UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN

Métodos de elementos finitos mixtos basado en pseudoesfuerzo conservativo de momentum y de masa para problemas de flujo de fluidos (Momentum and mass conservative pseudostress-based mixed finite element methods for fluid flow problems)

Tesis para optar al grado de
Magíster en Matemática Aplicada

KATHERINE PAOLA ROJO CONTRERAS

Profesora guía: Dra. Jessika Camaño Valenzuela
Cotutor: Dr. Ricardo Oyarzúa Vargas

FACULTAD DE INGENIERÍA
Departamento de Matemática y Física Aplicadas
Concepción - Chile
Abril - 2025

Momentum and mass conservative finite element methods for fluid flow problems

por

KATHERINE PAOLA ROJO CONTRERAS

Tesis presentada al Programa Magíster en Matemática Aplicada de la Facultad de Ingeniería de la Universidad Católica de la Santísima Concepción como parte de los requisitos para la obtención del grado de Magíster en Matemática Aplicada.

Esta tesis ha sido aprobada por:

Prof. Guía:	Dra. Jessika Camaño Valenzuela Universidad Católica de la Santísima Concepción, Chile.
Prof. Cotutor:	Dr. Ricardo Oyarzúa Vargas, Universidad del Bío-Bío, Chile.

Comisión Examinadora:

Evaluador Interno:	Dr. Sergio Caucao Paillán Universidad Católica de la Santísima Concepción, Chile.
Evaluador Externo:	Dr. Felipe Lepe Araya Universidad
Jefe de Programa:	Dr. Abner Poza Díaz Universidad Católica de la Santísima Concepción, Chile.

Concepción, Abril de 2025.

AGRADECIMIENTOS

Quiero empezar dando gracias a Dios por las oportunidades brindadas, por darme inteligencia, sabiduría y sobre todo la paciencia y resiliencia necesaria para iniciar y culminar esta etapa.

Agradezco también a mi directora de tesis, Jessika Camaño Valenzuela, por permitirme trabajar con ella, por todas sus enseñanzas, por su paciencia y sobre todo por su hospitalidad y buena disposición a la hora de resolver cada una de mis inquietudes, gracias por siempre tener una palabra de aliento. De igual manera, agradezco a mi codirector, Ricardo Oyarzúa Vargas, por aceptar trabajar conmigo y sumar su gran conocimiento a la presente investigación. Mi cariño y admiración para ustedes.

Gracias a mi familia, quienes en todo momento me han brindado su apoyo y amor incondicional.

A mis profesores del programa de Magíster en Matemática Aplicada, gracias por compartir sus conocimientos. En especial, al profesor Sergio Caucao, por su disposición, motivación y las interesantes conversaciones, y a la profesora Johanna García Saldaña, por estar siempre pendiente de mí y hacer espacio en su agenda para momentos de esparcimiento.

A mis compañeros y amigos del magíster, gracias por brindarme su amistad, en especial a Eduardo, con quien inicié este camino; y a Martín por ser mi sí a todo. También agradezco a Jesús y Adrián, quienes fueron apareciendo en el camino, infinitas gracias por soportar los días en los que no recordaba qué día era, esto sin ustedes hubiera sido más difícil.

Gracias a Patricia por toda la colaboración brindada siempre. Nunca había conocido una secretaria tan eficiente y amable.

Agradezco especialmente a la Unidad de Postgrado de la Universidad Católica de la Santísima Concepción por el financiamiento brindado mediante las becas de arancel y manutención.

Finalmente, y no menos importante, agradezco a ANID-Chile, por su apoyo a través de los proyectos Fondecyt Regular 1231336, Centro de Modelamiento Matemático (FB210005) y Anillo de Matemática Computacional para Procesos de Desalinización

(ACT210087). También mis agradecimientos a GIANuC^2 mediante la DI-UCSC con el proyecto FGII 06/2024.

Resumen

En esta tesis, proponemos y analizamos un nuevo método de elementos finitos mixtos basado en pseudoesfuerzo para los problemas de Stokes y Navier–Stokes, que permite conservación exacta de masa y momentum. En ambos casos, usamos una descomposición de Helmholtz para la velocidad y derivamos una formulación variacional mixta de tres campos, donde el pseudoesfuerzo, la velocidad y una incógnita adicional que representa la función nula son las principales incógnitas del sistema.

Para el problema de Stokes, demostramos que el método está bien planteado y obtenemos tasas de convergencia teóricas, incluyendo un resultado de superconvergencia para la aproximación del gradiente de la velocidad. Una ventaja clave del método propuesto es su eficiencia computacional, ya que resulta ligeramente menos costoso que el enfoque clásico basado en pseudoesfuerzo estudiado en [11, 26], y al mismo tiempo asegura la conservación de masa y cantidad de movimiento. Además, extendemos nuestro análisis al problema de Stokes con condiciones de borde mixtas.

Para el problema de Navier–Stokes, a diferencia del caso de Stokes, es necesario utilizar una descomposición de Helmholtz para la velocidad, pero en espacios de Banach, lo cual solo es válido en dos dimensiones. El análisis de los problemas continuo y discreto se lleva a cabo utilizando el teorema de Banach–Nečas–Babuška y el teorema del punto fijo de Banach, bajo el supuesto de datos suficientemente pequeños. También derivamos la correspondiente estimación a priori del error y proporcionamos la tasa de convergencia teórica. Otras variables de interés, como la presión del fluido, la vorticidad y el gradiente de velocidad del fluido, pueden aproximarse fácilmente mediante un simple postprocesamiento de las soluciones del método de elementos finitos, con la misma tasa de convergencia.

Para ambos problemas, presentamos varios ejemplos numéricos que validan los resultados teóricos, demostrando la efectividad y precisión del método propuesto, el

cual ofrece varias ventajas, incluyendo facilidad de implementación y compatibilidad con paquetes de software existentes para la resolución de ecuaciones en derivadas parciales.

Abstract

In this thesis, we propose and analyze a new pseudosstres-based mixed finite element method for the Stokes and Navier–Stokes problems allowing exact conservation of mass and momentum. In both cases, we decompose the velocity using a Helmholtz decomposition and derive a three-field mixed variational formulation, where the pseudostress, the velocity, and an additional unknown representing the null function are the main unknowns of the system.

For the Stokes problem, we establish that the method is well-posed and obtain theoretical convergence rates, including a superconvergence result for the approximation of the velocity gradient. A key advantage of the proposed method is its computational efficiency, as it is slightly less expensive than the classical pseudostress-based approach studied in [11, 26], while ensuring mass and momentum conservation. Additionally, we extend our analysis to the Stokes problem with mixed boundary conditions.

For the Navier–Stokes problem, and unlike the Stokes problem, it is necessary to use a Helmholtz decomposition to decompose the velocity, but in Banach spaces, which is valid only in two dimensions. The analysis of the continuous and discrete problems is carried out using the Banach–Nečas–Babuška theorem and Banach's fixed-point theorem, under the assumption of sufficiently small data. We also derive the corresponding a priori error estimate and provide the theoretical convergence rate. Other variables of interest, such as fluid pressure, fluid vorticity, and the fluid velocity gradient, can be easily approximated as a simple post-processing of the finite element solutions, with the same convergence rate.

For both problems we provide several numerical examples to validate the theoretical results, demonstrating the effectiveness and accuracy of the proposed method, which offers several advantages, including ease of implementation and compatibility

with existing software packages for solving partial differential equations.

Contenido

Introducción

El problema de Navier–Stokes (NS) tiene una importancia fundamental en el campo de la dinámica de fluidos. Como es bien sabido, sus ecuaciones describen el movimiento de fluidos viscosos y gobiernan una amplia gama de fenómenos, como la atmósfera terrestre, las corrientes oceánicas, el flujo alrededor de vehículos y proyectiles y, de forma más general, cualquier proceso relacionado con los fluidos. Este problema ha encontrado amplias aplicaciones en diversos campos científicos y de ingeniería, como la aerodinámica, la meteorología, la oceanografía, el diseño de sistemas de ventilación e inclusive como se puede ver en [38] se han utilizado ampliamente en videojuegos para modelar muchos fenómenos naturales. La complejidad matemática del problema de NS, hace que no sea posible hallar su solución analítica, salvo ciertos tipos de flujo y situaciones muy concretas, lo que lleva a la necesidad de desarrollar métodos numéricos eficientes para obtener soluciones aproximadas. Por esta razón, la comunidad de análisis numérico, lleva décadas desarrollando métodos numéricos para aproximar la solución de NS (ver por ejemplo [28] y [39]).

Entre los métodos numéricos utilizados para simular flujos gobernados por las ecuaciones de NS, los métodos de elementos finitos destacan por su flexibilidad y precisión en la aproximación de soluciones a problemas complejos con geometrías diversas y condiciones de contorno variadas. Estos métodos han sido objeto de un amplio desarrollo teórico y práctico, abordando desafíos clave como la estabilidad, la convergencia y la precisión de las soluciones obtenidas. En cuanto a las formulaciones mixtas para las ecuaciones de NS, los trabajos de Farhloul et al. (en [20] y [19]) destacan por extender el análisis de formulaciones mixtas duales desde las ecuaciones de Stokes al problema de NS. En [20], los autores introducen el tensor de deformación, y en [19], definen el tensor del gradiente de velocidad como las principales incógnitas de los sistemas correspondientes, proponiendo métodos

numéricos cuasi-óptimos para el problema de flujo de fluidos. Por su parte, Cai et al., en [8, 10, 9], han extendido el análisis de los métodos mixtos basados en el pseudoesfuerzo para el problema de Stokes al problema de NS. Introducen y analizan un método conforme para una formulación mixta basada en el pseudoesfuerzo, obteniendo también un esquema numérico cuasi-óptimo. Posteriormente, Howell y Walkington, en [33], introducen un nuevo método de elementos finitos mixtos dual antisimétrico para el problema en estudio, considerando como incógnitas principales el gradiente de velocidad (en L^2), la velocidad (en L^2) y un tensor de pseudoesfuerzo modificado (en $\mathbf{H}(\mathrm{div})$) que vincula el gradiente de velocidad y la presión con el término convectivo. Además, dado que el análisis se centra en condiciones inf-sup no estándar, los autores proponen nuevas familias de elementos finitos, obtenidas mediante el enriquecimiento de familias bien conocidas diseñadas para problemas de elasticidad, como los elementos de Arnold–Falk–Winther y PEERS. Usando estos espacios, se puede demostrar una convergencia óptima, aunque con un costo computacional relativamente alto. Un enfoque similar se presenta en [13], donde se introduce un método de elementos finitos mixtos aumentado para las ecuaciones de NS. En este método, de manera similar a [33], se define un tensor de pseudoesfuerzo no estándar que relaciona el gradiente de velocidad con el término convectivo. Este pseudoesfuerzo (en $\mathbf{H}(\mathrm{div})$), junto con la velocidad en H^1, son las únicas incógnitas del sistema, mientras que la presión y otras variables de interés pueden recuperarse mediante un procedimiento de post-proceso. La solución de los problemas continuo y discreto, así como la demostración de la convergencia óptima, se logran mediante la incorporación de términos de mínimos cuadrados en la formulación. Este enfoque les permite evitar la necesidad de demostrar condiciones inf-sup y, como resultado, relajar las hipótesis sobre los subespacios discretos correspondientes. Sin embargo, este procedimiento incrementa significativamente la complejidad y el costo computacional.

En cuanto a los métodos conservativos, podemos destacar [30], donde los autores introducen una familia de elementos finitos conformes y a divergencia nula para el problema de Stokes en mallas triangulares generales en dos dimensiones, garantizando la conservación de la masa en la formulación conforme velocidad-presión. La propiedad de divergencia nula se logra mediante el enriquecimiento adecuado del espacio polinómico para la velocidad, lo que hace que la implementación computa-

cional sea más desafiante y aumente el costo computacional. La extensión de este trabajo al caso tridimensional se puede encontrar en [31].

Otra posibilidad para obtener esquemas numéricos conservativos es reformular las ecuaciones. En [17] (véase también [2]), los autores introducen la vorticidad como una incógnita adicional, obteniendo así una formulación variacional para el problema de Stokes con la velocidad en $\mathbf{H}(\mathrm{div})$. Este enfoque permite que la velocidad sea aproximada mediante elementos finitos conformes en $\mathbf{H}(\mathrm{div})$, obteniendo así un esquema numérico conforme y conservativo de masa. Sin embargo, este enfoque no se ha extendido al problema de NS, ya que el término convectivo impide el uso de los mismos espacios para las variables que en [17].

Por otro lado, podemos mencionar [27], donde el gradiente de velocidad se introduce como una incógnita adicional, permitiendo que la velocidad sea aproximada mediante elementos finitos conformes en $\mathbf{H}(\mathrm{div})$ y logrando así la conservación exacta de la masa. Similar a [17], el término convectivo dificulta la extensión de este enfoque al problema de NS.

Por su parte, en [12], los autores han introducido y analizado un método de elementos finitos mixtos conservativo de momentum para el problema estacionario de NS con viscosidad constante, formulado en espacios de Banach. Derivan una formulación mixta en la que el tensor de pseudoesfuerzo, introducido en [33], y la velocidad son las incógnitas principales del sistema. Se emplean elementos de Raviart–Thomas de grado k y elementos polinomiales discontinuos a trozos de grado k para aproximar el tensor de pseudoesfuerzo y la velocidad, respectivamente. Con esta elección de espacios, la ecuación de equilibrio se satisface exactamente cuando la fuerza externa pertenece al espacio discreto de velocidades, asegurando así la conservación del momentum.

Recientemente, en [14], los autores han propuesto una reformulación de la formulación basada en pseudoesfuerzo para el problema de Stokes (consúltese [11], [24], [25] y [26] para más detalles), obteniendo un método de elementos finitos conforme basado en pseudoesfuerzo que conserva masa para resolver el problema de Stokes. Descomponen la velocidad mediante una descomposición de Helmholtz, obteniendo una formulación variacional mixta en la que el pseudoesfuerzo, la velocidad y una incógnita adicional son las principales incógnitas del sistema. Para obtener un problema discreto bien planteado en [14], es necesario incorporar términos adecuados

en la formulación, los cuales surgen de las mismas ecuaciones.

Motivados por lo anterior, en el Capítulo 1 comenzamos el análisis para el problema de Stokes. Introducimos la formulación variacional y el esquema discreto, analizamos el buen planteamiento tanto del problema continuo como del discreto, y obtenemos las tasas teóricas de convergencia. Finalmente, presentamos cuatro ejemplos numéricos que ilustran el desempeño del método. Luego, en el Capítulo 2, extendemos el análisis desarrollado en el Capítulo 1 al problema de NS formulado en espacios de Banach. El análisis de los problemas continuo y discreto, al igual que en [12], se lleva a cabo utilizando el teorema de Banach–Nečas–Babuška y el teorema del punto fijo de Banach, bajo el supuesto de datos suficientemente pequeños. Este riguroso marco matemático garantiza la estabilidad y la convergencia del método. A su vez, obtenemos las correspondientes estimaciones a priori del error y demostramos la convergencia óptima de nuestro método, lo cual es confirmado mediante pruebas numéricas. A diferencia de [14], en lugar de enriquecer la formulación con términos adicionales para establecer el buen planteamiento del problema discreto, adoptamos una descomposición de Helmholtz discreta, evitando así el procedimiento que incrementa el costo computacional.

Es importante enfatizar que, en el caso de Stokes, todo el análisis se ha realizado en d dimensiones, con $d \in \{2, 3\}$; mientras que, para NS, solo se ha desarrollado en dos dimensiones. Esto se debe a que la descomposición de Helmholtz de L^p sobre dominios Lipschitz sólo es válida para rangos específicos de p, los cuales varían según la dimensión considerada. Desafortunadamente, en tres dimensiones L^4 no puede descomponerse de la misma manera que en dos dimensiones. Este problema representa un tema de investigación futura.

Introduction

The Navier–Stokes (NS) problem is of fundamental importance in the field of fluid dynamics. As is well known, its equations describe the motion of viscous fluids and govern a broad range of phenomena, including Earth's atmosphere, ocean currents, flow around vehicles and projectiles, and, more generally, any fluid-related processes. This problem has found widespread applications in various scientific and engineering fields, such as aerodynamics, meteorology, oceanography, ventilation system design, and, as noted in [38], even in video games to model natural phenomena. The mathematical complexity of the NS problem makes it impossible to obtain an analytical solution in most cases, except for specific types of flow or highly constrained scenarios. As a result, there has long been a need to develop efficient numerical methods for approximating solutions. Consequently, the numerical analysis community has been working for decades to develop methods that can accurately approximate the NS solution (see for instance [28] and [39]).

Among the numerical methods used to simulate flows governed by the NS equations, finite element methods stand out for their flexibility and accuracy in approximating solutions to complex problems with diverse geometries and boundary conditions. These methods have been the subject of extensive theoretical and practical development, addressing key challenges such as stability, convergence, and the accuracy of the resulting solutions. Regarding mixed formulations for the NS equations, the works of Farhloul et al. (in [20] and [19]) stand out for extending the analysis of mixed dual formulations from the Stokes equations to the NS problem. In [20], the authors introduce the strain tensor, and in [19], they define the velocity gradient tensor as the main unknowns of the corresponding systems, proposing quasi-optimal numerical methods for the fluid flow problem. For their part, Cai et al. in [8, 10, 9], have extended the analysis of pseudostress-based mixed methods for

the Stokes problem to the NS problem. They introduce and analyze a conforming method for a pseudostress-based mixed formulation, also obtaining a quasi-optimal numerical scheme. Subsequently, Howell and Walkington in [33], introduce a new skew-symmetric dual-mixed finite element method for the problem under study, considering the velocity gradient (in L^2), the velocity (in L^2) and a modified pseudostress tensor (in $\mathbf{H}(\mathrm{div})$) that links the velocity gradient and pressure to the convective term, as the main unknowns of the system. Additionally, since the analysis focuses on nonstandard inf-sup conditions, the authors propose new families of finite elements, derived by enriching well-known families designed for elasticity problems, such as the Arnold–Falk–Winther and PEERS elements. Using these spaces, optimal convergence can be demonstrated, albeit at a relatively high computational cost. A similar approach is presented in [13], where an augmented mixed finite element method for the NS equations is introduced. In this method, similar to [33], a non-standard pseudostress tensor is defined that relates the velocity gradient to the convective term. This pseudostress (in $\mathbf{H}(\mathrm{div})$), along with the velocity in H^1, are the only unknowns in the system, while the pressure and other variables of interest can be recovered through a post-processing procedure. The solution to both the continuous and discrete problems, as well as the proof of optimal convergence, are achieved through the incorporation of least-squares terms into the formulation. This approach allows them to circumvent the necessity of proving inf-sup conditions, and as a result, to relax the hypotheses on the corresponding discrete subspaces. However, this procedure increases the complexity and computational cost significantly.

Concerning conservative methods, we can highlight [30] where the authors introduce a family of divergence-free conforming finite elements for the Stokes problem on general triangular meshes in two dimensions, ensuring mass conservation in the conforming velocity-pressure formulation. The divergence-free property is achieved by appropriately enriching the polynomial space for the velocity, which makes the computational implementation more challenging and increases the computational cost. We can find the extension of this work to the three-dimensional case in [31]. Another possibility to obtain conservative numerical schemes is to reformulate the equations. In [17] (see also [2]), the authors introduce the vorticity as an additional unknown, thus obtaining a variational formulation for the Stokes problem with the velocity in $\mathbf{H}(\mathrm{div})$. This approach allows the velocity to be approximated using

$\mathbf{H}$(div)-conforming finite elements, thereby obtaining a conforming mass conservative numerical scheme. However, this approach has not been extended to the NS problem, as the convective term prevents the use of the same spaces for the variables as in [17]. On the other hand, we can mention [27], where the velocity gradient is introduced as an additional unknown, and the velocity can be approximated using $\mathbf{H}$(div)-conforming finite elements, thereby achieving exact mass conservation. Similar to [17], the convective term hinders the extension of this approach to the NS problem. For their part, in [12] the authors have introduced and analyzed a momentum conservative mixed finite element method for the stationary NS problem with constant viscosity, posed in Banach spaces. They derive a mixed formulation in which the pseudostress tensor, as introduced in [33], and the velocity are the main unknowns of the system. Raviart–Thomas elements of degree k and discontinuous piecewise polynomial elements of degree k are used to approximate the pseudostress tensor and the velocity, respectively. With this choice of spaces, the equilibrium equation is exactly satisfied when the external force belongs to the discrete velocity space, thereby ensuring momentum conservation. Recently, in [14] the authors have proposed a reformulation of the pseudostress-based formulation for the Stokes problem (refer to [11], [24], [25] and [26] for details) obtaining a pseudostress-based conformal finite element method which conserves mass to solve the Stokes problem. They decompose the velocity by means of a Helmholtz decomposition obtaining a mixed variational formulation, where the pseudostress, the velocity and an additional unknown are the main unknowns of the system. In order to obtain a well-posed discrete problem in [14], it is necessary to incorporate suitable terms in the formulation which arise from the same equations.

Motivated by the above in Chapter 1 we begin the analysis for the Stokes problem. We introduce the variational formulation, the discrete scheme, we analyze well-posedness of both, the continuous and discrete problems and derive the theoretical rates of convergence. Finally, we provide 4 numerical examples that illustrate the perfomance of the method. Next, in Chapter 2 we extende the analysis provided in Chapter 1 to the NS problem, posed in Banach spaces. The analysis of the continuous and discrete problems, as in [12], is performed using Banach–Nečas–Babuška and Banach's fixed-point theorem, under the assumptions of sufficiently small data. This rigorous mathematical framework guarantees the stability and convergence of

the method. In turn, we derive the corresponding error estimates and prove the optimal convergence of our method, which is confirmed by the respective numerical tests. Unlike [14], instead of enriching the formulation with additional terms to establish the well-posedness of the discrete problem, we adopt a discrete Helmholtz decomposition, thereby avoiding the procedure that increases computational cost.

It is important to emphasize that in the case of Stokes, all the analysis has been carried out in d dimensions, with $d \in \{2, 3\}$; whereas in NS it has only been developed in two dimensions. This is because the Helmholtz decomposition of L^p into Lipschitz domains is valid only for specific ranges of p, which vary depending on the dimension considered. Unfortunately, L^4 cannot be decomposed in three dimensions as it can be done in two dimensions. This issue represents a topic for future research.

Chapter 1

A momentum and mass conservative pseudostress-based mixed finite element method for the Stokes problem

In this chapter, we propose and analyze a mass- and momentum-conservative numerical scheme for a pseudostress-based formulation of the Stokes problem.

1.1 Preliminaries

To provide more details on our approach, we begin by recalling the classical Stokes problem, governed by the following system of partial differential equations:

$$-\nu\Delta\mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega, \quad \operatorname{div}\mathbf{u} = 0 \quad \text{in } \Omega, \quad \mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma, \quad \int_\Omega p = 0. \quad (1.1.1)$$

This system describes the motion of an incompressible fluid with velocity $\mathbf{u} = (u_1, \ldots, u_d)^t$, pressure p, and viscosity $\nu > 0$ in a region $\Omega \subseteq \mathbb{R}^d$ $(d = 2, 3)$, subjected to a source force $\mathbf{f} = (f_1, \ldots, f_d)^t$ and a prescribed velocity $\mathbf{u}_D = (u_{D,1}, \ldots, u_{D,d})^t$ on the boundary $\Gamma := \partial\Omega$, satisfying the compatibility condition:

$$\int_\Gamma \mathbf{u}_D \cdot \mathbf{n} = 0. \quad (1.1.2)$$

Here, $\mathbf{n} = (n_1, \ldots, n_d)^t$ denotes the outward unit normal vector on Γ.

In [11] and [26] (see also [24] for a similar approach), the so-called pseudostress tensor given by:

$$\boldsymbol{\sigma} := \nu \nabla \mathbf{u} - p\,\mathbb{I} \quad \text{in } \Omega, \tag{1.1.3}$$

is introduced to reformulate (1.1.1) as follows:

$$\boldsymbol{\sigma}^{\mathrm{d}} = \nu \nabla \mathbf{u} \quad \text{in } \Omega, \quad -\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{f} \quad \text{in } \Omega, \quad \mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma, \quad \int_\Omega \mathrm{tr}\,(\boldsymbol{\sigma}) = 0, \tag{1.1.4}$$

where $\mathbb{I}$ denotes the identity matrix, $\mathrm{tr}\,(\boldsymbol{\sigma})$ is the trace of the tensor $\boldsymbol{\sigma}$, $\boldsymbol{\sigma}^{\mathrm{d}} := \boldsymbol{\sigma} - \frac{1}{d}\mathrm{tr}\,(\boldsymbol{\sigma})\mathbb{I}$ denotes the deviatoric part of $\boldsymbol{\sigma}$, and $\mathbf{div}\,\boldsymbol{\tau}$ is the divergence operator div acting along the rows of $\boldsymbol{\tau}$ for any tensor field $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,d}$.

Based on (1.1.4), the works [11] and [26] study conforming numerical discretizations for the following variational problem: Find $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$ and $\mathbf{u} \in \mathbf{L}^2(\Omega) = [\mathrm{L}^2(\Omega)]^d$ such that

$$\begin{aligned} \frac{1}{\nu}(\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega + (\mathbf{u}, \mathbf{div}\,\boldsymbol{\tau})_\Omega &= \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D \rangle_\Gamma\,, \quad \forall \boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}\,;\Omega), \\ (\mathbf{v}, \mathbf{div}\,\boldsymbol{\sigma})_\Omega &= -(\mathbf{f}, \mathbf{v})_\Omega, \qquad \forall \mathbf{v} \in \mathbf{L}^2(\Omega), \end{aligned} \tag{1.1.5}$$

where, for simplicity, we use the following notation:

$$(v, w)_\Omega := \int_\Omega vw, \quad (\mathbf{v}, \mathbf{w})_\Omega := \int_\Omega \mathbf{v} \cdot \mathbf{w}, \quad (\boldsymbol{\tau}, \boldsymbol{\zeta})_\Omega := \sum_{i,j=1}^{d} \int_\Omega \tau_{ij}\zeta_{ij},$$

for any scalars v, w, vectors $\mathbf{v} = (v_i)_{i=1,d}$, $\mathbf{w} = (w_i)_{i=1,d}$, and tensor fields $\boldsymbol{\zeta} = (\zeta_{ij})_{i,j=1,d}$, $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1,d}$. In addition, $\langle \cdot, \cdot \rangle_\Gamma$ denotes the duality pairing between the trace space $\mathbf{H}^{1/2}(\Gamma)$ and its dual $\mathbf{H}^{-1/2}(\Gamma)$, which coincides with the $\mathrm{L}^2(\Gamma)$-inner product when applied to functions in $\mathrm{L}^2(\Gamma)$.

Here, $\mathbb{H}(\mathbf{div}\,;\Omega)$ denotes the space of tensors whose rows belong to

$$\mathbf{H}(\mathrm{div};\Omega) := \{\mathbf{v} \in \mathbf{L}^2(\Omega) : \mathrm{div}\,\mathbf{v} \in \mathrm{L}^2(\Omega)\},$$

and $\mathbb{H}_0(\mathbf{div}\,;\Omega)$ is the subspace of $\mathbb{H}(\mathbf{div}\,;\Omega)$ given by

$$\mathbb{H}_0(\mathbf{div}\,;\Omega) := \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}\,;\Omega) : (\mathrm{tr}\,(\boldsymbol{\tau}), 1)_\Omega = 0\}.$$

In [11] and [26], it is shown that selecting Raviart–Thomas elements of degree $k \geq 0$ or Brezzi–Douglas–Marini (BDM) elements of order $k+1$ for $\mathbb{H}(\mathbf{div}\,;\Omega)$ and piecewise polynomials of degree k for $\mathbf{L}^2(\Omega)$ leads to a well-posed and optimally convergent conforming discretization of (1.1.5). In addition, from the second equation of (1.1.5) it can be seen that the equilibrium equation $\mathbf{div}\,\boldsymbol{\sigma} = -\mathbf{f}$ is exactly satisfied if $\mathbf{f}$ belongs to the same discrete space as the velocity. Consequently, the method preserves momentum. However, since the condition $\operatorname{div}\mathbf{u} = 0$ in Ω cannot be ensured at the discrete level, the method lacks mass conservation (see Example 4 in Section 1.5). To overcome this lack of mass conservation, in [14] is introduced the following variational formulation based on (1.1.4): Find $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$, $\mathbf{u} \in \mathbf{H}(\operatorname{div}^0;\Omega)$ and $\varphi \in \mathrm{H}^1_0(\Omega)$, such that

$$\begin{aligned} \frac{1}{\nu}(\boldsymbol{\sigma}^{\mathrm{d}},\boldsymbol{\tau}^{\mathrm{d}})_\Omega + \frac{1}{\nu}(\mathbf{div}\,\boldsymbol{\sigma},\mathbf{div}\,\boldsymbol{\tau})_\Omega + (\mathbf{u}+\nabla\varphi,\mathbf{div}\,\boldsymbol{\tau})_\Omega &= \langle \boldsymbol{\tau}\mathbf{n},\mathbf{u}_D\rangle_\Gamma - \frac{1}{\nu}(\mathbf{f},\mathbf{div}\,\boldsymbol{\tau})_\Omega, \\ (\mathbf{v}+\nabla\psi,\mathbf{div}\,\boldsymbol{\sigma})_\Omega &= -(\mathbf{f},\mathbf{v}+\nabla\psi)_\Omega, \end{aligned} \tag{1.1.6}$$

for all $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$ and $(\mathbf{v},\psi) \in \mathbf{H}(\operatorname{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega)$, where

$$\mathrm{H}^1_0(\Omega) := \{\psi \in \mathrm{H}^1(\Omega) : \psi = 0 \text{ on } \Gamma\} \quad \text{and}$$

$$\mathbf{H}(\operatorname{div}^0;\Omega) := \{\mathbf{v} \in \mathbf{H}(\operatorname{div};\Omega) : \operatorname{div}\mathbf{v} = 0 \text{ in } \Omega\}.$$

In [14], it is proven that $\varphi = 0$ in Ω and consequently, problems (1.1.6) and (1.1.5) are equivalent, with the key argument relying on the Helmholtz decomposition:

$$\mathbf{L}^2(\Omega) = \mathbf{H}(\operatorname{div}^0;\Omega) \oplus \mathrm{H}^1_0(\Omega).$$

Furthermore, by selecting Raviart–Thomas elements of degree $k \geq 0$ or BDM elements of degree $k+1$ for the tensor $\boldsymbol{\sigma}$, Raviart–Thomas elements of degree k for the velocity $\mathbf{u}$, and continuous piecewise polynomials of degree $k+1$ for $\mathrm{H}^1_0(\Omega)$, it is shown that the resulting conforming Galerkin discretization is well-posed, optimally convergent, and mass-conservative. However, the approach proposed in [14] fails to ensure momentum conservation. This is because the equation

$$(\mathbf{v}+\nabla\psi,\mathbf{div}\,\boldsymbol{\sigma}+\mathbf{f})_\Omega = 0, \tag{1.1.7}$$

for all discrete functions $\mathbf{v}$ and ψ in their respective discrete spaces does not necessarily imply that the equilibrium equation $\mathbf{div}\,\boldsymbol{\sigma} = -\mathbf{f}$ is exactly satisfied in Ω, even when $\mathbf{f}$ is approximated by piecewise polynomials.

