ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

MECCANICA E SCIENZE AVANZATE DELL'INGEGNERIA

Ciclo 36

Settore Concorsuale: 09/C2 - FISICA TECNICA E INGEGNERIA NUCLEARE

Settore Scientifico Disciplinare: ING-IND/19 - IMPIANTI NUCLEARI

NUMERICAL ALGORITHMS FOR MASS CONSERVATION APPLICATIONS

Presentata da: Giacomo Barbi

Coordinatore Dottorato

Lorenzo Donati

Supervisore

Sandro Manservisi

Esame finale anno 2024

Contents

Abstract

This dissertation focuses on the well-known issue of mass conservation in the context of the finite element technique for computational fluid dynamic simulations. Specifically, non-conventional finite element families for solving Navier-Stokes equations are investigated to address the mathematical constraint of incompressible flows. Raviart-Thomas finite elements are employed for the achievement of a discrete free-divergence velocity. Quadrilateral and hexahedral finite element spaces are considered to investigate the error convergence of different variables. In particular, the proposed algorithm projects the velocity field into the discrete free-divergence space by using the lowest-order Raviart-Thomas element. This decomposition is applied in the context of the projection method, a numerical algorithm employed for solving Navier-Stokes equations. Numerical examples validate the approach's effectiveness, considering different types of computational grids.

Additionally, the dissertation considers an interface advection problem using marker approximation, in the context of multiphase flow simulations. A `C++` library is presented, where the implemented algorithm is able to initialize, advect, and reconstruct a marker cloud performing a best-fit quadratic interpolation. Several numerical tests, equipped with an analytical velocity field for the surface advection, are presented to demonstrate the robustness of the algorithm. Lastly, a comparison with an interpolated velocity by using Raviart-Thomas basis functions is shown, with the aim of maintaining zero divergence, mitigating the classical issue of finite element mass loss.

Introduction

In the context of incompressible flow simulations, Computational Fluid Dynamics codes play a pivotal role. In particular, divergence-free fields are crucial for mass conservation in numerical simulations of engineering applications (multiphase flows, porous-media flows, etc.). Typically, commercial codes employ finite element (FEM) and finite volume (FVM) methods to approximate solutions of the physics problem, both widely applicable across various engineering domains.

While FEM and FVM methods share popularity and similar computational costs, the finite volume method is usually preferred for fluid dynamic simulations due to concerns about mass conservation related to the finite element method. Nevertheless, achieving the desired exact divergence-free field across the discrete domain remains challenging for both methods.

This study addresses the challenge of obtaining a finite element approximation of the solution for the Navier-Stokes system, specifically aiming for a divergence-free velocity field over the discrete domain. This is achieved through the utilization of Raviart-Thomas basis functions [1].

When dealing with fluid dynamics simulations, the mathematical analysis of the specific partial differential equations is fundamental for obtaining reliable numerical results [2]. Focusing on the incompressible Navier-Stokes equations, the literature of the past fifty years has increasingly emphasized the use of mixed finite elements [3, 4, 5, 6], where different types of numerical discretizations are employed to represent different variables. Indeed, this spatial discretization approach has proven to yield convergent numeri-

cal schemes, theoretical convergence rates, and other specific advantageous properties.

Naturally, in the framework of finite element discretization, the flexibility of mixed methods arises from the relaxation of the divergence constraint [7]. However, this relaxation entails a cost, which becomes evident when considering a standard error estimate for the considered equation.

To provide a brief example and without giving extensive details, that are explained in the following chapters, when dealing with error estimate for the velocity field in the Navier-Stokes system, non-divergence mixed finite elements present a connection between the discretized velocity and the continuous pressure [8]. For certain finite element families, this mathematical connection became a numerical drawback, since convergence issues may arise.

In order to describe divergence-free discretization, various approaches have been proposed in the literature, such as the Scott-Vogelius element and the discontinuous Galerkin method, [9, 10, 11, 12, 13, 14, 15]. Despite these efforts, the mathematical problem, commonly referred to as poor mass conservation [16], remains a subject of ongoing interest. Indeed, some stabilization techniques have been proposed in order to overcome this issue, such as the grad-div stabilization [17] or a transformation on the continuous pressure [18]. In this context, a few examples dealing with poor mass conservation are reported in Chapter 4.

Moreover, some remarks can be pointed out regarding the set of equations described by the incompressible Navier-Stokes system. The existence of the solution depends on the divergence operator. In particular, certain surjectivity properties are required, such as the well-known inf-sup condition, which is a sufficient condition in order to find a unique solution for a saddle point problem.

In addition, the problem must also preserve the invariance property, ensuring that a change in the external body force by adding a gradient field, has an effect only in the pressure solution, but not in the velocity.

As pointed out in [19], a lack of L^2-orthogonality between discretely-divergence-free vector and irrotational fields, such as the pressure gradient, may generate a poor momentum balance which translates into a misleading term, causing a poor mass conservation.

This thesis aims to analyze a finite element family designed to address the numerical challenges associated with maintaining a divergence-free velocity field. Specifically, within the framework of mixed problems, the Raviart-

Thomas finite element family assumes an important role, suitable for resolving partial differential equations subject to divergence constraints.

The examination of this problem is carried out in the context of both coupled and split pressure-velocity formulations of the momentum equation. Indeed, the coupled incompressible Navier-Stokes system is known for its high computational effort, leading to the development of various numerical algorithms for treating the split system, resolving velocity and pressure fields separately [20] decreasing the computational burden. We recall that the coupling between velocity and pressure is due to the incompressibility constraint, resulting in saddle-point matrices in discrete form.

To address these challenges, Chorin and Temam introduced the projection method, splitting the Navier-Stokes system into two distinct steps: one for resolving the velocity field and another for the pressure field [21, 22]. This approach has proven to reduce the computational effort and has consequently encouraged the development of various projection methods in recent years [23].

Moreover, the requirement for a divergence-free velocity gains significance in certain fluid dynamic simulations, especially in scenarios involving multiphase flows where mass conservation is crucial for reliable numerical outcomes. This thesis aims to address this mathematical constraint by leveraging a divergence-free representation of the velocity field, in the context of multiphase flows. Specifically, the multiphase problem considered involves the surface advection of a single phase using marker technique approximation.

This methodology has gained interest, particularly in interface tracking within multiphase flow simulations. Different approaches have been explored [24, 25], focusing on marker reconstruction. This thesis introduces a novel numerical algorithm for surface advection on two-dimensional domains, emphasizing marker reconstruction by using the best-fit quadric interpolating equation. Various functions designed to manage marker positions and define interfaces between phases are described, and a comparison between two different types of discrete velocity representation is presented.

The thesis is structured as follows: Chapter 1 outlines the mathematical framework with a description of the problem related to the divergence-free constraint. Chapter 2 provides an overview of finite element discretization, with specific attention to the Raviart-Thomas finite elements within quadrilateral and hexahedral elements. In Chapter 3 an overview of the projection method is presented, including the algorithm proposed for the resolution of

the Navier-Stokes system. Chapter 4 presents the numerical results focusing on the Raviart-Thomas approximation in the context of the velocity orthogonal decomposition and for the projection technique. Finally, Chapter 5 focuses on the multiphase flow problem, explaining the algorithm for surface marker advection and presenting numerical test results, including the code coupling for the employment of Raviart-Thomas velocity representation.

CHAPTER 1

Divergence-Free Mixed Methods

In the following chapter, the main basic properties regarding the numerical solution of the Navier-Stokes equations will be introduced. In particular, the mathematical framework will be described, with a major effort on the divergence-free constraint.

1.1 Mathematical Notations

To present and describe the Navier-Stokes equation mathematically, the typical weak formulation requires foundational concepts from variational calculus, such as appropriate Sobolev spaces [26]. We can start by introducing the vector space of infinitely differentiable scalar function f as $C^\infty(\Omega)$. Here, differentiability is understood in the classical sense, with the domain Ω assumed to be bounded. For the functions with compact support within the domain Ω, we can define the subspace $C_0^\infty(\Omega)$. Employing bold symbols, we denote similar spaces for vector-valued functions, defining therefore $\boldsymbol{C}^\infty(\Omega)$ and $\boldsymbol{C}_0^\infty(\Omega)$. Moving forward, we consider $\Omega \subset \mathbb{R}^2$ as a Lipschitz region, where $\partial\Omega$ is locally represented by a Lipschitz function. This contextual

description leads us to introduce

$$L^2(\Omega) := \left\{ f : \int_\Omega |f^2| d\Omega < \infty \right\} , \tag{1.1}$$

the space comprising square-integrable functions with norm

$$||f||_0 := \left(\int_\Omega |f^2| d\Omega \right)^{\frac{1}{2}} . \tag{1.2}$$

The inner product on the Hilbert space $L^2(\Omega)$ is represented as

$$(f, g) = \int_\Omega f \, g \, d\Omega \, , \tag{1.3}$$

where, naturally, $||f||_0^2 = (f, f)$. Additionally, we denote the space of functions that are almost everywhere bounded as

$$L^\infty(\Omega) := \left\{ f : \operatorname*{ess\,sup}_{\boldsymbol{x} \in \Omega} |f(\boldsymbol{x})| < \infty \right\} , \tag{1.4}$$

with the norm

$$||f||_\infty := \operatorname*{ess\,sup}_{\boldsymbol{x} \in \Omega} |f(\boldsymbol{x})| \, . \tag{1.5}$$

Furthermore, we define $L_0^2(\Omega)$ as the space of square-integrable functions with vanishing mean.

The Sobolev spaces for integers $k \geq 0$, are defined as follows:

$$H^k(\Omega) := \left\{ f \in L^2(\Omega) : \partial^\alpha f \in L^2(\Omega), \text{ for } |\alpha| \leq k \right\} , \tag{1.6}$$

$$W^\infty(\Omega) := \left\{ f \in L^\infty(\Omega) : \partial^\alpha f \in L^\infty(\Omega), \text{ for } |\alpha| \leq k \right\} , \tag{1.7}$$

where the non-negative integer indices $\alpha = (\alpha_1, \alpha_2)$ denote the order of the partial derivatives, i.e. $|\alpha| := \alpha_1 + \alpha_2$, and $\partial^\alpha f = \partial_x^{\alpha_1} \partial_y^{\alpha_2} f$. The respective norms and semi-norms are defined as

$$||f||_k := \left(\sum_{|\alpha| \leq k} ||\partial^\alpha f||_0^2 \right)^{\frac{1}{2}} , \tag{1.8}$$

$$|f|_k := \left(\sum_{|\alpha| = k} ||\partial^\alpha f||_0^2 \right)^{\frac{1}{2}} , \tag{1.9}$$

$$||f||_{k,\infty} := \max_{|\alpha| \leq k} ||\partial^\alpha f||_\infty \, , \tag{1.10}$$

$$|f|_{k,\infty} := \max_{|\alpha| = k} ||\partial^\alpha f||_\infty \, . \tag{1.11}$$

Additionally, by introducing the fractional-order Sobolev space on the boundary $\partial\Omega$ we have

$$H^{\frac{1}{2}}(\partial\Omega) := \left\{ f \in L^2(\partial\Omega) : |f|_{\frac{1}{2},\partial\Omega} < \infty \right\} , \tag{1.12}$$

with the corresponding norm and semi-norm given by

$$|f|_{\frac{1}{2}} := \left(\int_{\partial\Omega} \int_{\partial\Omega} \frac{|f(s) - f(t)|^2}{|s-t|^2} ds\, dt \right)^{\frac{1}{2}} , \tag{1.13}$$

$$||f||_{\frac{1}{2}} := \left(||f||^2_{0,\partial\Omega} + |f|^2_{\frac{1}{2},\partial\Omega} \right)^{\frac{1}{2}} . \tag{1.14}$$

For vector-valued functions such as $\boldsymbol{u} = (u_1, u_2)$, we can consider the following norms on $(H^k(\Omega))^2$ and $(W^k_\infty(\Omega))^2$

$$||\boldsymbol{u}||_k := (||u_1||^2_k + ||u_2||^2_k)^{\frac{1}{2}} , \tag{1.15}$$

$$||\boldsymbol{u}||_{k,\infty} := \max \left\{ ||u_1||_{k,\infty}, ||u_2||_{k,\infty} \right\} . \tag{1.16}$$

With the previously defined functional spaces, we can now introduce the space used for the mixed formulation of second-order elliptic problems. Hence, considering a vector-valued function, we define the space $\boldsymbol{H}(\mathrm{div}, \Omega)$ as

$$\boldsymbol{H}(\mathrm{div}, \Omega) = \left\{ \boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{\nabla} \cdot \boldsymbol{v} \in L^2(\Omega) \right\} , \tag{1.17}$$

that is a Hilbert space equipped with the norm

$$||\boldsymbol{v}||_{\boldsymbol{H}(\mathrm{div})} = \left(||\boldsymbol{v}||^2_0 + ||\boldsymbol{\nabla} \cdot \boldsymbol{v}||^2_0 \right)^{\frac{1}{2}} . \tag{1.18}$$

It is important to note that $\boldsymbol{H}^1(\Omega)$ is continuously embedded in $\boldsymbol{H}(\mathrm{div}, \Omega)$.

For the sake of completeness, some basic definitions are also reported for the space $\boldsymbol{H}(\mathbf{curl}, \Omega)$. Giving a domain $\Omega \in \mathbb{R}^3$, it is possible to define

$$\boldsymbol{H}(\mathbf{curl}, \Omega) := \left\{ \boldsymbol{\chi} \in (L^2(\Omega))^3 : \mathbf{curl}\, \boldsymbol{\chi} \in (L^2(\Omega))^3 \right\} . \tag{1.19}$$

In this case, the **curl** operator is defined as usual

$$\mathbf{curl}\, \boldsymbol{\chi} = \boldsymbol{\nabla} \times \boldsymbol{\chi} := \det \begin{pmatrix} i & j & k \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ \chi_1 & \chi_2 & \chi_3 \end{pmatrix} , \tag{1.20}$$

with its standard norm defined as

$$||\boldsymbol{\chi}||^2_{\mathbf{curl},\Omega} := |\boldsymbol{\chi}|^2_{0,\Omega} + |\,\mathbf{curl}\, \boldsymbol{\chi}|^2_{0,\Omega} . \tag{1.21}$$

It is now possible to introduce the trace map with a trace lemma [27, 28].

Lemma 1.1. *The trace map γ, with $\gamma f = f|_{\partial\Omega}$, for a smooth function f, can be extended as a continuous map from $H^1(\Omega)$ onto $H^{\frac{1}{2}}(\partial\Omega)$*

$$\gamma : H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)\,, \tag{1.22}$$

with a positive constant $C(\Omega)$ such that

$$||\gamma f||_{\frac{1}{2},\partial\Omega} \leq C(\Omega)||f||_{1,\Omega}\,. \tag{1.23}$$

Considering that γ is a continuous map between $H^1(\Omega)$ and $H^{\frac{1}{2}}(\partial\Omega)$, following [29] we get a continuous right inverse

$$E: \quad H^{\frac{1}{2}}(\partial\Omega) \to H^1(\Omega) \tag{1.24}$$

with

$$\gamma E g = g\,, \quad \forall\, g \in H^{\frac{1}{2}}(\partial\Omega)\,. \tag{1.25}$$

It is crucial to consider the normal component of a vector-valued function concerning the boundary of the given domain. With a Lipschitz domain Ω, we affirm the existence of the unit outward normal $\boldsymbol{\nu}$ to the boundary $\partial\Omega$ at almost every point. Consequently, for a smooth vector function $\boldsymbol{f}$ defined on $\overline{\Omega}$, the closure of Ω, we establish the existence of $\boldsymbol{f}\cdot\boldsymbol{\nu}$ at almost every point $\partial\Omega$. For Dirichlet boundary values problems, we introduce the following lemma [30].

Lemma 1.2. *Let $g \in H^{\frac{1}{2}}(\partial\Omega)$ and denote*

$$\langle g, \boldsymbol{f}\cdot\boldsymbol{\nu}\rangle := \int_{\partial\Omega} g\boldsymbol{f}\cdot\boldsymbol{\nu} ds, \quad \boldsymbol{f} \in \boldsymbol{H}(\mathrm{div},\Omega)\,. \tag{1.26}$$

Hence, we have that

$$|\langle g, \boldsymbol{f}\cdot\boldsymbol{\nu}\rangle| \leq ||Eg||_{1,\Omega}||\boldsymbol{f}||_{H(\mathrm{div})}, \quad \forall\, \boldsymbol{f} \in \boldsymbol{H}(\mathrm{div},\Omega), \tag{1.27}$$

with E defined in Lemma 1.1.

Proof. Following Lemma 1.1, consider the extension of g into $H^1(\Omega)$ as $p = Eg$. Thus,

$$\boldsymbol{\nabla}\cdot(p\boldsymbol{f}) = p\boldsymbol{\nabla}\cdot\boldsymbol{f} + \boldsymbol{\nabla}p\cdot\boldsymbol{f}\,. \tag{1.28}$$

If we integrate over the domain Ω and apply the divergence theorem, we get

$$\int_{\partial\Omega} g\boldsymbol{f}\cdot\boldsymbol{\nu} ds = \int_{\Omega} (p\boldsymbol{\nabla}\cdot\boldsymbol{f} + \boldsymbol{\nabla}p\cdot\boldsymbol{f})\, d\Omega\,. \tag{1.29}$$

Applying the Cauchy-Schwartz inequality we obtain

$$\left|\int_{\partial\Omega} g\boldsymbol{f}\cdot\boldsymbol{\nu} ds\right| \leq ||p||_{1,\Omega}||\boldsymbol{f}||_{H(\mathrm{div})}\,. \tag{1.30}$$

□

Before giving an extensive mathematical description of the problem, we recall some points already stated in the introduction, in order to better understand the reason of the following analysis. Therefore, consider a domain $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$ and a time interval $(0, T)$, with $T < \infty$. Hence, the Stokes equations can be expressed as

$$-\nu \Delta \boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{f}\,, \tag{1.31a}$$

$$-\boldsymbol{\nabla} \cdot \boldsymbol{u} = g\,, \tag{1.31b}$$

$$\boldsymbol{u}|_{\partial\Omega} = 0\,, \tag{1.31c}$$

where $g = 0$ for an incompressible fluid. For this type of equation, an a priori error estimate reads as,

$$||\boldsymbol{u} - \boldsymbol{u}_h||_{1,h} \leq C_1 h^k |\boldsymbol{u}|_{k+1} + \frac{C_2}{\nu} h^k |p|_k\,, \tag{1.32}$$

where polynomial finite elements are considered. In particular, the order k is chosen for the velocity discretization, while an order $k-1$ is set for the discrete pressure [3, 4]. The term $\frac{C_2}{\nu} h^k |p|_k$ represents the drawback when dealing with non-divergence mixed finite elements, showing the connection between the discrete velocity $\boldsymbol{u}_h$ and the continuous pressure p. In addition, we recall that the inf-sup condition must be satisfied since we aim to find a unique solution for the equations' system. Specifically, this condition reads as

$$\inf_{q \in L^2_0(\Omega) \setminus \{0\}} \sup_{\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega) \setminus \{0\}} \frac{(\boldsymbol{\nabla} \cdot \boldsymbol{u}, q)}{||\boldsymbol{\nabla} \boldsymbol{u}||_{L^2(\Omega)} ||q||_{L^2(\Omega)}} \geq \beta > 0\,. \tag{1.33}$$

Differently, the constraint $-\boldsymbol{\nabla} \cdot \boldsymbol{u} = g$ does not hold.

Moreover, as it will be described in the next sections, the invariance property must be preserved, ensuring that the velocity solution is not affected as the pressure by a change in an external body force, which is a gradient field. Hence, we have that

$$\boldsymbol{f} \to \boldsymbol{f} + \boldsymbol{\nabla} \psi \Rightarrow (\boldsymbol{u}, p) \to (\boldsymbol{u}, p + \psi)\,. \tag{1.34}$$

In fact, the additional force, given by $\boldsymbol{\nabla} \psi$, is balanced by the pressure gradient, since no-slip boundary conditions do not affect the pressure.

After this brief example, we proceed now to introduce the saddle-point variational formulation of a generic second-order elliptic problem [31] in order to generalize the Stokes equation. Initially, we define

$$\boldsymbol{V} := \boldsymbol{H}(\mathrm{div}, \Omega) \quad \text{and} \quad \boldsymbol{V}^0 := \{\boldsymbol{v} \in \boldsymbol{V} : \boldsymbol{v} \cdot \boldsymbol{\nu} = 0 \ \text{ on } \ \partial\Omega\}\,, \tag{1.35}$$

where $\boldsymbol{V}^0$ represents a subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$, possessing the same norm as $\boldsymbol{H}(\mathrm{div}, \Omega)$. Moreover, it is known that for a Neumann boundary value problem, such as the Stokes equation, the pressure p is uniquely defined in $L^2(\Omega)/\mathbb{R}$, implying that the pressure is unique up to a constant. Hence, we introduce

$$W \equiv L^2(\Omega) \qquad \text{and} \qquad W^0 \equiv W/\mathbb{R}\,, \tag{1.36}$$

where W^0, a subspace of $L^2(\Omega)$, is equipped with the same norm as $L^2(\Omega)$.

We can now introduce a generic second-order elliptic problem, written in the following form,

$$\begin{cases} -\boldsymbol{\nabla} \cdot (a \boldsymbol{\nabla} p) = f & \text{in } \Omega\,, \\ p = 0 & \text{in } \partial\Omega\,, \end{cases} \tag{1.37}$$

where the function $a = a(\boldsymbol{x})$ is a function bounded above and below by a positive constant. For several applications, the variable of interest has the form of

$$\boldsymbol{u} = -a \boldsymbol{\nabla} p\,, \tag{1.38}$$

and thus a mixed finite element method is preferred in order to approximate simultaneously the variables $\boldsymbol{u}$ and p. Specifically, the problem (1.37) can be reformulated into the following first-order system

$$\begin{cases} \boldsymbol{u} + a \boldsymbol{\nabla} p = 0 & \text{in } \Omega\,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} = f & \text{in } \Omega\,, \\ p = 0 & \text{in } \partial\Omega\,. \end{cases} \tag{1.39}$$

The system of equation (1.39), is often denoted as the Darcy problem and represents a simplified model for a single phase flow considering a porous media. This formulation is often preferred since in the context of porous media flow the vector field $\boldsymbol{u}$ is more important than the pressure p. Hence, the need to accurately compute a locally conservative velocity approximation is desired.

If we define $\mu(x) = 1/a(x)$, it is possible to rewrite the first equation of (1.39) in the following form

$$\mu \boldsymbol{u} + \boldsymbol{\nabla} p = 0 \qquad \text{in } \Omega. \tag{1.40}$$

Therefore, integrating by parts after the multiplication by suitable test functions, we obtain the standard weak formulation of a mixed problem

$$\int_\Omega \mu \boldsymbol{u} \cdot \boldsymbol{v}\, d\boldsymbol{x} - \int_\Omega p \boldsymbol{\nabla} \cdot \boldsymbol{v}\, d\boldsymbol{x} = 0 \qquad \forall \boldsymbol{v} \in \boldsymbol{V}^0\,, \tag{1.41a}$$

$$\int_\Omega w \boldsymbol{\nabla} \cdot \boldsymbol{u}\, d\boldsymbol{x} = \int_\Omega f w\, d\boldsymbol{x} \qquad \forall w \in W^0\,. \tag{1.41b}$$

Referring to [31], two bilinear forms are now introduced

$$\begin{aligned} (a)\, M(\boldsymbol{u}, \boldsymbol{v}) &= (\mu \boldsymbol{u}, \boldsymbol{v})\,, \\ (b)\, B(\boldsymbol{u}, w) &= (\boldsymbol{\nabla} \cdot \boldsymbol{u}, w)\,. \end{aligned} \tag{1.42}$$

We exploit (1.42), to rewrite the weak formulation in (1.41), in order to obtain the equivalent saddle-point variational problem:

Problem 1.1. *find* $\{\boldsymbol{u}, p\} \in \boldsymbol{V}^0 \times W^0$ *such that*

$$\begin{aligned} (a)\, M(\boldsymbol{u}, \boldsymbol{v}) - B(\boldsymbol{v}, p) &= 0\,, & \boldsymbol{v} &\in \boldsymbol{V}^0\,, \\ (b)\, B(\boldsymbol{u}, w) &= (f, w)\,, & w &\in W^0\,. \end{aligned} \tag{1.43}$$

Following the work of [32, 33] we can state the theorem for existence and uniqueness.

Theorem 1.1. *Let* M *and* B *defined as* (1.42)*, and consider a constant* $\alpha > 0$ *such that*

$$M(\boldsymbol{v}, \boldsymbol{v}) \geq \alpha ||\boldsymbol{v}||^2_{H(\mathrm{div})}\,, \qquad \forall \boldsymbol{v} \in \boldsymbol{V}^0\,, \tag{1.44}$$

i.e., $M(\cdot, \cdot)$ *is* $\boldsymbol{V}_0$*-elliptic, with*

$$\boldsymbol{V}_0 := \left\{ \boldsymbol{v} \in \boldsymbol{V}^0 : B(\boldsymbol{v}, w) = 0,\ \forall w \in W^0 \right\}\,. \tag{1.45}$$

Suppose also that there exists a constant $\beta > 0$ *such that*

$$\sup_{\boldsymbol{v} \in \boldsymbol{V}^0} \frac{B(\boldsymbol{v}, w)}{||\boldsymbol{v}||_{H(\mathrm{div})}} \geq \beta ||w||_0\,, \qquad \forall w \in W^0\,, \tag{1.46}$$

i.e., $B(\cdot, \cdot)$ *satisfies the inf-sup condition [32]. Then, problem* (1.1) *is well posed, therefore it has a unique solution* $(\boldsymbol{u}, p) \in \boldsymbol{V}^0 \times W^0$*, satisfying*

$$||\boldsymbol{u}||_{H(\mathrm{div})} + ||p||_0 \leq C(\alpha, \beta) ||f||_{W'}\,, \tag{1.47}$$

where

$$||f||_{W'} := \sup_{w \in W^0} \frac{|I|}{||w||_0}\,, \tag{1.48}$$

and

$$I = \int_\Omega f\, w\, d\Omega\,. \tag{1.49}$$

Note that the constant $C(\alpha, \beta)$ depends only on α and β.

For the non-homogeneous Dirichlet boundary value problem of the Stokes equation, we can state an equivalent saddle-point variational form:

Problem 1.2. *find $\{\boldsymbol{u}, p\} \in \boldsymbol{V} \times W$ such that*

$$(a) M(\boldsymbol{u}, \boldsymbol{v}) - B(\boldsymbol{v}, p) = \langle g, \boldsymbol{v} \cdot \boldsymbol{\nu} \rangle\,, \qquad \boldsymbol{v} \in \boldsymbol{V}\,, \tag{1.50}$$

$$(b) B(\boldsymbol{u}, w) \qquad\qquad = (f, w)\,, \qquad w \in W\,. \tag{1.51}$$

Moreover, for this formulation, it is possible to derive theorems concerning existence and uniqueness [32, 33]. Specifically, the inf-sup condition described as

$$\sup_{\boldsymbol{v} \in \boldsymbol{V}} \frac{B(\boldsymbol{v}, w)}{||\boldsymbol{v}||_{H(\mathrm{div})}} \geq \beta ||w||_0\,, \qquad \forall w \in W\,, \tag{1.52}$$

holds for a constant $\beta > 0$.

1.2 Helmholtz-Hodge decomposition

In this section, we expand upon the previous results, by introducing additional properties and statements regarding the Stokes equation. Subsequently, we introduce the Helmholtz-Hodge decomposition and projector, which will have a major role in the context of mixed finite element methods for the Stokes discretization. The forthcoming results are taken from [8], where the interested reader can find also the detailed proofs, which, for brevity, will be omitted here.

Revisiting our framework, we consider a bounded domain Ω equipped with a Lipschitz continuous boundary $\partial\Omega$. Let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ be the body force and the right-hand side of the continuity equation defined as $g \in L^2(\Omega)$. The well-established variational solution of the incompressible Stokes equations (1.31) is the pair $(\boldsymbol{u}, p) \in \boldsymbol{X} \times Y := \boldsymbol{H}^1_0(\Omega) \times L^2_0(\Omega)$ that satisfies the following equations set

$$a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \forall \boldsymbol{v} \in \boldsymbol{X}\,, \tag{1.53a}$$

$$b(\boldsymbol{u}, q) = (g, q) \qquad \forall q \in Y\,, \tag{1.53b}$$

where the bilinear forms are defined as $a(\boldsymbol{u}, \boldsymbol{v}) = \nu(\boldsymbol{\nabla u}, \boldsymbol{\nabla v})$ and $b(\boldsymbol{u}, q) = -(\boldsymbol{\nabla} \cdot \boldsymbol{u}, q)$. Naturally, the operator $(\cdot, \cdot)$ represents the L^2 inner product over Ω and ν is the kinematic viscosity. We recall that the subscript 0 stands for functional spaces with vanishing mean.

We can introduce now the *distributional divergence* and the *weak divergence*.

Lemma 1.3. *Let $\psi \in H^1(\Omega)$ and $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega)$. The following expression holds*

$$\int_\Omega \psi \boldsymbol{\nabla} \cdot \boldsymbol{w} \, d\boldsymbol{x} = -\int_\Omega \boldsymbol{\nabla}\psi \cdot \boldsymbol{w} \, d\boldsymbol{x} + \int_{\partial\Omega} \psi \boldsymbol{w} \cdot \boldsymbol{n} \, d\boldsymbol{s} \,. \tag{1.54}$$

In this case, the surface integral represents the duality pairing between the space $H^{\frac{1}{2}}(\partial\Omega)$ and $H^{-\frac{1}{2}}(\partial\Omega)$.

The previous lemma allows defining the *distributional divergence*.

Definition 1.1. *Let $\boldsymbol{w} \in \boldsymbol{L}^1(\Omega)$ a vector field. We call distributional divergence of $\boldsymbol{w}$ the mapping $C_0^\infty(\Omega) \to \mathbb{R}$ defined as*

$$\psi \mapsto -\int_\Omega \boldsymbol{\nabla}\psi \cdot \boldsymbol{w} \, d\boldsymbol{x} \,. \tag{1.55}$$

In order to introduce the next definition, we need to define the space $L^1_{\text{loc}}(\Omega)$, which is the space of locally integrable function. With the term 'locally' we mean that the function is integrable on every compact subset of its definition domain. We can define now the *weak divergence*.

Definition 1.2. *Let $\boldsymbol{w} \in \boldsymbol{L}^p(\Omega)$ with $p > 1$ a vector field. If there exists a function $\rho \in L^1_{loc}(\Omega)$ such that is possible to represent the distributional divergence as*

$$-\int_\Omega \boldsymbol{\nabla}\psi \cdot \boldsymbol{w} \, d\boldsymbol{x} = \int_\Omega \psi\rho \, d\boldsymbol{x} \quad \forall \psi \in C_0^\infty(\Omega) \,, \tag{1.56}$$

the function ρ is defined as the weak divergence of $\boldsymbol{w}$, i.e. $\rho := \boldsymbol{\nabla} \cdot \boldsymbol{w}$. Moreover, if the vector field $\boldsymbol{w}$ is divergence-free, we have

$$\int_\Omega \boldsymbol{\nabla}\psi \cdot \boldsymbol{w} \, d\boldsymbol{x} = 0 \quad \forall \psi \in C_0^\infty(\Omega) \,. \tag{1.57}$$

Note that if we consider a divergence-free vector field, it will be orthogonal in the $L^2(\Omega)$ scalar product to a gradient field with compact support. However, this property is not usually satisfied for classical mixed methods,

which rely on the conventional discrete inf-sup condition. Consequently, the lack of pressure-robustness may arise for certain mixed methods.

Regarding the divergence constraint equation (1.53b), if we consider a test function $q \in L^2_0(\Omega) \cap C^\infty_0(\Omega)$ and we apply the divergence theorem we obtain

$$\int_\Omega \boldsymbol{\nabla} q \cdot \boldsymbol{u}\, d\boldsymbol{x} = \int_\Omega q\, g\, d\boldsymbol{x}\,. \tag{1.58}$$

In this case, we consider $-g$ as the weak divergence of $\boldsymbol{u}$. In addition, if the field $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ satisfies no-slip boundary conditions, applying the divergence theorem we obtain

$$0 = \int_{\partial\Omega} \boldsymbol{u} \cdot \boldsymbol{n}\, d\boldsymbol{s} = \int_\Omega \boldsymbol{\nabla} \cdot \boldsymbol{u}\, d\boldsymbol{x}\,, \tag{1.59}$$

through which, we can state that $\boldsymbol{\nabla} \cdot \boldsymbol{u} \in L^2_0(\Omega)$. Therefore, choosing $q = \boldsymbol{\nabla} \cdot \boldsymbol{u} - g$ such that

$$||\boldsymbol{\nabla} \cdot \boldsymbol{u} + g||_{L^2(\Omega)} = 0\,, \tag{1.60}$$

we have that $\boldsymbol{\nabla} \cdot \boldsymbol{u} = -g$ in the sense of $L^2(\Omega)$.

The $\boldsymbol{H}(\mathrm{div}, \Omega)$ space can be also described with the following definition, through which for functions $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ Lemma 1.3 is still valid.

Definition 1.3. *Consider the Hilbert space of vector fields equipped with a weak divergence. This space is defined as*

$$\boldsymbol{H}(\mathrm{div}, \Omega) := \left\{ \boldsymbol{u} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{\nabla} \cdot \boldsymbol{u} \in L^2(\Omega) \right\}, \tag{1.61}$$

where $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ is understood in the sense of Definition 1.2, since $\forall p \geq 1$ L^p is a subspace of L^1_{loc}.

It is now possible to introduce the Helmholtz-Hodge decomposition.

Lemma 1.4. *Let Ω be a connected domain and consider a generic vector field $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$. For every $\boldsymbol{f}$ exists a vector field $\boldsymbol{f}_0 \in \boldsymbol{H}(\mathrm{div}, \Omega)$ and a scalar function $\phi \in H^1(\Omega)/\mathbb{R}$ such that the following statements hold*

1. $\boldsymbol{f} = \boldsymbol{f}_0 + \boldsymbol{\nabla}\phi$,
2. $\boldsymbol{\nabla} \cdot \boldsymbol{f}_0 = 0$,
3. $(\boldsymbol{f}_0, \boldsymbol{\nabla}\psi) = 0$ *, for all* $\psi \in H^1(\Omega)$.

Moreover, the decomposition is unique.

Definition 1.4. *We call the Helmholtz-Hodge projector of* $\boldsymbol{f}$ *the function* $\boldsymbol{f}_0 := \mathbb{P}(\boldsymbol{f})$.

The Helmholtz-Hodge projector allows us to establish results concerning existence, uniqueness, and stability estimates.

Lemma 1.5. *Consider a function* $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ *and* $g \in L^2_0(\Omega)$. *For the Stokes problem in equation* (1.31) *we have a unique solution equipped with the following stability estimates:*

$$||\boldsymbol{\nabla u}||_{L^2(\Omega)} \leq \frac{C_P}{\nu}||\mathbb{P}(\boldsymbol{f})||_{L^2(\Omega)} + \frac{1}{\beta}||g||_{L^2(\Omega)}\,, \tag{1.62}$$

$$||p||_{L^2(\Omega)} \leq \frac{C_P}{\beta}||\boldsymbol{f}||_{L^2(\Omega)} + \frac{1}{\beta^2}||g||_{L^2(\Omega)}\,, \tag{1.63}$$

$$\tag{1.64}$$

where β *is the usual inf-sup constant in equation* (1.52).

We can state a corollary regarding the Helmholtz-Hodge projector.

Corollary 1.1. *A gradient field* $\boldsymbol{\nabla}\phi$ *with* $\phi \in H^1(\Omega)/\mathbb{R}$ *is equipped with an Helmholtz projector* $\mathbb{P}(\boldsymbol{\nabla}\phi)$, *such that* $\mathbb{P}(\boldsymbol{\nabla}\phi) = \boldsymbol{0}$.

Considering the first statement of Lemma 1.4 we can derive another lemma, which confirms the fact that the pressure field is determined up to an additive constant, expressed by the scalar field ψ.

Lemma 1.6. *Consider the right-hand side of the Stokes equation,* $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$. *If we change* $\boldsymbol{f} \to \boldsymbol{f} + \boldsymbol{\nabla}\psi$, *with* $\psi \in H^1(\Omega)/\mathbb{R}$, *the solution will change as* $(\boldsymbol{u}, p) \to (\boldsymbol{u}, p + \psi)$.

1.3 Error analysis of mixed methods

In this section, we summarize the primary concept regarding the error analysis for the velocity error $||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}$, and its relationship with the pressure in standard mixed methods. The finite element solution of a generic elliptic problem, such as the Stokes equations, arises from a pair of finite-dimensional spaces defined by piecewise polynomials. Considering a partition τ_h of Ω, we can consider the pair of conforming piecewise polynomials spaces as the set given by $\boldsymbol{X}_h \times Y_h \subset \boldsymbol{X} \times Y$. A standard solution for a Galerkin

finite element method for the Stokes equation is the pair $(\boldsymbol{u}_h, p_h) \in \boldsymbol{X}_h \times Y_h$ such that

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h) \qquad \forall \boldsymbol{v}_h \in \boldsymbol{X}_h\,, \tag{1.65a}$$

$$b(\boldsymbol{u}_h, q_h) = (g, q_h) \qquad \forall q \in Y_h\,. \tag{1.65b}$$

The problem in equation (1.65) describes a mixed finite element method, distinguished by the inclusion of two finite element spaces. It is widely acknowledged in the literature that in order to reach the existence and uniqueness of a solution, the choice of the finite element spaces $\boldsymbol{X}_h$ and Y_h must follow compatibility criteria. For the Stokes equations, these criteria are reflected in the discrete inf-sup condition. Specifically, ensuring the existence and stability of a solution to the preceding problem necessitates the fulfillment of

$$\inf_{q_h \in Y_h \setminus \{0\}} \sup_{\boldsymbol{v}_h \in \boldsymbol{X}_h \setminus \{\boldsymbol{0}\}} \frac{(\boldsymbol{\nabla} \cdot \boldsymbol{v}_h, q_h)}{||\boldsymbol{\nabla} \boldsymbol{v}_h||_{L^2(\Omega)} ||q_h||_{L^2(\Omega)}} \geq \beta_h > 0\,. \tag{1.66}$$

Moreover, ensuring stability and achieving optimal convergence rates necessitates $\beta_h \geq \beta_0 > 0$ as $h \to 0^+$. Furthermore, due to the problem's linearity, we can guarantee the uniqueness of the velocity, regardless of the finite element space chosen for the velocity discretization. Extending this notion, the uniqueness of the pressure can also be ensured. Given the linearity and finite dimensionality of the problem, the existence of a solution is a consequence of this uniqueness. The discrete divergence operator $\boldsymbol{\nabla} \cdot_h : \boldsymbol{X}_h \to Y_h$ is defined by using the L^2-projection

$$(\boldsymbol{\nabla} \cdot_h \boldsymbol{v}_h, q_h) = (\boldsymbol{\nabla} \cdot \boldsymbol{v}_h, q_h) \quad \forall q_h \in Y_h\,. \tag{1.67}$$

Throughout the years, various finite element pairs have been investigated to fulfill the discrete inf-sup condition (1.66). One popular family of finite elements among these is the Taylor-Hood finite element pair $\mathcal{P}_k/\mathcal{P}_{k-1}$, with $k \geq 2$.

In the finite element framework, the representation of the divergence constraint exists exclusively through the second equation in (1.65). While for the standard Taylor-Hood pair, the relationship $\boldsymbol{\nabla} \cdot_h \boldsymbol{X}_h = Y_h$ holds, it is generally untrue that $\boldsymbol{\nabla} \cdot \boldsymbol{X}_h \not\subset Y_h$. This latter relationship is crucial for the vanishing of the divergence norm, i.e. $||\boldsymbol{\nabla} \cdot \boldsymbol{u}_h||_{L^2(\Omega)} = 0$. References in literature, such as [34], have demonstrated that for standard Taylor-Hood element of type $\mathcal{P}_2/\mathcal{P}_1$, the quantity $||\boldsymbol{\nabla} \cdot \boldsymbol{u}_h||_{L^2(\Omega)}$ can become large.

We define now the finite and infinite space of a weakly divergence-free vector. If we consider a function $\boldsymbol{u} \in \boldsymbol{X}$, the divergence-free subspace is described as

$$\boldsymbol{X}_{\text{div}} := \{\boldsymbol{u} \in \boldsymbol{X} : b(\boldsymbol{u}, q) = 0 \quad \forall q \in Y\} \,. \tag{1.68}$$

The discrete version follows as

$$\boldsymbol{X}_{h,\text{div}}(g) := \{\boldsymbol{u}_h \in \boldsymbol{X}_h : (\boldsymbol{\nabla} \cdot \boldsymbol{u}_h, q_h) = (g, q_h) \quad \forall q_h \in Y_h\} \,. \tag{1.69}$$

For cases with $g = 0$, we consider $\boldsymbol{X}_{h,\text{div}} = \boldsymbol{X}_{h,\text{div}}(0)$, defining the space of discrete divergence-free functions. Moreover, since $\boldsymbol{\nabla} \cdot \boldsymbol{X}_h \not\subset Y_h$, a function $\boldsymbol{f} \in \boldsymbol{X}_{h,\text{div}}$ will be not in general divergence-free in the sense of $L^2(\Omega)$, and therefore we can say that $\boldsymbol{X}_{h,\text{div}} \not\subset \boldsymbol{X}_{\text{div}}$. On the other hand, we know that the inclusion $\boldsymbol{X}_{h,\text{div}} \subset \boldsymbol{X}_h \subset \boldsymbol{X}$ holds, and therefore we can use the space $\boldsymbol{X}_{h,\text{div}}$ to take the test function for the problems in the equation (1.53) and its finite-dimensional version of (1.65). Since we are dealing with these test functions, we can derive the error equation subtracting the previous equations (infinite and finite version of the Stokes problem, i.e. equations (1.53) and (1.65)) and considering that $q_h = 0$, a necessary condition to have $b(\boldsymbol{v}_h, q_h) = 0$ with $||\boldsymbol{\nabla} \cdot \boldsymbol{v}_h||_{L^2(\Omega)} \neq 0$ always. Therefore, we obtain

$$a(\boldsymbol{u} - \boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p - p_h) = 0 \quad \forall \boldsymbol{v}_h \in \boldsymbol{X}_{h,\text{div}} \,. \tag{1.70}$$

The specific choice of the test functions allows removing the discrete pressure from the equation due to the orthogonality, i.e., $b(\boldsymbol{v}_h, p_h) = 0$. However, we can not remove the continuous pressure from the error equation because $\boldsymbol{X}_{h,\text{div}} \not\subset \boldsymbol{X}_{\text{div}}$, resulting in a dependency between the velocity error and pressure. The objective now is to obtain the best error estimate for the velocity field while explicitly considering the dependency on the pressure error approximation. We start with a decomposition of the velocity error introducing an arbitrary $\tilde{\boldsymbol{u}}_h \in \boldsymbol{X}_{h,\text{div}}$,

$$\boldsymbol{u} - \boldsymbol{u}_h = (\boldsymbol{u} - \tilde{\boldsymbol{u}}_h) - (\boldsymbol{u}_h - \tilde{\boldsymbol{u}}_h) := \boldsymbol{\eta} - \boldsymbol{\phi}_h \,. \tag{1.71}$$

If we take $\boldsymbol{v}_h = \boldsymbol{\phi}_h$ as a test function and we add to the right-hand side of the error equation the term $b(\boldsymbol{v}_h, q_h)$ since it is equal to zero for arbitrary $q_h \in Y_h$, we obtain

$$a(\boldsymbol{\eta}, \boldsymbol{\phi}_h) - a(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) + b(\boldsymbol{\phi}_h, p) - b(\boldsymbol{\phi}_h, q_h) = 0 \quad \forall q_h \in Y_h \,. \tag{1.72}$$

After, rearranging these terms we have

$$\nu||\boldsymbol{\nabla}\boldsymbol{\phi}_h||^2_{L^2(\Omega)} = \nu(\boldsymbol{\nabla}\boldsymbol{\eta}, \boldsymbol{\nabla}\boldsymbol{\phi}_h) - (\boldsymbol{\nabla} \cdot \boldsymbol{\phi}_h, p - q_h) \quad \forall q_h \in Y_h \,. \tag{1.73}$$

We can apply the Cauchy-Schwarz inequality to the right-hand side terms obtaining

$$\nu(\boldsymbol{\nabla}\boldsymbol{\eta}, \boldsymbol{\nabla}\boldsymbol{\phi}_h) \leq \nu||\boldsymbol{\nabla}\boldsymbol{\eta}||_{L^2(\Omega)}||\boldsymbol{\nabla}\boldsymbol{\phi}_h||_{L^2(\Omega)}\,, \tag{1.74}$$

$$-(\boldsymbol{\nabla}\cdot\boldsymbol{\phi}_h, p - q_h) \leq ||\boldsymbol{\nabla}\cdot\boldsymbol{\phi}_h||_{L^2(\Omega)}||p - q_h||_{L^2(\Omega)}\,. \tag{1.75}$$

Moreover, considering that for homogeneous Dirichlet boundary conditions with constant equal to 1 holds $||\boldsymbol{\nabla}\cdot\boldsymbol{\phi}_h||_{L^2(\Omega)} \leq ||\boldsymbol{\nabla}\boldsymbol{\phi}_h||_{L^2(\Omega)}$, we can divide by $\nu||\boldsymbol{\nabla}\boldsymbol{\phi}_h||_{L^2(\Omega)} \neq 0$ obtaining

$$||\boldsymbol{\nabla}\boldsymbol{\phi}_h||_{L^2(\Omega)} \leq ||\boldsymbol{\nabla}\boldsymbol{\eta}||_{L^2(\Omega)} + \nu^{-1}||p - q_h||_{L^2(\Omega)} \quad \forall q_h \in Y_h\,. \tag{1.76}$$

If we apply the triangle inequality to $\boldsymbol{\phi}_h$ and $\boldsymbol{\eta}$ we have

$$\begin{aligned}||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)} \leq 2 \inf_{\tilde{\boldsymbol{u}}_h \in \boldsymbol{X}_{h,\mathrm{div}}} ||\boldsymbol{\nabla}(\boldsymbol{u} - \tilde{\boldsymbol{u}}_h)||_{L^2(\Omega)} + \\ + \nu^{-1} \inf_{q_h \in Y_h} ||p - q_h||_{L^2(\Omega)}\,.\end{aligned} \tag{1.77}$$

From the error bound presented in (1.77), it becomes evident that the optimal approximation error for the pressure is connected to the velocity error. Specifically, the error tends to increase if the pressure approximation is large or if we have a small viscosity ν.

At this point, we can summarize the different interpretations regarding divergence-free functions that have been introduced.

- *Strong divergence-free*: refers to the classical strong formulation of the Navier-Stokes equations, where the velocity $\boldsymbol{u}$ is pointwise divergence-free considering classical derivatives.
- *Weak divergence-free*: refers to the weak formulation of the Navier-Stokes equation, for which the function $\boldsymbol{v} \in \boldsymbol{X}$ satisfies the relation $||\boldsymbol{\nabla}\cdot\boldsymbol{v}||_{L^2(\Omega)} = 0$.
- *Discrete weakly divergence-free*: describes a function $\boldsymbol{v}_h$ that is in $\boldsymbol{X}_{h,\mathrm{div}}$.

It holds that a strong divergence-free function is weak and discrete divergence-free, and a weak divergence-free function is also discrete divergence-free. However, usually, a discrete divergence-free function might not be strong or weak divergence-free.

Remark 1.1. *The divergence error of the velocity is bounded by the gradient error, i.e.,* $||\boldsymbol{\nabla}\cdot(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)} \leq ||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}$. *Consequently, the incompressibility constraint leads to* $||\boldsymbol{\nabla}\cdot\boldsymbol{u}_h||_{L^2(\Omega)} \leq ||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}$. *Therefore, we can conclude that large errors in* $||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}$ *result in a poor mass conservation.*

1.4 Pressure-Robust mixed method

In this subsection, we aim to address the constraints arising from the relationship between $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h$ and Y_h, i.e. $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h \not\subset Y_h$. In particular, we focus on stable conforming element pairs that satisfy $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h \subset Y_h$, ensuring $\boldsymbol{X}_h \subset \boldsymbol{X}$ and $Y_h \subset Y$, with $\boldsymbol{X}, Y$ defined as in (1.53). Our primary objective is to demonstrate the independence of the velocity error from the pressure norm by strictly including $\boldsymbol{X}_{h,\mathrm{div}}$ in $\boldsymbol{X}_{\mathrm{div}}$. To achieve these results, following [35], we can introduce the Fortin operator π_F, as a consequence of the inf-sup stability of the element pair $(\boldsymbol{X}_h, Y_h)$. Precisely, we define $\pi_F : \boldsymbol{X} \to \boldsymbol{X}_h$ such that

$$(\boldsymbol{\nabla}\cdot\boldsymbol{v}, q_h) = (\boldsymbol{\nabla}\cdot\pi_F(\boldsymbol{v}), q_h) \quad \forall \boldsymbol{v}\in\boldsymbol{X}\,, \forall q_h \in Y_h\,, \tag{1.78}$$

where $||\boldsymbol{\nabla}\pi_F(\boldsymbol{v})||_{L^2(\Omega)} \leq C_F||\boldsymbol{\nabla}\boldsymbol{v}||_{L^2(\Omega)}$.

We now present similar stability estimates, as introduced in Lemma 1.5, but for the discrete solution.

Lemma 1.7. *Consider a finite element pair solution for the Stokes problem $(\boldsymbol{u}_h, p_h) \in \boldsymbol{X}_h \times Y_h$, with $\boldsymbol{X}_h \subset \boldsymbol{X}$ and $Y_h \subset Y$, that satisfies the discrete inf-sup condition (1.66). If $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h \subset Y_h$, for $\boldsymbol{f}\in\boldsymbol{L}^2(\Omega)$ and $g\in Y$, the pair $(\boldsymbol{u}_h, p_h)$ is a solution for the Stokes problem with the following stability estimates*

$$||\boldsymbol{\nabla}\boldsymbol{u}_h||_{L^2(\Omega)} \leq \frac{C_P}{\nu}||\mathbb{P}(\boldsymbol{f})||_{L^2(\Omega)} + \frac{1}{\beta_h}||g||_{L^2(\Omega)}\,, \tag{1.79}$$

$$||p_h||_{L^2(\Omega)} \leq \frac{C_P}{\beta_h}||\boldsymbol{f}||_{L^2(\Omega)} + \frac{1}{\beta_h^2}||g||_{L^2(\Omega)}\,. \tag{1.80}$$

In this case $\mathbb{P}(\cdot)$ represents the Helmholtz-Hodge projector previously introduced in the Definition 1.4.

