Universidad de Cantabria

Facultad
de
Ciencias

Introducción al método de los elementos finitos y aplicación al problema de Stokes

(Introduction to the Finite Element Method and application to the Stokes problem)

Trabajo de Fin de Grado

para acceder al

GRADO EN MATEMÁTICAS

Author: Emilio José Escobedo Sevilla
Director: Diana Stan
Codirector: Max von Danwitz
Date: February 2026

Agradecimientos

Por su apoyo incondicional e inquebrantable, a mis padres, Ignacio y Nuria, y a mi hermano Javier.

Por la más completa felicidad, a mi pareja y compañera, Theresa.

Por su tiempo y su paciencia, a Max y a Marco.

Por hacer posible que aprovechase esta oportunidad, a Diana.

Abstract

This work presents a rigorous introduction to the Finite Element Method and its application to the two–dimensional stationary Stokes problem. The main objective is to establish a solid mathematical basis that allows the understanding of the numerical formulation and discretization of the problem, starting from its partial differential equation form, its variational formulation, and its discretization through the Galerkin method.

The study is developed in the context of a research project linked to the German Aerospace Center, related to the structural health monitoring of wind turbine blades. In particular, it arises from the need to analyze the relationship between the aerodynamic pressure distribution and the angle of attack, with the aim of a future formulation of an inverse problem oriented to the identification of structural damage.

First, the functional foundations required for the variational formulation of partial differential equation problems are introduced, including Sobolev spaces. Afterwards, the formal definition of a finite element and the construction of nodal bases and interpolation operators, both local and global, are presented.

On this basis, the discretization of the Stokes problem with mixed boundary conditions is developed using the Galerkin method. Finally, numerical results corresponding to the simulation of a two–dimensional viscous flow around an aerodynamic profile on conforming triangular meshes are included. These results illustrate the behavior of the velocity and pressure fields.

This work does not aim to completely solve the inverse problem that motivates the application context, but to provide the theoretical framework necessary for its future development, as well as to serve as a bridge between rigorous mathematical analysis and numerical simulation in fluid mechanics.

KEYWORDS: finite element method, Stokes problem, Galerkin method, Lagrange elements, Taylor–Hood elements.

Resumen

Este trabajo presenta una introducción rigurosa al Método de los Elementos Finitos y su aplicación al problema estacionario de Stokes en dos dimensiones. El objetivo principal es establecer una base matemática sólida que permita comprender la formulación y discretización numérica del problema, partiendo de su expresión en derivadas parciales, su posterior formulación variacional y su discretización mediante el método de Galerkin.

El estudio se enmarca en el contexto de un proyecto de investigación vinculado al German Aerospace Center, relacionado con la monitorización estructural de palas de aerogeneradores. En particular, surge de la necesidad de analizar la relación entre la distribución de presión aerodinámica y el ángulo de ataque, con vistas a la futura formulación de un problema inverso orientado a la identificación de daños estructurales.

En primer lugar, se establecen los fundamentos funcionales necesarios para la formulación variacional de problemas en derivadas parciales, introduciendo los espacios de Sobolev. Posteriormente, se presenta la definición formal de elemento finito y la construcción de bases nodales e interpolantes, tanto locales como globales.

Sobre esta base, se desarrolla la discretización del problema de Stokes con condiciones de contorno mixtas mediante el método de Galerkin. Finalmente, se incluyen resultados numéricos correspondientes a la simulación del flujo bidimensional alrededor de un perfil aerodinámico en régimen viscoso, sobre mallas triangulares conformes. Estos resultados ilustran el comportamiento del campo de velocidades y de la presión.

El trabajo no pretende resolver de forma completa el problema inverso que motiva el contexto de aplicación, sino proporcionar el marco teórico necesario para su desarrollo futuro, así como servir de puente entre el análisis matemático riguroso y la simulación numérica en mecánica de fluidos.

PALABRAS CLAVE: Método de los elementos finitos, problema de Stokes, método de Galerkin, elementos de Lagrange, elementos Taylor–Hood.

Contents

List of Figures

Chapter 1

Introduction

1.1 Exposé

Nowadays renewable energies have become an essential part of the electricity supply system in Europe. Present and future are bound to the improvement and protection of these energy sources and technologies. In this context, we find wind turbines, which currently generate approximately 39% out of the total renewable electricity in the EU [8].