Building on the above discussion and aiming to contribute to the development of numerical schemes for fluid flow problems that preserve conservation laws, we propose a new pseudostress-based numerical scheme that exactly preserves mass and momentum, where the latter holds for $\mathbf{f}$ in a suitably chosen piecewise polynomial space.

To achieve this, we discretize an equivalent reduced version of the three-field variational formulation (1.1.6), employing BDM elements of order 1 or Raviart–Thomas elements of order 0 for $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$, Raviart–Thomas elements of order 0 for $\mathbf{u}$ and $\mathbf{v}$, and the lowest-order Crouzeix–Raviart element (see [15]) for φ and ψ.

The key argument for ensuring momentum conservation is the discrete Helmholtz decomposition of piecewise constant functions into divergence-free Raviart–Thomas elements of order zero and gradients of Crouzeix–Raviart elements, as established in [3] (see also [41]). The above allows us to conclude from (1.1.7) that the momentum equation is exactly satisfied in Ω if $\mathbf{f}$ is piecewise constant.

It is important to note that, in this approach, as well as in the previous works [11], [26] and [14], the momentum equation $\mathbf{div}\,\boldsymbol{\sigma} = -\mathbf{f}$ is imposed in the L^2 sense at both the continuous and discrete levels. As discussed in [37], this enforcement comes at the expense of losing pressure robustness, which can be regarded as a trade-off for achieving momentum conservation.

The rest of the chapter is organized as follows: In Section 1.2, we introduce the three-field continuous problem and analyze its well-posedness. Then, in Section 1.3, we propose the numerical scheme and study its well-posedness and convergence. In Section 1.4 we address the extension to the Stokes problem with mixed boundary conditions. Finally, in Section 1.5 we illustrate the performance of the method by providing some numerical examples.

We conclude this section by introducing some notations and definitions. Throughout this work, we adopt standard notation for the Lebesgue and Sobolev spaces $\mathrm{L}^2(\Omega)$ and $\mathrm{H}^1(\Omega)$, equipped with the norms $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{1,\Omega}$, respectively. The seminorm $|\cdot|_{1,\Omega}$ is also used for $\mathrm{H}^1(\Omega)$ and serves as a norm in the subspace $\mathrm{H}^1_0(\Omega)$ introduced earlier.

Furthermore, we use $\mathbf{S}$ and $\mathbb{S}$ to represent the vectorial and tensorial counterparts of a generic scalar function space S. For a vector field $\mathbf{v} = (v_i)_{i=1,d}$, the differential operators $\nabla\mathbf{v}$ and $\operatorname{div}\mathbf{v}$ employed above are defined as

$$\nabla\mathbf{v} := \left(\frac{\partial v_i}{\partial x_j}\right)_{i,j=1,d}, \quad \operatorname{div}\mathbf{v} := \sum_{j=1}^{d} \frac{\partial v_j}{\partial x_j}.$$

As usual, the spaces $\mathbf{H}(\operatorname{div};\Omega)$ and $\mathbb{H}(\mathbf{div};\Omega)$ are equipped with the norms

$$\|\mathbf{v}\|_{\operatorname{div},\Omega} := \left(\|\mathbf{v}\|_{0,\Omega}^2 + \|\operatorname{div}\mathbf{v}\|_{0,\Omega}^2\right)^{1/2}, \quad \|\boldsymbol{\tau}\|_{\mathbf{div},\Omega} := \left(\|\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\mathbf{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2\right)^{1/2},$$

respectively. Moreover, using [22, Lemma 2.3], it can be shown that the seminorm

$$|\boldsymbol{\tau}|_{\mathbf{div},\Omega} := \left(\|\boldsymbol{\tau}^{\mathrm{d}}\|_{0,\Omega}^2 + \|\mathbf{div}\,\boldsymbol{\tau}\|_{0,\Omega}^2\right)^{1/2}$$

is also a norm in $\mathbb{H}_0(\mathbf{div};\Omega)$, equivalent to $\|\cdot\|_{\mathbf{div},\Omega}$, that is, there exist $c_1, c_2 > 0$, such that

$$c_1\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega} \leq |\boldsymbol{\tau}|_{\mathbf{div},\Omega} \leq c_2\|\boldsymbol{\tau}\|_{\mathbf{div},\Omega}, \qquad \forall\,\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div};\Omega). \tag{1.1.8}$$

Finally, throughout our analysis, we will use C and c, with or without subscripts, bars, tildes, or hats, to denote generic positive constants independent of the discretization parameters. These constants may take different values in different contexts.

1.2 Continuous Problem

As previously mentioned, we introduce an equivalent reduced version of (1.1.6). To this end, we first redefine the pseudostress tensor as

$$\boldsymbol{\sigma} := \nabla\mathbf{u} - \frac{1}{\nu}p\,\mathbb{I} \quad \text{in } \Omega, \tag{1.2.1}$$

which leads to the ν-scaled version of (1.1.4):

$$\boldsymbol{\sigma}^{\mathrm{d}} = \nabla \mathbf{u} \quad \text{in } \Omega, \quad -\mathbf{div}\, \boldsymbol{\sigma} = \frac{1}{\nu}\mathbf{f} \quad \text{in } \Omega, \quad \mathbf{u} = \mathbf{u}_D \quad \text{on } \Gamma, \quad (\mathrm{tr}\,(\boldsymbol{\sigma}), 1)_\Omega = 0. \tag{1.2.2}$$

Notice that the pseudostress (1.2.1) is essentially a viscosity-scaled version of the tensor (1.1.3) introduced in [11] and [26]. We adopt this new definition of $\boldsymbol{\sigma}$ because, as we will see in Section 1.2, it ensures that the stability estimates for the associated bilinear forms remain valid with constants independent of the viscosity. Furthermore, this property guarantees that the theoretical convergence rates for all unknowns are achieved with constants independent of ν. Based on this reasoning, for the remainder of this paper, we consider $\boldsymbol{\sigma}$ as defined in (1.2.1) and focus on deriving a finite element scheme that satisfies the conservation laws:

$$\mathrm{div}\, \mathbf{u} = 0 \quad \text{in } \Omega, \quad \text{and} \quad \mathbf{div}\, \boldsymbol{\sigma} = -\frac{1}{\nu}\mathbf{f} \quad \text{in } \Omega.$$

To achieve this, we introduce the following variational formulation based on (1.2.2): Find $\boldsymbol{\sigma} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$, $\mathbf{u} \in \mathbf{H}(\mathrm{div}^0;\Omega)$, and $\varphi \in \mathrm{H}^1_0(\Omega)$ such that

$$\begin{aligned} (\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega + (\mathbf{u} + \nabla\varphi, \mathbf{div}\, \boldsymbol{\tau})_\Omega &= \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D\rangle_\Gamma\,, \\ (\mathbf{v} + \nabla\psi, \mathbf{div}\, \boldsymbol{\sigma})_\Omega &= -\frac{1}{\nu}(\mathbf{f}, \mathbf{v} + \nabla\psi)_\Omega, \end{aligned} \tag{1.2.3}$$

for all $\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$ and $(\mathbf{v}, \psi) \in \mathbf{H}(\mathrm{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega)$.

From the fact that $\mathbf{L}^2(\Omega) = \mathbf{H}(\mathrm{div}^0;\Omega) \oplus \mathrm{H}^1_0(\Omega)$, it is clear that, after scaling by ν, problems (1.1.6) and (1.2.3) are equivalent and consequently, problem (1.2.3) is well-posed. However, for the sake of completeness, in what follows we establish its unique solvability and stability.

As usual in the context of mixed problems, first we introduce the bilinear forms $\mathbf{a} : \mathbb{H}(\mathbf{div}\,;\Omega) \times \mathbb{H}(\mathbf{div}\,;\Omega) \to \mathbb{R}$, $\mathbf{b} : \mathbb{H}(\mathbf{div}\,;\Omega) \times (\mathbf{H}(\mathrm{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega)) \to \mathbb{R}$ and the functionals $F : \mathbb{H}(\mathbf{div}\,;\Omega) \to \mathbb{R}$ and $G : \mathbf{H}(\mathrm{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega) \to \mathbb{R}$, as follows:

$$\mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) := (\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega, \quad \mathbf{b}(\boldsymbol{\tau}, (\mathbf{v}, \psi)) := (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{v} + \nabla\psi)_\Omega, \tag{1.2.4}$$

$$F(\boldsymbol{\tau}) := \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D\rangle_\Gamma \quad \text{and} \quad G(\mathbf{v}, \psi) := -\frac{1}{\nu}(\mathbf{f}, \mathbf{v} + \nabla\psi)_\Omega. \tag{1.2.5}$$

Then, problem (1.2.3) can be rewritten with a mixed structure as follows: Find $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{H}_0(\mathbf{div}\,;\Omega)\times(\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega))$, such that:

$$\begin{aligned}\mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau})+\mathbf{b}(\boldsymbol{\tau},(\mathbf{u},\varphi)) &= F(\boldsymbol{\tau}) \qquad \forall\,\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}\,;\Omega),\\ \mathbf{b}(\boldsymbol{\sigma},(\mathbf{v},\psi)) &= G(\mathbf{v},\psi) \quad \forall\,(\mathbf{v},\psi)\in\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega).\end{aligned}\tag{1.2.6}$$

The following theorem establishes the well-posedness of problem (1.2.6).

Theorem 1.2.1 *There exists a unique* $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{H}_0(\mathbf{div}\,;\Omega)\times(\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega))$ *solution to* (1.2.6) *with* $\varphi=0$ *in* Ω. *Furthermore, there exists* $C>0$, *independent of* ν, *such that*

$$|\boldsymbol{\sigma}|_{\mathbf{div},\Omega}+\|\mathbf{u}\|_{0,\Omega}\leq C\left(\frac{\|\mathbf{f}\|_{0,\Omega}}{\nu}+\|\mathbf{u}_D\|_{1/2,\Gamma}\right).\tag{1.2.7}$$

Proof. In what follows, we apply the classical Babuška–Brezzi theory (see [22, Theorem 2.3]) to establish the well-posedness of (1.2.6).

We begin by noting that, using the Cauchy–Schwarz inequality, estimate (1.1.8) and [22, Theorem 1.7], we can readily deduce that the bilinear forms $\mathbf{a}$ and $\mathbf{b}$, as well as the functionals G and F, satisfy the following boundedness estimates:

$$\begin{aligned}&|\mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau})|\leq|\boldsymbol{\sigma}|_{\mathbf{div},\Omega}|\boldsymbol{\tau}|_{\mathbf{div},\Omega}, && |\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))|\leq|\boldsymbol{\tau}|_{\mathbf{div},\Omega}(\|\mathbf{v}\|_{\mathrm{div},\Omega}+|\psi|_{1,\Omega}),\\ &|G(\mathbf{v},\psi)|\leq\frac{1}{\nu}\|\mathbf{f}\|_{0,\Omega}(\|\mathbf{v}\|_{\mathrm{div},\Omega}+|\psi|_{1,\Omega}), && |F(\boldsymbol{\tau})|\leq C\|\mathbf{u}_D\|_{1/2,\Gamma}|\boldsymbol{\tau}|_{\mathbf{div},\Omega},\end{aligned}\tag{1.2.8}$$

where $C>0$ is a constant independent of ν. Furthermore, a simple rescaling by ν allows us to deduce from [14, Lemma 2.3] that the following inf-sup condition holds:

$$\sup_{\mathbf{0}\neq\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}\,;\Omega)}\frac{\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))}{|\boldsymbol{\tau}|_{\mathbf{div},\Omega}}\geq\beta(\|\mathbf{v}\|_{\mathrm{div},\Omega}+|\psi|_{1,\Omega})\qquad\forall\,(\mathbf{v},\psi)\in\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega),\tag{1.2.9}$$

where $\beta>0$ is a constant independent of ν.

Next, we define the kernel of $\mathbf{b}$ as

$$\mathbb{V}:=\{\boldsymbol{\tau}\in\mathbb{H}_0(\mathbf{div}\,;\Omega):\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))=0\quad\forall\,(\mathbf{v},\psi)\in\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega)\}.$$

Using the Helmholtz decomposition $\mathbf{L}^2(\Omega)=\mathbf{H}(\mathrm{div}^0;\Omega)\oplus\mathrm{H}^1_0(\Omega)$, it follows that $\mathbb{V}$

can be characterized as

$$\mathbb{V} = \{\boldsymbol{\tau} \in \mathbb{H}_0(\mathbf{div}\,;\Omega) : \mathbf{div}\,\boldsymbol{\tau} = 0 \text{ in } \Omega\}.$$

Moreover, for each $\boldsymbol{\tau} \in \mathbb{V}$, the bilinear form $\mathbf{a}$ satisfies

$$\mathbf{a}(\boldsymbol{\tau}, \boldsymbol{\tau}) = |\boldsymbol{\tau}|^2_{\mathbf{div},\Omega}, \tag{1.2.10}$$

which establishes the ellipticity of $\mathbf{a}$ on $\mathbb{V}$.

In this way, from (1.2.8), (1.2.9), (1.2.10) and the classical Babuška–Brezzi theory, we readily obtain the unique solvability of problem (1.2.6).

Now, to deduce that $\varphi = 0$ in Ω, given $\psi \in \mathrm{H}^1_0(\Omega)$, we simply take $\boldsymbol{\tau} = (\psi - |\Omega|^{-1}(\psi, 1)_\Omega)\,\mathbb{I} \in \mathbb{H}_0(\mathbf{div}\,;\Omega)$ in the first equation of (1.2.6) and recall that $\langle \mathbf{n}, \mathbf{u}_D \rangle_\Gamma = 0$ (see (1.1.2)), to obtain

$$(\nabla\psi, \nabla\varphi)_\Omega = 0,$$

which together with the fact that ψ is arbitrary, implies that $\varphi = 0$ in Ω.

We conclude the proof by observing that estimate (1.2.7) is a direct consequence of the Babuška–Brezzi theory and the fact that $\varphi = 0$. □

1.3 Galerkin scheme

In this section, we introduce and analyze the mass and momentum conservative Galerkin scheme for the mixed formulation (1.2.6). Moreover, we derive the corresponding theoretical rates of convergence.

1.3.1 Discrete scheme

Let $\mathcal{T}_h$ be a regular family of regular triangulations of the polygonal region $\overline{\Omega}$ by triangles T in $\mathbb{R}^2$ or tetrahedra in $\mathbb{R}^3$ of diameter h_T, such that $\overline{\Omega} = \cup\{T : T \in \mathcal{T}_h\}$ and define $h := \max\{h_T : T \in \mathcal{T}_h\}$. Given an integer $l \geq 0$ and a subset S of $\mathbb{R}^d$, we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S. Hence, for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of lowest order and the Brezzi–Douglas–Marini (BDM) element of order 1, respectively as (see, for

instance [6]):

$$\mathbf{RT}_0(T) := [P_0(T)]^d \oplus P_0(T)\mathbf{x} \quad \text{and} \quad \mathbf{BDM}_1(T) = [P_1(T)]^d$$

where $\mathbf{x} := (x_1, \ldots, x_d)^t$ is a generic vector of $\mathbb{R}^d$.

In addition, we let $\mathcal{E}_h$ be the set of edges (in 2D) or faces (in 3D) of $\mathcal{T}_h$, whose corresponding diameters are denoted h_e, and define

$$\mathcal{E}_h(\Omega) := \big\{e \in \mathcal{E}_h : \quad e \subseteq \Omega\big\} \quad \text{and} \quad \mathcal{E}_h(\Gamma) := \big\{e \in \mathcal{E}_h : \quad e \subseteq \Gamma\big\}.$$

We also let $[\![\cdot]\!]$ be the usual jump operator across internal edges or faces defined for piecewise continuous functions v, by

$$[\![v]\!] \;=\; (v\big|_{T_+})\big|_e \;-\; (v\big|_{T_-})\big|_e \quad \text{with} \quad e \;=\; \partial T_+ \cap \partial T_-,$$

where T_+ and T_- are the elements of $\mathcal{T}_h$ having e as a common edge or face. Then, we introduce the well-known Crouzeix–Raviart space (see [15]):

$$\begin{aligned}\Psi_h^{\varphi} := \quad \Big\{ & v_h : \Omega \to \mathbb{R} : v_h|_T \in P_1(T),\, \forall\, T \in \mathcal{T}_h, \quad \int_e [\![v_h]\!] = 0, \quad \forall\, e \in \mathcal{E}_h(\Omega) \\ & \text{and} \quad \int_e v_h = 0, \quad \forall e \in \mathcal{E}_h(\Gamma)\Big\},\end{aligned} \tag{1.3.1}$$

equipped with the norm

$$|v_h|_h = \left(\sum_{T \in \mathcal{T}_h} |v_h|_{1,T}^2 \right)^{1/2}, \quad \forall v_h \in \Psi_h^{\varphi}.$$

In this way, defining the discrete spaces

$$\begin{aligned}\mathbb{H}_h^{\boldsymbol{\sigma}} &:= \{\boldsymbol{\tau}_h \in \mathbb{H}(\mathbf{div}\,;\Omega) : \mathbf{c}^t\boldsymbol{\tau}_h \in \mathbf{BDM}_1(T) \quad \forall\, \mathbf{c} \in \mathbb{R}^d, \quad \forall\, T \in \mathcal{T}_h\}, \\ \mathbf{H}_h^{\mathbf{u}} &:= \{\mathbf{z}_h \in \mathbf{H}(\mathrm{div};\Omega) : \mathbf{z}_h|_T \in \mathbf{RT}_0(T), \quad \forall\, T \in \mathcal{T}_h\}, \\ \mathbb{H}_{h,0}^{\boldsymbol{\sigma}} &:= \mathbb{H}_h^{\boldsymbol{\sigma}} \cap \mathbb{H}_0(\mathbf{div}\,;\Omega), \quad \mathbf{H}_{h,0}^{\mathbf{u}} := \mathbf{H}_h^{\mathbf{u}} \cap \mathbf{H}(\mathrm{div}^0;\Omega),\end{aligned} \tag{1.3.2}$$

the Galerkin scheme associated to (1.2.6) reads: Find $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{H}_{h,0}^{\boldsymbol{\sigma}} \times (\mathbf{H}_{h,0}^{\mathbf{u}} \times \Psi_h^{\varphi})$, such that:

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{u}_h, \varphi_h)) &= F(\boldsymbol{\tau}_h) \qquad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,0}, \\ \mathbf{b}_h(\boldsymbol{\sigma}_h, (\mathbf{v}_h, \psi_h)) &= G_h(\mathbf{v}_h, \psi_h) \quad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h, \end{aligned} \tag{1.3.3}$$

where the form $\mathbf{a}$ and the functional F are defined in (1.2.4) and (1.2.5), respectively, whereas $\mathbf{b}_h : \mathbb{H}_0(\mathbf{div}\,;\Omega) \times \mathrm{H}(h) \to \mathbb{R}$ and the functional $G_h : \mathrm{H}(h) \to \mathbb{R}$ are defined as follows

$$\mathbf{b}_h(\boldsymbol{\tau}, (\mathbf{v}, \psi_h)) := (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{v} + \nabla_h \psi_h)_\Omega, \tag{1.3.4}$$

$$G_h(\mathbf{v}, \psi_h) := -\frac{1}{\nu}(\mathbf{f}, \mathbf{v} + \nabla_h \psi_h)_\Omega, \tag{1.3.5}$$

where ∇_h is the discrete gradient for discontinuous functions, that is, $\nabla_h \psi_h|_T = \nabla(\psi_h|_T), \forall\, T \in \mathcal{T}_h$ and $\mathrm{H}(h) := \mathbf{H}(\mathrm{div}^0; \Omega) \times (\mathrm{H}^1_0(\Omega) + \Psi^{\varphi}_h)$.

1.3.2 Well-posedness of the discrete scheme

In what follows we address the unique solvability and stability of problem (1.3.3) by adapting to the discrete case the analysis described in Section 1.2. We begin by noticing that using Hölder inequality, the form $\mathbf{b}_h$ and the functional G_h are bounded with the same constants as for $\mathbf{b}$ and G in (1.2.8).

Now we let $\mathbb{V}_h$ be discrete kernel of $\mathbf{b}_h$, that is

$$\mathbb{V}_h := \{\boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_h : \ \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)) = 0, \quad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h\}.$$

Recalling from [3, Theorem 4.1] and [41, Theorem 4.9] that the following orthogonal decomposition holds:

$$\mathbf{Q}_h = \mathbf{H}^{\mathbf{u}}_{h,0} \oplus \nabla_h \Psi^{\varphi}_h, \tag{1.3.6}$$

where $\mathbf{Q}_h$ is the corresponding vectorial counterpart of the space

$$Q_h := \{q \in L^2(\Omega) \ : \ q|_T \in P_0(T), \quad \forall\, T \in \mathcal{T}_h\},$$

and

$$\nabla_h \Psi^{\varphi}_h := \{\mathbf{s}_h|_T \in \mathbf{P}_0(T) \ : \exists\, v_h \in \Psi^{\varphi}_h \ \text{ such that } \ \mathbf{s}_h|_T = \nabla(v_h|_T), \quad \forall\, T \in \mathcal{T}_h\},$$

we observe that for each $\boldsymbol{\tau}_h \in \mathbb{V}_h$,

$$(\mathbf{div}\ \boldsymbol{\tau}_h, \mathbf{v}_h + \nabla_h \psi_h)_\Omega = 0, \quad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi_h^{\varphi},$$

is equivalent to

$$(\mathbf{div}\ \boldsymbol{\tau}_h, \mathbf{z}_h)_\Omega = 0, \quad \forall\, \mathbf{z}_h \in \mathbf{Q}_h,$$

which implies that $\mathbb{V}_h$ can be characterized as follows

$$\mathbb{V}_h := \{\boldsymbol{\tau}_h \in \mathbb{H}_h^{\boldsymbol{\sigma}} : \mathbf{div}\, \boldsymbol{\tau}_h = \mathbf{0} \quad \text{in} \quad \Omega\}.$$

In this way, $\mathbf{a}$ satisfies

$$\mathbf{a}(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h) = |\boldsymbol{\tau}_h|^2_{\mathbf{div},\Omega} \qquad \forall \boldsymbol{\tau}_h \in \mathbb{V}_h,$$

that is, $\mathbf{a}$ is elliptic on the kernel of $\mathbf{b}_h$.

Now we establish the discrete inf-sup condition of $\mathbf{b}_h$. To that end, we recall from [6, Section 2.5] that there exist interpolator operators $\Pi_h^{RT} : \mathbf{H}^1(\Omega) \to \mathbf{H}^{\mathbf{u}}_h$ and $\Pi_h^{BDM} : \mathbf{H}^1(\Omega) \to \mathbf{X}_h := \{\tau_h \in \mathbf{H}(\mathrm{div};\Omega) : \tau_h|_T \in \mathbf{BDM}_1(T), \quad \forall\, T \in \mathcal{T}_h\}$, satisfying the approximation property

$$\|\Pi_h^\star(\tau) - \tau\|_{0,T} \leq c h_T^m |\tau|_{m,T}, \quad \forall\, \tau \in \mathbf{H}^m(T), \quad \forall\, T \in \mathcal{T}_h, \tag{1.3.7}$$

for all $1 \leq m \leq l_\star$ and $\star \in \{RT, BDM\}$, with $l_{RT} = 1$ and $l_{BDM} = 2$, and the commutative property

$$\mathrm{div}(\Pi_h^\star(\tau)) = \mathcal{P}_h(\mathrm{div}\,\tau), \quad \forall \tau \in \mathbf{H}^1(\Omega), \quad \forall \star \in \{RT, BDM\}, \tag{1.3.8}$$

where $\mathcal{P}_h$ is the L^2-projection on Q_h, which satisfies

$$(v - \mathcal{P}_h(v), z_h)_\Omega = 0 \quad \forall\, z_h \in Q_h,$$

and the local approximation property

$$\|v - \mathcal{P}_h(v)\|_{0,T} \leq C h^m |v|_{m,T}, \quad \forall\, T \in \mathcal{T}_h, \tag{1.3.9}$$

for all $0 \leq m \leq 1$ and for all $v \in \mathrm{H}^m(\Omega)$. Notice that from (1.3.8) and (1.3.9) we have that

$$\|\mathrm{div}\,\tau - \mathrm{div}(\Pi_h^\star(\tau))\|_{0,T} \leq Ch^m |\mathrm{div}\,\tau|_{m,T}, \quad \forall\, T \in \mathcal{T}_h,$$

for all $0 \leq m \leq 1$ and for all $\tau \in \mathbf{H}^1(\Omega)$ with $\mathrm{div}\,\tau \in \mathrm{H}^m(\Omega)$.

In what follows we will employ a tensor version of Π_h^{BDM}, denoted by $\mathbf{\Pi}_h^{BDM} : \mathbb{H}^1(\Omega) \to \mathbb{H}_h^{\boldsymbol{\sigma}}$, which is defined row-wise by Π_h^{BDM}, and the vector version of $\mathcal{P}_h$, denoted by $\mathbf{P}_h : \mathbf{L}^2(\Omega) \to \mathbf{Q}_h$, defined component-wise by $\mathcal{P}_h$.

Now we are in position of establishing the inf-sup condition of $\mathbf{b}_h$.

Lemma 1.3.1 *There exists $\tilde{\beta} > 0$, independent of h and ν, such that*

$$\sup_{\mathbf{0} \neq \boldsymbol{\tau}_h \in \mathbb{H}_{h,0}^{\boldsymbol{\sigma}}} \frac{\mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h))}{|\boldsymbol{\tau}_h|_{\mathbf{div},\Omega}} \geq \tilde{\beta}(\|\mathbf{v}_h\|_{\mathrm{div},\Omega} + |\psi_h|_h) \qquad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{H}_{h,0}^{\mathbf{u}} \times \Psi_h^{\varphi}.$$

Proof. We proceed similary to the proof of [14, Lemma 3.2]. In fact, we let $B \subseteq \mathbb{R}^d$ be a bounded and open convex domain such that $\overline{\Omega} \subset B$, and given $(\mathbf{v}_h, \psi_h) \in \mathbf{H}_{h,0}^{\mathbf{u}} \times \Psi_h^{\varphi}$, we let $\mathbf{z} \in \mathbf{H}_0^1(B)$ be the unique weak solution of the auxiliary problem

$$-\Delta \mathbf{z} = \mathbf{h}(\mathbf{v}_h, \psi_h) \quad \text{in} \quad B, \quad \mathbf{z} = \mathbf{0} \quad \text{on} \quad \partial B,$$

with

$$\mathbf{h}(\mathbf{v}_h, \psi_h) := \begin{cases} \mathbf{v}_h + \nabla_h \psi_h, & \text{in} \quad \Omega, \\ \mathbf{0}, & \text{in} \quad B \backslash \overline{\Omega}. \end{cases}$$

It is well known that $\mathbf{z} \in \mathbf{H}^2(B)$ (see [29]) and

$$\|\mathbf{z}\|_{2,\Omega} \leq C\|\mathbf{h}(\mathbf{v}_h, \psi_h)\|_{0,B} = C\|\mathbf{v}_h + \nabla_h \psi_h\|_{0,\Omega} \leq C\left(\|\mathbf{v}_h\|_{\mathrm{div},\Omega} + |\psi_h|_h\right). \quad (1.3.10)$$

Note that as $\mathbf{v}_h \in \mathbf{H}_{h,0}^{\mathbf{u}}$, then in accordance with [6, Corollary 2.3.1], $\mathbf{v}_h \in \mathbf{Q}_h$. Now, we define

$$\widehat{\boldsymbol{\tau}}_h := -\mathbf{\Pi}_h^{BDM}(\nabla \mathbf{z}|_\Omega) + \frac{1}{d|\Omega|}\left(\mathrm{tr}\left(\mathbf{\Pi}_h^{BDM}(\nabla \mathbf{z}|_\Omega)\right), 1\right)_\Omega \mathbb{I} \quad \text{in} \quad \Omega,$$

and observe from (1.3.8) and (1.3.10) that

$$\mathbf{div}\,\widehat{\boldsymbol{\tau}}_h = \mathbf{v}_h + \nabla_h \psi_h \in \mathbf{Q}_h \quad \text{and} \quad |\widehat{\boldsymbol{\tau}}_h|_{\mathbf{div},\Omega} \leq \hat{C}(\|\mathbf{v}_h\|_{\mathrm{div},\Omega} + |\psi_h|_h).$$

From the latter, we obtain

$$\sup_{\mathbf{0}\neq\boldsymbol{\tau}_h\in\mathbb{H}^{\boldsymbol{\sigma}}_{h,0}}\frac{\mathbf{b}_h(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))}{|\boldsymbol{\tau}_h|_{\mathbf{div},\Omega}}\geq\frac{\mathbf{b}(\widehat{\boldsymbol{\tau}}_h,(\mathbf{v}_h,\psi_h))}{|\widehat{\boldsymbol{\tau}}_h|_{\mathbf{div},\Omega}}\geq\hat{C}^{-1}\frac{\|\mathbf{v}_h\|^2_{\mathrm{div},\Omega}+|\psi_h|^2_h}{\|\mathbf{v}_h\|_{\mathrm{div},\Omega}+|\psi_h|_h}\\ \geq\tilde{\beta}(\|\mathbf{v}_h\|_{\mathrm{div},\Omega}+|\psi_h|_h). \tag{1.3.11}$$

with $\tilde{\beta}>0$ independent of h and ν.

□

These properties and the Babuška–Brezzi theory allow us to conclude the well-posedness of (1.3.3). This result is summarized in the following theorem.

Theorem 1.3.2 *There exists a unique* $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))\in\mathbb{H}^{\boldsymbol{\sigma}}_{h,0}\times(\mathbf{H}^{\mathbf{u}}_{h,0}\times\Psi^{\varphi}_h)$ *solution to the Galerkin scheme* (1.3.3). *In addition, there exists* $C>0$, *independent of* h *and* ν, *such that*

$$|\boldsymbol{\sigma}_h|_{\mathbf{div},\Omega}+\|\mathbf{u}_h\|_{0,\Omega}+|\varphi_h|_h\leq C\left(\frac{\|\mathbf{f}\|_{0,\Omega}}{\nu}+\|\mathbf{u}_D\|_{1/2,\Omega}\right).$$

Remark 1.3.3 *Observe that the discrete space* $\mathbf{H}^{\mathbf{u}}_{h,0}$ *becomes*

$$\mathbf{H}^{\mathbf{u}}_{h,0}=\{\mathbf{v}_h\in\mathbf{H}^{\mathbf{u}}_h:\operatorname{div}\mathbf{v}_h=0\ \text{in}\ \Omega\},$$

which implies that the numerical scheme (1.3.3) *produces exactly divergence-free approximations for the velocity* $\mathbf{u}$.