Lemma 1.8. *Consider the finite element spaces $\boldsymbol{X}_h \subset \boldsymbol{X}$ and $Y_h \subset Y$, with $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h \subset Y_h$ and the discrete inf-sup condition satisfied. Then, for all $\boldsymbol{w}\in\boldsymbol{X}$ and $\boldsymbol{w}_h \in \boldsymbol{X}_{h,div}(g)$ we have*

$$\inf_{\boldsymbol{w}_h\in\boldsymbol{X}_{h,div}(g)} ||\boldsymbol{\nabla}(\boldsymbol{w}-\boldsymbol{w}_h)||_{L^2(\Omega)} \leq (1+C_F)\inf_{\boldsymbol{v}_h\in\boldsymbol{X}_h} ||\boldsymbol{\nabla}(\boldsymbol{w}-\boldsymbol{v}_h)||_{L^2(\Omega)}\,. \tag{1.81}$$

Lemma 1.9. *Consider the finite element spaces $\boldsymbol{X}_h \subset \boldsymbol{X}$ and $Y_h \subset Y$, with $\boldsymbol{\nabla}\cdot\boldsymbol{X}_h \subset Y_h$ and the discrete inf-sup condition satisfied. Consider the L^2 projection of $p\in Y$, i.e. $\pi_{Y_h}p \in Y_h$ defined as*

$$(p-\pi_{Y_h}p, q_h) = 0 \quad \forall\, q_h \in Y_h\,. \tag{1.82}$$

Then, we have the following a priori error estimates for the unique discrete pair solution $(\boldsymbol{u}_h, p_h)$ *of the equation* (1.65)

$$||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)} \leq 2 \inf_{\boldsymbol{w}_h \in \boldsymbol{X}_{h,div}(g)} ||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{w}_h)||_{L^2(\Omega)} \tag{1.83}$$

$$\leq 2(1 + C_F) \inf_{\boldsymbol{w}_h \in \boldsymbol{X}_h} ||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{w}_h)||_{L^2(\Omega)}\,, \tag{1.84}$$

$$||\pi_{Y_h} p - p_h||_{L^2(\Omega)} \leq \frac{\nu}{\beta_h} ||\boldsymbol{\nabla}\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)}\,, \tag{1.85}$$

$$||p - p_h||_{L^2(\Omega)} \leq ||p - \pi_{Y_h} p||_{L^2(\Omega)} + \frac{\nu}{\beta_h} ||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}\,. \tag{1.86}$$

The previous lemma states an important result. The first set of inequality estimates demonstrates that the velocity error remains independent of the pressure error. This characteristic distinguishes pressure-robust mixed methods, a feature absent in classical mixed methods. Moreover, the second inequality illustrates that continuous pressure can be approximated by discrete pressure, with an additional term dependent only on velocity. Note that the discrete inf-sup constant only appears in the pressure estimates, and due to its scaling by ν, its value is usually negligible. Concerning velocity estimates, the classical constant $1/\beta_h$ is replaced by the Fortin constant C_F.

Lastly, we present the discrete version of Lemma 1.6.

Lemma 1.10. *Consider the finite element spaces* $\boldsymbol{X}_h \subset \boldsymbol{X}$ *and* $Y_h \subset Y$*, with* $\boldsymbol{\nabla} \cdot \boldsymbol{X}_h \subset Y_h$ *and the discrete inf-sup condition satisfied. For the pair* $(\boldsymbol{u}_h, p_h)$*, discrete solution of the problem* (1.65)*, the change* $\boldsymbol{f} \to \boldsymbol{f} + \boldsymbol{\nabla}\psi$ *in the right-hand side, with* $\psi \in H^1(\Omega)/\mathbb{R}$*, affects only the discrete pressure solution as* $(\boldsymbol{u}_h, p_h) \to (\boldsymbol{u}_h, p_h + \pi_{Y_h}\psi)$*.*

1.5 Remark about vorticity equation

In this section, we explore another aspect differentiating classical mixed methods from pressure-robust methods, considering the vorticity equation in both continuous and discrete settings.

1.5.1 Continuous setting

Consider $\boldsymbol{v} \in \boldsymbol{X}_{\text{div}} \cap \boldsymbol{C}_0^\infty(\Omega)$ as a divergence-free vector in a 3-dimensional space. This implies the existence of a vector potential $\boldsymbol{\xi} \in \boldsymbol{C}_0^\infty(\Omega)$, that satisfies the relation $\boldsymbol{v} = \boldsymbol{\nabla} \times \boldsymbol{\xi}$. In addition, let the pair solution for the

Stokes system $(\boldsymbol{u}, p) \in \boldsymbol{H}^3(\Omega) \cap H^1(\Omega)$, with the curl of body force square-integrable, i.e. $\boldsymbol{\nabla} \times \boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$.

Using $\boldsymbol{v}$ to test the momentum equation of the Stokes system (1.53) and integrating by part we obtain

$$(-\nu\Delta\boldsymbol{u}, \boldsymbol{\nabla} \times \boldsymbol{\xi}) + (\boldsymbol{\nabla} p, \boldsymbol{\nabla} \times \xi) = (\boldsymbol{f}, \boldsymbol{\nabla} \times \xi)\,, \tag{1.87}$$

where the vorticity $\boldsymbol{\omega}$ has been introduced, which is defined as $\boldsymbol{\omega} := \boldsymbol{\nabla} \times \boldsymbol{u}$. Considering that $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} p) \equiv 0$, we obtain

$$(-\nu\Delta\boldsymbol{\omega}, \xi) = (\boldsymbol{\nabla} \times \boldsymbol{f}, \xi)\,, \tag{1.88}$$

that can be seen as a classic diffusion equation

$$-\nu\Delta\boldsymbol{\omega} = \boldsymbol{\nabla} \times \boldsymbol{f}\,. \tag{1.89}$$

Note that the vorticity equation perfectly satisfies the invariance property defined in Lemma 1.6 because, considering two body forces, i.e. $\boldsymbol{f}$ and $\boldsymbol{f} + \boldsymbol{\nabla}\phi$, we obtain the same right-hand side for equation (1.89) since

$$\boldsymbol{\nabla} \times (\boldsymbol{f} + \boldsymbol{\nabla}\phi) = \boldsymbol{\nabla} \times \boldsymbol{f}\,. \tag{1.90}$$

This establishes that, as in the case of the equation for the explicit velocity $\boldsymbol{u}$, also for the vorticity equation $\boldsymbol{\omega}$ the addition of a force described by the gradient of a scalar field $\boldsymbol{\nabla}\phi$, produces only an influence on the pressure solution.

We underline that the formal vorticity equation in (1.89) has been derived starting from the strong form corresponding to the weak formulation of the velocity equation, i.e.

$$a(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) = (\mathbb{P}(\boldsymbol{f}), \boldsymbol{v}) \quad \forall \boldsymbol{v} \in \boldsymbol{X}_{\text{div}}\,. \tag{1.91}$$

The latter equation, with the addition of the condition $-\boldsymbol{\nabla} \cdot \boldsymbol{u} = g$, uniquely defines the velocity field $\boldsymbol{u}$. The Helmholtz projector in equation (1.91) shows that the dependence of the formal vorticity equation in (1.89) relies on $\boldsymbol{\nabla} \times \boldsymbol{f}$ and not on $\boldsymbol{f}$.

Moreover, note that for classical mixed methods, the L^2-orthogonality between discrete divergence-free test functions and gradient fields is relaxed.

Remark 1.2. *Following a theorem presented in [36], we underline the importance of the operators already presented, i.e. the divergence and the curl of a vector field.*

Consider a simply connected bounded region $\Omega \in \mathbb{R}^3$, where the surface $\partial\Omega$ is described by the union of a finite number of disjoint closed surfaces in C^2. Hence, there is a unique vector field $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$, which satisfies the following expressions

$$\boldsymbol{\nabla} \cdot \boldsymbol{v} = g \quad \text{in } \Omega, \qquad \boldsymbol{\nabla} \times \boldsymbol{v} = \boldsymbol{\omega} \quad \text{in } \Omega, \qquad \boldsymbol{v} \cdot \boldsymbol{n} = 0 \quad \text{in } \partial\Omega\,, \tag{1.92}$$

where $g \in L^2_0(\Omega)$, $\boldsymbol{\omega} \in \boldsymbol{C}^1(\Omega)$ with $\boldsymbol{\nabla} \cdot \boldsymbol{\omega} = 0$. The preceding theorem states that a vector field can be completely determined by having information about its divergence, and its curl, together with additional boundary data.

1.5.2 Discrete setting

Following the analogy of the continuous equation, testing the vorticity equation with a discrete divergence-free function leads to a discrete vorticity equation formulated by a weak application of the curl operator.

We introduce now the discrete Helmholtz projector $\mathbb{P}_h : \boldsymbol{L}^2(\Omega) \to \boldsymbol{X}_{h,\mathrm{div}}$, defined as the L^2-projection onto $\boldsymbol{X}_{h,\mathrm{div}}$. Therefore, if we are dealing with a pressure-robust test function, i.e. $\boldsymbol{X}_{h,\mathrm{div}} \subset \boldsymbol{X}_{\mathrm{div}}$, since we have that $\mathbb{P}_h(\boldsymbol{\nabla}\phi) = 0$ for all $\phi \in H^1(\Omega)$, a discrete vorticity equation can be obtained as

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = -\nu(\Delta\boldsymbol{u}, \boldsymbol{v}_h)\,. \tag{1.93}$$

Hence, the latter equation is pressure-independent for all $\boldsymbol{v}_h \in \boldsymbol{X}_{h,\mathrm{div}}$.

On the other hand, for classical mixed methods which hold $\boldsymbol{X}_{h,\mathrm{div}} \not\subset \boldsymbol{X}_{\mathrm{div}}$, we have

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = (\mathbb{P}_h(\boldsymbol{f}), \boldsymbol{v}_h) = -\nu(\Delta\boldsymbol{u}, \boldsymbol{v}_h) + (\mathbb{P}_h(\boldsymbol{\nabla}p), \boldsymbol{v}_h)\,, \tag{1.94}$$

for all $\boldsymbol{v}_h \in \boldsymbol{X}_{h,\mathrm{div}}$. In addition, considering the definition of $\mathbb{P}_h$, integrating by part we have that $(\pi_{Y_h}p, \boldsymbol{\nabla} \cdot \boldsymbol{v}_h) = 0$, which leads to the approximation estimate for the L^2-projection

$$\begin{aligned} |(\mathbb{P}_h(\boldsymbol{\nabla}p), \boldsymbol{v}_h)| = |(\boldsymbol{\nabla}p, \boldsymbol{v}_h)| = |(p, \boldsymbol{\nabla} \cdot \boldsymbol{v}_h)| = |(p - \pi_{Y_h}p, \boldsymbol{\nabla} \cdot \boldsymbol{v}_h)| \leq \\ \leq Ch^k |p|_{H^k(\Omega)} ||\boldsymbol{\nabla} \cdot \boldsymbol{v}_h||_{L^2(\Omega)}\,, \end{aligned} \tag{1.95}$$

where in this case the pressure p is assumed to be sufficiently regular.

Note that for classical mixed method, we have an additional term in equation (1.94), i.e. $(\mathbb{P}_h(\boldsymbol{\nabla}p), \boldsymbol{v}_h)$ with $\boldsymbol{v}_h \in \boldsymbol{X}_{h,\mathrm{div}}$. This is a discrepancy from a pressure-robust method, for which equation (1.93) holds. Naturally,

the previous consistency error decreases with optimal order in the case of sufficient regularity of p, even though the flow problem configuration may produce arbitrarily large errors.

CHAPTER 2

Finite Element Discretization

2.1 Basic Properties

2.1.1 Partition of Domain

In this section, fundamental insights into domain partitioning will be presented. Establishing a finite element approximation within a domain Ω demands maintaining specific continuity properties at the interfaces between its constituent elements.

Consider a domain Ω partitioned into subdomains, such that $\Omega = \cup_{r=1}^{m} K_r$, where the generic element K_r can take the form of a triangle or a quadrilateral in the case of a bi-dimensional domain (tetrahedron or hexahedron for three dimensions). An example of a generic domain Ω is depicted in Figure 2.1. Denote with $\mathcal{L}_h$ the partition of Ω in m elements, each equipped with e_i edges (for instance e_i, with $i = 1, 2, 3$ or $i = 1, 2, 3, 4$). We shall also denote the interface between elements K_i and K_j as

$$e_{ij} = \partial K_i \cap \partial K_j \,. \tag{2.1}$$

Moreover, we define also ε_h as

$$\varepsilon_h = \bigcup_{ij} e_{ij} \bigcup \Gamma_h = \bigcup_{i} \partial K_i \,, \tag{2.2}$$

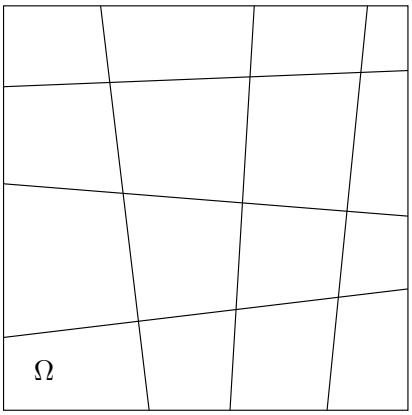

Figure 2.1: Generic domain Ω divided in m non-regular quadrilateral elements.

where Γ_h is the set of the boundary edges or faces. It is important to note that in this work only compatible meshes are taken into account, avoiding situations where a mesh contains hanging nodes. Additionally, following the literature notation, the mesh size is represented by the index h, which is also used for the maximum diameter of the element. We now reformulate the functional spaces, previously described, but in the context of a partitioned domain. Starting with a scalar field v, we define $X(\Omega)$ as

$$X(\Omega) := \left\{ v \in L^2(\Omega) : v|_{K_i} \in H^1(K_i)\,, \forall i \right\} = \prod_r H^1(K_r)\,, \tag{2.3}$$

equipped with the norm

$$||v||^2_{X(\Omega)} := \sum_r ||v||^2_{1,K_r}\,. \tag{2.4}$$

Regarding vector-valued fields, we define also the space $\boldsymbol{Y}(\Omega)$ as

$$\begin{aligned} \boldsymbol{Y}(\Omega) :&= \left\{ \boldsymbol{q} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{q}|_{K_i} \in \boldsymbol{H}(\mathrm{div}, K_i)\,, \forall i \in \mathbb{N}, i \leq m \right\} = \\ &= \prod_r \boldsymbol{H}(\mathrm{div}, K_r)\,, \end{aligned} \tag{2.5}$$

with the norm

$$||\boldsymbol{q}||^2_{\boldsymbol{Y}(\Omega)} := \sum_r ||\boldsymbol{q}||^2_{\mathrm{div},K_r}\,. \tag{2.6}$$

Lastly, we define the space $\boldsymbol{W}(\Omega)$ as

$$\begin{aligned} \boldsymbol{W}(\Omega) :&= \left\{ \boldsymbol{\chi} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{\chi}|_{K_i} \in \boldsymbol{H}(\mathbf{curl}, K_i)\,, \forall i \in \mathbb{N}, i \leq m \right\} = \\ &= \prod_r \boldsymbol{H}(\mathbf{curl}, K_r)\,, \end{aligned} \tag{2.7}$$

with the norm equal to

$$||\boldsymbol{\chi}||^2_{\boldsymbol{W}(\Omega)} := \sum_r ||\boldsymbol{\chi}||^2_{\mathbf{curl},K_r} \,. \tag{2.8}$$

At this point, it is important to introduce other spaces, that are subsets of the spaces introduced above. Therefore, we can consider a partition of the domain boundary $\Gamma = D \cup N | D \cap N = \emptyset$, where the boundary Γ has been divided considering different types of boundary conditions, i.e. D for Dirichlet and N for Neumann. Hence, we define

$$H^1_{0,D}(\Omega) := \left\{ v \in H^1(\Omega) : v|_D = 0 \right\} \subset X(\Omega)\,, \tag{2.9}$$

where naturally we have $H^1_{0,D}(\Omega) = H^1_0(\Omega)$ if $D = \Gamma$ and $H^1_{0,D}(\Omega) = H^1(\Omega)$ if $D = \emptyset$. Likewise, we have

$$\boldsymbol{H}_{0,N}(\mathrm{div},\Omega) := \left\{ \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\Omega) : \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle = 0\,, \forall v \in H^1_{0,D}(\Omega) \right\} \subset \boldsymbol{Y}(\Omega)\,, \tag{2.10}$$

with $\boldsymbol{H}_0(\mathrm{div},\Omega) = \boldsymbol{H}_{0,N}(\mathrm{div},\Omega)$ if $N = \Gamma$. Lastly, we recall another subspace of $\boldsymbol{H}(\mathrm{div},\Omega)$ defined as

$$\boldsymbol{H}_0(\mathrm{div},\Omega) := \{ \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\Omega) : \boldsymbol{\nabla}\cdot\boldsymbol{q} = 0 \}\,. \tag{2.11}$$

Considering now a function $v \in H^1(\Omega)$ and a function $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div},\Omega)$, denoting the vector normal to $\Gamma_r = \partial K_r$ as $\boldsymbol{n}_r$, we have that

$$\sum_r \langle \boldsymbol{q}\cdot\boldsymbol{n}_r, v\rangle_{\Gamma_r} = \langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_\Gamma\,, \tag{2.12}$$

where the operator $\langle\cdot,\cdot\rangle$ represents the duality product between $H^{\frac{1}{2}}(\Gamma_r)$ and $H^{-\frac{1}{2}}(\Gamma_r)$. Therefore, inside each element, we can apply the Green formula obtaining

$$\langle \boldsymbol{q}\cdot\boldsymbol{n}, v\rangle_\Gamma = \sum_r \left\{ \int_{K_r} \boldsymbol{\nabla}\cdot\boldsymbol{q} v dx + \int_{K_r} \boldsymbol{q}\cdot\boldsymbol{\nabla} v dx \right\}\,. \tag{2.13}$$

As described in [37] we have the following propositions:

Proposition 2.1. *For the space $H^1_{0,D}(\Omega)$ we have that*

$$H^1_{0,D}(\Omega) = \left\{ v \in X(\Omega) : \sum_r \langle \boldsymbol{q}\cdot\boldsymbol{n}_r, v\rangle = 0,\ \forall \boldsymbol{q} \in \boldsymbol{H}_{0,N}(\mathrm{div},\Omega) \right\}\,. \tag{2.14}$$

Proposition 2.2. *For the space $\boldsymbol{H}_{0,N}(\mathrm{div},\Omega)$ we have that*

$$\boldsymbol{H}_{0,N}(\mathrm{div},\Omega) = \left\{ \boldsymbol{q} \in \boldsymbol{Y}(\Omega) : \sum_r \langle \boldsymbol{q}\cdot\boldsymbol{n}_r, v\rangle = 0,\ \forall v \in H^1_{0,D}(\Omega) \right\}\,. \tag{2.15}$$

We can conclude that a function $\boldsymbol{q} \in \boldsymbol{Y}(\Omega)$ is in $\boldsymbol{H}(\mathrm{div},\Omega)$ if and only if the normal trace is continuous at the interface.

2.1.2 Change of variable

In the framework of finite element approximation, the coordinate changes play a fundamental role, allowing the use of a reference element in which the computation is performed. In this section, the Piola's transformation is introduced, in order to equip the finite element discretization with an essential component.

Consider a reference domain $\hat{K} \subset \mathbb{R}^n$. Let $\partial\hat{K}$ be its boundary and denote by $\hat{\boldsymbol{n}}$ the outward-oriented normal. Additionally, since we are dealing with integral relations, we define $d\hat{\boldsymbol{x}}$ the Lebesgue measure on $\hat{K}$ and with $d\hat{\boldsymbol{\sigma}}$ the superficial measure on $\partial\hat{K}$. We can introduce a smooth mapping $\mathcal{F} : \mathbb{R}^n \to \mathbb{R}^n$, where 'smooth' implies at least C^1 continuity. Consequently, the mapping between the element and the reference element is given by

$$K = \mathcal{F}(\hat{K})\,. \tag{2.16}$$

In this context, it can be asserted that the element K is the image of $\hat{K}$ under the diffeomorphism $\mathcal{F}$.

We define with $\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})$ the Jacobian matrix of the transformation, assuming that $\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})$ is invertible for any $\hat{\boldsymbol{x}}$ and that $\mathcal{F}$ is globally invertible on the element K. Therefore, the following relationship holds

$$\mathcal{D}\mathcal{F}^{-1}(\hat{\boldsymbol{x}}) = (\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}}))^{-1}\,. \tag{2.17}$$

In the case where the transformation $\mathcal{F}(\hat{\boldsymbol{x}})$ is linear, i.e. $\mathcal{F}(\hat{\boldsymbol{x}}) = \boldsymbol{x}_0 + B\hat{\boldsymbol{x}}$, the map is said affine. In addition, the Jacobian matrix is constant, $\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}}) = B$. We report now some definitions and properties regarding the transformation and its related Jacobian matrix. Specifically, we define the $L^\infty(\hat{K})$ norm of a function $\hat{\boldsymbol{x}} \to ||\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})||$, namely the matrix norm of $\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})$, as

$$||\mathcal{D}\mathcal{F}||_\infty := \sup_{\hat{\boldsymbol{x}}\in\hat{K}} \left(\sup_{\xi\in\mathbb{R}^n} \frac{|\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})\xi|}{|\xi|_{\mathbb{R}^n}} \right)\,. \tag{2.18}$$

Likewise, we have the norm of the inverse matrix defined as

$$||\mathcal{D}\mathcal{F}^{-1}||_\infty := \sup_{\boldsymbol{x}\in K} \left(\sup_{\xi\in\mathbb{R}^n} \frac{|(\mathcal{D}\mathcal{F}^{-1}(\boldsymbol{x}))\xi|}{|\xi|_{\mathbb{R}^n}} \right)\,. \tag{2.19}$$

Finally, we define the determinant of the matrix as

$$\mathcal{J}(\hat{\boldsymbol{x}}) := |\det \mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})|\,, \tag{2.20}$$

where, for $\hat{\boldsymbol{x}} \in \partial\hat{K}$ we also have

$$\mathcal{J}_{\hat{\boldsymbol{n}}}(\hat{\boldsymbol{x}}) := \mathcal{J}(\hat{\boldsymbol{x}})||(\mathcal{D}\mathcal{F}^{-1})^t\hat{\boldsymbol{n}}||_{\mathbb{R}^n} . \tag{2.21}$$

We now apply the previous definitions, with functions belonging to the classical spaces already introduced.

Sobolev Spaces $H^s(\Omega)$ Functions

We introduce now some properties related to generic functions $v \in H^s(\Omega)$, where the index s represents an integer number. Consider a scalar function defined on $\hat{K}$, i.e. $\hat{v}(\hat{\boldsymbol{x}})$. On the real element K, we define $v(\boldsymbol{x}) = F(\hat{v}(\hat{\boldsymbol{x}}))$ as

$$v := \hat{v} \circ \mathcal{F}^{-1} . \tag{2.22}$$

Therefore, we have that the gradient reads as

$$\boldsymbol{\nabla} v = (\mathcal{D}\mathcal{F}^{-1})^t \boldsymbol{\nabla}\hat{v} \circ \mathcal{F}^{-1} = F((\mathcal{D}\mathcal{F}^{-1})^t \boldsymbol{\nabla}\hat{v}) , \tag{2.23}$$

with the following integral relations

$$\int_K F(\hat{v})\, d\boldsymbol{x} = \int_{\hat{K}} \hat{v}\mathcal{J}\, d\hat{\boldsymbol{x}} , \tag{2.24}$$

$$\int_{\partial K} F(\hat{v})\, d\boldsymbol{\sigma} = \int_{\partial\hat{K}} \hat{v}\mathcal{J}_{\hat{\boldsymbol{n}}}\, d\hat{\boldsymbol{\sigma}} . \tag{2.25}$$

Without the proof, some inequalities regarding the semi-norms on the element and the reference element are now reported.

Lemma 2.1. *The mapping described by F is an isomorphism from $L^2(\hat{K})$ onto $L^2(K)$ and from $H^1(\hat{K})$ onto $H^1(K)$, where the following inequalities hold:*

$$|v|_{0,K} \leq \left(\sup_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}})\right)^{1/2} |\hat{v}|_{0,\hat{K}} , \tag{2.26}$$

$$|\hat{v}|_{0,\hat{K}} \leq \left(\inf_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}})\right)^{-1/2} |v|_{0,K} , \tag{2.27}$$

$$|v|_{1,K} \leq \left(\sup_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}})\right)^{1/2} ||\mathcal{D}\mathcal{F}^{-1}||_\infty |\hat{v}|_{1,\hat{K}} , \tag{2.28}$$

$$|\hat{v}|_{1,\hat{K}} \leq \left(\inf_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}})\right)^{-1/2} ||\mathcal{D}\mathcal{F}||_\infty |v|_{0,K} . \tag{2.29}$$

Moreover, if the mapping $\mathcal{F}$ between real and reference element is affine, following [38], we have that

$$|v|_{m,K} \leq c(\det B)^{-\frac{1}{2}}||B||^m|\hat{v}|_{m,\hat{K}}\,, \tag{2.30}$$

and in the same fashion

$$|\hat{v}|_{m,\hat{K}} \leq c(\det B)^{-\frac{1}{2}}||B||^m|v|_{m,K}\,, \tag{2.31}$$

with the constant c depending only on the space dimension n and m.

Space $H(\mathrm{div},\Omega)$ Functions

One of the main features of functions in $H(\mathrm{div},\Omega)$ is the use of normal components as degrees of freedom. In fact, remembering the Proposition 2.2, we understand how a function belonging in $\boldsymbol{H}_{0,N}(\mathrm{div},\Omega)$ is defined as the summation over every element such that the dot product between the fluxes of the function, i.e. $\boldsymbol{v}\cdot\boldsymbol{n}_\Gamma$, and the test function is equal to zero. On the other hand, the transformation previously described does not preserve the normal components, and besides, we are not able to map $H(\mathrm{div},\hat{K})$ into $H(\mathrm{div},K)$. For this reason, it is necessary to introduce the Piola's transformation, which is a contravariant transformation. For a vector-valued function $\hat{\boldsymbol{q}}\in(H^1(\hat{K}))^n$ we consider the mapping

$$\mathcal{G}(\hat{\boldsymbol{q}})(\boldsymbol{x}) := \frac{1}{\mathcal{J}(\hat{\boldsymbol{x}})}\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\,, \quad \text{with} \quad \boldsymbol{x}=\mathcal{F}(\hat{\boldsymbol{x}})\,. \tag{2.32}$$

In a bi-dimensional case, the previous definition takes the following form

$$\begin{pmatrix} \frac{\partial q_1}{\partial x} & \frac{\partial q_1}{\partial y} \\ \frac{\partial q_2}{\partial x} & \frac{\partial q_2}{\partial y} \end{pmatrix} = \frac{1}{\mathcal{J}}(\mathcal{D}\mathcal{F}) \begin{pmatrix} \frac{\partial \hat{q}_1}{\partial x} & \frac{\partial \hat{q}_1}{\partial y} \\ \frac{\partial \hat{q}_2}{\partial x} & \frac{\partial \hat{q}_2}{\partial y} \end{pmatrix} (\mathcal{D}\mathcal{F}^{-1})\,, \tag{2.33}$$

where, since we have an invariance of the trace matrix with a change of variable, the following relation holds

$$\boldsymbol{\nabla}\cdot\boldsymbol{q} = \frac{1}{\mathcal{J}}\boldsymbol{\nabla}\cdot\hat{\boldsymbol{q}}\,. \tag{2.34}$$

Following the idea presented in [39, 40] we report other important lemmas.

Lemma 2.2. *Consider the transformations $v = F(\hat{v})$ and $\boldsymbol{q} = \mathcal{G}(\hat{\boldsymbol{q}})$. The following integral relations hold*

$$\int_K \boldsymbol{q} \cdot \boldsymbol{\nabla} v \, d\boldsymbol{x} = \int_{\hat{K}} \hat{\boldsymbol{q}} \cdot \boldsymbol{\nabla} \hat{v} \, d\hat{\boldsymbol{x}} \qquad \forall\, \hat{v} \in L^2(\hat{K})\,, \tag{2.35}$$

$$\int_K v \boldsymbol{\nabla} \cdot \boldsymbol{q} \, d\boldsymbol{x} = \int_{\hat{K}} \hat{v} \boldsymbol{\nabla} \cdot \hat{\boldsymbol{q}} \, d\hat{\boldsymbol{x}} \qquad \forall\, \hat{v} \in L^2(\hat{K})\,, \tag{2.36}$$

$$\int_{\partial K} \boldsymbol{q} \cdot \boldsymbol{n} v \, d\boldsymbol{\sigma} = \int_{\partial \hat{K}} \hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{n}} \hat{v} \, d\hat{\boldsymbol{\sigma}} \qquad \forall\, \hat{v} \in L^2(\partial \hat{K})\,. \tag{2.37}$$

Note that, from the last relation in Lemma 2.2, with the new transformation $\mathcal{G}$ the normal trace in $H^{-\frac{1}{2}}$ is preserved.

Lemma 2.3. *The mapping $\mathcal{G}$ is an isomorphism of $H(\mathrm{div}, \hat{K})$ onto $H(\mathrm{div}, K)$ and of $H^0(\mathrm{div}, \hat{K})$ onto $H^0(\mathrm{div}, K)$, for which holds that*

$$|\boldsymbol{q}|_{0,K} \leq \left(\inf_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}}) \right)^{-1/2} ||\mathcal{DF}||_\infty |\hat{\boldsymbol{q}}|_{0,\hat{K}}\,, \tag{2.38}$$

$$|\hat{\boldsymbol{q}}|_{0,\hat{K}} \leq \left(\sup_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}}) \right)^{1/2} ||\mathcal{DF}^{-1}||_\infty |\boldsymbol{q}|_{0,K}\,, \tag{2.39}$$

$$|\boldsymbol{\nabla} \cdot \boldsymbol{q}|_{0,K} \leq \left(\inf_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}}) \right)^{-1/2} |\boldsymbol{\nabla} \cdot \hat{\boldsymbol{q}}|_{0,\hat{K}}\,, \tag{2.40}$$

$$|\boldsymbol{\nabla} \cdot \hat{\boldsymbol{q}}|_{0,\hat{K}} \leq \left(\sup_{\hat{\boldsymbol{x}}} \mathcal{J}(\hat{\boldsymbol{x}}) \right)^{1/2} |\boldsymbol{\nabla} \cdot \boldsymbol{q}|_{0,K}\,. \tag{2.41}$$

In the case the transformation $\mathcal{F}$ is affine and $\boldsymbol{q} \in \boldsymbol{H}^m(\mathrm{div}, \Omega)$, i.e.

$$\boldsymbol{H}^m(\mathrm{div}, \Omega) := \{\boldsymbol{q} \in (H^m(\Omega))^n : \boldsymbol{\nabla} \cdot \boldsymbol{q} \in H^m\}\,, \tag{2.42}$$

the following lemma holds

Lemma 2.4. *Having a vector function $\boldsymbol{q} \in \boldsymbol{H}^m(\mathrm{div}, \Omega)$ and an affine mapping $\mathcal{F}$, since $B = \mathcal{DF}$, we have that*

$$|\boldsymbol{q}|_{m,K} \leq (\det B)^{-\frac{1}{2}} ||B^{-1}||^m ||B|| |\hat{\boldsymbol{q}}|_{m,\hat{K}}\,, \tag{2.43}$$

$$|\boldsymbol{\nabla} \cdot \boldsymbol{q}|_{m,K} \leq (\det B)^{-\frac{1}{2}} ||B^{-1}||^m |\boldsymbol{\nabla} \cdot \hat{\boldsymbol{q}}|_{m,\hat{K}}\,. \tag{2.44}$$

Having the definition of the transformation $\mathcal{F}$, we can now define also the normal vector and the tangent vector on ∂K. Considering the unit normal $\hat{\boldsymbol{n}}$ to $\partial \hat{K}$, we have that

$$\boldsymbol{n}(\boldsymbol{x}) = \frac{[\mathcal{DF}]^{-T} \cdot \hat{\boldsymbol{n}}(\hat{\boldsymbol{x}})}{|| \, [\mathcal{DF}]^{-T} \cdot \hat{\boldsymbol{n}}(\hat{\boldsymbol{x}}) \, ||}\,. \tag{2.45}$$

In the same way, considering the unit tangent vector $\hat{\boldsymbol{t}}$ to $\partial\hat{K}$, we have that

$$\boldsymbol{t}(\boldsymbol{x}) = \frac{\mathcal{D}\mathcal{F}\cdot\hat{\boldsymbol{t}}(\hat{\boldsymbol{x}})}{||\mathcal{D}\mathcal{F}\cdot\hat{\boldsymbol{t}}(\hat{\boldsymbol{x}})||}\,. \tag{2.46}$$

2.1.3 De Rham Diagram

With the name of De Rham Diagram we refer to a mathematical tool that has been investigated and developed for the description of the setting for functional space and finite element approximation. An important review of this topic can be found in [41, 42, 43, 44].

Following [8], we start considering the sequence of mapping calling de Rham complex in the two-dimensional case, considering a minimal L^2 smoothness

$$\mathbb{R} \to H^1(\Omega) \xrightarrow{\mathbf{curl}} \boldsymbol{H}(\mathrm{div},\Omega) \xrightarrow{\mathrm{div}} L^2(\Omega) \to 0\,, \tag{2.47}$$

where, naturally, the curl operator is defined as $\mathbf{curl} := (\partial/\partial x_2, -\partial/\partial x_1)^t$. Following classical literature results [3], the complex is said exact if the domain Ω is simply connected, i.e. each operator has a range that is the kernel of the succeeding one. To fix the main ideas, we have that

- a curl-free function $f \in H^1(\Omega)$ implies that f is constant;
- a solenoidal vector-valued function $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega)$ is equal to the curl of f, i.e. $\boldsymbol{v} = \mathbf{curl} f$, with $f \in H^1(\Omega)$;
- the map operator div: $\boldsymbol{H}(\mathrm{div},\Omega) \to L^2(\Omega)$ is a surjective function.

Regarding a finite element analysis, we can define a subcomplex of (2.47), considering a finite element space $\Upsilon_h \subset H^1(\Omega)$, $\boldsymbol{W}_h \subset \boldsymbol{H}(\mathrm{div},\Omega)$ and $Q_h \subset L^2(\Omega)$ characterized by the following mapping

$$\mathbb{R} \to \Upsilon_h \xrightarrow{\mathbf{curl}} \boldsymbol{W}_h \xrightarrow{\mathrm{div}} Q_h \to 0\,. \tag{2.48}$$

Naturally, considering standard conforming finite element spaces we obtain a discrete complex of (2.47) [44, 45]. This subcomplex, when the domain Ω is simply connected, becomes exact as the continuous setting. This means that if we consider a finite element pair $\Upsilon_h \times \boldsymbol{W}_h$ ($\boldsymbol{W}_h \times Q_h$), we obtain a stable finite element pair with respect to the curl (divergence) operator. For instance, consider the well-known Lagrange finite element space as Υ_h, the Raviart-Thomas or Brezzi-Douglas-Marini finite element space (which are

described in the next sections) for $\boldsymbol{W}_h$ space and the space of discontinuous piecewise polynomial functions for Q_h. Hence, being equipped with exactness properties of the subcomplex means that, for example, the operator div : $\boldsymbol{W}_h \to Q_h$ is a surjective function, that also has a bounded right-inverse independent of h. Thus, the inf-sup condition follows as

$$\sup_{\boldsymbol{w}\in\boldsymbol{W}_h} \frac{(\boldsymbol{\nabla}\cdot\boldsymbol{w}, q)}{||\boldsymbol{w}||_{H(\mathrm{div})}} \geq \beta||q||_{L^2(\Omega)} \quad \forall q \in Q_h. \tag{2.49}$$

Because of the minimal smoothness equipping the Hilbert spaces, the complex (2.47) does not apply well to the Stokes problem. In fact, referring to [13, 46, 47], a Stokes complex has been introduced and defined as

$$\mathbb{R} \to H^2(\Omega) \xrightarrow{\mathbf{curl}} \boldsymbol{H}^1(\Omega) \xrightarrow{\mathrm{div}} L^2(\Omega) \to 0\,, \tag{2.50}$$

where the exactness of the subcomplex results, as above, from a simply connected domain. In particular, we have that for an incompressible fluid, divergence-free functions $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ are connected to the stream-function $z \in H^2(\Omega)$, since $\boldsymbol{v} = \mathbf{curl} z$ holds. Also in this case, the generic inf-sup condition (1.33) is valid, since the operator div : $\boldsymbol{H}^1(\Omega) \to L^2(\Omega)$ is a surjection.

Similarly to the complex (2.47), it is possible to find the finite element setting for the Stokes complex (2.50), described with the following subcomplex of standard finite element spaces

$$\mathbb{R} \to \Sigma_h \xrightarrow{\mathbf{curl}} \boldsymbol{X}_h \xrightarrow{\mathrm{div}} Y_h \to 0\,, \tag{2.51}$$

where, naturally we have that $\Sigma_h \subset H^2(\Omega)$, $\boldsymbol{X}_h \subset H^1(\Omega)$ and $Y_h \subset L^2(\Omega)$. This last subcomplex, if exact, produces a finite element pair $\boldsymbol{X}_h \times Y_h$ satisfying the discrete inf-sup condition with a bounded right-inverse. A mapping such as (2.51) satisfies the property that $\boldsymbol{\nabla}\boldsymbol{X}_h = W_h$, and hence the considered finite element pair implies divergence-free approximations. The previous mapping can be exploited to develop a pair $\boldsymbol{X}_h \times Y_h$ satisfying these properties. Note the presence, for this pair, of $H^2(\Omega)$-conforming relatives that describe the local and global properties of these spaces. Therefore, the subcomplex (2.50) is satisfied by every divergence-free finite element pair equipped with an H^2 relative.

2.2 Finite Element Approximation

2.2.1 Space $H^1(\Omega)$

Standard method on Affine Meshes In this section, we revisit fundamental properties related to finite element approximation within $H^1(\Omega)$, particularly concerning conforming methods. With these methods, we mean the use of finite dimensional subspaces of the approximated function space, hence also said as an internal approximation (for the nonconforming method an extension of the function space to approximate is performed in order to find a new finite element subspace, and therefore in this sense is considered as an external approximation).

If we consider a standard partition of the domain Ω with polygonal elements, the conforming approximation in $H^1(\Omega)$ is translated to a space of continuous functions characterized by a finite number of degrees of freedom. In particular, this condition is satisfied by using piecewise polynomial functions, through which a function approximation inside an element reads as

$$v_h|_K = \hat{v} \circ \mathcal{F}^{-1}\,, \tag{2.52}$$

where the standard transformation to the real element is represented by $K = \mathcal{F}(\hat{K})$ and $\hat{v}$ is the approximating polynomial function. If the continuity is reached considering a specific degree of freedom for $\hat{v}$, the resulting $v_h|_K$ is still a polynomial function only with affine transformation. Given an element

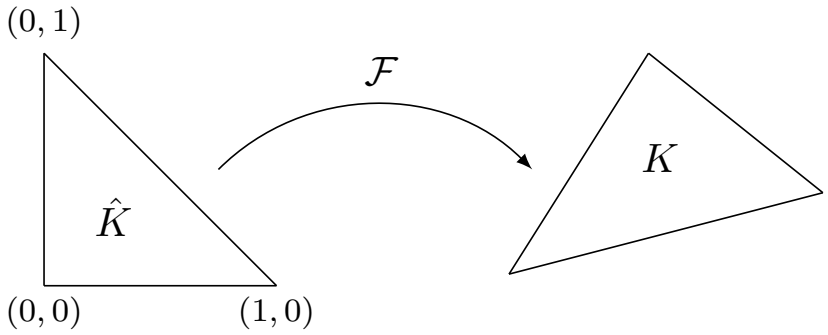

Figure 2.2: Affine transformation of a triangle.

K and the space $\mathcal{P}_k(K)$ of polynomial functions with degree $\leq k$, an affine finite element can be represented for instance in Figure 2.2 by using the affine

transformation

$$\mathcal{F}(\hat{\boldsymbol{x}}) = \boldsymbol{x}_0 + B\hat{\boldsymbol{x}}\,, \tag{2.53}$$

for which, since $\det B \neq 0$, the real element K is still a triangle. Choosing a suitable order of degree of freedom, the reference polynomial function reads $\hat{\mathcal{P}} = \mathcal{P}_k(\hat{K})$. Moreover, the dimension of $\mathcal{P}_k(K)$ for a bi-dimensional case is equal to $\frac{1}{2}(k+1)(k+2)$. Note that, when triangular elements are considered, the transformation is always an affine isomorphism.

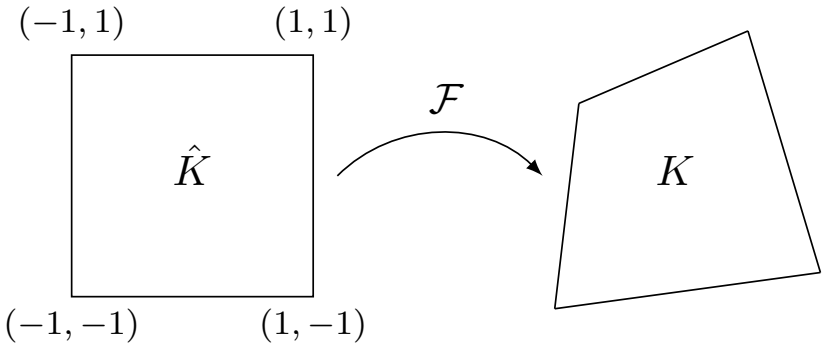

Figure 2.3: Affine transformation of a quadrilateral. Specifically, it is represented by a $\mathcal{Q}_1$ isoparametric element.

Another classical finite element family is the isoparametric quadrilateral element. The reference element in Figure 2.3 is the square $\hat{K} = [-1,1] \times [-1,1]$. In addition, we can define with $n = 2$, the space of polynomials with degree $\leq k_1$ for x_1 and $\leq k_2$ for x_2 as

$$\mathcal{P}_{k_1,k_2}(K) := \left\{ p(x_1,x_2) : p(x_1,x_2) = \sum_{i\leq k_1, j\leq k_2} a_{ij} x_1^i x_2^j \right\}. \tag{2.54}$$

Therefore, defining also $\mathcal{Q}_k(K) := \mathcal{P}_{k,k}(K)$, always for a bidimensional case, the previous quadrilateral is equipped with $\hat{\mathcal{P}} = \mathcal{Q}_k(\hat{K})$, where the transformation $\mathcal{F}$ is applied to each component of $\mathcal{Q}_k(\hat{K})$. Note that, in order to have a general straight-side quadrilateral as in Figure 2.3, the transformation can be considered as $\mathcal{F} \in (\mathcal{Q}_1(\hat{K}))^2$. Furthermore, consider that an affine isomorphism applied on a quadrilateral element can produce only parallelograms, therefore we are not able to mesh every domain (for instance by using parallelograms it is not possible to mesh a triangle). On the contrary, a generic bilinear transformation of a square can produce arbitrary convex quadrilaterals, that can be useful in order to mesh generic polygons [48].

Some considerations are now reported concerning the analysis of a function approximation by using finite element spaces. Consider a general set of degrees of freedom $\{\hat{l}_i\}$ on $\hat{K}$. The interpolate $\hat{r}_h\hat{v}$ of a function v can be defined as

$$\hat{l}_i(\hat{r}_h\hat{v}) := M(\hat{v}) \quad 1 \leq i \leq \dim \hat{\mathcal{P}}\,, \tag{2.55}$$

where a well-defined continuous form is required for the operator M. Therefore, we have the following definition

$$(\hat{r}_h\hat{v})(\hat{a}_i) = \hat{v}(\hat{a}_i)\,, \tag{2.56}$$

which holds only if $\hat{v}$ is a continuous function. For example, considering Lagrange type elements in $\mathbb{R}^2$ or $\mathbb{R}^3$, we have that $\hat{v} \in H^2(\hat{K})$ is a sufficient condition for (2.56), and consequently $\hat{r}\hat{v}$ represents the Lagrange interpolate of $\hat{v}$ in the classical sense. At this point, it is possible to state the following proposition

Proposition 2.3. *Consider an affine mapping $\mathcal{F}$, i.e. $\mathcal{F}(\hat{\boldsymbol{x}}) = \boldsymbol{x}_0 + B\hat{\boldsymbol{x}}$. If $r_h p_k = p_k$, for any $p_k \in \mathcal{P}_k(K)$, the following inequality holds for $v \in H^s(\Omega)$*

$$|v - r_h v|_{m,K} \leq c||B^{-1}||^m||B||^s|v|_{s,K}\,, \tag{2.57}$$

with $m \leq s$, $1 < s \leq k+1$.

An important lemma is now reported

Lemma 2.5 (Bramble-Hilbert's lemma)**.** *Consider a continuous linear form L on $H^{k+1}(\Omega)$, such that $L(p_k) = 0$, for any $p_k \in \mathcal{P}_k(\Omega)$. Hence, there exists a constant c, which depends on L and Ω, such that*

$$|L(v)| \leq c|v|_{k+1,\Omega}\,. \tag{2.58}$$

In order to expand this discussion, it is important to introduce some variables related to finite element discretization. For this reason, giving h_K the diameter of K and ρ_K the diameter of the largest ball that can be inscribed in the element K, the regularity assumption for the finite element requires that

$$\frac{h_K}{\rho_K} \leq \sigma\,, \tag{2.59}$$

where σ is called therefore regularity constant. In particular, the family partition of the domain $\mathcal{T}_h$ is said shape-regular uniformly in h if

$$\sigma_K \leq C \quad \forall K \in \mathcal{T}_h\,. \tag{2.60}$$

The previous condition can also be reformulated as a condition on the angle element, excluding degenerate ones. The following corollary holds

Corollary 2.1. *Consider a regular family of affine partitions* $(\mathcal{T}_h)_{h\geq 0}$. *Hence, there exists a constant* c *which depends on* k *and* σ *such that*

$$|v - r_h v|_{1,K} \leq c h^k |v|_{k+1,\Delta K}\,, \tag{2.61}$$

where ΔK *is defined as*

$$\Delta K := \left\{ K' : \bar{K}' \cap \bar{K} \neq 0 \right\}\,. \tag{2.62}$$

Quadrilateral Finite Element on Non-Affine Meshes In this section, we underline some results regarding non-affine meshes and how they can potentially lead to problematic properties in the finite element approximation of elliptic equations. A domain triangulation denoted as $\mathcal{T}_h$, serves as the mesh, and it is defined as affine if the mapping $\mathcal{F}_K$ from the reference element $\hat{K}$ to the real element K is represented by affine functions for every element $K \in \mathcal{T}_h$. For such cases, the Jacobian matrix $\mathcal{D}\mathcal{F}_K$ becomes constant. However, non-affine meshes, equipped therefore with a nonconstant Jacobian matrix associated with the elements, may result in a degradation of the approximation properties. In this context, only quadrilateral or hexahedral elements can be characterized by this problematic aspect. Specifically, this topic has been investigated in the context of quadrilateral finite element in $\boldsymbol{H}(\mathrm{div}, \Omega)$ for vector-valued functions in a bidimensional domain in [49] and in $H(\mathrm{div}, \Omega)$ for scalar functions in [48].

Consider $\hat{K}$ a reference triangle, with its image $K = \mathcal{F}_k(\hat{K})$. If we assume a smooth function $\hat{v} : \hat{K} :\to \mathbb{R}$, we obtain the corresponding mapped function

$$v = \hat{v} \circ F_K^{-1} : K \to \mathbb{R}\,, \tag{2.63}$$

equipped with the linear interpolant $r_h v : K \to \mathbb{R}$. In the case of an affine mapping F_K (triangle with straight sides), since the Jacobian matrix is constant, i.e. $\mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}}) = B$, considering the Corollary 2.1 we have

$$|v - r_h v|_{1,K} \leq c h_K |v|_{2,K}\,. \tag{2.64}$$

Furthermore, we know from the estimate (2.31) that the following estimate holds

$$|\hat{v}|_{2,K} \leq c|\det(B)|^{-1/2}||B||^2 |v|_{2,K} \leq c|\det(B)|^{-1/2} h_K^2 |v|_{2,K}\,. \tag{2.65}$$

On the contrary, for non-affine mesh, since the relation (2.31) is not valid, the nonconstant Jacobian gives the following estimate

$$|\hat{v}|_{2,K} \leq c|\inf_{\hat{\boldsymbol{x}}}(\mathcal{J}(\hat{\boldsymbol{x}}))|^{-1/2}\left(||\mathcal{DF}||^2_{L^\infty(\hat{K})}|v|_{2,K} + ||\mathcal{D}^2\mathcal{F}||_{L^\infty(\hat{K})}|v|_{1,K}\right), \quad (2.66)$$

where, since the term $||\mathcal{DF}||^2_{L^\infty(\hat{K})} \sim O(h_K^2)$, the optimal order approximation can incur in some problems, even though $||\mathcal{D}^2\mathcal{F}||_{L^\infty(\hat{K})}$, i.e. the norm of the Hessian matrix, can be a lower order term.

2.2.2 Space $H(\mathrm{div}, \Omega)$

In this section, we aim to describe a suitable criterion to understand if a finite element space is a subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$. Therefore, considering the shape-regular triangulation $\mathcal{T}_h$ of the domain Ω, we define with ε_h the set of edges or faces, considering respectively a bi-dimensional or a three-dimensional domain. Moreover, we define the set of boundary edges/faces with $\varepsilon_h^B \subset \varepsilon_h$, namely the variable $e \in \varepsilon_h^B$ if $e \cap \partial\Omega \neq 0$, and the set of interior edges/faces with $\varepsilon_h^I := \varepsilon_h \backslash \varepsilon_h^B$. An important lemma regarding the normal component of finite element subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$ is now reported without giving the proof [8].

