0.89	1.04	1.02

Table 4.2: [Moving Square] Errors for the dG(0) scheme and successive mesh refinements.

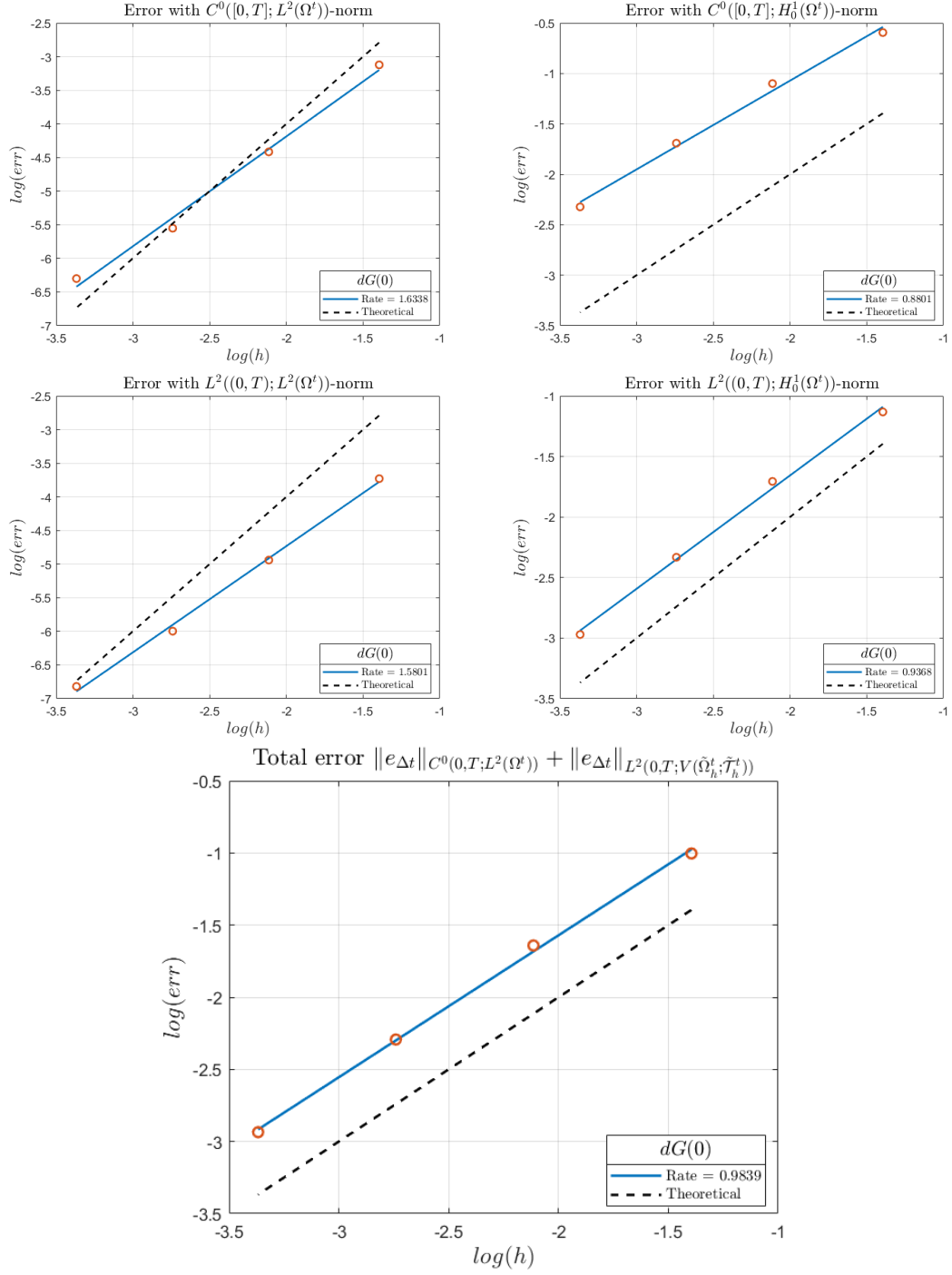


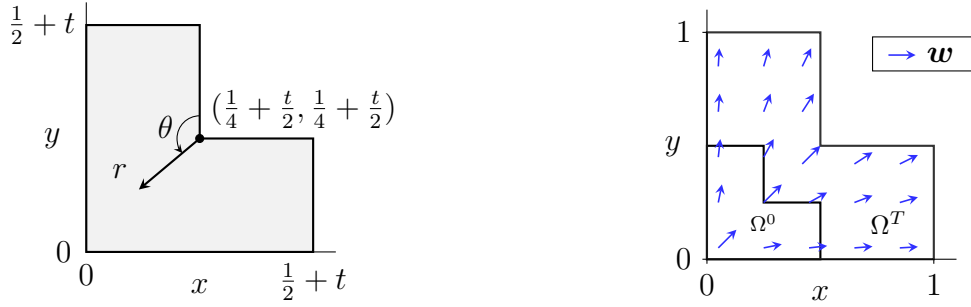
Figure 4.10: [Moving Square] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with $dG(0)$.