Furthermore, from the second equation of (1.3.3) *and the discrete Helmholtz decomposition* (1.3.6), *we deduce that*

$$\left(\mathbf{div}\ \boldsymbol{\sigma}_h+\nu^{-1}\mathbf{f},\mathbf{z}_h\right)_\Omega=0\quad\forall\,\mathbf{z}_h\in\mathbf{Q}_h.$$

This implies that $\mathbf{div}\ \boldsymbol{\sigma}_h=-\nu^{-1}\mathbf{P}_h(\mathbf{f})$, *meaning that the method exactly preserves the discrete equilibrium equation when* $\mathbf{f}\in\mathbf{Q}_h$. *In other words, the scheme is momentum conservative whenever* $\mathbf{f}\in\mathbf{Q}_h$. *Moreover, if* $\mathbf{f}\in\mathbf{H}^1(\Omega)$, *from* (1.3.9) *we obtain the estimate*

$$\|\nu^{-1}\mathbf{f}+\mathbf{div}\ \boldsymbol{\sigma}_h\|_{0,\Omega}=\nu^{-1}\|\mathbf{f}-\mathbf{P}_h(\mathbf{f})\|_{0,\Omega}\leq c\nu^{-1}h\|\mathbf{f}\|_{1,\Omega}, \tag{1.3.12}$$

showing that, for sufficiently smooth $\mathbf{f}$, *the momentum equation is approximated with*

an optimal rate of convergence.

We also emphasize that the achieved momentum conservation comes at the cost of losing pressure robustness. More specifically, since our method enforces the equilibrium equation in $\mathbf{L}^2(\Omega)$*, the gradient component of the Helmholtz decomposition of the source term* $\mathbf{f}$ *affects the solution, leading to a lack of pressure robustness, as discussed in [37].*

On the other hand, note that if $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h))$ *is a solution to* (1.3.3)*,* φ_h *is not necessarily identically zero in* Ω*. However, as shown in Theorem 1.3.4,* φ_h *converges to zero (see* (1.3.16)*). Furthermore, as demonstrated in Example 4, Section 1.5, despite introducing an additional unknown, the proposed numerical scheme remains slightly less expensive than the standard formulation studied in [11] and [26]. The computational cost can be further reduced by employing the exactly divergence-free discrete basis for* $\mathbf{H}^{\mathbf{u}}_{h,0}$ *introduced in [1], optimizing implementation efficiency.*

Finally, we note that, to the best of the authors' knowledge, the Helmholtz decomposition (1.3.6) *is only available in the literature for the lowest-order case. This limitation prevents a straightforward extension of the above analysis to higher-order cases.*

1.3.3 Convergence analysis

We now analyze the convergence of (1.3.1) and establish the corresponding theoretical rate of convergence. We begin by noting that the gradient operator ∇ and its discrete counterpart ∇_h coincide in $\mathrm{H}^1_0(\Omega)$, which implies that

$$\mathbf{b}_h(\boldsymbol{\tau}, (\mathbf{v}, \psi)) = \mathbf{b}(\boldsymbol{\tau}, (\mathbf{v}, \psi)), \qquad \forall\, (\boldsymbol{\tau}, (\mathbf{v}, \psi)) \in \mathbb{H}(\mathbf{div}\,; \Omega) \times (\mathbf{H}(\mathrm{div}^0; \Omega) \times \mathrm{H}^1_0(\Omega)), \tag{1.3.13}$$

and

$$G_h(\mathbf{v}, \psi) = G(\mathbf{v}, \psi), \qquad \forall\, (\mathbf{v}, \psi) \in \mathbf{H}(\mathrm{div}^0; \Omega) \times \mathrm{H}^1_0(\Omega). \tag{1.3.14}$$

Thus, if $(\boldsymbol{\sigma}, (\mathbf{u}, \varphi)) \in \mathbb{H}(\mathbf{div}\,; \Omega) \times (\mathbf{H}(\mathrm{div}^0; \Omega) \times \mathrm{H}^1_0(\Omega))$ is the unique solution of (1.2.6), we can replace $\mathbf{b}$ with $\mathbf{b}_h$ and G with G_h in (1.2.6) without altering the validity of the equations.

The following theorem establishes the theoretical rate of convergence for the numerical scheme (1.3.3). Instead of relying on the *a priori* error estimate given,

for instance, in [22, Theorem 2.6], we derive the estimate from scratch, utilizing the orthogonality property of the scheme. This approach allows us to obtain error bounds with constants independent of ν, which, as we shall see, differs from the results in [14]. Moreover, it allows us to establish a superconvergence result for the deviatoric part of $\boldsymbol{\sigma}_h$.

Theorem 1.3.4 *Let* $(\boldsymbol{\sigma},(\mathbf{u},0)) \in \mathbb{H}_0(\mathbf{div}\,;\Omega) \times (\mathbf{H}(\mathrm{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega))$ *and* $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)) \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,0} \times \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h$ *be the unique solutions of* (1.2.6) *and* (1.3.3)*, respectively, and assume that the exact solution satisfies* $\boldsymbol{\sigma} \in \mathbb{H}^2(\Omega)$ *and* $\mathbf{u} \in \mathbf{H}^1(\Omega)$*. Then, there exist positive constants* c_1*,* c_2 *and* c_3*, independent of* ν *and* h*, such that,*

$$\|\boldsymbol{\sigma}^{\mathrm{d}} - \boldsymbol{\sigma}_h^{\mathrm{d}}\|_{0,\Omega} \leq c_1 h^2 |\boldsymbol{\sigma}|_{2,\Omega} \tag{1.3.15}$$

and

$$\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} + |\varphi_h|_h \leq c_2 h |\boldsymbol{\sigma}|_{2,\Omega} + c_3 h |\mathbf{u}|_{1,\Omega}. \tag{1.3.16}$$

Proof. From (1.2.6), (1.3.3), (1.3.13), (1.3.14) and the fact that $\varphi = 0$ in Ω, we readily obtain the orthogonality property:

$$\begin{aligned} ((\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)^{\mathrm{d}}, \boldsymbol{\tau}_h^{\mathrm{d}})_\Omega + (\mathbf{u} - \mathbf{u}_h - \nabla_h \varphi_h, \mathbf{div}\,\boldsymbol{\tau}_h)_\Omega &= 0, \quad \forall \boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,0}, \\ (\mathbf{v}_h + \nabla_h \psi_h, \mathbf{div}\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h))_\Omega &= 0, \quad \forall (\mathbf{v}_h, \psi_h) \in \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h. \end{aligned} \tag{1.3.17}$$

Now, let $\widehat{\boldsymbol{\sigma}}_h := \boldsymbol{\Pi}_h^{BDM}(\boldsymbol{\sigma})$ and $\widehat{\mathbf{u}}_h := \Pi_h^{RT}(\mathbf{u})$. From (1.3.8) and using that $\mathbf{div}\,\boldsymbol{\sigma} = -\frac{1}{\nu}\mathbf{f}$ in Ω and $\mathbf{div}\,\boldsymbol{\sigma}_h = -\frac{1}{\nu}\mathbf{P}_h(\mathbf{f})$ in Ω (see Remark 1.3.3), it is clear that

$$\mathbf{div}\,(\widehat{\boldsymbol{\sigma}}_h) = \mathbf{P}_h(\mathbf{div}\,\boldsymbol{\sigma}) = -\frac{1}{\nu}\mathbf{P}_h(\mathbf{f}) = \mathbf{div}\,(\boldsymbol{\sigma}_h) \quad \text{in} \quad \Omega,$$

thus $\mathbf{div}\,(\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h) = \mathbf{0}$ in Ω. Then, adding and subtracting $\widehat{\boldsymbol{\sigma}}_h$ in the first equation of (1.3.17) and taking $\boldsymbol{\tau}_h = \widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h$, we deduce that

$$((\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h)^{\mathrm{d}}, (\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h)^{\mathrm{d}})_\Omega = -((\boldsymbol{\sigma} - \widehat{\boldsymbol{\sigma}}_h)^{\mathrm{d}}, (\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h)^{\mathrm{d}})_\Omega,$$

which implies

$$\|(\widehat{\boldsymbol{\sigma}}_h - \boldsymbol{\sigma}_h)^{\mathrm{d}}\|_{0,\Omega} \leq \|(\boldsymbol{\sigma} - \widehat{\boldsymbol{\sigma}}_h)^{\mathrm{d}}\|_{0,\Omega}.$$

In this way, from the latter, the triangle inequality, estimate (1.3.7) and the regularity of the mesh, we deduce that

$$\|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)^{\mathrm{d}}\|_{0,\Omega} \leq 2\|(\boldsymbol{\sigma}-\widehat{\boldsymbol{\sigma}}_h)^{\mathrm{d}}\|_{0,\Omega} \leq 2\|\boldsymbol{\sigma}-\widehat{\boldsymbol{\sigma}}_h\|_{0,\Omega} \leq c_1 h^2 \|\boldsymbol{\sigma}\|_{2,\Omega},$$

with $c_1 > 0$ in dependent of h and ν.

Now, to deduce (1.3.16) we add and subtract $\widehat{\mathbf{u}}_h$ in the first equation of (1.3.17) to obtain

$$(\widehat{\mathbf{u}}_h - \mathbf{u}_h - \nabla_h \varphi_h, \mathbf{div}\,\boldsymbol{\tau}_h)_\Omega = -(\mathbf{u}-\widehat{\mathbf{u}}_h, \mathbf{div}\,\boldsymbol{\tau}_h)_\Omega - ((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)^{\mathrm{d}}, \boldsymbol{\tau}_h^{\mathrm{d}})_\Omega,$$

for all $\boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,0}$. Then, from this identity and the discrete inf-sup condition (1.3.11) we obtain

$$\begin{aligned}\widetilde{\beta}(\|\widehat{\mathbf{u}}_h - \mathbf{u}_h\|_{0,\Omega} + |\varphi_h|_h) &\leq \sup_{\mathbf{0}\neq\boldsymbol{\tau}_h\in\mathbb{H}^{\boldsymbol{\sigma}}_{h,0}} \frac{\left|(\mathbf{u}-\widehat{\mathbf{u}}_h, \mathbf{div}\,\boldsymbol{\tau}_h)_\Omega + ((\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)^{\mathrm{d}}, \boldsymbol{\tau}_h^{\mathrm{d}})_\Omega\right|}{|\boldsymbol{\tau}_h|_{\mathbf{div},\Omega}},\\ &\leq C(\|\mathbf{u}-\widehat{\mathbf{u}}_h\|_{0,\Omega} + \|(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)^{\mathrm{d}}\|_{0,\Omega}),\end{aligned}$$

which combined with (1.3.7), (1.3.15) and the triangle inequality imply (1.3.16). □

Remark 1.3.5 *Recalling that $\boldsymbol{\sigma}^{\mathrm{d}} = \nabla\mathbf{u}$, the previous theorem confirms that the method provides a superconvergent approximation for the velocity gradient. Specifically, we obtain the following estimate:*

$$\|\nabla\mathbf{u} - \boldsymbol{\sigma}_h^{\mathrm{d}}\|_{0,\Omega} \leq c_1 h^2 |\boldsymbol{\sigma}|_{2,\Omega},$$

where c_1 is independent of ν.

1.4 Stokes problem with mixed boundary conditions

Now we briefly discuss how to extend the method for the case of mixed boundary conditions. To that end, now we let $\Gamma_D \subseteq \partial\Omega$ and $\Gamma_N \subseteq \partial\Omega$ satisfying $|\Gamma_N| \neq 0$, $|\Gamma_D| \neq 0$, $\Gamma_D \cap \Gamma_N = \emptyset$ and $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial\Omega$, and consider the following Stokes problem

with mixed boundary conditions:

$$-\nu\Delta\mathbf{u}+\nabla p=\mathbf{f} \quad \text{in } \Omega, \quad \operatorname{div}\mathbf{u}=0 \quad \text{in } \Omega, \quad \mathbf{u}=\mathbf{u}_D \quad \text{on } \Gamma_D, \quad (\nu\nabla\mathbf{u}-p\mathbb{I})\mathbf{n}=\mathbf{0} \quad \text{on } \Gamma_N,$$

where the last equation on Γ_N is the so-called *do-nothing* condition (see eg. [34, Section 2.4]).

Introducing the pseudostress tensor (1.2.1) the equations above can be rewritten equivalently as follows:

$$\boldsymbol{\sigma}^{\mathrm{d}}=\nabla\mathbf{u} \quad \text{in } \Omega, \quad -\mathbf{div}\,\boldsymbol{\sigma}=\frac{1}{\nu}\mathbf{f} \quad \text{in } \Omega, \quad \mathbf{u}=\mathbf{u}_D \quad \text{on} \quad \Gamma_D, \quad \boldsymbol{\sigma}\mathbf{n}=\mathbf{0} \quad \text{on} \quad \Gamma_N, \tag{1.4.1}$$

which lead to the variational formulation: Find $\boldsymbol{\sigma}\in\mathbb{H}_N(\mathbf{div}\,;\Omega)$, $\mathbf{u}\in\mathbf{H}(\mathrm{div}^0;\Omega)$ and $\varphi\in\mathrm{H}^1_0(\Omega)$, such that

$$\begin{aligned}\mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau})+\mathbf{b}(\boldsymbol{\tau},(\mathbf{u},\varphi)) &= F_D(\boldsymbol{\tau}), \quad \forall\boldsymbol{\tau}\in\mathbb{H}_N(\mathbf{div}\,;\Omega),\\ \mathbf{b}(\boldsymbol{\sigma},(\mathbf{v},\psi)) &= G(\mathbf{v},\psi), \quad \forall(\mathbf{v},\psi)\in\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega),\end{aligned} \tag{1.4.2}$$

where $\mathbf{a}$, $\mathbf{b}$ and G are defined in (1.2.4) and (1.2.5), whereas $F_D:\mathbb{H}_N(\mathbf{div}\,;\Omega)\to\mathbb{R}$ is given by

$$F_D(\boldsymbol{\tau}):=\langle\boldsymbol{\tau}\mathbf{n},\mathbf{u}_D\rangle_{\Gamma_D}, \tag{1.4.3}$$

with

$$\mathbb{H}_N(\mathbf{div}\,;\Omega):=\{\boldsymbol{\tau}\in\mathbb{H}(\mathbf{div}\,;\Omega)\ :\ \boldsymbol{\tau}\mathbf{n}=\mathbf{0} \quad \text{on} \quad \Gamma_N\},$$

and $\langle\cdot,\cdot\rangle_{\Gamma_D}$ denoting the product of duality between the trace space $\mathbf{H}^{1/2}(\Gamma_D)$ and its dual $\mathbf{H}^{-1/2}(\Gamma_D)$.

The well-posedness of (1.4.2) is established next.

Theorem 1.4.1 *There exists a unique* $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{H}_N(\mathbf{div}\,;\Omega)\times(\mathbf{H}(\mathrm{div}^0;\Omega)\times\mathrm{H}^1_0(\Omega))$ *solution to* (1.4.2) *with* $\varphi=0$ *in* Ω. *Furthermore, there exists* $C>0$, *independent of* ν, *such that*

$$|\boldsymbol{\sigma}|_{\mathbf{div},\Omega}+\|\mathbf{u}\|_{0,\Omega}\leq C\left(\frac{\|\mathbf{f}\|_{0,\Omega}}{\nu}+\|\mathbf{u}_D\|_{1/2,\Gamma_D}\right). \tag{1.4.4}$$

Proof. Employing similar arguments to those utilized in the proof of Theorem 1.2.1 it is possible to prove existence and uniqueness of solution of problem (1.4.2) and

estimate (1.4.4). In particular, the ellipticity of the bilinear form $\mathbf{a}$ on the kernel of $\mathbf{b}$ can be derived by combining (1.1.8) and [22, Lemma 2.5]. In addition, by taking $\boldsymbol{\tau} = \psi\mathbb{I} \in \mathbb{H}_N(\mathbf{div}\,;\Omega)$ with $\psi \in \mathrm{H}^1_0(\Omega)$ in the second equation of (1.4.2), and proceeding analogously to the proof of Theorem 1.2.1 we easily prove that $\varphi = 0$ in Ω. We omit further details. □

As for the Galerkin discretization of problem (1.4.2), here we consider the spaces defined in (1.3.1) and (1.3.2) and additionally let

$$\mathbb{H}^{\boldsymbol{\sigma}}_{h,N} := \mathbb{H}^{\boldsymbol{\sigma}}_h \cap \mathbb{H}_N(\mathbf{div}\,;\Omega).$$

Then, the mass and momentum conservative Galerkin scheme associated to (1.4.2) reads: Find $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,N} \times (\mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h)$, such that

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) + \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{u}_h, \varphi_h)) &= F_D(\boldsymbol{\tau}_h) \qquad \forall\, \boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,N}, \\ \mathbf{b}_h(\boldsymbol{\sigma}_h, (\mathbf{v}_h, \psi_h)) &= G_h(\mathbf{v}_h, \psi_h) \quad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h, \end{aligned} \tag{1.4.5}$$

where the form $\mathbf{a}$ is defined in (1.2.4), $\mathbf{b}_h$ and G_h are defined in (1.3.4) and (1.3.5), respectively, and the functional F_D is given by (1.4.3).

The well-posedness of (1.4.5), along with the derivation of the optimal convergence rates, can be established with minor modifications to the analysis presented in Sections 1.3.2 and 1.3.3. The following theorem summarizes these results, with the proofs omitted for brevity.

Theorem 1.4.2 *There exists a unique* $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,N} \times (\mathbf{H}^{\mathbf{u}}_{h,0} \times \Psi^{\varphi}_h)$ *solution to the Galerkin scheme* (1.3.3). *In addition, there exists* $C > 0$, *independent of* h *and* ν, *such that*

$$|\boldsymbol{\sigma}_h|_{\mathbf{div}\,,\Omega} + \|\mathbf{u}_h\|_{0,\Omega} + |\varphi_h|_h \leq C\left(\frac{\|\mathbf{f}\|_{0,\Omega}}{\nu} + \|\mathbf{u}_D\|_{1/2,\Gamma_D}\right).$$

In addition, if $(\boldsymbol{\sigma}, (\mathbf{u}, 0)) \in \mathbb{H}_N(\mathbf{div}\,;\Omega) \times (\mathbf{H}(\mathrm{div}^0;\Omega) \times \mathrm{H}^1_0(\Omega))$ *is the unique solution of* (1.4.2) *satisfying* $\boldsymbol{\sigma} \in \mathbb{H}^2(\Omega)$ *and* $\mathbf{u} \in \mathbf{H}^1(\Omega)$, *then there exist positive constants* $\tilde{c}_1$, $\tilde{c}_2$ *and* $\tilde{c}_3$ *independent of* ν *and* h, *such that,*

$$\|\boldsymbol{\sigma}^{\mathrm{d}} - \boldsymbol{\sigma}^{\mathrm{d}}_h\|_{0,\Omega} \leq \tilde{c}_1 h^2 |\boldsymbol{\sigma}|_{2,\Omega}$$

and

$$\|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}+|\varphi_h|_h \leq \tilde{c}_2 h|\boldsymbol{\sigma}|_{2,\Omega} + \tilde{c}_3 h|\mathbf{u}|_{1,\Omega}.$$

1.5 Numerical tests

In this section we present four numerical examples illustrating the performance of our finite element scheme and confirming the theoretical rates of convergence. We begin by mentioning that the numerical results that follow are attained by imposing the condition of $(\mathrm{tr}\,(\boldsymbol{\sigma}_h),1)_\Omega = 0$ through a penalty strategy using a scalar Lagrange multiplier (adding one row and one column to the system). Also, the divergence-free constraint for the velocity is imposed by means of an appropriate Lagrange multiplier $r_h \in Q_h$. More precisely, we replace the numerical scheme (1.3.3) by the system: Find $(\boldsymbol{\sigma}_h,\mathbf{u}_h,\varphi_h,r_h,\lambda_h)\in \mathbb{H}^{\boldsymbol{\sigma}}_{h,0}\times\mathbf{H}^{\mathbf{u}}_h\times\Psi_h\times Q_h\times\mathbb{R}$, such that:

$$\begin{aligned}
\mathbf{a}(\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h) + \mathbf{b}_h(\boldsymbol{\tau}_h,(\mathbf{u}_h,\varphi_h)) + \lambda_h(\mathrm{tr}\,(\boldsymbol{\tau}_h),1)_\Omega &= F(\boldsymbol{\tau}_h),\\
\mathbf{b}_h(\boldsymbol{\sigma}_h,(\mathbf{v}_h,\psi_h)) + (r_h,\mathrm{div}\,\mathbf{v}_h)_\Omega &= G(\mathbf{v}_h,\psi_h),\\
(s_h,\mathrm{div}\,\mathbf{u}_h)_\Omega &= 0,\\
\eta_h(\,\mathrm{tr}\,(\boldsymbol{\sigma}_h),1)_\Omega &= 0,
\end{aligned}$$

for all $(\boldsymbol{\tau}_h,\mathbf{v}_h,\psi_h,s_h,\eta_h)\in\mathbb{H}^{\boldsymbol{\sigma}}_h\times\mathbf{H}^{\mathbf{u}}_{h,0}\times\Psi^{\varphi}_h\times Q_h\times\mathbb{R}$. Our implementation is based on *Freefem++* code (see [32]), in conjunction with the direct linear solver UMFPACK (see [16]).

Now we introduce some additional notations. In what follows, N stands for the total number of degrees of freedom defining $\mathbb{H}^{\boldsymbol{\sigma}}_h\times\mathbf{H}^{\mathbf{u}}_{h,0}\times\Psi^{\varphi}_h\times Q_h\times\mathbb{R}$ associated to the system (1.3.3), or $\mathbb{H}^{\boldsymbol{\sigma}}_{h,N}\times\mathbf{H}^{\mathbf{u}}_{h,0}\times\Psi^{\varphi}_h\times Q_h$ for the system (1.4.5). We denote the individual errors by

$$\mathsf{e}(\boldsymbol{\sigma}^{\mathrm{d}}) := \|\boldsymbol{\sigma}^{\mathrm{d}}-\boldsymbol{\sigma}^{\mathrm{d}}_h\|_{0,\Omega}\,,\quad \mathsf{e}(\mathbf{u}) := \|\mathbf{u}-\mathbf{u}_h\|_{0,\Omega}\,,\quad \mathsf{e}(p) := \|p-p_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\varphi) := |\varphi_h|_h\,,\quad \mathsf{e}(\mathbf{f}) := \|\mathbf{f}-\mathbf{P}_h(\mathbf{f})\|_{0,\Omega}$$

where p is the exact pressure that can be recovered through the identity $p = -\frac{\nu}{d}\,\mathrm{tr}\,(\boldsymbol{\sigma})$ and the approximate pressure p_h is computed through the postprocessing

formula $p_h = -\frac{\nu}{d}\mathrm{tr}\,(\boldsymbol{\sigma}_h)$.

In addition, we let $r(\%)$ be the experimental rate of convergence given by

$$r(\%) := \frac{\log(\mathsf{e}(\%)/\mathsf{e}'(\%))}{\log(h/h')},$$

where $\mathsf{e}(\%)$ is any of the errors defined above and h and h' are two consecutive meshsizes with errors e and e'.

EXAMPLE 1: RATES OF CONVERGENCE

The first example focuses on illustrating the performance of the two dimensional mixed finite element scheme under a quasi-uniform refinement, by considering manufactured exact solution $(\mathbf{u}, p)$ in the domain $\Omega = (0,1)^2$ given by

$$\mathbf{u}(x_1,x_2) = \begin{pmatrix} \pi\exp(x_1)\cos(\pi x_2) \\ \\ -\exp(x_1)\sin(\pi x_2) \end{pmatrix}, \quad p(x_1,x_2) = x_1^3 + x_2^3 - 0.5\,.$$

In this case,

$$\mathbf{f} = \begin{pmatrix} 3x_1^2 + \nu\pi\exp(x_1)\cos(\pi x_2)(\pi^2-1) \\ \\ 3x_2^2 - \nu\exp(x_1)\sin(\pi 2x_2)(\pi^2-1) \end{pmatrix}.$$

In Table 1.1, we summarize the convergence history for a sequence of quasi-uniform triangulations, considering the viscosity $\nu = 1$ and ν =1.0E-3. We see there that the rate of convergence provided by Theorem 1.3.4 is attained by all the unknowns. We emphasize that, when comparing the errors obtained for $\nu = 1$ and ν=1.0E-3, the numerical results indicate that these errors are not amplified by the factor ν^{-1}. In addition, the l^∞–norm of $\mathrm{div}\,\mathbf{u}_h$ in each mesh is close to 0 which shows that this method is mass conserving. From the columns corresponding to $\|\mathbf{div}\,\boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}\|_{0,\Omega}$ and $\mathsf{e}(\mathbf{f})$, we observe that for $\nu = 1.0$, both columns exhibit similar magnitudes. In contrast, for ν=1.0E-3, the column corresponding to $\|\mathbf{div}\,\boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}\|_{0,\Omega}$ matches the column for $\mathsf{e}(\mathbf{f})$ scaled by a factor of ν^{-1}, and the convergence rate is of order 1. This numerical evidence confirms the theoretical prediction given in (1.3.12).

EXAMPLE 2: MOMENTUM CONSERVATION

The second example addresses the momentum conservation of the method when

NUMERICAL RESULTS FOR ν=1.0

N	h	$\mathsf{e}(\boldsymbol{\sigma}^{\mathrm{d}})$	$r(\boldsymbol{\sigma}^{\mathrm{d}})$	$\mathsf{e}(\mathbf{u})$	$r(\mathbf{u})$	$\mathsf{e}(p)$	$r(p)$	$\mathsf{e}(\varphi)$	$r(\varphi_h)$
409	0.372	3.941E-1	–	1.138	–	3.253E-1	–	0.765	–
1637	0.190	0.741E-1	2.483	0.487	1.261	0.557E-1	2.619	0.323	1.279
6213	0.095	0.192E-1	1.944	0.246	0.987	0.148E-1	1.909	0.160	1.017
24445	0.049	0.046E-1	2.178	0.122	1.058	0.036E-1	2.132	0.080	1.052
97129	0.024	0.012E-1	1.898	0.062	0.970	0.009E-1	1.993	0.040	0.960
391577	0.014	0.003E-1	2.581	0.031	1.268	0.002E-1	2.447	0.020	1.280

$\|\operatorname{div}\mathbf{u}_h\|_{l^\infty}$	$\|\mathbf{div}\,\boldsymbol{\sigma}_h+\nu^{-1}\mathbf{f}\|_{0,\Omega}$	$\mathsf{e}(\mathbf{f})$	$r(\mathbf{f})$
7.1E-15	7.500	7.531	–
5.7E-14	3.244	3.248	1.249
8.7E-14	1.662	1.662	0.966
2.3E-13	0.823	0.823	1.063
4.5E-13	0.416	0.417	0.977
9.1E-13	0.206	0.206	1.258

NUMERICAL RESULTS FOR ν=1.0E-3

N	h	$\mathsf{e}(\boldsymbol{\sigma}^{\mathrm{d}})$	$r(\boldsymbol{\sigma}^{\mathrm{d}})$	$\mathsf{e}(\mathbf{u})$	$r(\mathbf{u})$	$\mathsf{e}(p)$	$r(p)$	$\mathsf{e}(\varphi)$	$r(\varphi_h)$
409	0.372	6.083	–	1.148	–	1.418E-2	–	0.777	–
1637	0.190	1.276	2.320	0.488	1.261	0.300E-2	2.307	0.322	1.311
6213	0.095	0.323	1.983	0.246	0.987	0.072E-2	2.049	0.159	1.014
24445	0.049	0.078	2.139	0.122	1.058	0.017E-2	2.159	0.080	1.049
97129	0.024	0.019	1.999	0.062	0.970	0.005E-2	1.934	0.040	0.960
391577	0.014	0.005	2.526	0.031	1.268	0.001E-2	2.537	0.020	1.280

$\|\operatorname{div}\mathbf{u}_h\|_{l^\infty}$	$\|\mathbf{div}\,\boldsymbol{\sigma}_h+\nu^{-1}\mathbf{f}\|_{0,\Omega}$	$\mathsf{e}(\mathbf{f})$	$r(\mathbf{f})$
9.3E-15	288.009	0.288	–
2.8E-14	132.239	0.132	1.158
5.7E-14	65.772	0.065	1.008
1.7E-13	32.422	0.032	1.068
3.4E-13	16.399	0.016	0.989
9.1E-13	8.119	0.008	1.257

Table 1.1: EXAMPLE 1: Degrees of freedom, mesh sizes, errors, convergence rates, L^∞-norm of $\operatorname{div}\mathbf{u}_h$, L^2-norm of the discrete momentum equation, error in the projection of $\mathbf{f}$, and its convergence rate for the Galerkin scheme with $\nu = 1.0$ and $\nu =$ 1.0E-3.

the datum $\mathbf{f} \in \mathbf{Q}_h$. To that end, we consider the manufactured solution $(\mathbf{u}, p)$ in the

domain $\Omega = (0,1)^2$ given by:

$$\mathbf{u}(x_1, x_2) = \begin{pmatrix} x_2^2 \\ \\ -x_1^2 \end{pmatrix}, \quad p(x_1, x_2) = x_1 + x_2 - 1\,,$$

so that the datum $\mathbf{f}$ becomes,

$$\mathbf{f} = \begin{pmatrix} 1 - 2\nu \\ \\ 1 + 2\nu \end{pmatrix} \in \mathbf{Q}_h.$$

We run the code for viscosity values $\nu = 1.0$ and $\nu = 1.0\text{E-}3$, using a sequence of quasi-uniform triangulations.

Table 1.2 presents the l^∞-norm of $\operatorname{div} \mathbf{u}_h$ and $\mathbf{div}\, \boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}$. From these results, we observe that both quantities remain close to zero, confirming that the method is mass conservative and preserves momentum when $\mathbf{f} \in \mathbf{Q}_h$.

Furthermore, by comparing both tables, we observe that the l^∞-norm of $\mathbf{div}\, \boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}$ scales with ν^{-1}, whereas the l^∞-norm of $\operatorname{div} \mathbf{u}_h$ appears to remain unaffected by ν.