For this reason, protecting the existing infrastructure is also a critical mission. This is exactly the goal of the DLR PI (German Aerospace Center, Institute for the Protection of Terrestrial Infrastructures), the institution under which this work has been developed.

The evolution of the wind turbine industry is leading us to bigger devices pursuing higher efficiency, thus materials and structural components of the blades are being pushed to the limit. In this situation, structural health monitoring and predictive maintenance is required more than ever for preventing a local flaw to generate a complete failure.

To develop strategies in this field, a line of research is being carried out. From laboratory experiments [10], aerodynamic pressure measurements on an oscillating NACA 633418 airfoil seem to contain a damage sensitive feature, related to the rotatory vibrations, that allows classification of damage states. These experiments are performed in a wind tunnel under different wind speeds, angles of attack, excitation frequencies and structural states, providing time-resolved pressure data for both undamaged and damaged configurations.

Damage to the structure leads to changes in its vibration behavior. These changes affect the aerodynamic loading of the airfoil and, as a result, the measured pressure distribution. The angle of attack plays an important role in the interaction between structural motion and aerodynamic forces. Being able to estimate this parameter, together with its variations over time, from pressure measurements is therefore important for a clearer interpretation of the damage-sensitive features observed in the experiments.

For this reason, an important step in this research line is the formulation of an inverse problem, together with a reliable numerical solver, to infer the angle of attack from the pressure distribution over the blade surface. In this context, a differentiable finite element solver for the forward flow problem has already been implemented within a JAX-based

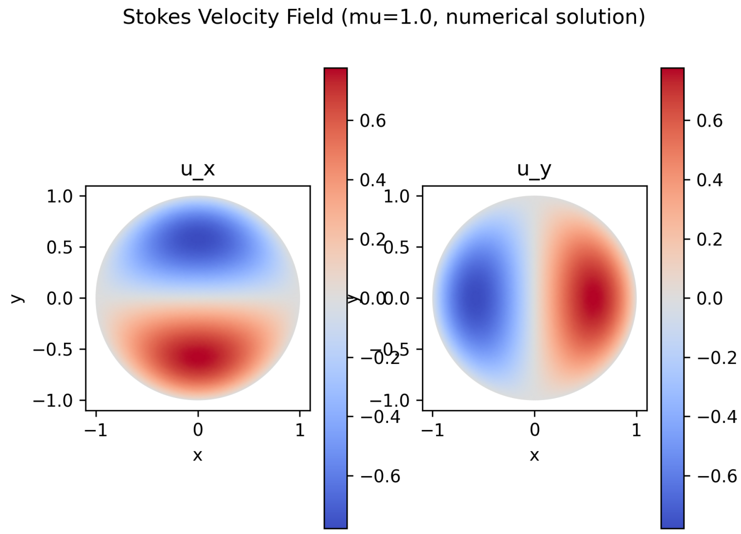


Figure 1.1: Polynomial solution on a circular domain with differentiable parameterized source term

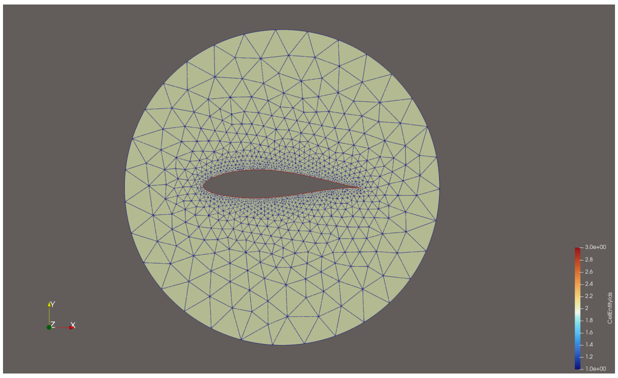


Figure 1.2: Discretization of flow domain around NACA 633418 airfoil.

computational framework.

The present thesis does not focus on the full development of the inverse identification procedure. Instead, its aim is to provide a clear and rigorous mathematical formulation of the forward finite element solver on which the numerical framework is based. In particular, the variational formulation of the stationary Stokes problem is developed in detail, together with the corresponding functional setting and the construction of conforming finite element discretizations.

By making the mathematical structure of the forward problem explicit, this work establishes the theoretical basis required for the consistent use of the existing differentiable solver within future inverse formulations.

As a first implementation example of this work, Figure 1.1 illustrates a polynomial solution on a circular domain with parameterized source term. The current JAX implementation provides automatic differentiation of the solution with respect to the source term parameters. The next step is to extend the approach to a domain given by the airfoil geometry (see Figure 1.2) and implement the angle of attack as parameter input.