4.6.3 L-shaped domain

In this section, the L-shaped domain test case is analyzed. Specifically, we extend the numerical test for stationary problems presented in [21] to the case of moving domains. The Eulerian velocity is prescribed and coincides with that of the square domain test described in (4.100). During the time interval $t \in [0, \frac{1}{2}]$, the domain evolves from $\Omega^0 = (0, \frac{1}{2})^2 \setminus (\frac{1}{4}, \frac{1}{2})^2$ to $\Omega^T = (0, 1)^2 \setminus (\frac{1}{2}, 1)^2$. The solution is chosen to exhibit a singularity at the reentrant corner of the L-shaped domain, ensuring that it belongs to $H^{1+\alpha-\epsilon}(\Omega^t)$ for any $\epsilon > 0$ and for all t . Letting $\omega = \frac{3}{2}\pi$ denote the angle of the reentrant corner, we define $\alpha = \frac{\pi}{\omega}$. In local polar coordinates, the exact solution reads

$$u = \sin(t) r^\alpha \sin(\alpha\theta),$$

where $\sin(t)$ acts as a time-dependent component, while the term $r^\alpha \sin(\alpha\theta)$ is responsible for the spatial singularity. We recall that, as illustrated in Figure 4.11 (a), the local coordinate system (r, θ) is centered at the reentrant corner at every time t . The Nitsche penalty parameter is set to 10.



(a) Polar coordinates in the L-shaped domain.

(b) Expanding L-shaped domain.

Figure 4.11: L-shaped domain: polar coordinates and its motion according to the velocity field w .

A comparison between the numerical and exact solutions is presented in Figure 4.12. At each time step, the method provides a satisfactory approximation of the solution, despite the relatively coarse mesh and time step.

Table 4.3 reports the computed errors and the expected orders of convergence. The theoretical rates for a stationary L-shaped domain are $\alpha \simeq 0.67$ for the H^1 -seminorm and $2\alpha \simeq 1.33$ for the L^2 -norm. At the finest mesh refinement level, we confirm the theoretical convergence rates, bearing in mind that the time semi-discretization error converges linearly. Consequently, we correctly observe a convergence rate of 0.66 for the $C^0([0, T]; H_0^1(\tilde{\Omega}_h^t))$ and $L^2(0, T; H_0^1(\tilde{\Omega}_h^t))$ norms,

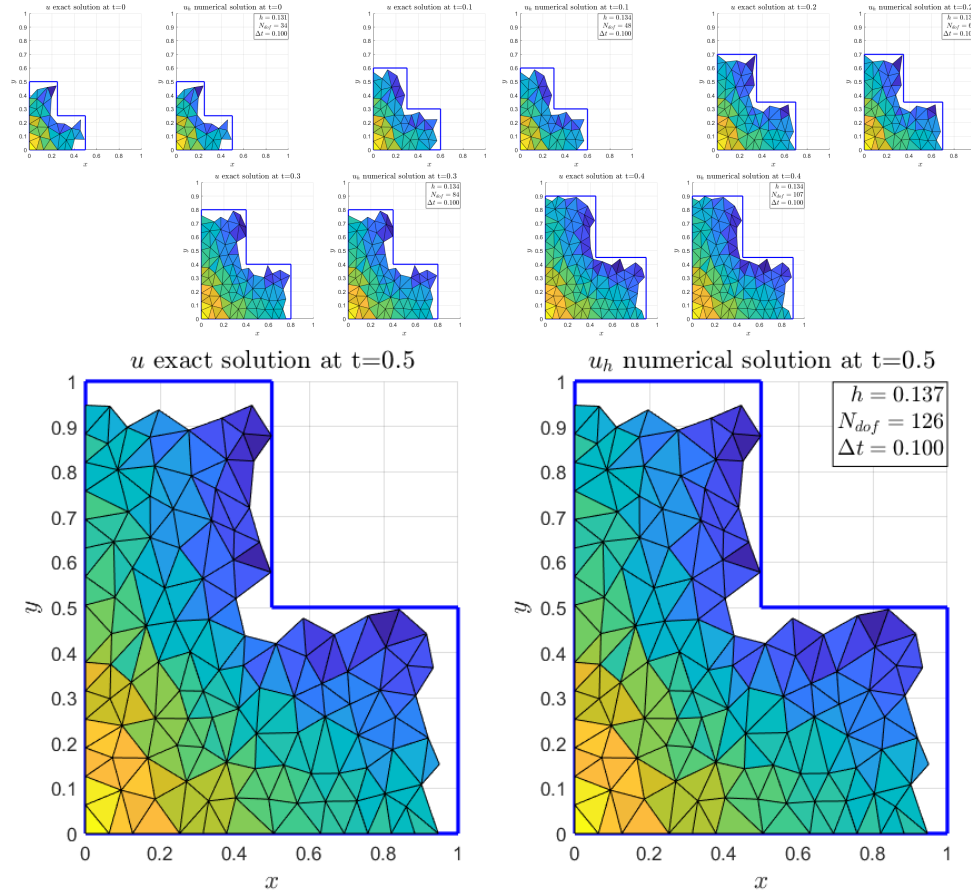


Figure 4.12: [Moving L-shaped domain] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137, \Delta t = 0.1$.

and a rate of approximately 1 for the total error, as well as for the $C^0([0, T]; L^2(\tilde{\Omega}_h^t))$ and $L^2(0, T; L^2(\tilde{\Omega}_h^t))$ norms.