Lemma 2.6. *Let $\mathcal{T}_h$ the partitioned domain Ω and consider a space of piecewise polynomials function $\boldsymbol{W}_h$. Since $\boldsymbol{W}_h \subset \boldsymbol{H}(\mathrm{div}, \Omega)$, we have continuity across interelement boundaries $e \in \varepsilon_h^I$ of the normal components.*

Naturally, the space $\boldsymbol{W}_h$ represents in the context of the Stokes equation the finite element space for the velocity field. Furthermore, this lemma does not hold necessarily for the tangential components. Among the finite element spaces satisfying the previous lemma, i.e. conforming subspaces of $\boldsymbol{H}(\mathrm{div}, \Omega)$, we can introduce the Raviart-Thomas space of order $k \geq 0$ ($\boldsymbol{\mathcal{RT}}_k$) and the Brezzi-Douglas-Marini space of degree $k \geq 1$ ($\boldsymbol{\mathcal{BDM}}_k$) [50, 51, 52, 53]. Given an element K of the partitioned domain Ω, the local Raviart-Thomas space of order $k \geq 0$ is defined as

$$\boldsymbol{\mathcal{RT}}_k(K) = \mathcal{P}_k(K)^n + \boldsymbol{x}\mathcal{P}_k(K)\,. \quad (2.67)$$

Therefore, we define the spaces as

$$\boldsymbol{\mathcal{RT}}_k := \{\boldsymbol{w}_h \in \boldsymbol{H}_0(\mathrm{div}, \Omega) : \boldsymbol{w}_h|_K \in \boldsymbol{\mathcal{RT}}_k(K) \quad \forall K \in \mathcal{T}_h\}\,, \quad (2.68)$$

$$\boldsymbol{\mathcal{BDM}}_k := \{\boldsymbol{w}_h \in \boldsymbol{H}_0(\mathrm{div}, \Omega) : \boldsymbol{w}_h|_K \in \boldsymbol{\mathcal{P}}_k(K) \quad \forall K \in \mathcal{T}_h\}\,, \quad (2.69)$$

where, we have used the space $\boldsymbol{H}_0(\mathrm{div},\Omega) = \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega) : \boldsymbol{v}\cdot\boldsymbol{n}|_{\partial\Omega} = 0\}$. Moreover, $\boldsymbol{\mathcal{P}}_k = \mathcal{P}_k^n$ ($\mathcal{P}_k$) represents the space of globally continuous vector (scalar)-valued piecewise polynomials of degree not exceeding k.

Specifically, to introduce these new approximating spaces a triangular element is considered now as a reference element. However, note that the same spaces on rectangular elements are equipped with analogous properties, therefore the error estimates that are presented are valid for both elements.

We report now some basic properties regarding the space $\boldsymbol{\mathcal{RT}}_k(K)$. Let F_i, $i = 1,\dots,n+1$ the edges of the element K, with $\boldsymbol{n}_i$ their outward unit normal. The following lemma holds

Lemma 2.7. *The following statements hold:*

(a)

$$\begin{aligned} dim\boldsymbol{\mathcal{RT}}_k(K) &= n\binom{k+n}{k} + \binom{k+n-1}{k} = \\ &= \begin{cases} (k+1)(k+3) & \text{for } n=2\,, \\ \frac{1}{2}(k+1)(k+2)(k+4) & \text{for } n=3\,. \end{cases} \end{aligned} \tag{2.70}$$

(b) If $\boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k(K)$, we have that $\boldsymbol{v}\cdot\boldsymbol{n}_i \in \mathcal{P}_k(F_i)$ for $i = 1,\dots,n+1$.

(c) If $\boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k(K)$, satisfies the relation $\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0$, we have that $\boldsymbol{v} \in \mathcal{P}_k^n$.

The idea is now to construct an interpolator operator able to map these basis functions, i.e. we define Π_K as

$$\Pi_K : H^1(K)^n \to \boldsymbol{\mathcal{RT}}_k(K)\,. \tag{2.71}$$

It is notable that we consider the interpolation of $H^1(\Omega)^n$ functions, even though this remains valid for less regular functions. In particular, the definition of Π_K is related to the following lemma

Lemma 2.8. *Let a function $\boldsymbol{v} \in \boldsymbol{H}^1(K)$. Therefore, there exists a unique interpolator $\Pi_K\boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k(K)$ that satisfies*

$$\int_{F_i} \Pi_K\boldsymbol{v}\cdot\boldsymbol{n}_i p_k\,ds = \int_{F_i} \boldsymbol{v}\cdot\boldsymbol{n}_i p_k\,ds \quad \forall\, p_k \in \mathcal{P}_k(F_i),\ i = 1,\dots,n+1\,. \tag{2.72}$$

Moreover, if $k \geq 1$ we have

$$\int_K \Pi_K\boldsymbol{v}\cdot\boldsymbol{p}_{k-1}\,d\boldsymbol{x} = \int_K \boldsymbol{v}\cdot\boldsymbol{p}_{k-1}\,d\boldsymbol{x} \quad \forall\, \boldsymbol{p}_{k-1} \in \mathcal{P}_k^n(K)\,. \tag{2.73}$$

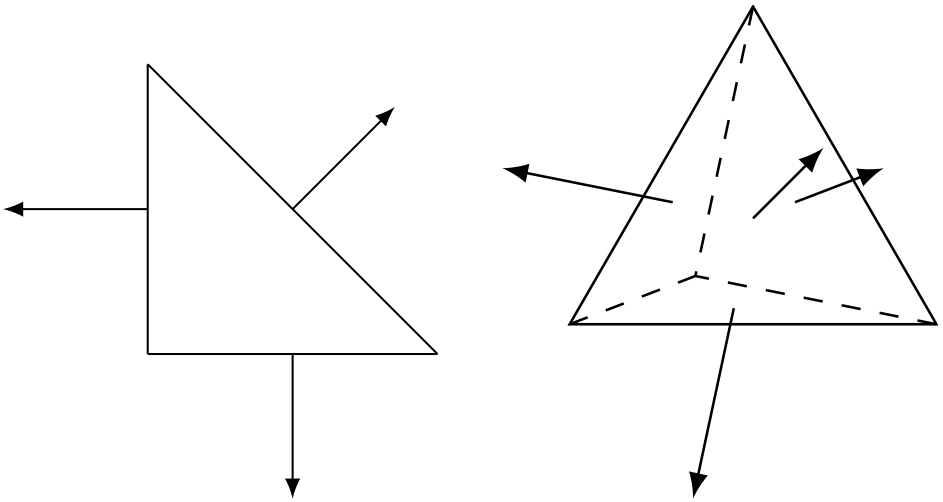

Figure 2.4: Lowest order Raviart-Thomas triangular element on the left and tetrahedral element on the right.

We present now the theorem [50] concerning the optimal order estimates for the interpolation with the Raviart-Thomas basis function.

Theorem 2.1. *For every function* $\boldsymbol{v} \in \boldsymbol{H}^m(K)$ *with* $1 \leq m \leq k+1$*, there exists a constant* C *which depends on* k*,* n *(the dimension of* $\boldsymbol{v}$*) and the regularity constant* σ *such that*

$$||\boldsymbol{v} - \Pi_K \boldsymbol{v}||_{L^2(K)} \leq C h_K^m ||\boldsymbol{\nabla}^m \boldsymbol{v}||_{L^2(K)} \,. \tag{2.74}$$

From the previous definitions, the global Raviart-Thomas finite element spaces can be introduced. In this case, the mesh-size h has to be defined as $h = \max_{K \in \mathcal{T}_h} h_K$. Moreover, we assume that the hypothesis of regularity condition (2.59) is satisfied in any element $K \in \mathcal{T}_h$. Therefore, the global space can be defined as

$$\boldsymbol{\mathcal{RT}}_k(\mathcal{T}_h) := \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega) : \boldsymbol{v}|_K \in \boldsymbol{\mathcal{RT}}_k(K) \quad \forall k \in \mathcal{T}_h\} \,. \tag{2.75}$$

We now introduce the global operator

$$\Pi_h : \boldsymbol{H}(\mathrm{div}, \Omega) \cap \prod_{K \in \mathcal{T}_h} H^1(K)^n \to \boldsymbol{\mathcal{RT}}_k \,, \tag{2.76}$$

where specifically

$$\Pi_h \boldsymbol{v}|_K = \Pi_K \boldsymbol{v} \quad \forall K \in \mathcal{T}_h \,. \tag{2.77}$$

Recalling Lemma 2.8, it is important to note that by definition we have $\Pi_K \boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k(K)$. To verify whether $\Pi_h \boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k$, we need to ensure that $\Pi_h \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega)$. In particular, applying the divergence theorem, we know that the continuity of the normal component across the edges elements is a necessary condition for a piecewise polynomial vector-valued function to be in $\boldsymbol{H}(\mathrm{div}, \Omega)$. Therefore, the continuity of the normal component of $\Pi_h \boldsymbol{v}$, since $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega)$, follows from the property (b) of Lemma 2.7 considering the degrees of freedom described in (2.72) defining Π_K.

In order to complete the description of the finite element subspaces required for the approximation of the mixed problem, it is important to introduce the approximating subspace of the scalar variable, which, in our context, represents the pressure. Note that the formulation resulting from the classical Darcy problem does not have any derivatives of the pressure (1.41). In addition, even in the Stokes equation, it is possible to overcome any derivatives of the pressure, since with a standard procedure the strong form of the equation is integrated and changed in the weak one, which does not include the pressure gradient after integration by part of that term. Consequently, the requirement of a continuity property for the pressure does not hold. Given these considerations, we introduce the standard space of piecewise polynomial functions with degree k, not necessarily continuous, as

$$\mathcal{P}_k^d(\mathcal{T}_h) := \{q \in L^2(\Omega) : q|_K \in \mathcal{P}_k(K) \quad \forall K \in \mathcal{T}_h\}\,, \tag{2.78}$$

where the superscript d considers the possibility of discontinuous functions. We report now another lemma regarding properties of Π_h and $\mathcal{P}_k^d$

Lemma 2.9. *Let Π_h be the global Raviart-Thomas operator. Thus, it satisfies*

$$\int_\Omega \boldsymbol{\nabla} \cdot (\boldsymbol{v} - \Pi \boldsymbol{v}) q \, d\boldsymbol{x} = 0 \tag{2.79}$$

$\forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega) \cap \prod_{K \in \mathcal{T}_h} H^1(K)^n$ and $\forall q \in \mathcal{P}_k^d$. In addition, the following identity holds

$$\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{RT}}_k = \mathcal{P}_k^d\,. \tag{2.80}$$

The previous properties and definitions can also be summarized by using the commutative diagram tool. In particular, considering the L^2-projection

$P_h : L^2(\Omega) \to \mathcal{P}_k^d$, we have that

$$\begin{array}{ccccc} H^1(\Omega)^n & \xrightarrow{\text{div}} & L^2(\Omega) & & \\ \Pi_h \downarrow & & \downarrow P_h & & \\ \boldsymbol{\mathcal{RT}}_k & \xrightarrow{\text{div}} & \mathcal{P}_k^d & \longrightarrow 0\,. \end{array} \tag{2.81}$$

Note that in this case to simplify the notation, the subspace $H^1(\Omega)^n$ has been used instead of the more complete space $H(\text{div}, \Omega) \cap \prod_{K \in \mathcal{T}_h} H^1(K)^n$. Having introduced the finite element subspace for the approximation of the elliptic mixed problem (1.41), it is possible now to report some error estimates results for the solution pair $(\boldsymbol{u}_h, p_h) \in \boldsymbol{\mathcal{RT}}_k \times \mathcal{P}_k^d$, for which we recall the finite element discretization

$$\int_\Omega \mu \boldsymbol{u}_h \cdot \boldsymbol{v}\, d\boldsymbol{x} - \int_\Omega p_h \boldsymbol{\nabla} \cdot \boldsymbol{v}\, d\boldsymbol{x} = 0 \qquad \forall \boldsymbol{v} \in \boldsymbol{\mathcal{RT}}_k \tag{2.82a}$$

$$\int_\Omega q \boldsymbol{\nabla} \cdot \boldsymbol{u}_h\, d\boldsymbol{x} = \int_\Omega f q\, d\boldsymbol{x} \qquad \forall q \in \mathcal{P}_k^d\,. \tag{2.82b}$$

Lemma 2.10. *Let $\boldsymbol{u}$ and $\boldsymbol{u}_h$ be the solution of the infinite (1.41) and finite (2.82) mixed problem respectively. Then, the following error estimates hold*

$$||\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)} \leq (1 + ||a||_{L^\infty(\Omega)} ||\mu||_{L^\infty(\Omega)}) ||\boldsymbol{u}_h - \Pi_h \boldsymbol{u}||_{L^2(\Omega)}\,. \tag{2.83}$$

Naturally, in this case, μ can be connected to the dynamic viscosity and a to the density of the Darcy/Stokes equations. Both constants are employed for the proof of theorems and lemmas [50]. From this, it is possible to derive an order error estimate for the velocity field $\boldsymbol{u}$

Theorem 2.2. *Let $\boldsymbol{u} \in \boldsymbol{H}^m(\Omega)$ the solution of the problem (1.39), with $1 \leq m \leq k+1$. If there exists a constant C that depends on $||a||_{L^\infty(\Omega)}$, $||\mu||_{L^\infty(\Omega)}$, k, n (the dimension of $\boldsymbol{u}$) and the regularity constant σ, we have*

$$||\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)} \leq C h^m ||\boldsymbol{\nabla}^m \boldsymbol{u}||_{L^2(\Omega)}\,. \tag{2.84}$$

Now, since Theorem 2.1 holds, a particular case leads to

$$||\boldsymbol{v} - \Pi_h \boldsymbol{v}||_{L^2(\Omega)} \leq C h ||\boldsymbol{v}||_{H^1(\Omega)} \quad \forall \boldsymbol{v} \in H^1(\Omega)\,, \tag{2.85}$$

from which we have that [50]

$$||\Pi_h \boldsymbol{v}||_{L^2(\Omega)} \leq ||\boldsymbol{v}||_{H^1(\Omega)}\,. \tag{2.86}$$

Now, starting from the inequalities already introduced in Section 1.4, we add some considerations about the error estimate for the scalar variable p. Indeed, following [50], we present the same error estimate that we found for the pressure-robust mixed method, knowing that now the finite element space for the pressure, namely Y or $\mathcal{P}_k^d$, is represented by piecewise polynomial functions with degree k.

Lemma 2.11. *Consider the infinite and finite pair solution of the Stokes problem* $(\boldsymbol{u}, p)$ *and* $(\boldsymbol{u}_h, p_h)$*. If there exists a constant* C *that depends on* $||a||_{L^\infty(\Omega)}$, $||\mu||_{L^\infty(\Omega)}$, k, n *and the regularity constant* σ*, we have*

$$||p - p_h||_{L^2(\Omega)} \leq C\{||p - P_h p||_{L^2(\Omega)} + ||\boldsymbol{u} - \Pi_h \boldsymbol{u}||_{L^2(\Omega)}\}\,. \tag{2.87}$$

Therefore, for the pressure p we have that

Theorem 2.3. *Let* $(\boldsymbol{u}, p) \in \boldsymbol{H}^m(\Omega) \times H^m(\Omega)$ *the solution of the problem (1.39), with* $1 \leq m \leq k+1$*. If there exists a constant* C *that depends on* $||a||_{L^\infty(\Omega)}$, $||\mu||_{L^\infty(\Omega)}$, k, n *and the regularity constant* σ*, we have*

$$||p - p_h||_{L^2(\Omega)} \leq Ch^m\{||\boldsymbol{\nabla}^m \boldsymbol{u}||_{L^2(\Omega)} + ||\boldsymbol{\nabla}^m p||_{L^2(\Omega)}\}\,. \tag{2.88}$$

Moreover, if the coefficient a is smooth enough, considering the domain Ω as a convex polygon, the a priori estimate

$$||p||_{H^2(\Omega)} \leq C_0 ||f||_{L^2(\Omega)}\,, \tag{2.89}$$

leads to a better estimate of the error $||P_h p - p_h||_{L^2(\Omega)}$.

Lemma 2.12. *Consider the estimate in (2.89) and the coefficient* $a \in W^{1,\infty}$*. If there exists a constant* C *that depends on* $||a||_{W^{1,\infty}}$, $||\mu||_{L^\infty(\Omega)}$, k, n, C_0 *and the regularity constant* σ*, we have*

$$||P_h p - p_h||_{L^2(\Omega)} \leq Ch\{||\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)} + ||\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}\}\,. \tag{2.90}$$

Theorem 2.4. *Consider the estimate in (2.89), the coefficient* $a \in W^{1,\infty}$*, the field* $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega)$ *and the body force* $f \in H^{k+1}(\Omega)$*. If there exists a constant* C *that depends on* $||a||_{W^{1,\infty}}$, $||\mu||_{L^\infty(\Omega)}$, k, n, C_0 *and the regularity constant* σ*, we have*

$$||P_h p - p_h||_{L^2(\Omega)} \leq Ch^{k+2}\{||\boldsymbol{\nabla}^{k+1} \boldsymbol{u}||_{L^2(\Omega)} + ||\boldsymbol{\nabla}^{k+1} f||_{L^2(\Omega)}\}\,. \tag{2.91}$$

Nevertheless, for the sake of completeness, we also define the rectangular Raviart-Thomas element. In particular, giving the non-negative integer k, m, the space of polynomials $\mathcal{Q}_{k,m}$ is defined as

$$q(x,y) = \sum_{i=0}^{k} \sum_{j=0}^{m} a_{ij} x^i y^j \,. \tag{2.92}$$

Thus, the Raviart-Thomas space on a rectangular element R is defined as

$$\boldsymbol{\mathcal{RT}}_k(R) = \mathcal{Q}_{k+1,k}(R) \times \mathcal{Q}_{k,k+1}(R)\,, \tag{2.93}$$

where for the scalar variable the space takes the form of $\mathcal{Q}_k(R)$.

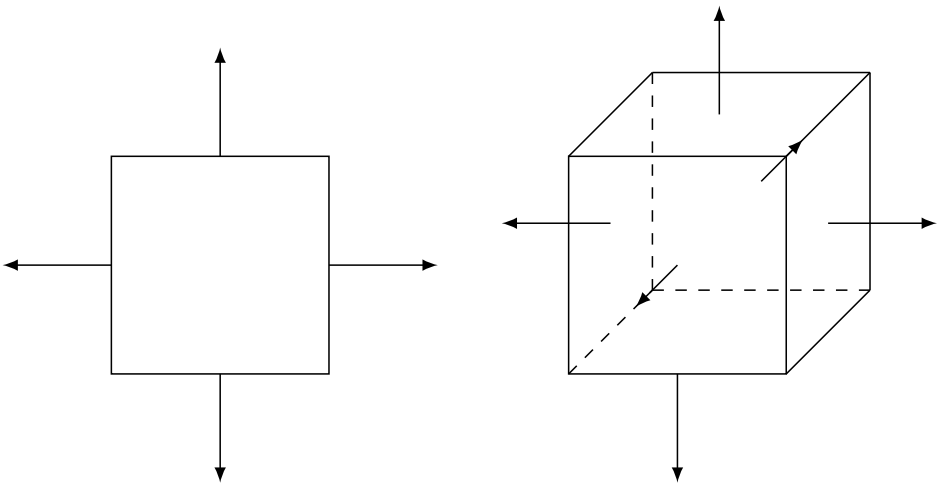

Figure 2.5: Lowest order Raviart-Thomas quadrilateral element on the left and hexahedral element on the right.

Considering also the three-dimensional case (generic hexahedral), the dimension of the space reads

$$dim\boldsymbol{\mathcal{RT}}_k(R) = \begin{cases} 2(k+1)(k+2) & \text{for } n=2\,, \\ 3(k+1)^2(k+2) & \text{for } n=3\,. \end{cases} \tag{2.94}$$

If we denote the four edges of R with $l_i, i = 1,2,3,4$, the interpolant operator is equipped with the following degree of freedom

$$\int_{l_i} \Pi_T \boldsymbol{v} \cdot \boldsymbol{n}_i p_k \, dl = \int_{l_i} \boldsymbol{v} \cdot \boldsymbol{n}_i p_k \, dl \quad \forall p_k \in \mathcal{P}_k(l_i),\ i = 1,2,3,4\,, \tag{2.95}$$

and for $k \geq 1$ we have

$$\int_R \Pi_T \boldsymbol{v} \cdot \phi_k \, d\boldsymbol{x} = \int_R \boldsymbol{v} \cdot \phi_k \, d\boldsymbol{x} \quad \forall \phi_k \in \mathcal{Q}_{k-1,k}(R) \times \mathcal{Q}_{k,k-1}(R)\,. \tag{2.96}$$

Indeed, given the polynomial basis function $\boldsymbol{q} \in \boldsymbol{\mathcal{RT}}_k(R)$, we have that

$$\begin{cases} \boldsymbol{\nabla} \cdot \boldsymbol{q}_k & \in \mathcal{Q}_k \,, \\ \boldsymbol{q} \cdot \boldsymbol{n}_k|_{l_i} & \in \mathcal{P}_k(l_i) \quad \text{for } n = 2 \,, \\ \boldsymbol{q} \cdot \boldsymbol{n}_k|_{f_i} & \in \mathcal{Q}_k(f_i) \quad \text{for } n = 3 \,. \end{cases} \tag{2.97}$$

Moreover, defining the particular space case of

$$\boldsymbol{\mathcal{RT}}_k^0(R) := \{\boldsymbol{q} \in \boldsymbol{\mathcal{RT}}_k : \boldsymbol{\nabla} \cdot \boldsymbol{q} = 0\} \,, \tag{2.98}$$

we have the following lemma.

Lemma 2.13. *Consider a two-dimensional domain and a function* $\boldsymbol{q} \in \boldsymbol{\mathcal{RT}}_k^0(\hat{R})$. *Thus, there exist a function* $\psi \in \mathcal{Q}_{k+1}(\hat{R})$, *that satisfies*

$$\boldsymbol{q} = \boldsymbol{\nabla} \times \psi \,. \tag{2.99}$$

In this case we have that $dim\boldsymbol{\mathcal{RT}}_k^0(\hat{R}) = (k+1)(k+3)$.

In the following analysis, general quadrilateral elements are used instead of triangular ones. In particular, having a generic convex quadrilateral Q as a partition of the domain, the reference element is usually a specific rectangular $R = [-1, 1] \times [-1, 1]$. The coordinate transformation used is the Piola transformation described in Section 2.1.2. Therefore, in order to define the Raviart-Thomas space on the (real) element Q, consider the bilinear transformation $\mathcal{F}$ between the vertices of R and Q, i.e. $\mathcal{F} : R \to Q$. Finally, the local space $\boldsymbol{\mathcal{RT}}_k(Q)$ is defined as

$$\boldsymbol{\mathcal{RT}}_k(Q) := \left\{ \boldsymbol{v} : Q \to \mathbb{R}^2 : \boldsymbol{v}(\boldsymbol{x}) = \mathcal{J}^{-1} \mathcal{D}\mathcal{F}(\hat{\boldsymbol{x}}) \hat{\boldsymbol{v}}(\hat{\boldsymbol{x}}) \right\} \,, \tag{2.100}$$

where naturally $\mathcal{J}$ is the determinant of the Jacobian matrix $\mathcal{D}\mathcal{F}$ associated to the transformation $\mathcal{F}$ and $\hat{\boldsymbol{v}} \in \boldsymbol{\mathcal{RT}}_k(R)$.

The extension of triangular and rectangular finite elements to tetrahedral and hexahedral elements has been deeply described in [54, 51].

In the last part of this section, the finite element families described above are discussed in the framework of free-divergence finite element discretization for the Stokes equation. Indeed, these spaces allow the construction of inf-sup stable pairs, considering naturally an appropriate pressure space.

For instance, we consider again the space of discontinuous piecewise polynomials $\mathcal{P}_k^d$, with the degree k if the velocity space is $\boldsymbol{W}_h = \boldsymbol{\mathcal{RT}}_k$ or degree

$k-1$ if $\boldsymbol{W}_h = \mathcal{BDM}_k$. If in addition, the space $\mathcal{P}_k^d$ has vanishing mean, we have a finite element pair $\boldsymbol{W}_h \times \mathcal{P}_k^d$, such that the inf-sup stability condition holds

$$\inf_{q_h \in \mathcal{P}_k^d \setminus \{0\}} \sup_{\boldsymbol{w}_h \in \boldsymbol{W}_h \setminus \{0\}} \frac{\int_\Omega (\boldsymbol{\nabla} \cdot \boldsymbol{w}_h) q_h d\boldsymbol{x}}{||\boldsymbol{w}_h||_{\boldsymbol{H}(\mathrm{div},\Omega)} ||q_h||_{L^2(\Omega)}} \geq \beta_h \,, \tag{2.101}$$

where we have a uniform bound from below for β_h. Note that, from the space definition, we have that $\boldsymbol{\nabla} \cdot \boldsymbol{W}_h \subseteq \mathcal{P}_k^d$, and therefore discrete divergence-free functions are also weak divergence-free, i.e. the relation

$$\left\{ \boldsymbol{w}_h \in \boldsymbol{W}_h : \int_\Omega (\boldsymbol{\nabla} \cdot \boldsymbol{w}_h) q_h \, d\boldsymbol{x} = 0 \qquad \forall q_h \in \mathcal{P}_k^d \right\} \tag{2.102}$$

is equal to

$$\{ \boldsymbol{w}_h \in \boldsymbol{W}_h : \boldsymbol{\nabla} \cdot \boldsymbol{w}_h \equiv 0 \} \,. \tag{2.103}$$

However, we cannot apply these functions to the Stokes problem because of the lack of smoothness, even though the spaces are equipped with the inf-sup stability considering the $\boldsymbol{H}(\mathrm{div}, \Omega)$ norm and the discrete divergence-free functions are solenoidal. In fact, both subspaces $\boldsymbol{\mathcal{RT}}$ and $\boldsymbol{\mathcal{BDM}}$ do not satisfy the inclusion $\boldsymbol{W}_h \not\subset \boldsymbol{H}_0^1(\Omega)$, and hence we do not have the conformity property with respect to $\boldsymbol{H}_0^1(\Omega)$. Indeed, since the gradient of functions in $\boldsymbol{W}_h$ does not exist globally, we cannot define well the Stokes problem. Moreover, by replacing the gradient in the Stokes formulation (1.65), we obtain that

$$-\int_\Omega \Delta \boldsymbol{v} \cdot \boldsymbol{w}_h \neq a(\boldsymbol{v}, \boldsymbol{w}_h) \tag{2.104}$$

if we consider general functions $\boldsymbol{v} \in \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}_0^1(\Omega)$ and $\boldsymbol{w}_h \in \boldsymbol{W}_h$. Therefore, the method even if non-singular is not convergent because it is not consistent.

2.2.3 Quadrilateral approximation of scalar functions

Since the main topic of this thesis is the use of quadrilateral finite elements in the context of divergence-free simulation, it is important to give some details about the geometrical shape of a generic quadrilateral. Indeed, certain geometrical aspects are fundamental when error convergence of the approximated fields is discussed.

For this reason, it is possible to quantify the deviation of an arbitrary quadrilateral from a parallelogram element [55]. We can measure this deviation introducing σ_k as

$$\sigma_k := \max(|\pi - \theta_1|, |\pi - \theta_2|) \,. \tag{2.105}$$

In this case, the angles θ_1 and θ_2 are respectively the angles between the outward normals of two opposite sides of R. Therefore, we have $0 \leq \sigma_k < \pi$, where $\sigma_k = 0$ only if R is a parallelogram. Giving the bilinear isomorphism $\mathcal{F}$, we have $|\mathcal{F}|_{W^2_\infty(\hat{R})} \leq Ch_k(h_k + \sigma_k)$. This enables the definition of a family of elements which converges to an asymptotic parallelogram. In particular, the grid must be characterized in having $\sigma_k = O(h_k)$, thus the value σ_k/h_k remains bounded across all the quadrilateral elements in the domain. It is noteworthy that any polygonal shape can be meshed considering an asymptotically parallelogram, which is also a shape regular family of meshes where the mesh size h is going to zero. In fact, if we consider a mesh built with convex quadrilaterals, we can reach an asymptotically parallelogram, that is also shape-regular if we refine every quadrilateral by dividing by four and connecting the opposite edges.

We start with the discussion considering now a scalar function and then move on to the next section with a discussion on vector-valued fields. Hence, we can consider a smooth function $u : \Omega \to \mathbb{R}$, approximated by using a space family V_h of finite elements. If we define with k the order of the polynomial of the reference element and with $\boldsymbol{\nabla}_h$ the element-by-element gradient, the optimal approximation properties read as

$$\inf_{v_h \in V_h} ||u - v_h||_{L^2(\Omega)} = O(h^{k+1}) \,, \tag{2.106}$$

$$\inf_{v_h \in V_h} ||\boldsymbol{\nabla}_h(u - v_h)||_{L^2(\Omega)} = O(h^k) \,. \tag{2.107}$$

In order to satisfy the optimal approximation of the error gradient, we know from [32] that the space of polynomial functions with order less than or equal to k in each variable, i.e. $Q_k(\hat{R})$ must be in the reference element space. Therefore, the error estimates take the form of

$$\inf_{v_h \in V_h} ||u - v_h||_{L^2(\Omega)} \leq ch^{k+1}|u|_{k+1,\Omega} \,, \tag{2.108}$$

$$\inf_{v_h \in V_h} ||\boldsymbol{\nabla}_h(u - v_h)||_{L^2(\Omega)} \leq ch^k|\boldsymbol{\nabla} u|_{k,\Omega} \,. \tag{2.109}$$

Moreover, the previous estimates represent also a necessary condition, that has been proved in [48] for a bidimensional domain and extend to a three-dimensional domain in [56]. The previous analysis has been performed by using and introducing a trapezoidal mesh, such as in Figure 2.6.

In particular, Boffi et al. [48] show that for certain finite element families such as serendipity elements (for which a degree of freedom is removed from

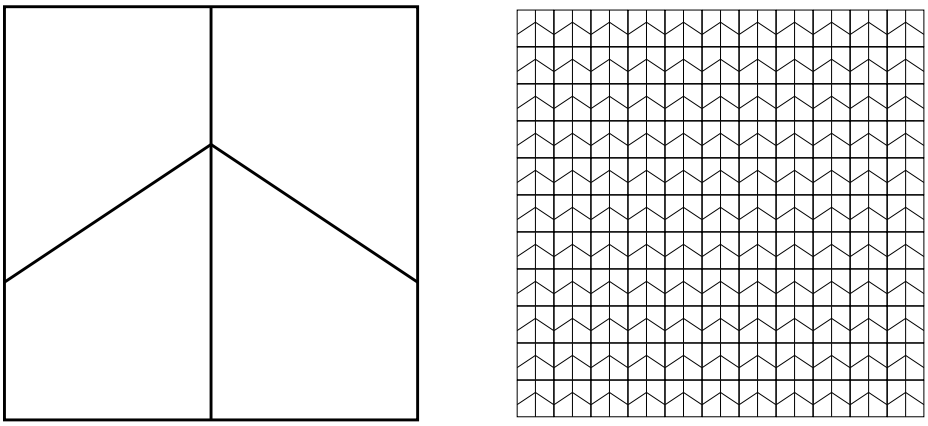

Figure 2.6: Trapezoidal mesh.

the element center) an optimal approximation order can not be achieved. They suggest that the reason lies in the mesh distortion which is maintained also for the refined grids as we can see in the trapezoidal mesh. Despite that, if we consider the same initial trapezoidal mesh, its asymptotically affine partition is still able to achieve an optimal approximation order.

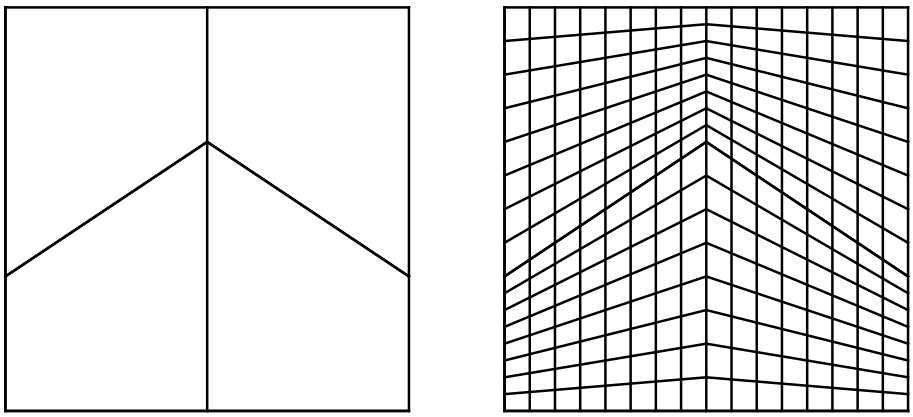

Figure 2.7: Asymptotically trapezoidal mesh.

2.2.4 Quadrilateral approximation of vector-valued functions

Necessary and sufficient conditions for optimal order approximation of vector-valued function in $\boldsymbol{H}(\mathrm{div})$ have been extensively discussed [49]. In particular,

having a smooth vector function $\boldsymbol{q} : \Omega \to \mathbb{R}^2$ and considering a finite element space $\boldsymbol{X}_h$, an optimal convergence in $L^2(\Omega)$ can be achieved if

$$\inf_{\boldsymbol{p}_h \in \boldsymbol{X}_h} ||\boldsymbol{q} - \boldsymbol{p}_h||_{L^2(\Omega)} = O(h^{k+1}) \,, \tag{2.110}$$

where k is the usual degree of the polynomial basis on the reference element. Similarly, the optimal convergence in the $\boldsymbol{H}(\mathrm{div})$ semi-norm for the space $\boldsymbol{X}_h$ is reached if

$$\inf_{\boldsymbol{p}_h \in \boldsymbol{X}_h} ||\boldsymbol{\nabla}_h \cdot (\boldsymbol{q} - \boldsymbol{p}_h)||_{L^2(\Omega)} = O(h^{k+1}) \,, \tag{2.111}$$

where $\boldsymbol{\nabla}_h$ denotes the divergence operator for every element.

Indeed, if we consider general convex quadrilateral, the accuracy of some well-known finite elements defined in $\boldsymbol{H}(\mathrm{div})$ is lower with respect to affine isomorphism (triangular or general parallelogram). In the context of a generic diffeomorphism $\mathcal{F}$ which maps the element $\hat{K}$ to the element K, we know that a linear transformation leads to a parallelogram, or in the case of a diagonal linear diffeomorphism to a rectangle. However, a generic bilinear diffeomorphism can produce an arbitrary convex quadrilateral. Following again [49], we know that, giving $\hat{\boldsymbol{V}}$ the finite-dimensional space of vector field on $\hat{K}$, which is a subset of $\boldsymbol{H}(\mathrm{div})$, e.g. the Raviart-Thomas space $\boldsymbol{\mathcal{RT}}_k$, the space $\boldsymbol{\mathcal{P}}_k(\hat{K})$ is strictly embedded in $\hat{\boldsymbol{V}}$. We recall that $\boldsymbol{\mathcal{P}}_k(\hat{K})$ is the space of the polynomial vector function on $\hat{K}$ with degree at most k. However, $\hat{\boldsymbol{V}}$ does not contain the space $\boldsymbol{\mathcal{P}}_{k+1}(\hat{K})$. An example of basis functions for the space $\boldsymbol{\mathcal{RT}}_k$, considering the $2(k+1)(k+2)$ vector fields, is given by

$$(\hat{x}_1^i \hat{x}_2^j, 0) \,, \quad (0, \hat{x}_1^j \hat{x}_2^i) \,, \quad 0 \leq i \leq r+1 \,, 0 \leq j \leq r \,. \tag{2.112}$$

We have seen at the beginning of this subsection that it is possible to estimate the error of the best approximation of a given vector (smooth) function, by using the approximating finite element function with the power of h, i.e. the mesh size or the maximum element diameter. Therefore, with the usual interpolant π_h of the space $\boldsymbol{\mathcal{RT}}_k$ we have that

$$||\boldsymbol{u} - \pi_h \boldsymbol{u}||_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{u}|_{H^{r+1}(\Omega)} \,, \tag{2.113}$$

$$||\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \pi_h \boldsymbol{u})||_{L^2(\Omega)} \leq C h^{r+1} |\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{H^{r+1}(\Omega)} \,. \tag{2.114}$$

For general shape-regular and convex quadrilateral, in [39] it has been demonstrated that we can have

$$||\boldsymbol{u} - \pi_h \boldsymbol{u}||_{L^2(\Omega)} \leq C h^{r+1} [|\boldsymbol{u}|_{H^{r+1}(\Omega)} + h|\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{H^{r+1}(\Omega)}] \,, \tag{2.115}$$

$$||\boldsymbol{\nabla} \cdot (\boldsymbol{u} - \pi_h \boldsymbol{u})||_{L^2(\Omega)} \leq C h^{r} |\boldsymbol{\nabla} \cdot \boldsymbol{u}|_{H^{r+1}(\Omega)} \,. \tag{2.116}$$

Note that, an additional regularity is required for the L^2 estimate of $\boldsymbol{u}$, despite the order in h being the same as for the parallelogram. However, we have a degradation of the order in h considering the divergence of $\boldsymbol{u}$.

In some scenarios, the standard shape functions are modified in order to approximate vector fields. For instance, a lowest-order element in a bidimensional domain is analyzed in [57], without however employing the Piola transform.

The idea presented in [49], is to extend the work discussed in [48] to vector functions, with the aim to seek the necessary conditions to have an L^2-approximation of order $r+1$ for the vector field and its divergence. From a technical perspective, the best approximation of $\boldsymbol{u} \in \boldsymbol{V}_h$ with an L^2 error can be of order $r+1$ only if the space $\hat{\boldsymbol{V}}$ contains the space $\boldsymbol{\mathcal{S}}_r$. In this case, the space $\boldsymbol{\mathcal{S}}_r$ is defined as the subspace of codimension one of $\boldsymbol{\mathcal{RT}}_k$, described with the same vector fields in (2.112), but with the single vector field $(\hat{x}_1^{r+1}\hat{x}_2^r, -\hat{x}_1^r\hat{x}_2^{r+1})$ substituting the two vector fields $(\hat{x}_1^{r+1}\hat{x}_2^r, 0)$ and $(0, \hat{x}_1^r\hat{x}_2^{r+1})$. In particular, each time that $\boldsymbol{\mathcal{S}}_r \nsubseteq \hat{\boldsymbol{V}}$, a smooth vector field $\boldsymbol{u}$ defined in the domain Ω satisfies

$$\inf_{\boldsymbol{v} \in \boldsymbol{V}_h} ||\boldsymbol{u} - \boldsymbol{v}||_{L^2(\Omega)} \neq o(h^r) \,. \tag{2.117}$$

Concerning the divergence of the field $\boldsymbol{u}$, in order to achieve an approximation of order $r+1$, we define also the space $\mathcal{R}_r$, which is the subspace of codimension one of $\mathcal{Q}_{r+1}$, i.e. the polynomial functions space of order $\leq r+1$ in each variable, with the same monomials of $\mathcal{Q}_{r+1}$ except $\hat{x}_1^{r+1}\hat{x}_2^{r+1}$. In this case, for an $r+1$ order of approximation $\boldsymbol{\nabla} \cdot \hat{\boldsymbol{V}} \supseteq \mathcal{R}_r$ must hold.

Consider now a simple square with the edges parallel to the reference axes, with the bilinear transformation

$$\mathcal{F}_K(\hat{\boldsymbol{x}}) = \boldsymbol{x}_K + h_K \hat{\boldsymbol{x}} \,, \tag{2.118}$$

where the point $\boldsymbol{x}_K \in \mathbb{R}^2$ represents the lower left corner of K and the parameter h_K is the side length. By using the Piola transformation $\boldsymbol{P}_{F_K}$ on

the approximating value function on the reference element, i.e. $\hat{\boldsymbol{u}} \in \boldsymbol{L}^2(\hat{K})$, we obtain

$$(\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-1}\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})\,, \tag{2.119}$$

with naturally $\boldsymbol{x} = \mathcal{F}_K\hat{\boldsymbol{x}}$. Moreover, we have

$$\boldsymbol{\nabla}\cdot(\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-2}\hat{\boldsymbol{\nabla}}\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})\,, \tag{2.120}$$

with the relation $||\boldsymbol{P}_{F_K}\hat{\boldsymbol{u}}||_{L^2(K)} = ||\hat{\boldsymbol{u}}||_{L^2(\hat{K})}$.

Consider now a domain partition of Ω in $\mathcal{T}_h$, with a uniform mesh of n^2 subsquares with length side equal to $h = 1/n$, and n a positive integer. We define now the space

$$\boldsymbol{V}_h := \{\boldsymbol{u} : \Omega \to \mathbb{R}^2 |\, \boldsymbol{u}|_K \in \boldsymbol{P}_{F_K}\hat{\boldsymbol{V}} \quad \forall K \in \mathcal{T}_h\}\,, \tag{2.121}$$

where $\hat{\boldsymbol{V}}$ is a subspace of $\boldsymbol{L}^2(\hat{K})$. With the previous definition, we are not imposing any interelement continuity, in the sense that $\boldsymbol{u}$ is defined on the element $\hat{K}$ in agreement with functions in $\boldsymbol{P}_{F_K}\hat{\boldsymbol{V}}$ almost everywhere. We have the following theorems

Theorem 2.5. *Consider the finite-dimensional space $\hat{\boldsymbol{V}}$, subset of $\boldsymbol{L}^2(\hat{K})$, and a non-negative integer r. The following conditions are equivalent:*

(i) For all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$, there exists a constant C such that

$$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h} ||\boldsymbol{u}-\boldsymbol{v}||_{L^2(\Omega)} \le Ch^{r+1}|\boldsymbol{u}|_{H^{r+1}(\Omega)}\,. \tag{2.122}$$

(ii) $\inf_{\boldsymbol{v}\in\boldsymbol{V}_h} ||\boldsymbol{u}-\boldsymbol{v}||_{L^2(\Omega)} = o(h^r)$ *for all* $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_r(\Omega)$.

(iii) $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_r(\hat{K})$.

Theorem 2.6. *Consider the finite-dimensional space $\hat{\boldsymbol{V}}$, subset of $\boldsymbol{L}^2(\hat{K})$, and a non-negative integer r. The following conditions are equivalent:*

(i) For all $\boldsymbol{u} \in \boldsymbol{H}^{r+1}(\Omega)$ with $\boldsymbol{\nabla}\cdot\boldsymbol{u} \in H^{r+1}(\Omega)$, there exists a constant C such that

$$\inf_{\boldsymbol{v}\in\boldsymbol{V}_h} ||\boldsymbol{\nabla}\cdot\boldsymbol{u}-\boldsymbol{\nabla}\cdot\boldsymbol{v}||_{L^2(\Omega)} \le Ch^{r+1}|\boldsymbol{\nabla}\cdot\boldsymbol{u}|_{H^{r+1}(\Omega)}\,. \tag{2.123}$$

(ii) $\inf_{\boldsymbol{v}\in\boldsymbol{V}_h} ||\boldsymbol{\nabla}\cdot\boldsymbol{u}-\boldsymbol{\nabla}\cdot\boldsymbol{v}||_{L^2(\Omega)} = o(h^r)$ *for all* $\boldsymbol{u}$ *with* $\boldsymbol{\nabla}\cdot\boldsymbol{u} \in \mathcal{P}_r(\Omega)$.

(iii) $\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{V}} \supseteq \mathcal{P}_r(\hat{K})$.

Note that, without imposing the interelement continuity inside the definition of $\boldsymbol{V}_h$, the divergence must be considered applied only elementwise to $\boldsymbol{v} \in \boldsymbol{V}_h$.

Arnold et al. [49] have shown that for more general quadrilateral mesh in a partitioned domain, stronger conditions on $\hat{\boldsymbol{V}}$ are required. In particular, they consider a specific bilinear mapping between the unit square $\hat{K}$ and the generic quadrilateral K^α, defined by the four vertices $(0,0)$, $(0,1)$, $(1,\alpha+1)$ and $(0,\alpha)$. With the mesh $\mathcal{T}_h$ composed by subsquares depending on α, as previously described, they have extended approximations properties of finite element subspace in $\boldsymbol{H}(\mathrm{div})$ from rectangular elements to generic quadrilaterals. Indeed, since by definition, we have that $\boldsymbol{\mathcal{S}}_r \subseteq \boldsymbol{\mathcal{RT}}_k$, the Raviart-Thomas space of order r is equipped with an L^2 approximation of order $r+1$ also for generic quadrilateral meshes. Nevertheless, if we apply the divergence operator, we obtain that $\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{RT}}_k = \mathcal{Q}_r$ does not contain $\mathcal{R}_r$ even though contains $\mathcal{R}_{r-1}$. For this reason, considering the Raviart-Thomas space of order r, the best approximation order for the divergence operator in L^2 is only r, regarding generic quadrilateral mesh. This value is one degree lower with respect to the more regular rectangular meshes. Thus, the space $\boldsymbol{\mathcal{RT}}_0$ lacks convergence for the divergence operator, and therefore the best approximation in the space $\boldsymbol{H}(\mathrm{div},\Omega)$ for $\hat{\boldsymbol{V}} = \boldsymbol{\mathcal{RT}}_0$ is not convergent for general quadrilateral meshes in $\boldsymbol{H}(\mathrm{div},\Omega)$.

2.2.5 Bochev technique

In this subsection, we recall the technique presented in [58] by Bochev et al. in order to restore the use of lowest order Raviart-Thomas finite element for the approximation of vector-valued functions on (shape regular) non-affine quadrilateral elements. Bochev et al. based their new formulation considering the mimetic finite difference method [59]. The idea is to introduce different operators in order to replace the canonical divergence and gradient. In particular, following [60, 61], we can introduce the natural mimetic divergence DIV and the derived gradient **GRAD**. In particular, the new divergence operator is defined as

$$\boldsymbol{\nabla} \cdot \boldsymbol{u}(\boldsymbol{x}) := \lim_{\boldsymbol{x} \in k; \mu(k) \to 0} \frac{\int_{\partial k} \boldsymbol{u} \cdot \boldsymbol{n}\, dS}{\mu(k)}, \tag{2.124}$$

where k represents a bounded region, $\mu(k)$ its measure and ∂k its boundary. With this definition, we mean a discrete operator that maps the fluxes

of $\boldsymbol{u}$, which are face-based values, onto cell-based constant values. Bochev suggests that the problem with the standard divergence operator depends on the information collected on the fluxes, and how we compute them. Indeed, for non-affine grids the operator $\boldsymbol{\nabla}\cdot$ acts on the flux data only in an indirect way, because of the Piola transformation on the basis functions. Therefore, for mixed methods, the result depends on the element shape. On the other hand, the DIV operator is able to act directly on the flux degrees of freedom, allowing the independence of the element shape.

Considering the bilinear function $\mathcal{F}_K(\hat{\boldsymbol{x}})$ that maps the reference element $\hat{K}$ to the generic quadrilateral element K, and its derivative (the Jacobian matrix) $\mathcal{DF}_K(\hat{\boldsymbol{x}})$, we define the measure μ as

$$\mu(k) = \det \mathcal{DF}(0,0)\mu(\hat{K})\,, \tag{2.125}$$

where naturally the point of coordinates (0,0) represents the center of the reference element.

Consider now the standard mixed elliptic problem, involving for instance velocity and pressure approximation, such as Stokes or Darcy equation. Moreover, consider the use of the lowest-order Raviart-Thomas space, $\boldsymbol{\mathcal{RT}}_0$, for the velocity approximation and the piecewise constant space, Q_0, for the pressure approximation. Given the standard partition domain $\mathcal{T}_h$ considering convex quadrilateral elements K, we also define the set of oriented faces $\boldsymbol{f}_i \in \mathcal{T}_h$ as F_h. Let now F_h^* and $\mathcal{T}_h^*$ the duals of F_h and $\mathcal{T}_h$, namely the collection of real numbers $\{F_f\}$, $\{K_K\}$ that are associated to the number of faces and cells of the mesh. Since F_h^* and $\mathcal{T}_h^*$ are isomorphic to $\boldsymbol{\mathcal{RT}}_0$ and Q_0, we can use the same symbols for their elements. It is possible now to define the natural divergence as the mapping DIV: $F_h^* \to \mathcal{T}_h^*$

$$\mathrm{DIV}(\boldsymbol{u}_h)|_K := \frac{1}{\mu(K)} \sum_{\boldsymbol{f}\in F_h(K)} \sigma_f F_f\,, \quad \forall K \in \mathcal{T}_h\,, \tag{2.126}$$

with $\boldsymbol{u}_h \in F_h^*$. Moreover, $F_h(K)$ is the set of the oriented faces on the element K where

$$\sigma_f = \begin{cases} 1 & \text{if } \boldsymbol{n}_f = \boldsymbol{n}_K \\ -1 & \text{if } \boldsymbol{n}_f = -\boldsymbol{n}_K\,. \end{cases} \tag{2.127}$$

We can understand that the functions F_f are indeed the degrees of freedom defining a vector field function in the space $\boldsymbol{\mathcal{RT}}_0$, i.e.

$$\boldsymbol{u}_h = \sum_{f\in F_h} F_f \boldsymbol{u}_f\,, \quad \forall \boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0\,. \tag{2.128}$$

With the newly introduced operator, the discrete divergence $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0$ can be replaced by using the formula in (2.126). Therefore, we reformulate the mixed problem using the operator DIV:$\boldsymbol{\mathcal{RT}}_0 \rightarrow Q_0$. In particular for the basis functions $\boldsymbol{u}_f \in \boldsymbol{\mathcal{RT}}_0$ we have

$$\mathrm{DIV}(\boldsymbol{u}_f) = \frac{\sigma_f}{\mu(K)}\,, \quad \forall f \in F_h\,. \tag{2.129}$$

Introducing the $\boldsymbol{\mathcal{RT}}_0$ interpolator $\mathcal{I}_{\boldsymbol{\mathcal{RT}}_0}$, and the interpolator $\mathcal{I}_{Q_0}$ for the space Q_0, we report the following lemmas:

Lemma 2.14. *The commuting diagram property can be applied pointwise to the natural divergence DIV*

$$\begin{array}{ccc} W(\Omega) & \xrightarrow{\boldsymbol{\nabla}\cdot} & L^2(\Omega) \\ \mathcal{I}_{\boldsymbol{\mathcal{RT}}_0}\downarrow & & \downarrow\mathcal{I}_{Q_0} \\ \boldsymbol{\mathcal{RT}}_k & \xrightarrow{DIV} & Q_0\,. \end{array} \tag{2.130}$$

Lemma 2.15. *The operator DIV is a surjection between $\boldsymbol{\mathcal{RT}}_0$ onto Q_0, equipped with a continuous lifting from Q_0 into $\boldsymbol{\mathcal{RT}}_0$. Thus, for every $q_h \in Q_0$ there exists a function $\boldsymbol{u}_{h,q} \in \boldsymbol{\mathcal{RT}}_0$ that satisfies*

$$q_h = DIV(\boldsymbol{u}_{h,q})\,, \tag{2.131}$$

$$||\boldsymbol{u}_{h,q}||_{L^2(\Omega)} + ||DIV(\boldsymbol{u}_{h,q})||_{L^2(\Omega)} \leq C||q_h||_{L^2(\Omega)}\,. \tag{2.132}$$

Note that for standard non-affine quadrilateral elements, we usually have that $\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{RT}}_0 \neq Q_0$ and therefore the surjective property between the domain and the divergence range, i.e. Q_0, is not satisfied. However, the property can be restored by using the DIV operator instead of the analytic divergence. In this sense, Bochev asserts that "*DIV provides a better approximation of* $\boldsymbol{\nabla}\cdot$ *on* $\boldsymbol{\mathcal{RT}}_0$ *than the usual finite element practice of restricting the analytic operator to the finite element space*". As we said before, the surjective property equipped by the DIV operator depends on the fact the DIV acts only on the flux data and is not influenced by the element shape, unlike the $\boldsymbol{\nabla}\cdot$ operator which acts considering the Piola transform via basis functions. Another important lemma is reported now to show the equivalence with the standard formulation of the divergence operator

Lemma 2.16. *Let $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0$. Thus, we have*

$$\int_K \boldsymbol{\nabla}\cdot\boldsymbol{u}_h\, d\boldsymbol{x} = \int_K DIV(\boldsymbol{u}_h)\, d\boldsymbol{x}\,, \quad \textit{with } K \in \mathcal{T}_h\,. \tag{2.133}$$

Moreover, we also have that

Lemma 2.17. *Consider a shape-regular partition $\mathcal{T}_h$ of the domain Ω. Thus, there exists a positive constant C_D such that*

$$||\boldsymbol{\nabla} \cdot \boldsymbol{u}_h||_{L^2(\Omega)} \leq C_D ||DIV(\boldsymbol{u}_h)||_{L^2(\Omega)} \,. \tag{2.134}$$

Note that considering the reference quadrilateral element $\hat{K}$, the standard basis functions are

$$\hat{\boldsymbol{u}}_f = \frac{1}{4}\begin{bmatrix} 1 \pm x \\ 0 \end{bmatrix} \quad \text{or} \quad \hat{\boldsymbol{u}}_f = \frac{1}{4}\begin{bmatrix} 0 \\ 1 \pm y \end{bmatrix} , \tag{2.135}$$

which leads to

$$\boldsymbol{\nabla}_{\hat{\boldsymbol{x}}} \cdot \hat{\boldsymbol{u}}_f = \frac{1}{4} \,. \tag{2.136}$$

Bochev also suggests how to implement this new formulation in the framework of a finite element code. In particular, considering $\mu(\hat{K}) = 4$, we have that

$$\mathrm{DIV}(\boldsymbol{u}_f)|_K = \frac{\sigma_f}{\mu(K)} = \frac{\sigma_f}{4 \det(\mathcal{DF}(0,0))} \,, \tag{2.137}$$

where naturally the case of a bidimensional domain has been considered. Following this suggestion, we can avoid the computation of the standard $\boldsymbol{\nabla}\cdot$ operator for every quadrature point and the computation of $\det \mathcal{DF}$ at the same points. Instead, a single call to a function that computes $\det(\mathcal{DF}(0,0))$ is necessary, paying attention to the value of σ_f which depends on the face orientation.