Furthermore, although this lies beyond the scope of the present thesis and belongs to future work, the differentiable flow solver opens the possibility of developing FEM-based neural networks. Such approaches may allow a better preconditioning of physics-informed loss functions, leading to improved convergence properties and greater stability of the training process compared to purely data-driven methods [12].

More generally, the integration of measurement data and physical models through differentiable solvers provides a framework for sensitivity analysis, parameter identification, and hybrid modelling. This is particularly relevant in structural health monitoring applications, where measurement data are limited, affected by noise, and where maintaining physical consistency of the solution is essential.

1.2 Structure of the document

The thesis consists of four main chapters and an appendix, in addition to the present introductory chapter, which progressively develop the mathematical foundations of the finite element method and culminate in its application to the two–dimensional Stokes problem within the context described in the Exposé.

Chapter 2 presents the variational framework underlying the boundary value problems considered in this work. It begins with symmetric coercive variational problems, including existence and uniqueness results and their Galerkin approximation, using the one–dimensional Poisson equation as a model problem. The chapter then introduces mixed and non–symmetric variational problems and formulates the stationary two–dimensional Stokes problem in weak form, highlighting the functional setting.

Chapter 3 develops the finite element method from a rigorous mathematical perspective. It starts with the formal definition of a finite element based on a triple (K, P, N), including unisolvence, nodal bases and local interpolation operators. The geometric framework is then introduced through conforming triangular meshes and affine mappings from the reference triangle. Concrete elements are presented, including the Lagrange P_1 and P_2 elements and the Taylor–Hood P_2/P_1 pair. The chapter concludes with the Galerkin discretization of the Stokes problem and the derivation of the associated algebraic linear system.

Chapter 4 contains the numerical results. The computational domain, geometry and boundary conditions are specified, including the discretization of the flow domain around the NACA 633418 airfoil.

Chapter 5 summarizes the main conclusions of the work and outlines the directions for future research.

The Appendix collects the preliminary definitions and fundamental results from linear algebra, Hilbert and Banach spaces, Sobolev spaces and weak derivatives that support the theoretical developments of the main text.

Chapter 2

Variational theory

This chapter develops the variational framework underlying the boundary value problems studied in this work. The finite element method is fundamentally designed to approximate solutions of such problems in their variational (weak) form, rather than in their classical differential formulation. In this sense, the variational setting is not a technical reformulation, but the natural analytical framework in which existence, uniqueness, and stability of solutions can be rigorously established. Moreover, it is precisely this structure that enables the systematic construction of conforming finite element approximations.

The precise statements of the continuous models considered in this work, can be found in [6]. The fundamental analytical results concerning well-posedness and stability required for their variational formulation are presented in [1], which also provides the theoretical foundation for the finite element discretization discussed throughout this work.

2.1 Symmetric coercive variational problems

Although the main problem addressed in this work is the two–dimensional Stokes equation, the one–dimensional Poisson problem is included as a preliminary model problem. From the point of view of variational analysis, the Stokes equation give rise to mixed and non-symmetric problems, whereas the Poisson problem leads to a symmetric and coercive variational formulation, which is mathematically simpler. For this reason, the Poisson problem provides a convenient setting to introduce the basic variational concepts and analytical tools that will later be reused in the mixed framework valid for the Stokes problem.

2.1.1 Problem statement and existence of a solution

Definition 2.1.1 (Symmetric coercive variational problem)**.** *Let $(V, (\cdot,\cdot)_V)$ be a real Hilbert space as in Definition A.2.6. Let $a : V \times V \to \mathbb{R}$ be a bilinear form which is bounded, symmetric, and coercive on V in the sense of Definition A.2.14. Let $F \in V'$ be a continuous linear functional, where V' denotes the topological dual of V as in Definition A.2.12. The* *symmetric coercive variational problem consists of finding $u \in V$ such that*

$$a(u,v) = F(v) \qquad \forall v \in V. \tag{SCV}$$

Theorem 2.1.2 (Existence and uniqueness of the solution to the symmetric variational problem)**.** *The symmetric coercive variational problem posed in Definition 2.1.1 always admits a unique solution $u \in V$.*