Example 3: Lack of pressure robustness

The third example examines the lack of pressure robustness of the method. To this end, we use the data presented in [35, Example 1.1]. Specifically, we consider the exact solutions:

$$\mathbf{u}(x_1, x_2) = \mathbf{0} \quad \text{and} \quad p(x_1, x_2) = Ra \left(x_2^3 - \frac{1}{2} x_2^2 + x_2 - \frac{7}{12} \right),$$

which results in the forcing term:

$$\mathbf{f} = \begin{pmatrix} 0 \\ \\ Ra(1 - x_2 + 3x_2^2) \end{pmatrix} \in \mathbf{Q}_h,$$

where $Ra > 0$ is a given parameter that affects only the pressure and will be assigned different values in the analysis.

In Table 1.3 we present the errors for each variable for different values of Ra.

Mass and momentum conservation for ν=1.0

h	$\|\mathrm{div}\,\mathbf{u}_h\|_{l^\infty}$	$\|\mathbf{div}\,\boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}\|_{l^\infty}$
0.372	1.78E-15	3.55E-15
0.190	4.44E-15	1.24E-14
0.095	1.42E-14	2.84E-14
0.049	2.84E-14	1.14E-13
0.024	5.68E-14	2.27E-10
0.014	1.42E-13	4.55E-10

Mass and momentum conservation for ν=1.0E-3

h	$\|\mathrm{div}\,\mathbf{u}_h\|_{l^\infty}$	$\|\mathbf{div}\,\boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}\|_{l^\infty}$
0.372	1.78E-15	2.96E-12
0.190	3.55E-15	7.05E-12
0.095	1.07E-14	1.82E-11
0.049	2.84E-14	4.37E-11
0.024	5.68E-14	1.02E-10
0.014	1.26E-13	2.91E-10

Table 1.2: Example 2: Mesh sizes, L^∞-norm of $\mathrm{div}\,\mathbf{u}_h$ and $\mathbf{div}\,\boldsymbol{\sigma}_h + \nu^{-1}\mathbf{f}$, with ν = 1.0 and ν = 1.0E-3.

We consider a fix triangulation of size $h = 0.0244$. There, it can be appreciated that when Ra increases, all the errors increase in the same order, which confirms the aforementioned lack of pressure robustness of the method.

Lack of pressure robustness

Ra	$\mathsf{e}(\boldsymbol{\sigma}^{\mathrm{d}})$	$\mathsf{e}(\mathbf{u})$	$\mathsf{e}(p)$	$\mathsf{e}(\varphi)$
1.0E+0	1.593E-5	8.903E-8	2.759E-5	4.123E-8
1.0E+1	1.593E-4	8.903E-7	2.759E-4	4.123E-7
1.0E+2	1.593E-3	8.903E-6	2.759E-3	4.123E-6
1.0E+3	1.593E-2	8.903E-5	2.759E-2	4.123E-5
1.0E+4	1.593E-1	8.903E-4	2.759E-1	4.123E-4

Table 1.3: Example 3: Errors for each variable, considering different values of Ra.

Example 4: Backward-facing step flow

Finally, the fourth example examines mass loss in the standard backward-facing step flow test, similarly as in [5]. For this test, we consider a rectangular domain $\Omega = [0, 10] \times [0, 1]$ with a re-entrant corner at $(2, 0.5)$. The boundary Γ is partitioned

into three segments: the inflow boundary Γ_{in}, the outflow boundary Γ_{out}, and the wall boundary Γ_{wall}, where $\Gamma_{wall} = \Gamma \setminus \left(\overline{\Gamma}_{in} \cup \overline{\Gamma}_{out}\right)$ (see Figure 1.1). We consider

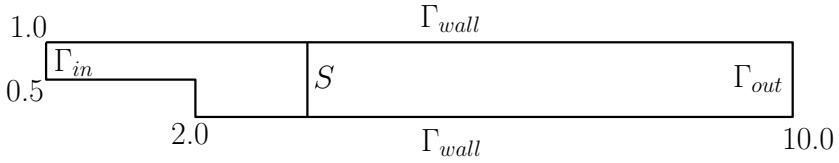

Figure 1.1: Example 4: Geometry for the backward-facing step flow test.

$\nu = 1.0$ and $\mathbf{f}(x_1, x_2) = \mathbf{0}$ in Ω. The boundary conditions are prescribed as follows: a parabolic inflow profile $\mathbf{u}_D(x_1, x_2) = (8(x_2 - 0.5)(1 - x_2), 0)^t$ on Γ_{in}, a no-slip condition $\mathbf{u}_D(x_1, x_2) = \mathbf{0}$ on Γ_{wall}, and, unlike [5], a *do-nothing* boundary condition is imposed on Γ_{out}, given by $\boldsymbol{\sigma}\mathbf{n} = \mathbf{0}$ on Γ_{out}.

To evaluate mass conservation in the discrete solution, we measure the total mass flow across a sequence of vertical surfaces connecting the top and bottom boundaries of the computational domain. The line labeled "S" in Figure 1.1 illustrates a representative example of such a surface for the test problem.

Since $\operatorname{div} \mathbf{u} = 0$ in Ω, from the divergence theorem it follows that

$$\int_{\Gamma_{in}} \mathbf{u} \cdot \mathbf{n} = \int_{S} \mathbf{u} \cdot \mathbf{n}_S,$$

for any S connecting the top and bottom walls of the domain. Then, suggested by the above, in what follows, mass conservation in the discrete solution will be quantified by the percentage mass loss across the surface S, defined as

$$\%m_{loss} := 100 \frac{\left| \displaystyle\int_{\Gamma_{in}} \mathbf{u}_h \cdot \mathbf{n} - \int_{S} \mathbf{u}_h \cdot \mathbf{n}_S \right|}{\left| \displaystyle\int_{\Gamma_{in}} \mathbf{u}_h \cdot \mathbf{n} \right|}. \tag{1.5.1}$$

We compare the mass loss in (1.4.5) against the standard discrete pseudostress-based scheme for (1.4.1), formulated as follows: find $\boldsymbol{\sigma}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,N}$ and $\mathbf{u}_h \in \mathbf{Q}_h$ such that

$$\begin{aligned} (\boldsymbol{\sigma}_h^{\mathrm{d}}, \boldsymbol{\tau}_h^{\mathrm{d}})_\Omega + (\mathbf{u}_h, \mathbf{div}\, \boldsymbol{\tau}_h)_\Omega &= \langle \boldsymbol{\tau}_h \mathbf{n}, \mathbf{u}_D \rangle_{\Gamma_D}, \quad \forall \boldsymbol{\tau}_h \in \mathbb{H}^{\boldsymbol{\sigma}}_{h,N}, \\ (\mathbf{v}_h, \mathbf{div}\, \boldsymbol{\sigma}_h)_\Omega &= -\frac{1}{\nu}(\mathbf{f}, \mathbf{v}_h)_\Omega, \quad \forall \mathbf{v}_h \in \mathbf{Q}_h, \end{aligned} \tag{1.5.2}$$

with $\Gamma_N = \Gamma_{out}$ and $\Gamma_D = \Gamma \setminus \Gamma_{out}$.

We use a computational grid consisting of 120.926 triangles with a mesh size of $h = 0.0212$, leading to a total of $N = 1.214.540$ degrees of freedom for the mass-conservative scheme (1.4.5) and $N = 1.274.123$ for (1.5.2). Notice that, despite the fact that (1.4.5) involves three unknowns while (1.5.2) considers only two, the former is slightly less computationally expensive than the latter.

The results of our study are summarized in Figure 1.2, where we compare the mass losses obtained using formulation (1.4.5) with those from (1.5.2), considering 100 lines S equally distributed in Ω. The figure clearly demonstrates a significant improvement in mass conservation, as quantified by the percent mass loss formula (1.5.1). Specifically, in (1.4.5), the maximum mass loss remains below 0.1%, whereas in (1.5.2), it exceeds 1.0%.

In Figure 1.3, we present the pressure distribution (top panel), velocity magnitude (center panel), and streamlines (bottom panel) obtained using the scheme (1.4.5). The results exhibit the expected behavior: high pressure near the inlet, a characteristic parabolic velocity profile throughout the full length of the domain with higher velocity near the inlet, and the formation of the expected vortex below the re-entrant corner.

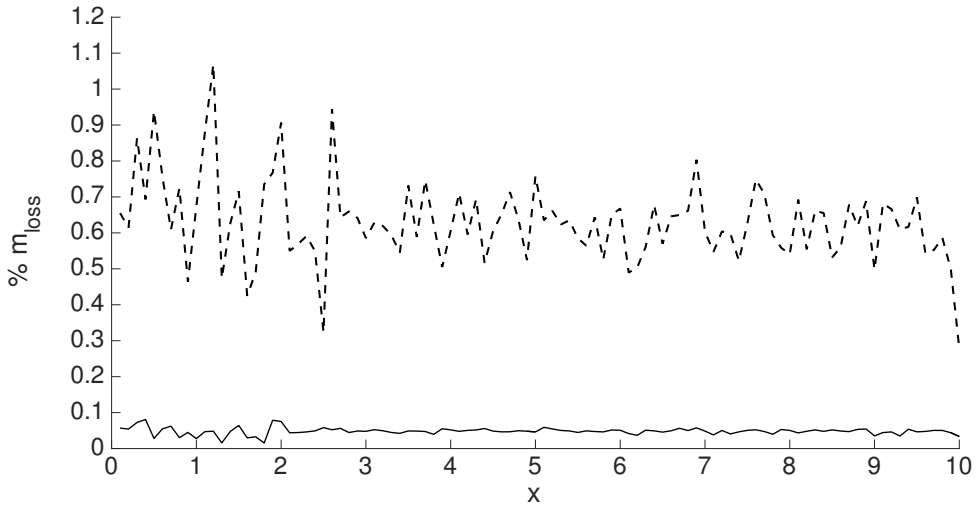

Figure 1.2: Example 4: Comparison of mass losses in (1.4.5) and (1.5.2) using 100 equally spaced lines in Ω.

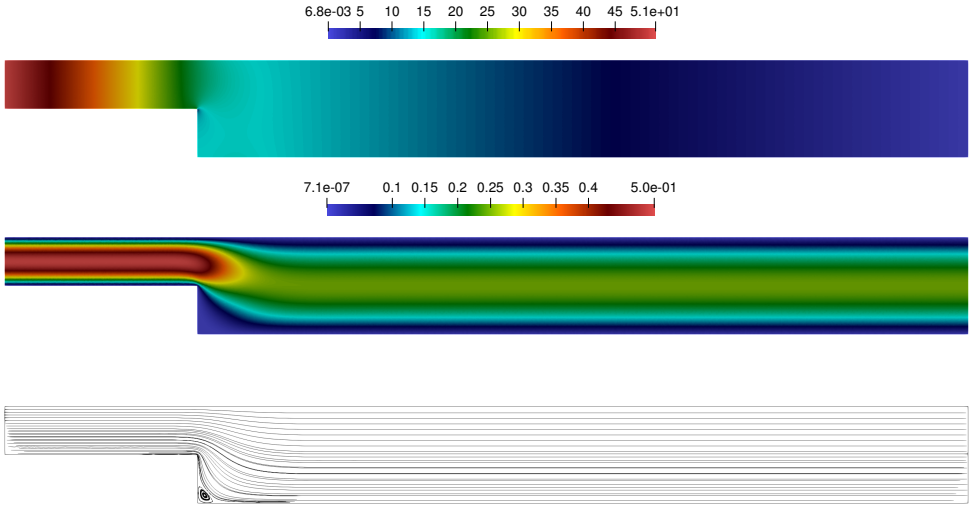

Figure 1.3: Example 4: Pressure (top panel), velocity magnitude (center panel) and streamlines (bottom panel).

Chapter 2

A momentum and mass conservative pseudostress-based mixed finite element method for the Navier–Stokes problem

In this chapter, we extend the analysis provided in Chapter 1 for the Stokes problem to the Navier–Stokes problem.

2.1 Preliminaries

In this section we introduce some notations and definitions that will be employed for the rest of the chapter.

Let $\Omega \subseteq \mathbb{R}^2$, be a bounded domain with a polygonal connected boundary Γ, and let $\mathbf{n}$ be the outward unit normal vector on Γ. We will use standard notations for the Lebesgue spaces $\mathrm{L}^p(\Omega)$, with $p > 1$ and the Sobolev spaces $\mathrm{W}^{r,p}(\Omega)$ with $r \geq 0$, endowed with the norms $\|\cdot\|_{0,p;\Omega}$ and $\|\cdot\|_{r,p;\Omega}$, respectively. Note that $\mathrm{W}^{0,p}(\Omega) = \mathrm{L}^p(\Omega)$ and if $p = 2$, we write $\mathrm{H}^r(\Omega)$ instead of $\mathrm{W}^{r,2}(\Omega)$, with the corresponding Lebesgue and Sobolev norms, denoted by $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{r,\Omega}$, respectively. We also write $|\cdot|_{r,\Omega}$ for the H^r-seminorm. On the other hand, $\mathrm{W}_0^{r,p}(\Omega)$ denotes the closure in $\mathrm{W}^{r,p}(\Omega)$ of all distribution with compact support in Ω which belongs to $\mathrm{W}^{r,p}(\Omega)$. In addition, $\mathrm{H}^{1/2}(\Gamma)$ is the trace space of functions of $\mathrm{H}^1(\Omega)$, and $\mathrm{H}^{-1/2}(\Gamma)$ is its

dual. The corresponding duality product between $\mathrm{H}^{1/2}(\Gamma)$ and $\mathrm{H}^{-1/2}(\Gamma)$ is denoted by $\langle\cdot,\cdot\rangle_\Gamma$. Furthermore, we will denote by $\mathbf{S}$ and $\mathbb{S}$ the corresponding vectorial and tensorial counterparts of the generic scalar functional space S. In turn, for any two vector fields $\mathbf{v}=(v_i)_{i=1,2}$ and $\mathbf{w}=(w_i)_{i=1,2}$ we define the gradient, divergence and tensor product operators, as follows

$$\nabla\mathbf{v}:=\left(\frac{\partial v_i}{\partial x_j}\right)_{i,j=1,2},\quad \mathrm{div}\,\mathbf{v}:=\sum_{j=1}^{2}\frac{\partial v_j}{\partial x_j},\quad \text{and}\quad \mathbf{v}\otimes\mathbf{w}:=(v_iw_j)_{i,j=1,2}.$$

Furthermore, for arbitrary tensor fields $\boldsymbol{\tau}=(\tau_{ij})_{i,j=1,2}$ and $\boldsymbol{\zeta}=(\zeta_{ij})_{i,j=1,2}$, $\mathbf{div}\,\boldsymbol{\tau}$ is defined as the divergence operator div acting along the rows of $\boldsymbol{\tau}$. The transpose, trace, tensor product, and deviatoric tensor are defined respectively as

$$\boldsymbol{\tau}^t:=(\boldsymbol{\tau}_{ji})_{j,i=1,2},\quad \mathrm{tr}\,(\boldsymbol{\tau}):=\sum_{i=1}^{2}\tau_{ii},\quad \boldsymbol{\tau}:\boldsymbol{\zeta}:=\sum_{i,j=1}^{2}\tau_{ij}\zeta_{ij},\quad \text{and}\quad \boldsymbol{\tau}^{\mathrm{d}}:=\boldsymbol{\tau}-\frac{1}{2}\mathrm{tr}\,(\boldsymbol{\tau})\mathbb{I},$$

where $\mathbb{I}$ is the identity tensor in $\mathbb{R}^{2\times 2}$. For simplicity, in what follows, we will denote

$$(v,w)_\Omega:=\int_\Omega vw,\quad (\mathbf{v},\mathbf{w})_\Omega:=\int_\Omega\mathbf{v}\cdot\mathbf{w},\quad \text{and}\quad (\boldsymbol{\tau},\boldsymbol{\zeta})_\Omega:=\int_\Omega\boldsymbol{\tau}:\boldsymbol{\zeta}.$$

On the other hand, given $p>1$, we define the Banach space $\mathbf{H}^p(\mathrm{div};\Omega)$ as

$$\mathbf{H}^p(\mathrm{div};\Omega):=\{\mathbf{u}\in\mathbf{L}^p(\Omega):\quad \mathrm{div}\,\mathbf{u}\in\mathrm{L}^2(\Omega)\},$$

equipped with the norm

$$\|\mathbf{u}\|_{p,\mathrm{div};\Omega}:=\left\{\|\mathbf{u}\|_{0,p;\Omega}^2+\|\mathrm{div}\,\mathbf{u}\|_{0,\Omega}^2\right\}^{1/2}.$$

In particular, for the case $p=2$, we simply denote $\mathbf{H}(\mathrm{div};\Omega)=\mathbf{H}^2(\mathrm{div};\Omega)$, and define the subspace

$$\mathbf{H}^p(\mathrm{div}^0;\Omega):=\{\mathbf{u}\in\mathbf{H}^p(\mathrm{div};\Omega):\quad \mathrm{div}\,\mathbf{u}=0\quad\text{in}\quad\Omega\}.$$

Additionally, we make use of the tensor version of $\mathbf{H}(\mathrm{div};\Omega)$, namely

$$\mathbb{H}(\mathbf{div}\,;\Omega):=\{\boldsymbol{\tau}\in\mathbb{L}^2(\Omega):\quad \mathbf{div}\,\boldsymbol{\tau}\in\mathbf{L}^2(\Omega)\},$$

whose norm will be denoted $\|\cdot\|_{\mathbf{div}_p;\Omega}$. In turn, given $p > 1$, in what follows we will also use the non-standard Banach space $\mathbb{H}(\mathbf{div}_p;\Omega)$, defined by

$$\mathbb{H}(\mathbf{div}_p;\Omega) := \{\boldsymbol{\tau} \in \mathbb{L}^2(\Omega) : \quad \mathbf{div}\ \boldsymbol{\tau} \in \mathbf{L}^p(\Omega)\},$$

endowed with the norm

$$\|\boldsymbol{\tau}\|_{p,\mathbf{div}_p;\Omega} := \left\{\|\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\mathbf{div}\,\boldsymbol{\tau}\|_{0,p;\Omega}^2\right\}^{1/2}.$$

The rest of this chapter is organized as follows: In Section 2.2 we present the main aspects of the continuous problem. We reformulate the problem as an equivalent set of equations and derive the mixed variational formulation. In Section 2.3 we introduce the fixed-point strategy and apply, firstly, the classical Banach–Nečas–Babuška theorem, and secondly, the Banach fixed-point theorem, to show that the associated fixed-point operator is well-defined and that the continuous problem is uniquely solvable, respectively. Next, in Section 2.4 we introduce and analyze the associated Galerkin scheme by mimicking the theory developed for the continuous problem. Furthermore, we establish the corresponding a priori error estimate and prove the optimal convergence of the method. Finally, in Section 2.5 we present several numerical results illustrating the good performance of our scheme.

2.2 The model problem and its variational formulation

In this section we present the model problem and derive the variational formulation. We consider the stationary Navier–Stokes equations, that is

$$\begin{aligned} -\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p &= \mathbf{f} && \text{in} \quad \Omega, \\ \operatorname{div}\mathbf{u} &= 0 && \text{in} \quad \Omega, \\ \mathbf{u} &= \mathbf{u}_D && \text{on} \quad \Gamma, \\ (p,1)_\Omega &= 0, \end{aligned} \tag{2.2.1}$$

where the unknowns are the velocity field $\mathbf{u}$ and the pressure p of a fluid occupying the region Ω. The given data are $\nu > 0$ describing the fluid viscosity, an external force $\mathbf{f}$ acting on Ω, and the boundary velocity $\mathbf{u}_D$ on Γ. Note that $\mathbf{u}_D$ must satisfy the compatibility condition

$$(\mathbf{u}_D \cdot \mathbf{n}, 1)_\Gamma = 0, \tag{2.2.2}$$

which comes from the incompressibility condition of the fluid.

Now, in order to derive our mixed approach, we begin by introducing the pseudo-stress tensor

$$\boldsymbol{\sigma} := \nabla \mathbf{u} - \frac{1}{\nu}(\mathbf{u} \otimes \mathbf{u}) - \frac{1}{\nu} p\, \mathbb{I} \quad \text{in} \quad \Omega. \tag{2.2.3}$$

Observe that from the incompressibility condition $\mathrm{tr}\,(\nabla \mathbf{u}) = \mathrm{div}\,\mathbf{u} = 0$ in Ω, there hold

$$\mathbf{div}\,(\mathbf{u} \otimes \mathbf{u}) = (\mathbf{u} \cdot \nabla)\mathbf{u} \quad \text{in} \quad \Omega \quad \text{and} \quad \mathrm{tr}\,(\boldsymbol{\sigma}) = -\frac{1}{\nu}(\mathrm{tr}\,(\mathbf{u} \otimes \mathbf{u}) + 2p) \quad \text{in} \quad \Omega.$$

According to the above, we can rewrite equations (2.2.1), equivalently, as follows

$$\begin{aligned} \boldsymbol{\sigma}^{\mathrm{d}} = \nabla \mathbf{u} - \frac{1}{\nu}(\mathbf{u} \otimes \mathbf{u})^{\mathrm{d}} \quad \text{in} \quad \Omega, \quad -\mathbf{div}\ \boldsymbol{\sigma} = \frac{1}{\nu}\mathbf{f} \quad \text{in} \quad \Omega, \\ \mathbf{u} = \mathbf{u}_D \quad \text{on} \quad \Gamma, \quad (\mathrm{tr}\,(\boldsymbol{\sigma}), 1)_\Omega = -\frac{1}{\nu}(\mathrm{tr}\,(\mathbf{u} \otimes \mathbf{u}), 1)_\Omega, \end{aligned} \tag{2.2.4}$$

where the unknowns are the velocity $\mathbf{u}$ and the tensor $\boldsymbol{\sigma}$. The pressure can be easily calculated as a postprocess of the solution using

$$p = -\frac{1}{2}(\mathrm{tr}\,(\mathbf{u} \otimes \mathbf{u}) + \nu\, \mathrm{tr}\,(\boldsymbol{\sigma})) \quad \text{in} \quad \Omega. \tag{2.2.5}$$

We now proceed to derive our momentum and mass conservative mixed variational formulation. Initially we recall from [12, Section 2.3] the weak formulation of (2.2.4). To that end, we define the spaces $\mathbb{X} := \mathbb{H}(\mathbf{div}_{4/3}; \Omega)$, $\mathbf{M} := \mathbf{L}^4(\Omega)$ and

$$\mathbb{X}_0 := \mathbb{H}_0(\mathbf{div}_{4/3}; \Omega) := \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}_{4/3}; \Omega) : \quad (\mathrm{tr}\,(\boldsymbol{\tau}), 1)_\Omega = 0\},$$

and observe that the following descomposition holds:

$$\mathbb{X} = \mathbb{X}_0 \oplus P_0(\Omega)\mathbb{I}, \tag{2.2.6}$$

where $P_0(\Omega)$ is the space of constant polynomials on Ω.

Equivalently, each $\boldsymbol{\tau} \in \mathbb{X}$ can be descomposed as $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + d\mathbb{I}$, with

$$\boldsymbol{\tau}_0 := \boldsymbol{\tau} - \left(\frac{1}{2|\Omega|}(\mathrm{tr}\ \boldsymbol{\tau}, 1)_\Omega\right)\mathbb{I} \in \mathbb{X}_0 \quad \text{and} \quad d := \frac{1}{2|\Omega|}(\mathrm{tr}\ \boldsymbol{\tau}, 1)_\Omega \in \mathbb{R}. \tag{2.2.7}$$

In particular, decomposing $\boldsymbol{\sigma}$ in (2.2.4) as $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + c\mathbb{I}$, with $\boldsymbol{\sigma}_0 \in \mathbb{X}_0$, we deduce from (2.2.7) and the last equation in (2.2.4) that c is given explicity in terms of $\mathbf{u}$ as

$$c = -\frac{1}{2\nu|\Omega|}(\mathrm{tr}\,(\mathbf{u}\otimes\mathbf{u}), 1)_\Omega,$$

and thus,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 - \left(\frac{1}{2\nu|\Omega|}(\mathrm{tr}\ (\mathbf{u}\otimes\mathbf{u}), 1)_\Omega\right)\mathbb{I}. \tag{2.2.8}$$

In this way, since $\boldsymbol{\sigma}^{\mathrm{d}} = \boldsymbol{\sigma}_0^{\mathrm{d}}$ and $\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{div}\,\boldsymbol{\sigma}_0$, throughout the rest of the paper we rename $\boldsymbol{\sigma}_0$ as $\boldsymbol{\sigma} \in \mathbb{X}_0$ and realize that the first, second and third equations of (2.2.4) remain unchanged.

Thus, multiplying the first equation of (2.2.4) by a test function $\boldsymbol{\tau} \in \mathbb{X}$, integrating by parts, utilizing the Dirichlet boundary condition $\mathbf{u} = \mathbf{u}_D$ on Γ, and the identity $\boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{\tau} = \boldsymbol{\sigma}^{\mathrm{d}} : \boldsymbol{\tau}^{\mathrm{d}}$, we obtain

$$(\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega + (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{u})_\Omega + \frac{1}{\nu}(\mathbf{u}\otimes\mathbf{u}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega = \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D\rangle_\Gamma, \quad \forall\, \boldsymbol{\tau} \in \mathbb{X}.$$

In addition, the equilibrium equation $-\mathbf{div}\ \boldsymbol{\sigma} = \dfrac{1}{\nu}\mathbf{f}$ is imposed weakly as follows

$$(\mathbf{div}\ \boldsymbol{\sigma}, \mathbf{v})_\Omega = -\frac{1}{\nu}(\mathbf{f}, \mathbf{v})_\Omega \quad \forall\, \mathbf{v} \in \mathbf{M}.$$

Therefore, the variational formulation of (2.2.4) is as follows: Find $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$, such that

$$\begin{aligned} (\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega + (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{u})_\Omega + \frac{1}{\nu}(\mathbf{u}\otimes\mathbf{u}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega &= \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D\rangle_\Gamma, \\ (\mathbf{div}\ \boldsymbol{\sigma}, \mathbf{v})_\Omega &= -\frac{1}{\nu}(\mathbf{f}, \mathbf{v})_\Omega, \end{aligned} \tag{2.2.9}$$

for all $(\boldsymbol{\tau}, \mathbf{v}) \in \mathbb{X} \times \mathbf{M}$.

Now, let us consider the Helmholtz–Weyl decomposition (see [40, Theorem 4.4]):

$$\mathbf{L}^4(\Omega) = \mathbf{H}^4(\mathrm{div}^0;\Omega) \oplus \nabla W_0^{1,4}(\Omega). \tag{2.2.10}$$

Our aim is to reformulate (2.2.9) to obtain a numerical scheme that conserves mass and momentum, using the decomposition (2.2.10). Given that $\mathbf{u},\mathbf{v} \in \mathbf{L}^4(\Omega)$ and thanks to the previous decomposition (2.2.10), we know that there exist $\mathbf{w},\mathbf{z} \in \mathbf{H}^4(\mathrm{div}^0;\Omega)$ and $\varphi,\psi \in W_0^{1,4}(\Omega)$ such that $\mathbf{u} = \mathbf{w} + \nabla\varphi$ and $\mathbf{v} = \mathbf{z} + \nabla\psi$. Therefore, our problem reduces to: Find $(\boldsymbol{\sigma},(\mathbf{w},\varphi)) \in \mathbb{X}_0 \times (\mathbf{V}_0 \times \Psi_0)$ such that:

$$\begin{aligned}(\boldsymbol{\sigma}^{\mathrm{d}},\boldsymbol{\tau}^{\mathrm{d}})_\Omega + (\mathbf{div}\ \boldsymbol{\tau},\mathbf{w}+\nabla\varphi)_\Omega + \frac{1}{\nu}((\mathbf{w}+\nabla\varphi)\otimes(\mathbf{w}+\nabla\varphi),\boldsymbol{\tau}^{\mathrm{d}})_\Omega &= \langle\boldsymbol{\tau}\mathbf{n},\mathbf{u}_D\rangle_\Gamma,\\ (\mathbf{div}\ \boldsymbol{\sigma},\mathbf{z}+\nabla\psi)_\Omega &= -\frac{1}{\nu}(\mathbf{f},\mathbf{z}+\nabla\psi)_\Omega,\end{aligned} \tag{2.2.11}$$

for all $(\boldsymbol{\tau},(\mathbf{z},\psi)) \in \mathbb{X} \times (\mathbf{V}_0 \times \Psi_0)$, where

$$\mathbf{V}_0 := \mathbf{H}^4(\mathrm{div}^0;\Omega) \quad \text{and} \quad \Psi_0 := \mathrm{W}_0^{1,4}(\Omega).$$

On the other hand, we recall from [18, Lemma B.66] the classic Poincaré inequality

$$\|v\|_{1,p;\Omega} \leq C_{p,\Omega}\|\nabla v\|_{0,p;\Omega} \quad \forall\, v \in \mathrm{W}_0^{1,p}(\Omega), \tag{2.2.12}$$

with $1 \leq p < +\infty$ and $C_{p,\Omega} > 0$; for $p = 2$, we denote $C_{2,\Omega} = C_\Omega$. This inequality proves that on $\mathrm{W}_0^{1,p}(\Omega)$, the seminorm $|v|_{1,p;\Omega}$ is equivalent to the usual norm $\|v\|_{1,p;\Omega}$, i.e., $\|v\|_{1,p;\Omega} \sim |v|_{1,p;\Omega} = \|\nabla v\|_{0,p;\Omega}$, which will be used in the subsequent results.

Additionally, we will make use of the following result to prove the equivalence of the problems (2.2.9) and (2.2.11).