Another way to reach a first-order error approximation for the divergence operator has been introduced in [57] and is presented in the next section. In particular, the idea proposed is a method to restore a constant divergence in any element $K \in \mathcal{T}_h$ imposing a correction on the $\boldsymbol{\mathcal{RT}}_0$ basis function on $\hat{K}$.

The previous mimetic reformulation of the mixed problem leads to the same linear system of equations considering the unknown coefficients for the velocity $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0$ and for the pressure $p_h \in Q_0$. Therefore, as demonstrated in [58] the solution of the new method is the same as the standard mixed method. Moreover, rigorous stability proof has been proved for the mixed problem equipped with the DIV operator. Indeed, we have

Lemma 2.18. *Consider the space $\boldsymbol{Z}_h := \{\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0 | DIV(\boldsymbol{u}_h) = 0\}$, i.e. the null-space of DIV. Thus, we have that the classical form $a_h(\cdot,\cdot)$ is coercive on $\boldsymbol{Z}_h \times \boldsymbol{Z}_h$ with*

$$C_a ||\boldsymbol{v}_h||^2 \leq a_h(\boldsymbol{v}_h, \boldsymbol{v}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{Z}_h \,. \tag{2.138}$$

In addition, we have that the inf-sup condition is satisfied by the form $b_h(\cdot,\cdot)$ *with*

$$\sup_{\boldsymbol{v}_h \in \boldsymbol{\mathcal{RT}}_0} \frac{b_h(\boldsymbol{v}_h, p_h)}{||\boldsymbol{v}_h||_{DIV}} \geq C_b ||p_h||_{L^2(\Omega)} \quad \forall p_h \in Q_0 \,. \tag{2.139}$$

Finally, in [58] an optimal error estimate has been proved with the following theorem

Theorem 2.7. *Consider a shape regular partition* $\mathcal{T}_h$*, for which holds*

$$\frac{1}{\alpha}\mu(K) \leq ||\det \mathcal{DF}_K||_{\infty,\hat{K}} \leq \alpha\mu(K) \quad \forall K \in \mathcal{T}_h \,, \tag{2.140}$$

where α *is a positive number. Consider also, the pair solution of the mixed problem* (1.39) *as* $(\boldsymbol{u}, p) \in \boldsymbol{H}_N(\mathrm{div}, \Omega) \cap \boldsymbol{H}^2(\Omega) \times H^1_D(\Omega)$*, where* $H^1_D(\Omega)$ *is the subspace of* $H^1(\Omega)$ *where the functions vanish on* Γ_D *and* $\boldsymbol{H}_N(\mathrm{div}, \Omega)$ *is the subspace of* $\boldsymbol{H}(\mathrm{div}, \Omega)$ *equipped with* $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ *on* Γ_N*. Thus, considering the pair solution of the discretized problem* $(\boldsymbol{u}_h, p_h)$*, the following error bound holds*

$$\begin{aligned} ||\boldsymbol{\nabla} \cdot \boldsymbol{u} - DIV(\boldsymbol{u}_h)||_{L^2(\Omega)} + ||\boldsymbol{u} - \boldsymbol{u}_h||_{L^2(\Omega)} + ||p - p_h||_{L^2(\Omega)} \leq \\ \leq Ch \left(||\boldsymbol{u}||_{H^2(\Omega)} + ||p||_{H^1(\Omega)} \right) . \end{aligned} \tag{2.141}$$

The theoretical conclusions have also been demonstrated with numerical results, for which the first-order accuracy of the divergence velocity approximation is restored by using the natural mimetic divergence operator, naturally considering the case of non-affine grids. For the sake of completeness, in the paper described above other types of finite element families have been investigated. Specifically, for least square finite elements, the presence of a non-affine mesh can produce a degradation of the order error approximation. On the other hand, as we have already discussed, mixed finite elements such as $\boldsymbol{\mathcal{RT}}_0$ can benefit from the introduction of the new operators, especially when dealing with not asymptotically affine grids.

2.2.6 Shen-Boffi technique

We present now the solution proposed in [57] for the resolution of the Darcy equation and in [62] for the eigenvalue problem of the same equation, in the context of distorted rectangular grids. The idea is to introduce and describe a pair solution of the mixed elliptic problem, that is related to a discrete

velocity space $\boldsymbol{V}_h$ and a pressure space W_h satisfying the relation

$$\begin{aligned} \boldsymbol{V}_h \subset \boldsymbol{V} \subset \boldsymbol{H}(\mathrm{div}, \Omega)\,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{V}_h \equiv W_h \subset W \subset L^2(\Omega)\,. \end{aligned} \tag{2.142}$$

Note that the suggested technique conserves the same number of unknowns as the implementation of the lowest-order Raviart-Thomas. Moreover, if the generic quadrilateral elements degenerate back to parallelogram elements, the proposed mixed method matches with the lowest-order Raviart-Thomas finite element applied to orthogonal rectangular grids. Finally, this technique avoids any problematic issues related to the existence and uniqueness of the solution since it satisfies the inf-sup condition.

Consider now the characteristic function of the domain K, i.e. $\chi(K)$. Therefore, the space of the discretized pressure can be defined as $W_h = \mathrm{span}\{\chi(K)\}_{K \in \mathcal{T}_h}$. In other words, the space W_h is the space of the piecewise constant function, and therefore we have that $W_h \subset L^2(\Omega)$.

Concerning the velocity space, we have to add some considerations. First of all, we want to preserve the continuity of the normal component on the element edges. Therefore, considering the nodal bases $\boldsymbol{b}_i \in \boldsymbol{V}_h$, $i = 1, 2, 3, 4$ and their restriction to the element K we want that

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{b}_i &= \text{constant}\,, \\ \boldsymbol{b}_i \cdot \boldsymbol{n}_j|_{e_j} &= \delta_{ij}\,, \end{aligned} \tag{2.143}$$

where e_j represents the element edges. Recalling the properties of the Piola transform, the previous relations are equivalent to

$$\begin{aligned} \hat{\boldsymbol{\nabla}} \cdot \hat{\boldsymbol{b}}_i &= \mathcal{J} \cdot \text{constant}\,, \\ \hat{\boldsymbol{b}}_i \cdot \hat{\boldsymbol{n}}_j|_{\hat{e}_j} &= |e_i| \delta_{ij}\,, \end{aligned} \tag{2.144}$$

where $\mathcal{J}$, the well-known determinant of the Jacobian matrix, is a linear function of $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, meanwhile $|e_i|$ represents the length of the edge i. Indeed, we can consider the bilinear transformation between the real and the reference element, i.e. $\mathcal{F} : \hat{K} \to K$, defined as

$$\boldsymbol{x} = \mathcal{F}(\hat{\boldsymbol{x}}) = \boldsymbol{x}_1 + \boldsymbol{x}_{21}\hat{x} + \boldsymbol{x}_{41}\hat{y} + (\boldsymbol{x}_{32} - \boldsymbol{x}_{41})\hat{x}\hat{y}\,, \tag{2.145}$$

where $\boldsymbol{x}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j$ represent the four vertices of the element K (or respectively $\hat{K}$). Naturally, we have that

$$\boldsymbol{x}_i = \mathcal{F}(\hat{\boldsymbol{x}}_i)\,, \quad i = 1, 2, 3, 4\,. \tag{2.146}$$

Recalling that the Jacobian matrix and its determinant take the following form

$$\mathcal{DF} = \begin{pmatrix} x_{\hat{x}} & x_{\hat{y}} \\ y_{\hat{x}} & y_{\hat{y}} \end{pmatrix}, \quad \mathcal{J} = |\det \mathcal{DF}| = x_{\hat{x}} y_{\hat{y}} - x_{\hat{y}} y_{\hat{x}} > 0\,, \tag{2.147}$$

with

$$\mathcal{DF}^{-1} = \begin{pmatrix} \hat{x}_x & \hat{x}_y \\ \hat{y}_x & \hat{y}_x \end{pmatrix} = \frac{1}{\mathcal{J}} \begin{pmatrix} y_{\hat{y}} & -x_{\hat{y}} \\ -y_{\hat{x}} & x_{\hat{x}} \end{pmatrix}, \tag{2.148}$$

the determinant can be written also as a linear function of $\hat{x}$ and $\hat{y}$

$$\mathcal{J} = \alpha + \beta\hat{x} + \gamma\hat{y}\,. \tag{2.149}$$

Moreover, it is easy to compute also the area of the element K as

$$|K| = \int_K dK = \int_{\hat{K}} \mathcal{J} d\hat{K} = \int_0^1 \int_0^1 \mathcal{J} d\hat{x} d\hat{y} = \alpha + \frac{\beta}{2} + \frac{\gamma}{2}\,. \tag{2.150}$$

With this information, we report a lemma proved by Shen in [57].

Lemma 2.19. *There exists a function $\hat{\boldsymbol{b}}_0 \in \mathbb{R}^2$, defined on the reference element $\hat{K}$ that satisfies*

$$\hat{\boldsymbol{\nabla}} \cdot \hat{\boldsymbol{b}}_0 = \frac{\mathcal{J}}{K} - 1\,, \tag{2.151}$$

$$\hat{\boldsymbol{b}}_0 \cdot \hat{\boldsymbol{n}}_j|_{\hat{j}} = 0\,, \quad j = 1, 2, 3, 4\,, \tag{2.152}$$

and then,

$$\int_{\hat{K}} \hat{\boldsymbol{\nabla}} \cdot \hat{\boldsymbol{b}}_0 d\hat{K} = 0\,. \tag{2.153}$$

Note that, since the vector function $\hat{\boldsymbol{b}}_0$ is defined as

$$\hat{\boldsymbol{b}}_0 = \begin{pmatrix} \zeta \\ \eta \end{pmatrix}, \tag{2.154}$$

where the two components are described as follows

$$\zeta = \frac{\beta\hat{x}(\hat{x} - 1)}{2|K|}, \quad \eta = \frac{\gamma\hat{y}(\hat{y} - 1)}{2|K|}\,, \tag{2.155}$$

this new term is described with a quadratic function in the reference coordinates and depends only on $\mathcal{J}$.

Now, we can introduce the new basis functions defined on $\hat{K}$

$$\begin{aligned}\hat{\boldsymbol{b}}_1 &= |e_1|\begin{pmatrix}1-\hat{x}\\0\end{pmatrix} - |e_1|\hat{\boldsymbol{b}}_0, \quad \hat{\boldsymbol{b}}_3 = |e_3|\begin{pmatrix}\hat{x}\\0\end{pmatrix} - |e_3|\hat{\boldsymbol{b}}_0,\\ \hat{\boldsymbol{b}}_2 &= |e_2|\begin{pmatrix}0\\1-\hat{y}\end{pmatrix} - |e_2|\hat{\boldsymbol{b}}_0, \quad \hat{\boldsymbol{b}}_4 = |e_4|\begin{pmatrix}0\\\hat{y}\end{pmatrix} - |e_4|\hat{\boldsymbol{b}}_0\,,\end{aligned} \tag{2.156}$$

and consider that

$$\boldsymbol{b}_i = \frac{1}{\mathcal{J}}\mathcal{D}\mathcal{F}\,\hat{\boldsymbol{b}}_{\boldsymbol{i}}\,, \quad i = 1,2,3,4\,. \tag{2.157}$$

It can be easily seen that every relation in (2.143) and (2.144) is satisfied, because of the choice of the new basis functions defined in (2.156). Hence, for any $\boldsymbol{u} \in \boldsymbol{V}_h$, the corresponding restriction on the real element K is determined by

$$\begin{aligned}\boldsymbol{u}|_K &= \sum_{i=1}^{4} u_i \boldsymbol{b}_i\,,\\ u_i &= \boldsymbol{u}\cdot\boldsymbol{n}_i|_{e_i}\,, \quad i = 1,2,3,4\,.\end{aligned} \tag{2.158}$$

Note now that $\boldsymbol{\nabla}\cdot\boldsymbol{V}_h \subset W_h$, and since the operator $\boldsymbol{\nabla}\cdot : \boldsymbol{V}_h \to W_h$ is surjective, we have that $W_h \subset \boldsymbol{\nabla}\cdot\boldsymbol{V}_h$, that leads to $W_h \equiv \boldsymbol{\nabla}\cdot\boldsymbol{V}_h$. Hence, the new numerical space $\boldsymbol{V}_h \times W_h$ satisfies the relation in (2.142).

This result has been achieved through the addition of the $\hat{\boldsymbol{b}}_0$ function, which is equipped with quadratic terms in the $\hat{x}$ and $\hat{y}$ components. In fact, despite the boundary fluxes on the edges are the same without the function $\hat{\boldsymbol{b}}_{\boldsymbol{0}}$, unfortunately, the divergence of the discrete velocity space would not be included anymore in the space of the discretized pressure, namely the piecewise constant space.

Boffi and Gastaldi gave their interpretation in [62] about the function $\hat{\boldsymbol{b}}_0$, through which they denote this new element as Modified Raviart Thomas ($\boldsymbol{\mathcal{MRT}}$). First, this function vanishes if the element reduces to a parallelogram (therefore if we consider an affine mapping). Additionally, its normal component is always equal to zero on four sides of the reference element, and for this reason, the function can be interpreted as a bubble in $\boldsymbol{H}(\mathrm{div}, \hat{K})$. As already stated, the four reference shape functions in the case of affine mapping are equal to the standard lowest-order Raviart-Thomas basis. Therefore, thanks to this new term which is added to every basis function we are able to obtain always a constant divergence of the velocity inside each element, properties that in general are not true for meshes with a generic bilinear mapping if we consider standard $\boldsymbol{\mathcal{RT}}_0$ elements.

Furthermore, they reported an example in which, with this kind of basis function, it is possible to represent a constant vector inside the element, even though the Piola transformation acts on non-affine elements. This is a key point in order to reach an optimal order approximation of the velocity divergence. In fact, it is sufficient to prove that, by using a suitable combination of the shape functions $\hat{\boldsymbol{b}}_i$, the reference vector is the Piola transform of the vector function in the real element (for instance $\boldsymbol{u} = (1,0)^T$). If we consider the reported bidimensional case, we have the bilinear mapping between K and $\hat{K}$ defined by

$$\mathcal{F}(\hat{x},\hat{y}) = \begin{cases} a_1 + a_2\hat{x} + a_3\hat{y} + a_4\hat{x}\hat{y} \\ b_1 + b_2\hat{x} + b_3\hat{y} + b_4\hat{x}\hat{y} \end{cases}, \tag{2.159}$$

and consequently, the Jacobian matrix reads as

$$\mathcal{DF}(\hat{x},\hat{y}) = \begin{pmatrix} a_2 + a_4\hat{y} & a_3 + a_4\hat{x} \\ b_2 + b_4\hat{y} & b_3 + b_4\hat{x} \end{pmatrix}. \tag{2.160}$$

By inverting now the Piola transform for the vector $\boldsymbol{u}$ we have

$$\hat{\boldsymbol{u}} = \mathcal{J}\mathcal{D}\mathcal{F}^{-1}\boldsymbol{u} = \begin{pmatrix} b_3 \\ -b_2 \end{pmatrix} + b_4 \begin{pmatrix} \hat{x} \\ -\hat{y} \end{pmatrix}. \tag{2.161}$$

Finally, the expression of the vector function in the reference element is the following

$$\begin{aligned} \hat{\boldsymbol{u}} =& \frac{b_3}{|e_1||e_3|}(|e_1|\hat{\boldsymbol{b}}_3 - |e_3|\hat{\boldsymbol{b}}_1) - \frac{b_2}{|e_2||e_4|}(|e_2|\hat{\boldsymbol{b}}_4 - |e_4|\hat{\boldsymbol{b}}_2) + \\ &+ \frac{b_4}{|e_3||e_4|}(|e_4|\hat{\boldsymbol{b}}_3 - |e_3|\hat{\boldsymbol{b}}_4)\,. \end{aligned} \tag{2.162}$$

Note the difference between the coefficients of the bilinear transformation b_i and the shape functions $\hat{\boldsymbol{b}}_i$. In the same fashion, we can find the $\hat{\boldsymbol{u}}$ corresponding to $\boldsymbol{u} = (0,1)^T$.

Without providing a detailed description, we mention that in [63] an extension of the previous idea is presented. In particular, the author suggests that the approach of Shen [57] relies on the assumption that the velocity field $\boldsymbol{u}$ is at most linear, and therefore the employment of a Taylor series expansion can be justified. For this reason, in [63], the case in which $\boldsymbol{u}$ is not linear is considered, achieving an optimal error estimate for the new proposed element maintaining the same degree of freedom of standard $\boldsymbol{\mathcal{RT}}$ elements.

2.2.7 Hexahedral elements

In the last year, many authors have tried to address the problem of considering general hexahedral grids in the context of mixed finite elements. In particular, lowest-order Raviart-Thomas finite elements seem to lack convergence for general hexahedral [64, 65, 66]. In fact, as it is described later, different authors [67, 68, 69] suggest that on general hexahedrons constant vector fields cannot be represented by $\boldsymbol{\mathcal{RT}}_0$ functions. On the other hand, considering hexahedrons with flat faces or grids with elements tending asymptotically to parallelepiped elements, $\boldsymbol{H}(\mathrm{div})$ convergence can still be reached [70]. For example, in [69] a three-dimensional general hexahedron with flat surfaces has been considered. They implement a composite method, by which each hexahedron is subdivided into five tetrahedral elements, avoiding any convergence issue. Furthermore, Brezzi et al. [65] proposed a method based on the mimetic finite difference by which even hexahedral with non-planar faces convergence can be obtained.

In this subsection, the extension of the previous results is described in the framework of a three-dimensional problem. In particular, if the major part of the discussion presented in the previous sections was related to quadrilateral elements, now we focus on generic hexahedral elements, to discuss the approximation properties of finite element subspaces in $\boldsymbol{H}(\mathrm{div}, \Omega)$, with $\Omega \in \mathbb{R}^3$. An important analysis of this situation has been presented in [71], from which is taken the following description.

We recall shortly the main results concerning the approximation theory of vector fields in a three-dimensional case, but considering a uniform mesh, in order to explain better than the difference with a generic hexahedral mesh. Thus, consider a cube K, that has by definition the edges parallel to the reference axes. In this case, the generic trilinear diffeomorphism which maps $K = \boldsymbol{D}_K(\hat{K})$ is linear and defined as

$$\boldsymbol{D}_K(\hat{\boldsymbol{x}}) = \boldsymbol{x}_K + h_K \hat{\boldsymbol{x}}\,, \tag{2.163}$$

where the corner of smallest coordinates (x, y, z) of the real element K is represented by $\boldsymbol{x}_K \in \mathbb{R}^3$ with the side length $h_K > 0$. In this case, the Piola transform $\boldsymbol{P}_{D_K}$ applied to the vector function $\hat{\boldsymbol{u}} \in \boldsymbol{L}^2(\hat{K})$ reads as

$$(\boldsymbol{P}_{D_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-2}\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})\,, \tag{2.164}$$

where $\boldsymbol{x} = \boldsymbol{D}_K(\hat{\boldsymbol{x}})$. Instead, for the divergence operator we obtain

$$\boldsymbol{\nabla} \cdot (\boldsymbol{P}_{D_K}\hat{\boldsymbol{u}})(\boldsymbol{x}) = h_K^{-3}\hat{\boldsymbol{\nabla}} \cdot \hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})\,. \tag{2.165}$$

We can define now the domain partition $\mathcal{T}_h$ considering n^3 subcubes, with n a positive integer, where the side length is equal to $h = 1/n$. Now, defining the space $\boldsymbol{V}(K) = \boldsymbol{P}_{D_K}\hat{\boldsymbol{V}}_K$, where $\hat{\boldsymbol{V}}_K$ is a subspace of $\boldsymbol{L}^2(\hat{K})$, we introduce the space

$$\tilde{\boldsymbol{V}}_h := \left\{ \boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}_K \in \boldsymbol{V}(K), \quad \forall K \in \mathcal{T}_h \right\}. \tag{2.166}$$

The following theorems are in analogy with Section 2.2.4 regarding the bidimensional case.

Theorem 2.8. *Consider the finite dimensional subspace $\hat{\boldsymbol{V}}$ of $\boldsymbol{L}^2(\hat{K})$. The following statements are equivalent:*

(i) For all $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, there exists a constant C such that

$$\inf_{\boldsymbol{v}\in\tilde{\boldsymbol{V}}_h} ||\boldsymbol{u} - \boldsymbol{v}||_{L^2(\Omega)} \leq Ch||\boldsymbol{\nabla u}||_{L^2(\Omega)}.$$

(ii) $\inf_{\boldsymbol{v}\in\tilde{\boldsymbol{V}}_h} ||\boldsymbol{u} - \boldsymbol{v}||_{L^2(\Omega)} = o(1)$ for all $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_0(\Omega)$.

(iii) $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_0(\hat{K})$.

In the previous theorem, the space $\boldsymbol{\mathcal{P}}_0(\cdot)$ represents the space of polynomial vector functions of order zero, i.e. a constant function.

In addition, we have the same theorem for the divergence operator.

Theorem 2.9. *Consider the finite dimensional subspace $\hat{\boldsymbol{V}}$ of $\boldsymbol{L}^2(\hat{K})$. The following statements are equivalent:*

(i) For all $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, there exists a constant C such that

$$\inf_{\boldsymbol{v}\in\tilde{\boldsymbol{V}}_h} ||\boldsymbol{\nabla}\cdot\boldsymbol{u} - \boldsymbol{\nabla}\cdot\boldsymbol{v}||_{L^2(\Omega)} \leq Ch||\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u})||_{L^2(\Omega)}.$$

(ii) $\inf_{\boldsymbol{v}\in\tilde{\boldsymbol{V}}_h} ||\boldsymbol{\nabla}\cdot\boldsymbol{u} - \boldsymbol{\nabla}\cdot\boldsymbol{v}||_{L^2(\Omega)} = o(1)$ for all $\boldsymbol{u}$ with $\boldsymbol{\nabla}\cdot\boldsymbol{u} \in \mathcal{P}_0(\Omega)$.

(iii) $\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{V}} \supseteq \mathcal{P}_0(\hat{K})$.

In this case, since the velocity divergence is considered, the space $\mathcal{P}_0(\cdot)$ represents the space of polynomial scalar functions of order zero. Because in this case the interelement continuity is not imposed in the definition of $\tilde{\boldsymbol{V}}_h$, the divergence of the velocity is to be interpreted as the elementwise divergence of $\boldsymbol{v}$.

We consider now general hexahedral elements, for which the shape regularity property holds. The idea is to find the necessary condition for having an $O(h)$ approximation of a vector field function considering a generic trilinear mapping. In particular, this mapping can be described as

$$F_1 = a_1 + b_1\hat{x} + c_1\hat{y} + d_1\hat{z} + e_1\hat{x}\hat{y} + f_1\hat{y}\hat{z} + g_1\hat{z}\hat{x} + h_1\hat{x}\hat{y}\hat{z}\,, \tag{2.167a}$$

$$F_2 = a_2 + b_2\hat{x} + c_2\hat{y} + d_2\hat{z} + e_2\hat{x}\hat{y} + f_2\hat{y}\hat{z} + g_2\hat{z}\hat{x} + h_2\hat{x}\hat{y}\hat{z}\,, \tag{2.167b}$$

$$F_3 = a_3 + b_3\hat{x} + c_3\hat{y} + d_3\hat{z} + e_3\hat{x}\hat{y} + f_3\hat{y}\hat{z} + g_3\hat{z}\hat{x} + h_3\hat{x}\hat{y}\hat{z}\,. \tag{2.167c}$$

Note that, with this generic transformation, planar faces are not necessary on the real element K. In fact, even if the element edges are always straight segments, a planar face is obtained if and only if its four vertices lie on the same plane.

In order to reach an approximation estimate of the type $o(1)$ for the velocity field $\boldsymbol{u} \in \boldsymbol{\mathcal{P}}_0(\Omega)$, we have seen that for affine subcubes of a domain is sufficient to have $\hat{\boldsymbol{V}} \supseteq \boldsymbol{\mathcal{P}}_0(\hat{K})$. Therefore, in the case of a general trilinear mapping, the objective is to verify if, after the application of the Piola transform between the unit reference cube and the arbitrary hexahedron, we still obtain on the hexahedron constant vectors, i.e. the functions $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$. In order to verify if we are able to represent these vector field functions we apply the inverse Piola transform to the function $\boldsymbol{u}$ obtaining

$$\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}}) = \boldsymbol{P}_F^{-1}\boldsymbol{u}(\boldsymbol{x}) = \mathcal{J}(\hat{\boldsymbol{x}})\mathcal{DF}^{-1}(\hat{\boldsymbol{x}})\boldsymbol{u}(\boldsymbol{x})\,. \tag{2.168}$$

Thus, the Jacobian matrix reads

$$\mathcal{DF}(\hat{\boldsymbol{x}}) = \begin{pmatrix} b_1 + e_1\hat{y} + g_1\hat{z} + h_1\hat{y}\hat{z} & c_1 + e_1\hat{y} + f_1\hat{z} + h_1\hat{y}\hat{z} & d_1 + f_1\hat{y} + g_1\hat{z} + h_1\hat{y}\hat{z} \\ b_2 + e_2\hat{y} + g_2\hat{z} + h_2\hat{y}\hat{z} & c_2 + e_2\hat{y} + f_2\hat{z} + h_2\hat{y}\hat{z} & d_2 + f_2\hat{y} + g_2\hat{z} + h_2\hat{y}\hat{z} \\ b_3 + e_3\hat{y} + g_3\hat{z} + h_3\hat{y}\hat{z} & c_3 + e_3\hat{y} + f_3\hat{z} + h_3\hat{y}\hat{z} & d_3 + f_3\hat{y} + g_3\hat{z} + h_3\hat{y}\hat{z} \end{pmatrix}. \tag{2.169}$$

Therefore, if we consider the vector $\boldsymbol{u}(\boldsymbol{x}) = (1,0,0)^T$ we obtain for the respective value on the reference element

$$\begin{aligned} \hat{u}_1(\boldsymbol{x}) &= (c_2 + e_2\hat{x} + f_2\hat{z} + h_2\hat{x}\hat{z})(d_3 + f_3\hat{y} + g_3\hat{x} + h_3\hat{x}\hat{y}) \\ &\quad - (d_2 + f_2\hat{y} + g_2\hat{x} + h_2\hat{x}\hat{y})(c_3 + e_3\hat{x} + f_3\hat{z} + h_3\hat{x}\hat{z})\,, \end{aligned} \tag{2.170a}$$

$$\begin{aligned} \hat{u}_2(\boldsymbol{x}) &= -(b_2 + e_2\hat{y} + g_2\hat{z} + h_2\hat{y}\hat{z})(d_3 + f_3\hat{y} + g_3\hat{x} + h_3\hat{x}\hat{y}) \\ &\quad + (d_2 + f_2\hat{y} + g_2\hat{x} + h_2\hat{x}\hat{y})(b_3 + e_3\hat{y} + g_3\hat{z} + h_3\hat{y}\hat{z})\,, \end{aligned} \tag{2.170b}$$

$$\begin{aligned} \hat{u}_3(\boldsymbol{x}) &= (b_2 + e_2\hat{y} + g_2\hat{z} + h_2\hat{y}\hat{z})(c_3 + e_3\hat{x} + f_3\hat{z} + h_3\hat{x}\hat{z}) \\ &\quad - (c_2 + e_2\hat{x} + f_2\hat{z} + h_2\hat{x}\hat{z})(b_3 + e_3\hat{y} + g_3\hat{z} + h_3\hat{y}\hat{z})\,. \end{aligned} \tag{2.170c}$$

Finally, we have for $\hat{\boldsymbol{u}}(\hat{\boldsymbol{x}})$

$$\hat{u}_1(\hat{\boldsymbol{x}}) = A_1 + (D_3 - C_2)\hat{x} + C_1\hat{y} + D_1\hat{z} - (E_2 + G_2)\hat{x}\hat{y} + (E_3 - G_3)\hat{x}\hat{z} + G_1\hat{x}^2 + H_3\hat{x}^2\hat{y} + H_2\hat{x}^2\hat{z}\,, \tag{2.171a}$$

$$\hat{u}_2(\hat{\boldsymbol{x}}) = A_2 + B_2\hat{x} + (B_1 - D_3)\hat{y} + D_2\hat{z} + (E_1 - G_1)\hat{y}\hat{x} - (E_3 + G_3)\hat{y}\hat{z} + G_2\hat{y}^2 - H_3\hat{x}\hat{y}^2 + H_1\hat{y}^2\hat{z}\,, \tag{2.171b}$$

$$\hat{u}_3(\hat{\boldsymbol{x}}) = A_3 + B_3\hat{x} + C_3\hat{y} + (C_2 - B_1)\hat{z} - (E_1 + G_1)\hat{z}\hat{x} + (E_2 - G_2)\hat{z}\hat{y} + G_3\hat{z}^2 + H_2\hat{x}\hat{z}^2 - H_1\hat{y}\hat{z}^2\,, \tag{2.171c}$$

where the coefficients of the capital letter with subscript i are combinations of the coefficients, lowercase letter, present in the trilinear form. Just to make an example we can define the first coefficients A_1 as the combination of the constant terms present in the first component of the reference velocity $\hat{u}_1(\boldsymbol{x})$, i.e.

$$A_1 = c_2 d_3 - d_2 c_3\,, \tag{2.172}$$

meanwhile, for the linear term in $\hat{x}$ we have for instance

$$D_3 - C_2 = (e_2 d_3 - e_3 d_2) - (g_2 c_3 - c_2 g_3)\,, \tag{2.173}$$

and so on. The other coefficients follow the same idea in order to gather the same order of local coordinates. In the same way, we can obtain the resulting vectors by choosing the constant vectors $(0,1,0)$ or $(0,0,1)$. In particular, we can think about $A_i^2, B_i^2, \ldots$, for the vector $(0,1,0)$ and the coefficients $A_i^3, B_i^3, \ldots$, for the vector $(0,0,1)$. The linear space that we have obtained is equipped with 20 independent parameters, and we can define it as $\hat{\boldsymbol{S}}_0^-$. Note that even though there are 21 parameters, if we sum the coefficients B_1, C_2, and D_3, we obtain zero. Therefore, having only two linearly independent coefficients, we have 20 coefficients instead of 21. It is easy to understand, that in this context the lowest-order Raviart-Thomas element is not equipped with a sufficient number of basis for the representation of a constant function. In other words, with only linear basis functions a simple constant vector as $(1,0,0)$ can not be represented considering generic trilinear transformation.

Likewise, it has also been shown that to obtain an approximate estimate for the divergence of the form

$$\inf_{\boldsymbol{v}\in\tilde{\boldsymbol{V}}_h} ||\boldsymbol{\nabla}\cdot(\boldsymbol{u}-\boldsymbol{v})||_{L^2(\Omega)} = o(1)\,, \tag{2.174}$$

we should consider a sequence of a partitioned domain $\mathcal{T}_h$ expressed with meshes of subcubes of side $h = 1/n$, satisfying the relation $\hat{\boldsymbol{\nabla}}\cdot\hat{\boldsymbol{V}} \supseteq \mathcal{P}_0(\hat{K})$.

On the other hand, for more general hexahedral elements, stronger conditions on $\hat{\boldsymbol{V}}$ are required. In particular, a necessary condition is that, after the mapping by using the Piola transform between the unit reference cube and the general hexahedron, it is possible to represent a constant that must be contained in the divergence of the resulting function on the real element.

The main result is that, in order to satisfy the continuity of $\hat{\boldsymbol{u}} \cdot \boldsymbol{n}$ across the elements that share a common face, other 19 coefficients must be added to the space $\hat{\boldsymbol{S}}_0^-$, thus leading to a final space dimension equal to 40. The technical details can be found in [71], where necessary conditions for the optimal approximation of the divergence are derived.

On the other hand, by restricting the set of all the possible geometric configurations derived by a generic trilinear transformation, the number of coefficients necessary to represent a constant vector can change. One of these configurations is represented by hexahedrons with flat faces. However, preliminary results in [71] seem to suggest that a linear space for the vector $\hat{\boldsymbol{u}}$ can not be obtained. Indeed, considering the more restrictive case described by a hexahedron with flat boundary and midplane faces, the best that can be achieved is the removal of the cubic and quartic terms. Hence, quadratic terms are still present in the expression for the local representation of the constant vector.

Another important discussion on the lowest-order hexahedral Raviart-Thomas element has been proposed in [64]. Specifically, a measure of the 'deviation' of a hexahedron concerning a parallelepiped has been introduced. In particular, the mapping $\mathcal{F}$ is decomposed as the sum of the affine part and another component that depends on four geometrical vector parameters, $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\boldsymbol{\varepsilon}_3$, $\boldsymbol{\varepsilon}_4$, $\in \mathbb{R}^3$.

To explain the meaning of these four parameters, consider a parallelepiped for which three vertices represent the canonical vector of the reference system, as depicted in Figure 2.8. Thus, by displacing the remaining vertex from the parallelepiped configuration, we obtain a generic hexahedron. The distance between the new positions of these vertices and their original positions is represented by the parameter $\boldsymbol{\varepsilon}_i$. Hence, the mapping can be rewritten as

$$\mathcal{F}_K(\hat{\boldsymbol{x}}) = \mathcal{F}_K^l(\hat{\boldsymbol{x}}) + \mathcal{F}_K^d(\hat{\boldsymbol{x}})\,. \tag{2.175}$$

The linear part of the mapping is defined as

$$\mathcal{F}_K^l(\hat{\boldsymbol{x}}) = \hat{x}_1\boldsymbol{\alpha}_1 + \hat{x}_2\boldsymbol{\alpha}_2 + \hat{x}_3\boldsymbol{\alpha}_3\,, \tag{2.176}$$

where $\boldsymbol{\alpha}_i := \mathcal{F}(\boldsymbol{e}_i)$, with $\boldsymbol{e}_i$ which is the canonical vector with $i = 1, 2, 3$. The other component is defined as

$$\mathcal{F}_K^d(\hat{\boldsymbol{x}}) = \hat{x}_2\hat{x}_3\boldsymbol{\varepsilon}_1 + \hat{x}_1\hat{x}_3\boldsymbol{\varepsilon}_2 + \hat{x}_1\hat{x}_2\boldsymbol{\varepsilon}_3 + \hat{x}_1\hat{x}_2\hat{x}_3(\boldsymbol{\varepsilon}_4 - \boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_3)\,. \quad (2.177)$$

An important aspect is that if the faces on the real element K are planar, then

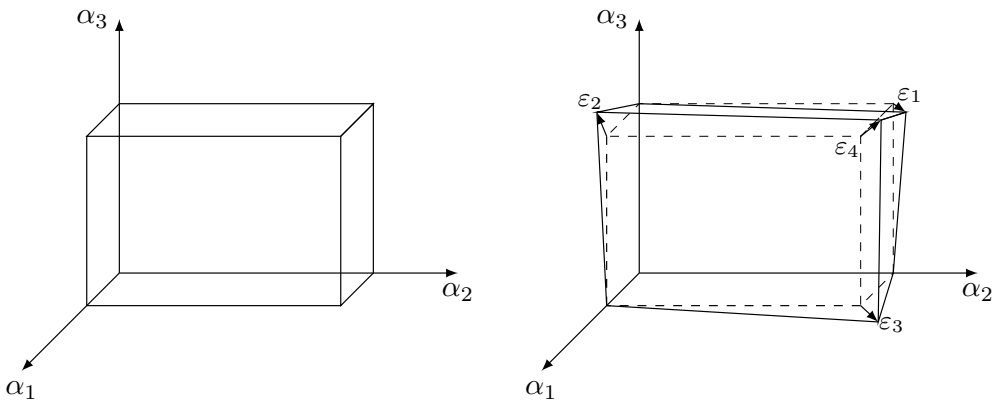

Figure 2.8: Parallelepiped on the left and corresponding non-affine hexahedron on the right, obtained applying the ε_i corner shifts.

$\boldsymbol{\varepsilon}_4 = \boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \boldsymbol{\varepsilon}_3$, and consequently, the cubic term of the mapping vanishes. In other words, the deviation of the vertices is not entirely arbitrary, since choosing the new positions of three vertices determines the position of the fourth. This situation is represented on the right in Figure 2.8, where the hexahedron is equipped only with planar faces.

To better understand the previous technical discussion, we refer now to [69, 68], where a trivial example of a non-regular hexahedron with a uniform constant velocity field is analyzed. Consider the geometry depicted in Figure 2.9, i.e. a truncated pyramid P with the two horizontal bases orthogonal to the z-axis. Thus, the bases are square with side lengths equal to s_0 and s_1, respectively. Consequently, the areas of the two bases are s_0^2 and s_1^2. Now, if we consider a uniform constant vector field $\boldsymbol{u}_1(\boldsymbol{x}) = (0, 0, 1)^T\,, \forall \boldsymbol{x} \in P$, the only nonzero component is orthogonal to the pair of parallel faces of the pyramid. Having a constant field, the flux through an interior horizontal section B_z of the element P can be easily computed since it is equal to the area of that section. Therefore, for a generic section B_z with $0 \leq z \leq 1$ we have

$$\int_{B_z} \boldsymbol{u}_1 \cdot \boldsymbol{n}_z = ((1 - z)s_0 + zs_1)^2\,, \quad (2.178)$$

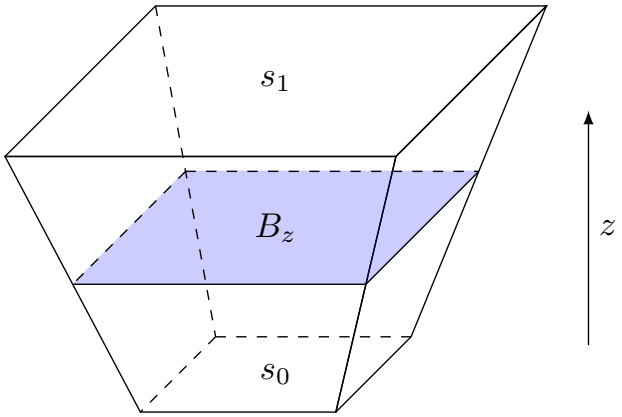

Figure 2.9: .

where naturally the side of the section B_z varies linearly with the coordinate z. Then, consider a reference cube where the cross-sectional area is denoted by $\hat{B}_z$, and compute the flux of a vector field $\boldsymbol{u}_2 \in \boldsymbol{\mathcal{RT}}_0$ by using the Piola transformation. Hence, we have

$$\int_{\hat{B}_z} \boldsymbol{u}_2 \cdot \boldsymbol{n}_z = (1-z)s_0^2 + zs_1^2\,, \tag{2.179}$$

because it is well-known that the flux varies linearly when considering vector-valued functions in $\boldsymbol{\mathcal{RT}}_0$. Therefore, since for a three-dimensional domain (2.178) is different from (2.179), the Piola transformation of a $\boldsymbol{\mathcal{RT}}_0$ function is not able to exactly represent a uniform flow field, even if the considered element is as simple as a truncated pyramid. Referring to the previous space $\hat{\boldsymbol{S}}_0^-$, in this simplified configuration, we can reduce the dimension from 40 to 28 [71].

Mesh construction

A remark on meshes discretized with generic hexahedral elements is now reported. The creation of a hexahedral mesh with planar faces is not as straightforward as in regular three-dimensional geometries. The deviation from a cubic geometry must satisfy the flatness condition for every face. Therefore, it is not sufficient to arbitrarily modify the eight vertices that define a cube. While there are several ways to build such a computational grid, it cannot be achieved by uniformly dividing a cube with several refinements. Instead, a replication of the 'fundamental unit grid' has to be performed in every spatial direction, with a scaling transformation to obtain the same

volume as the initial cube. For this thesis, the 'fundamental unit grid' is composed of eight non-affine hexahedrons, that form a unit cube. Specifically, each of the eight sub-cubes can be described with a two sub-set of faces:

- three faces are perpendicular to each other, and they represent and make up the external surface of the unit cube;
- the other three faces are internal to the unit cube, and they induce the non-affinity of the hexahedron. Indeed, these faces are not parallel to the corresponding faces and, therefore, they are not orthogonal to each other.

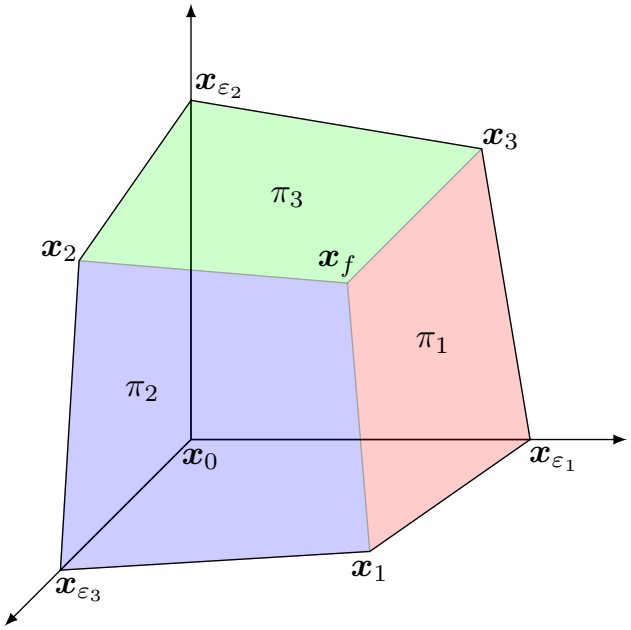

Figure 2.10: Single hexahedron formed by its eight vertices. Note that the vertex coordinates of $\boldsymbol{x}_f$ correspond to the intersection of three planes (red, blue and green).

To construct the height sub-cubes, it is sufficient to create the first one and generate the other cubes while respecting the constraint of the flat faces. For instance, by fixing four vertices $(\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$, it is possible to deform the sub-cube by moving only three vertices $(\boldsymbol{x}_{\varepsilon_1}, \boldsymbol{x}_{\varepsilon_2}, \boldsymbol{x}_{\varepsilon_3})$. Note that this movement must be executed within the plane defined by the original set of points. The eighth point $(\boldsymbol{x}_f)$ is determined by the intersection of the three planes (π_1, π_2, π_3) that are created after the shift of the three vertices. Indeed,

we know that for three points, a unique plane can be found. Therefore a flat face can only be constructed if the fourth vertex lies on the plane formed by the initial three vertices. A visual representation of this scenario is presented in Figure 2.10.

With this information, the other sub-cubes can be determined, as the coordinates of the first cube also provide the coordinates of the free vertices for the other cubes. The procedure for these points follows the same path, by finding three planes and solving their intersection. Moreover, if the initial deformation ε for the first three points has the same value and if these points lie on the hexahedron edges, it becomes possible to easily determine all 27 points coordinates of the unit cube. This can be achieved by exploiting some geometrical relation between the coordinates inside a cube, without performing the intersection between the planes. In particular, given the displacement ε we can define three expressions by which all coordinates can be found

$$\begin{aligned} d_1 &= 1+\varepsilon\,, & d_2 &= \frac{1+\varepsilon}{1+2\varepsilon}\,, \\ d_3 &= \frac{1+3\varepsilon}{1+2\varepsilon}\,, & d_4 &= \frac{1+5\varepsilon}{1+4\varepsilon}\,. \end{aligned} \tag{2.180}$$

In Figure 2.11 an example of the unit cube is reported with the connectivity map of its 27 nodes on the left in the initial configuration. On the right, the resulting geometry is depicted after applying the displacement ε to three vertices. Moreover, in (2.181) the points coordinates are reported following the expressions in (2.180), considering the unit cube of dimensions $2\times 2\times 2$. It is also possible to consider generic displacement for the first three vertices, allowing for multiple ε_i to be imposed. However, the geometric expressions relating the coordinates of the points to the shifts become more complex.

The eight non-affine elements are reported in Figure 2.12. Their combination produces the unit cube, as shown in Figure 2.11. Subsequently, the unit cube needs to be mirrored three times, considering symmetry with respect to the Cartesian planes, i.e. the planes OX, OY and OZ, to double the number of elements in each direction. This operation is necessary since the nodes on the external faces of the unit cube are not symmetric with respect to two opposite faces. The procedure can be repeated several times to construct the standard refined grids, such as $4\times 4\times 4$, $8\times 8\times 8$ and so on. Additionally, a scaling factor can be applied to adjust the geometry volume to a fixed value.

A similar approach to constructing a unit reference cube with eight non-affine hexahedrons has been proposed in [72], where three internal planes

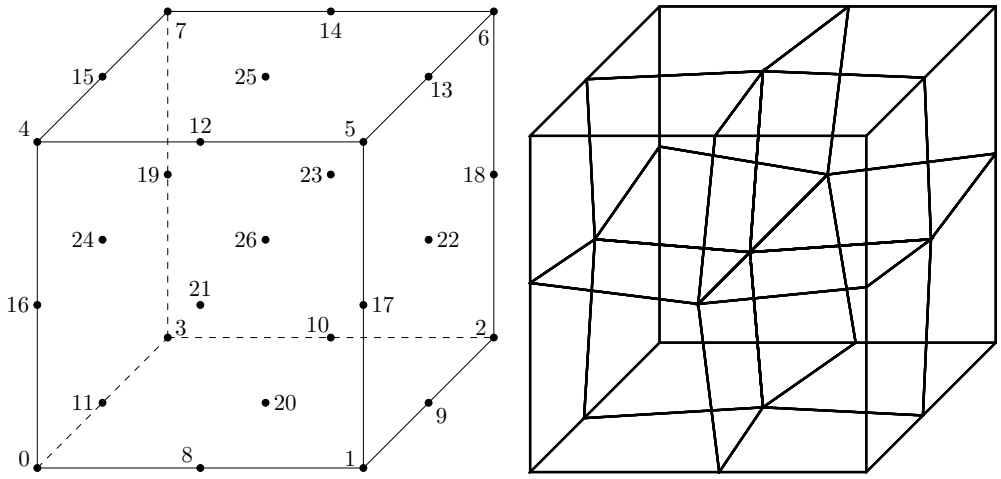

Figure 2.11: Unit reference cube on the left with the connectivity map for its 27 nodes. On the right, the same unit cube is built as the composition of 8 sub-cubes, which are the non-affine hexahedra elements.

are rotated around three axes. Hence, their intersections represent the edges of the hexahedrons. In this case as well, the authors recommend repeating the unit cube sub-mesh pattern over the domain, by applying appropriate reflections to maintain mesh conformity.

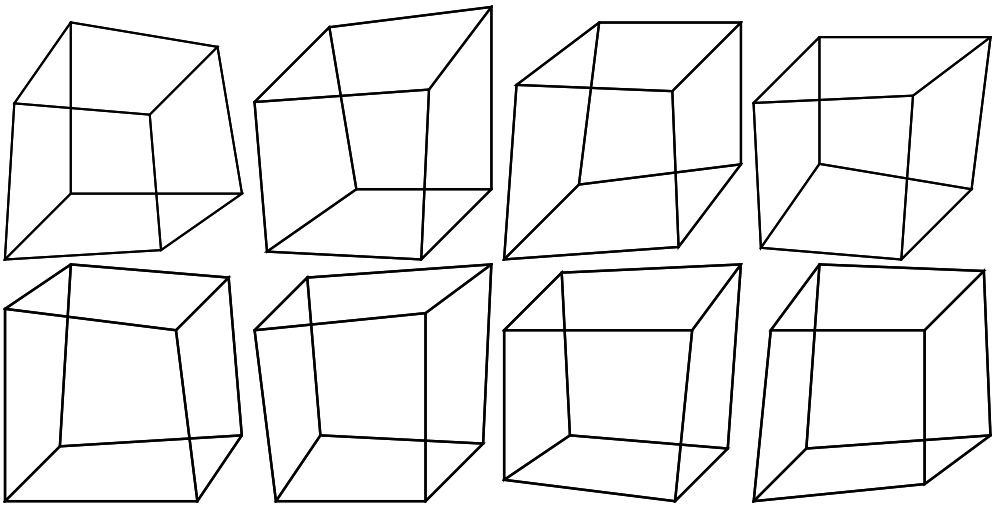

Figure 2.12: The eight non-affine hexahedrons equipped with six planar faces.