Proof. **1. (V, a) is a Hilbert Space**

In the first place, we prove that (V, a) is also a Hilbert space. Since $a : V \times V \to \mathbb{R}$ is bilinear and symmetric by assumption, it only remains to verify positive definiteness in order to conclude that a defines an inner product. By coercivity, there exists $\alpha > 0$ such that

$$a(v,v) \geq \alpha \|v\|_V^2 \qquad \forall v \in V.$$

In particular, $a(v,v) \geq 0$ for all $v \in V$, and $a(v,v) = 0$ implies $\|v\|_V = 0$, hence $v = 0$. Therefore, a satisfies the axioms of Definition A.2.1 and defines an inner product on V, which we denote by

$$(v,w)_a := a(v,w).$$

By Proposition A.2.4, the inner product $(\cdot,\cdot)_a$ induces a norm on V, given by

$$\|v\|_a := \sqrt{a(v,v)}, \qquad v \in V,$$

and hence a metric $d_a(u,v) := \|u - v\|_a$, which allows us to study convergence and completeness (Definition A.2.5).

Let $(v_n) \subset V$ be a Cauchy sequence with respect to the norm $\|\cdot\|_a$. Using coercivity again, we obtain for all $w \in V$ that

$$\|w\|_a^2 = a(w,w) \geq \alpha \|w\|_V^2,$$

which implies

$$\|w\|_V \leq \frac{1}{\sqrt{\alpha}} \|w\|_a.$$

Applying this estimate to $w = v_n - v_m$, it follows that (v_n) is also a Cauchy sequence in the norm $\|\cdot\|_V$. Since $(V, (\cdot,\cdot)_V)$ is a Hilbert space by Definition A.2.6, there exists $v \in V$ such that

$$v_n \to v \quad \text{in } (V, \|\cdot\|_V).$$

To conclude convergence with respect to $\|\cdot\|_a$, we use the boundedness of a (Definition A.2.14), which ensures the existence of a constant $M > 0$ such that

$$|a(w,w)| \leq M \|w\|_V^2 \qquad \forall w \in V.$$

Hence,

$$\|w\|_a \leq \sqrt{M}\, \|w\|_V \qquad \forall w \in V.$$

Applying this inequality to $w = v_n - v$ yields

$$\|v_n - v\|_a \leq \sqrt{M}\, \|v_n - v\|_V \xrightarrow[n\to\infty]{} 0.$$

Therefore, every Cauchy sequence in $(V, \|\cdot\|_a)$ converges in V, and we conclude that $(V, (\cdot,\cdot)_a)$ is a Hilbert space.

2. Existence of a solution for the (SCV) problem

In the second place, we argue directly in this Hilbert space and we show the existence of the solution. Let $F \in V'$ be the continuous linear functional given by assumption. If F is identically nule, then $u = 0_V$ satisfies $a(u, v) = F(v)$ for all $v \in V$, and uniqueness is immediate. Assume henceforth that exist $v \in V$ such that $F(v) \neq 0_V$ and set

$$M := \ker(F) = \{ v \in V \mid F(v) = 0 \}.$$

M is a sub-Hilbert space of $(V, (\cdot,\cdot)_a)$. Since F is linear, M is a vector subspace of V. Moreover, F is continuous (Definition A.2.9), so if $(m_n) \subset M$ and $m_n \to m$ in $(V, \|\cdot\|_a)$, then by continuity

$$F(m) = \lim_{n\to\infty} F(m_n) = 0,$$

hence $m \in M$. Therefore M is closed in $(V, \|\cdot\|_a)$, and consequently it is a Hilbert space with the induced inner product (Remark A.2.8).

Since M is a closed subspace of the Hilbert space $(V, (\cdot,\cdot)_a)$, we may use the orthogonal decomposition theorem (Theorem 2.3.5 in [1]) and write

$$V = M \oplus M^{\perp_a}, \qquad M^{\perp_a} := \{ w \in V \mid a(w, m) = 0 \ \forall m \in M \}.$$

Because $F \neq 0$, we have $M \neq V$, hence $M^{\perp_a} \neq \{0\}$. So we can pick $w \in M_a^{\perp}$, $w \neq 0$.