Figure 4.13 displays the logarithmic regression of the errors. The computed slopes do not perfectly match the theoretical asymptotic rates, as they are affected by the pre-asymptotic behavior of the initial mesh refinements.

h_{input}	1/16	1/32	1/64	1/128
h	0.1155	0.0636	0.0345	0.0170
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h^t))}$	7.16×10^{-4}	5.19×10^{-4}	3.85×10^{-4}	1.56×10^{-4}
	–	0.54	0.49	1.28
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h^t))}$	2.69×10^{-2}	2.38×10^{-2}	1.78×10^{-2}	1.11×10^{-2}
	–	0.21	0.47	0.66
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h^t))}$	3.98×10^{-4}	2.23×10^{-4}	1.33×10^{-4}	6.04×10^{-5}
	–	0.97	0.84	1.12
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h^t))}$	1.25×10^{-2}	9.01×10^{-3}	6.15×10^{-3}	4.01×10^{-3}
	–	0.55	0.62	0.61
Total error	1.32×10^{-2}	9.53×10^{-3}	6.53×10^{-3}	4.16×10^{-3}
	–	0.55	0.62	0.64

Table 4.3: [Moving L-shaped domain] Errors for the dG(0) scheme and successive mesh refinements.

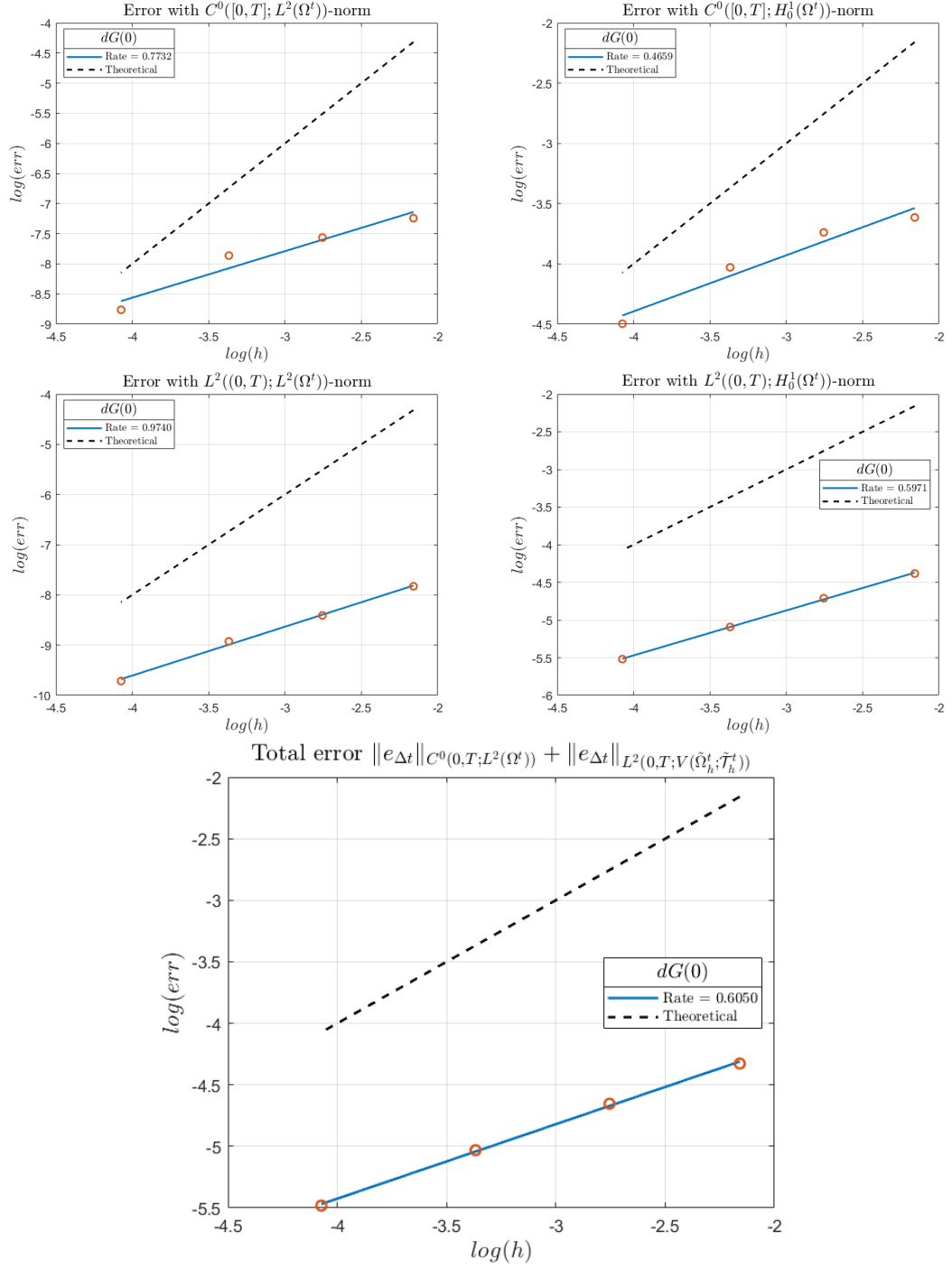


Figure 4.13: [Moving L-shaped domain] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with $dG(0)$.