Points coordinates:

$$\begin{aligned}
&\{P_0(0,0,0)\,,\ P_1(2,0,0)\,,\ P_2(2,2,0)\,,\ P_3(0,2,0)\,,\ P_4(0,0,2)\,,\\
&P_5(2,0,2)\,,\ P_6(2,2,2)\,,\ P_7(0,2,2)\,,\ P_8(d_1,0,0)\,,\ P_9(2,d_3,0)\,,\\
&P_{10}(d_3,2,0)\,,\ P_{11}(0,d_1,0)\,,\ P_{12}(d_3,0,2)\,,\ P_{13}(2,d_4,2)\,,\ P_{14}(d_4,2,2)\,,\\
&P_{15}(0,d_3,2)\,,\ P_{16}(0,0,d_1)\,,\ P_{17}(2,0,d_3)\,,\ P_{18}(2,2,d_4)\,,\ P_{19}(0,2,d_3)\,,\\
&P_{20}(1,1,0)\,,\ P_{21}(1,0,1)\,,\ P_{22}(2,1,1)\,,\ P_{23}(1,2,1)\,,\ P_{24}(0,1,1)\,,\\
&P_{25}(1,1,2)\,,\ P_{26}(d_2,d_2,d_2)\}\,.
\end{aligned} \tag{2.181}$$

A non-affinity check has been performed to measure the actual skewness of these elements. Specifically, certain coefficients of the trilinear transformation have been computed, such as the coefficient expressed in (2.172). Considering that flat faces are the result of a transformation for which the cubic term vanishes, the coefficients related to the deformation from a regular cube to a non-affine hexahedron are G_{ij} with $i,j = 1,2,3$. Indeed, for a cube or for affine hexahedral elements, every G_{ij} is equal to zero. In (2.182) the expressions for the nine G_{ij} coefficients are reported

$$\begin{aligned}
&G_{11} = e_2g_3 - g_2e_3\,, \quad G_{12} = f_2e_3 - e_2f_3\,, \quad G_{13} = g_2f_3 - f_2g_3\,,\\
&G_{21} = e_3g_1 - g_3e_1\,, \quad G_{22} = f_3e_1 - e_3f_1\,, \quad G_{23} = g_3f_1 - f_3g_1\,,\\
&G_{31} = e_1g_2 - g_1e_2\,, \quad G_{32} = f_1e_2 - e_1f_2\,, \quad G_{33} = g_1f_2 - f_1g_2\,.
\end{aligned} \tag{2.182}$$

Furthermore, these coefficients depend only on the eight vertices that define the polyhedron, facilitating their computation. For each element, all the G_{ij} are non-zero, confirming the effectiveness of the algorithm used for the non-affine mesh creation.

CHAPTER 3

Projection method for Navier-Stokes equation

This chapter presents an overview of the projection method used for solving the Navier-Stokes equations in incompressible flow simulations, also known in literature as the fractional step method. This method broadly falls into three schemes: pressure-correction, velocity-correction, and consistent splitting methods. The following sections offer a brief review of these methods to better understand how this numerical scheme correlates with the finite element discretization described in earlier chapters. In particular, some steps present in the projection algorithm are essentially the same as the orthogonal decomposition of the velocity. Therefore, since the decomposition is treated by using the mixed finite element tools presented in Chapter 2, we establish a connection to fully resolve the Navier-Stokes system of equations. The following description draws heavily from the comprehensive work of Guermond et al. [23].

This numerical scheme originates from the incompressibility constraint, which forms the central theme of this thesis. In particular, one of the main issues related to the numerical solution of momentum and mass conservation equations is the coupling between velocity and pressure, imposed by the condition of zero divergence of velocity. The solution of the coupled velocity-pressure system (3.1) can be expensive from a computational point of view,

and this has led to the development of different numerical algorithms for the treatment of the split system for velocity and the pressure fields. In order to overcome these problems, many authors introduced the projection method by which the Navier-Stokes system is subdivided into two separate steps, one for the velocity and one for the pressure.

The first idea that appeared in literature is related to the work of Chorin and Temam [21, 22], where the time-dependent solution of incompressible viscous flow has been proposed. Specifically, the idea is to build a sequence of decoupled elliptic equations for the velocity and the pressure and solve them at each time step. Indeed, this numerical method leads to an efficient simulation and reduces the computational effort.

Since the convergence rate of the splitting error is not affected by the nonlinear term of the Navier-Stokes equations, the following analysis takes into account the time-dependent Stokes equations that, considering constant physical properties, read as

$$\begin{cases} \dfrac{\partial \boldsymbol{u}}{\partial t} - \nu \Delta \boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{f} \quad \text{in } \Omega \times [0,T]\,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \quad \text{in } \Omega \times [0,T]\,, \\ \boldsymbol{u}|_{\Gamma} = 0 \quad \text{in } \Omega \times [0,T]\,, \\ \boldsymbol{u}|_{t=0} = \boldsymbol{u}_0 \quad \text{in } \Omega\,, \end{cases} \tag{3.1}$$

where the right-hand side of the momentum equation is a smooth source and $\boldsymbol{u}_0$ denotes the initial value of the velocity field. Note that in this case, we consider homogeneous Dirichlet boundary conditions, since the use of natural boundary conditions needs more attention regarding the nonlinear term.

In order to define the time-dependent solutions, we can consider the time step $\Delta t > 0$ and set $t^k = k\Delta t$ for $0 \leq k \leq K = [T/\Delta t]$, where T is the upper bound of the time interval considered, $[0, T]$.

3.1 Pressure-correction schemes

For every algorithm that is described in the following subsection, the error estimates and theorem proofs can be found in [23] and references therein.

3.1.1 Non-incremental pressure-correction scheme

In this scheme, proposed originally in [21, 73], the implicit Euler scheme is adopted. Therefore, starting with the initial value of the velocity as $\boldsymbol{u}^0 = \boldsymbol{u}_0$,

at every step $k \geq 0$ we determine $\tilde{\boldsymbol{u}}^{k+1}, \boldsymbol{u}^{k+1}, p^{k+1}$ by solving the following equations

$$\frac{1}{\Delta t}(\tilde{\boldsymbol{u}}^{k+1} - \boldsymbol{u}^{k}) - \nu\Delta\tilde{\boldsymbol{u}}^{k+1} = \boldsymbol{f}(t^{k+1}), \quad \tilde{\boldsymbol{u}}^{k+1}|_{\Gamma} = 0, \tag{3.2}$$

$$\begin{cases} \dfrac{1}{\Delta t}(\boldsymbol{u}^{k+1} - \tilde{\boldsymbol{u}}^{k+1}) + \boldsymbol{\nabla} p^{k+1} = 0, , \\ \boldsymbol{\nabla}\cdot\boldsymbol{u}^{k+1} = 0 \quad \boldsymbol{u}^{k+1}\cdot\boldsymbol{n}|_{\Gamma} = 0. \end{cases} \tag{3.3}$$

Note that $\tilde{\boldsymbol{u}}$ is a fictitious velocity field, which is added (and subtracted) to the momentum equation of the Stokes system (3.1), in order to remove the pressure dependence from the velocity. In fact, the resulting momentum equation is then divided into two parts: the first one corresponds to the equation (3.2), while the other one is represented by (3.3). After that, by applying the divergence operator to the first equation of (3.3), we obtain,

$$\boldsymbol{\nabla}\cdot\tilde{\boldsymbol{u}}^{k+1} = -\Delta t\Delta p^{k+1}, \tag{3.4}$$

by which we are able to find the new value of $\tilde{\boldsymbol{u}}^{k+1}$ by solving a laplacian equation for the pressure. Subsequently, we can correct the value of the initial velocity $\boldsymbol{u}$ using (3.2). If the first step (3.2), takes into account the viscous effects, the second one (3.3), known also as the projection step, accounts for the incompressibility constraint. Note that the Laplacian equation applied to the pressure requires an additional boundary condition. Typically, an artificial Neumann boundary condition of the form $\boldsymbol{\nabla} p^{k+1}\cdot\boldsymbol{n}|_{\Gamma} = 0$ is applied, although this imposition lacks direct physical reasoning. An extensive analysis of this aspect has been presented in [74].

3.1.2 Standard incremental pressure-correction scheme

To enhance the accuracy of the model, Goda in [75] suggested adding the old value of the pressure gradient, i.e. $\boldsymbol{\nabla} p^{k}$, in the viscous equation, i.e. the first sub-step. Moreover, Van Kan in [76], proposed a second-order incremental pressure correction scheme, by using the backward difference formula of second order for the time derivative. In this case, the algorithm can be expressed as follows

$$\frac{1}{2\Delta t}(3\tilde{\boldsymbol{u}}^{k+1} - 4\boldsymbol{u}^{k} + \boldsymbol{u}^{k+1}) - \nu\Delta\tilde{\boldsymbol{u}}^{k+1} + \boldsymbol{\nabla} p^{k} = \boldsymbol{f}(t^{k+1}), \quad \tilde{\boldsymbol{u}}^{k+1}|_{\Gamma} = 0, \tag{3.5}$$

$$\begin{cases} \dfrac{1}{2\Delta t}(3\boldsymbol{u}^{k+1} - 3\tilde{\boldsymbol{u}}^{k+1}) + \boldsymbol{\nabla}(p^{k+1} - p^k) = 0\,, \\ \boldsymbol{\nabla}\cdot\boldsymbol{u}^{k+1} = 0 \quad \boldsymbol{u}^{k+1}\cdot\boldsymbol{n}|_\Gamma = 0\,. \end{cases} \tag{3.6}$$

Similar to the previous schemes, the presence of the hidden Laplacian equation for the pressure introduces a numerical boundary layer which is a non-physical Neumann boundary condition on the pressure. Naturally, this factor affects the accuracy of the scheme.

3.1.3 Rotational incremental pressure-correction scheme

An improvement to the previous schemes, specifically addressing the issue of artificial boundary conditions for pressure, was proposed in [77]. While the viscous step remains unchanged in comparison with the incremental-pressure scheme, the projection sub-step is modified as follows

$$\begin{cases} \dfrac{1}{2\Delta t}(3\boldsymbol{u}^{k+1} - 3\tilde{\boldsymbol{u}}^{k+1}) + \boldsymbol{\nabla}\phi^{k+1} = 0\,, \\ \boldsymbol{\nabla}\cdot\boldsymbol{u}^{k+1} = 0 \quad \boldsymbol{u}^{k+1}\cdot\boldsymbol{n}|_\Gamma = 0\,, \end{cases} \tag{3.7}$$

with the additional variable ϕ is defined as

$$\phi^{k+1} = p^{k+1} - p^k + \nu\boldsymbol{\nabla}\cdot\boldsymbol{u}^{k+1}\,. \tag{3.8}$$

Manipulating the previous relations, we obtain for the pressure the following boundary condition

$$\left.\frac{\partial p^{k+1}}{\partial n}\right|_\Gamma = (\boldsymbol{f}(t^{k+1}) - \nu\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\boldsymbol{u}^{k+1})\cdot\boldsymbol{n}|_\Gamma\,, \tag{3.9}$$

which is a more consistent boundary condition for the pressure. In particular, we have now only an inexact tangential boundary condition on the velocity, and since it is present in the operator $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times$ this scheme is denoted as the incremental pressure correction scheme in rotational form.

For the previous schemes, some remarks can be stated: note that, even though the velocity $\boldsymbol{u}$ is divergence-free, the appropriate boundary conditions are not satisfied. The situation is opposite considering $\tilde{\boldsymbol{u}}$, for which the Dirichlet boundary condition holds but is not divergence-free. On the other hand, it has been demonstrated that both variables are equipped with the same error estimates, i.e. the same accuracy.

3.2 Other type schemes

As the comprehensive examination of projection methods for the Navier-Stokes equations is beyond the scope of this thesis, we briefly outline other types of algorithms that, although not employed in our simulations, can be used to decouple velocity and pressure.

3.2.1 Non-incremental velocity-correction scheme

In this approach, the roles of pressure and velocity are interchanged, and therefore the first and second sub-steps are exchanged. For the references, the interested reader can consult [78, 79]. Giving the initial value of the velocity $\tilde{\boldsymbol{u}}^0 = \boldsymbol{u}_0$, we find $\tilde{\boldsymbol{u}}^{k+1}, \boldsymbol{u}^{k+1}, p^{k+1}$ for $k \geq 0$ as

$$\begin{cases} \dfrac{1}{\Delta t}(\boldsymbol{u}^{k+1} - \tilde{\boldsymbol{u}}^k) + \boldsymbol{\nabla} p^{k+1} = \boldsymbol{f}(t^{k+1})\,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u}^{k+1} = 0 \quad \boldsymbol{u}^{k+1} \cdot \boldsymbol{n}|_\Gamma = 0\,, \end{cases} \tag{3.10}$$

$$\frac{1}{\Delta t}(\tilde{\boldsymbol{u}}^{k+1} - \boldsymbol{u}^{k+1}) - \nu \Delta \tilde{\boldsymbol{u}}^{k+1} = 0\,, \quad \tilde{\boldsymbol{u}}^{k+1}|_\Gamma = 0\,. \tag{3.11}$$

Notably, compared to the standard Chorin-Temam algorithm, we enforce the following relations

$$\frac{\partial p^{k+1}}{\partial n}|_\Gamma = \boldsymbol{f}(t^{k+1})\,, \tag{3.12}$$

$$\Delta \tilde{\boldsymbol{u}}^{k+1} \cdot \boldsymbol{n}|_\Gamma = 0\,, \tag{3.13}$$

instead of having

$$\frac{\partial p^{k+1}}{\partial n}|_\Gamma = 0\,, \tag{3.14}$$

$$\Delta \tilde{\boldsymbol{u}}^{k+1} \cdot \boldsymbol{n}|_\Gamma = \boldsymbol{f}(t^{k+1})\,. \tag{3.15}$$

Another scheme discussed in [23] is the standard incremental velocity-correction scheme, which is analogous to the incremental pressure-correction algorithm. In this scheme, the viscous step follows the projection step, leading to the concept of calling them projection-diffusion methods. Additionally, the rotational form can be reformulated within the framework of incremental velocity-correction schemes.

Lastly, consistent splitting schemes compute velocity in the first step using an explicit formulation for pressure. Subsequently, the pressure is updated

by considering the L^2-inner product of the momentum equation with the gradient of the test function, i.e. $\boldsymbol{\nabla} q$, which reads

$$\int_\Omega \boldsymbol{\nabla} p \cdot \boldsymbol{\nabla} q = \int_\Omega (\boldsymbol{f} + \nu \Delta \boldsymbol{u}) \cdot \boldsymbol{\nabla} q \,, \quad \forall q \in H^1(\Omega) \,, \tag{3.16}$$

noting that $(\boldsymbol{u}_t, \boldsymbol{\nabla} q) = -(\boldsymbol{\nabla} \cdot \boldsymbol{u}_t, q) = 0$.

3.2.2 The implemented algorithm

In this section, the algorithm employed for the resolution of the Navier-Stokes system in the finite element library FEMuS [80] is described. This technique is applied for the solution of numerical problems that are presented in Section 4, where a multigrid approach is available in the numerical code. On the other hand, several physical problems can be addressed with the FEMuS code, from fluid-structure interaction, optimal control problems, turbulence modeling and so on. Numerical examples of the physical problems addressed by FEMuS can be found in [80, 81] and references therein. In this thesis, we report numerical examples regarding only the resolution of the Navier-Stokes system, since the topic of this work is related to numerical methods for the equation solution. In particular, the projection technique is specifically developed for the resolution of the momentum equation, and therefore the employment of Raviart-Thomas finite elements for the resolution of different physical problems has not already been investigated.

We recall the Navier-Stokes system for a generic incompressible fluid flow, with constant physical properties

$$\begin{aligned} \frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \boldsymbol{\nabla}) \boldsymbol{u} &= -\boldsymbol{\nabla} p + \nu \Delta \boldsymbol{u} + \boldsymbol{f} \,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0 \,. \end{aligned} \tag{3.17}$$

To describe the adopted pressure-velocity split algorithm, the simple Euler scheme for the time discretization is used. Specifically, a fictitious velocity field $\boldsymbol{u}^{n*}$ is introduced leading to

$$\frac{\boldsymbol{u}^{n*} - \boldsymbol{u}^{n-1*}}{\Delta t} + \frac{\boldsymbol{u}^{n} - \boldsymbol{u}^{n-1}}{\Delta t} + (\boldsymbol{u}^n \cdot \boldsymbol{\nabla}) \boldsymbol{u}^n = -\boldsymbol{\nabla} p^n + \nu \Delta \boldsymbol{u}^n + \boldsymbol{f} \,. \tag{3.18}$$

After that, following a standard approach, the equation can be subdivided in

two different equations by introducing $p^n = p^{n-1} + \delta p^n$

$$\frac{\boldsymbol{u}^n - \boldsymbol{u}^{n*}}{\Delta t} = -\boldsymbol{\nabla}\delta p^n \,, \tag{3.19}$$

$$\frac{\boldsymbol{u}^{n*} - \boldsymbol{u}^{n-1}}{\Delta t} = -(\boldsymbol{u}^n \cdot \boldsymbol{\nabla})\boldsymbol{u}^n - \boldsymbol{\nabla} p^{n-1} + \nu \Delta \boldsymbol{u}^n + \boldsymbol{f} \,. \tag{3.20}$$

In this thesis, a different velocity-pressure split is proposed and analyzed with the objective of exploiting the orthogonal decomposition of the velocity field [82]. Specifically, the first equation, which solves the velocity remains the same, i.e.

$$\frac{\boldsymbol{u}^{n*} - \boldsymbol{u}^{n-1}}{\Delta t} = -(\boldsymbol{u}^{n*} \cdot \boldsymbol{\nabla})\boldsymbol{u}^{n*} - \boldsymbol{\nabla} p^{n-1} + \nu \Delta \boldsymbol{u}^{n*} + \boldsymbol{f} \,, \tag{3.21}$$

while the pressure equation is not transformed into the standard laplacian equation, thus we have

$$\boldsymbol{u}^n - \boldsymbol{u}^{n*} + \Delta t \boldsymbol{\nabla}\delta p^n = 0 \,. \tag{3.22}$$

The idea is to apply the orthogonal decomposition to (3.22), by using the properties of the Raviart-Thomas finite element family. In particular, in this thesis, Equation (3.22) is solved by considering an optimal minimization problem. The objective is to find a field $\boldsymbol{u}$ which is the closest velocity to $\boldsymbol{u}^*$, under the constraint of the divergence equal to zero.

Therefore, for the velocity $\boldsymbol{u}^*$ the solution space can be found considering the standard solution of the Navier-Stokes equations. In particular, consider the Taylor-Hood finite element space $\boldsymbol{X}_h \subset \boldsymbol{H}^1(\Omega)$. Thus, if we have $\boldsymbol{u}_h^* \in \boldsymbol{X}_h$, the objective is to find the velocity $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0 \subset \boldsymbol{H}(\mathrm{div}, \Omega)$ by minimizing

$$\mathcal{F}(\boldsymbol{u}_h) = \frac{1}{2}\int_\Omega (\boldsymbol{u}_h - \boldsymbol{u}_h^*)^2 d\Omega, \tag{3.23}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u}_h = 0 \,, \tag{3.24}$$

over the linear function subspace $\boldsymbol{\mathcal{RT}}_0 \subseteq \boldsymbol{H}(\mathrm{div}, \Omega)$.

Naturally, for $\boldsymbol{u}_h^* \in \boldsymbol{X}_h \subset \boldsymbol{H}^1(\Omega)$, the velocity approximation is defined as $\boldsymbol{u}_h^* = \sum_j \boldsymbol{u}_{jh}^* \varphi_j(\boldsymbol{x})$, where $\boldsymbol{u}_{jh}^*$ is the velocity field at the points j and $\varphi_j(\boldsymbol{x})$ indicates the Lagrangian quadratic polynomial basis functions. Instead, for $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0 \subset \boldsymbol{H}(\mathrm{div}, \Omega)$, i.e. the lowest-order Raviart-Thomas subspace, the approximation velocity reads as $\boldsymbol{u}_h = \sum_f p_f \boldsymbol{N}_f(\boldsymbol{x})$, where p_f represents the fluxes through the faces and $\boldsymbol{N}_f(\boldsymbol{x})$ corresponds to the Raviart-Thomas vector basis functions.

In order to derive the final system to be solved, we can consider now a piecewise constant pressure, i.e. $k_h = \delta p_h \in S_h \subset L^2(\Omega)$, $\boldsymbol{u}_h^n \in \boldsymbol{\mathcal{RT}}_0 \subset \boldsymbol{H}(\mathrm{div}, \Omega)$, which satisfies $\boldsymbol{\nabla} \cdot \boldsymbol{u}_h^n = 0$. Therefore, the system to be solved is described with the following two equations

$$\int_\Omega \delta k_h \boldsymbol{\nabla} \cdot \boldsymbol{u}_h^n \, d\Omega = 0 \qquad \forall \delta k_h \in S_h \subset L^2(\Omega)\,, \tag{3.25a}$$

$$\int_\Omega (\boldsymbol{u}_h^n - \boldsymbol{u}_h^{*n}) \delta \boldsymbol{u}_h^n \, d\Omega + \int_\Omega k_h \boldsymbol{\nabla} \cdot \delta \boldsymbol{u}_h^n \, d\Omega = 0 \qquad \forall \delta \boldsymbol{u}_h^n \in \boldsymbol{\mathcal{RT}}_0(\Omega)\,, \tag{3.25b}$$

where k_h can be interpreted as the Lagrange multiplier of the divergence equation.

It is worth noting that the set of equations (3.25), is only a step of the split algorithm. Indeed, the velocity equation (3.21) is still solved by considering standard Lagrangian finite elements, in particular the Taylor-Hood type in our case.

On the other hand, the previous algorithm has the benefit of decreasing the computational effort. Indeed, we recall that by using the lowest-order $\boldsymbol{\mathcal{RT}}$ element the velocity is determined only with four (six) degrees of freedom for a bidimensional (three-dimensional) domain. Naturally, in this discussion, only quadrilateral and hexahedral elements have been considered.

Hence, the local matrix is equipped only with five (seven) rows that represent the four (six) fluxes through the faces and the central value for the pressure field, instead of using, for example, the nine nodes of a biquadratic quadrilateral with Lagrangian basis functions. For this reason, the total number of degrees of freedom of the simulation is smaller, gaining benefits from a computational point of view. Exploiting velocity fluxes through the element faces as unknowns, allows to ease the global sparsity pattern of the matrix since every row is equipped with a smaller number of entries.

CHAPTER 4

Divergence free applications

In this chapter, our goal is to showcase and discuss fluid flow simulations developed within the theoretical framework outlined in the previous chapters. Thus, we apply the mixed finite element discretization discussed in Chapter 2 to validate arguments regarding the well-known constraint of a divergence-free velocity field. We not only restrict the analysis to the implementation of the divergence-free finite element but also explore how this specific configuration is suitable for applying projection algorithms to solve the Navier-Stokes equations.

In the first part, we present a trivial numerical discussion about the definition of affine, non-affine, and asymptotically affine mesh, with the aim of better understanding the well-known theoretical results. Subsequently, some numerical tests related to the orthogonal velocity decomposition are presented, considering bidimensional and three-dimensional geometries. Error norms are also reported in order to verify the theoretical order of convergence by using Raviart-Thomas finite elements. After that, a few examples regarding the application of the projection method are shown.

4.1 Remark on affine quadrilateral elements

In this section, an example of an asymptotically affine quadrilateral element is presented considering a simple geometrical example. The goal is to numerically illustrate the reason why a multigrid refinement can overcome the issue related to non-affine meshes. A theoretical explanation of this problem has been provided in Section 2.2.3. In particular, we recall that a quadrilateral mesh is asymptotically affine if $\sigma_K = O(h_K)$ [48], i.e. the ratio σ_K/h_K is uniformly bounded for every element of the mesh. Naturally, h_K represents the maximum element diameter, while σ_K has been defined as $\max(|\pi - \theta_1|, |\pi - \theta_2|)$, where θ_1 and θ_2 are the angles between the outward normal vectors of the two pairs of opposite sides.

Note that, with a multigrid refinement, we refer to the standard technique of finding the four child elements by equally dividing by two the four edges of the parent element. An example of this mesh is shown in Figure 2.7, where an asymptotically trapezoidal affine mesh is shown. Otherwise, to maintain the non-affinity characteristic throughout the entire domain and across the different levels of refinement, it is necessary to perform a 'remesh' of the discretized grid: the initial pattern of non-affinity must be replicated for every direction and preserved for the entire geometry like in Figure 2.6. Note that in this case, the ratio between the length of opposite edges does not change with the refinement, since the angles between adjacent sides are preserved.

Consider a simple trapezoidal element $\mathcal{T}$ as shown in Figure 4.1. It can be noticed that the left and the right edges are not parallel since the point $\boldsymbol{x}_2$ does not have the same x coordinate of point $\boldsymbol{x}_1$. In particular, we can consider a shift toward the point $\boldsymbol{x}_3$ by a value ε. The idea is to analyze the component values of the Jacobian matrix in comparison with a standard square element $\mathcal{Q}$.

To fix the idea, consider a square $\mathcal{Q}$ of dimension $[0, 1]^2$ and the corresponding trapezoidal element $\mathcal{T}$, where $\varepsilon = 0.2$. A homogeneous refinement has been performed on the trapezoidal mesh as depicted in Figure 4.2, and for every level, the Jacobian matrix of the element equipped with the vertex $\boldsymbol{x}_0$ has been computed. The comparison has been done considering the same

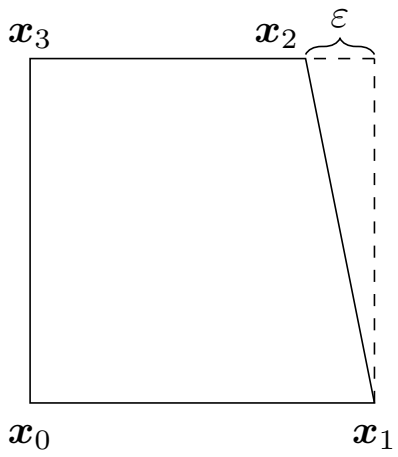

Figure 4.1: Trapezoidal element deriving from a unit square by applying a ε shift to a corner.

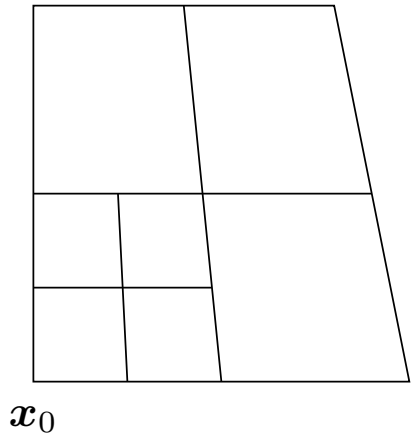

Figure 4.2: Homogeneous refinement of the trapezoidal element.

child square of the initial $\mathcal{Q}$. At the first level, the Jacobian matrices read,

$$\mathcal{DF}_{\mathcal{Q}_0} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \qquad \mathcal{DF}_{\mathcal{T}_0} = \begin{pmatrix} \frac{9-\eta}{20} & \frac{-1-\xi}{20} \\ 0 & \frac{1}{2} \end{pmatrix}, \tag{4.1}$$

where the respective determinants are $\mathcal{J}_{\mathcal{Q}_0} = 1/4$ and $\mathcal{J}_{\mathcal{T}_0} = (9-\eta)/40$. As expected, the presence of an off-diagonal element in the matrix $\mathcal{DF}_{\mathcal{T}_0}$, confirms the non-affinity for the trapezoidal element. At the second level,

the same child element has the following Jacobian matrices

$$\mathcal{DF}_{\mathcal{Q}_1} = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}, \qquad \mathcal{DF}_{\mathcal{T}_1} = \begin{pmatrix} \frac{19-\eta}{80} & \frac{-1-\xi}{80} \\ 0 & \frac{1}{4} \end{pmatrix}, \tag{4.2}$$

where, naturally, the determinants become $\mathcal{J}_{\mathcal{Q}_1} = 1/16$ and $\mathcal{J}_{\mathcal{T}_1} = (19 - \eta)/320$. Even if trivial, this example shows the different scaling rates of the elements' matrix. In fact, when we examine the off-diagonal element of the matrix $\mathcal{DF}_{\mathcal{T}_1}$ the scaling factor is twice that of the other element in the same column (the first goes with 4^n, while the second with 2^n, where n is level number). The computations performed at further levels have confirmed this result.

It is possible to generalize the previous analysis considering a generic shift ε and a refinement level r. In fact, the bottom-left refined element can be described with the following vertices: (0,0), $(\frac{1}{2^r}, 0)$, $(\frac{1}{2^r} - \frac{\varepsilon}{2^r}, \frac{1}{2^r})$, $(0, \frac{1}{2^r})$. With this notation, the Jacobian matrix can be reformulated in a generic way as

$$\begin{aligned} \mathcal{DF}_{\mathcal{T}_r} &= \begin{pmatrix} 2^{-2(r+1)}(2^{r+1} - (\eta+1)\varepsilon) & -2^{-2(r+1)}(\xi+1)\varepsilon \\ 0 & 2^{-r-1} \end{pmatrix} = \\ &= \frac{1}{2^{r+1}} \begin{pmatrix} 1 - \frac{(\eta+1)\varepsilon}{2^{r+1}} & -\frac{(\xi+1)\varepsilon}{2^{r+1}} \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{4.3}$$

From this notation, it is easy to understand that the off-diagonal term of the Jacobian matrix goes faster to zero by increasing the refinement r, considering a generic shift ε.

It is noticeable that, if the off-diagonal represents the deformation from a squared shape, this kind of deformation tends to decrease faster with respect to the diagonal elements, which represent the scaling factor. Therefore, if graphically a refined trapezoid looks more and more similar to a parallelogram shape, i.e. the generic configuration of an affine quadrilateral, the trend of the Jacobian matrix elements values proves from a numerical perspective the graphical intuition.

The same analysis has been carried out for a complete non-affine mesh such as the one depicted in Figure 2.6. Also in this case, the analyzed element

$\mathcal{C}$ is the bottom left trapezoid of the base mesh, defined initially with coordinates $(0,0),(0.5,0),(0.5,2/3),(0,1/3)$. The Jacobian matrix for the first and second levels of refinement are,

$$\mathcal{DF}_{\mathcal{C}_1} = \begin{pmatrix} \frac{1}{4} & 0 \\ \frac{1+\eta}{12} & \frac{3+\xi}{12} \end{pmatrix}, \qquad \mathcal{DF}_{\mathcal{C}_2} = \begin{pmatrix} \frac{1}{8} & 0 \\ \frac{1+\eta}{24} & \frac{3+\xi}{24} \end{pmatrix}, \tag{4.4}$$

where the corresponding determinants read as $\mathcal{J}_{\mathcal{C}_1} = (3+\xi)/48$ and $\mathcal{J}_{\mathcal{C}_2} = (3+\xi)/192$. Hence, we can notice that the scaling rate is the same for every element of the matrix since, for this mesh, the geometrical pattern of the parent element is preserved in the four sub-elements. For this reason, the non-affinity property does not vanish with the grid refinement and characterizes the Jacobian matrix of every element.

4.2 Preliminary results about divergence-free flows

In this section, we present a computational example that, while trivial, holds significant value in understanding issues related to the divergence constraint in standard mixed methods.

4.2.1 No-flow problem for the Stokes equations

In this numerical example, we aim to solve the stationary Stokes equations which reads

$$-\nu\Delta\boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{f} \tag{4.5}$$

$$-\boldsymbol{\nabla}\cdot\boldsymbol{u} = g\,, \tag{4.6}$$

where we consider no-slip boundary conditions at the domain boundary, i.e. $\boldsymbol{u}|_{\partial\Omega} = 0$. Additionally, we impose a pressure outlet boundary condition on the top side of the domain, setting the pressure value to zero. In particular, we want to solve this system of equations considering a square domain, $\Omega = (0,1)^2$, where naturally the incompressibility constraint holds, thus $g = 0$, and with the kinematic viscosity $\nu = 1$. The geometry chosen for this simulation is shown in Figure 4.3, where $l = 1$.

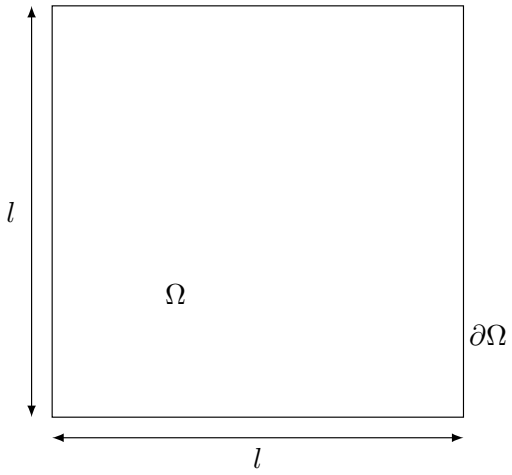

Figure 4.3: Square domain Ω of dimension $l \times l$, with boundary $\partial\Omega$.

We fix the body force at the right-hand side by $\boldsymbol{f} = (0, Ra(1-y+3y^2))^T$, where x, y are the domain coordinates and Ra is a positive parameter. The domain is discretized with standard biquadratic elements, starting with a mesh discretization of 4 elements, 2 for each direction. The convergence analysis has been performed using a standard multigrid algorithm.

We find analytically and numerically that the solution to the equations gives $\boldsymbol{u} = 0$ and $p = Ra(y^3 - y^2/2 + y - 3/2)$. While changing the parameter Ra on the right-hand side only affects the pressure solution, we notice an influence on the discrete velocity when using a standard finite element pair of the Taylor-Hood type (i.e., $\mathcal{P}_2 - \mathcal{P}_1$).

From a numerical perspective, the simulations have been performed by varying the Ra number. For each Ra the L^2-norm of the velocity gradient error $||\boldsymbol{\nabla}(\boldsymbol{u} - \boldsymbol{u}_h)||_{L^2(\Omega)}$ has been computed, where the analytical solution of the velocity is $\boldsymbol{u} = 0$. This behavior, known in the literature as a lack of robustness, is related to the discretization of the divergence constraint for mixed methods. In fact, we recall that for the standard finite element family of the Taylor-Hood type, if we consider the pair $\mathcal{P}_2 - \mathcal{P}_1$, we know that, calling the space of the discretized velocity $\boldsymbol{X}_h^2$ and X_h^1 for the pressure, $\boldsymbol{\nabla} \cdot \boldsymbol{X}_h^2 \not\subseteq X_h^1$.

In Figure 4.4, the plot illustrates that a change in the Ra parameter, which affects only the right-hand side, also influences the gradient of the velocity error. Consequently, we are not able to separate the behavior of velocity from pressure when using a standard Taylor-Hood finite element

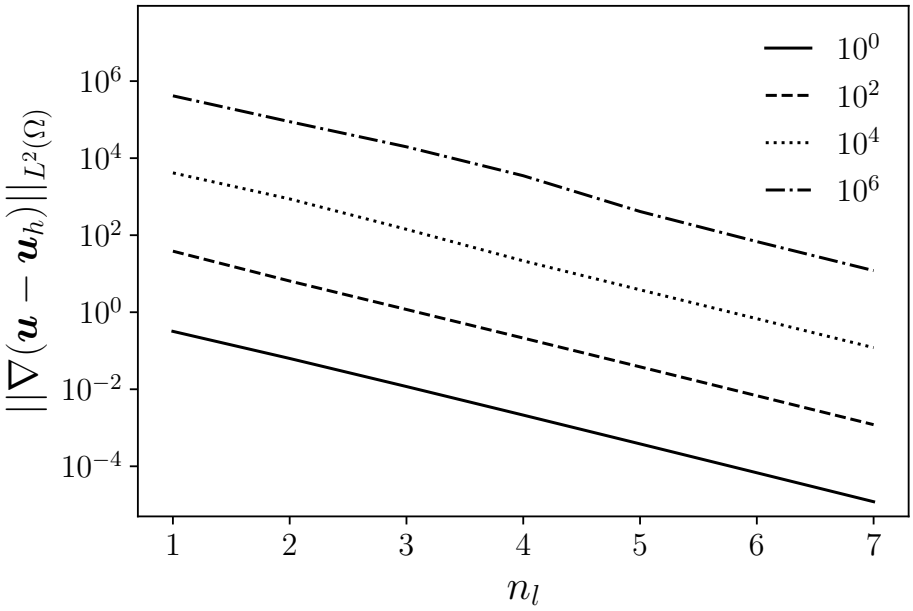

Figure 4.4: Gradient of velocity error with different values of Ra (from 1 up to 10^6), as a function of the level refinement n_l.

discretization, since we have shown that the error on velocity depends on the error on pressure.

4.3 Divergence-free velocity decomposition

In this section, the numerical results related to the presented orthogonal decomposition of the velocity are shown. In particular, we aim to decompose a vector field by following the Helmholtz theorem.

Two types of problems have been investigated to compare the results with numerical simulations presented in the literature. The first problem is the classical Darcy problem, which is a homogeneous Dirichlet-Neumann boundary elliptic problem described as follows

$$\begin{cases} \boldsymbol{s} = -\mathcal{K}\boldsymbol{\nabla} p & \text{in } \Omega\,, \\ \boldsymbol{\nabla}\cdot\boldsymbol{s} = -f & \text{in } \Omega\,, \\ \boldsymbol{s}\cdot\boldsymbol{n} = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{4.7}$$

where $\mathcal{K}$ is the permeability tensor divided by the viscosity and Ω represents a polygonal domain as described in the previous sections. Therefore, consid-

ering the spaces $\boldsymbol{V}^0 \subset \boldsymbol{H}(\mathrm{div}, \Omega)$ as defined in (1.35) and $Q = L^2_0(\Omega)$ as in (1.36), we can reformulate the problem in the following form: find $\boldsymbol{s} \in \boldsymbol{V}^0$ and $p \in Q$ such that

$$\begin{cases} (\mathcal{K}^{-1}\boldsymbol{s}, \boldsymbol{v}) + (\boldsymbol{\nabla} \cdot \boldsymbol{v}, p) &= 0 & \forall \boldsymbol{v} \in \boldsymbol{V}^0 \\ (\boldsymbol{s}, q) &= -(f, q) & \forall q \in Q\,. \end{cases} \tag{4.8}$$

Using the discretized finite element spaces of the previous sections, and considering the spaces $\boldsymbol{V}_h \subset \boldsymbol{V}^0$ and $Q_h \subset Q$, the discrete problem reads: find $\boldsymbol{s}_h \in \boldsymbol{V}_h$ and $p_h \in Q_h$ such that

$$\begin{cases} (\mathcal{K}^{-1}\boldsymbol{s}_h, \boldsymbol{v}) + (\boldsymbol{\nabla} \cdot \boldsymbol{v}, p_h) &= 0 & \forall \boldsymbol{v} \in \boldsymbol{V}_h \\ (\boldsymbol{s}_h, q) &= -(f, q) & \forall q \in Q_h\,. \end{cases} \tag{4.9}$$

Another formulation of the problem (4.7), with the hypothesis of free divergence for the velocity field, is obtained setting by $\boldsymbol{s} = \boldsymbol{u} - \boldsymbol{u}^*$. In particular, the problem is strictly connected to the Helmholtz decomposition of a vectorial field and takes the following form

$$\begin{cases} \boldsymbol{u} - \boldsymbol{u}^* = -\mathcal{K}\boldsymbol{\nabla} p & \text{in } \Omega\,, \\ \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 & \text{in } \Omega\,, \\ \boldsymbol{u} \cdot \boldsymbol{n} = 0 & \text{on } \partial\Omega\,. \end{cases} \tag{4.10}$$

According to the Helmholtz decomposition in Lemma 1.4 and following the functional spaces used for the system (4.7), we have that $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, \Omega)$, meanwhile $\boldsymbol{u}^* \in \boldsymbol{L}^2_0(\Omega)$. Specifically, $\boldsymbol{u}$ belongs to a subspace of $\boldsymbol{H}(\mathrm{div}, \Omega)$ since the condition $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$ holds.

On the other hand, as explained in Section 3.2.2, the velocity decomposition can be understood as a part of the second step of the projection algorithm. For this reason, we can restrict the space of the variable $\boldsymbol{u}^*$ and consider $\boldsymbol{H}^1(\Omega)$. We recall that the Stokes equation is equipped with the Laplace operator, which, after integrating by part for the weak formulation, presents the gradient operator. Thus, the standard mathematical space to be considered is $\boldsymbol{H}^1(\Omega)$. On the other hand, the Darcy equation in (4.7) presents only the divergence operator, and for this reason, we can extend the functional space for the velocity field considering a less regular solution, i.e. the space $\boldsymbol{H}(\mathrm{div}, \Omega)$.

In the following discussion, we always consider $\boldsymbol{u}^* \in \boldsymbol{H}^1(\Omega)$ and $\boldsymbol{u} \in \boldsymbol{H}(\mathrm{div}, \Omega)$ with the usual $L^2_0(\Omega)$ space for the pressure. In particular, we

have $\boldsymbol{u} \in \boldsymbol{V}^0$ since the condition $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ holds, while its orthogonal space $\boldsymbol{V}^\perp$ can be introduced for the pressure gradient. In fact, since $\boldsymbol{V}^0$ is a closed subspace of $\boldsymbol{L}^2$, we have the following decomposition

$$\boldsymbol{L}^2(\Omega) = \boldsymbol{V}^0 + \boldsymbol{V}^{0,\perp}\,, \tag{4.11}$$

considering a standard scalar product between the spaces. Naturally, in our context, the decomposition of the field must be applied to $\boldsymbol{u}^*$ which still belongs to $\boldsymbol{L}^2(\Omega)$ since it is in $\boldsymbol{H}^1(\Omega)$. Thus, $\boldsymbol{u}^*$ is expressed as the sum of a gradient of a scalar field ($\boldsymbol{\nabla} p$) and a free-divergence field ($\boldsymbol{u}$).

The objective of the following numerical test is to find a discretized velocity $\boldsymbol{u}_h \in \boldsymbol{\mathcal{RT}}_0 \subset \boldsymbol{V}^0 \subset \boldsymbol{H}(\mathrm{div}, \Omega)$ corresponding to the divergence-free part of the velocity field $\boldsymbol{u}^*$ which is given analytically. Several cases have been evaluated, differing by geometry and boundary conditions. The computations are performed by using FEMuS code [80].

Lastly, note that the following numerical results have been obtained while considering the theoretical issues concerning error convergence of the considered field data. Several types of discretized geometries have been investigated, including both affine and non-affine meshes. The theoretical results presented in Chapter 2 have been verified in both two and three-dimensional geometries. For this reason, examples characterized by non-affine meshes have not been reported. On the other hand, numerical examples where asymptotically affine elements have been found are reported.

4.3.1 Two-dimensional problems

Non-affine mesh test

In this section, we present the numerical results concerning the systems (4.7) in the context of a non-affine mesh. Specifically, a trapezoidal mesh has been employed, ensuring a pair of not parallel edges for the quadrilateral element. Figure 4.5 illustrates the discretized square geometry.

Since the first test is based on (4.7), a pressure field p has been fixed, from which the vector $\boldsymbol{s}$ has been derived. In fact, we have that

$$p = 4x(1-x)y(1-y)\,, \tag{4.12}$$

from which

$$\boldsymbol{s} = \nabla p = \begin{bmatrix} 4(1-2x)y(1-y) \\ 4(1-2y)x(1-x) \end{bmatrix}. \tag{4.13}$$

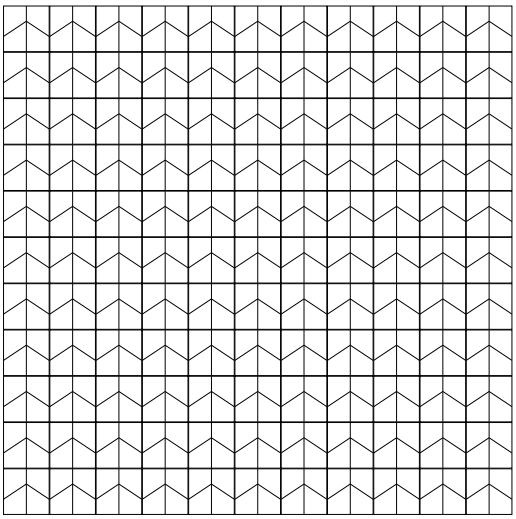

Figure 4.5: Trapezoidal mesh.

Since $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$ is imposed, we have that $\boldsymbol{u}^*$ is represented only by an irrotational field, and therefore we have that $\boldsymbol{\nabla} \cdot \boldsymbol{s} = -\boldsymbol{\nabla} \cdot \boldsymbol{u}^*$. The computations have been performed by using different refined meshes, which maintain the same trapezoidal pattern. The error norm of the velocity, the pressure, and the divergence of velocity have been reported in Table 4.1, as a function of the refinements levels. In particular, the coarsest grid has 4×4 elements, while the finest 128×128. We recall that to be precise, the velocity divergence $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ should be understood in the sense of mimetic divergence DIV.

In fact, as described in Chapter 2, non-affine meshes, such as the one considered in this case, may lead to a non-convergence of the divergence error. On the other hand, the mimetic finite difference (i.e. the Bochev technique 2.2.5) has been implemented and exploited for the computation. We recall that, with this method, the Jacobian matrix is not computed varying the $\hat{\boldsymbol{x}}$ coordinates, but $\mathcal{JF}$ is evaluated at the center of the element.

The orders of convergence have been reported, where p_u represents the convergence rate for the velocity error, p_p for the pressure error, and p_d for the divergence error. These convergence rates have been computed following the relation

$$p_i = \frac{ln\left(\frac{||f - f_h||_{0,l-1}}{||f - f_h||_{0,l}}\right)}{ln(2)}, \tag{4.14}$$

l	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	$\|\boldsymbol{\nabla}\cdot\boldsymbol{u}-\boldsymbol{\nabla}\cdot\boldsymbol{u}_h\|_0$	p_u	p_p	p_d
2	6.86E-08	1.86E-03	2.37E-01	-	-	-
3	3.31E-10	5.01E-04	6.11E-02	3.85	0.94	0.98
4	1.54E-12	1.28E-04	1.54E-02	3.87	0.99	0.99
5	7.00E-15	3.21E-05	3.86E-03	3.89	1.00	1.00
6	3.11E-17	8.03E-06	9.64E-04	3.91	1.00	1.00
7	1.37E-19	2.01E-06	2.41E-04	3.91	1.00	1.00

Table 4.1: Convergence trend of the error norm for problem (4.7), considering a non-affine mesh, by using mimetic divergence. In the last columns, the orders of convergence of the computed variables are reported.

where f represents the variable for which the error norm is computed, thus $\boldsymbol{u}, p$ and $\boldsymbol{\nabla}\cdot\boldsymbol{u}$. Naturally, the L^2-norm subscripts stand for the error evaluated at the levels l and $l-1$.

We can notice that the implementation of the mimetic divergence can overcome the numerical issues related to non-affine elements since an evident convergence can be observed for each variable. Additionally, note that also $\boldsymbol{\nabla}\cdot\boldsymbol{u}$ is capable of converging since mimetic finite differences have been used. In fact, as expected the order of convergence tends to 1 as described by the theoretical discussion in Equation (2.114) and [58]. In fact, also in the latter reference, the convergence of the velocity divergence shows a linear trend.

Plane channel test

This test considers a plane channel geometry, as shown in Figure 4.6. The dimensions are reported in the Figure with the variables L_x and L_y. The geometry boundaries are denoted with Γ, where Γ_i and Γ_o represent the inlet and the outlet of the channel, while with Γ_w we denote the other two boundaries.

Standard quadrilateral affine elements have been considered for mesh discretization, avoiding the numerical issue related to distorted grids. In fact, the finite element library has been equipped only with the lowest-order Raviart-Thomas family, as one of the simulation objectives was to reduce the computational effort associated with the number of degrees of freedom

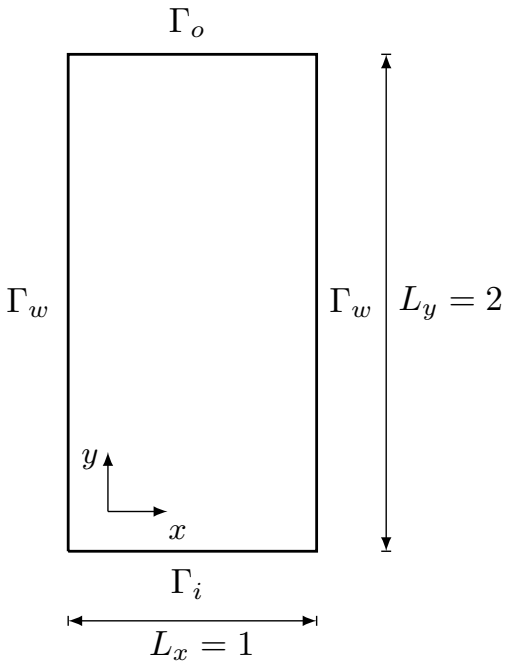

Figure 4.6: Channel geometry for the bidimensional velocity decomposition test.

for every element. We recall that in this case, the number of unknowns is only 5 for every element, i.e. the four fluxes over the faces and the piecewise constant value of the pressure.