Let $v \in V$ be arbitrary. By the orthogonal decomposition there exist unique $m \in M$ and $\lambda \in \mathbb{R}$ such that

$$v = m + \lambda w.$$

Applying F and using that $M = \ker(F)$, we obtain

$$F(v) = F(m) + \lambda F(w) = \lambda F(w).$$

On the other hand, since $w \in M^{\perp_a}$, we have

$$a(w, v) = a(w, m) + \lambda a(w, w) = \lambda a(w, w).$$

Combining both identities yields

$$F(v) = \lambda F(w) = \frac{F(w)}{a(w, w)}\, a(w, v) \qquad \forall v \in V.$$

We now define

$$u := \frac{F(w)}{a(w, w)}\, w.$$

It is immediate that $u \in V$. Then, for every $v \in V$, we have

$$a(u, v) = \frac{F(w)}{a(w, w)}\, a(w, v) = F(v),$$

which proves the existence of a solution to *(SCV)*.

3. Uniqueness

Assume that $u_1, u_2 \in V$ satisfy $F(v) = a(u_1, v) = a(u_2, v)$ for all $v \in V$. Then $a(u_1 - u_2, v) = 0$ for all $v \in V$. Taking $v = u_1 - u_2$ gives

$$a(u_1 - u_2, u_1 - u_2) = 0.$$

By coercivity of a (Definition A.2.14), this implies $u_1 - u_2 = 0$, hence $u_1 = u_2$.

In particular, there exists a unique $u \in V$ such that

$$a(u, v) = F(v) \qquad \forall v \in V,$$

which establishes existence and uniqueness of the solution to *(SCV)*. □

2.1.2 Galerkin approximation of the symmetric coercive problem

The variational formulation *(SCV)* is, in general, an infinite-dimensional problem, which means that the solution is sought in a Hilbert space V of infinite dimension. Consequently, the problem cannot be solved directly by finite linear algebra techniques. The Galerkin method replaces the infinite-dimensional space V by a finite-dimensional subspace and seeks an approximate solution within that space.

Definition 2.1.3 (Galerkin approximation of the symmetric coercive problem)**.** *Assume the hypotheses of Definition 2.1.1 and let $V_h \subset V$ be a finite-dimensional subspace. The Galerkin approximation of* (SCV) *consists of finding $u_h \in V_h$ such that*

$$a(u_h, v_h) = F(v_h) \qquad \forall v_h \in V_h. \tag{SCV$_h$}$$

Theorem 2.1.4 (Existence and uniqueness of the Galerkin solution)**.** *Under the hypotheses of Definition 2.1.1, for every finite-dimensional subspace $V_h \subset V$, the discrete problem* (SCV$_h$) *admits a unique solution $u_h \in V_h$.*

Proof. The restriction $a|_{V_h \times V_h}$ is bilinear, symmetric, and bounded. Moreover, since a is coercive on V, there exists $\alpha > 0$ such that

$$a(v_h, v_h) \geq \alpha \|v_h\|_V^2 \qquad \forall v_h \in V_h,$$

so a is also coercive on V_h. Hence, a defines an inner product on V_h and, since V_h is finite-dimensional, (V_h, a) is a Hilbert space.

Because $F \in V'$ is continuous on V, its restriction $F|_{V_h}$ is a continuous linear functional on V_h. Therefore, by applying the same argument used in Theorem 2.1.2 to the Hilbert space (V_h, a), there exists a unique $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h) \qquad \forall v_h \in V_h.$$

□

The practical effectiveness of the Galerkin method depends on the choice of the finite-dimensional subspace V_h. Finite element methods consist precisely in constructing suitable spaces $V_h \subset V$ that can achieve good approximation properties, computational efficiency, and conformity with the functional setting of the variational problem.

2.1.3 The one–dimensional Poisson problem

This section introduces the one–dimensional Poisson problem as a model example of a symmetric and coercive variational formulation. We first state the boundary value problem in strong form, then derive its variational formulation by testing and integrating by parts, and finally prove the equivalence between both formulations under appropriate regularity assumptions. We also show that the resulting variational problem is a particular case of the abstract symmetric coercive variational framework introduced in Section 2.1.