4.6.4 Rotating ellipse

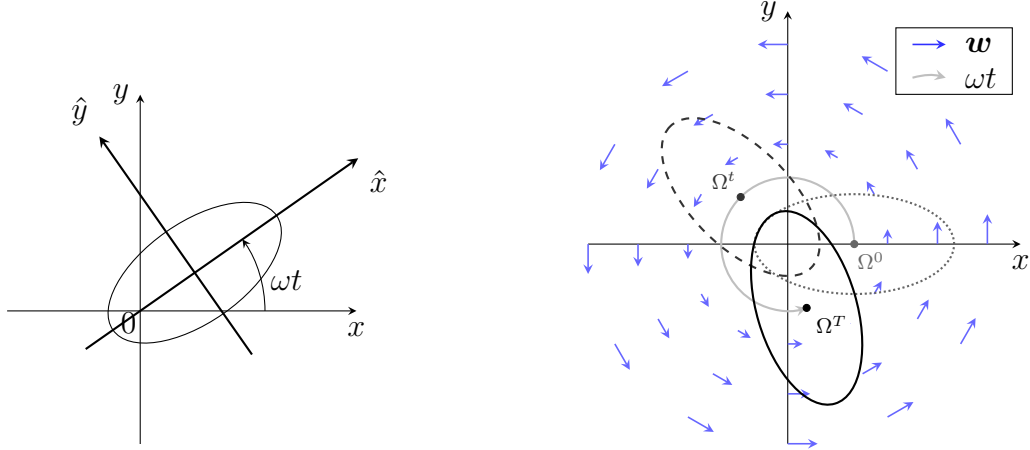
This final test considers an elliptic domain undergoing a rigid rotation about the origin. During the time interval $t \in [0,1]$, the ellipse, initially centered at $(x_0, y_0) = (\frac{1}{5}, 0)$ with semi-axes $a = \frac{1}{2}$ and $b = \frac{1}{4}$, rotates with an angular velocity $\omega = 5$. The local coordinates $(\hat{x}, \hat{y})$, attached to the moving center of the ellipse, are given by

$$\begin{cases} \hat{x} = x \cos(\omega t) + y \sin(\omega t) - x_0, \\ \hat{y} = -x \sin(\omega t) + y \cos(\omega t) - y_0, \end{cases}$$

and, consequently, the exact solution reads

$$u = t^2 \cos \left[\frac{\pi}{2} \left(\frac{\hat{x}^2}{a^2} + \frac{\hat{y}^2}{b^2} \right) \right].$$

Since the solution vanishes on the moving boundary, homogeneous Dirichlet boundary conditions are imposed. The forcing term is then computed by substituting u into the heat equation. The Nitsche penalty parameter is set to 10.



(a) Local coordinates on the rotated ellipse.

Figure 4.14: Rotating elliptic domain: local coordinate system tracked along the center's trajectory (left) and rigid rotation around the origin alongside its velocity field $\mathbf{w}$ (right).

The qualitative comparison between the exact and numerical solutions is illustrated in Figure 4.15. The numerical scheme successfully tracks the rigid motion of the domain at each time step, yielding an accurate approximation even on a relatively coarse spatio-temporal grid.