An analytical vector field $\boldsymbol{u}^* = (u^*, v^*) \in \boldsymbol{H}^1(\Omega)$, has been considered in order to test the orthogonal decomposition, as explained in the introduction of the Section 4.3. Four cases (a, b, c, d) have been investigated, changing the boundary conditions and the vector field. Every vector field can be decomposed by considering

$$\boldsymbol{u}^* = \boldsymbol{u}_0 + \boldsymbol{u}_0^{\perp}\,, \tag{4.15}$$

where $\boldsymbol{u}_0 \in \boldsymbol{V}^0$ and $\boldsymbol{u}_0^{\perp} \in \boldsymbol{V}^{0,\perp}$. Two different $\boldsymbol{u}_0$ vectors have been considered, one for cases a and b, and another one for cases c and d. Specifically, we have

$$\boldsymbol{u}_0^{a,b} = \begin{bmatrix} \pi \sin^2(\pi x) \sin\left(\pi \dfrac{y}{2}\right) \cos\left(\pi \dfrac{y}{2}\right) \\ -2\pi \sin(\pi x) \cos(\pi x) \sin^2\left(\pi \dfrac{y}{2}\right) \end{bmatrix}, \tag{4.16}$$

and

$$\boldsymbol{u}_0^{c,d} = \begin{bmatrix} \pi \sin^2\left(\pi\dfrac{x}{2}\right) \sin\left(\pi\dfrac{y}{2}\right) \cos\left(\pi\dfrac{y}{2}\right) \\ \pi \sin\left(\pi\dfrac{x}{2}\right) \cos\left(\pi\dfrac{x}{2}\right) \sin^2\left(\pi\dfrac{y}{2}\right) \end{bmatrix} . \tag{4.17}$$

Regarding the value of $\boldsymbol{u}_0^{\perp}$ the same expression has been considered for the cases a, b and c, while for the case d we fix $\boldsymbol{u}_0^{\perp} = 0$. Specifically, $\boldsymbol{u}_0^{\perp}$ is described with a potential function that is representative of the pressure gradient $\boldsymbol{\nabla} p_0$. In particular, we consider a pressure field described as

$$p_0 = \frac{1}{4} xy(x-1)(y-2)\,, \tag{4.18}$$

from which it is easy to obtain the vector gradient which reads

$$\boldsymbol{\nabla} p_0 = \begin{bmatrix} \left(x - \dfrac{1}{2}\right)\left(\dfrac{y(y-2)}{2}\right) \\ (y-1)\left(\dfrac{x(x-2)}{2}\right) \end{bmatrix} . \tag{4.19}$$

Hence, these expressions are the components of the vector $\boldsymbol{u}_0^{\perp}$. Therefore, to identify the vector $\boldsymbol{u}^*$ that has been considered for every simulation, it is sufficient to add $\boldsymbol{u}_0$ and $\boldsymbol{u}_0^{\perp}$. As an example, we report the full expression of $\boldsymbol{u}^*$ for the case a

$$\boldsymbol{u}^* = \begin{bmatrix} u^* \\ v^* \end{bmatrix} = \begin{bmatrix} \pi \sin^2(\pi x) \sin\left(\pi\dfrac{y}{2}\right) \cos\left(\pi\dfrac{y}{2}\right) + \left(x - \dfrac{1}{2}\right)\left(\dfrac{y(y-2)}{2}\right) \\ -2\pi \sin(\pi x)\cos(\pi x) \sin^2\left(\pi\dfrac{y}{2}\right) + (y-1)\left(\dfrac{x(x-2)}{2}\right) \end{bmatrix} . \tag{4.20}$$

The boundary conditions imposed for the four cases are reported in Table 4.2. Note that for these simple tests, only two kinds of boundary conditions have been investigated, i.e. a non-homogeneous Dirichlet boundary condition for the velocity field and a homogeneous Dirichlet boundary condition for the pressure. In fact, since the pressure field is defined up to an additive constant, we recall that it is necessary to fix the value of the pressure at least at one point of the domain, which can also be done by fixing the pressure value on an entire boundary or on a single point.

The numerical pair solution $(\boldsymbol{u}_h, p_h)$ has been computed by solving the system (3.25), and by exploiting a standard multigrid technique for the treatment of the refined grids.

case	Γ_i	Γ_o	Γ_w	Γ_w
a	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$
b	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$
c	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$
d	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\|_\Gamma$

Table 4.2: Boundary conditions for the four different cases of the velocity field decomposition in a bidimensional plane channel geometry.

In the following Tables 4.3, 4.4, 4.5, 4.6 the velocity and pressure error norms are reported, including the convergence rate. The L^2 error norms are shown changing the mesh refinement, from two up to eight levels, corresponding to a number of elements varying from 256 up to 1048576.

In particular, the convergence error at the level l, which is reported in the fifth column, can be defined as the sum of the L^2 norm of velocity and pressure error, i.e.

$$\varepsilon_l = ||\boldsymbol{u}_0 - \boldsymbol{u}_h||_{L^2(\Omega)} + ||p_0 - p_h||_{L^2(\Omega)} \,. \tag{4.21}$$

By taking the logarithm of the ratio between the errors at two different levels, i.e. ε_l and ε_{l+1} (sixth column) we obtain an estimate of the convergence order p, which is reported in the last column of the tables.

l	n_{el}	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	$\|p - p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	2.56E+02	0.4850	0.0093	0.4943	1.9792	0.9849
3	1.02E+03	0.2435	0.0040	0.2476	1.9965	0.9975
4	4.10E+03	0.1217	0.0019	0.1236	2.0031	1.0023
5	1.64E+04	0.0609	0.0010	0.0619	1.9977	0.9984
6	6.55E+04	0.0305	0.0005	0.0309	1.9996	0.9997
7	2.62E+05	0.0152	0.0002	0.0155	2.0000	1.0000
8	1.05E+06	0.0076	0.0001	0.0077	1.9990	0.9993

Table 4.3: Plane channel geometry - case a: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of elements n_{el}.

l	n_{el}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	2.56E+02	0.4854	0.0091	0.4945	1.9833	0.9879
3	1.02E+03	0.2435	0.0041	0.2477	1.9969	0.9978
4	4.10E+03	0.1219	0.0020	0.1238	2.0000	1.0000
5	1.64E+04	0.0610	0.0010	0.0619	2.0001	1.0001
6	6.55E+04	0.0305	0.0005	0.0310	2.0000	1.0000
7	2.62E+05	0.0152	0.0002	0.0155	2.0000	1.0000
8	1.05E+06	0.0076	0.0001	0.0077	2.0000	1.0000

Table 4.4: Plane channel geometry - case *b*: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of elements n_{el}.

l	n_{el}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	2.56E+02	0.2431	0.0091	0.2522	1.9951	0.9965
3	1.02E+03	0.1218	0.0041	0.1259	2.0029	1.0021
4	4.10E+03	0.0609	0.0019	0.0629	2.0021	1.0015
5	1.64E+04	0.0305	0.0010	0.0314	2.0006	1.0005
6	6.55E+04	0.0152	0.0005	0.0157	2.0002	1.0001
7	2.62E+05	0.0076	0.0002	0.0079	2.0000	1.0000
8	1.05E+06	0.0038	0.0001	0.0039	2.0000	1.0000

Table 4.5: Plane channel geometry - case *c*: velocity and pressure error norm and convergence rate for different levels l of grid refinement and the corresponding number of elements n_{el}.

We can notice a convergence rate order equal to one for every case, showing therefore a reliable numerical solution for the orthogonal decomposition considered. In fact, we know that the theoretical error norms satisfy the following inequalities

$$\|\boldsymbol{u}_0-\boldsymbol{u}_h\|_0 \leq Ch\|\boldsymbol{u}^*\|_1\,, \quad \|p_0-p_h\|_0 \leq Ch\|p_0\|_1\,. \tag{4.22}$$

Hence, the results confirm how the order convergence rate follows a linear

l	n_{el}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	2.56E+02	0.2431	0.009623	0.2527	2.0098	1.0070
3	1.02E+03	0.1218	0.002652	0.1245	2.0302	1.0216
4	4.10E+03	0.0609	0.000683	0.0616	2.0197	1.0141
5	1.64E+04	0.0305	0.000172	0.0306	2.0107	1.0077
6	6.55E+04	0.0152	0.000043	0.0153	2.0056	1.0040
7	2.62E+05	0.0076	0.000011	0.0076	2.0028	1.0020
8	1.05E+06	0.0038	0.000003	0.0038	2.0014	1.0010

Table 4.6: Plane channel geometry - case d: velocity and pressure error norm and convergence rate for different levels l of grid refinement and the corresponding number of elements n_{el}.

behavior considering the mesh size h. In Figure 4.7, the velocity and pressure norms trends for the four cases are shown including a comparison with a linear and a quadratic behavior with respect to the mesh size h, as a function of the number of points n_p of the grid. The error norms are represented by marker, circular for the velocity and triangular for the pressure, while solid and dashed lines represent the linear and the quadratic behavior respectively. As already noted in the previous tables, it can be noticed a good agreement with the linear trend for the velocity error norm, for every considered case. Concerning the pressure error norm, we notice the same linear trend for the first three cases. On the other hand, a quadratic trend seems to appear for the fourth case. This behavior can be justified by the fact that the case d is the only case in which we have an orthogonal decomposition that does not consider $\boldsymbol{u}^{\perp}$, and therefore the function $\boldsymbol{u}^*$ is defined only by the divergence-free part without the gradient of a scalar function.

In Figure 4.8 the components of the velocity are reported for different mesh refinements in order to verify the grid convergence. In particular, on the left, the u component is shown for the case a considering a plot over the x coordinates located at $y = 0.5$. Regarding the figure on the right, the v component is shown for the case c considering a plot over the y coordinates located at $x = 0.5$.

By using circular markers, the analytical velocity is represented, while the results for different levels of refinement are represented with four different

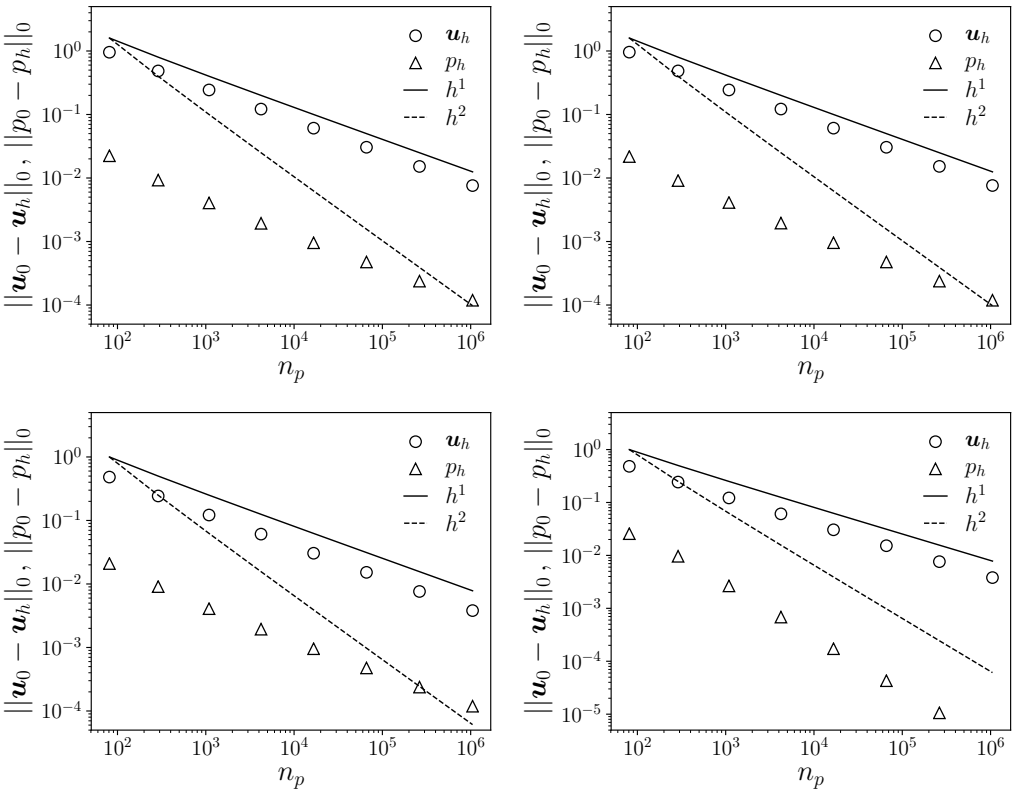

Figure 4.7: Velocity and pressure error norms for the channel geometry: case a (top left), case b (top right), case c (bottom left) and case d (bottom right).

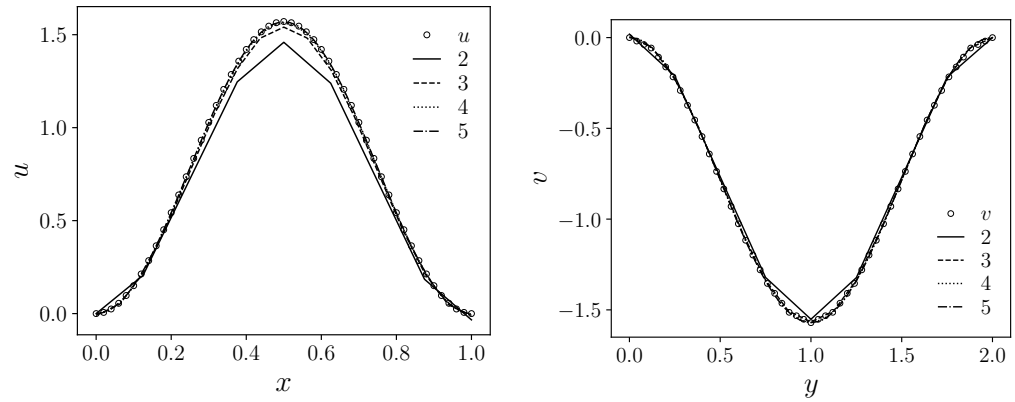

Figure 4.8: Grid convergence for the velocity components: u for case a on the left and v for case c on the right. Circular markers represent the analytical velocity solution.

types of lines. In fact, since a good approximation result has been obtained with few refinements, only four numerical results have been shown, starting from two up to five levels. We recall that the analytical velocity solution represented by the circular white markers coincides with the theoretical $\boldsymbol{u}_0$ defined in (4.16)-(4.17), since our objective was to decompose the analytical velocity $\boldsymbol{u}^*$ in the divergence-free part $\boldsymbol{u}_0$ and its orthogonal part (the gradient of a scalar, i.e. the pressure p).

Circular geometry

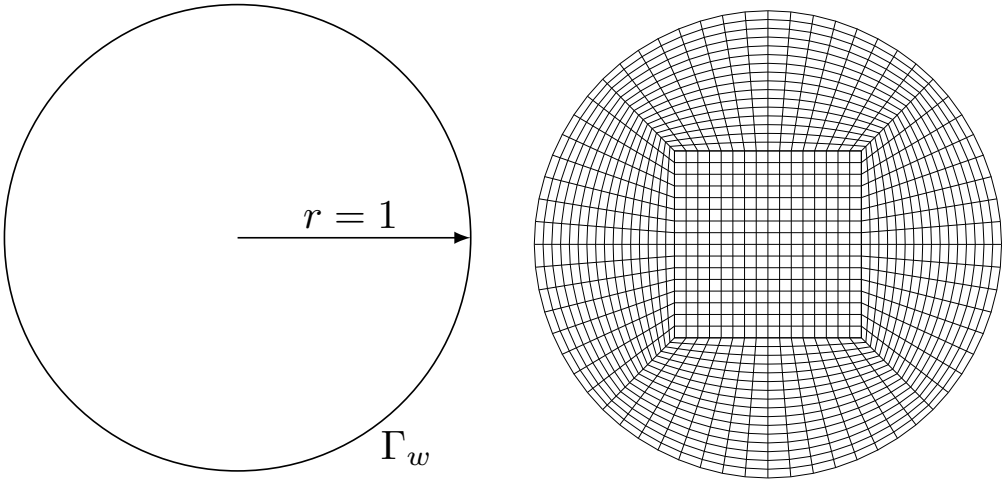

Figure 4.9: On the left circular geometry for the bidimensional velocity decomposition test, on the right corresponding mesh for the two levels of refinement.

In this test, we consider a circle as the computational domain, as shown in Figure 4.9. Hence, a circular domain has been considered with a radius $r = 1$, and with the geometry boundary defined only by Γ_w. For a circular domain, the quadrilateral elements are not equipped with four rectangle corners. In fact, these elements are equipped with only one pair of opposite parallel edges, and therefore they can be denoted as non-affine elements. On the other hand, the grid refinement has been performed on the same mesh, leading to asymptotically affine elements.

Two cases have been considered (a, b), changing the boundary conditions and the analytical vector function. Also in this case, the decomposition for

the vector field $\boldsymbol{u}^* = (u^*, v^*)$ can be described as

$$\boldsymbol{u}^* = \boldsymbol{u}_0 + \boldsymbol{u}_0^{\perp} \,, \tag{4.23}$$

where $\boldsymbol{u}_0 \in \boldsymbol{V}$ and $\boldsymbol{u}_0^{\perp} \in \boldsymbol{V}^{\perp}$. The same $\boldsymbol{u}_0$ function has been considered for the two cases, which reads

$$\boldsymbol{u}_0 = \begin{bmatrix} 4(x^2 + y^2 - 1)y \\ \\ -4(x^2 + y^2 - 1)x \end{bmatrix} \,, \tag{4.24}$$

while the orthogonal term, $\boldsymbol{u}_0^{\perp}$, is present only for the second case b. We have already seen that the orthogonal part is represented by the gradient of a potential function that is the pressure p_0. Hence, for the case b we have

$$p_0 = \frac{(x+1)^3}{3} \,, \tag{4.25}$$

from which the gradient vector reads as

$$\boldsymbol{\nabla} p_0 = \begin{bmatrix} (x+1)^2 \\ 0 \end{bmatrix} \,. \tag{4.26}$$

Therefore for the second test case, we have that $\boldsymbol{u}_0^{\perp} = ((x+1)^2, 0)$. Regarding the boundary conditions, for the case a we impose a vanishing pressure on the boundary, i.e. $p = 0$ on Γ_w. If for the case a, a homogeneous Dirichlet boundary condition has been imposed, for the case b we impose a non-homogeneous Dirichlet boundary condition. Specifically, we have $\boldsymbol{u}|_{\Gamma} = \boldsymbol{u}^*|_{\Gamma}$. Moreover, the pressure has been set equal to 0 at the point $\boldsymbol{x} = (-1, 0)$ to fix the constant value. The system (3.25) is solved considering $\boldsymbol{\mathcal{RT}}_0 \times S_h^0$ finite elements for the pair solution $(\boldsymbol{u}_h, p_h)$.

In the following Tables, 4.7 and 4.8, the velocity and pressure error norms are reported, including the convergence rate. The L^2 error norms are shown changing the mesh refinement, from two up to eight levels, corresponding to a grid size varying from 320 up to 43680 elements. The convergence rate order p is defined as in the channel test, and the same considerations can be done for the values in the table columns.

From the error Tables, 4.7 and 4.8, a perfect agreement with the value 1 is found for the order p, in both cases. The same error norms are depicted in Figure 4.10, where a comparison with a linear and a quadratic trend is

l	n_{nodes}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
1	1.31E+03	0.4373	0.01037	0.4477	-	-
2	5.19E+03	0.2176	0.00317	0.2208	2.0277	1.0198
3	2.06E+04	0.1087	0.00087	0.1095	2.0158	1.0113
4	8.22E+04	0.0543	0.00023	0.0545	2.0082	1.0059
5	3.28E+05	0.0272	0.00006	0.0272	2.0042	1.0030
6	1.31E+06	0.0136	0.00002	0.0136	2.0017	1.0012

Table 4.7: Circular geometry - case a: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of nodes n_{nodes}.

l	n_{nodes}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
1	1.31E+03	0.4373	0.0873	0.5246	-	-
2	5.19E+03	0.2176	0.0437	0.2613	2.0075	1.0054
3	2.06E+04	0.1087	0.0219	0.1305	2.0021	1.0015
4	8.22E+04	0.0543	0.0109	0.0652	2.0005	1.0004
5	3.28E+05	0.0272	0.0055	0.0326	2.0001	1.0001
6	1.31E+06	0.0136	0.0027	0.0163	2.0000	1.0000

Table 4.8: Circular geometry - case b: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of nodes n_{nodes}.

performed. The employed symbols are the same as the plane channel test. Naturally, n_p and n_{nodes} have the same meaning.

Also in this case, the simulation with a vanishing pressure term (case a), presents a quadratic behavior for the pressure error norm. The velocity error norms follow a linear trend considering the mesh size, confirming the theoretical error estimate.

In Figure 4.11, the u component of the velocity is shown as a function of the y coordinates, considering a fixed x equal to 0, for two different levels of grid refinement and for the case a. In particular, the grid convergence is

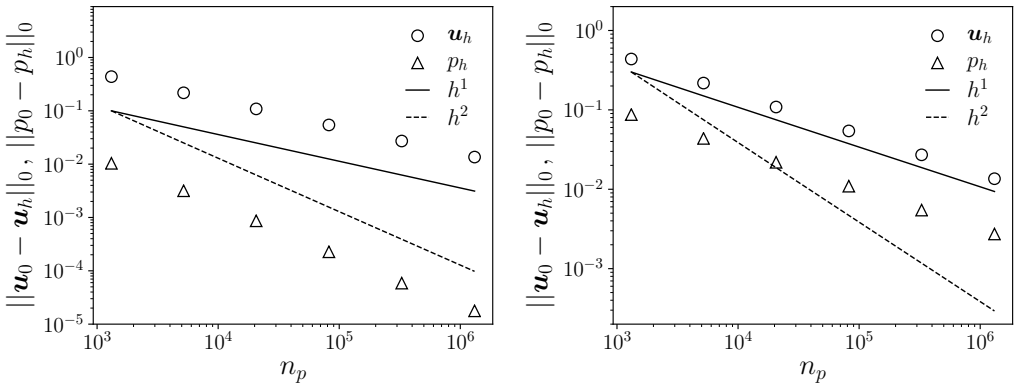

Figure 4.10: Convergence rate for velocity and pressure error for the circular geometry: case a on the left, and case b on the right.

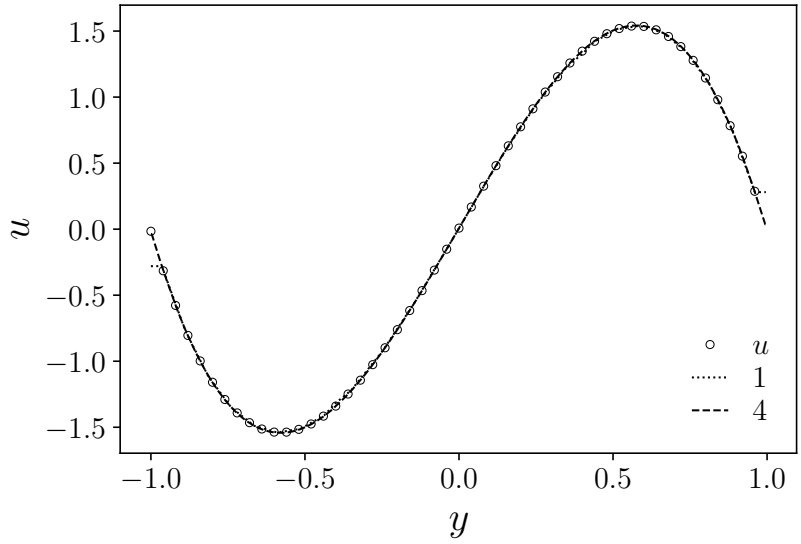

Figure 4.11: Grid convergence for the velocity component u, considering 1 and 4 levels of refinement for the case a.

shown with a comparison with the analytical velocity, which is represented using circular markers. Only two levels are reported since the mesh with zero refinement is already in good agreement with the analytical result.

Note that, for the case with no levels of refinement the solution seems to be different from zero at the boundary, despite homogeneous Dirichlet boundary conditions have been imposed. This issue is related only to the post-processing of the solution. In fact, we recall that the degrees of freedom

of the system of equations correspond to the four fluxes through the element faces. Therefore, in order to represent these fluxes on the grid nodes we represent the velocity field as a piecewise constant solution over the cell. This value is computed as the mean of the fluxes, where naturally the velocity is obtained from the flux by applying the normalization by using the length of the element edges. By increasing the grid refinement this small discrepancy from zero in the elements close to the boundary tends naturally to zero.

4.3.2 Three-dimensional problems

Channel geometry

We now consider a three-dimensional test, which has been carried out on the domain shown in Figure 4.12. The characteristic dimensions of the channel are $L_x = 1$, $L_y = 1$ and $L_z = 4$. Moreover, the geometry boundaries are defined with Γ, where specifically we have that the inlet and the outlet are represented by Γ_i and Γ_o, the top and bottom faces with Γ_t and Γ_b, while the lateral faces are Γ_l and Γ_r. Standard hexahedral elements have been employed for the geometrical discretization, avoiding irregularity related to surface distortion and hence non-affine properties.

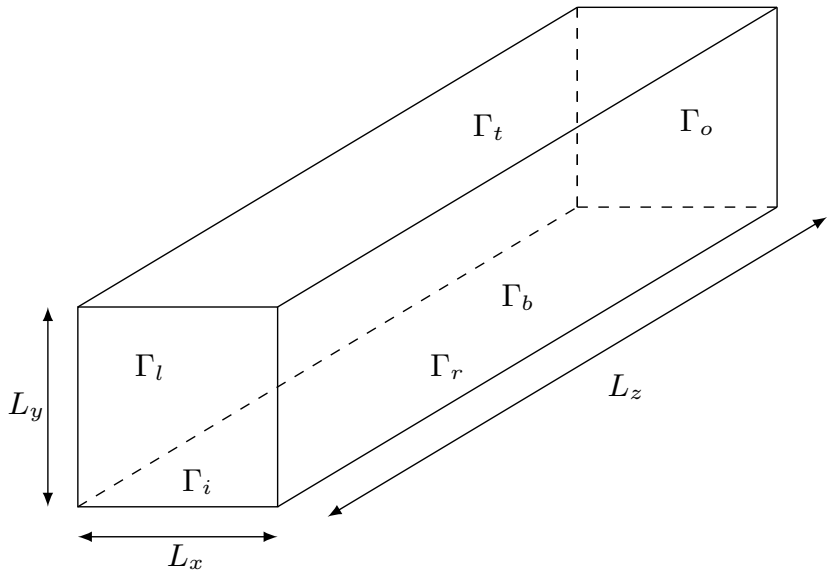

Figure 4.12: Three-dimensional channel geometry for the three-dimensional velocity decomposition test.

This test is similar to the circle test since only two cases (a, b) have been investigated, characterized by the same $\boldsymbol{u}_0$. In this case, we have

$$\boldsymbol{u}_0 = \begin{bmatrix} 2\pi \sin^2(\pi x) \sin(\pi y) \cos(\pi y) z(4-z) \\ -\pi \sin(\pi y) \sin(2\pi x) z(4-z) + \dfrac{\pi}{4} \sin^2(\pi y) \sin\left(\dfrac{\pi z}{2}\right) x(1-x) \\ -4\pi \sin^2(\pi z/4) \sin(\pi y) \cos(\pi y) x(1-x) \end{bmatrix}, \tag{4.27}$$

where naturally the decomposition of the vector $\boldsymbol{u}^*$ follows the same path as the previous tests. The only difference between the two cases relies on the $\boldsymbol{u}^\perp$ term, which is present only in the case b. Specifically, we have a pressure field defined as

$$p_0 = 25(x^2 - x)(y^2 - y)\,, \tag{4.28}$$

from which the gradient vector reads as

$$\boldsymbol{\nabla} p_0 = \begin{bmatrix} 50(x - 0.5)(y^2 - y) \\ 50(x^2 - x)(y - 0.5) \\ 0 \end{bmatrix}. \tag{4.29}$$

The set of equations in (3.25) is employed to solve $(\boldsymbol{u}_h, p_h) \in \boldsymbol{\mathcal{RT}}_0 \times S_h^0$. Concerning the boundary conditions, we impose a non-homogeneous Dirichlet boundary condition for the velocity on every boundary, while a homogeneous Dirichlet boundary condition for the pressure has been imposed on Γ_o. The boundary conditions are summarized in Table 4.9.

Γ_i	Γ_o	Γ_l	Γ_r	Γ_t	Γ_b
$\boldsymbol{u} = \boldsymbol{u}^*\vert_\Gamma$	$p = 0$	$\boldsymbol{u} = \boldsymbol{u}^*\vert_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\vert_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\vert_\Gamma$	$\boldsymbol{u} = \boldsymbol{u}^*\vert_\Gamma$

Table 4.9: Boundary conditions for the three-dimensional channel test.

In addition, the constant for the pressure has been fixed, considering $p_0 = p_h = 0$ at the point $\boldsymbol{x} = (0, 0, 0)$. With the same notation of previous tests, in Tables 4.10 the error norms are reported for the velocity and pressure field, as a function of the refinement levels. For this test, the grid size of the mesh changes from 512 (first level), up to $2.09 \cdot 10^6$ (fifth level) elements. The computed convergence rate order p is very close to 1, confirming the goodness of the numerical results.

l	n_{nodes}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	4.91E+03	0.3818	0.00153	0.3834	1.9433	0.9585
3	3.59E+04	0.1924	0.00041	0.1929	1.9878	0.9912
4	2.75E+05	0.0964	0.00010	0.0965	1.9982	0.9987
5	2.15E+06	0.0482	0.00002	0.0483	2.0000	1.0000

Table 4.10: 3-dimensional channel geometry - case a: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of nodes n_{nodes}.

l	n_{nodes}	$\|\boldsymbol{u}-\boldsymbol{u}_h\|_0$	$\|p-p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
2	4.91E+03	0.3818	0.1551	0.5369	2.0689	1.0489
3	3.59E+04	0.1924	0.0705	0.2629	2.0419	1.0299
4	2.75E+05	0.0964	0.0341	0.1305	2.0151	1.0109
5	2.15E+06	0.0482	0.0169	0.0651	2.0044	1.0032

Table 4.11: 3-dimensional channel geometry - case b: velocity and pressure error norm and convergence rate for different levels l of grid refinement and corresponding number of nodes n_{nodes}.

In Figure 4.13, the trend of the error norms is shown in comparison with a linear and a quadratic behavior with respect to the mesh size h. The employed symbol and the conclusions are similar to the previous cases: if for the velocity error norm a linear trend can be noted, for the case a the pressure error norm seems to follow a quadratic behavior. We recall that the case a is characterized by a vanishing $\boldsymbol{u}^{\perp}$ term.

Cylindrical geometry

The last test focuses on a cylindrical geometry, as shown in Figure 4.14. In particular, the considered cylinder is described by the radius r and the height H, for which the values are reported in the Figure. Naturally, the cylinder axis coincides with the z-axis. The boundaries are denoted with Γ_i and Γ_o for the inlet and outlet section, while the side wall is denoted with Γ_w. For this

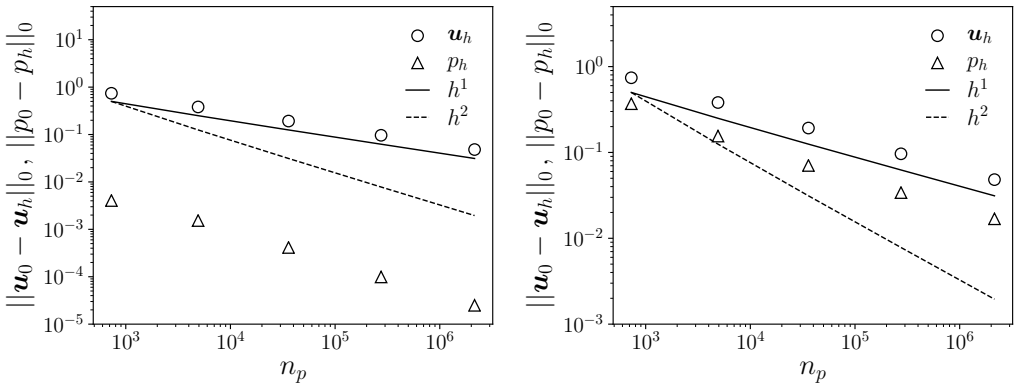

Figure 4.13: Convergence rate for velocity and pressure error for the 3-dimensional channel geometry: case *a* on the left, and case *b* on the right.

geometry, the mesh discretization is based on the employing of hexahedral elements.

It is easy to understand that, in order to represent a curved surface, the hexahedral elements are equipped with a pair of opposite faces that are not parallel to each other. On the other hand, since the refinement is performed on the same mesh, the faces become more and more parallel to each other, avoiding any problematic convergence issues.

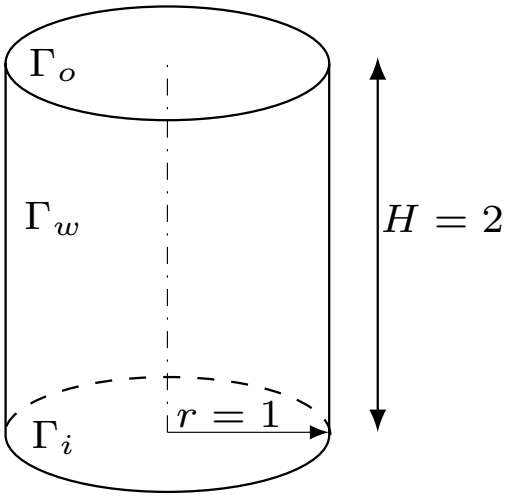

Figure 4.14: Cylindrical geometry for the three-dimensional velocity decomposition test.

A single case has been investigated, where the tested vector function is

described only with the divergence-free term, i.e. $\boldsymbol{u}^* = \boldsymbol{u}_0$. Therefore, the potential function, i.e. the pressure field, is equal to zero. In particular, we have

$$\boldsymbol{u}_0 = \begin{bmatrix} \pi \cos^2\left(\pi\frac{x}{2}\right) \sin\left(\pi\frac{y}{2}\right) \cos\left(\pi\frac{y}{2}\right) \\ -\pi \cos^2\left(\pi\frac{y}{2}\right) \sin\left(\pi\frac{x}{2}\right) \cos\left(\pi\frac{x}{2}\right) \\ 0 \end{bmatrix}. \tag{4.30}$$

Regarding the boundary conditions, the pressure has been fixed equal to zero on the outlet section, while on the remaining boundaries, a non-homogenous boundary condition for the velocity has been imposed, $\boldsymbol{u} = \boldsymbol{u}^*|_\Gamma$. In Table 4.12 the error norms for the velocity and pressure field are reported. In this case, the mesh size changes from the first level equal to $1.08 \cdot 10^3$ elements up to the fourth level with $5.53 \cdot 10^5$ elements.

l	n_{nodes}	$\|\boldsymbol{u} - \boldsymbol{u}_h\|_0$	$\|p - p_h\|_0$	ε_l	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
1	1.31E+03	0.8303	0.0044	0.8557	-	-
2	9.69E+03	0.4357	0.0015	0.4371	1.9575	0.9690
3	7.32E+04	0.2193	0.0004	0.2197	1.9895	0.9924
4	5.69E+05	0.1099	0.0001	0.1100	1.9980	0.9985

Table 4.12: Cylindrical geometry: velocity and pressure error norm and convergence rate for different levels l of grid refinement and the corresponding number of elements n_{nodes}.

The convergence rate order p tends to 1 with increasing the mesh refinement, confirming the validity of the numerical results. In Figure 4.15 the error norm values are shown in comparison with a linear and a quadratic trend considering the mesh size h. Despite the few points, it can be noticed a difference between the velocity and pressure error norm. In fact, as expected, the velocity error norm follows a linear behavior, while for the pressure a quadratic trend can be observed.

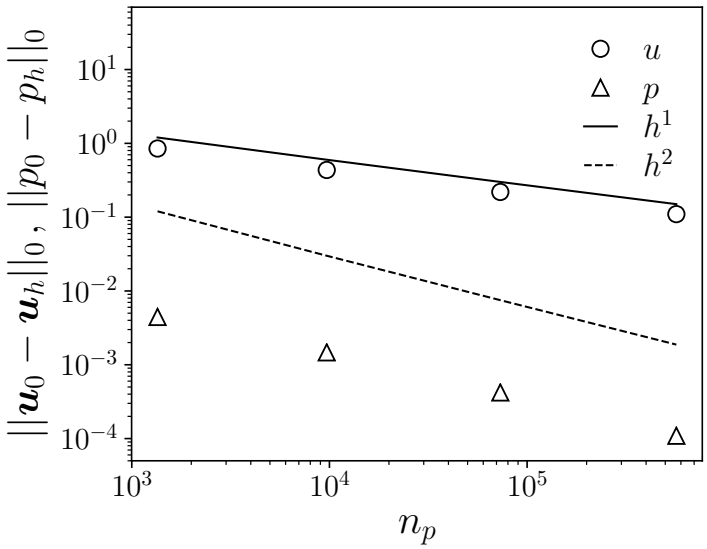

Figure 4.15: Convergence rate for velocity and pressure error for the cylindrical geometry.

4.4 Projection method for Navier-Stokes equations

In this section, we aim to test the employment of the Raviart-Thomas finite element family in the framework of the projection method for the resolution of the Navier-Stokes equation. Specifically, the system of equations (3.25), has been adopted where we recall that the orthogonal decomposition of the velocity field is considered as the second step of the split algorithm, i.e. the resolution of the pressure equation.

Three cases have been investigated, considering a bidimensional and a three-dimensional geometry: a channel, both in two and three dimensions and a bidimensional cavity. For these tests, the same comparison has been performed, considering three different algorithms for the numerical solution. The first one consists of a standard coupled algorithm for the resolution of the velocity and pressure, and by using classical lagrangian finite elements $\mathcal{P}_2 - \mathcal{P}_1$ for the field discretization. The second technique is based on the same finite element family but employs a standard projection algorithm for the Navier-Stokes system. Lastly, the third method is characterized by the employment of the Raviart-Thomas finite element family for the resolution

of the pressure equation in the context of the split technique.

4.4.1 Two-dimensional channel

The first analyzed case focuses on the resolution of the Navier-Stokes equation characterized by a low Reynolds number inside a bidimensional channel, i.e. a laminar flow. It is well known that the solution of this kind of configuration is expressed by the Poiseuille profile for the streamwise component of the velocity. Indeed, we expect to obtain the classical parabolic profile.

In Figure 4.16 the investigated mesh discretizations of the channel geometry are reported. Note that, on the right, is reported a channel discretization where non-affine elements have been employed. On the other hand, the multigrid refinement produces elements that converge to a parallelogram shape, which is an affine element since the opposite edges are parallel. The geometry dimensions and the boundaries are the same as the case described in Figure 4.6.

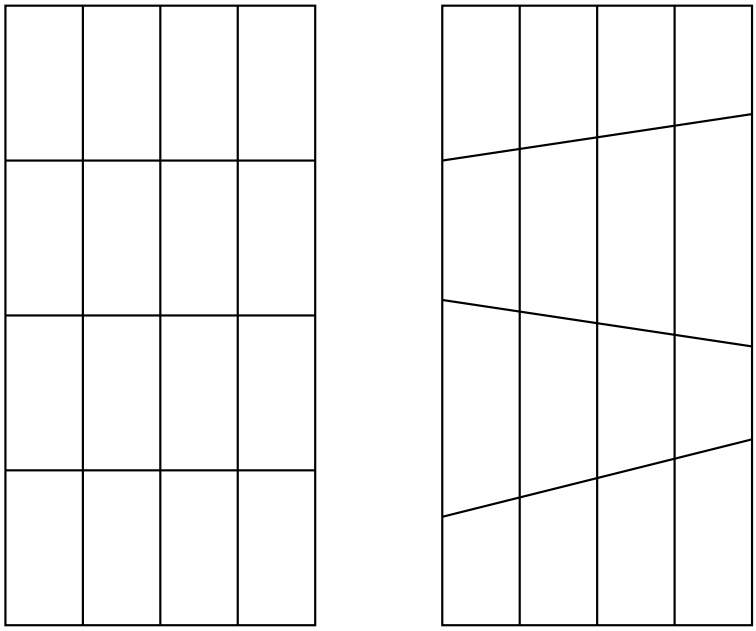

Figure 4.16: Two-dimensional Poiseuille flow: regular and irregular coarsest mesh for the channel flow.

Considering the boundary conditions, at the inlet section Γ_i a fixed velocity has been imposed, a standard no-slip boundary condition has been

imposed on the wall side Γ_w, while an outlet-type condition has been imposed at the outlet section Γ_0 to fix the pressure value. In Table 4.13 the boundary conditions are summarized.

Γ_i	Γ_o	Γ_w
$\boldsymbol{u} = \boldsymbol{u}_0$	$p = 0$	$\boldsymbol{u} = 0$

Table 4.13: Boundary conditions for the laminar channel test.

Otherwise the previous examples, for this kind of numerical simulation it is not possible to have an analytical solution of the velocity field. In fact, even though a Poiseuille flow type is searched, if we consider an inlet boundary condition the analytical solution for the Navier-Stokes system does not exist. For this reason, the velocity error norm with respect to a reference solution does not have a numerical significance. On the other hand, in order to compare qualitatively the solutions with the three types of numerical discretizations the L^2 norm of the velocity field over the entire domain is reported, for different grid refinements. In particular, for both regular and irregular meshes, we compute the velocity norm as reported in Table 4.14. The different methods are denoted with $\mathcal{P}^c$, $\mathcal{P}^s$ respectively, for the coupled and split technique with Taylor-Hood finite elements, while with $\mathcal{RT}_0$ the Raviart-Thomas finite elements.

l	n_{el}	Regular mesh			Irregular mesh		
		$\mathcal{P}^c$	$\mathcal{P}^s$	$\mathcal{RT}_0$	$\mathcal{P}^c$	$\mathcal{P}^s$	$\mathcal{RT}_0$
1	6.40E+01	1.551	1.408	1.515	1.516	1.408	1.526
2	2.56E+02	1.544	1.470	1.530	1.525	1.470	1.531
3	1.02E+03	1.539	1.502	1.532	1.530	1.501	1.532
4	4.10E+03	1.536	1.517	1.532	1.532	1.517	1.532
5	1.64E+04	1.535	1.525	1.532	1.533	1.525	1.531
6	6.55E+04	1.534	1.529	1.530	1.533	1.529	1.528

Table 4.14: Two-dimensional channel test: L^2-norm of the velocity field for different levels l of grid refinement with n_{el} number of elements, for regular and irregular mesh.

We can notice that with the increasing of the grid refinement, the veloc-

ity norm values tend to the same value for every method, ensuring reliable numerical results for this test. Consider that, even if the second mesh is characterized by non-affine quadrilateral elements, the numerical solution is sought considering a multigrid approach. Therefore, the initial irregular mesh tends to asymptotically affine quadrilateral elements, allowing the use of the standard Raviart-Thomas family of order 0. The last column of Table 4.14 confirms that, with this geometry, this spatial discretization is able to find the right numerical solution.

Naturally, different numerical aspects should be considered in order to perform an extensive comparison between the three methods. In fact, from a computational point of view, the simulation duration should be discussed, since with a split technique a decrease in the computational effort is sought. This kind of analysis will be addressed in future works.

In addition, considering the framework of a laminar flow inside a channel, the main variable of interest is the streamwise component of the velocity field, denoted with v. For this reason, in Figure 4.17 the velocity profile of v is reported for the three different algorithms, as a function of the x coordinates. The plot has been done, considering a fixed y equal to 1. Since the numerical solutions between the two different grids are very similar, we report only the case with regular quadrilateral elements.

Considering Figure 4.17, the type of algorithm to solve the Navier-Stokes system is denoted with the subscript, c for the coupled system and s for the split one. With the superscript is denoted the type of finite element family employed for the velocity-pressure discretization, $\mathcal{P}$ for standard $\mathcal{P}_2 - \mathcal{P}_1$ lagrangian elements and with $\mathcal{RT}$ the Raviart-Thomas elements. The solid line represents the reference numerical result, that is, the numerical solution obtained with a coupled algorithm and classical Taylor-Hood lagrangian elements. With the markers, the numerical solutions of the split system are reported: the triangular markers represent the case where $\mathcal{P}_2 - \mathcal{P}_1$ elements have been employed in every equation, while the circular markers represent the solution obtained by using an $\mathcal{RT}$ approximation for the resolution of the pressure equation. We can notice a good agreement of the velocity profiles obtained by employing the three different algorithms.

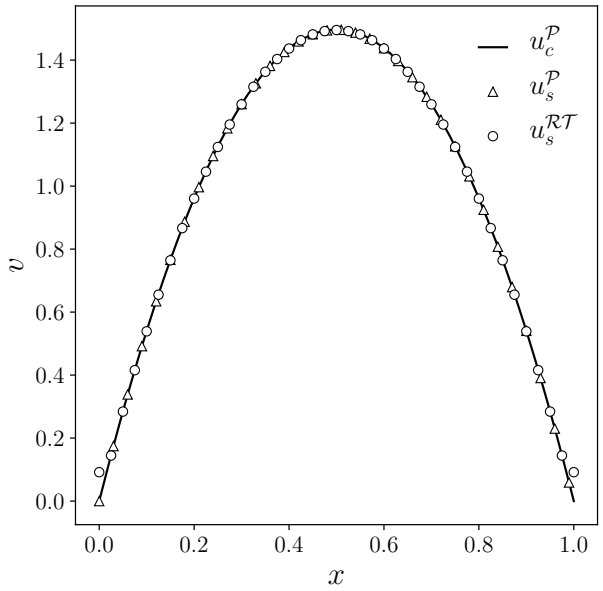

Figure 4.17: Two-dimensional channel flow (regular mesh): streamwise velocity component v for coupled algorithm, split algorithm, and split algorithm with Raviart-Thomas approximation.

4.4.2 Two-dimensional cavity

The aim of this test is to verify the convergence error rate, by considering a cavity configuration, i.e. a rectangular closed geometry. Specifically, the domain is described with the same channel of the previous cases, i.e. the bidimensional channel described in Figure 4.6, where the elements employed for the spatial discretization are standard regular quadrilateral elements. Regarding the velocity field, the boundary condition imposed on every edge is a homogeneous Dirichlet boundary condition in order to have $\boldsymbol{u} = 0$ on the walls. Naturally, Raviart-Thomas finite elements have been employed for this numerical simulation, in order to discuss the solution goodness in a different setting.

In order to compute the L^2 norm of the velocity error, the steady exact Navier-Stokes solution has been imposed on the right-hand side of the equation. In fact, given a generic operator A which represents the left-hand side terms of the Navier-Stokes equation, we aim to solve

$$A\boldsymbol{u} = A\boldsymbol{u}^* \tag{4.31}$$

for a specified $\boldsymbol{u}^*$, which represents the desired solution. Specifically, the

exact solution for the velocity components reads as

$$\boldsymbol{u}^* = \begin{bmatrix} u^* \\ v^* \end{bmatrix} = \begin{bmatrix} \frac{\pi}{2}\sin^2(\pi x)\sin\left(\pi\frac{y}{2}\right)\cos\left(\pi\frac{y}{2}\right) \\ -\pi\sin^2\left(\pi\frac{y}{2}\right)\sin(\pi x)\cos(\pi x) \end{bmatrix}. \tag{4.32}$$

In Table 4.15 the velocity error norm and the corresponding order of convergence for different levels of refinement are reported, for the $\mathcal{RT}_0$ finite element approximation. The order of convergence p is reported in the last column, computed as in the previous tests. We can notice that also in this case, a linear trend for the velocity error norm can be observed for the Raviart-Thomas velocity approximation.

l	n_{el}	$\mathcal{RT}_0$		
		$\|\boldsymbol{u}^* - \boldsymbol{u}_h\|_0$	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	p
1	6.40E+01	9.59E-01	-	-
2	2.56E+02	4.86E-01	1.976	0.997
3	1.02E+03	2.44E-01	1.994	0.997
4	4.10E+03	1.22E-01	1.998	1.002
5	1.64E+04	6.10E-02	2.000	0.998
6	6.55E+04	3.05E-02	2.000	0.999
7	2.62E+05	1.52E-02	1.999	1.000

Table 4.15: Two-dimensional cavity test with projection method: velocity error norm and convergence rate for different levels l of grid refinement and corresponding number of elements n_{el}, for $\mathcal{RT}_0$ finite element approximation.

The same test has been computed considering also the standard Taylor-Hood lagrangian basis function, both with a coupled and a split algorithm. The results are reported in Table 4.16. We can notice a good convergence trend for the velocity error for both methods, even though the order of convergence p is not reported. In fact, despite these parameters appearing to be equal to 3 for both simulations, we recall that the velocity error should be considered together with the pressure error norm. For coupled systems or for lagrangian-type basis functions, that are not pointwise divergence-free,

we know that an error in the pressure field produces an effect also the error in the velocity. An example of this problem is reported at the beginning of this chapter, in section 4.2.

l	n_{el}	Coupled $\mathcal{P}_2 - \mathcal{P}_1$		Split $\mathcal{P}_2 - \mathcal{P}_1$	
		$\|\boldsymbol{u}^* - \boldsymbol{u}_h\|_0$	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	$\|\boldsymbol{u}^* - \boldsymbol{u}_h\|_0$	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$
1	6.40E+01	4.60E-02	-	4.56E-02	-
2	2.56E+02	5.81E-03	7.925	5.77E-03	7.905
3	1.02E+03	7.25E-04	8.018	7.23E-04	7.982
4	4.10E+03	9.05E-05	8.007	9.06E-05	7.986
5	1.64E+04	1.13E-05	8.002	1.13E-05	8.011
6	6.55E+04	1.43E-06	7.929	1.41E-06	7.995

Table 4.16: Two-dimensional cavity test with Taylor-Hood finite element approximation: velocity error norm for different levels l of grid refinement and corresponding number of elements n_{el}, for coupled and split algorithm.

An interesting result has been found considering the behavior of the velocity divergence. As expected, for the coupled system solved by using $\mathcal{RT}_0$ approximation, the velocity divergence reaches values close to machine precision, indicating that this technique is equipped with an exact zero divergence in every point. Considering the other two methods, the results have been reported in Table 4.17.

It is worth noting that the velocity divergence error is different from zero as expected. On the other hand, these values seem to converge with a quadratic order.

4.4.3 Three-dimensional channel

This test represents the extension of the bidimensional Poiseuille flow to a three-dimensional configuration, i.e. a regular parallelepiped. The considered domain is the same as the three-dimensional test for the orthogonal velocity decomposition (represented in Figure 4.12), where the characteristic lengths have the same values ($L_x = 1$, $L_y = 1$ and $L_z = 4$). Moreover, regular hexahedral elements have been employed for the domain discretization.

l	n_{el}	Coupled $\mathcal{P}_2 - \mathcal{P}_1$		Split $\mathcal{P}_2 - \mathcal{P}_1$	
		$\|\boldsymbol{\nabla} \cdot \boldsymbol{u}^* - \boldsymbol{\nabla} \cdot \boldsymbol{u}_h\|_0$	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$	$\|\boldsymbol{\nabla} \cdot \boldsymbol{u}^* - \boldsymbol{\nabla} \cdot \boldsymbol{u}_h\|_0$	$\frac{\varepsilon_{l-1}}{\varepsilon_l}$
1	6.40E+01	4.33E-01	-	4.42E-01	-
2	2.56E+02	1.13E-01	3.839	1.13E-01	3.92
3	1.02E+03	2.83E-02	3.978	2.83E-02	3.98
4	4.10E+03	7.09E-03	3.996	7.09E-03	4.00
5	1.64E+04	1.77E-03	3.999	1.77E-03	4.00
6	6.55E+04	4.44E-04	3.990	4.43E-04	4.00

Table 4.17: Two-dimensional cavity test with Taylor-Hood finite element approximation: velocity divergence error norm for different levels l of grid refinement and corresponding number of elements n_{el}, for coupled and split algorithm.

The flow configuration follows the same path of the bidimensional case, and thus the fluid enters at the inlet section Γ_i, and exits through the outlet section Γ_o. On the remaining boundaries, the standard no-slip boundary condition has been imposed.