Definition 2.1.5 (Strong Poisson problem in one dimension). *Let $\Omega = (a, b) \subset \mathbb{R}$ be an open bounded interval and let $f \in C^0(\overline{\Omega})$ be given. The one–dimensional Poisson problem with homogeneous Dirichlet boundary conditions consists of finding a function $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ such that*

$$\begin{cases} -u''(x) = f(x), & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega, \end{cases} \tag{P$_{strong}$}$$

A function u satisfying ($\mathrm{P}_{\mathrm{strong}}$) *is called a* strong solution *of the one–dimensional Poisson problem.*

Definition 2.1.6 (Variational problem of Poisson in one dimension). *Let $\Omega = (a, b) \subset \mathbb{R}$ be an open bounded interval and let $f \in L^2(\Omega)$ be given. Set $V := H_0^1(\Omega)$ introduced in Definition A.3.20. The variational formulation of the one–dimensional Poisson problem consists of finding $u \in V$ such that*

$$a(u, v) = F(v) \qquad \forall v \in V, \tag{P$_{\mathrm{weak}}$}$$

where the bilinear form $a : V \times V \to \mathbb{R}$ and the linear functional $F \in V'$ are defined by

$$a(u, v) := \int_\Omega u'(x)\, v'(x)\, \mathrm{d}x, \qquad F(v) := \int_\Omega f(x)\, v(x)\, \mathrm{d}x,$$

and where the derivatives u' and v' are understood in the sense of Definition A.3.15. Since $u, v \in H_0^1(\Omega) \subset H^1(\Omega)$, their weak derivatives are well defined and belong to $L^2(\Omega)$.

A function $u \in V$ satisfying ($\mathrm{P}_{\mathrm{weak}}$) is called a weak solution *of the one–dimensional Poisson problem.*

Proposition 2.1.7 (Strong solutions are weak solutions). *Let $\Omega = (a, b) \subset \mathbb{R}$ be an open bounded interval and let $f \in C^0(\overline{\Omega})$ be given. If $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ is a solution of the strong Poisson problem* ($\mathrm{P}_{\mathrm{strong}}$)*, then u is also a solution of the variational problem* ($\mathrm{P}_{\mathrm{weak}}$).

Proof. Since $u \in C^0(\overline{\Omega})$ and $\overline{\Omega}$ is compact, the function u is bounded on Ω. As Ω has finite measure, this implies $u \in L^2(\Omega)$. Since $f \in C^0(\overline{\Omega})$ and $-u'' = f$ in Ω, the second derivative u'' is bounded on Ω. Fix any $x_0 \in \Omega$. For all $x \in \Omega$, the fundamental theorem of calculus gives

$$u'(x) = u'(x_0) + \int_{x_0}^{x} u''(s)\, \mathrm{d}s,$$

and therefore

$$|u'(x) - u'(x_0)| \le \int_{x_0}^{x} |u''(s)|\,\mathrm{d}s \le C|x - x_0|,$$

for some constant $C > 0$. This estimate shows that the variation of u' is controlled by the distance between points, and in particular implies that u' cannot blow up when approaching the boundary. Consequently, the limits

$$\lim_{x\to a^+} u'(x) \quad \text{and} \quad \lim_{x\to b^-} u'(x)$$

exist and are finite, so that u' admits a continuous extension to $\overline{\Omega}$. Since $\overline{\Omega}$ is compact, this extension is bounded, and hence $u' \in L^2(\Omega)$. Together with $u \in C^0(\overline{\Omega}) \subset L^2(\Omega)$, this yields $u \in H^1(\Omega)$.

Finally, u vanishes on $\partial\Omega = \{a, b\}$ in the classical sense. Since $H_0^1(\Omega)$ is defined as the closure of $C_c^\infty(\Omega)$ in the H^1 norm, and functions in $H^1(\Omega)$ that vanish at the boundary of a bounded interval can be approximated in $H^1(\Omega)$ by smooth functions with compact support, it follows that $u \in H_0^1(\Omega)$.

Moreover, since $u' \in C^1(\Omega)$ and $u'' \in C^0(\Omega)$, the classical derivatives of u belong to $L^1_{\mathrm{loc}}(\Omega)$ and therefore coincide with the weak derivatives in the sense of Definition A.3.15. Let $v \in H_0^1(\Omega)$ be arbitrary. Multiplying the differential equation $-u'' = f$ by v and integrating over Ω, we obtain

$$\int_\Omega f\,v\,\mathrm{d}x = -\int_\Omega u''\,v\,\mathrm{d}x.$$

An integration by parts yields

$$-\int_\Omega u''\,v\,\mathrm{d}x = \int_\Omega u'\,v'\,\mathrm{d}x - \big[u'(x)v(x)\big]_{\partial\Omega}.$$