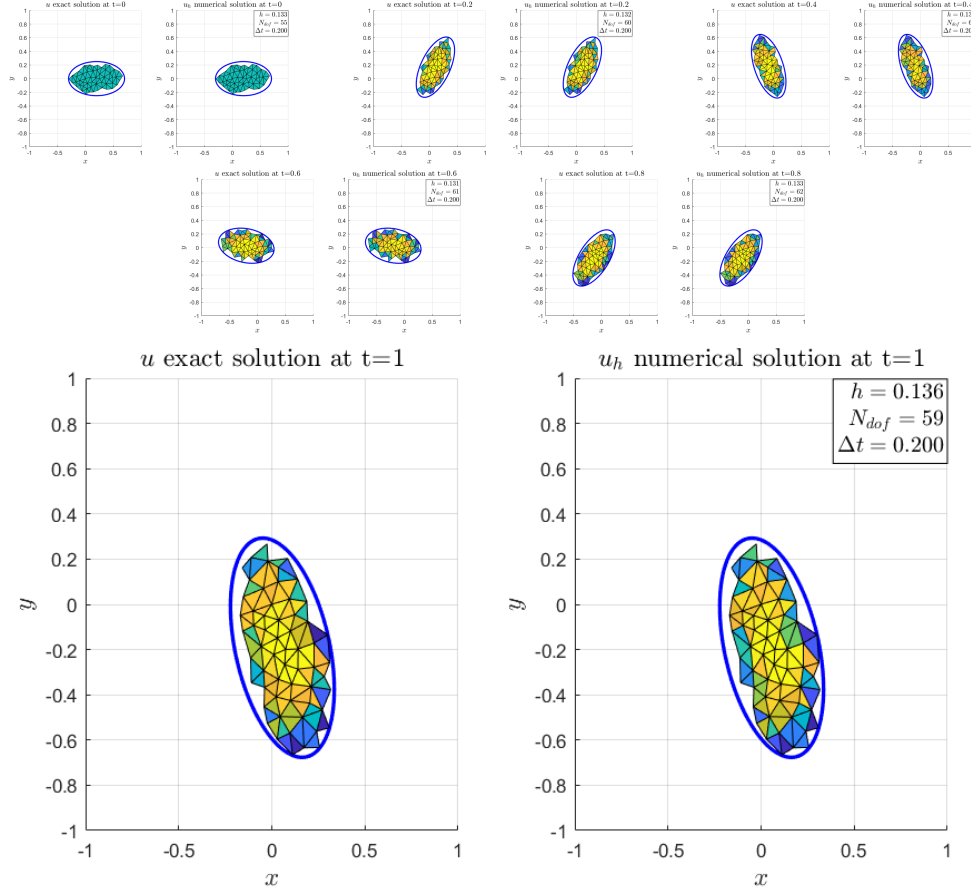


Figure 4.15: [Rotating Ellipse] Comparison at each time step between the exact solution u and the numerical solution u_h using dG(0) for $h = 0.137$, $\Delta t = 0.2$.

Table 4.4 reports the computed errors and the expected orders of convergence. At the finest mesh refinement level, we confirm the theoretical convergence rates, bearing in mind that the time semi-discretization error converges linearly. Consequently, we correctly observe a convergence rate of 1 for the H^1 Bochner norms. It is worth noting that, as in the square domain test case, the theoretical quadratic spatial convergence rate of the Bochner norms involving the L^2 error is dominated by the first-order accuracy of the dG(0) scheme.

Figure 4.16 illustrates the logarithmic regression lines of the computed errors. The empirical slopes slightly deviate from the exact theoretical asymptotic rates, as they are influenced by the pre-asymptotic regime associated with the coarser initial meshes.

h_{input}	1/10	1/20	1/40	1/80
h	0.1992	0.1021	0.0499	0.0255
$\ e_{\Delta t}\ _{C^0([0,T];L^2(\tilde{\Omega}_h^t))}$	1.32×10^{-1}	2.91×10^{-2}	8.44×10^{-3}	2.79×10^{-3}
	–	2.26	1.72	1.65
$\ e_{\Delta t}\ _{C^0([0,T];H_0^1(\tilde{\Omega}_h^t))}$	8.78×10^{-1}	5.17×10^{-1}	2.73×10^{-1}	1.44×10^{-1}
	–	0.79	0.89	0.95
$\ e_{\Delta t}\ _{L^2(0,T;L^2(\tilde{\Omega}_h^t))}$	6.92×10^{-2}	1.32×10^{-2}	3.81×10^{-3}	1.24×10^{-3}
	–	2.48	1.73	1.68
$\ e_{\Delta t}\ _{L^2(0,T;H_0^1(\tilde{\Omega}_h^t))}$	5.33×10^{-1}	2.41×10^{-1}	1.30×10^{-1}	6.65×10^{-2}
	–	1.19	0.86	1.00
Total error	6.65×10^{-1}	2.70×10^{-1}	1.38×10^{-1}	6.93×10^{-2}
	–	1.35	0.93	1.03

Table 4.4: [Rotating elliptic domain] Errors for the dG(0) scheme and successive mesh refinements.