Regarding the reliability of the numerical solutions, the same comments of the two-dimensional channel case can also be drawn for the case of a three-dimensional channel. Therefore, since it is not possible to compare the numerical solution with an analytical field, also in this case the L^2-norm of the velocity field has been computed for the three methods as an indicator of the solution goodness. For this reason, in Table 4.18 the velocity norm values have been reported considering three levels of refinements, for the three techniques. The notation is the same as the two-dimensional channel test.

The L^2-norm of the velocity field tends to the same value for each different algorithm, confirming the good behavior of the numerical solution, as already shown in the bidimensional case.

Also in this case, the same qualitative comparison between three types of algorithms has been performed on the numerical solution. Therefore, in Figure 4.18 the streamwise velocity component w has been reported, where the employed notation and symbols are the same as the bidimensional channel. The w component is represented as a function of the x coordinate, with a

l	n_{el}	$\mathcal{P}^c$	$\mathcal{P}^s$	$\mathcal{RT}_0$
1	2.08E+02	2.147	1.962	2.301
2	4.10E+03	2.236	2.139	2.312
3	3.28E+04	2.280	2.231	2.323

Table 4.18: Three-dimensional channel test: velocity error norm for different levels l of grid refinement and the corresponding number of elements n_{el}.

plot performed with a fixed y coordinate equal to 0.5.

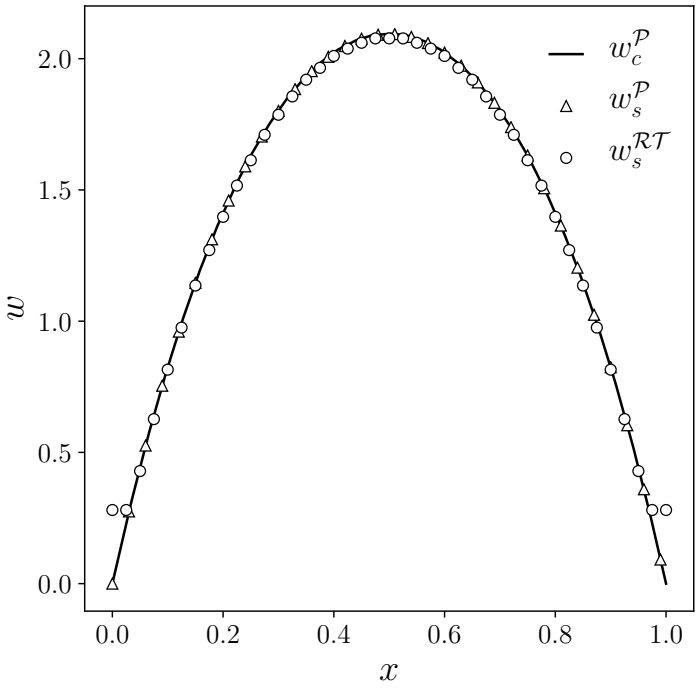

Figure 4.18: Three-dimensional channel flow: streamwise velocity component w for coupled algorithm, split algorithm, and split algorithm with Raviart-Thomas approximation.

Figure 4.18 shows good accuracy for the computed numerical solution. Indeed, the velocity profile w obtained with a Raviart-Thomas approximation of the pressure equation in a split system (3.25) has a similar trend to the velocity profile obtained with a standard coupled (3.17) and split algorithm (3.19).

CHAPTER 5

Code coupling for Multiphase simulations

The objective of this chapter is to exploit theoretical and numerical considerations about the velocity field approximation characterized by the incompressibility constraint. In particular, we try to apply the resolution of the Navier-Stokes equations previously described, in the context of multiphase flow simulations. In fact, it is well known that mass conservation is a key feature when numerical simulations of multiple-phase flows are considered. This feature can be easily understood if we suppose that the density of the considered phases is represented by a constant value. Therefore, the classical constraint of the vanishing velocity divergence is a natural consequence, that must be addressed with the most suitable numerical tools.

In order to exploit the divergence-free velocity approximation, we have to describe the physical application that has been considered. Therefore, this chapter is organized as follows: in the first part we describe the multiphase problem and we discuss how to manage the advection of the interface of one phase; after that, some standard benchmark applications are presented where a comparison between finite element families for the velocity interpolation is performed. In fact, the idea is to take advantage of a discretized divergence-free velocity field in the context of multiphase simulations.

The numerical codes that have been used for this purpose are the finite el-

ement libraries FEMuSTTU [83] and ProXPDE [84]. These codes are numerical libraries built for the resolution of partial differential equations developed at the Department of Industrial Engineering of the University of Bologna and at the Department of Mathematics and Statistics of the Texas Tech University respectively. Indeed, this work has been performed by exploiting the collaboration between the research group of Nuclear Power Plants, at UNIBO, and the one of Numerical Analysis, at TTU. Each of the mentioned codes is written in C++ and is available on the GitHub website.

5.1 Surface advection with marker technique

As mentioned earlier, the subject of the numerical simulations presented in this chapter is related to the multiphase simulations. Additionally, our study is focused on addressing the treatment of the interface between the multiphase phases. Although numerous numerical techniques and algorithms have been developed over the years in the literature in order to address this issue (VOF, Level Set method, etc) [85, 86, 87], our attention is drawn to the Front Tracking Method [88]. Indeed, various approaches have already been investigated considering the marker technique for the surface reconstruction [24, 25]. The basic idea of this technique is to advect a set of points, called markers, which represent the interface, from the initial configuration and with the aim of preserving the topology information of the interface during the simulation transient.

In particular, a numerical library has been developed capable of initializing, advecting, and rebuilding the geometric position of a set of markers to represent the surface with the best-fit quadric equation. These actions are computed in every cell of the domain, where we take into account the information of the neighboring cells for the computation of the different variables in the algorithm, such as the color function and other parameters connected to the markers (curvature, normal vector, etc.). We underline that the presented library has been developed and initially tested only for bidimensional problems, and at this stage we are able to handle only cartesian grids. On the other hand, three-dimensional examples will be investigated in the future.

In the first part of this chapter, the implemented algorithm is described, explaining the several functions necessary for the surface advection. After that, some standard numerical tests are presented with the aim of verifying the ability of the markers to represent a multiphase surface movement.

These tests are performed by applying analytical velocity fields on standard geometric shapes, such as a bubble, i.e. a circumference. Finally, a comparison with two types of different finite element families for the approximation of the specific analytical velocity field is presented.

5.1.1 Marker geometry initialization

The first step of the algorithm involves the implementation of an initialization function, to define the set of the markers and their related parameters. This function is specifically designed to define the initial geometry of the marker cloud, representing one of the two phases within the domain. In order to achieve this, the library is equipped with two possible methods:

i) providing an explicit set of points with related parameters;

ii) providing a generic quadric equation.

Concerning the first method, it is sufficient to provide a complete set of points along with their geometric coordinates, normal vectors and curvature for each marker. This allows the initialization of interface shapes that cannot be represented with a quadric equation.

Alternatively, the second method involves providing the equation of a generic conic equation in the form:

$$f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F\,, \tag{5.1}$$

where the chosen initial geometry of the marker cloud defines the coefficients A, B, C, D, E and F. This function also allows the possibility of initializing multiple geometries based on different chosen quadric equations. Indeed, it is sufficient to change only the coefficients to define another initial configuration. From an implementation perspective, specific functions have been provided for ellipse-type geometry (which can be transformed into a circumference as well). For example, given the center coordinates of the ellipse $\boldsymbol{x}_c$ and defining the semi-axes length with (a, b), the quadric coefficients are defined as

$$\begin{aligned} &A = b^2, && B = 0, && C = a^2, \\ &D = -2x_c b^2, && E = -2y_c a^2, && F = x_c^2 b^2 + y_c^2 a^2 + 1\,. \end{aligned}$$

Having the quadric equation, markers can be inserted onto it by providing the desired number of points for every cell. Therefore, giving n_p markers for

each cell with their corresponding points coordinates $\boldsymbol{x}_p = (x_p, y_p)$, the unit normal vector $\hat{\boldsymbol{n}}$ can be calculated as

$$\hat{\boldsymbol{n}} = \left\langle \frac{n_x}{\sqrt{n_x^2 + n_y^2}}, \frac{n_y}{\sqrt{n_x^2 + n_y^2}} \right\rangle , \tag{5.2}$$

where

$$n_x = 2Ax_p + By_p + D\,, \qquad n_y = 2Cy_p + Bx_p + E\,. \tag{5.3}$$

In addition, the curvature value k is computed for each marker by using the following expression [89]

$$k = \frac{\mathcal{A} - \mathcal{B}}{\mathcal{C}^{3/2}}\,, \tag{5.4}$$

where

$$\mathcal{A} = 8AC^2y_p^2 + 2C\left[(D + 2Ax_p)^2 + 4A(E + Bx_p)y_p - B^2y_p^2\right]\,, \tag{5.5}$$

$$\mathcal{B} = -2(E + Bx_p)\left[-AE + B(D + Ax_p + By_p)\right]\,, \tag{5.6}$$

$$\mathcal{C} = \left[(D + 2Ax_p + By_p)^2 + (E + Bx_p + 2Cy_p)^2\right]\,. \tag{5.7}$$

Note that this formulation for the curvature k derives from the well-known formula [90]

$$k = -\frac{d\hat{\boldsymbol{n}}}{ds} \cdot \hat{\boldsymbol{t}}\,, \tag{5.8}$$

where $\hat{\boldsymbol{t}}$ represents the unit tangent vector and s is the arc length of the curve.

Moreover, the arc length ds associated with every marker of the quadric is considered. To evaluate it, an osculating circle is built on the quadric connecting the arc length of a circular sector to ds. Specifically, we have

$$ds = Rd\theta\,. \tag{5.9}$$

The radius R and the angle $d\theta$ are evaluated with other functions, that are described in the marker rebuilding section.

In addition to the marker initialization, another feature has been provided to identify different phases inside the domain. In particular, at every node of the grid, the sign of the quadric surface is considered. If a cell is equipped with a negative sign at every node that region is regarded as the first phase, and an internal marker is placed at the center of the cell. The advection step is also applied to this new marker inside the cell, in order to manage the color function variable. Otherwise, if the sign of the quadric is positive at every grid node, it indicates the second phase. The third option arises when

a cell has nodes with different signs of the quadric, representing the set of the boundary cells, that are indeed cut by the surface quadric equation. An example of boundary and interior markers is reported in Figure 5.1, showing also the set of normal vectors for the marker present in a single cell. Note that the interior markers are placed only in the cells for which every node is considered inside the reference phase, i.e. the nodes equipped with the same negative sign concerning the quadric equation.

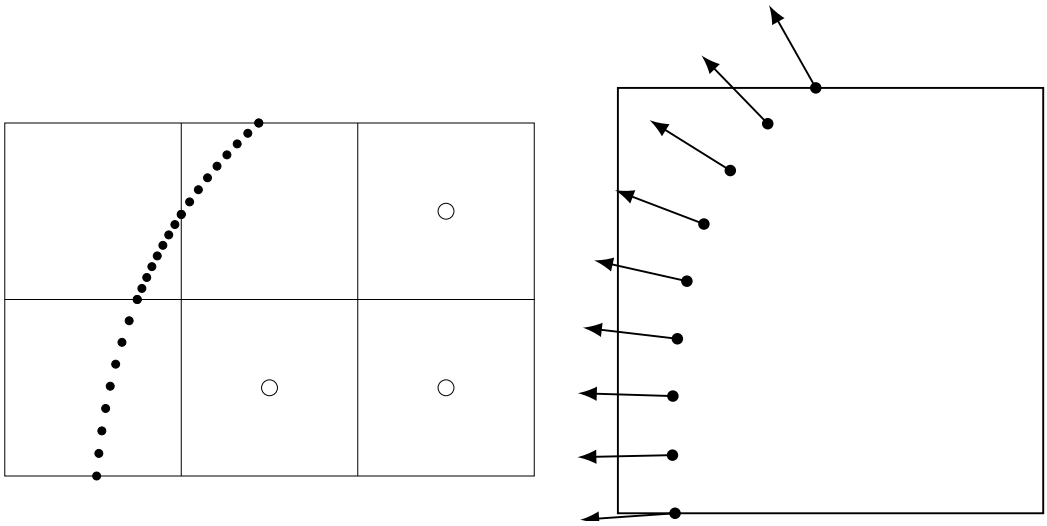

Figure 5.1: On the left, initialization of a few cells with boundary and interior markers. On the right, a generic cell with nine markers and their respective normal vectors for an ellipse-type geometry initialization.

5.1.2 Color function evaluation

In the context of multiphase simulations, the tracking of the two phases is an essential requirement. To address this need, the library includes a function capable of evaluating the value of the color function C for every cell. This variable serves as a tool for monitoring the phase under consideration and represents the ratio between the area occupied by one of the two phases A_c and the total area A of the respective cell. Indeed, the value of C varies between 0 and 1, with the extremes of this interval representing the different phases. To illustrate, considering a bubble (phase 1) inside a domain (phase 2), the interior phase can be associated with $C = 1$, while the exterior phase

is represented by $C = 0$. Therefore, C can be defined as

$$C = \frac{A_c}{A}\,. \tag{5.10}$$

The value of A_c is computed by implementing a linear fit of the interface cell markers, providing the best approximation of these points in the cell with a line. In order to evaluate the area of the interior phase, the linear fit is employed using an exact subdomain polynomial integration [91] implemented for finite elements with linear cuts.

As a result, the value of C is piecewise constant for every cell in the domain. On the other hand, the library also provides the evaluation of a pointwise color function C_n, defined on the grid nodes. Note that, since the advection of a phase can produce thin filaments, the evaluation of C_n requires some attention. Initially, every node inside a cut cell is assigned with a value equal to 0.75 if the sign of the quadric equation is positive. Otherwise, for the grid nodes with negative quadric sign we fix the value at 0.25. These values on the grid nodes remain unchanged for boundary cells, i.e. the cells containing the markers. However, for the internal cells, described by $C = 1$, a check on the nodes is performed in order to assign $C_n = 1$ on every node.

In Figure 5.2 a representation of C and C_n is reported for the same set of four cells. Note the difference between the computed value of C by using a cut-fem integration and the corresponding approximated value of C_n depending mainly on the sign of the quadric. Regarding the latter, since the example reported deals with quadrilateral finite elements, the C_n variable is placed at the biquadratic nodes of the cell.

5.1.3 Runge-Kutta advection scheme

After the initialization routine, the marker located on the quadric equation for each boundary cell and the central marker are ready to be advected by a velocity field. In particular, in order to test the library the preliminary simulations have been performed considering an analytical velocity field, without implementing and resolving the Navier-Stokes type equations. In addition, using specific velocity fields, such as periodic stream functions, allows for the verification of the accuracy of the algorithm, as demonstrated in the results section of this part. On the other hand, the numerical scheme adopted for the advection does not change in the presence of a fully solved velocity field, which derives for example from the resolution of the Navier-Stokes system in the context of multiphase flow simulations.

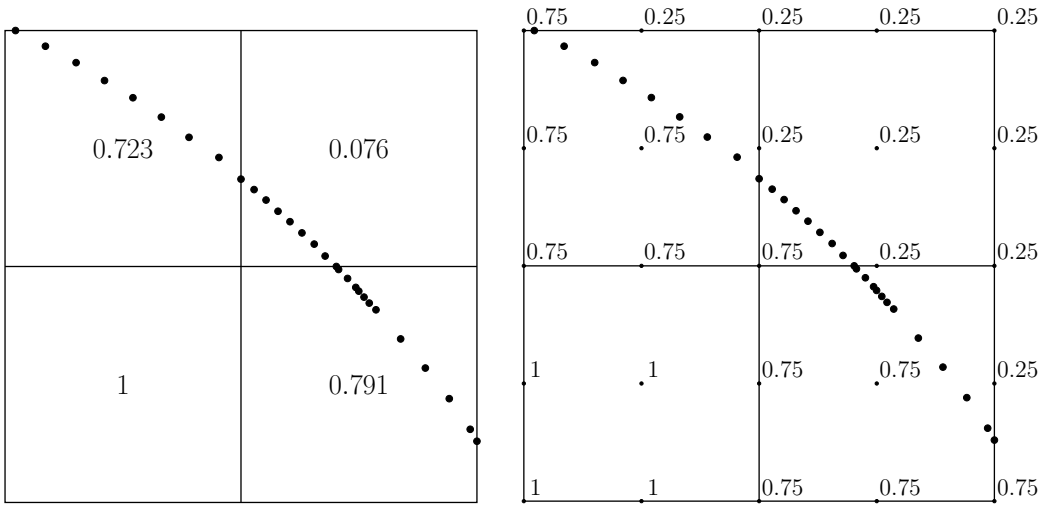

Figure 5.2: Four cells of a generic quadric surface for a markers cloud. On the left the piecewise constant value of C and on the right the value of C_n for every cell node.

In the library, the standard Runge-Kutta method [92] has been implemented in order to move the markers in the cell. Although the simulations performed use a 4-th order Runge-Kutta scheme, the library provides the flexibility to perform a generic n-th order scheme. In this section, the 4th-order scheme is described to recall standard results present in the literature about discretized advection schemes. Therefore, considering an initial time t_0 and a fixed time step interval Δt, the initial value problem is defined as

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}(\boldsymbol{x}, t)\,, \qquad \text{with} \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0\,, \tag{5.11}$$

where $\boldsymbol{v}$ is the velocity field function of the space $\boldsymbol{x}$ and the time t, and $\boldsymbol{x}_0$ represents the initial position.

To evaluate the position, i.e., the spatial coordinates $\boldsymbol{x}_i$ for the i-th marker at the time t, we need to solve the following relation

$$\boldsymbol{x}_i(t) = \boldsymbol{x}_i(t_{n-1} + \Delta t) = \boldsymbol{x}_i(t_{n-1}) + \frac{1}{6}(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4)\Delta t\,, \tag{5.12}$$

where the coefficients $\boldsymbol{k}_i$ are defined as

$$
\begin{aligned}
\boldsymbol{k}_1 &= \boldsymbol{v}\left(\boldsymbol{x}(t_{n-1}), t_{n-1}\right) , \\
\boldsymbol{k}_2 &= \boldsymbol{v}\left(\boldsymbol{x}(t_{n-1}) + \boldsymbol{k}_1 \frac{\Delta t}{2}, t_{n-1} + \frac{\Delta t}{2}\right) , \\
\boldsymbol{k}_3 &= \boldsymbol{v}\left(\boldsymbol{x}(t_{n-1}) + \boldsymbol{k}_2 \frac{\Delta t}{2}, t_{n-1} + \frac{\Delta t}{2}\right) , \\
\boldsymbol{k}_4 &= \boldsymbol{v}\left(\boldsymbol{x}(t_{n-1}) + \boldsymbol{k}_3 \Delta t, t_{n-1} + \Delta t\right) .
\end{aligned}
$$

This numerical scheme is performed at every time step and in every cell to represent the surface motion by moving the markers. This step serves as the initial configuration for the next section, where, starting from the advection result, the rebuilding of the marker position is performed when certain conditions are not satisfied. The new parameters related to the markers are computed, and once the new markers are found from the best-fit quadric of the cell, another time step with the Runge-Kutta method is performed.

5.1.4 Best-fit quadric equation

After the marker advection for every cell, a crucial aspect of the algorithm is the computation of a quadric equation that represents the best-fit approximation of the marker positions. This is achieved by exploiting the information about the marker of a single cell at a specific time step, i.e. position, normal, along with the markers from neighboring cells. The resulting quadric equation is then selected from an ellipse, a hyperbola, or a parabola. In particular, the least-square minimization approach is employed to determine the coefficients corresponding to these types of conics, and a criterion is implemented to choose the best one among the three. To briefly summarize, the algorithm for seeking the best-fit quadric equation relies solely on the relative positions of the markers. Other aspects, such as incorporating information about the normal vector in the minimization process, are still under investigation for library development but are beyond the scope of this thesis.

The first step of the minimization process involves evaluating the barycenter in the cell, considering a weighted set of markers instead of the classical geometrical barycenter. This step is performed using the arc length variable s of each marker, allowing the detection and penalization of regions with a high density of markers (i.e. marker cluster resulting from the advection routine).

In particular, the weighted markers are evaluated with a Gaussian distribution curve that depends on the distance between the specific marker and the barycenter, considering also the marker of the neighboring cells. Therefore, the square distance between the barycenter $\boldsymbol{x}_c$ and the i-th with coordinate $\boldsymbol{x}_i$ is defined as

$$d_i^2 = (x_i - x_c)^2 + (y_i - y_c)^2 \,. \tag{5.13}$$

The variance related to the Gaussian curve is computed as

$$\sigma^2 = \frac{\sum_{i=1}^{n} d_i^2}{n f_\sigma} \,, \tag{5.14}$$

where n represents the total number of considered markers, including the markers of the neighboring cells. Additionally, the parameter f_σ can be chosen in order to penalize markers that are distant from the barycenter differently. Therefore, the weight for the i-th marker can be defined using the variance as

$$w_i = s\, e^{-\frac{d_i^2}{2\sigma^2}} \,. \tag{5.15}$$

Having the weighted marker, the Jacobi Singular Value Decomposition (SVD) is utilized in order to find the quadric coefficients. In particular, a new barycenter point $\boldsymbol{x}_g$ is determined, resulting from the weighted marker as

$$\boldsymbol{x}_g = \sum_{i=1}^{n} \boldsymbol{x}_i w_i \,. \tag{5.16}$$

Let $\delta_i = \sqrt{(x_i - x_g)^2 + (y_i - y_g)^2}$, and $\delta_{max} = \max(\delta_i)$, $i = 1, \cdots, n$ and introduce the variables

$$\chi_i = \frac{x_i - x_g}{\delta_{max}} \,, \qquad \psi_i = \frac{y_i - y_g}{\delta_{max}} \,. \tag{5.17}$$

With these new variables, we can build a new quadric cost function to be minimized, i.e. $\sum_{i=1}^{n} w_i (A\chi_i^2 + B\chi_i\psi_i + C\psi_i^2 + D\chi_i + E\psi_i + F)^2$. Hence, we want to find the coefficients $X = [A, B, C, \dots]$ related to the new entries of the cost function matrix M. Therefore, for every marker we can build a corresponding row $M(i,)$ that fills the matrix M which is then decomposed with the Jacobi SVD technique [93]. In particular, for a quadric interpolation, the six matrix entries are expressed as

$$M(i,) = \left[\chi_i^2, \quad \chi_i\psi_i, \quad \psi_i^2, \quad \chi_i, \quad \psi_i, \quad 1\right] \,. \tag{5.18}$$

Since the matrix M is built, the Jacobi decomposition $M = USV^*$ is performed by exploiting the linear algebra library Eigen [94]. In particular, considering the right singular vector V matrix, the quadric coefficients X that represent the best-fit approximation of the marker position are taken from the last column of V, which represents the minimum eigenvector of the matrix [95].

On the other hand, our analysis has shown that the previous steps are not sufficient to find the best quadric coefficients in every cell. Neglecting information about the normal vector can lead to quadrics that are in good agreement with the marker positions, but are not representative of the real interface position. For example, considering a parabola quadric, a method to mitigate this aspect is to align the quadric with the principal direction of the marker cloud. Specifically, we can define the $n \times 2$ matrix X, for which the i-th row can be filled with

$$X(i,) = [x_i - x_g, y_i - y_g] \ . \tag{5.19}$$

Therefore, if we consider the square matrix B, equal to $B = X^T X$ and with dimension 2×2, the eigenvector corresponding to the minimum eigenvalue of B represents the principal direction.

After that, we consider for every marker its outer normal, i.e. $\hat{\boldsymbol{n}}_i^{old}$ and also $\hat{\boldsymbol{n}}_{i,quad}$ representing the outer normal vector of the three evaluated quadric (ellipse, hyperbola, parabola). Therefore, an additional cost function is built such that, for each quadric, computes the value of the distance between the markers and the best-fit quadric with a penalizing term. This one is evaluated considering the dot product between the previously cited normal vectors

$$\hat{\boldsymbol{n}}_i^{old} \cdot \hat{\boldsymbol{n}}_{i,quad} \, . \tag{5.20}$$

Finally, the best-fit quadric corresponds to the one that is able to preserve the direction of the normal vectors most efficiently, i.e. the conic equipped with the lower value of the cost function previously described.

In Figure 5.3, a cell with two quadric surfaces for the markers' approximation is shown. The dotted line represents the quadric surface equation for an ellipse, while the dashed one represents a parabola. Despite the good approximation for both quadrics, the cost value computed for the ellipse is lower than the parabola one. This comparison is performed in every cell in order to determine the best quadric for the markers' approximation, leading to the recreation of the markers in the same cell avoiding clusters. In Figure

5.4, an advected circle is depicted with different colored markers cell by cell, representing the three types of quadrics used in this function. In particular, black markers represent the parabola, red the ellipse and green the hyperbola.

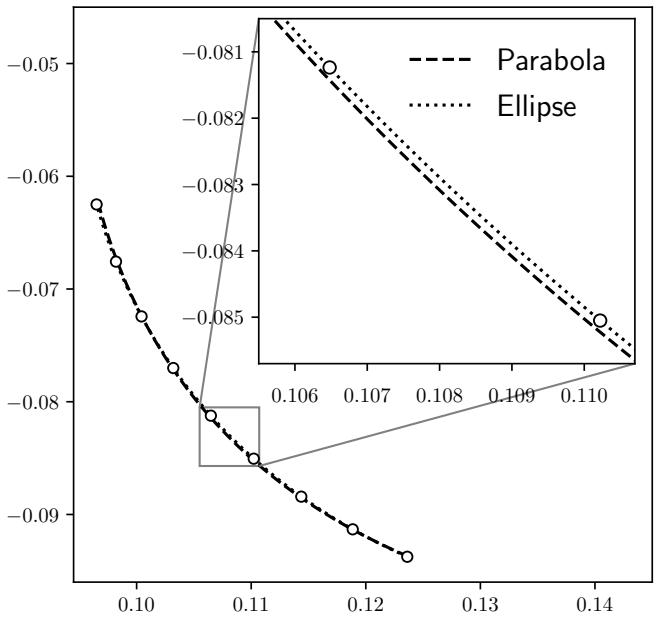

Figure 5.3: Cell with nine markers and their approximation with two kinds of quadrics. The dashed line represents a parabola, meanwhile the dotted one is an ellipse. The latter is chosen because of a lower cost value.

5.1.5 Rebuilding marker position

Once the best-fit quadric equation based on the advected marker positions is found, we have the option to rearrange the markers inside the cell. This feature gains interest since it is not possible to control the marker distribution directly after advection. Indeed, the new marker position depends solely on the imposed velocity field. Consequently, even if we start with a homogeneous distribution of the marker inside the cell, this homogeneity may be lost after advection, leading to regions with very high or low marker density. Thus, the ability to regenerate markers with a better distribution based on the best-fit quadric can be a useful computational tool.

Furthermore, although the distribution can be managed with the rebuilding operation, we are also interested in maintaining a controlled number of

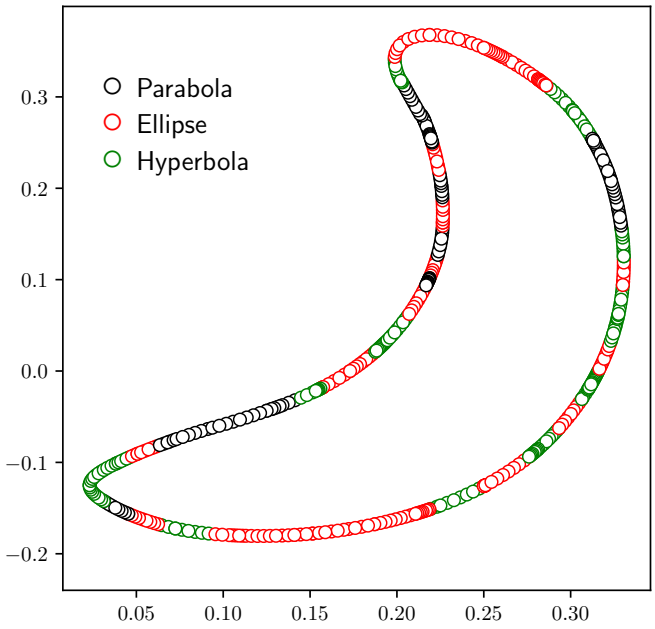

Figure 5.4: Different kinds of quadric surfaces employed to approximate the markers in the cells.

markers inside every cell. Hence, the library is equipped with two parameters n_{min} and n_{max}, representing the minimum and maximum number of possible markers inside the cell. Therefore, if the advection results in a number of markers n for a cell that satisfies $n_{min} \leq n \leq n_{max}$, we do not apply the rebuilding algorithm. Otherwise, we regenerate a fixed number of markers n_0 inside the cell starting from the best-fit quadric.

However, the resulting best-fit quadric does not allow for a direct remeshing of the markers inside the cell. In certain situations, the quadric obtained may have more than two intersections with the cell edges, such as four intersections. This can occur, for instance, in cases where thin filaments exist, like the tail of a very stretched bubble. From a quadric point of view, this translates into situations such as ellipses with one semi-axis much smaller than the other, or where both branches of a hyperbola are included in the cell, creating four intersections. To avoid issues arising from these situations, the library is equipped with an adaptive refinement of the cell. The square or triangular cell is divided into four sub-cells using this algorithm, and the process is repeated until we have only two intersections of the quadric with the edges of the sub-cells. To subdivide the cell, projection matrices, $\mathcal{P}_Q$ and

$\mathcal{P}_T$, have been implemented in the library for the mapping between the cell and its four sub-cells, both for quadrilateral and triangular elements. An example of this situation is depicted in Figure 5.5, for both types of finite elements.

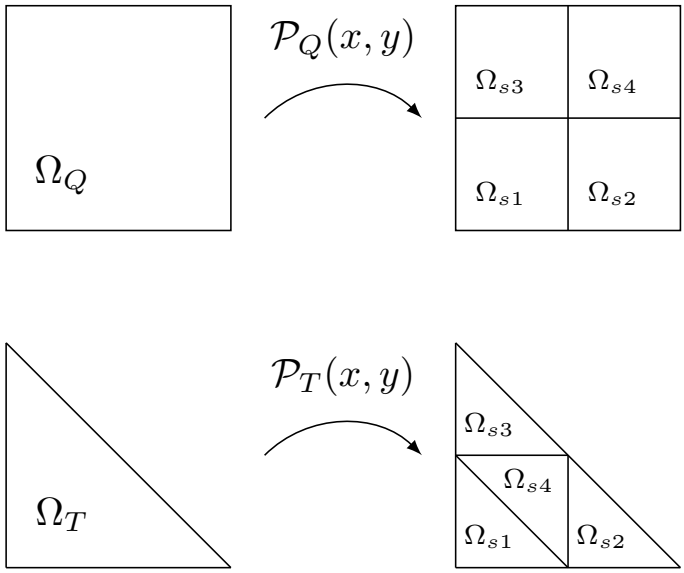

Figure 5.5: Refinement into four sub-cells for quadrilateral and triangular finite elements.

After the adaptive refinement, each cell has only a double intersection, represented by $\boldsymbol{x}_0 = (x_0, y_0)$, $\boldsymbol{x}_1 = (x_1, y_1)$, with the edges. The goal is to find the coordinates of the new set of markers located on the quadric. We start by finding the center coordinates of the osculating circle to the quadric, as

$$\boldsymbol{x}_c = (x_c, y_c) = \boldsymbol{x}_m - \frac{\hat{\boldsymbol{n}}}{k}\,. \tag{5.21}$$

In this case $\boldsymbol{x}_m$ represents the mid-point between the intersections $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$, $\hat{\boldsymbol{n}}$ is the outward normal vector, and k is the quadric curvature evaluated at $\boldsymbol{x}_m$. After that, we evaluate the angles between the center of the osculating circle $\boldsymbol{x}_c$ and the intersections $\boldsymbol{x}_1$ and $\boldsymbol{x}_0$ as

$$\theta_0 = \arctan\left(\frac{y_0 - y_c}{x_0 - x_c}\right), \qquad \theta_1 = \arctan\left(\frac{y_1 - y_c}{x_1 - x_c}\right). \tag{5.22}$$

In addition, we also define other two variables representing the differences

between these angles as

$$\Delta\theta_0 = \begin{cases} \theta_1 - \theta_0 & \text{if} \quad \theta_1 > \theta_0 \\ 2\pi + \theta_1 - \theta_0 & \text{if} \quad \theta_1 < \theta_0 \,, \end{cases} \tag{5.23}$$

$$\Delta\theta_1 = \begin{cases} \theta_0 - \theta_1 & \text{if} \quad \theta_0 > \theta_1 \\ 2\pi + \theta_0 - \theta_1 & \text{if} \quad \theta_0 < \theta_1 \,. \end{cases} \tag{5.24}$$

Therefore, we introduce the angle θ_s and the angle variation $\Delta\theta$ with the following conditions

$$\begin{aligned} &\text{if } \Delta\theta_0 \geq \Delta\theta_1 \quad \rightarrow \quad \Delta\theta = \Delta\theta_1 \qquad \theta_s = \theta_1 \,, \\ &\text{if } \Delta\theta_0 < \Delta\theta_1 \quad \rightarrow \quad \Delta\theta = \Delta\theta_0 \qquad \theta_s = \theta_0 \,. \end{aligned}$$

We define now $\boldsymbol{v} = \langle v_x, v_y \rangle$ as

$$v_x = R_0 \cos\left(\theta_s + \frac{\Delta\theta}{2}\right) \,, \qquad v_y = R_0 \sin\left(\theta_s + \frac{\Delta\theta}{2}\right) \,, \tag{5.25}$$

where $R_0 = \sqrt{(x_0 - x_c)^2 + (y_0 - y_c)^2}$ represents the radius of the osculating circle. Then, we can compute the intersection between the quadric equation and the line passing through the circle center $\boldsymbol{x}_c$ and with the direction given by the vector $\boldsymbol{v}$. Finally, the new marker positions are found by solving the following relation

$$\boldsymbol{x}_n = \boldsymbol{x}_c + t_n \boldsymbol{v} \,, \tag{5.26}$$

where $t_n = \min(t_1, t_2)$, and t_1 and t_2 are the solutions of $a^* t^2 + b^* t + c^* = 0$. In this case, the normalized coefficients, defined with the $*$, are defined as follows

$$a^* = \frac{a}{\sqrt{a^2 + b^2 + c^2}} \,, \tag{5.27}$$

$$b^* = \frac{b}{\sqrt{a^2 + b^2 + c^2}} \,, \tag{5.28}$$

$$c^* = \frac{c}{\sqrt{a^2 + b^2 + c^2}} \,, \tag{5.29}$$

$$\tag{5.30}$$

where the a, b, c coefficients are defined considering the quadric coefficients, the point $\boldsymbol{x}_c$ and the vector $\boldsymbol{v}$

$$a = Av_x^2 + Bv_x v_y + Cv_y^2 \,, \tag{5.31}$$

$$b = 2Av_x x_c + B(v_y x_c + v_x y_c) + 2Cv_y y_c + Dv_x + Ev_y \,, \tag{5.32}$$

$$c = Ax_c^2 + Bx_c y_c + Cy_c^2 + Dx_c + Ey_c + F \,. \tag{5.33}$$

Moreover, the algorithm is also able to update the arc-length variable ds for the new marker as

$$ds = \frac{R_0 + R_n}{2}\Delta\theta\,, \tag{5.34}$$

where R_n is the distance between $\boldsymbol{x}_n$ and $\boldsymbol{x}_c$ defined as

$$R_n = \sqrt{(t_n v_x)^2 + (t_n v_y)^2}\,. \tag{5.35}$$

In Figure 5.6 a cell with two sets of markers is reported: the first one in black is the set of markers advected from the previous iteration; in red are reported the rebuilt markers. To fix the ideas, in the reported example we set $n_{min} = 9$ and $n_{max} = 12$. In this case, the reconstruction of the markers is performed because $n_{old,RK4} = 16 > n_{max}$. Therefore, we reduce their population inside the cell, placing the new n_0 markers (in this case $n_0 = 9$) on the interpolating quadric. Unlike the black markers, which are in the position obtained from the RK4 scheme, note that the new markers are built by placing the endpoints on the edges of the cell. The other internal markers are located homogeneously on the approximating quadric as described before, creating $n - 1$ circular sectors.

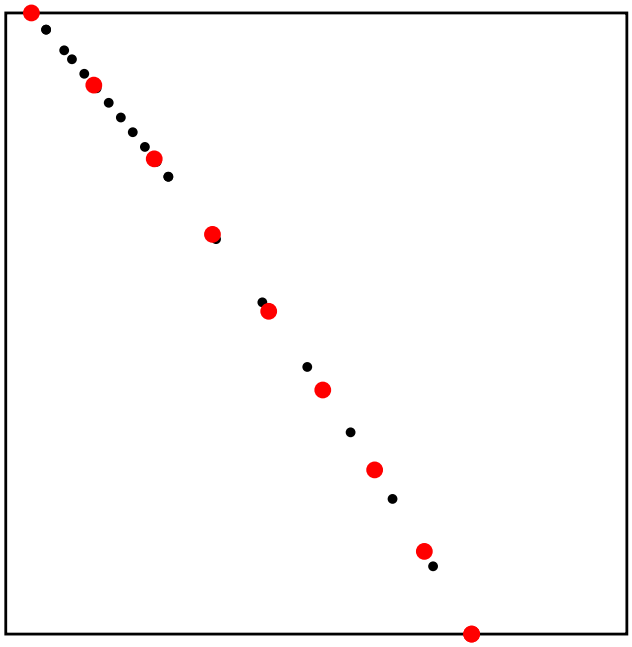

Figure 5.6: Reconstruction of the markers from the black one (the result of the advection) to the red one (to be advected in the new time step).

5.2 Interface advection test

In this section, numerical results are presented in order to test the algorithm outlined above. In particular, standard tests from the literature are considered benchmark results to evaluate the interface tracking. Flow fields characterized by uniform translations and rotations are employed to move and advect the markers in the cells. The numerical tests discussed in [24, 25] are considered as the reference benchmark, even though other cases have been tested and presented later. Other interface tracking tests found in the literature, such as those in [96, 97], share a similar concept. The simple advection movements are designed to displace smooth fluid bodies within a domain and verify the conservation of surface shape and volume.

All the tests presented are two-dimensional problems, with a velocity field designed so that the resulting vorticity is not uniform in the domain. Consequently, significant distortions occur in the fluid body interface, making the maintenance of the interface not straightforward. Finally, cosinusoidal time-dependence analytical velocity functions are employed, in order to achieve the re-establishment of the initial configuration at the end of the period [98].

To facilitate quantitative comparisons of the results, the L_1 error norms presented in [24] have been computed. The *relative mass error* $E_m(t_1)$ is defined to compare the total volume of a phase, specifically the reference volume, at the initial time t_0 and the subsequent time t_1

$$E_m(t_1) = \frac{|\sum_{i=1}^{N_{el}} A_i C_i(t_1) - A_i C_i(t_0)|}{\sum_{i=1}^{N_{el}} A_i C_i(t_0)} \,. \tag{5.36}$$

In this case, the color function value at the cell i at time t is represented by $C_i(t)$, A_i represents the area of the cell i meanwhile the total number of the cells is N_{el}. Another error, termed *geometrical error* $E_g(t_1)$, is introduced as

$$E_g(t_1) = \sum_{i=1}^{N_{el}} A_i |C_i(t_1) - C_i(t_0)| \,. \tag{5.37}$$

As mentioned earlier, the objective is to verify whether the final shape aligns with the initial configuration. To achieve this, we introduced and computed another type of error, considering a circular geometry as the initial shape. Therefore, given the center (x_c, y_c) and the radius R, the distance between a marker m with position (x_m, y_m) and the center is computed and compared

with the radius R:

$$E_{al} = \sum_{m=1}^{N_m} \left| \sqrt{(x_m - x_c)^2 + (y_m - y_c)^2} - R \right| s_m \,, \tag{5.38}$$

where, for each marker, the arc length is denoted by s_m.

In all computations presented, the $CFL = u\Delta t/h$ has been considered, with u representing the velocity component along the x-axis, Δt the time step simulation and h the grid spacing.

Single vortex test The first test conducted to evaluate the algorithm's performance involves the advection of a single bubble within a square domain. The domain is defined as $\Omega = [-L/2, L/2] \times [-H/2, H/2]$, where $H = L = 1$. The reference phase is represented by a circular geometry centered at $(0, 0.25)$ with a radius of $R = 0.15$. The simulations were carried out using both triangular and quadrangular mesh discretizations. For the analytical velocity field, a sinusoidal stream-function was employed, described by the equation

$$\psi = \frac{1}{\pi} \sin^2(\pi x) \sin^2(\pi y) \cos\left(\frac{\pi t}{T}\right) \,, \tag{5.39}$$

that leads to

$$u = \frac{\partial \psi}{\partial y} = 2 \sin^2(\pi x) \sin(\pi y) \cos(\pi y) \cos\left(\frac{\pi t}{T}\right) \,,$$
$$v = -\frac{\partial \psi}{\partial x} = -2 \sin(\pi x) \cos(\pi x) \sin^2(\pi y) \cos\left(\frac{\pi t}{T}\right) \,,$$

where T represents the vortex period. The simulation starts at $t = 0$ and progresses until reaching the maximum deformation point at $t = T/2$, ultimately returning to the initial configuration at $t = T$. The positions of the markers at the initial and final configurations are then compared to assess the errors introduced during the simulation.

The chosen test has been selected based on other tests described in the literature. Consequently, in Table 5.1, the errors values have been compared with those in [24], as the same test was presented in that paper and was taken as a reference case. As expected, both errors decrease with an increase in grid refinement. Regarding the relative mass error E_m, the results are consistent with the ones reported in [24]. In fact, it can be noticed that with finest grid we reach the same order of magnitude, equal to 10^{-5}. In addition, the geometrical error E_g presents a better behavior considering the

N_{lev}	E_m	E_g	$E_m(AMS)$	$E_g(AMS)$	p_m	p_g
4	$9.42 \cdot 10^{-2}$	$6.56 \cdot 10^{-3}$	$9.42 \cdot 10^{-3}$	$2.53 \cdot 10^{-2}$	-	-
5	$1.03 \cdot 10^{-2}$	$7.25 \cdot 10^{-4}$	$1.94 \cdot 10^{-3}$	$2.78 \cdot 10^{-3}$	3.20	3.18
6	$7.25 \cdot 10^{-4}$	$5.12 \cdot 10^{-5}$	$2.53 \cdot 10^{-4}$	$4.78 \cdot 10^{-4}$	3.82	3.82
7	$6.47 \cdot 10^{-5}$	$4.57 \cdot 10^{-6}$	$5.95 \cdot 10^{-5}$	$1.16 \cdot 10^{-4}$	3.49	3.49

Table 5.1: Single vortex field test with $T = 8\,s$: values of the relative mass error E_m (5.36) and the geometrical error E_g (5.37) with $CFL = 1$ for different grid resolutions. The comparison with the results presented in [24] (AMS) is also reported. The corresponding rates of convergence p_m, p_g are reported in the last columns.

values in [24]. In fact, for every level of refinement, the error is one order of magnitude lower than $E_g(AMS)$. In the last columns, we report the rates of convergence corresponding to E_g and E_m. In particular, these parameters have been defined as

$$p_m \approx \frac{\ln\left(\frac{E_m(h)}{E_m(h/2)}\right)}{\ln(2)}, \qquad p_g \approx \frac{\ln\left(\frac{E_g(h)}{E_g(h/2)}\right)}{\ln(2)}, \tag{5.40}$$

where h represents the grid spacing for the considered mesh. For both orders of convergence, a value close to 3 is obtained, showing approximately a cubic convergence of the implemented algorithm.

Note that the error values are reported based on the level of mesh refinement, where level 4 corresponds to a 32×32 grid, level 5 to a 64×64 grid, level 6 to a 128×128 grid, and level 7 to a 256×256 grid.

In Table 5.2, the circular shape error E_{al} introduced by the definition in (5.38) is also reported, quantifying the geometrical inaccuracy of the tracking algorithm. The table illustrates the convergence of the error with an increase in grid refinement. The last column presents the rate of convergence based on the considered grids. This parameter is defined as

$$p_{al} \approx \frac{\ln\left(\frac{E_{al}(h)}{E_{al}(h/2)}\right)}{\ln(2)}, \tag{5.41}$$

N_{lev}	E_{al}	p_{al}
4	$5.68 \cdot 10^{-3}$	-
5	$7.39 \cdot 10^{-4}$	2.94
6	$5.24 \cdot 10^{-5}$	3.38
7	$6.07 \cdot 10^{-6}$	3.29

Table 5.2: Single vortex field test with $T = 8\,s$: values of the newly introduced geometrical error E_{al} (5.38) with the corresponding rate of convergence p_{al}.

where h represents the grid spacing for the considered mesh. This value is approximately equal to 3, indicating the good behavior of the method, as the convergence is almost cubic.

Additionally, Figure 5.7 displays the final configuration after the advection of the marker. This position corresponds to the time $t = T = 8$ seconds, i.e., the stream function at the final period after that the bubble has reached the maximum distortion point at $t = T/2 = 4$ seconds (see Figure 5.8). It is evident that increasing the refinement produces final configurations that are in good agreement with the initial geometry represented by the dashed line, which is the initial circumference. Only for coarse grids, some errors can be noticed in reaching the initial geometrical configuration.

Double bubble vortex test The next test is conducted within the same domain as the single bubble, i.e., $\Omega = [-L/2, L/2] \times [-H/2, H/2]$, with $H = L = 1$. The objective is to advect multiple regions representing the reference phase inside Ω, starting from two initial circular geometries. The centers of the circumferences have been fixed at $(0, 0.25)$ and $(0, -0.25)$, and the same radius $R = 0.15$ has been considered. In this test, recalling the algorithm described in Section 5.1.5, we have set the minimum and maximum number of markers for the rebuilding technique, i.e., $n_{min} = 9$ and $n_{max} = 12$. The number of new markers n_0 has been set equal to n_{min}. Two types of computational grids have been considered for the numerical simulation: biquadratic quadrilateral elements (QUAD9) and biquadratic triangular elements (TRI6), in order to test the library with different space discretizations. Regarding the velocity field, as described above, a sinusoidal stream function has been chosen to compare the initial and final geometry config-

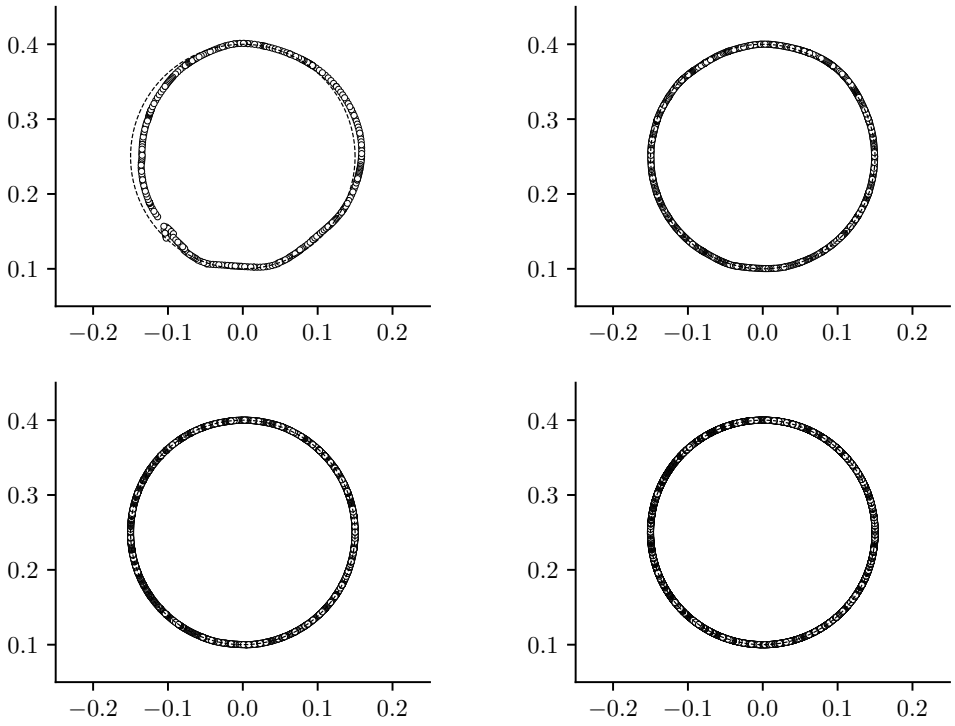

Figure 5.7: Single bubble test: comparison of the final interface position at $t = 8\,s$ (circular marker) with the initial circular geometry (dashed black line) for the meshes with 32×32 (top left), 64×64 (top right), 128×128 (bottom left) and 256×256 (bottom right) cells.

urations over multiple periods. In particular, the following stream function has been exploited

$$\psi = -\frac{1}{\pi}\sin^2(\pi x)\sin^2(\pi y)\cos\left(\frac{\pi t}{T}\right), \tag{5.42}$$

leading to the respective velocity components defined as

$$u = \frac{\partial \psi}{\partial y} = -2\sin^2(\pi x)\sin(\pi y)\cos(\pi y)\cos\left(\frac{\pi t}{T}\right), \tag{5.43}$$

$$v = -\frac{\partial \psi}{\partial x} = 2\sin^2(\pi x)\sin(\pi y)\cos(\pi y)\cos\left(\frac{\pi t}{T}\right). \tag{5.44}$$

Also in this case, by using a sinusoidal function we expect to obtain the same geometric shape at the final time step $t = T$ as in the initial one at $t = 0$. This aspect is considered as an index to check the quality of the newly proposed algorithm. Unlike the single bubble test, the period T for this simulation

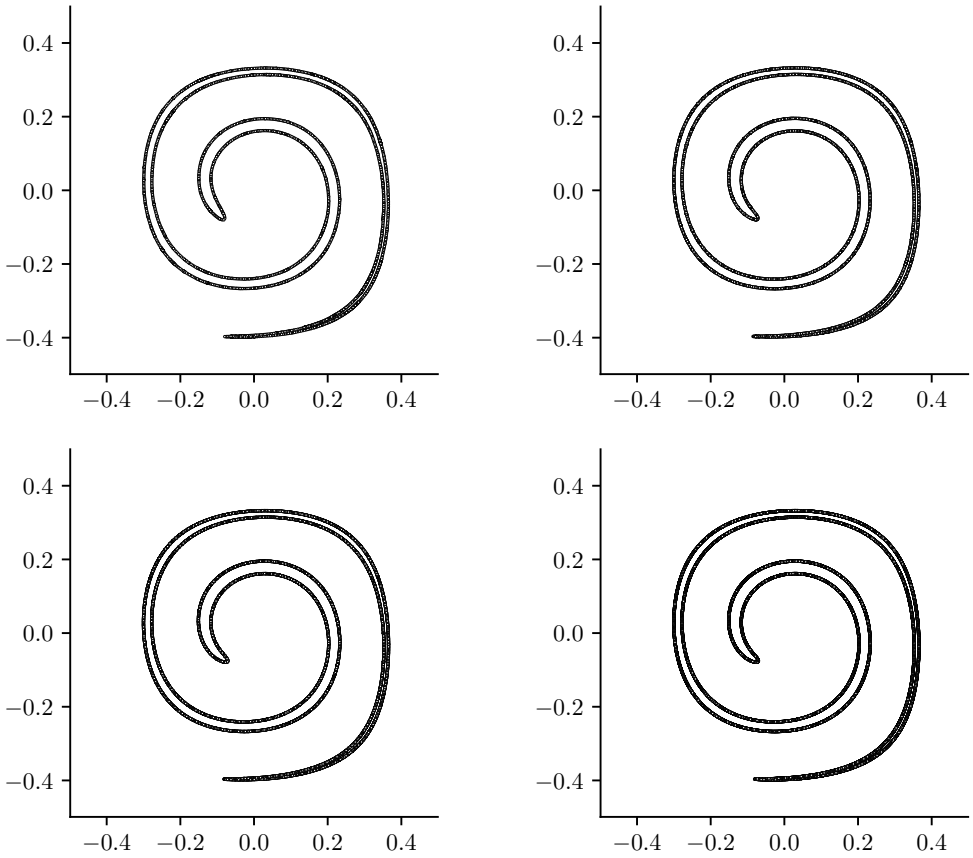

Figure 5.8: Single bubble test: interface position at maximum deformation ($t = 4\,s$) for the meshes with 32×32 (top left), 64×64 (top right), 128×128 (bottom left) and 256×256 (bottom right) cells.

has been set equal to $4s$, and thus the maximum deformation is reached at $t = 2s$.