Lemma 2.2.1 *There exists $\beta_0 > 0$, such that*

$$\sup_{0\neq\psi\in W_0^{1,4/3}(\Omega)} \frac{(\nabla\psi,\nabla\varphi)_\Omega}{|\psi|_{1,4/3;\Omega}} \geq \beta_0|\varphi|_{1,4;\Omega} \quad \forall\varphi \in \mathrm{W}_0^{1,4}(\Omega). \tag{2.2.13}$$

Proof. In fact, given $\varphi \in \mathrm{W}_0^{1,4}(\Omega)$, we observe that $\nabla\varphi|\nabla\varphi|^2 \in \mathbf{L}^{4/3}(\Omega)$, because

$$(|\nabla\varphi|\nabla\varphi|^2|^{4/3},1)_\Omega = (|\nabla\varphi|^{4/3}|\nabla\varphi|^{8/3},1)_\Omega = (|\nabla\varphi|^4,1)_\Omega < +\infty.$$

Then, again thanks to [40, Theorem 4.4], we know that

$$\mathbf{L}^{4/3}(\Omega) = \mathbf{H}^{4/3}(\mathrm{div}^0;\Omega) \oplus \nabla \mathrm{W}_0^{1,4/3}(\Omega),$$

so that there exist $\mathbf{t} \in \mathbf{H}^{4/3}(\mathrm{div}^0;\Omega)$ and $\chi \in \mathrm{W}_0^{1,4/3}(\Omega)$ such that

$$\nabla\varphi|\nabla\varphi|^2 = \mathbf{t} + \nabla\chi \tag{2.2.14}$$

and also, since the decomposition is stable (see [23, Lemma 3.20]), there exists a constant $c > 0$ such that

$$\|\mathbf{t}\|_{4/3,\mathrm{div};\Omega} + |\chi|_{1,4/3;\Omega} \leq c\|\nabla\varphi|\nabla\varphi|^2\|_{0,4/3;\Omega} = c|\varphi|^3_{1,4;\Omega}. \tag{2.2.15}$$

So, using (2.2.14), integration by parts and (2.2.15) we obtain that

$$\begin{aligned}
\sup_{0\neq\psi\in W_0^{1,4/3}(\Omega)} \frac{(\nabla\psi,\nabla\varphi)_\Omega}{|\psi|_{1,4/3;\Omega}} &\geq \frac{(\nabla\chi,\nabla\varphi)_\Omega}{|\chi|_{1,4/3;\Omega}} = \frac{((\nabla\varphi|\nabla\varphi|^2 - \mathbf{t}),\nabla\varphi)_\Omega}{|\chi|_{1,4/3;\Omega}} \\
&\geq \frac{(\nabla\varphi|\nabla\varphi|^2,\nabla\varphi)_\Omega - (\mathbf{t},\nabla\varphi)_\Omega}{c|\varphi|^3_{1,4;\Omega}} = \frac{|\varphi|^4_{1,4;\Omega}}{c|\varphi|^3_{1,4;\Omega}} \\
&= \beta_0|\varphi|_{1,4;\Omega}.
\end{aligned}$$

where $\beta_0 := 1/c$. □

Observe that if $(\boldsymbol{\sigma},(\mathbf{w},\varphi)) \in \mathbb{X}_0 \times (\mathbf{V}_0 \times \Psi_0)$ is a solution to (2.2.11), then taking $\boldsymbol{\tau} = \psi\,\mathbb{I}$, with $\psi \in \mathrm{W}_0^{1,4/3}(\Omega)$ in the first equation of (2.2.11), and using the fact that $\boldsymbol{\tau}^{\mathrm{d}} = (\psi\,\mathbb{I})^{\mathrm{d}} = 0$, $\langle\psi\mathbf{n},\mathbf{u}_D\rangle_\Gamma = 0$ and $\mathbf{div}\,(\psi\,\mathbb{I}) = \nabla\psi$, it follows that

$$(\nabla\psi,\nabla\varphi)_\Omega = 0, \qquad \forall\,\psi \in \mathrm{W}_0^{1,4/3}(\Omega).$$

Then, thanks to the inf-sup condition (2.2.13), it follows that $\nabla\varphi = 0$ in Ω, i.e. φ is constant in Ω. However, as $\varphi \in \mathrm{W}_0^{1,4/3}(\Omega)$ it follows that $\varphi = 0$ in Ω and $(\boldsymbol{\sigma},\mathbf{u})$ is a solution of (2.2.9).

Conversely, if $(\boldsymbol{\sigma},\mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ is a solution to (2.2.9), from the Helmholtz descomposition (2.2.10), the velocity $\mathbf{u}$ can descomposed as follows

$$\mathbf{u} = \mathbf{w} + \nabla\varphi \quad \text{in} \quad \Omega,$$

with $\mathbf{w} \in \mathbf{V}_0$ and $\varphi \in \Psi_0$. Therefore, $(\boldsymbol{\sigma}, (\mathbf{w}, \varphi))$ satisfies (2.2.11). Proceeding exactly as before, one can deduce that $\varphi = 0$ in Ω, which implies that $\mathbf{u} = \mathbf{w} \in \mathbf{V}_0$ and thus $(\boldsymbol{\sigma}, (\mathbf{u}, \varphi))$ is a solution to (2.2.11). In this way, we have proved the following Lemma

Lemma 2.2.2 *If $(\boldsymbol{\sigma}, \mathbf{u}) \in \mathbb{X}_0 \times \mathbf{M}$ is a solution to (2.2.9), then $\mathbf{u} \in \mathbf{V}_0$ and $(\boldsymbol{\sigma}, (\mathbf{u}, 0))$ is a solution to (2.2.11). Conversely, if $(\boldsymbol{\sigma}, (\mathbf{u}, \varphi)) \in \mathbb{X}_0 \times (\mathbf{V}_0 \times \Psi_0)$ is a solution to (2.2.11), then $\varphi = 0$ in Ω and $(\boldsymbol{\sigma}, \mathbf{u})$ is a solution to (2.2.9).*

As a consequence of the previous lemma, in what follows we focus on studying the system (2.2.11). Furthermore, from the decomposition (2.2.6), the compatibility condition (2.2.2), observe that both sides of the first equation of (2.2.11) are explicitly cancelled when $\boldsymbol{\tau} \in P_0(\Omega)\,\mathbb{I}$. Therefore, we see that testing the first equation of (2.2.11) with $\boldsymbol{\tau} \in \mathbb{X}$ is equivalent to testing with $\boldsymbol{\tau} \in \mathbb{X}_0$.

In this way, defining the forms $\mathbf{a} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$, $\mathbf{b} : \mathbb{X} \times \mathbf{N} \to \mathbb{R}$, $\mathbf{c} : \mathbf{N} \times \mathbf{N} \times \mathbb{X} \to \mathbb{R}$ and the functionals $F : \mathbb{X} \to \mathbb{R}$ and $G : \mathbf{N} \to \mathbb{R}$, respectively as follows

$$\mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) := (\boldsymbol{\sigma}^{\mathrm{d}}, \boldsymbol{\tau}^{\mathrm{d}})_\Omega, \quad \mathbf{b}(\boldsymbol{\tau}, (\mathbf{v}, \psi)) := (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{v} + \nabla\psi)_\Omega, \tag{2.2.16}$$

$$\mathbf{c}((\mathbf{w}, \varphi); (\mathbf{v}, \chi), \boldsymbol{\tau}) := \frac{1}{\nu}((\mathbf{w} + \nabla\varphi) \otimes (\mathbf{v} + \nabla\chi), \boldsymbol{\tau}^{\mathrm{d}})_\Omega, \tag{2.2.17}$$

$$F(\boldsymbol{\tau}) := \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_D \rangle_\Gamma \quad \text{and} \quad G(\mathbf{v}, \psi) := -\frac{1}{\nu}(\mathbf{f}, \mathbf{v} + \nabla\psi)_\Omega, \tag{2.2.18}$$

where

$$\mathbf{N} := \mathbf{V}_0 \times \Psi_0,$$

we rewrite (2.2.11) equivalently as the variational problem: Find $(\boldsymbol{\sigma}, (\mathbf{u}, \varphi)) \in \mathbb{X}_0 \times \mathbf{N}$, such that:

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + \mathbf{b}(\boldsymbol{\tau}, (\mathbf{u}, \varphi)) + \mathbf{c}((\mathbf{u}, \varphi); (\mathbf{u}, \varphi), \boldsymbol{\tau}) &= F(\boldsymbol{\tau}) \qquad \forall\, \boldsymbol{\tau} \in \mathbb{X}_0, \\ \mathbf{b}(\boldsymbol{\sigma}, (\mathbf{v}, \psi)) &= G(\mathbf{v}, \psi) \quad \forall\, (\mathbf{v}, \psi) \in \mathbf{N}. \end{aligned} \tag{2.2.19}$$

Remark 2.2.3 *Observe that according to (2.2.5) and (2.2.8), the post-processing formula for the pressure p reduces to*

$$p = -\frac{1}{2}\left(\nu\,\mathrm{tr}\,(\boldsymbol{\sigma}) + \mathrm{tr}\,(\mathbf{u} \otimes \mathbf{u}) - \frac{1}{|\Omega|}(\mathrm{tr}\,(\mathbf{u} \otimes \mathbf{u}), 1)_\Omega\right) \quad \textit{in} \quad \Omega.$$

2.3 Analysis of the continuous problem

In this section we prove the well-posedness of problem (2.2.19), to that end we apply the Banach fixed-point theorem to obtain both, existence and uniqueness of solution of (2.2.19).

We first recall from [42, Theorem 1.3.3] the following classical estimate that will be used in the next results:

$$\|w\|_{0,r,\Omega} \leq C_{Sob}\|w\|_{1,\Omega} \quad \forall\, w \in \mathrm{H}^1(\Omega), \text{ for } r \geq 1 \tag{2.3.1}$$

with $C_{Sob} > 0$ depending only on $|\Omega|$.

In turn, from [12, Lemma 3.1], exists $C_d > 0$, such that the following inequality holds

$$C_d\|\boldsymbol{\tau}\|^2_{0,\Omega} \leq \|\boldsymbol{\tau}^{\mathrm{d}}\|^2_{0,\Omega} + \|\mathbf{div}\ \boldsymbol{\tau}\|^2_{0,4/3;\Omega} \quad \forall\, \boldsymbol{\tau} \in \mathbb{X}_0. \tag{2.3.2}$$

which in particular implies that the seminorm

$$|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega} := \left\{\|\boldsymbol{\tau}^{\mathrm{d}}\|^2_{0,\Omega} + \|\mathbf{div}\,\boldsymbol{\tau}\|^2_{0,4/3;\Omega}\right\}^{1/2} \qquad \forall\, \boldsymbol{\tau} \in \mathbb{X}_0, \tag{2.3.3}$$

is a norm in $\mathbb{X}_0$, equivalent to the norm $\|\cdot\|_{4/3,\mathbf{div};\Omega}$. According to this, in what follows we equip the space $\mathbb{X}_0$ with the norm $|\cdot|_{4/3,\mathbf{div};\Omega}$. In turn, we endow the product space $\mathbf{N}$ with the norm

$$\|(\mathbf{v},\psi)\|_{\mathbf{N}} := \|\mathbf{v}\|_{4,\mathrm{div};\Omega} + |\psi|_{1,4;\Omega}.$$

Thus, from the Hölder inequality

$$|(f,g)_\Omega| \leq \|f\|_{0,p;\Omega}\|g\|_{0,q;\Omega} \quad \forall\, f \in \mathrm{L}^p(\Omega), \forall\, g \in \mathrm{L}^q(\Omega), \quad \text{with} \quad \frac{1}{p}+\frac{1}{q} = 1,$$

we can easily deduce that

$$|\mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau})| \leq |\boldsymbol{\sigma}|_{4/3,\mathbf{div};\Omega}|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega} \quad \forall\, \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathbb{X}_0, \tag{2.3.4}$$

$$|\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))| \leq |\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega}\|(\mathbf{v},\psi)\|_{\mathbf{N}} \quad \forall\, \boldsymbol{\tau} \in \mathbb{X}_0, \forall\,(\mathbf{v},\psi) \in \mathbf{N}, \tag{2.3.5}$$

$$|\mathbf{c}((\mathbf{w},\varphi);(\mathbf{v},\psi),\boldsymbol{\tau})| \leq \frac{1}{\nu}\|(\mathbf{w},\varphi)\|_{\mathbf{N}}\|(\mathbf{v},\psi)\|_{\mathbf{N}}|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbb{X}_0, \forall\,(\mathbf{w},\varphi),\ (\mathbf{v},\psi)\in\mathbf{N}, \tag{2.3.6}$$

and

$$|G(\mathbf{v},\psi)| \leq \frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\|(\mathbf{v},\psi)\|_{\mathbf{N}}. \tag{2.3.7}$$

where we use the fact that

$$\|\mathbf{v}+\nabla\psi\|_{0,4;\Omega} \leq \|\mathbf{v}\|_{4,\mathrm{div};\Omega} + |\psi|_{1,4;\Omega} = \|(\mathbf{v},\psi)\|_{\mathbf{N}} \quad \forall\,(\mathbf{v},\psi)\in\mathbf{N}.$$

Moreover, thanks to [12, Lemma 3.5], it follows that

$$|F(\boldsymbol{\tau})| \leq C_F\|\mathbf{u}_D\|_{1/2,\Gamma}|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega}, \tag{2.3.8}$$

where C_F is a positive constant depending on C_{Sob} and C_d.

We now let $\mathbb{V}$ be the kernel of $\mathbf{b}$, that is

$$\mathbb{V} := \{\boldsymbol{\tau}\in\mathbb{X}_0 \,:\, \mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi)) = 0, \quad \forall\,(\mathbf{v},\psi)\in\mathbf{N}\}$$

$$= \{\boldsymbol{\tau}\in\mathbb{X}_0 \,:\, (\mathbf{div}\ \boldsymbol{\tau},\mathbf{v}+\nabla\psi)_\Omega = 0, \quad \forall\,(\mathbf{v},\psi)\in\mathbf{N}\}.$$

It is clear that thank to (2.2.10), $\mathbb{V}$ can be characterized as follows

$$\mathbb{V} = \{\boldsymbol{\tau}\in\mathbb{X}_0 \,:\, \mathbf{div}\ \boldsymbol{\tau} = \mathbf{0} \quad \text{in} \quad \Omega\}.$$

In turn, it follows easily from the definition of the bilinear form $\mathbf{a}$ (cf. (2.2.16)) that

$$\mathbf{a}(\boldsymbol{\tau},\boldsymbol{\tau}) = |\boldsymbol{\tau}|^2_{4/3,\mathbf{div};\Omega} \quad \forall\,\boldsymbol{\tau}\in\mathbb{V}, \tag{2.3.9}$$

thus, $\mathbf{a}$ is elliptic on $\mathbb{V}$.

Next, we establish the inf-sup condition of the bilinear form $\mathbf{b}$ through the following lemma.

Lemma 2.3.1 *There exists $\beta>0$, such that*

$$\sup_{\mathbf{0}\neq\boldsymbol{\tau}\in\mathbb{X}_0} \frac{\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))}{|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega}} \geq \beta\|(\mathbf{v},\psi)\|_{\mathbf{N}} \qquad \forall\,(\mathbf{v},\psi)\ \in\ \mathbf{N}. \tag{2.3.10}$$

Proof. We proceed similarly to the proof of Reference [12, Lemma 3.3]. In fact, given

$(\mathbf{v},\psi)\in\mathbf{N}$, we let $\mathbf{h}(\mathbf{v},\psi):=|\mathbf{v}+\nabla\psi|^2(\mathbf{v}+\nabla\psi)$ and observe that

$$(|\mathbf{h}(\mathbf{v},\psi)|^{4/3},1)_\Omega=(|\mathbf{v}+\nabla\psi|^{8/3}|\mathbf{v}+\nabla\psi|^{4/3},1)_\Omega=(|\mathbf{v}+\nabla\psi|^4,1)_\Omega<+\infty,$$

which implies that $\mathbf{h}(\mathbf{v},\psi)\in\mathbf{L}^{4/3}(\Omega)$. Then, defining

$$\hat{\boldsymbol{\tau}}=-\nabla\mathbf{z}+\frac{1}{2|\Omega|}(\operatorname{div}\mathbf{z},1)_\Omega\mathbb{I}\in\mathbb{L}^2(\Omega),\tag{2.3.11}$$

with $\mathbf{z}\in\mathbf{H}^1_0(\Omega)$ being the unique solution of the variational problem

$$(\nabla\mathbf{z},\nabla\mathbf{w})_\Omega=(\mathbf{h}(\mathbf{v},\psi),\mathbf{w})_\Omega\quad\forall\,\mathbf{w}\in\mathbf{H}^1_0(\Omega),\tag{2.3.12}$$

it follows that

$$\mathbf{div}\;\hat{\boldsymbol{\tau}}=\mathbf{h}(\mathbf{v},\psi)\in\mathbf{L}^{4/3}(\Omega),\quad(\operatorname{tr}(\hat{\boldsymbol{\tau}}),1)_\Omega=0,\tag{2.3.13}$$

and, consequently, $\hat{\boldsymbol{\tau}}\in\mathbb{X}_0$.

On the other hand, from (2.3.12) with $\mathbf{w}=\mathbf{z}$ and the Hölder inequality, we obtain

$$|\mathbf{z}|^2_{1,\Omega}=(\nabla\mathbf{z},\nabla\mathbf{z})_\Omega=(\mathbf{h}(\mathbf{v},\psi),\mathbf{z})_\Omega\leq\|\mathbf{h}(\mathbf{v},\psi)\|_{0,4/3;\Omega}\|\mathbf{z}\|_{0,4;\Omega}$$

which together with (2.3.1) with $r=4$ and the Poincaré inequality (2.2.12) with $p=2$, implies

$$\|\mathbf{z}\|_{1,\Omega}\leq C_{Sob}C^2_\Omega\|\mathbf{h}(\mathbf{v},\psi)\|_{0,4/3;\Omega}=C_{Sob}C^2_\Omega\|\mathbf{v}+\nabla\psi\|^3_{0,4;\Omega}.$$

Based on the latter and in accordance with (2.3.3) and (2.3.11), it readily follows that

$$\begin{aligned}|\hat{\boldsymbol{\tau}}|_{4/3,\mathbf{div};\Omega}\leq\left\{\|\hat{\boldsymbol{\tau}}\|^2_{0,\Omega}+\|\mathbf{div}\,\hat{\boldsymbol{\tau}}\|^2_{0,4/3;\Omega}\right\}^{1/2}&\leq(C+1)^{1/2}\|\mathbf{h}(\mathbf{v},\psi)\|_{0,4/3;\Omega}\\&=\tilde{C}\|\mathbf{v}+\nabla\psi\|^3_{0,4;\Omega},\end{aligned}\tag{2.3.14}$$

with $\tilde{C}=(C+1)^{1/2}$. In addition, we observe that by (2.2.10), and [23, Definition 3.13, Lemma 3.20], there exist a constant $C>0$, such that

$$\|(\mathbf{v},\psi)\|_{\mathbf{N}}=\|\mathbf{v}\|_{4,\operatorname{div};\Omega}+|\psi|_{1,4;\Omega}\leq C\|\mathbf{v}+\nabla\psi\|_{0,4;\Omega}.\tag{2.3.15}$$

In this way, from (2.3.13), (2.3.14), and the identity (2.3.15), we obtain

$$\sup_{\mathbf{0}\neq\boldsymbol{\tau}\in\mathbb{X}_0}\frac{\mathbf{b}(\boldsymbol{\tau},(\mathbf{v},\psi))}{|\boldsymbol{\tau}|_{4/3,\mathbf{div};\Omega}}\geq\frac{\mathbf{b}(\hat{\boldsymbol{\tau}},(\mathbf{v},\psi))}{|\hat{\boldsymbol{\tau}}|_{4/3,\mathbf{div};\Omega}}\geq\tilde{C}^{-1}\frac{\|\mathbf{v}+\nabla\psi\|^4_{0,4;\Omega}}{\|\mathbf{v}+\nabla\psi\|^3_{0,4;\Omega}}=\tilde{C}^{-1}\|\mathbf{v}+\nabla\psi\|_{0,4;\Omega}$$
$$\geq\beta\|(\mathbf{v},\psi)\|_{\mathbf{N}},$$

with $\beta=\tilde{C}^{-1}C^{-1}>0$. □

2.3.1 A fixed point strategy

We begin the solvability analysis of (2.2.19) by introducing the bounded set

$$\mathbf{K}:=\left\{(\mathbf{v},\psi)\in\mathbf{N}\ :\ \|(\mathbf{v},\psi)\|_{\mathbf{N}}\leq\frac{2}{\gamma}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right)\right\},\tag{2.3.16}$$

with γ and C_F being the positive constant defined below in (2.3.21) and (2.3.8), respectively. Then, we define our operator as

$$\mathbf{T}:\mathbf{K}\to\mathbf{K},\quad(\mathbf{w},\phi)\to\mathbf{T}(\mathbf{w},\phi)=(\mathbf{u},\varphi),\tag{2.3.17}$$

where given $(\mathbf{w},\phi)\in\mathbf{K}$, $(\mathbf{u},\varphi)$ is the second component of the solution of the linearized version of problem (2.2.19): Find $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{X}_0\times\mathbf{N}$, such that

$$\begin{aligned}\mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau})+\mathbf{b}(\boldsymbol{\tau},(\mathbf{u},\varphi))+\mathbf{c}((\mathbf{w},\phi);(\mathbf{u},\varphi),\boldsymbol{\tau})&=F(\boldsymbol{\tau})\qquad\forall\,\boldsymbol{\tau}\in\mathbb{X}_0,\\ \mathbf{b}(\boldsymbol{\sigma},(\mathbf{v},\psi))&=G(\mathbf{v},\psi)\quad\forall\,(\mathbf{v},\psi)\in\mathbf{N}.\end{aligned}\tag{2.3.18}$$

Hence, it is not difficult to see that $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{X}_0\times\mathbf{N}$ is a solution of (2.2.19) if and only if $(\mathbf{u},\varphi)$ is a fixed-point of $\mathbf{T}$, that is

$$\mathbf{T}(\mathbf{u},\varphi)=(\mathbf{u},\varphi).$$

In this way, in what follows we focus on proving that $\mathbf{T}$ possesses a unique fixed-point. Before proceeding with the solvability analysis, we prove the well-definiteness of the fixed-point operator, for which, according to the definition of $\mathbf{T}$, it is sufficient to prove that the problem (2.3.18) is well-posed.

Let us now define the bilinear form $\mathbf{A} : (\mathbb{X}_0 \times \mathbf{N}) \times (\mathbb{X}_0 \times \mathbf{N}) \to \mathbb{R}$ given by

$$\mathbf{A}((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi))) := \mathbf{a}(\boldsymbol{\sigma},\boldsymbol{\tau}) + \mathbf{b}(\boldsymbol{\tau},(\mathbf{u},\varphi)) + \mathbf{b}(\boldsymbol{\sigma},(\mathbf{v},\psi)). \tag{2.3.19}$$

Notice that, from now on, for simplicity, the norm for the product space $\mathbb{X} \times \mathbf{N}$, will be denoted by $\|(\cdot,\cdot)\| = |\cdot|_{4/3,\mathbf{div};\Omega} + \|\cdot\|_{\mathbf{N}}$. From (2.3.4) and (2.3.5), it is clear that $\mathbf{A}$ is bounded. Furthermore, based on (2.3.9), (2.3.10), and [18, Proposition 2.36] it is not difficult to see that the following inf-sup condition holds:

$$\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \frac{\mathbf{A}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|} \geq \gamma\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\| \quad \forall\,(\boldsymbol{\rho},(\mathbf{z},\zeta)) \in \mathbb{X}_0 \times \mathbf{N}, \tag{2.3.20}$$

with

$$\gamma := \frac{\beta\,\min\{1,\beta\}}{4(\beta+1)}. \tag{2.3.21}$$

Next, we establish the well-definiteness of $\mathbf{T}$.

Theorem 2.3.2 *Assume that*

$$\frac{4}{\nu\gamma^2}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right) \leq 1. \tag{2.3.22}$$

Then, given $(\mathbf{w},\phi) \in \mathbf{K}$*, there exists a unique* $(\mathbf{u},\varphi) \in \mathbf{K}$ *such that* $\mathbf{T}(\mathbf{w},\phi) = (\mathbf{u},\varphi)$.

Proof. In fact, given $(\mathbf{w},\phi) \in \mathbf{K}$, we begin by defining the bilinear form

$$\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi))) := \mathbf{A}((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi))) + \mathbf{c}((\mathbf{w},\phi);(\mathbf{u},\varphi),\boldsymbol{\tau}), \tag{2.3.23}$$

where $\mathbf{A}$ and $\mathbf{c}$ are the forms defined in (2.3.19) and (2.2.17), respectively.

Thus, the problem (2.3.18) can be equivalently rewritten as: Find $(\boldsymbol{\sigma},(\mathbf{u},\varphi)) \in \mathbb{X}_0 \times \mathbf{N}$, such that

$$\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi))) = F(\boldsymbol{\tau}) + G(\mathbf{v},\psi) \quad \forall\,(\boldsymbol{\tau},(\mathbf{v},\psi)) \in \mathbb{X}_0 \times \mathbf{N}. \tag{2.3.24}$$

Therefore, to prove the well-definiteness of $\mathbf{T}$, in what follows we prove that the above problem is well-posed by means of the Banach–Nečas–Babuška theorem [18, Theorem 2.6].

First, we observe that given $(\boldsymbol{\rho},(\mathbf{z},\zeta))$, $(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi})) \in \mathbb{X}_0 \times \mathbf{N}$ with $(\hat{\mathbf{v}},\hat{\psi}) \neq \mathbf{0}$,

from (2.3.6), it follows that

$$\begin{aligned}\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}&\geq\frac{|\mathbf{A}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi})))|}{\|(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi}))\|}\\&\quad-\frac{|\mathbf{c}((\mathbf{w},\phi);(\mathbf{z},\zeta),\hat{\boldsymbol{\tau}})|}{\|(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi}))\|}\\&\geq\frac{|\mathbf{A}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi})))|}{\|(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi}))\|}\\&\quad-\frac{1}{\nu}\|(\mathbf{w},\phi)\|_{\mathbf{N}}\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|,\end{aligned}$$

which, together with (2.3.20) and given that $(\hat{\boldsymbol{\tau}},(\hat{\mathbf{v}},\hat{\psi}))$ is arbitrary, implies

$$\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\geq\left(\gamma-\frac{1}{\nu}\|(\mathbf{w},\phi)\|_{\mathbf{N}}\right)\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|. \tag{2.3.25}$$

Therefore, from the definition of the set $\mathbf{K}$ (cf. (2.3.16)), and the assumption (2.3.22), we get

$$\frac{1}{\nu}\|(\mathbf{w},\phi)\|_{\mathbf{N}}\leq\frac{2}{\nu\gamma}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right)\leq\frac{\gamma}{2}, \tag{2.3.26}$$

which, combined with (2.3.25), results in

$$\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\rho},(\mathbf{z},\zeta)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\geq\frac{\gamma}{2}\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|\quad\forall\,(\boldsymbol{\rho},(\mathbf{z},\zeta))\in\mathbb{X}_0\times\mathbf{N}. \tag{2.3.27}$$

In addition, we observe that given $(\boldsymbol{\rho},(\mathbf{z},\zeta))\in\mathbb{X}_0\times\mathbf{N}$,

$$\begin{aligned}&\sup_{(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))\\&\qquad\geq\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\\&\qquad=\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))+\mathbf{c}((\mathbf{w},\phi);(\mathbf{v},\psi),\boldsymbol{\rho})}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|},\end{aligned}$$

so that,

$$\begin{aligned}\sup_{(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} & \mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))\\ &\geq \frac{|\mathbf{A}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))+\mathbf{c}((\mathbf{w},\phi);(\mathbf{v},\psi),\boldsymbol{\rho})|}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\\ &\geq \frac{|\mathbf{A}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))|}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}-\frac{|\mathbf{c}((\mathbf{w},\phi);(\mathbf{v},\psi),\boldsymbol{\rho})|}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|},\end{aligned}$$

for all $(\boldsymbol{\tau},(\mathbf{v},\psi))\in(\mathbb{X}_0\times\mathbf{N})\backslash\{\mathbf{0}\}$, which together with (2.3.6), implies

$$\begin{aligned}\sup_{(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} & \mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))\\ &\geq \sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \frac{\mathbf{A}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}-\frac{1}{\nu}\|(\mathbf{w},\phi)\|_{\mathbf{N}}\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|.\end{aligned} \tag{2.3.28}$$

Then, since $\mathbf{A}(\cdot,\cdot)$ is symmetric, and based on (2.3.20) and (2.3.28), we obtain:

$$\sup_{(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))\geq\gamma\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|-\frac{1}{\nu}\|(\mathbf{w},\phi)\|_{\mathbf{N}}\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|,$$

which together with (2.3.26), results in

$$\sup_{(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \mathbf{A}_{\mathbf{w},\phi}((\boldsymbol{\tau},(\mathbf{v},\psi)),(\boldsymbol{\rho},(\mathbf{z},\zeta)))\geq\frac{\gamma}{2}\|(\boldsymbol{\rho},(\mathbf{z},\zeta))\|\quad\forall\,(\boldsymbol{\rho},(\mathbf{z},\zeta))\in\mathbb{X}_0\times\mathbf{N}. \tag{2.3.29}$$

Thus, from (2.3.27) and (2.3.29), we get that $\mathbf{A}_{\mathbf{w},\phi}(\cdot,\cdot)$ satisfies the hypothesis of the Banach–Nečas–Babuška theorem (see eg. [18, Theorem 2.6]), which enables us to deduce the existence of a unique $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{X}_0\times\mathbf{N}$ solution to (2.3.18), or equivalently, the existence of a unique $(\mathbf{u},\varphi)\in\mathbf{N}$ such that $\mathbf{T}(\mathbf{w},\phi)=(\mathbf{u},\varphi)$. Furthermore, from (2.3.27) with $(\boldsymbol{\rho},(\mathbf{z},\zeta))=(\boldsymbol{\sigma},(\mathbf{u},\varphi))$ and (2.3.24), we easily obtain that

$$\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\leq\|(\boldsymbol{\sigma},(\mathbf{u},\varphi))\|\leq\frac{2}{\gamma}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right).$$

Therefore, $(\mathbf{u},\varphi)$ belongs to $\mathbf{K}$, and the proof is concluded. □

2.3.2 Well-posedness of the continuous problem

Having proved the well-posedness of the problem (2.3.18), which ensure that the operator $\mathbf{T}$ is well defined, we now aim to establish the existence of a unique fixed point of the operator $\mathbf{T}$. For this purpose, in what follows we verify the hypothesis of the Banach fixed-point theorem.