Since $v \in H_0^1(\Omega)$, its trace vanishes on $\partial\Omega$, and the boundary term is zero. Therefore,

$$\int_\Omega u'\,v'\,\mathrm{d}x = \int_\Omega f\,v\,\mathrm{d}x \qquad \forall v \in H_0^1(\Omega),$$

which is precisely the variational formulation *(P_{weak})*. □

Proposition 2.1.8 (Weak solutions are strong solutions under additional regularity)**.** *Let $\Omega \subset \mathbb{R}$ be an open bounded interval and let $f \in L^2(\Omega)$ be given. Assume that $u \in H_0^1(\Omega)$ is a solution of the variational problem* ($\mathrm{P}_{\mathrm{weak}}$)*. If, in addition,*

$$u \in C^2(\Omega) \cap C^0(\overline{\Omega}),$$

then u is a solution of the strong Poisson problem ($\mathrm{P}_{\mathrm{strong}}$)*.*

Proof. Assume that $u \in H_0^1(\Omega)$ satisfies the variational problem *(P_{weak})*, and assume in addition that $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$.

Let $\varphi \in C_c^\infty(\Omega)$. Since $C_c^\infty(\Omega) \subset H_0^1(\Omega)$, we may choose $v = \varphi$ as a test function in *(P_{weak})* and obtain

$$\int_\Omega u'(x)\,\varphi'(x)\,\mathrm{d}x = \int_\Omega f(x)\,\varphi(x)\,\mathrm{d}x.$$

Because $u \in C^2(\Omega)$, we have $u' \in C^1(\Omega)$ and $u'' \in C^0(\Omega)$, so the classical derivatives belong to $L^1_{\text{loc}}(\Omega)$ and coincide with the weak derivatives. Moreover, since φ has compact support in Ω, integration by parts yields

$$\int_\Omega u'(x)\,\varphi'(x)\,\mathrm{d}x = -\int_\Omega u''(x)\,\varphi(x)\,\mathrm{d}x.$$

Combining both identities, we obtain

$$-\int_\Omega u''(x)\,\varphi(x)\,\mathrm{d}x = \int_\Omega f(x)\,\varphi(x)\,\mathrm{d}x \qquad \forall \varphi \in C_c^\infty(\Omega),$$

that is,

$$\int_\Omega \big(-u''(x) - f(x)\big)\,\varphi(x)\,\mathrm{d}x = 0 \qquad \forall \varphi \in C_c^\infty(\Omega).$$

Since u'' and f are continuous in Ω, it follows that

$$-u'' - f = 0 \quad \text{in } \Omega.$$

This is a direct consequence of Lemma A.3.28.

Finally, since $u \in H^1_0(\Omega)$, its trace on $\partial\Omega$ vanishes. Because we also have $u \in C^0(\overline{\Omega})$, this implies

$$u(x) = 0 \qquad \text{for all } x \in \partial\Omega.$$

Therefore, u satisfies the strong Poisson problem *(*P_{strong}*)*. □

Proposition 2.1.9 (Poisson 1D as a symmetric coercive variational problem)**.** *The variational Poisson problem in one dimension defined in Definition 2.1.6 fits into the abstract framework of symmetric coercive variational problems introduced in Definition 2.1.1.*

Proof. First, we recall that the space $V = H^1_0(\Omega)$ is a real Hilbert space when endowed with the norm

$$\|v\|_{H^1(\Omega)} := \left(\|v\|^2_{L^2(\Omega)} + \|v'\|^2_{L^2(\Omega)}\right)^{1/2},$$

and the associated inner product. This is a classical result; see, for instance, Chapter 1 of [1].

The bilinear form

$$a(u,v) = \int_\Omega u'(x)\,v'(x)\,\mathrm{d}x$$

is clearly bilinear and symmetric.

To prove boundedness, let $u, v \in V$. By the Cauchy–Schwarz inequality,

$$|a(u,v)| = \left|\int_\Omega u'(x)\,v'(x)\,\mathrm{d}x\right| \le \|u'\|_{L^2(\Omega)}\,\|v'\|_{L^2(\Omega)} \le \|u\|_{H^1(\Omega)}\,\|v\|_{H^1(\Omega)}.$$

To prove coercivity, we use the Poincaré inequality on $H^1_0(\Omega)$, see Theorem A.3.27, which ensures the existence of a constant $C_P > 0$ such that

$$\|v\|_{L^2(\Omega)} \le C_P\,\|v'\|_{L^2(\Omega)} \qquad \forall v \in H^1_0(\Omega).$$

Therefore,

$$\|v\|_{H^1(\Omega)}^2 = \|v\|_{L^2(\Omega)}^2 + \|v'\|_{L^2(\Omega)}^2 \le (C_P^2 + 1)\,\|v'\|_{L^2(\Omega)}^2 = (C_P^2 + 1)\,a(v,v),$$

which implies

$$a(v,v) \ge \alpha\,\|v\|_{H^1(\Omega)}^2 \qquad \forall v \in V, \quad \text{with } \alpha := \frac{1}{C_P^2 + 1} > 0.$$

Hence, a is coercive on V.