In Table 5.3, the E_m and E_g errors are reported for different levels of domain discretization, ranging from a 32×32 grid up to a 256×256 one. A decrease in both error norms can be observed with an increase in grid refinement, confirming the good behavior of the implemented algorithm.

Additionally, in the last columns, the order of convergence for E_m and E_g are reported. For both errors, the rates of convergence p_m and p_g, defined in (5.40), seem to have a value between 2 and 3.

From a graphical perspective, we can compare the geometric positions of the markers at the beginning and at the end of the periodic simulation. In fact, we know that the advection of these points with the rebuilding routine

N_{el}	E_m	E_g	p_m	p_g
32×32	$1.30 \cdot 10^{-2}$	$1.83 \cdot 10^{-3}$	-	-
64×64	$1.17 \cdot 10^{-3}$	$1.66 \cdot 10^{-4}$	3.46	3.46
128×128	$1.85 \cdot 10^{-4}$	$2.61 \cdot 10^{-5}$	2.67	2.67
256×256	$3.07 \cdot 10^{-5}$	$4.34 \cdot 10^{-6}$	2.59	2.59

Table 5.3: Double bubble test: values of the E_m error and the E_g error for different grids, with corresponding rates of convergence p_m and p_g.

may lead to markers deviating from the theoretical trajectory given by the analytical velocity field. Figure 5.9 presents the final positions of the markers after being advected for a period $t = T = 4s$, allowing for a comparison with the initial configuration at $t = 0$. The dashed black line represents the initial geometric configuration, considering the equations of the two circumferences on which the markers are initially located. The circular markers represent the markers at the end of the simulation, showing good agreement with the initial position for every grid considered, except for the coarser grid, i.e., the 32×32 one, where some discrepancies can be noted.

Furthermore, in Figure 5.10, the configuration at maximum distortion, i.e., the situation at $t = T/2$, is presented to illustrate the thin filaments of the bubble tails. Specifically, the maximum stretch for the 32×32 and 128×128 computational grids is shown.

5.3 Advection test by using Raviart-Thomas element

This section introduces the initial attempt to incorporate Raviart-Thomas basis functions into the surface marker reconstruction framework. The primary goal is to compare the advection test using two different representations of the velocity field. As the Navier-Stokes equation is not solved in this application, the velocity field is imposed using an analytical function, chosen as sinusoidal for consistency with previous tests. The comparison focuses on how the analytical velocity field is approximated in the code. Despite having the exact solution for the velocity, the code is structured for the resolution

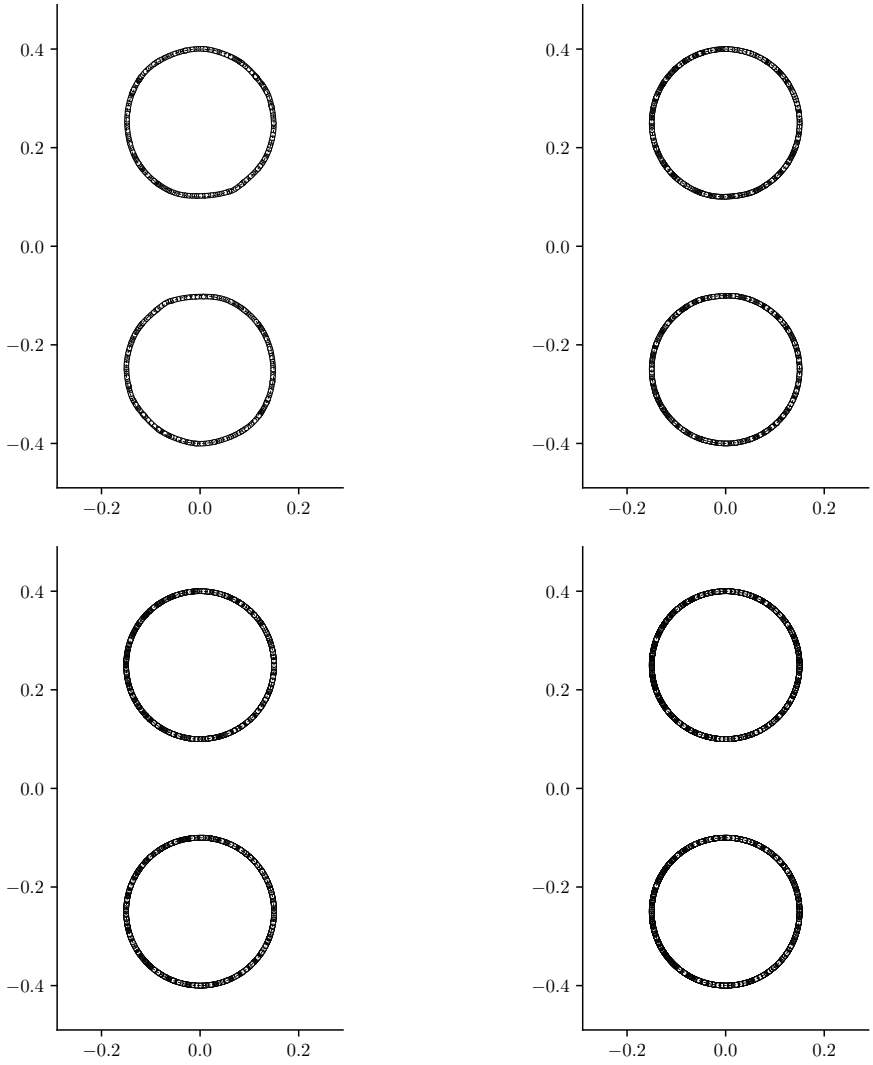

Figure 5.9: Double bubble test: reconstructed interface back to the initial position at $t = 4\,s$ for the meshes with 32×32 (top left), 64×64 (top right), 128×128 (bottom left) and 256×256 (bottom right) cells. The initial circular configuration is reported in a dashed line.

of dynamic simulations. These tests involve resolving the velocity field while considering surface tension modeling on the right-hand side of the equation.

Indeed, the numerical approximation of the velocity is naturally stored through finite element discretization, such as the classical 9 points of a bi-quadratic quadrilateral element. In the same way, the analytical velocity is stored based on the specific data structure arising from mesh discretization.

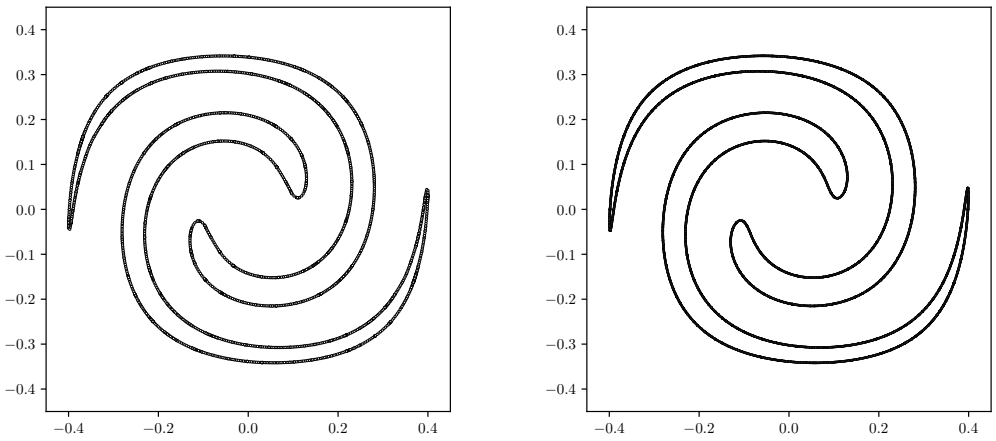

Figure 5.10: Double bubble test: reconstructed interface at maximum deformation at $t = 2\,s$ for the 32×32 (left) and 128×128 (right) grids.

The exact values of that field are found only in the standard degrees of freedom of the mesh elements. Notably, the library for marker reconstruction is designed to handle both kinematic and dynamic two-phase flow simulations, accommodating scenarios where the velocity field is either imposed or fully solved.

In situations where we need to compute the velocity at points beyond the nodes of the elements (such as marker coordinates), interpolation is used. This is essential for marker advection, where the Runge-Kutta method requires varying velocity values at different positions. This section aims to compare two types of finite element interpolation techniques for determining the velocity field at marker locations. The standard approach involves Lagrangian interpolation, while the alternative method employs Raviart-Thomas interpolation.

Denoting $\boldsymbol{u}_{mrk}$ as the velocity field of the i-th marker in the cell with coordinates $\boldsymbol{x}_p$, the comparison is performed using the following formulations

$$\boldsymbol{u}_{mrk} = \sum_{i=1}^{n_{dof}} \boldsymbol{u}_i \varphi_i(\boldsymbol{x}_p) \qquad \boldsymbol{u}_{mrk} = \sum_{i=1}^{n_{faces}} \boldsymbol{b}_i(\boldsymbol{x}_p) p_i \,, \tag{5.45}$$

where n_{dof} represents the biquadratic nodes for a quadrilateral element, and n_{faces} is the number of faces of the same element. Naturally, φ_i represents the biquadratic Lagrangian basis functions, while $\boldsymbol{b}_i$ are the Raviart-Thomas basis functions that act on the fluxes faces p_i. It is worth noting that, since the computational domain is a classical bidimensional Cartesian mesh, any con-

vergence issues related to using Raviart-Thomas finite elements, as described in previous chapters, are not present in this context.

The reason behind the use of the Raviart-Thomas basis function for the velocity field interpolation is motivated by two closely connected aspects:

- Mass Conservation in Multiphase Flow: In the context of multiphase flow simulation, maintaining mass conservation is crucial. Since the density of both phases is assumed constant, it is imperative to satisfy the mass conservation equation. Ensuring a divergence-free velocity is a key requirement and is managed through appropriate discretization techniques. The use of a divergence-free representation of the numerical velocity, facilitated by Raviart-Thomas finite element discretization, addresses this constraint effectively;

- Divergence-Free Analytical Velocity Field: The analytical velocity fields implemented for advection tests are sinusoidal functions derived from stream functions, which are, by definition, divergence-free. Therefore, approximating these velocity fields using $\boldsymbol{H}_0(\mathrm{div})$ basis functions should provide an exact representation at every physical point in the domain. This approach helps avoid approximation errors that may arise when using standard finite element Lagrangian interpolation.

In summary, from a computational perspective, only the function responsible for Runge-Kutta advection has been modified to allow for both types of interpolations. The application has been developed by integrating two finite element libraries: the FEMuSTTU library [83], where routines for surface marker reconstruction are implemented, and the ProXPDE library [84], which provides the Raviart-Thomas finite element interpolation.

The approach for testing the Raviart-Thomas basis functions for velocity field interpolation involves a single bubble advection test, similar to the initial one. The analytical velocity field for this simulation is given by

$$u = 2\sin^2(\pi x)\sin(\pi y)\cos(\pi y)\cos\left(\frac{\pi t}{T}\right),$$
$$v = -2\sin(\pi x)\cos(\pi x)\sin^2(\pi y)\cos\left(\frac{\pi t}{T}\right).$$

The domain considered remains $\Omega = [-L/2, L/2] \times [-H/2, H/2]$, with $H = L = 1$, and the initial circular geometry is located at $(0, 0.25)$ with a radius $R = 0.15$. The period T for this test is set to $4s$, implying that the maximum stretch of the bubble is reached at $t = 2s$. The purpose is to compare

the performance of standard Lagrangian interpolation with Raviart-Thomas interpolation for the velocity field in the context of the surface marker reconstruction algorithm. The simulation is performed to evaluate how well each interpolation scheme maintains the accuracy and conservation properties of the marker positions during advection.

N_{el}	E_m^Q	E_g^Q	E_{al}^Q	p_m	p_g	p_{al}
32×32	$1.64 \cdot 10^{-2}$	$1.00 \cdot 10^{-3}$	$9.66 \cdot 10^{-4}$	-	-	-
64×64	$9.88 \cdot 10^{-4}$	$6.69 \cdot 10^{-5}$	$8.52 \cdot 10^{-5}$	4.05	3.91	3.50
128×128	$2.35 \cdot 10^{-4}$	$1.65 \cdot 10^{-5}$	$1.36 \cdot 10^{-5}$	2.07	2.02	2.64

Table 5.4: Values of the E_m, E_g and E_{al} errors for different grids, with respective rates of convergence for the Q_2 velocity interpolation.

In Table 5.4, the error values for the Lagrangian-type interpolation, denoted with the superscript Q, are presented based on the number of mesh elements N_{el}, ranging from a 32×32 grid to a 128×128 grid. Every computed error shows a decrease by increasing the mesh refinement, showing the same behavior as the previous tests. Regarding the rates of convergence of the errors, we can notice that for p_m and p_g we reach a value close to 2, while p_{al} seems to have a value slightly bigger. We recall that the E_{al} error provides a geometrical measure of the difference between a circular initial configuration, and the corresponding final shape. For this reason, an order of convergence of p_{al} between 2 and 3 indicates a good behavior of the algorithm. Note the different error values with respect to Table 5.1, since a different period for the same single bubble simulation has been adopted.

N_{el}	E_m^{RT}	E_g^{RT}	E_{al}^{RT}	p_m	p_g	p_{al}
32×32	$6.43 \cdot 10^{-2}$	$3.94 \cdot 10^{-3}$	$2.24 \cdot 10^{-3}$	-	-	-
64×64	$1.57 \cdot 10^{-2}$	$1.07 \cdot 10^{-3}$	$5.79 \cdot 10^{-4}$	2.03	1.89	1.95
128×128	$3.73 \cdot 10^{-3}$	$2.59 \cdot 10^{-4}$	$1.42 \cdot 10^{-4}$	2.08	2.04	2.03

Table 5.5: Values of the E_m, E_g and E_{al} errors for different grids, with respective rates of convergence for $\boldsymbol{\mathcal{RT}}_0$ velocity interpolation.

The same error analysis has been performed for the $\boldsymbol{\mathcal{RT}}_0$ velocity interpolation, and the results are reported in Table 5.5. For these simulations, similar conclusions can be drawn regarding the evolution of errors: every computed error decreases with the grid refinement. Regarding the rates of convergence, similar values have been obtained for the $\boldsymbol{\mathcal{RT}}_0$, except for p_{al} which seems to be slightly lower than the one computed with the lagrangian interpolation.

In terms of the comparison between the two techniques, it appears that the Q_2 Lagrangian interpolation provides better error results. For every error, the order of magnitude is almost lower by one compared to the corresponding $\boldsymbol{\mathcal{RT}}_0$ interpolation on the same grid. However, it is important to note that these errors consider the entire simulation, taking into account the final values of the color function and the marker positions. As such, these errors, as defined, do not serve as a direct indicator of the approximation quality of the velocity.

The numerical algorithm incorporates several functions for marker reconstruction. Therefore, while the velocity interpolations of the marker may be quite similar inside the cell, a small variation in the velocity field can lead to a small change in the final position of the marker with the Runge-Kutta advection scheme. Additionally, it is on this final position that every routine of the reconstruction library acts, and thus the resulting best-fit approximating quadric can be different.

Moreover, it should be considered that using a $\boldsymbol{H}(\mathrm{div})$ velocity field the conservation of the tangential component is not satisfied. Although this aspect has not been thoroughly investigated, it could be a significant concern, particularly for markers lying on the edges of the cell. As a reminder, to rebuild the marker inside the cell on the best-fit quadric, two markers are located at the intersections between the quadric and the cell. Naturally, these two markers are also advected, computing the velocity interpolation on points lying on cell edges.

In Figure 5.11, the final positions of the markers are reported for different grid refinements obtained with the standard Lagrangian Q_2 interpolation. It is noticeable that, only for the coarser grid (32×32), some discrepancies from the analytical circular geometry are present in the marker positions. As the grid refinement increases, the markers reach a configuration at the end of the period that is in perfect agreement with the initial circumference.

The same comments can be made for the $\boldsymbol{\mathcal{RT}}_0$ interpolation, which is

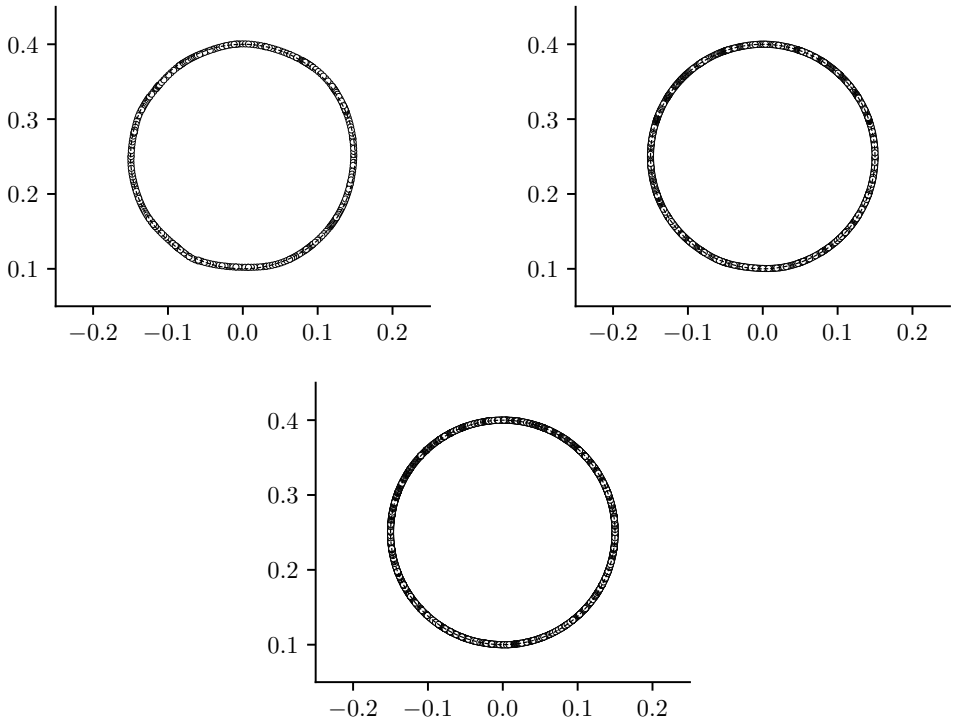

Figure 5.11: Single bubble test with Q_2 velocity interpolation: comparison of the final interface position at $t = 4\,s$ (circular marker) with the initial circular geometry (dashed black line) for the meshes with 32×32 (top left), 64×64 (top right) and 128×128 (bottom) cells.

shown in Figure 5.12. Only for the 32×32 grid, the final marker positions do not match perfectly with the initial circular geometry. It is important to note that this aspect is not merely a graphical consideration; it is represented by the E_{al} error, which measures the discrepancies from the initial shape. For both interpolation techniques, we observe orders of magnitude ranging from 10^{-3} up to 10^{-5}, confirming the good behavior of the algorithm in regaining the initial circumference.

As previously described, it is crucial to verify the shape configuration when the maximum deformation is reached. For this reason, we now present the bubble configuration at half of the period T, i.e., at $t = 2$ seconds, to check if the closed surface is preserved. Despite the thin filaments in the bubble tail, we observe that the algorithm is still able to reconstruct the surface and rebuild the marker positions. This situation is depicted for both techniques, in Figure 5.13 for the Q_2 interpolation and in Figure 5.14 for the

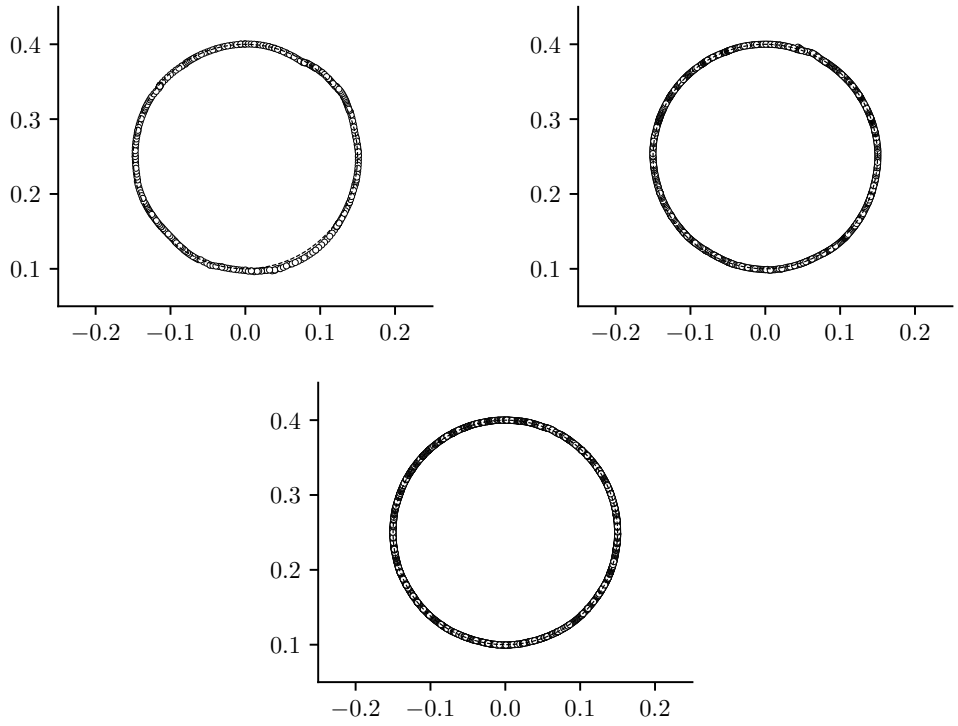

Figure 5.12: Single bubble velocity with $\boldsymbol{\mathcal{RT}}_0$ velocity interpolation: comparison of the final interface position at $t = 4\,s$ (circular marker) with the initial circular geometry (dashed black line) for the meshes with 32×32 (top left), 64×64 (top right) and 128×128 (bottom) cells.

$\boldsymbol{\mathcal{RT}}_0$ interpolation, across the four tested grids.

In conclusion, the coupling by using Raviart-Thomas interpolated velocity has shown promising preliminary results, despite a slight deterioration in the computed error norm. Further investigations are needed for the improvement of the numerical library, in order to obtain more accurate simulation results.

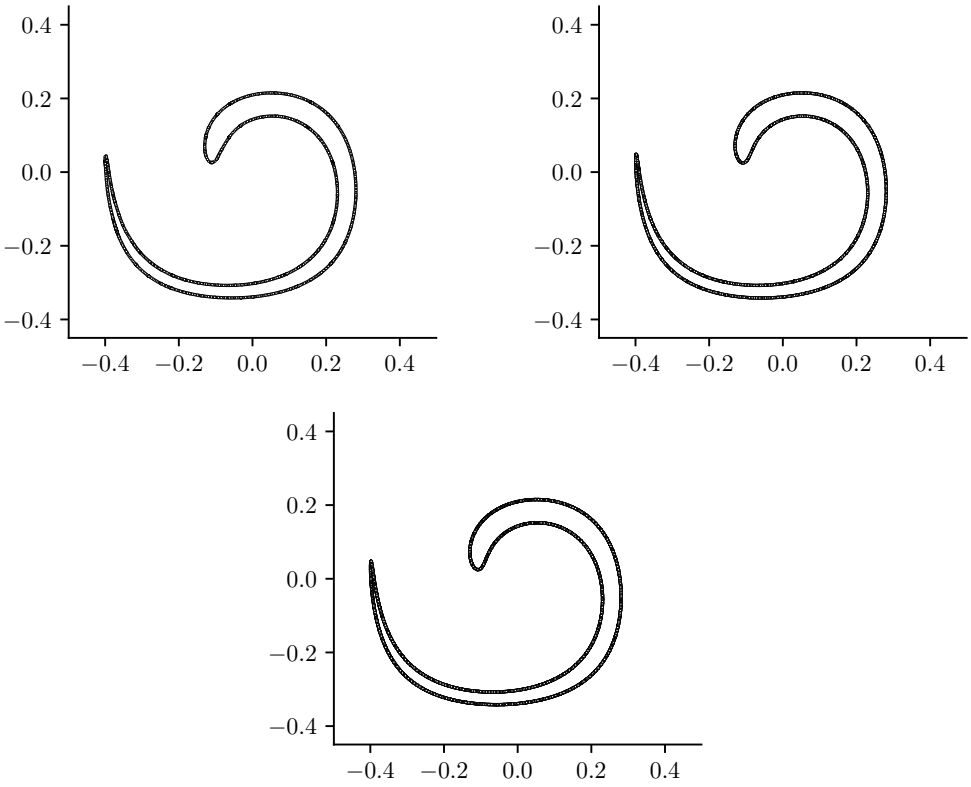

Figure 5.13: Single bubble test with Q_2 velocity interpolation: interface position at maximum deformation ($t = 2\,s$) for the meshes with 32×32 (top left), 64×64 (top right) and 128×128 (bottom) cells.

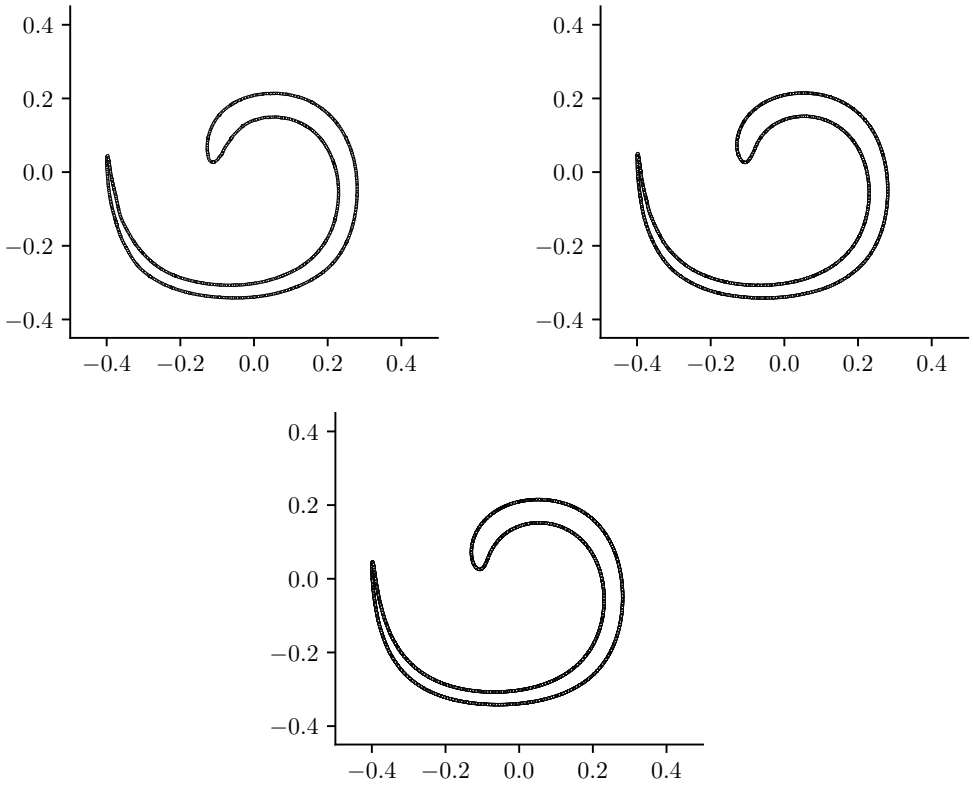

Figure 5.14: Single bubble test with $\boldsymbol{\mathcal{RT}}_0$ velocity interpolation: interface position at maximum deformation ($t = 2\,s$) for the meshes with 32×32 (top left), 64×64 (top right) and 128×128 (bottom) cells.

Conclusions

This dissertation has focused on the challenges posed by mass conservation in incompressible flow simulations, employing a finite element approach, with a particular focus on the use of Raviart-Thomas basis functions. The main goal was to achieve a divergence-free velocity field, crucial for accurate numerical solutions in engineering applications, ranging from multiphase flows to porous-media flows.

In the first chapter, an overview of the problem related to the divergence constraint was explained. After the introduction regarding the notation, several aspects have been presented, ranging from the Helmholtz decomposition to the error analysis for mixed problems, in order to understand the different meanings of the divergence-free constraint.

Subsequently, a wide overview of finite element discretization has been presented. Naturally, the emphasis has been mainly placed on Raviart-Thomas finite elements built over quadrilateral and hexahedral elements. Specifically, this thesis has taken into account only the lowest-order Raviart-Thomas elements. Several approaches described in the literature have been reported to better understand convergence issues related to affine and non-affine finite elements. Both advantages and disadvantages have been described, with a specific focus on the three-dimensional setting.

In Chapter 3 the projection method has been described, with the aim of reducing the computational effort of standard coupled algorithms for the resolution of the Navier-Stokes system. Moreover, the orthogonal decomposition by using Raviart-Thomas elements has been included in a standard split

method to take advantage of a divergence-free representation of the velocity, for the solution of the pressure equation.

In Chapter 4 numerical results concerning the orthogonal velocity decomposition and the employment of the projection method have been shown. Computed convergence rates for velocity and pressure error have been proven to agree with the theoretic discussion. All the presented results have employed computational grids for which convergence of the computed variables can be obtained. Non-affine meshes have also been tested, where naturally an error convergence has not been obtained.

The thesis also extended its focus to multiphase flow simulations, emphasizing the significance of maintaining a divergence-free velocity in scenarios where mass conservation is critical. The novel algorithm introduced for surface advection using marker techniques has shown promising results, particularly in the context of interface tracking within multiphase flow simulations. Specifically, we provide the possibility to advect a marker cloud representing a multiphase interface, and by exploiting a best-fit interpolated quadric it is possible to create a new set of markers that represents the surface in the new position. An interesting comparison has been performed, exploiting specific characteristics of lagrangian and Raviart-Thomas finite elements.

Regarding future challenges, we aim to extend the application of Raviart-Thomas finite elements considering different physical problems, such as electromagnetism and optimal control problems. Additionally, an improvement of the marker library is expected in order to extend the capability of the surface reconstruction, for the analysis of physical multiphase tests such as interface instabilities.

In summary, this dissertation has contributed to the understanding of maintaining divergence-free fields in incompressible flow simulations. The Raviart-Thomas finite element family emerged as a valuable tool in addressing this challenge, with applications ranging from the projection method for the Navier-Stokes equations to the advection of interfaces in multiphase flows.

List of Figures

List of Tables

Bibliography

[1] P.-A. Raviart and J. M. Thomas, "Introduction à l'analyse numérique des équations aux dérivées partielles," 1983. 3

[2] A. Linke, "On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime," *Computer methods in applied mechanics and engineering*, vol. 268, pp. 782–800, 2014. 3

[3] V. Girault and P.-A. Raviart, "Finite element methods for Navier-Stokes equations: theory and algorithms," *NASA STI/Recon Technical Report A*, vol. 87, p. 52227, 1986. 3, 11, 34

[4] M. Fortin and F. Brezzi, *Mixed and hybrid finite element methods*, vol. 2. New York: Springer-Verlag, 1991. 3, 11

[5] H. Elman, D. Silvester, and A. Wathen, *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics*. Oxford university press, 2014. 3

[6] W. Layton, *Introduction to the numerical analysis of incompressible viscous flows*. SIAM, 2008. 3

[7] M. Gunzburger, "The inf-sup condition in mixed finite element methods with application to the Stokes system," *Collected lectures on the preservation of stability under discretization*, pp. 93–121, 2002. 4

[8] V. John, A. Linke, C. Merdon, M. Neilan, and L. G. Rebholz, "On the divergence constraint in mixed finite element methods for incompressible flows," *SIAM review*, vol. 59, no. 3, pp. 492–544, 2017. 4, 14, 34, 40

[9] M. Vogelius, "A right-inverse for the divergence operator in spaces of piecewise polynomials: Application to the p-version of the finite element method," *Numerische Mathematik*, vol. 41, no. 1, pp. 19–37, 1983. 4

[10] J. Qin, *On the convergence of some low order mixed finite elements for incompressible fluids*. The Pennsylvania State University, 1994. 4

[11] B. Cockburn, G. Kanschat, and D. Schötzau, "A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations," *Journal of Scientific Computing*, vol. 31, pp. 61–73, 2007. 4

[12] J. Evans and T. Hughes, "Isogeometric divergence-conforming B-splines for the unsteady Navier–Stokes equations," *Journal of Computational Physics*, vol. 241, pp. 141–167, 2013. 4

[13] R. Falk and M. Neilan, "Stokes complexes and the construction of stable finite elements with pointwise mass conservation," *SIAM Journal on Numerical Analysis*, vol. 51, no. 2, pp. 1308–1326, 2013. 4, 35

[14] J. Wang and X. Ye, "New finite element methods in computational fluid dynamics by H (div) elements," *SIAM Journal on Numerical Analysis*, vol. 45, no. 3, pp. 1269–1286, 2007. 4

[15] S. Zhang, "A new family of stable mixed finite elements for the 3D Stokes equations," *Mathematics of computation*, vol. 74, no. 250, pp. 543–554, 2005. 4

[16] A. Linke, "Collision in a cross-shaped domain: A steady 2d Navier–Stokes example demonstrating the importance of mass conservation in CFD," *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 41-44, pp. 3278–3286, 2009. 4

[17] M. Olshanskii and A. Reusken, "Grad-div stabilization for Stokes equations," *Mathematics of Computation*, vol. 73, no. 248, pp. 1699–1718, 2004. 4

[18] S. Turek, A. Ouazzi, and J. Hron, "On pressure separation algorithms (PSepA) for improving the accuracy of incompressible flow simulations," *International Journal for Numerical Methods in Fluids*, vol. 59, no. 4, pp. 387–403, 2009. 4

[19] A. Linke, "A divergence-free velocity reconstruction for incompressible flows," *Comptes Rendus Mathematique*, vol. 350, no. 17-18, pp. 837–840, 2012. 4

[20] S. Turek, "On discrete projection methods for the incompressible Navier-Stokes equations: An algorithmical approach," *Computer Methods in Applied Mechanics and Engineering*, vol. 143, no. 3-4, pp. 271–288, 1997. 5

[21] A. J. Chorin, "Numerical solution of the Navier-Stokes equations," *Mathematics of computation*, vol. 22, no. 104, pp. 745–762, 1968. 5, 76

[22] R. Temam, *Navier-Stokes equations: theory and numerical analysis*, vol. 343. American Mathematical Soc., 2001. 5, 76

[23] J.-L. Guermond, P. Minev, and J. Shen, "An overview of projection methods for incompressible flows," *Computer methods in applied mechanics and engineering*, vol. 195, no. 44-47, pp. 6011–6045, 2006. 5, 75, 76, 79

[24] E. Aulisa, S. Manservisi, and R. Scardovelli, "A mixed markers and volume-of-fluid method for the reconstruction and advection of interfaces in two-phase and free-boundary flows," *Journal of Computational Physics*, vol. 188, no. 2, pp. 611–639, 2003. 5, 120, 134, 135, 136, 159

[25] E. Aulisa, S. Manservisi, and R. Scardovelli, "A surface marker algorithm coupled to an area-preserving marker redistribution method for three-dimensional interface tracking," *Journal of Computational Physics*, vol. 197, no. 2, pp. 555–584, 2004. 5, 120, 134

[26] R. A. Adams and J. J. Fournier, *Sobolev spaces*. Elsevier, 2003. 7

[27] J. Necas, *Les méthodes directes en théorie des équations elliptiques*. Masson, 1967. 9

[28] P. Grisvard, *Elliptic problems in nonsmooth domains*. SIAM, 2011. 9

[29] V. Hutson, J. Pym, and M. Cloud, *Applications of functional analysis and operator theory*. Elsevier, 2005. 10

[30] P.-A. Raviart and J.-M. Thomas, "A mixed finite element method for 2-nd order elliptic problems," in *Mathematical Aspects of Finite Element Methods: Proceedings of the Conference Held in Rome, December 10–12, 1975*, pp. 292–315, Springer, 2006. 10

[31] J. Shen, *Mixed finite element methods: analysis and computational aspects*. University of Wyoming, 1992. 11, 13

[32] V. Girault and P.-A. Raviart, *Finite element methods for Navier-Stokes equations: theory and algorithms*, vol. 5. Springer Science & Business Media, 2012. 13, 14, 49

[33] F. Brezzi, "On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers," *Publications des séminaires de mathématiques et informatique de Rennes*, no. S4, pp. 1–26, 1974. 13, 14

[34] M. A. Case, V. J. Ervin, A. Linke, and L. G. Rebholz, "A connection between Scott–Vogelius and grad-div stabilized Taylor–Hood FE approximations of the Navier–Stokes equations," *SIAM Journal on Numerical Analysis*, vol. 49, no. 4, pp. 1461–1481, 2011. 18

[35] M. Fortin, "An analysis of the convergence of mixed finite element methods," *RAIRO. Analyse numérique*, vol. 11, no. 4, pp. 341–354, 1977. 21

[36] G. Auchmuty and J. Alexander, "L^2-well-posedness of 3d div-curl boundary value problems," *Quarterly of applied mathematics*, vol. 63, no. 3, pp. 479–508, 2005. 23

[37] D. Boffi, F. Brezzi, M. Fortin, *et al.*, *Mixed finite element methods and applications*, vol. 44. Springer, 2013. 29

[38] P. Ciarlet, *The finite element method for elliptic problems*. SIAM, 2002. 32

[39] J.-M. Thomas, *Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes*. PhD thesis, Université Pierre et Marie Curie, 1977. 32, 52

[40] P.-A. Raviart and J.-M. Thomas, "A mixed finite element method for second order elliptic problems," *Mathematical Aspects of the Finite Element Method, Rome*, 1975. 32

[41] D. Arnold, R. Falk, and R. Winther, "Differential complexes and stability of finite element methods I. The de Rham complex," in *Compatible spatial discretizations*, pp. 23–46, Springer, 2006. 34

[42] D. Arnold, R. Falk, and R. Winther, "Differential complexes and stability of finite element methods II: The elasticity complex," in *Compatible spatial discretizations*, pp. 47–67, Springer, 2006. 34

[43] D. Arnold, R. Falk, and R. Winther, "Finite element exterior calculus, homological techniques, and applications," *Acta numerica*, vol. 15, pp. 1–155, 2006. 34

[44] D. Arnold, R. Falk, and R. Winther, "Finite element exterior calculus: from Hodge theory to numerical stability," *Bulletin of the American Mathematical Society*, vol. 47, no. 2, pp. 281–354, 2010. 34

[45] P. Monk, *Finite element methods for Maxwell's equations*. Oxford University Press, 2003. 34

[46] J. Guzmán and M. Neilan, "Conforming and divergence-free Stokes elements on general triangular meshes," *Mathematics of Computation*, vol. 83, no. 285, pp. 15–36, 2014. 35

[47] K. A. Mardal, X.-C. Tai, and R. Winther, "A robust finite element method for Darcy–Stokes flow," *SIAM Journal on Numerical Analysis*, vol. 40, no. 5, pp. 1605–1631, 2002. 35

[48] D. Arnold, D. Boffi, and R. Falk, "Approximation by quadrilateral finite elements," *Mathematics of computation*, vol. 71, no. 239, pp. 909–922, 2002. 37, 39, 49, 52, 84

[49] D. Arnold, D. Boffi, and R. Falk, "Quadrilateral H (div) finite elements," *SIAM Journal on Numerical Analysis*, vol. 42, no. 6, pp. 2429–2451, 2005. 39, 50, 51, 52, 54

[50] R. Durán, "Mixed finite element methods," *Mixed finite elements, compatibility conditions, and applications*, pp. 1–44, 2008. 40, 42, 44, 45

[51] F. Brezzi, J. Douglas, R. Durán, and M. Fortin, "Mixed finite elements for second order elliptic problems in three variables," *Numerische Mathematik*, vol. 51, pp. 237–250, 1987. 40, 47

[52] F. Brezzi, M. Fortin, L. D. Marini, *et al.*, "Efficient rectangular mixed finite elements in two and three space variables," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 21, no. 4, pp. 581–604, 1987. 40

[53] F. Brezzi, J. J. Douglas, and D. Marini, "Recent results on mixed finite element methods for second order elliptic problems," *Vistas in applied mathematics. Numerical analysis, atmospheric sciences, immunology*, pp. 25–43, 1986. 40

[54] J.-C. Nédélec, "Mixed finite elements in R 3," *Numerische Mathematik*, vol. 35, pp. 315–341, 1980. 47

[55] R. Rannacher and S. Turek, "Simple nonconforming quadrilateral Stokes element," *Numerical Methods for Partial Differential Equations*, vol. 8, no. 2, pp. 97–111, 1992. 48

[56] G. Matthies, "Mapped finite elements on hexahedra. Necessary and sufficient conditions for optimal interpolation errors," *Numerical Algorithms*, vol. 27, pp. 317–327, 2001. 49

[57] J. Shen *et al.*, "Mixed finite element methods on distorted rectangular grids," *Techincal Report ISC*, pp. 94–13, 1994. 52, 57, 58, 60, 62

[58] P. Bochev and D. Ridzal, "Rehabilitation of the lowest-order Raviart–Thomas element on quadrilateral grids," *SIAM Journal on Numerical Analysis*, vol. 47, no. 1, pp. 487–507, 2009. 54, 57, 58, 93

[59] M. Shashkov, *Conservative finite-difference methods on general grids*, vol. 6. CRC press, 1995. 54

[60] J. Hyman and M. Shashkov, "Natural discretizations for the divergence, gradient, and curl on logically rectangular grids," *Computers & Mathematics with Applications*, vol. 33, no. 4, pp. 81–104, 1997. 54

[61] J. Hyman and M. Shashkov, "Adjoint operators for the natural discretizations of the divergence, gradient and curl on logically rectangular grids," *Applied Numerical Mathematics*, vol. 25, no. 4, pp. 413–442, 1997. 54

[62] D. Boffi and L. Gastaldi, "Some remarks on quadrilateral mixed finite elements," *Computers & structures*, vol. 87, no. 11-12, pp. 751–757, 2009. 58, 61

[63] D. Kwak and H. C. Pyo, "Mixed finite element methods for general quadrilateral grids," *Applied Mathematics and Computation*, vol. 217, no. 14, pp. 6556–6565, 2011. 62

[64] A. Bermúdez, P. Gamallo, M. R. Nogueiras, and R. Rodríguez, "Approximation properties of lowest-order hexahedral Raviart–Thomas finite elements," *Comptes Rendus Mathematique*, vol. 340, no. 9, pp. 687–692, 2005. 63, 67

[65] F. Brezzi, K. Lipnikov, M. Shashkov, and V. Simoncini, "A new discretization methodology for diffusion problems on generalized polyhedral meshes," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 37-40, pp. 3682–3692, 2007. 63

[66] J. Jaffré, J. Roberts, and A. Sboui, "Mixed hexahedral finite elements for Darcy flow calculations," in *Proceedings of the XVI International Conference on Computational Methods in Water Resources*, Citeseer, 2006. 63

[67] H. Hægland, H. Dahle, G. Eigestad, K.-A. Lie, and I. Aavatsmark, "Improved streamlines and time-of-flight for streamline simulation on irregular grids," *Advances in Water Resources*, vol. 30, no. 4, pp. 1027–1045, 2007. 63

[68] R. Naff, T. F. Russell, and J. D. Wilson, "Shape functions for velocity interpolation in general hexahedral cells," *Computational Geosciences*, vol. 6, pp. 285–314, 2002. 63, 68

[69] A. Sboui, J. Jaffré, and J. Roberts, *A composite mixed finite elements for general hexahedral grids*. PhD thesis, INRIA, 2007. 63, 68

[70] J. Nordbotten and H. Hægland, "On reproducing uniform flow exactly on general hexahedral cells using one degree of freedom per surface," *Advances in Water Resources*, vol. 32, no. 2, pp. 264–267, 2009. 63

[71] R. Falk, P. Gatto, and P. Monk, "Hexahedral H (div) and H (curl) finite elements," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 45, no. 1, pp. 115–143, 2011. 63, 67, 69

[72] T. Arbogast and Z. Tao, "Construction of H (div)-conforming mixed finite elements on cuboidal hexahedra," *Numerische Mathematik*, vol. 142, no. 1, pp. 1–32, 2019. 71

[73] R. Temam, "Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires (II)," *Archive for rational mechanics and analysis*, vol. 33, pp. 377–385, 1969. 76

[74] J.-L. Guermond and L. Quartapelle, "On the approximation of the unsteady Navier–Stokes equations by finite element projection methods," *Numerische mathematik*, vol. 80, pp. 207–238, 1998. 77

[75] K. Goda, "A multistep technique with implicit difference schemes for calculating two-or three-dimensional cavity flows," *Journal of computational physics*, vol. 30, no. 1, pp. 76–95, 1979. 77

[76] J. Van Kan, "A second-order accurate pressure-correction scheme for viscous incompressible flow," *SIAM journal on scientific and statistical computing*, vol. 7, no. 3, pp. 870–891, 1986. 77

[77] L. Timmermans, P. Minev, and F. Van De Vosse, "An approximate projection scheme for incompressible flow using spectral elements," *International journal for numerical methods in fluids*, vol. 22, no. 7, pp. 673–688, 1996. 78

[78] S. Orszag, M. Israeli, and M. Deville, "Boundary conditions for incompressible flows," *Journal of Scientific Computing*, vol. 1, pp. 75–111, 1986. 79

[79] G. Karniadakis, M. Israeli, and S. Orszag, "High-order splitting methods for the incompressible Navier-Stokes equations," *Journal of computational physics*, vol. 97, no. 2, pp. 414–443, 1991. 79

[80] G. Barbi, G. Bornia, D. Cerroni, A. Cervone, A. Chierici, L. Chirco, R. Viá, V. Giovacchini, S. Manservisi, R. Scardovelli, *et al.*, "Femus-platform: A numerical platform for multiscale and multiphysics code coupling," in *9th International Conference on Computational Methods for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2021*, pp. 1–12, International Center for Numerical Methods in Engineering, 2021. 80, 91

[81] "Femusplatform." https://github.com/FEMuSPlatform. 80

[82] G. Barbi, A. Cervone, A. Chierici, V. Giovacchini, S. Manservisi, R. Scardovelli, L. Sirotti, *et al.*, "A new projection method for navier-stokes equations by using raviart-thomas finite element," in *World Congress in Computational Mechanics and ECCOMAS Congress*, pp. 1–12, Scipedia SL, 2022. 81

[83] E. Aulisa, S. Bná, and G. Bornia, "Femus." https://github.com/eaulisa/MyFEMuS.git, 2014. 120, 143

[84] A. Cervone, "ProXPDE." https://github.com/capitalaslash/proxpde.git, 2015. 120, 143

[85] F. Chen and H. Hagen, "A survey of interface tracking methods in multiphase fluid visualization," in *Visualization of Large and Unstructured Data Sets-Applications in Geospatial Planning, Modeling and Engineering (IRTG 1131 Workshop)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2011. 120

[86] C. Hirt and B. Nichols, "Volume of fluid (VOF) method for the dynamics of free boundaries," *Journal of computational physics*, vol. 39, no. 1, pp. 201–225, 1981. 120

[87] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988. 120

[88] S. O. Unverdi and G. Tryggvason, "A front-tracking method for viscous, incompressible, multi-fluid flows," *Journal of computational physics*, vol. 100, no. 1, pp. 25–37, 1992. 120

[89] R. Goldman, "Curvature formulas for implicit curves and surfaces," *Computer Aided Geometric Design*, vol. 22, no. 7, pp. 632–658, 2005. 122

[90] J. J. Stoker, *Differential geometry*, vol. 20. John Wiley & Sons, 2011. 122

[91] E. Aulisa and J. Loftin, "Exact subdomain and embedded interface polynomial integration in finite elements with planar cuts," *Numerical Algorithms*, pp. 1–36, 2023. 124

[92] W. Vetterling, W. Press, S. Teukolsky, and B. Flannery, *Numerical recipes example book (c++): The art of scientific computing*. Cambridge University Press, 2002. 125

[93] G. Stewart, "On the early history of the singular value decomposition," *SIAM review*, vol. 35, no. 4, pp. 551–566, 1993. 127

[94] G. Guennebaud, B. Jacob, *et al.*, "Eigen v3." http://eigen.tuxfamily.org, 2010. 128

[95] M. Lawiyuniarti, E. Rahmadiantri, I. Alamsyah, and G. Rachmaputri, "Application of least-squares fitting of ellipse and hyperbola for two dimensional data," in *Journal of Physics: Conference Series*, vol. 948, p. 012069, IOP Publishing, 2018. 128

[96] B. Lafaurie, C. Nardone, R. Scardovelli, S. Zaleski, and G. Zanetti, "Modelling merging and fragmentation in multiphase flows with SURFER," *Journal of computational physics*, vol. 113, no. 1, pp. 134–147, 1994. 134

[97] D. Harvie and D. Fletcher, "A new volume of fluid advection algorithm: the stream scheme," *Journal of Computational Physics*, vol. 162, no. 1, pp. 1–32, 2000. 134

[98] R. J. Leveque, "High-resolution conservative algorithms for advection in incompressible flow," *SIAM Journal on Numerical Analysis*, vol. 33, no. 2, pp. 627–665, 1996. 134