Theorem 2.3.3 *Let* $\mathbf{f} \in \mathbf{L}^{4/3}(\Omega)$ *and* $\mathbf{u}_D \in \mathbf{H}^{1/2}(\Gamma)$ *such that*

$$\frac{4}{\nu\gamma^2}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right) < 1. \tag{2.3.30}$$

Then, there exists a unique $(\boldsymbol{\sigma},(\mathbf{u},\varphi)) \in \mathbb{X}_0 \times \mathbf{N}$ *solution to* (2.2.19) *such that*

$$|\boldsymbol{\sigma}|_{4/3,\mathbf{div};\Omega} + \|(\mathbf{u},\varphi)\|_{\mathbf{N}} \leq \frac{2}{\gamma}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right).$$

Proof. In what follows we prove that $\mathbf{T}$ is a contraction mapping. We begin by noting that thanks to Theorem 2.3.2, the assumption (2.3.30) guarantees that $\mathbf{T}$ is well-defined. Now, let $(\mathbf{w}_1,\phi_1),(\mathbf{w}_2,\phi_2),(\mathbf{u}_1,\varphi_1),(\mathbf{u}_2,\varphi_2) \in \mathbf{K}$, such that $(\mathbf{u}_1,\varphi_1) = \mathbf{T}(\mathbf{w}_1,\phi_1)$ and $(\mathbf{u}_2,\varphi_2) = \mathbf{T}(\mathbf{w}_2,\phi_2)$. Then, from the definition of $\mathbf{T}$ (cf. (2.3.17)) and (2.3.24), it follows that there exist unique $\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2 \in \mathbb{X}_0$, such that for all $(\boldsymbol{\tau},(\mathbf{v},\psi)) \in \mathbb{X}_0 \times \mathbf{N}$, there hold

$$\mathbf{A}_{\mathbf{w}_1,\phi_1}((\boldsymbol{\sigma}_1,(\mathbf{u}_1,\varphi_1)),(\boldsymbol{\tau},(\mathbf{v},\psi))) = F(\boldsymbol{\tau}) + G(\mathbf{v},\psi),$$

and

$$\mathbf{A}_{\mathbf{w}_2,\phi_2}((\boldsymbol{\sigma}_2,(\mathbf{u}_2,\varphi_2)),(\boldsymbol{\tau},(\mathbf{v},\psi))) = F(\boldsymbol{\tau}) + G(\mathbf{v},\psi).$$

Then, subtracting both equations, adding and subtracting suitable terms, and recalling the definition of $\mathbf{A}_{\mathbf{w},\phi}$ (cf. (2.3.23)), we easily arrive at

$$\mathbf{A}_{\mathbf{w}_1,\phi_1}((\boldsymbol{\sigma}_1-\boldsymbol{\sigma}_2,(\mathbf{u}_1-\mathbf{u}_2,\varphi_1-\varphi_2)),(\boldsymbol{\tau},(\mathbf{v},\psi))) = -\mathbf{c}((\mathbf{w}_1-\mathbf{w}_2,\phi_1-\phi_2);(\mathbf{u}_2,\varphi_2),\boldsymbol{\tau}),$$

for all $(\boldsymbol{\tau},(\mathbf{v},\psi)) \in \mathbb{X}_0 \times \mathbf{N}$.

So, recalling that $(\mathbf{w}_1,\phi_1) \in \mathbf{K}$, from the above identity, (2.3.27) and (2.3.6), we

get

$$\begin{aligned}
&\frac{\gamma}{2}\|(\mathbf{u}_1-\mathbf{u}_2,\varphi_1-\varphi_2)\|_{\mathbf{N}}\\
&\leq \sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \frac{\mathbf{A}_{\mathbf{w}_1,\phi_1}((\boldsymbol{\sigma}_1-\boldsymbol{\sigma}_2,(\mathbf{u}_1-\mathbf{u}_2,\varphi_1-\varphi_2)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\\
&= \sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}} \frac{-\mathbf{c}((\mathbf{w}_1-\mathbf{w}_2,\phi_1-\phi_2);(\mathbf{u}_2,\varphi_2),\boldsymbol{\tau})}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\\
&\leq \frac{1}{\nu}\|(\mathbf{w}_1-\mathbf{w}_2,\phi_1-\phi_2)\|_{\mathbf{N}}\|(\mathbf{u}_2,\varphi_2)\|_{\mathbf{N}},
\end{aligned}$$

which together with $(\mathbf{u}_2,\varphi_2)\in\mathbf{K}$, implies

$$\|(\mathbf{u}_1-\mathbf{u}_2,\varphi_1-\varphi_2)\|_{\mathbf{N}}\leq\frac{4}{\nu\gamma^2}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right)\|(\mathbf{w}_1-\mathbf{w}_2,\phi_1-\phi_2)\|_{\mathbf{N}}.$$

Hence, and assumption (2.3.30) we conclude that $\mathbf{T}$ is a contraction.

Now, to derive the continuous dependence of the data, let $(\mathbf{u},\varphi)\in\mathbf{K}$ be the unique fixed point of $\mathbf{T}$ and let $\boldsymbol{\sigma}\in\mathbb{X}_0$ be such that $(\boldsymbol{\sigma},(\mathbf{u},\varphi))$ is the unique solution of (2.2.19). According to (2.3.27), with $(\boldsymbol{\rho},(\mathbf{z},\zeta))=(\boldsymbol{\sigma},(\mathbf{u},\varphi))$, (2.3.7) and (2.3.8) we obtain

$$\begin{aligned}
\|(\boldsymbol{\sigma},(\mathbf{u},\varphi))\| &\leq \frac{2}{\gamma}\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{\mathbf{A}_{\mathbf{u},\varphi}((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi)))}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|}\\
&= \frac{2}{\gamma}\sup_{\mathbf{0}\neq(\boldsymbol{\tau},(\mathbf{v},\psi))\in\mathbb{X}_0\times\mathbf{N}}\frac{F(\boldsymbol{\tau})+G(\mathbf{v},\psi)}{\|(\boldsymbol{\tau},(\mathbf{v},\psi))\|},
\end{aligned}$$

consequently,

$$|\boldsymbol{\sigma}|_{4/3,\mathbf{div};\Omega}+\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\leq\frac{2}{\gamma}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right).$$

□

We end this section by providing the converse of the derivation of (2.2.19).

Theorem 2.3.4 *Let $(\boldsymbol{\sigma},(\mathbf{u},\varphi))\in\mathbb{X}_0\times\mathbf{M}$ be the unique solution of (2.2.19). Then, $-\mathbf{div}\,\boldsymbol{\sigma}=\frac{1}{\nu}\mathbf{f}$ in Ω, $\varphi=0$ in Ω, $\mathbf{u}=\mathbf{u}_D$ on Γ and $\nabla\mathbf{u}=\boldsymbol{\sigma}^{\mathrm{d}}+\frac{1}{\nu}(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}$ in Ω, which implies that $\mathbf{u}\in\mathbf{H}^1(\Omega)$.*

Proof. The identity $-\mathbf{div}\,\boldsymbol{\sigma}=\frac{1}{\nu}\mathbf{f}$ in Ω follows from the second equation of (2.2.19) and the Helmholtz decomposition (2.2.10) whereas identity $\varphi=0$ in Ω follows from

Lemma 2.2.2. The rest of the identities follow from the first equation of (2.2.19), considering suitable test functions and integrating by parts backwardly. We omit further details. □

2.4 The Galerkin scheme

In this section, we introduce and analyze the corresponding Galerkin scheme for the mixed formulation (2.2.19). The solvability of this scheme is addressed following analogous tools to those employed throughout Section 2.3. Finally, we derive the corresponding a priori error estimate and rates of convergence of the Galerkin scheme.

2.4.1 Discrete scheme

Let $\mathcal{T}_h$ be a regular family of regular triangulations of the polygonal region $\overline{\Omega}$ by triangles T in $\mathbb{R}^2$ of diameter h_T, such that $\overline{\Omega} = \cup\{T : T \in \mathcal{T}_h\}$ and define $h := \max\{h_T : T \in \mathcal{T}_h\}$. Given an integer $l \geq 0$ and a subset S of $\mathbb{R}^2$, we denote by $P_l(S)$ the space of polynomials of total degree at most l defined on S. Hence, for each $T \in \mathcal{T}_h$, we define the local Raviart–Thomas space of lowest order (see for instance [7]), as:

$$\mathbf{RT}_0(T) := [P_0(T)]^2 \oplus P_0(T)\mathbf{x},$$

where $\mathbf{x} := (x_1, x_2)^t$ is a generic vector of $\mathbb{R}^2$.

In addition, we let $\mathcal{E}_h$ be the set of edges of $\mathcal{T}_h$, whose corresponding diameters are denoted h_e, and define

$$\mathcal{E}_h(\Omega) := \{ e \in \mathcal{E}_h : \quad e \subseteq \Omega \} \quad \text{and} \quad \mathcal{E}_h(\Gamma) := \{ e \in \mathcal{E}_h : \quad e \subseteq \Gamma \}.$$

We also let $[\![\cdot]\!]$ be the usual jump operator across internal faces defined for piecewise continuous functions v, by

$$[\![v]\!] = (v|_{T_+})|_e - (v|_{T_-})|_e \quad \text{with} \quad e = \partial T_+ \cap \partial T_-,$$

where T_+ and T_- are the elements of $\mathcal{T}_h$ having e as a common edge. Then, we

introduce the well-known Crouzeix–Raviart space (see, for instance, [15]):

$$\Psi_h := \Big\{ v_h : \Omega \to \mathbb{R} : v_h|_T \in P_1(T), \forall\, T \in \mathcal{T}_h, \quad (\llbracket v_h \rrbracket, 1)_e = 0, \quad \forall\, e \in \mathcal{E}_h(\Omega) \\ \text{and} \quad (v_h, 1)_e = 0, \quad \forall\, e \in \mathcal{E}_h(\Gamma) \Big\}.$$

In this way, defining the discrete spaces

$$\mathbb{X}_h := \{ \boldsymbol{\tau}_h \in \mathbb{X} \,:\, \mathbf{c}^t \boldsymbol{\tau}_h|_T \in \mathbf{RT}_0(T), \quad \forall\, \mathbf{c} \in \mathbb{R}^2 \quad \forall\, T \in \mathcal{T}_h \}\,,$$
$$\mathbb{X}_{h,0} := \mathbb{X}_h \cap \mathbb{H}_0(\mathbf{div}_{\,4/3}; \Omega),$$

$$\mathbf{V}_h := \{ \mathbf{z}_h \in \mathbf{H}^4(\mathrm{div}; \Omega) \,:\, \mathbf{z}_h|_T \in \mathbf{RT}_0(T), \quad \forall\, T \in \mathcal{T}_h \}\,,$$
$$\mathbf{V}_{h,0} := \mathbf{V}_h \cap \mathbf{H}^4(\mathrm{div}^0; \Omega), \quad \mathbf{N}_h := \mathbf{V}_{h,0} \times \Psi_h,$$

the Galerkin scheme associated with problem (2.2.19) reads: Find $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, such that:

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) \,+\, \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{u}_h, \varphi_h)) \,+\, \mathbf{c}_h((\mathbf{u}_h, \varphi_h); (\mathbf{u}_h, \varphi_h), \boldsymbol{\tau}_h) &= F(\boldsymbol{\tau}_h) \\ \mathbf{b}_h(\boldsymbol{\sigma}_h, (\mathbf{v}_h, \psi_h)) &= G_h(\mathbf{v}_h, \psi_h) \end{aligned} \tag{2.4.1}$$

for all $(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, where the form $\mathbf{a}$ and the functional F are defined in (2.2.16) and (2.2.18), respectively, and the forms $\mathbf{b}_h : \mathbb{X}_0 \times \mathbf{N}(h) \to \mathbb{R}$, $\mathbf{c}_h : \mathbf{N}(h) \times \mathbf{N}(h) \times \mathbb{X}_0 \to \mathbb{R}$ and the functional $G_h : \mathbf{N}(h) \to \mathbb{R}$ are defined as follows

$$\mathbf{b}_h(\boldsymbol{\tau}, (\mathbf{v}, \psi_h)) := (\mathbf{div}\ \boldsymbol{\tau}, \mathbf{v} + \nabla_h \psi_h)_\Omega,$$

$$\mathbf{c}_h((\mathbf{w}, \varphi_h); (\mathbf{v}, \chi_h), \boldsymbol{\tau}) := \frac{1}{\nu}((\mathbf{w} + \nabla_h \varphi_h) \otimes (\mathbf{v} + \nabla_h \chi_h), \boldsymbol{\tau}^{\mathrm{d}})_\Omega,$$

$$G_h(\mathbf{v}, \psi_h) := -\frac{1}{\nu}(\mathbf{f}, \mathbf{v} + \nabla_h \psi_h)_\Omega,$$

where ∇_h is the discrete gradient for discontinuous functions, that is, $\nabla_h \psi_h|_T = \nabla(\psi_h|_T), \forall\, T \in \mathcal{T}_h$ and $\mathbf{N}(h) := \mathbf{V}_0 \times (\Psi_0 + \Psi_h)$. Since Ψ_h is not included in Ψ_0, in the sequel, we equipped to the space Ψ_h with the following seminorm

$$|v_h|_h = \Big(\sum_{T \in \mathcal{T}_h} |v_h|^4_{1,4;T} \Big)^{1/4},$$

for all $v_h \in \Psi_h$ which can be proved to be a norm on Ψ_h (see, for instance, [15]) . In accordance with the above, for all $(\mathbf{v}_h, \psi_h) \in \mathbf{N}_h$, we define $\|(\mathbf{v}_h, \psi_h)\|_{\mathbf{N}_h} := \|\mathbf{v}_h\|_{4,\mathrm{div};\Omega} + |\psi_h|_h$ and as in the continuous case, we will denote by $\|(\cdot,\cdot)\| = |\cdot|_{4/3,\mathbf{div};\Omega} + \|\cdot\|_{\mathbf{N}_h}$ the norm for the product space $\mathbb{X} \times \mathbf{N}_h$. In addition, we can easily deduce by using Hölder inequality at local level that the forms $\mathbf{b}_h$, $\mathbf{c}_h$ and G_h are bounded with the same bounding constants that appear in (2.3.5), (2.3.6) and (2.3.7), respectively. Notice that the gradient operator and the discrete gradient operator coincide in Ψ_0 which implies that $\mathbf{b}_h$, $\mathbf{c}_h$ and G_h coincide with $\mathbf{b}$, $\mathbf{c}$ and G in $\mathbf{N}$, respectively.

Remark 2.4.1 *It is worth noting that, although the analysis in this section was conducted using* $\mathbf{RT}_0$ *elements to approximate the pseudostress, other viable options are also available. For example, another possibility would be to approximate the pseudostress using* $\mathbf{BDM}_1$ *elements. For more details about properties of* $\mathbf{BDM}_1$ *elements see, for instance, [6].*

2.4.2 Analysis of the discrete problem

Now we address the unique solvability of (2.4.1) by adapting the fixed-point strategy developed in Section 2.3. To that end, and analogously to the continuous case, we introduce a fixed-point operator associated with a linearized version of problem (2.4.1) and equivalently prove that this operator possess a unique fixed-point by means of the Banach fixed-point theorem. We begin introducing some preliminaries results.

Preliminary results

We begin by recalling from [3, Theorem 4.1], the following orthogonal decomposition:

$$[P_0(\mathcal{T}_h)]^2 = \mathbf{V}_{h,0} \oplus \nabla_h \Psi_h, \tag{2.4.2}$$

where $P_0(\mathcal{T}_h)$ are the piecewise constant functions in Ω and

$$\nabla_h \Psi_h := \{\mathbf{s}_h|_T \in [P_0(T)]^2 : \exists\, v_h \in \Psi_h \text{ such that } \mathbf{s}_h|_T = \nabla(v_h|_T), \quad \forall\, T \in \mathcal{T}_h\}.$$

Now, to prove the discrete version of (2.3.10) we first define the space

$$\mathbf{Z}_p := \{\tau \in \mathbf{H}(\mathrm{div}_p\,;\Omega) : \tau|_T \in \mathbf{W}^{1,p}(T), \quad \forall\, T \in \mathcal{T}_h\},$$

and let

$$\Pi_h^{RT} : \mathbf{Z}_p \to \mathbf{X}_h := \{\tau \in \mathbf{H}(\mathrm{div};\Omega) : \tau|_T \in \mathbf{RT}_0(T), \quad \forall\, T \in \mathcal{T}_h\},$$

be the Raviart–Thomas interpolator operator, which satisfies the identity holds

$$\mathrm{div}\big(\Pi_h^{RT}(\tau)\big) = \mathcal{P}_h(\mathrm{div}\,\tau), \quad \forall\, \tau \in \mathbf{Z}_p, \tag{2.4.3}$$

where $\mathcal{P}_h$ is the L^2-projection on $P_0(\mathcal{T}_h)$ which satisfies

$$(v - \mathcal{P}_h(v), z_h)_\Omega = 0 \quad \forall\, z_h \in P_0(\mathcal{T}_h),$$

and the following error estimate (see [18, Proposition 1.135, section 1.6.3]): For each $0 \leq m \leq 1$ and for each $w \in \mathrm{W}^{m,r}(\Omega)$, with $1 \leq r \leq \infty$, it holds that

$$\|w - \mathcal{P}_h(w)\|_{0,r;\Omega} \leq C h^m |w|_{m,r;\Omega}. \tag{2.4.4}$$

Moreover, the following lemma establishes the local approximation properties of Π_h^{RT}.

Lemma 2.4.2 *Let $r > 1$. Then, there exists $C_1 > 0$, independent of h, such that for each $\tau \in \mathbf{W}^{1,r}(T)$, and for each $0 \leq m \leq 1$, there holds*

$$|\tau - \Pi_h^{RT}(\tau)|_{m,r;T} \leq C_1 \frac{h_T^2}{\rho_T^{m+1}} |\tau|_{1,r;T}, \tag{2.4.5}$$

where h_T is the diameter of T, and ρ_T is the diameter of the largest circle contained in T. Additionally, there exists $C_2 > 0$, independent of h, such that for each $\tau \in \mathbf{W}^{1,r}(T)$, with $\mathrm{div}\,\tau \in \mathrm{W}^{1,r}(T)$ and for each $0 \leq m \leq 1$,

$$|\mathrm{div}\,\tau - \mathrm{div}(\Pi_h^{RT}(\tau))|_{m,r;T} \leq C_2 \frac{h_T}{\rho_T^m} |\mathrm{div}\,\tau|_{1,r;T}. \tag{2.4.6}$$

Proof. As it was explained in [12, Lemma 4.1], the proof follows by employing the

L^r-version of the Deny–Lions Lemma provided in Reference [18, Lemma B.67], the local estimates given in [18, Lemma 1.101], and proceeding analogously as in [22, Section 3.4.4] it can be proved that for any $r > 1$ the estimates (2.4.5) and (2.4.6) hold. We omit further details. □

Owing to the regularity of the mesh and of the estimates (2.4.5) and (2.4.6), it is not difficult see that the following global estimate holds

$$\|\tau - \Pi_h^{RT}(\tau)\|_{0,\Omega} + \|\mathrm{div}\,\tau - \mathrm{div}(\Pi_h^{RT}(\tau))\|_{0,p;\Omega} \leq ch\{|\tau|_{1,\Omega} + |\mathrm{div}\,\tau|_{1,p;\Omega}\}, \quad (2.4.7)$$

for all $\tau \in \mathbf{H}^1(\Omega)$ with $\mathrm{div}\,\tau \in \mathrm{W}^{1,p}(\Omega)$.

In what follows we will employ a tensor version of Π_h^{RT}, denoted by $\mathbf{\Pi}_h^{RT} : \mathbb{Z}_p \to \mathbb{X}_h$, which is defined row-wise by Π_h^{RT}, and the vector version of $\mathcal{P}_h$, denoted by $\mathbf{P}_h$, defined component-wise by $\mathcal{P}_h$.

Remark 2.4.3 *Observe that from the regularity of the mesh and from (2.4.5) with $m = 0$ and $m = 1$, one can easily obtain, respectively, that*

$$\|\tau - \Pi_h^{RT}(\tau)\|_{0,r;T} \leq C_1 \frac{h_T^2}{\rho_T} |\tau|_{1,r;T} \leq \tilde{C}_1 h_T |\tau|_{1,r;T}$$

and

$$|\tau - \Pi_h^{RT}(\tau)|_{1,r;T} \leq C_2 \frac{h_T^2}{\rho_T^2} |\tau|_{1,r;T} \leq \tilde{C}_2 |\tau|_{1,r;T},$$

which combined with [18, Lemma 1.101] and (2.3.1), yields

$$\|\tau - \Pi_h^{RT}(\tau)\|_{0,T} \leq C h_T^{1-\frac{2(2-r)}{2r}} |\tau|_{1,r;T} \quad \forall \tau \in \mathbf{W}^{1,r}(T),$$

with $r > 1$, in particular for $r = 4/3$, there holds

$$\|\tau - \Pi_h^{RT}(\tau)\|_{0,T} \leq C h_T^{1/2} |\tau|_{1,4/3;T} \quad \forall \tau \in \mathbf{W}^{1,4/3}(T). \quad (2.4.8)$$

The latter will be employed next in the proof of Lemma 2.4.4.

Fixed-point strategy and well-posedness analysis

Analogously to the continuous case, let us introduce the finite-dimensional bounded set

$$\mathbf{K}_h := \left\{ (\mathbf{v}_h, \psi_h) \in \mathbf{N}_h \ : \ \|(\mathbf{v}_h, \psi_h)\|_{\mathbf{N}_h} \leq \frac{2}{\tilde{\gamma}} \left(C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu} \|\mathbf{f}\|_{0,4/3;\Omega} \right) \right\},$$

with $\tilde{\gamma} > 0$ being the constant defined in (2.4.21), and define the discrete version of $\mathbf{T}$ (cf. (2.3.17)):

$$\mathbf{T}_h : \mathbf{K}_h \to \mathbf{K}_h, \quad (\mathbf{w}_h, \phi_h) \to \mathbf{T}_h(\mathbf{w}_h, \phi_h) = (\mathbf{u}_h, \varphi_h),$$

where given $(\mathbf{w}_h, \phi_h) \in \mathbf{K}_h$, $(\mathbf{u}_h, \varphi_h)$ is the second component of $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, solution to

$$\begin{aligned} \mathbf{a}(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) \ + \ \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{u}_h, \varphi_h)) \ + \ \mathbf{c}_h((\mathbf{w}_h, \phi_h); (\mathbf{u}_h, \varphi_h), \boldsymbol{\tau}_h) &= F(\boldsymbol{\tau}_h) \\ \mathbf{b}_h(\boldsymbol{\sigma}_h, (\mathbf{v}_h, \psi_h)) &= G_h(\mathbf{v}_h, \psi_h) \end{aligned} \tag{2.4.9}$$

for all $(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$. Similar to the continuous case, the following equivalence hold:

$$\mathbf{T}_h(\mathbf{u}_h, \varphi_h) = (\mathbf{u}_h, \varphi_h) \Leftrightarrow (\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h \text{ satisfies (2.4.1)}, \tag{2.4.10}$$

from which we deduce that to prove the well-posedness of problem (2.4.1), it is sufficient to prove that $\mathbf{T}_h$ has a unique fixed-point in $\mathbf{K}_h$. To that end, and analogously to the analysis of the continuous problem, we first focus on providing the necessary tools to prove that the operator $\mathbf{T}_h$ is well defined. We begin by observing that according to the decomposition (2.4.2) and the fact that $\mathbf{div}\,(\mathbb{X}_h) \subseteq [P_0(\mathcal{T}_h)]^2$, the kernel of $\mathbf{b}_h$,

$$\mathbb{V}_h := \{ \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} : \ \mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)) = 0, \quad \forall\, (\mathbf{v}_h, \psi_h) \in \mathbf{N}_h \},$$

can be characterized as follows

$$\mathbb{V}_h := \{ \boldsymbol{\tau}_h \in \mathbb{X}_{h,0} : \ \mathbf{div}\, \boldsymbol{\tau}_h = \mathbf{0} \quad \text{in} \quad \Omega \}.$$

Hence, it follows easily from the definition of the bilinear form $\mathbf{a}$ that

$$\mathbf{a}(\boldsymbol{\tau}_h, \boldsymbol{\tau}_h) = |\boldsymbol{\tau}_h|^2_{4/3,\mathbf{div};\Omega} \quad \forall\, \boldsymbol{\tau}_h \in \mathbb{V}_h, \tag{2.4.11}$$

as in (2.3.9).

Now we establish the discrete inf-sup condition of $\mathbf{b}_h$.

Lemma 2.4.4 *There exists $\tilde{\beta} > 0$, independent of h, such that*

$$\sup_{\mathbf{0}\neq\boldsymbol{\tau}_h\in\mathbb{X}_{h,0}} \frac{\mathbf{b}_h(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h))}{|\boldsymbol{\tau}_h|_{4/3,\mathbf{div};\Omega}} \geq \tilde{\beta}\|(\mathbf{v}_h, \psi_h)\|_{\mathbf{N}_h} \qquad \forall\, (\mathbf{v}_h, \psi_h) \,\in\, \mathbf{N}_h. \tag{2.4.12}$$

Proof. We let $B \subseteq \mathbb{R}^2$ be a bounded and open convex domain such that $\overline{\Omega} \subset B$, and given $(\mathbf{v}_h, \psi_h) \in \mathbf{N}_h$, we set

$$\mathbf{g}(\mathbf{v}_h, \psi_h) := \begin{cases} |\mathbf{v}_h + \nabla_h\psi_h|^2(\mathbf{v}_h + \nabla_h\psi_h), & \text{in } \Omega, \\ \mathbf{0}, & \text{in } B\backslash\overline{\Omega}. \end{cases}$$

Since $\mathbf{g}(\mathbf{v}_h, \psi_h) \in \mathbf{L}^{4/3}(\Omega)$, a well-known result on regularity of elliptic problems (see, for instance, [21]) implies that there exists a unique weak solution $\mathbf{z} \in \mathbf{W}^{2,4/3}(B) \cap \mathbf{W}^{1,4/3}_0(B)$ of the boundary value problem

$$-\Delta\mathbf{z} \,=\, \mathbf{g}(\mathbf{v}_h, \psi_h) \quad \text{in} \quad B, \qquad \mathbf{z} \,=\, \mathbf{0} \quad \text{on} \quad \partial B,$$

which satisfies

$$\|\mathbf{z}\|_{2,4/3;\Omega} \leq C\|\mathbf{g}(\mathbf{v}_h, \psi_h)\|_{0,4/3;B} = C\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}. \tag{2.4.13}$$

Therefore, we let $\hat{\boldsymbol{\tau}} = -\nabla\mathbf{z}|_\Omega \in \mathbb{W}^{1,4/3}(\Omega)$, and note from (2.4.13) that

$$\|\hat{\boldsymbol{\tau}}\|_{1,4/3;\Omega} \leq c_0\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}, \tag{2.4.14}$$

which together with the continuous embedding from $\mathrm{W}^{1,4/3}(\Omega)$ into $\mathrm{L}^2(\Omega)$, implies

$$\|\hat{\boldsymbol{\tau}}\|_{0,\Omega} \leq c_1\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}. \tag{2.4.15}$$

Now, we define $\hat{\boldsymbol{\tau}}_h = \boldsymbol{\Pi}^{RT}_h(\hat{\boldsymbol{\tau}}) - \dfrac{1}{2|\Omega|}(\mathrm{tr}\,(\boldsymbol{\Pi}^{RT}_h(\hat{\boldsymbol{\tau}})), 1)_\Omega\mathbb{I} \in \mathbb{X}_{h,0}$ and observe from

(2.4.3) that

$$\mathbf{div}\ \hat{\boldsymbol{\tau}}_h = \mathbf{P}_h(\mathbf{div}\ \hat{\boldsymbol{\tau}}) = \mathbf{P}_h(|\mathbf{v}_h + \nabla_h\psi_h|^2(\mathbf{v}_h + \nabla_h\psi_h)) = |\mathbf{v}_h + \nabla_h\psi_h|^2(\mathbf{v}_h + \nabla_h\psi_h). \tag{2.4.16}$$

In turn, proceeding as in [12, Lemma 4.3], that is, using the triangle inequality and estimates (2.4.8) and (2.4.15), we obtain

$$\begin{aligned}
\|\hat{\boldsymbol{\tau}}_h\|_{0,\Omega} &\leq \left\|\hat{\boldsymbol{\tau}} - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\hat{\boldsymbol{\tau}}),1)_\Omega\mathbb{I} - \hat{\boldsymbol{\tau}}_h\right\|_{0,\Omega} + \left\|\hat{\boldsymbol{\tau}} - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\hat{\boldsymbol{\tau}}),1)_\Omega\mathbb{I}\right\|_{0,\Omega} \\
&= \left\|\hat{\boldsymbol{\tau}} - \boldsymbol{\Pi}_h^{RT}(\hat{\boldsymbol{\tau}}) - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\hat{\boldsymbol{\tau}} - \boldsymbol{\Pi}_h^{RT}(\hat{\boldsymbol{\tau}})),1)_\Omega\mathbb{I}\right\|_{0,\Omega} + \left\|\hat{\boldsymbol{\tau}} - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\hat{\boldsymbol{\tau}}),1)_\Omega\mathbb{I}\right\|_{0,\Omega} \\
&\leq \|\hat{\boldsymbol{\tau}} - \boldsymbol{\Pi}_h^{RT}(\hat{\boldsymbol{\tau}})\|_{0,\Omega} + \|\hat{\boldsymbol{\tau}}\|_{0,\Omega} \\
&\leq c_2\left\{\sum_{T\in\mathcal{T}_h} h_T|\hat{\boldsymbol{\tau}}|^2_{1,4/3;T}\right\}^{1/2} + c_1\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega} \\
&\leq c_2 h^{1/2}|\hat{\boldsymbol{\tau}}|_{1,4/3;\Omega} + c_1\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega},
\end{aligned}$$

which together with (2.4.14), implies

$$\|\hat{\boldsymbol{\tau}}_h\|_{0,\Omega} \leq c_3\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}. \tag{2.4.17}$$

Then, using (2.4.16) and (2.4.17), we obtain easily

$$|\hat{\boldsymbol{\tau}}_h|_{4/3,\mathbf{div};\Omega} \leq \{\|\hat{\boldsymbol{\tau}}_h\|^2_{0,\Omega} + \|\mathbf{div}\,(\hat{\boldsymbol{\tau}}_h)\|^2_{0,4/3;\Omega}\}^{1/2} \leq \hat{c}\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}, \tag{2.4.18}$$

with $\hat{c} > 0$ independent of h.