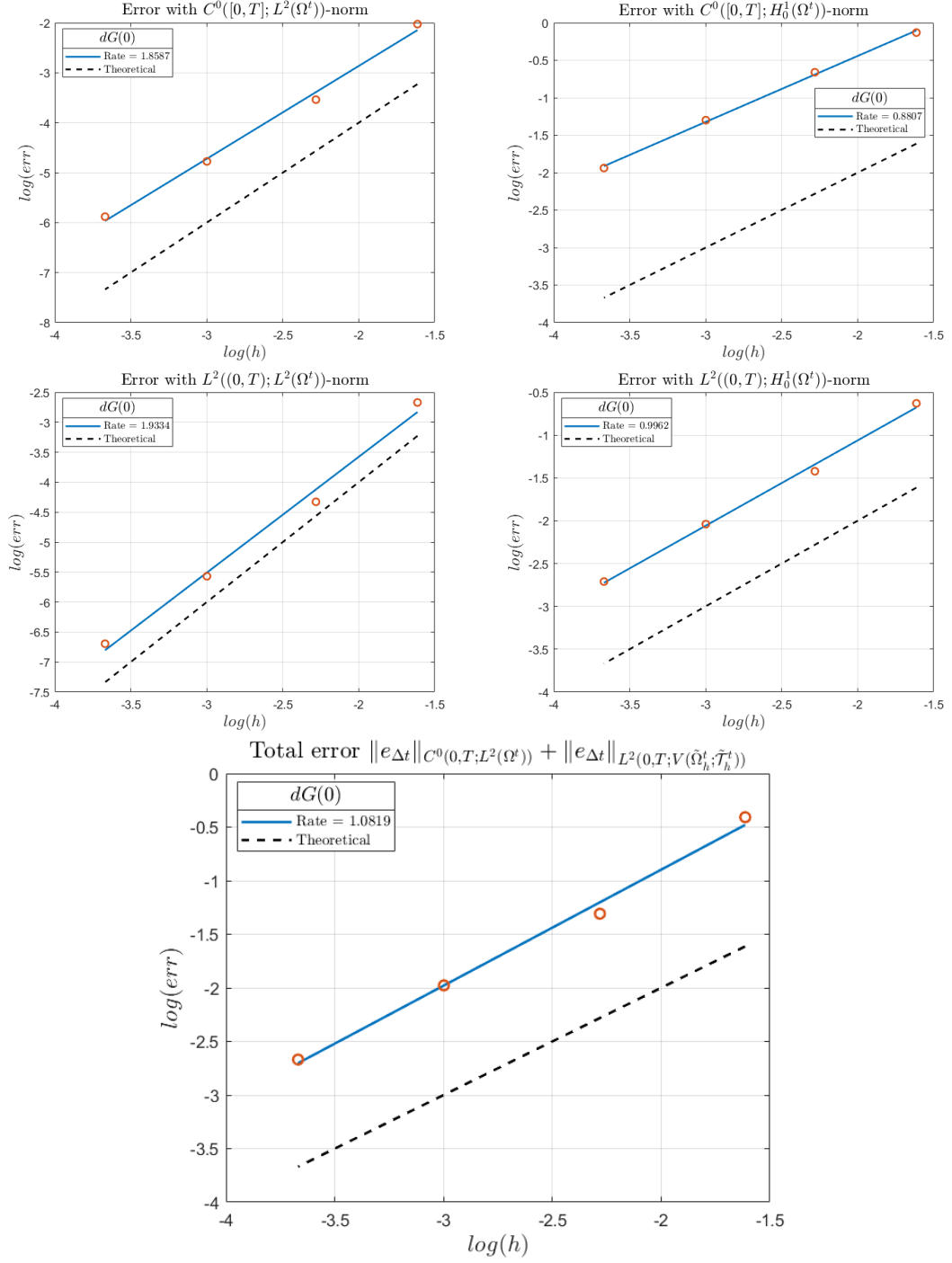


Figure 4.16: [Rotating Ellipse] Logarithmic error regressions of $u - u_h$ and their corresponding theoretical convergence lines with $dG(0)$.

Chapter 5

Conclusion

In this thesis, we have investigated and analyzed in detail the application of the Shifted Boundary Method (SBM) to the solution of differential problems, with a particular focus on the heat equation in stationary, time-dependent, and moving domains. The main objective was to provide a rigorous analysis of the stability, consistency, and convergence of the method, supported by an extensive numerical validation presented at the end of each chapter.

Summary of the Main Results

The work was structured across three levels of increasing complexity, yielding the following key results:

- **Stationary Problems:** We proved the well-posedness of the discrete problem and derived rigorous stability estimates. The analysis of the consistency error confirmed optimal convergence rates in the energy norm and provided an enhanced L^2 -norm estimate. Furthermore, we introduced and numerically validated an effective map to compute the shift for smooth boundaries, with a specific application to elliptical domains.
- **Time-Dependent Problems:** Extending the stationary results, we addressed the heat equation using the method of lines. By first applying a spatial semi-discretization, we integrated the resulting ordinary differential systems using the Backward Euler and Crank-Nicolson schemes. For the SBM-Backward Euler scheme, formal proofs of stability, consistency, and convergence were rigorously provided, and subsequently confirmed experimentally for various combinations of the spatial parameter h and the time step Δt .
- **Moving Domain Problems:** The main challenge of the thesis involved coupling the Arbitrary Lagrangian-Eulerian (ALE) framework with the SBM.

Adopting Rothe’s approach (transverse method of lines), we first applied the time discretization using dG(0) and subsequently solved the spatial problem via SBM at each time step. A highly significant theoretical result was the derivation of a stability estimate for the infinite-dimensional problem which, under homogeneous boundary conditions, is *ALE-free*, i.e., independent of the Eulerian velocity. This property is preserved in the ideal time discretization, confirming the robustness of the proposed approach. Moreover, we defined a discrete extension operator in order to properly state the fully discrete SBM-ALE problem. By leveraging two important properties of this operator, we proved the stability and consistency of the method, and consequently established its convergence. Finally, various numerical tests are presented to corroborate these theoretical results.

Advantages of the SBM-ALE Approach

The theoretical and numerical results obtained in this work confirm that the SBM represents an extremely powerful and flexible tool. As an *unfitted* method, the computational grid does not need to conform to the true domain, drastically reducing the computational costs associated with mesh generation, especially when dealing with complex geometries featuring corners or holes.

This characteristic proved to be a crucial advantage in moving domain problems: unlike classical ALE methods, which require mandatory and expensive *re-meshing* processes to prevent element stretching and the subsequent deterioration of mesh quality, the SBM-ALE framework developed in this thesis completely circumvents this necessity. The background mesh can be kept entirely fixed, requiring only the update of the map and the shift at each time $t \in [0, T]$.