On the other hand, thanks to (2.4.2), and [23, Definition 3.13, Lemma 3.20], there exist a constant $\hat{C} > 0$, such that

$$\|(\mathbf{v}_h,\psi_h)\|_{\mathbf{N}_h} = \|\mathbf{v}_h\|_{4,\mathrm{div};\Omega} + |\psi_h|_h \leq \hat{C}\|\mathbf{v}_h + \nabla_h\psi_h\|_{0,4;\Omega},$$

for all $(\mathbf{v}_h,\psi_h) \in \mathbf{V}_{h,0}\times\Psi_h$. Therefore, from estimates (2.4.16), (2.4.18) and the previous inequality, it follows that

$$\begin{aligned}
\sup_{\mathbf{0}\neq\boldsymbol{\tau}_h\in\mathbb{X}_{h,0}} \frac{\mathbf{b}_h(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))}{|\boldsymbol{\tau}_h|_{4/3,\mathbf{div};\Omega}} &\geq \frac{\mathbf{b}_h(\hat{\boldsymbol{\tau}}_h,(\mathbf{v}_h,\psi_h))}{|\hat{\boldsymbol{\tau}}_h|_{4/3,\mathbf{div};\Omega}} \geq \frac{1}{\hat{c}}\frac{\|\mathbf{v}_h + \nabla_h\psi_h\|^4_{0,4;\Omega}}{\|\mathbf{v}_h + \nabla_h\psi_h\|^3_{0,4;\Omega}} \\
&= \frac{1}{\hat{c}}\|\mathbf{v}_h + \nabla_h\psi_h\|_{0,4;\Omega} \geq \tilde{\beta}\|(\mathbf{v}_h,\psi_h)\|_{\mathbf{N}_h},
\end{aligned}$$

with $\tilde{\beta} = \hat{c}^{-1}\hat{C}^{-1} > 0$. □

To conclude the derivation of the necessary tools to prove the well-definiteness of the operator $\mathbf{T}_h$, analogously to the continuous case, from (2.4.11), (2.4.12) and [18, Proposition 2.36] it follows that the bilinear form $\mathbf{A}^h : (\mathbb{X}_0 \times \mathbf{N}(h)) \times (\mathbb{X}_0 \times \mathbf{N}(h)) \to \mathbb{R}$ defined as

$$\mathbf{A}^h((\boldsymbol{\sigma}, (\mathbf{u}, \varphi)), (\boldsymbol{\tau}, (\mathbf{v}, \psi))) := \mathbf{a}(\boldsymbol{\sigma}, \boldsymbol{\tau}) + \mathbf{b}_h(\boldsymbol{\tau}, (\mathbf{u}, \varphi)) + \mathbf{b}_h(\boldsymbol{\sigma}, (\mathbf{v}, \psi)), \tag{2.4.19}$$

satisfies the discrete inf-sup condition

$$\sup_{\mathbf{0} \neq (\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h} \frac{\mathbf{A}^h((\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)), (\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h)))}{\|(\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h))\|} \geq \tilde{\gamma} \|(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h))\|, \tag{2.4.20}$$

for all $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, with

$$\tilde{\gamma} := \frac{\tilde{\beta} \min\{1, \tilde{\beta}\}}{4(\tilde{\beta} + 1)}. \tag{2.4.21}$$

We are now in position to establishing the well-definiteness of $\mathbf{T}_h$.

Theorem 2.4.5 *Assume that*

$$\frac{4}{\nu \tilde{\gamma}^2} \left(C_F \|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu} \|\mathbf{f}\|_{0,4/3;\Omega} \right) \leq 1. \tag{2.4.22}$$

Then, given $(\mathbf{w}_h, \phi_h) \in \mathbf{K}_h$, *there exists a unique* $(\mathbf{u}_h, \varphi_h) \in \mathbf{K}_h$ *such that* $\mathbf{T}_h(\mathbf{w}_h, \phi_h) = (\mathbf{u}_h, \varphi_h)$.

Proof. In fact, given $(\mathbf{w}_h, \phi_h) \in \mathbf{K}_h$, we proceed analogously to the proof of Theorem 2.3.2 and use (2.3.6), (2.4.20) and the assumption (2.4.22) to deduce that the bilinear form

$$\begin{aligned} \mathbf{A}^h_{\mathbf{w}_h, \phi_h}((\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)), (\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h))) &:= \mathbf{A}^h((\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)), (\boldsymbol{\tau}_h, (\mathbf{v}_h, \psi_h))) \\ &\quad + \mathbf{c}_h((\mathbf{w}_h, \phi_h); (\mathbf{u}_h, \varphi_h), \boldsymbol{\tau}_h), \end{aligned} \tag{2.4.23}$$

satisfies the discrete inf-sup condition

$$\sup_{\mathbf{0}\neq(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h}\frac{\mathbf{A}^h_{\mathbf{w}_h,\phi_h}((\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)))}{\|(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))\|}\geq\frac{\tilde{\gamma}}{2}\|(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))\|. \tag{2.4.24}$$

Then, because subjectivity and injectivity are equivalent, in finite dimensional linear problems, from (2.4.24), the fact that $\mathbf{c}_h$ is bounded with the same bounding constant that apper in (2.3.6) and the Banach–Nečas–Babuška theorem we obtain that there exists a unique $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h$ which satisfies

$$\mathbf{A}^h_{\mathbf{w}_h,\phi_h}((\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)))=F(\boldsymbol{\tau}_h)+G_h(\mathbf{v}_h,\psi_h),$$

for all $(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h$, or equivalently (2.4.9), with $(\mathbf{u}_h,\varphi_h)\in\mathbf{K}_h$ and the proof is concluded.

□

The following theorem establishes the well-posedness of the Galerkin scheme (2.4.1).

Theorem 2.4.6 *Let* $\mathbf{f}\in\mathrm{L}^{4/3}(\Omega)$ *and* $\mathbf{u}_D\in\mathbf{H}^{1/2}(\Gamma)$ *such that*

$$\frac{4}{\nu\tilde{\gamma}^2}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right)<1. \tag{2.4.25}$$

Then, there exists a unique $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h$ *solution to* (2.4.1). *Furthermore, there holds*

$$|\boldsymbol{\sigma}_h|_{4/3,\mathbf{div};\Omega}+\|(\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h}\leq\frac{2}{\tilde{\gamma}}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma}+\frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right).$$

Proof. Using (2.4.10), (2.4.20) and (2.4.25), the proof still repeats exactly the same steps developed in the proof of Theorem 2.3.3. □

Remark 2.4.7 *Observe that from the second equation of* (2.4.1), *we have that*

$$\big(\mathbf{div}\ \boldsymbol{\sigma}_h+\nu^{-1}\mathbf{f},\mathbf{v}_h+\nabla_h\psi_h\big)_\Omega=0\quad\forall\,(\mathbf{v}_h,\psi_h)\in\mathbf{N}_h,$$

which thanks to (2.4.2) *leads to* $\mathbf{div}\ \boldsymbol{\sigma}_h=-\nu^{-1}\mathbf{P}_h(\mathbf{f})$, *and accordingly, our method preserves exactly the discrete equilibrium equation, so the method is momentum con-*

servative. Moreover, the discrete space $\mathbf{V}_{h,0}$ *becomes*

$$\mathbf{V}_{h,0} = \{\mathbf{v}_h \in \mathbf{V}_h : \operatorname{div} \mathbf{v}_h = 0 \quad in \quad \Omega\},$$

thus the numerical scheme (2.4.1) *produces exactly divergence-free aproximations for the velocity* $\mathbf{u}$.

2.4.3 A priori error estimate and theoretical rate of convergence

In this section, we study the convergence of the Galerkin scheme (2.4.1). More precisely, we first deduce the corresponding a priori error estimate and later on, under an extra regularity assumption of the exact solution, and employing the approximation properties of the discrete spaces introduced in (2.4.1), we derive the theoretical rate of convergence.

The following result establishes the aforementioned a priori error estimate.

Theorem 2.4.8 *Assume that*

$$\frac{4}{\nu\gamma\tilde{\gamma}}\left(C_F\|\mathbf{u}_D\|_{1/2,\Gamma} + \frac{1}{\nu}\|\mathbf{f}\|_{0,4/3;\Omega}\right) \leq \frac{1}{2}, \tag{2.4.26}$$

with γ *and* $\tilde{\gamma}$ *being the constants defined in* (2.3.21) *and* (2.4.21)*, respectively. Let* $(\boldsymbol{\sigma},(\mathbf{u},0)) \in \mathbb{X}_0 \times \mathbf{N}$ *and* $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$ *be the unique solutions of problems* (2.2.19) *and* (2.4.1)*, respectively. Then, there exists a positive constant* C*, independent of* h *and* ν*, such that*

$$|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h|_{4/3,\mathbf{div};\Omega}+\|(\mathbf{u}-\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h} \leq C\left\{\inf_{\boldsymbol{\tau}_h\in\mathbb{X}_{h,0}}|\boldsymbol{\sigma}-\boldsymbol{\tau}_h|_{4/3,\mathbf{div};\Omega}+\inf_{\mathbf{v}_h\in\mathbf{V}_{h,0}}\|\mathbf{u}-\mathbf{v}_h\|_{0,\Omega}\right\} \tag{2.4.27}$$

Proof. In order to simplify the subsequent analysis, we define

$$\mathbf{e}_{\boldsymbol{\sigma}} := \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \quad \mathbf{e}_{\mathbf{u}} := \mathbf{u} - \mathbf{u}_h, \quad \mathbf{e}_{\varphi} := \varphi - \varphi_h$$

and for any $(\boldsymbol{\rho}_h,(\mathbf{z}_h,\zeta_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, we write

$$\mathbf{e}_{\boldsymbol{\sigma}} = \xi_{\boldsymbol{\sigma}} + \chi_{\boldsymbol{\sigma}}, \quad \mathbf{e}_{\mathbf{u}} = \xi_{\mathbf{u}} + \chi_{\mathbf{u}}, \quad \mathbf{e}_{\varphi} = \xi_{\varphi} + \chi_{\varphi}, \tag{2.4.28}$$

where

$$\xi_{\boldsymbol{\sigma}} := \boldsymbol{\sigma} - \boldsymbol{\rho}_h, \quad \chi_{\boldsymbol{\sigma}} := \boldsymbol{\rho}_h - \boldsymbol{\sigma}_h, \quad \xi_{\mathbf{u}} := \mathbf{u} - \mathbf{z}_h, \quad \chi_{\mathbf{u}} := \mathbf{z}_h - \mathbf{u}_h,$$
$$\xi_{\varphi} := \varphi - \zeta_h, \quad \chi_{\varphi} := \zeta_h - \varphi_h.$$

According to the definition of the bilinear form $\mathbf{A}^h$ (see (2.4.19)), from (2.2.19), the fact that $-\mathbf{div}\,\boldsymbol{\sigma} = \dfrac{1}{\nu}\mathbf{f}$ in Ω and $\varphi = 0$ in Ω (see Theorem 2.3.4), and (2.4.1) we have that the following identities are satisfied

$$\mathbf{A}^h((\boldsymbol{\sigma},(\mathbf{u},\varphi)),(\boldsymbol{\tau},(\mathbf{v},\psi))) + \mathbf{c}_h((\mathbf{u},\varphi);(\mathbf{u},\varphi),\boldsymbol{\tau}) = F(\boldsymbol{\tau}) + G_h(\mathbf{v},\psi)$$

and

$$\mathbf{A}^h((\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))) + \mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\mathbf{u}_h,\varphi_h),\boldsymbol{\tau}_h) = F(\boldsymbol{\tau}_h) + G_h(\mathbf{v}_h,\psi_h),$$

for all $(\boldsymbol{\tau},(\mathbf{v},\psi)) \in \mathbb{X}_0 \times \mathbf{N}(h)$ and $(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$; from which follows the Galerkin orthogonality property

$$\mathbf{A}^h((\mathbf{e}_{\boldsymbol{\sigma}},(\mathbf{e}_{\mathbf{u}},\mathbf{e}_{\varphi})),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))) + \mathbf{c}_h((\mathbf{u},\varphi);(\mathbf{u},\varphi),\boldsymbol{\tau}_h) - \mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\mathbf{u}_h,\varphi_h),\boldsymbol{\tau}_h) = 0, \tag{2.4.29}$$

for all $(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$.

Now, using the decompositions (2.4.28), the definition of $\mathbf{A}^h_{\mathbf{w}_h,\phi_h}$ (cf. (2.4.23)), and the fact that

$$\mathbf{c}_h((\mathbf{u},\varphi);(\mathbf{u},\varphi),\boldsymbol{\tau}_h) = \mathbf{c}_h((\mathbf{u}-\mathbf{u}_h,\varphi-\varphi_h);(\mathbf{u},\varphi),\boldsymbol{\tau}_h) + \mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\mathbf{u},\varphi),\boldsymbol{\tau}_h),$$

from (2.4.29) we obtain that for all $(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$, it follows that

$$\mathbf{A}^h_{\mathbf{u}_h,\varphi_h}((\chi_{\boldsymbol{\sigma}},(\chi_{\mathbf{u}},\chi_{\varphi})),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))) = -\mathbf{A}^h((\xi_{\boldsymbol{\sigma}},(\xi_{\mathbf{u}},\xi_{\varphi})),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)))$$
$$- \mathbf{c}_h((\xi_{\mathbf{u}},\xi_{\varphi});(\mathbf{u},\varphi),\boldsymbol{\tau}_h) - \mathbf{c}_h((\chi_{\mathbf{u}},\chi_{\varphi});(\mathbf{u},\varphi),\boldsymbol{\tau}_h) - \mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\xi_{\mathbf{u}},\xi_{\varphi}),\boldsymbol{\tau}_h),$$

which, according to the definition of $\mathbf{A}^h$ (cf. (2.4.19)), implies

$$\begin{aligned}\mathbf{A}^h_{\mathbf{u}_h,\varphi_h}((\chi_{\boldsymbol{\sigma}},(\chi_{\mathbf{u}},\chi_{\varphi})),(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h)))&\\ = -\mathbf{a}(\xi_{\boldsymbol{\sigma}},\boldsymbol{\tau}_h)-\mathbf{b}_h(\boldsymbol{\tau}_h,(\xi_{\mathbf{u}},\xi_{\varphi}))-\mathbf{b}_h(\xi_{\boldsymbol{\sigma}},(\mathbf{v}_h,\psi_h))-\mathbf{c}_h((\xi_{\mathbf{u}},\xi_{\varphi});(\mathbf{u},\varphi),\boldsymbol{\tau}_h)&\\ -\mathbf{c}_h((\chi_{\mathbf{u}},\chi_{\varphi});(\mathbf{u},\varphi),\boldsymbol{\tau}_h)-\mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\xi_{\mathbf{u}},\xi_{\varphi}),\boldsymbol{\tau}_h),&\end{aligned} \tag{2.4.30}$$

for all $(\boldsymbol{\tau}_h,(\mathbf{v}_h,\psi_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h$. Then, because $(\mathbf{u}_h,\varphi_h)\in\mathbf{K}_h$, we apply the discrete inf-sup condition (2.4.20) on the left-hand side of (2.4.30) and the continuity properties of $\mathbf{a}$, $\mathbf{b}_h$ and $\mathbf{c}_h$, on the right-hand side of (2.4.30), to get

$$\begin{aligned}\|(\chi_{\boldsymbol{\sigma}},(\chi_{\mathbf{u}},\chi_{\varphi}))\|\leq&\frac{2}{\tilde{\gamma}}\left(2|\xi_{\boldsymbol{\sigma}}|_{4/3,\mathbf{div};\Omega}+\left(1+\frac{1}{\nu}\|(\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h}+\frac{1}{\nu}\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\right)\|(\xi_{\mathbf{u}},\xi_{\varphi})\|_{\mathbf{N}_h}\right.\\ &\left.+\frac{1}{\nu}\|(\chi_{\mathbf{u}},\chi_{\varphi})\|_{\mathbf{N}_h}\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\right),\end{aligned}$$

i.e.,

$$\begin{aligned}&|\chi_{\boldsymbol{\sigma}}|_{4/3,\mathbf{div};\Omega}+\left(1-\frac{2}{\nu\tilde{\gamma}}\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\right)\|(\chi_{\mathbf{u}},\chi_{\varphi})\|_{\mathbf{N}_h}\\ &\quad\leq\frac{2}{\tilde{\gamma}}\left(2|\xi_{\boldsymbol{\sigma}}|_{4/3,\mathbf{div};\Omega}+\left(1+\frac{1}{\nu}\|(\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h}+\frac{1}{\nu}\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\right)\|(\xi_{\mathbf{u}},\xi_{\varphi})\|_{\mathbf{N}_h}\right).\end{aligned} \tag{2.4.31}$$

Therefore, given that $(\mathbf{u},\varphi)\in\mathbf{K}$ and $(\mathbf{u}_h,\varphi_h)\in\mathbf{K}_h$, from the assumption (2.4.26) and (2.4.31), we obtain that, $\frac{1}{\nu}\|(\mathbf{u},\varphi)\|_{\mathbf{N}}\leq\frac{\tilde{\gamma}}{4}$ and $\frac{1}{\nu}\|(\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h}\leq\frac{\gamma}{4}$, thus

$$|\chi_{\boldsymbol{\sigma}}|_{4/3,\mathbf{div};\Omega}+\|(\chi_{\mathbf{u}},\chi_{\varphi})\|_{\mathbf{N}_h}\leq C(|\xi_{\boldsymbol{\sigma}}|_{4/3,\mathbf{div};\Omega}+\|(\xi_{\mathbf{u}},\xi_{\varphi})\|_{\mathbf{N}_h}), \tag{2.4.32}$$

where $C>0$ is independent of h and ν. Thus, from (2.4.28), (2.4.32) and the triangle inequality, we get

$$\|(\mathbf{e}_{\boldsymbol{\sigma}},(\mathbf{e}_{\mathbf{u}},\varphi_h))\|\leq\|(\chi_{\boldsymbol{\sigma}},(\chi_{\mathbf{u}},\chi_{\varphi}))\|+\|(\xi_{\boldsymbol{\sigma}},(\xi_{\mathbf{u}},\xi_{\varphi}))\|\leq(1+C)\|(\xi_{\boldsymbol{\sigma}},(\xi_{\mathbf{u}},\xi_{\varphi}))\|,$$

which combined with the fact that $(\boldsymbol{\rho}_h,(\mathbf{z}_h,\zeta_h))\in\mathbb{X}_{h,0}\times\mathbf{N}_h$ is arbitrary and the fact that $\varphi=0$ in Ω, concludes the proof.

□

From the estimates (2.4.4), (2.4.7) and the error estimate (2.4.27) we readily obtain the corresponding theorical rate of convergence of our mixed finite element

method. This result is established next.

Theorem 2.4.9 *Let* $(\boldsymbol{\sigma},(\mathbf{u},0)) \in \mathbb{X}_0 \times \mathbf{N}$ *and* $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$ *be the unique solutions of* (2.2.19) *and* (2.4.1)*, respectively, with* $\mathbf{f}$ *and* $\mathbf{u}_D$ *satisfying* (2.4.26)*; and assume that* $\boldsymbol{\sigma} \in \mathbb{H}^1(\Omega), \mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{W}^{1,4/3}(\Omega)$ *and* $\mathbf{u} \in \mathbf{W}^{1,4}(\Omega)$*. Then, there exist* $C>0$*, independent of* h *and* ν*, such that*

$$|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h|_{4/3,\mathbf{div};\Omega} + \|(\mathbf{u}-\mathbf{u}_h,\varphi_h)\|_{\mathbf{N}_h} \leq Ch\{|\boldsymbol{\sigma}|_{1,\Omega} + |\mathbf{div}\,\boldsymbol{\sigma}|_{1,4/3;\Omega} + |\mathbf{u}|_{1,4;\Omega}\}.$$

2.4.4 Computing other variables of interest

In this section, we introduce suitable approximations for other variables of interest, such as the pressure p, the velocity gradient $\mathbf{G} = \nabla\mathbf{u}$, the vorticity $\boldsymbol{\omega} = \frac{1}{2}(\nabla\mathbf{u} - \nabla\mathbf{u}^t)$, and the stress tensor $\tilde{\boldsymbol{\sigma}} := \nu(\nabla\mathbf{u} + \nabla\mathbf{u}^t) - p\mathbb{I}$, are all them written in terms of the solution of the discrete problem (2.4.1). In fact, using the Remark 2.2.3 and simple computations, we deduce that at the continuous level, there hold

$$\begin{aligned}
&p = -\frac{1}{2}\left(\nu\,\mathrm{tr}\,(\boldsymbol{\sigma}) + \mathrm{tr}\,(\mathbf{u}\otimes\mathbf{u}) - \frac{1}{|\Omega|}(\mathrm{tr}\,(\mathbf{u}\otimes\mathbf{u}),1)_\Omega\right),\\
&\mathbf{G} = \boldsymbol{\sigma}^{\mathrm{d}} + \frac{1}{\nu}(\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}}, \quad \boldsymbol{\omega} = \frac{1}{2}(\boldsymbol{\sigma}-\boldsymbol{\sigma}^t),\\
&\text{and} \quad \tilde{\boldsymbol{\sigma}} = \nu(\boldsymbol{\sigma}^{\mathrm{d}} + \boldsymbol{\sigma}^t) - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\mathbf{u}\otimes\mathbf{u}),1)_\Omega\mathbb{I} + \mathbf{u}\otimes\mathbf{u} + (\mathbf{u}\otimes\mathbf{u})^{\mathrm{d}},
\end{aligned}$$

provided the discrete solution $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h))$ of problem (2.4.1), we propose the following approximations for the mentioned variables:

$$\begin{aligned}
&p_h = -\frac{1}{2}\left(\nu\,\mathrm{tr}\,(\boldsymbol{\sigma}_h) + \mathrm{tr}\,(\mathbf{u}_h\otimes\mathbf{u}_h) - \frac{1}{|\Omega|}(\mathrm{tr}\,(\mathbf{u}_h\otimes\mathbf{u}_h),1)_\Omega\right),\\
&\mathbf{G}_h = \boldsymbol{\sigma}_h^{\mathrm{d}} + \frac{1}{\nu}(\mathbf{u}_h\otimes\mathbf{u}_h)^{\mathrm{d}}, \quad \boldsymbol{\omega}_h = \frac{1}{2}(\boldsymbol{\sigma}_h-\boldsymbol{\sigma}_h^t), \quad \text{and}\\
&\tilde{\boldsymbol{\sigma}}_h = \nu(\boldsymbol{\sigma}_h^{\mathrm{d}} + \boldsymbol{\sigma}_h^t) - \frac{1}{2|\Omega|}(\mathrm{tr}\,(\mathbf{u}_h\otimes\mathbf{u}_h),1)_\Omega\mathbb{I} + \mathbf{u}_h\otimes\mathbf{u}_h + (\mathbf{u}_h\otimes\mathbf{u}_h)^{\mathrm{d}}.
\end{aligned} \tag{2.4.33}$$

The following result, whose proof follows directly from Theorem 2.4.9 and suitable algebraic manipulations, establishes the corresponding approximation result for this post-processing procedure.

Corollary 2.4.10 *Let* $(\boldsymbol{\sigma},(\mathbf{u},0)) \in \mathbb{X}_0 \times \mathbf{N}$ *and* $(\boldsymbol{\sigma}_h,(\mathbf{u}_h,\varphi_h)) \in \mathbb{X}_{h,0} \times \mathbf{N}_h$ *be the unique solutions of* (2.2.19) *and* (2.4.1)*, respectively, with* $\mathbf{f}$ *and* $\mathbf{u}_D$ *satisfying* (2.4.26)*; and assume that* $\boldsymbol{\sigma} \in \mathbb{H}^1(\Omega)$, $\mathbf{div}\,\boldsymbol{\sigma} \in \mathbf{W}^{1,4/3}(\Omega)$ *and* $\mathbf{u} \in \mathbf{W}^{1,4}(\Omega)$*. Finally, assume that the hypotheses of the Theorem 2.4.9 be hold. Then, there exists* $C>0$*, independient of* h *and* ν*, such that*

$$\|p-p_h\|_{0,\Omega} + \|\tilde{\boldsymbol{\sigma}}-\tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega} + \|\mathbf{G}-\mathbf{G}_h\|_{0,\Omega} + \|\boldsymbol{\omega}-\boldsymbol{\omega}_h\|_{0,\Omega} \\ \leq Ch\Big\{|\boldsymbol{\sigma}|_{1,\Omega} + |\mathbf{div}\,\boldsymbol{\sigma}|_{1,4/3;\Omega} + |\mathbf{u}|_{1,4;\Omega}\Big\}.$$

2.5 Numerical Results

In this section we present three numerical examples illustrating the performance of our finite element scheme and confirm the theorical results. We begin by mentioning that the numerical results that follow are realised by imposing the condition of $(\mathrm{tr}\,(\boldsymbol{\sigma}_h),1)_\Omega = 0$ through a penalty strategy using a scalar Lagrange multiplier (adding only one row and one column to the system). Also, the divergence-free constraint for the velocity is imposed by means of an appropriate Lagrange multiplier $r_h \in P_0(\mathcal{T}_h)$. More precisely, we replace the numerical scheme (2.4.1) by the system: Find $(\boldsymbol{\sigma}_h,\mathbf{u}_h,\varphi_h,r_h,\lambda_h) \in \mathbb{X}_h \times \mathbf{V}_h \times \Psi_h \times P_0(\mathcal{T}_h) \times \mathbb{R}$, such that:

$$\begin{aligned}
\mathbf{a}(\boldsymbol{\sigma}_h,\boldsymbol{\tau}_h) + \mathbf{b}_h(\boldsymbol{\tau}_h,(\mathbf{u}_h,\varphi_h)) + \mathbf{c}_h((\mathbf{u}_h,\varphi_h);(\mathbf{u}_h,\varphi_h),\boldsymbol{\tau}_h) + \lambda_h(\mathrm{tr}\,(\boldsymbol{\tau}_h),1)_\Omega &= F(\boldsymbol{\tau}_h),\\
\mathbf{b}_h(\boldsymbol{\sigma}_h,(\mathbf{v}_h,\psi_h)) + (r_h,\mathrm{div}\,\mathbf{v}_h)_\Omega &= G(\mathbf{v}_h,\psi_h),\\
(s_h,\mathrm{div}\,\mathbf{u}_h)_\Omega &= 0,\\
\eta_h(\,\mathrm{tr}\,(\boldsymbol{\sigma}_h),1)_\Omega &= 0,
\end{aligned}$$

for all $(\boldsymbol{\tau}_h,\mathbf{v}_h,\psi_h,s_h,\eta_h) \in \mathbb{X}_h \times \mathbf{V}_h \times \Psi_h \times P_0(\mathcal{T}_h) \times \mathbb{R}$. Our implementation is based on *Freefem++* code (see [32]), in conjunction with the direct linear solver UMFPACK (see [16]). We apply a Newton's method with a fixed tolerance $tol = 1E-06$ and the iterations are terminated once the relative error of the entire coefficient vectors between two consecutive iterates is sufficiently small, i.e.,

$$\frac{|\mathrm{coeff}^{m+1} - \mathrm{coeff}^{m}|}{|\mathrm{coeff}^{m+1}|} \leq tol,$$

where $|\cdot|$ is the standard euclidean norm $\mathbb{R}^N$, with N stands for the total number of degree of freedom defining $\mathbb{X}_h \times \mathbf{V}_h \times \Psi_h \times P_0(\mathcal{T}_h) \times \mathbb{R}$. For each example shown

below, we simply take $(\boldsymbol{\sigma}_h, (\mathbf{u}_h, \varphi_h)) = (\mathbf{0}, (\mathbf{0}, 0))$ as initial guess. We denote the individual errors by

$$\mathsf{e}(\boldsymbol{\sigma}) := |\boldsymbol{\sigma} - \boldsymbol{\sigma}_h|_{4/3,\mathbf{div};\Omega}\,, \quad \mathsf{e}(\mathbf{u}) := \|\mathbf{u} - \mathbf{u}_h\|_{0,4;\Omega}\,, \quad \mathsf{e}(\varphi) := |\varphi - \varphi_h|_h\,,$$

$$\mathsf{e}(p) := \|p - p_h\|_{0,\Omega}\,, \quad \mathsf{e}(\boldsymbol{\omega}) := \|\boldsymbol{\omega} - \boldsymbol{\omega}_h\|_{0,\Omega}\,, \quad \mathsf{e}(\nabla\mathbf{u}) := \|\nabla\mathbf{u} - \mathbf{G}_h\|_{0,\Omega}\,,$$

$$\mathsf{e}(\tilde{\boldsymbol{\sigma}}_h) := \|\tilde{\boldsymbol{\sigma}} - \tilde{\boldsymbol{\sigma}}_h\|_{0,\Omega}\,,$$

where p_h, $\boldsymbol{\omega}_h$, $\mathbf{G}_h$, and $\tilde{\boldsymbol{\sigma}}_h$ are approximated, respectively, through the post-processing formulas (2.4.33).

Moreover, we let $r(\%)$ be the experimental rate of convergence given by

$$r(\%) := \frac{\log(\mathsf{e}(\%)/\mathsf{e}'(\%))}{\log(h/h')}$$

where $\mathsf{e}(\%)$ is any of the errors defined above and h and h' are two consecutive meshsizes with errors e and e'.

The first example focuses on the performance of the iterative method as a function of the viscosity ν, considering the analytical solution $(\mathbf{u}, p)$ obtained by Kovasznay in [36]. For the domain $\Omega := (-1/2, 3/2) \times (0, 2)$ and for a given ν, this solution is given by

$$\mathbf{u}(x_1, x_2) = \begin{pmatrix} 1 - e^{\lambda x_1}\cos(2\pi x_2) \\ \\ \dfrac{\lambda}{2\pi} e^{\lambda x_1}\sin(2\pi x_2) \end{pmatrix}, \quad p(x_1, x_2) = -\frac{1}{2}e^{2\lambda x_1} + \overline{p},$$

where

$$\lambda := \frac{-8\pi^2}{\nu^{-1} + \sqrt{\nu^{-2} + 16\pi^2}},$$

and the constant $\overline{p}$ is such that $(p, 1)_\Omega = 0$. Note that in this case $\mathbf{div}\,\boldsymbol{\sigma} = \mathbf{f} = \mathbf{0}$ in Ω.

In Table 2.1, we show the behavior of Newton's method as a function of the viscosity number, considering different mesh sizes and the finite element spaces introduced in Section 2.4. Here we observe that the lower the parameter ν the higher the number of iterations. Blanks spaces means that the iterative method takes more than 100 iterations. Next, in Table 2.2, we summarize the convergence history for

a sequence of quasi-uniform triangulations, considering the viscosity $\nu = 1$. We see there that the rate of convergence provided by Theorem 2.4.9 and Corollary 2.4.10 is attained by the unknowns and all the post-processed variables. In addition, the l^∞–norm of $\operatorname{div} \mathbf{u}_h$ and of $\mathbf{div}\ \boldsymbol{\sigma}_h$ in each mesh are close to 0 which shows that this method conserves mass and momentum.

NUMBER OF ITERATIONS

ν	$h = 0.7454$	$h = 0.3802$	$h = 0.1989$	$h = 0.0968$	$h = 0.0489$
1	9	9	7	5	4
0.1	–	–	14	6	5

ν	$h = 0.7454$	$h = 0.3802$	$h = 0.1989$	$h = 0.0968$	$h = 0.0489$
1	9	7	6	5	4
0.1	17	10	7	6	5
0.01	–	–	–	7	6

Table 2.1: EXAMPLE 1: Convergence behavior of the Newton's method with respect to the parameter ν and for the mixed $\mathbf{RT}_0 - \mathbf{RT}_0 - \Psi_h$ (top) and $\mathbf{BDM}_1 - \mathbf{BDM}_1 - \Psi_h$ (bottom) approximations of the Navier–Stokes problem.