Future Work

Despite the encouraging results, the present work paves the way for several future research directions:

- **Consistency Analysis:** Regarding the consistency analysis in Proposition 13, the assumption $D_t P_h^t u \in L^1(0, T; L^2(\Omega^t))$ requires further investigation. In particular, the properties of the operator P_h^t should be analyzed more thoroughly, supported by specific numerical tests.
- **Stability Conditions:** The assumption $ch \leq \Delta t$ appears highly restrictive, although this limitation has not been observed in our numerical experiments. Moreover, it is quite unusual that the assumption required to prove stability in (4.83) instead dictates $ch^2 \leq \Delta t$. Since the dG(0) time discretization is

structurally very similar to the Backward Euler method, future work should investigate whether the discontinuous Galerkin method in time can be rigorously proven to be *unconditionally stable* in this framework.

- **Reaction-Dominated Regimes:** Equation (4.72) can be interpreted as a generic diffusion-convection-reaction PDE at each time step. Because the other terms are scaled by Δt , another interesting avenue of research would be to analyze the stability of the SBM for reaction-dominated PDEs. The challenges associated with reaction-dominated problems using Nitsche’s approach are discussed in [28], providing a solid starting point for this extension.
- **Model Order Reduction (MOR):** A natural extension would consist of exploring model order reduction techniques, treating time (or domain deformation) as a parameter, to further reduce computational costs in long-term simulations.
- **Machine Learning for the Taylor Shift Map:** An innovative perspective concerns the application of Machine Learning techniques to explicitly learn and approximate the Taylor shift map. While frameworks such as *Physics-Informed Neural Networks* (PINNs) [29] are typically employed to solve the entire differential problem, neural networks could be trained exclusively to evaluate the geometric shift at the boundaries. This approach would be particularly advantageous when dealing with domains whose dynamic deformations are analytically complex to track, making the standard computation of the shift map unfeasible or computationally prohibitive.

Bibliography

- [1] Nabil M Atallah, Claudio Canuto, and Guglielmo Scovazzi. «Analysis of the Shifted Boundary Method for the Poisson problem in general domains». In: *arXiv preprint arXiv:2006.00872* (2020) (cit. on pp. 2, 4, 5, 10, 18, 19, 21, 25, 27).
- [2] Nabil M Atallah, Claudio Canuto, and Guglielmo Scovazzi. «Analysis of the shifted boundary method for the Stokes problem». In: *Computer Methods in Applied Mechanics and Engineering* 358 (2020), p. 112609 (cit. on p. 1).
- [3] Nabil M Atallah, Claudio Canuto, and Guglielmo Scovazzi. «The shifted boundary method for solid mechanics». In: *International Journal for Numerical Methods in Engineering* 122.20 (2021), pp. 5935–5970 (cit. on p. 1).
- [4] Daniele Boffi and Lucia Gastaldi. «A finite element approach for the immersed boundary method». In: *Computers & structures* 81.8-11 (2003), pp. 491–501 (cit. on p. 1).
- [5] Andrea Bonito, Irene Kyza, and Ricardo H Nochetto. «Time-discrete higher-order ALE formulations: stability». In: *SIAM Journal on Numerical Analysis* 51.1 (2013), pp. 577–604 (cit. on pp. 1, 87).
- [6] Erik Burman. «Ghost penalty». In: *Comptes Rendus Mathématique* 348.21-22 (2010), pp. 1217–1220 (cit. on p. 1).
- [7] Erik Burman, Susanne Claus, Peter Hansbo, Mats G Larson, and André Massing. «CutFEM: discretizing geometry and partial differential equations». In: *International Journal for Numerical Methods in Engineering* 104.7 (2015), pp. 472–501 (cit. on p. 1).
- [8] Erik Burman, Peter Hansbo, and Mats Larson. «A cut finite element method with boundary value correction». In: *Mathematics of Computation* 87.310 (2018), pp. 633–657 (cit. on p. 1).
- [9] Bernardo Cockburn and Manuel Solano. «Solving Dirichlet boundary-value problems on curved domains by extensions from subdomains». In: *SIAM Journal on Scientific Computing* 34.1 (2012), A497–A519 (cit. on p. 1).