N	h	$\mathsf{e}(\boldsymbol{\sigma})$	$r(\boldsymbol{\sigma})$	$\mathsf{e}(\mathbf{u})$	$r(\mathbf{u})$	$\mathsf{e}(p)$	$r(p)$	$\|\operatorname{div} \mathbf{u}_h\|_{l^\infty}$
285	0.7454	9.53E+01	–	8.80E+00	–	5.45E+01	–	1.78E-15
1129	0.3802	7.65E+01	0.3261	5.48E+00	0.7040	2.99E+01	0.8935	1.42E-14
4385	0.1989	4.58E+01	0.7927	3.08E+00	0.8909	1.74E+01	0.8381	5.68E-14
17309	0.0968	2.32E+01	0.9465	1.56E+00	0.9428	8.18E+00	1.0459	1.42E-13
68763	0.0489	1.15E+01	1.0237	7.81E-01	1.0113	3.98E+00	1.0529	3.41E-13

$\|\mathbf{div}\ \boldsymbol{\sigma}_h\|_{l^\infty}$	$\mathsf{e}(\boldsymbol{\omega})$	$r(\boldsymbol{\omega})$	$\mathsf{e}(\nabla\mathbf{u})$	$r(\nabla\mathbf{u})$	$\mathsf{e}(\tilde{\boldsymbol{\sigma}})$	$r(\tilde{\boldsymbol{\sigma}})$	$\|\varphi_h\|_{l^\infty}$	itt
5.68E-14	2.55E+01	–	5.29E+01	–	1.259E+02	–	5.71E-01	9
3.41E-13	3.63E+01	-0.5272	4.78E+01	0.1504	9.11E+01	0.4812	2.83E-01	9
9.09E-13	2.36E+01	0.6655	3.26E+01	0.5902	6.12E+01	0.6135	1.15E-01	7
2.73E-12	1.48E+01	0.6461	1.87E+01	0.7737	3.30E+01	0.8564	3.09E-02	5
1.46E-11	7.62E+00	0.9724	9.61E+00	0.9730	1.69E+01	0.9827	1.11E-02	4

Table 2.2: EXAMPLE 1: Degrees of freedom, meshsizes, errors, rates of convergence, L^∞–norm of $\operatorname{div} \mathbf{u}_h$, $\mathbf{div}\ \boldsymbol{\sigma}_h$ and φ_h for the for the mixed $\mathbf{RT}_0 - \mathbf{RT}_0 - \Psi_h$ approximation of the Navier–Stokes problem considering $\nu = 1.0$.

In our second example, we study a regularized driven cavity problem. The domain is given by the unit square $\Omega = (0, 1)^2$ and we consider a structured mesh with

meshsize $h = 0.0258$. The data are given as in Reference [43, Example D.4], that is, a null body force $\mathbf{f} = \mathbf{0}$ and the prescribed velocity boundary

$$\mathbf{u}(x, 1) = \begin{pmatrix} u_1(x) \\ 0 \end{pmatrix}$$

where

$$u_1(x) = \begin{cases} 1 - \dfrac{1}{4}\left(1 - \cos\left(\dfrac{x_1 - x}{x_1}\pi\right)\right)^2 & \text{for } x \in [0, x_1], \\ 1 & \text{for } x \in (x_1, 1 - x_1), \\ 1 - \dfrac{1}{4}\left(1 - \cos\left(\dfrac{x - (1 - x_1)}{x_1}\pi\right)\right)^2 & \text{for } x \in [1 - x_1, 1]. \end{cases}$$

In addition, all simulations in this example were carried out with $x_1 = 0.1$. In Figure 2.1, we show the velocity streamlines for Reynolds number $Re = 1$ (left) and $Re = 200$ (right), with $Re = 1/\nu$. It can be seen that, as expected, and similar to the results obtained in [43, Example D.4], in the case of low Reynolds number $Re = 1$, there is a large vortex whose centre is close to the upper boundary. On the other hand, for increasing Reynolds numbers, the main vortex moves towards the centre of the cavity. There are smaller counter-rotating vortices in both lower corners.

Finally, in our third example, we study an extension to a steady-state flow around a cylinder. This example focuses on the performance of our numerical scheme, even though the domain considered does not satisfy the connected boundary hypothesis which corresponds to a fundamental assumption for the development of the theory in this work. In fact, let Ω be a bounded polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$ as shown in Figure 2.2 and assume that the boundary is decomposed into an inlet Γ_{in}, an outlet Γ_{out}, the cylinder surface Γ_{cyl} and the upper and lower walls of the channel, denoted by Γ_1 and Γ_3, i.e., $\partial\Omega = \overline{\Gamma}_{\text{in}} \cup \overline{\Gamma}_{\text{out}} \cup \overline{\Gamma}_{\text{cyl}} \cup \overline{\Gamma}_1 \cup \overline{\Gamma}_3$. In addition, let us consider the incompressible steady-state Navier–Stokes equations with the

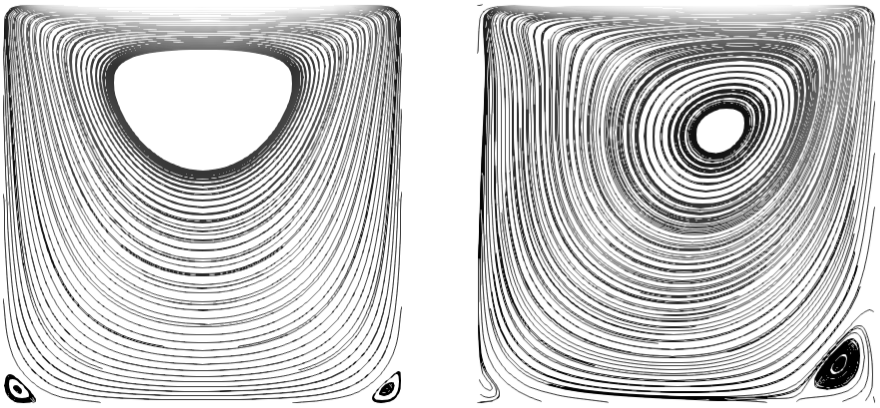

Figure 2.1: EXAMPLE 2: Velocity streamlines for $\nu = 1$ (left) and $\nu = 0.005$ (right) with $h = 0.0258$.

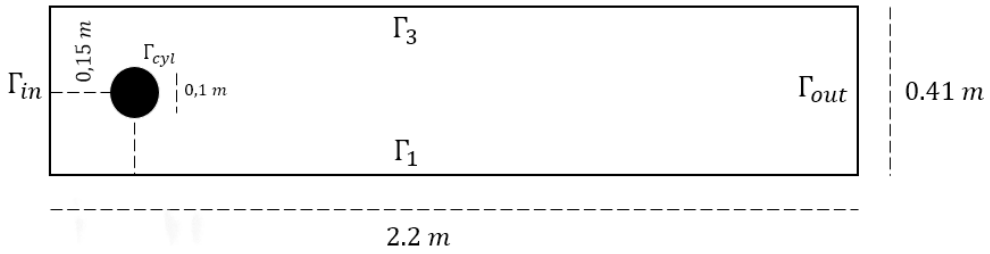

Figure 2.2: EXAMPLE 3: Sketch of a laminar flow around a cylinder.

boundary conditions specified below

$$
\begin{aligned}
-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in} \quad \Omega,\\
\operatorname{div}\mathbf{u} &= 0 \quad \text{in} \quad \Omega,\\
\mathbf{u} &= \mathbf{0} \quad \text{on} \quad \Gamma_1 \cup \Gamma_3 \cup \Gamma_{cyl},\\
\mathbf{u}(0,y) = 0.41^{-2}\begin{pmatrix} 1.2y(0.41-y) \\ 0 \end{pmatrix} &\quad \text{on} \quad \Gamma_{in},\\
\mathbf{u}(2.2,y) = 0.41^{-2}\begin{pmatrix} 1.2y(0.41-y) \\ 0 \end{pmatrix} &\quad \text{on} \quad \Gamma_{out},\\
(p,1)_\Omega &= 0.
\end{aligned}
$$

Proceeding in a completely similar way as in Section 2.2, we obtain the variational problem (2.2.19), with the functional $F : \mathbb{X} \to \mathbb{R}$ defined as follows

$$F(\boldsymbol{\tau}) := \langle \boldsymbol{\tau}\mathbf{n}, \mathbf{u}_{D,io} \rangle_{\Gamma},$$

where

$$\mathbf{u}_{D,io} = \begin{cases} \mathbf{0} & \text{on} \quad \Gamma_1 \cup \Gamma_3, \\ 0.41^{-2} \begin{pmatrix} 1.2y(0.41-y) \\ 0 \end{pmatrix} & \text{on} \quad \Gamma_{in} \cup \Gamma_{out}. \end{cases}$$

Note that the functional F is well-defined and bounded. For further details see, for instance, [4]. Thus, the discrete version of (2.2.19) corresponds to (2.4.1) with the functional F defined above.

In this example, we consider a flow in a two-dimensional domain with a two-dimensional cylinder (circle), whose diameter is $d = 0.1\,m$ (see Figure 2.2 for a schematic representation of this domain). The viscosity of the fluid is given by $\nu = 0.001$, its density by $\rho = 1\,kg/m^3$, and the mean inflow velocity is $U_{mean} = 0.2\,m/s$. Based on the above values, the Reynolds number of the flow is $Re = (dU_{mean})/\nu = 20$. There are no external forces acting on the flow, i.e., $\mathbf{f} = \mathbf{0}$; no slip boundary conditions are prescribed at the top and bottom of the channel and at the surface body of the cylinder. In addition, the parabolic inflow and outflow are defined by

$$\mathbf{u}(0, y) = \mathbf{u}(2.2, y) = 0.41^{-2} \begin{pmatrix} 1.2y(0.41-y) \\ 0 \end{pmatrix}, \quad 0 \leq y \leq 0.41.$$

Figures 2.3 and 2.4 demostrate that the configuration described above admits a steady-state solution. The velocity field figure reveals, as expected for a flow with these characteristics, the formation of two symmetrical vortices behind the cylinder. These vortices, associated with low-pressure regions, result directly from the flow separation as it interacts with the cylinder's surface, leading to a recirculation region at the rear. The symmetry of the vortices is typical of the laminar flow regime at these Reynolds numbers, where viscous forces dominate the flow dynamics.

Additionally, the velocity vectors illustrate how the fluid accelerates around the cylinder, reaching maximum speed at the upper and lower regions, while decelerating

significantly in the wake. This pattern reflects the interaction between the fluid and the cylinder: the flow slows at the front due to the obstruction, then accelerates at the sides.

In the pressure field figure, a high-pressure zone appears at the front of the cylinder, where the flow directly impinges, while a pressure drop is observed in the rear, indicating a low-pressure region behind the cylinder. This behavior is characteristic of stationary laminar flows and is directly linked to the velocity distribution and the forces acting on the body.

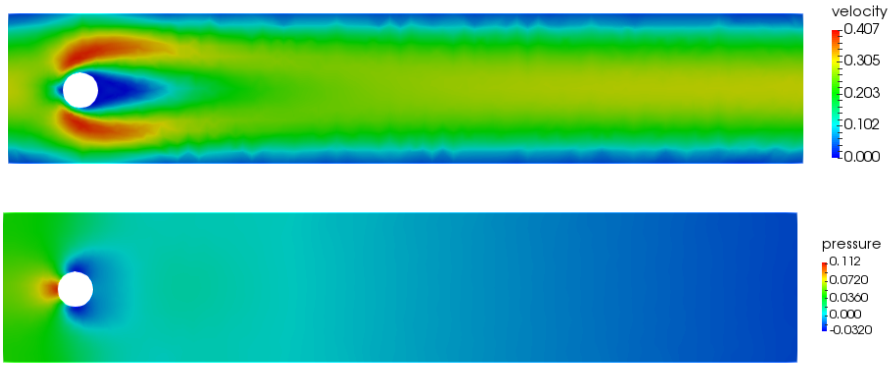

Figure 2.3: EXAMPLE 3: Absolute value of the velocity (top) and pressure (bottom) in the flow around a cylinder.

Also, in this example, we are interested in calculating three quantities: the pressure difference between the front and the back of the cylinder, that is, $\Delta p = p(0.15, 0.2) - p(0.25, 0.2)$ and the forces that are exerted on the body, that is, the drag and lift coefficients given respectively by

$$c_{\text{drag}} = -500\big((\nu\nabla\mathbf{u}, \nabla\mathbf{v}_d)_\Omega + (\mathbf{div}\,(\mathbf{u}\otimes\mathbf{u}), \mathbf{v}_d)_\Omega - (p, \text{div}(\mathbf{v}_d))_\Omega\big),$$

and

$$c_{\text{lift}} = -500\big((\nu\nabla\mathbf{u}, \nabla\mathbf{v}_l)_\Omega + (\mathbf{div}\,(\mathbf{u}\otimes\mathbf{u}), \mathbf{v}_l)_\Omega - (p, \text{div}(\mathbf{v}_l))_\Omega\big),$$

respectively, where $\mathbf{v}_d$ and $\mathbf{v}_l$ are arbitrary functions belonging to $\mathbf{H}^1(\Omega)$ such that $(\mathbf{v}_d)|_{\Gamma_{\text{cyl}}} = (1,0)^t$; $(\mathbf{v}_l)|_{\Gamma_{\text{cyl}}} = (0,1)^t$, and disappear on all other boundaries. Note that, thanks to the definition of $\boldsymbol{\sigma}$ (cf. (2.2.3)) we can easily obtain the drag and lift

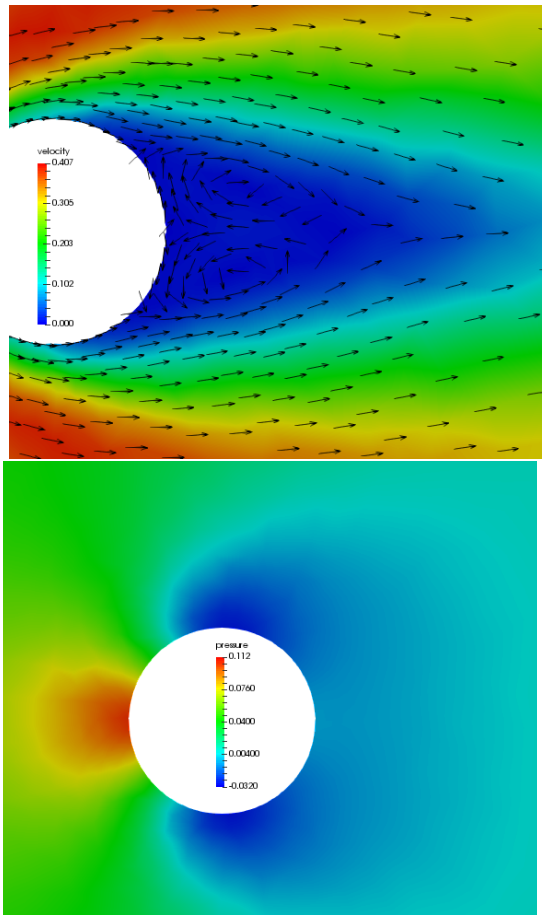

Figure 2.4: EXAMPLE 3: Velocity and pressure at the cylinder.

coefficients using the following expressions:

$$c_{\text{drag}} = -500\,\nu\,(\boldsymbol{\sigma}, \nabla \mathbf{v}_d)_\Omega \quad \text{and} \quad c_{\text{lift}} = -500\,\nu\,(\boldsymbol{\sigma}, \nabla \mathbf{v}_l)_\Omega,$$

respectively.

In this way, we get

$$\Delta p = 0.111109, \quad c_{\text{drag}} = 5.49542 \quad \text{and} \quad c_{\text{lift}} = 0.00441961.$$

These results confirm that the laminar regime is dominated by viscous forces and the stability of the flow around the cylinder. The low lift indicates that there are no relevant instabilities. However, the considerable drag suggests that the cylinder

generates significant drag, typical of laminar flows in this configuration.

Conclusions and future works

Conclusions

In this thesis we proposed, developed and analyzed two numerical schemes that, at the discrete level, are able to conserve mass and momentum for the Stokes and Navier–Stokes problems. The main conclusions of this work are:

1. We analyze a pseudostress-based mixed finite element method for the Stokes problem that ensures both mass and momentum conservation. Mass conservation is achieved by approximating the velocity using the lowest-order Raviart–Thomas elements, while momentum conservation is enforced through a discrete Helmholtz decomposition of the piecewise-constant vector space. We establish the well-posedness of the method and derive theoretical convergence rates, including a superconvergence result for the velocity gradient approximation. A key advantage of the proposed method is its computational efficiency, as it is slightly less expensive than the classical pseudostress-based approach studied in [11, 26], while also guaranteeing mass and momentum conservation. Additionally, we extend our analysis to the Stokes problem with mixed boundary conditions and present numerical experiments that confirm the theoretical results.

2. We extend the previous analysis to the stationary Navier–Stokes problem. To this end, we introduced a mixed formulation and subsequently decomposed the velocity using a Helmholtz decomposition. We then derived a three-field mixed variational formulation, where the pseudo-stress, velocity, and an additional unknown representing the null function are the main unknowns of the system. Additionally, other variables of interest, such as pressure, velocity

gradient, vorticity, and stress tensor, are recovered through a simple post-processing step. We developed the analysis of the continuous problem, where we employed Banach's fixed-point theorem and the Banach–Nečas–Babuška theorem to prove the unique solvability. We then introduced and analyzed the discrete problem, replicating the theory developed for the continuous problem. We established the *a priori* error analysis for the proposed mixed finite element method, obtaining stability estimates and an optimal order of convergence to the solution. We developed *FreeFEM++* codes to validate the theoretical results. Three numerical examples are reported to illustrate the performance of the mixed finite element method and the theoretical convergence rate.

As a consequence of the work developed in this thesis we have the following preprint under review:

- J. Camaño, R. Oyarzúa, K. Rojo. *A momentum and mass conservative pseudostress-based mixed finite element method for the Stokes problem.* Preprint 2025-06, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción, (2025).

Future works

The method developed and the results obtained in this thesis have motivated several future projects. Some of them are described below:

1. **Extension of the analysis of the Navier–Stokes problem to three dimensions**.

   As a complement to the analysis developed in this thesis, we are interested in extending our work to the three dimensional case.

2. **A posteriori error analysis of momentum and mass conservative finite element methods for fluid flow problems**.

   In order to improve the robustness of the error in problems involving complex geometries or solutions with high gradients, we are interested in conducting an *a posteriori* error analysis for the problems studied in this thesis.

3. **New numerical schemes for coupled problems (Boussinesq, Stokes–Darcy, Navier–Stokes–Darcy)**.

   Finally, we aim to use the theory developed in this thesis to design new conservative numerical schemes for coupled problems.

Conclusiones y trabajos futuros

Conclusiones

En esta tesis propusimos, desarrollamos y analizamos dos esquemas numéricos que, a nivel discreto, son capaces de conservar masa y momentum para los problemas de Stokes y Navier–Stokes. Las conclusiones principales de este trabajo son:

1. Analizamos un método de elementos finitos mixto basado en pseudoesfuerzo para el problema de Stokes que asegura la conservación tanto de masa como de momentum. La conservación de masa se consigue mediante la aproximación de la velocidad utilizando los elementos Raviart–Thomas de orden más bajo, mientras que la conservación de momentum se consigue mediante una descomposición de Helmholtz discreta. Establecemos que el método está bien planteado y obtenemos las tasas de convergencia teóricas, incluido un resultado de superconvergencia para la aproximación del gradiente de velocidad. Una ventaja clave del método propuesto es su eficiencia computacional, ya que es ligeramente menos costoso que la aproximación clásica basada en pseudoesfuerzo estudiada en [11, 26], a la vez que garantiza la conservación de masa y momentum. Además, extendemos nuestro análisis al problema de Stokes con condiciones de contorno mixtas y presentamos experimentos numéricos que confirman los resultados teóricos.

2. Extendemos el análisis anterior al problema estacionario de Navier–Stokes; para ello introdujimos una formulación mixta y posteriormente, descomponemos la velocidad mediante una descomposición de Helmholtz y derivamos una formulación variacional mixta de tres campos, donde el pseudoesfuerzo, la velocidad y una incógnita adicional que representa la función nula, son las

principales incógnitas del sistema. Adicionalmente, recuperamos por un simple postproceso, otras variables de interés tales como la presión, gradiente de velocidad, vorticidad y tensor de esfuerzo. Desarrollamos el análisis del problema continuo, en donde empleamos los teoremas de punto fijo de Banach y Banach–Nečas–Babuška, para demostrar la solvencia única. Luego introdujimos y analizamos el problema discreto, replicando la teoría desarrollada en el problema continuo. Establecimos el análisis de error *a priori* para el método de elementos finitos mixtos propuesto, obteniendo estimaciones de estabilidad y un orden óptimo de convergencia a la solución. Desarrollamos códigos *Freefem++* para validar los resultados teóricos. Se reportan tres ejemplos numéricos que ilustran el desempeño del método de elementos finitos mixtos y la tasa de convergencia teórica.

Como consecuencia del trabajo desarrollado en esta tesis, tenemos la siguiente prepublicación sometida a revisión:

- J. Camaño, R. Oyarzúa, K. Rojo. *A momentum and mass conservative pseudostress-based mixed finite element method for the Stokes problem.* Preprint 2025-06, Centro de Investigación en Ingeniería Matemática (CI^{2}MA), Universidad de Concepción, (2025).

Trabajos futuros

El método desarrollado y los resultados obtenidos en esta tesis han motivado varios proyectos a futuro. Algunos de ellos son descritos a continuación:

1. **Extensión del análisis del problema de Navier–Stokes a tres dimensiones**.

   Como complemento al análisis desarrollado en esta tesis, nos interesa extender nuestro trabajo al caso tridimensional.

2. **Análisis de error a posteriori de métodos de elementos finitos que conserven momentum y masa para problemas de flujo de fluidos**.

   Con el fin de mejorar la robustez del error en problemas que involucren geometríias complejas o soluciones con altos gradientes, es que estamos interesados

en llevar a cabo un análisis de error *a posteriori* para los problemas estudiados en esta tesis.

3. **Nuevos esquemas numéricos para problemas acoplados (Boussinesq, Stokes–Darcy, Navier–Stokes–Darcy)**.

   Finalmente, nuestro objetivo es utilizar la teoría desarrollada en esta tesis para diseñar nuevos esquemas numéricos conservativos para problemas acoplados.

Bibliography

[1] A. Alonso-Rodríguez, J. Camaño, R. Ghiloni and A. Valli, *Graphs, spanning trees and divergence-free finite elements in domains of general topology*. IMA J. Numer. Anal. 37 (2017), no. 4, 1986–2003.

[2] V. Anaya, D. Mora, R. Oyarzúa and R. Ruiz-Baier, *A priori and a posteriori error analysis of a mixed scheme for the Brinkman problem*. Numer. Math. 133 (2016), no. 4, 781—817.

[3] D.N. Arnold and R.S. Falk, *A uniformly accurate finite element method for the Reissner–Mindlin plate*. SIAM J. Numer. Anal. vol. 26, p. 1276–1290, 1989.

[4] I. Bermúdez, J. Camaño, R. Oyarzúa and M. Solano, *A conforming mixed finite element method for a coupled Navier–Stokes/transport system modeling reverse osmosis*. Computer Methods in Applied Mechanics and Engineering vol. 433, Paper No. 117527 (2025).

[5] P. Bochev, J. Lai and L. Olson, *A locally conservative, discontinuous least-squares finite element method for the Stokes equations*. Internat. J. Numer. Methods Fluids 68 (2012), no. 6, 782–804.

[6] D. Boffi, F. Brezzi and M. Fortin, Mixed Finite Element Methods and Applications. Springer Series in Computational Mathematics, vol. 44, 2013.

[7] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*. Springer Series in Computational Mathematics, Springer-Verlag, New-York, vol. 15, 1991.

[8] Z. Cai and S. Zhang, *Mixed methods for stationary Navier–Stokes equations based on pseudostress-pressure-velocity formulation*. Mathematics of Computation, 2012, vol. 81, no 280, p. 1903-1927.

[9] Z. Cai, C. Wang and S. Zhang, *Mixed finite element methods for incompressible flow: stationary Navier–Stokes equations*. SIAM Journal on Numerical Analysis, 2010, vol. 48, no 1, p. 79-94.

[10] Z. Cai and Y. Wang, *Pseudostress–velocity formulation for incompressible Navier–Stokes equations*. International journal for numerical methods in fluids, 2010, vol. 63, no 3, p. 341-356.

[11] Z. Cai, Ch. Tong, P.S. Vassilevski and Ch. Wang, *Mixed finite element methods for incompressible flow: stationary Stokes equations*. Numer. Methods Partial Differential Equations 26 (2010), no. 4, 957—978.

[12] J. Camaño, C. García and R. Oyarzúa, *Analysis of a momentum conservative mixed-FEM for the stationary Navier–Stokes problem*. Numerical Methods for Partial Differential Equations, 2021, vol. 37, no 5, p. 2895-2923.

[13] J. Camaño, R. Oyarzúa and G. Tierra, *Analysis of an augmented mixed-FEM for the Navier–Stokes problem*. Mathematics of Computation, 2017, vol. 86, no 304, p. 589-615.

[14] J. Camaño and R. Oyarzúa, *A conforming and mass conservative pseudostress-based mixed finite element method for Stokes*. Preprint 2023-15, Centro de Investigación en Ingeniería Matemática (CI2MA), UdeC.

[15] M. Crouzex and P.-A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations I*. R.A.I.R.O., vol. 7, no. R3, p. 33–75, 1973.

[16] T.A. Davis, *Algorithm 832: UMFPACK V4. 3—an unsymmetric-pattern multifrontal method*. ACM Transactions on Mathematical Software (TOMS), 2004, vol. 30, no 2, p. 196-199.

[17] F. Dubois, M. Salaün and S. Salmon, *First vorticity–velocity–pressure numerical scheme for the Stokes problem*. Comput. Methods Appl. Mech. Engrg. 192 (2003), no. 44–46, 4877—4907.

[18] A. Ern and J.-L.Guermond, Theory and Practice of Finite Elements. Applied Mathematical Sciences, Springer-Verlag, New-York, vol. 159, 2004.

[19] M. Farhloul, S. Nicaise, and L. Paquet, *A priori and a posteriori error estimations for the dual mixed finite element method of the Navier–Stokes problem*. Numerical Methods for Partial Differential Equations, 2009, vol. 25, no 4, p. 843-869.

[20] M. Farhloul, S. Nicaise, and L. Paquet, *A refined mixed finite-element method for the stationary Navier–Stokes equations with mixed boundary conditions*. IMA journal of numerical analysis, 2008, vol. 28, no 1, p. 25-45.

[21] S.J. Fromm, *Potential space estimates for Green potentials in convex domains*. Proceedings of the American Mathematical Society, 1993, vol. 119, no 1, p. 225-233.

[22] G.N. Gatica, A Simple Introduction to the Mixed Finite Element Method. Theory and Applications, Springer Briefs in Mathematics, Springer, Cham Heidelberg New York Dordrecht London, 2014.

[23] G.N. Gatica, Introducción al Análisis Funcional. Teoría y Aplicaciones, Reverté, 2021.

[24] G.N. Gatica, A. Márquez and M.A. Sánchez, *Analysis of a velocity-pressure-pseudostress formulation for the stationary Stokes equations*. Comput. Methods Appl. Mech. Engrg. 199 (2010), no. 17–20, 1064-–1079.

[25] G.N. Gatica, A. Márquez and M.A. Sánchez, *A priori and a posteriori error analyses of a velocity-pseudostress formulation for a class of quasi-Newtonian Stokes flows*. Comput. Methods Appl. Mech. Engrg. 200 (2011), no. 17–20, 1619–1636.

[26] G.N. Gatica, A. Márquez and M.A. Sánchez, *Pseudostress-based mixed finite element methods for the Stokes problem in R^n with Dirichlet boundary*

conditions. I: A priori error analysis. Commun. Comput. Phys. 12 (2012), no. 1, 109–134.

[27] J. Gopalakrishnan, P. Lederer and J. Schöberl, *A mass conserving mixed stress formulation for the Stokes equations*. IMA J. Numer. Anal. 40 (2020), no. 3, 1838–1874.

[28] V. Girault and P. A. Raviart Finite element methods for Navier-Stokes equations: theory and algorithms (Vol. 5). Springer Science & Business Media, 2012.

[29] P. Grisvard, Elliptic Problems in Nonsmooth Domains. Classics in Applied Mathematics, 69. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[30] J. Guzmán and M. Neilan, *Conforming and divergence-free Stokes elements on general triangular meshes*. Math. Comp. 83 (2014), no. 285, 15–36.

[31] J. Guzmán and M. Neilan, *Conforming and divergence-free Stokes elements in three dimensions*. IMA J. Numer. Anal. 34 (2014), no. 4, 1489–1508.

[32] F. Hecht, *New development in FreeFem++*. Journal of numerical mathematics, 2012, vol. 20, no 3-4, p. 251-266.

[33] J.S. Howell and N.J. Walkington, *Dual-mixed finite element methods for the Navier–Stokes equations*. ESAIM: Mathematical Modelling and Numerical Analysis, 2013, vol. 47, no 3, p. 789-805.

[34] V. John, Finite element methods for incompressible flow problems. Springer Series in Computational Mathematics, 51. Springer, Cham, 2016. xiii+812 pp.

[35] 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 Rev. 59 (2017), no. 3, 492–544.

[36] L.I.G. Kovasznay, *Laminar flow behind a two-dimensional grid*. Mathematical Proceedings of the Cambridge Philosophical Society, vol. 44, no. 1, p. 58-62.

[37] P.L. Lederer and C. Merdon, *Guaranteed upper bounds for the velocity error of pressure-robust Stokes discretisations* J. Numer. Math. 30 (2022), no. 4, 267–294.

[38] J. Stam, *Real-time fluid dynamics for games*. Proceedings of the game developer conference, 2003, vol. 18, no 11.

[39] R. Tenam, Navier–Stokes equations theory and numerical analysis, Vol. 343, American Mathematical Society, 2024.

[40] D. Mitrea, *Sharp L^p–Hodge decompositions for Lipschitz domains in* $\mathbb{R}^2$. Advances in Differential Equations, 2002, vol. 7, no 3, p. 343-364.

[41] P. Monk, *A mixed method for approximating Maxwell's equations*. SIAM Journal on Numerical Analysis, 1991, vol. 28, no 6, p. 1610-1634.

[42] A. Quarteroni and A. Valli, Numerical approximation of partial differential equations, Springer series in computational mathematics, Vol. 23, Springer-Verlag, Berlin, 1994.

[43] J. Volker, Finite element methods for incompressible flow problems, Springer International Publishing, Vol. 51, 2016.