- [10] Jean Donea, Antonio Huerta, J-Ph Ponthot, and Antonio Rodríguez-Ferran. «Arbitrary Lagrangian–Eulerian Methods». In: *Encyclopedia of computational mechanics* (2004) (cit. on pp. 73, 74).
- [11] David Eberly. «Distance from a point to an ellipse, an ellipsoid, or a hyperellipsoid». PhD thesis. Geometric Tools, 2020 (cit. on p. 9).
- [12] Anita Hansbo and Peter Hansbo. «An unfitted finite element method, based on Nitsche’s method, for elliptic interface problems». In: *Computer Methods in Applied Mechanics and Engineering* 191.47-48 (2002), pp. 5537–5552. DOI: 10.1016/S0045-7825(02)00524-8 (cit. on p. 1).
- [13] Klaus Höllig. *Finite element methods with B-splines*. Vol. 26. Frontiers in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2003 (cit. on p. 1).
- [14] David Kamensky, Ming-Chen Hsu, Yue Yu, John A Evans, Michael S Sacks, and Thomas J. R. Hughes. «Immersogeometric cardiovascular fluid-structure interaction analysis with divergence-conforming B-splines». In: *Computer Methods in Applied Mechanics and Engineering* 285 (2015), pp. 211–243. DOI: 10.1016/j.cma.2014.10.040 (cit. on p. 1).
- [15] Efthymios N Karatzas, Giovanni Stabile, Nabil Atallah, Guglielmo Scovazzi, and Gianluigi Rozza. «A reduced order approach for the embedded shifted boundary FEM and a heat exchange system on parametrized geometries». In: *IUTAM Symposium on Model Order Reduction of Coupled Systems, Stuttgart, Germany, May 22–25, 2018: MORCOS 2018*. Springer. 2019, pp. 111–125 (cit. on p. 1).
- [16] Efthymios N Karatzas, Giovanni Stabile, Leo Nouveau, Guglielmo Scovazzi, and Gianluigi Rozza. «A reduced basis approach for PDEs on parametrized geometries based on the shifted boundary finite element method and application to a Stokes flow». In: *Computer Methods in Applied Mechanics and Engineering* 347 (2019), pp. 568–587 (cit. on p. 1).
- [17] Efthymios N Karatzas, Giovanni Stabile, Leo Nouveau, Guglielmo Scovazzi, and Gianluigi Rozza. «A reduced-order shifted boundary method for parametrized incompressible Navier–Stokes equations». In: *Computer Methods in Applied Mechanics and Engineering* 370 (2020), p. 113273 (cit. on p. 1).
- [18] Buyang Li, Yinhua Xia, and Zongze Yang. «Optimal convergence of arbitrary Lagrangian–Eulerian iso-parametric finite element methods for parabolic equations in an evolving domain». In: *IMA Journal of Numerical Analysis* 43.1 (2023), pp. 501–534 (cit. on p. 76).
- [19] Alex Main and Guglielmo Scovazzi. «The shifted boundary method for embedded domain computations. Part I: Poisson and Stokes problems». In: *Journal of Computational Physics* 372 (2018), pp. 972–995 (cit. on p. 1).

- [20] Alex Main and Guglielmo Scovazzi. «The shifted boundary method for embedded domain computations. Part II: Linear advection–diffusion and incompressible Navier–Stokes equations». In: *Journal of Computational Physics* 372 (2018), pp. 996–1026 (cit. on p. 1).
- [21] William F Mitchell and MA McClain. «A collection of 2D elliptic problems for testing adaptive algorithms». In: *NISTIR* 7668 (2010), p. 104 (cit. on p. 116).
- [22] Joachim Nitsche. «Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind». In: *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*. Vol. 36. 1. Springer. 1971, pp. 9–15 (cit. on pp. 6, 44).
- [23] Leo Nouveau, Mario Ricchiuto, and Guglielmo Scovazzi. «High-order gradients with the shifted boundary method: an embedded enriched mixed formulation for elliptic PDEs». In: *Journal of Computational Physics* 398 (2019), p. 108898 (cit. on p. 1).
- [24] Robert Nürnberg. «Distance from a point to an ellipse». In: *URL: <http://wwwf.imperial.ac.uk/rn/distance2ellipse.pdf>* (2006) (cit. on p. 9).
- [25] R. Oyarzúa, M. Solano, and P. Zúñiga. «A high order mixed-FEM for diffusion problems on curved domains». In: *Journal of Scientific Computing* 79.1 (2019), pp. 49–78. DOI: 10.1007/s10915-018-0840-5 (cit. on p. 1).
- [26] J. Parvizian, A. Düster, and E. Rank. «Finite cell method: h- and p-extension for embedded domain problems in solid mechanics». In: *Computational Mechanics* 41.1 (2007), pp. 121–133. DOI: 10.1007/s00466-007-0173-y (cit. on p. 1).
- [27] C. S. Peskin. «Flow patterns around heart valves». In: *Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics*. Vol. 19. Lecture Notes in Physics. Berlin: Springer, 1973, pp. 214–221 (cit. on p. 1).
- [28] Frits de Prenter, Christoph Lehrenfeld, and André Massing. «A note on the stability parameter in Nitsche’s method for unfitted boundary value problems». In: *Computers & Mathematics with Applications* 75.12 (2018), pp. 4322–4336 (cit. on p. 126).
- [29] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. «Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations». In: *Journal of Computational physics* 378 (2019), pp. 686–707 (cit. on p. 126).

- [30] T. Rüberg and F. Cirak. «Subdivision-stabilised immersed b-spline finite elements for moving boundary flows». In: *Computer Methods in Applied Mechanics and Engineering* 209–212 (2012), pp. 266–283. DOI: 10.1016/j.cma.2011.10.007 (cit. on p. 1).
- [31] Ting Song, Alex Main, Guglielmo Scovazzi, and Mario Ricchiuto. «The shifted boundary method for hyperbolic systems: Embedded domain computations of linear waves and shallow water flows». In: *Journal of Computational Physics* 369 (2018), pp. 45–79 (cit. on p. 1).