Finally, the linear functional

$$F(v) = \int_\Omega f(x)\,v(x)\,\mathrm{d}x$$

is well defined on V and belongs to the dual space V'. Indeed, by the Cauchy–Schwarz inequality, Lemma A.2.3 , and the Poincaré inequality, Theorem A.3.27,

$$|F(v)| \le \|f\|_{L^2(\Omega)}\,\|v\|_{L^2(\Omega)} \le C_P\,\|f\|_{L^2(\Omega)}\,\|v'\|_{L^2(\Omega)} \le C_P\,\|f\|_{L^2(\Omega)}\,\|v\|_{H^1(\Omega)},$$

which shows, by Theorem A.2.10 that F is a continuous linear functional on V.

All the assumptions of Definition 2.1.1 are therefore satisfied, and the variational Poisson problem in one dimension fits into the class of symmetric coercive variational problems. □

Remark 2.1.10. *Consequently, the variational problem* ($\mathrm{P}_{\mathrm{weak}}$) *is well posed. More precisely, it admits a unique solution $u \in V$, whose existence and uniqueness follow from Theorem 2.1.2. Moreover, this solution can be approximated by the Galerkin method in any finite–dimensional subspace $V_h \subset V$ according to Definition 2.1.3.*

2.2 Mixed non-symmetric variational problems

Despite its importance, the symmetric coercive variational framework is not sufficiently general to cover problems involving multiple unknown fields. In particular, the requirement of global coercivity fails in the presence of incompressibility constraints, which are intrinsic to many problems in fluid mechanics. This limitation naturally leads to the introduction of mixed and saddle–point variational formulations. In such formulations, the unknown of the problem is composed of two or more fields, each approximated in a distinct function space, and the variational problem is characterized by the interaction of multiple bilinear forms. As a consequence, the standard Galerkin formulation can not tolerate all combinations of function spaces for the different unknown fields, because not all of them lead to stable and convergent approximations.

The mathematical analysis of mixed variational problems is therefore based on a different set of stability conditions. Pioneering works by Babuška and Brezzi established that the well-posedness of mixed problems relies on two key requirements: a coercivity condition on a suitable subspace and an additional stability condition, commonly referred to as the inf–sup or Ladyženskaya–Babuška–Brezzi (LBB) condition. These results provide the theoretical foundation for the analysis of mixed finite element methods and play a central role in the numerical approximation of the Stokes and Navier–Stokes equations.

2.2.1 The two–dimensional Stokes problem

Definition 2.2.1 (Strong Stokes problem with mixed boundary conditions). *Let $\Omega \subset \mathbb{R}^2$ be a bounded, non-empty, connected open set with Lipschitz boundary $\partial\Omega$. Assume that the boundary is decomposed into two disjoint parts*

$$\partial\Omega = \Gamma_D \cup \Gamma_N, \qquad \Gamma_D \cap \Gamma_N = \varnothing,$$

with Γ_D of positive boundary measure. Let $\mathbf{n}$ denote the outward unit normal to $\partial\Omega$. Let $\mu > 0$ denote the dynamic viscosity of the fluid and $\rho > 0$ its mass density. Let $\mathbf{f} \in C^0(\overline{\Omega})^2$ be a given body force per unit mass, and let $\mathbf{g} \in C^2(\overline{\Omega})^2$ be prescribed Dirichlet data.

The stationary incompressible Stokes problem with non-homogeneous Dirichlet conditions on Γ_D and natural (do-nothing) outflow conditions on Γ_N consists of finding a velocity field

$$\mathbf{u} \in C^2(\overline{\Omega})^2$$

and a pressure field

$$p \in C^1(\overline{\Omega})$$

such that

$$\begin{cases} -\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p) = \rho\,\mathbf{f}, & \textit{in } \Omega, \\ \nabla \cdot \mathbf{u} = 0, & \textit{in } \Omega, \\ \mathbf{u} = \mathbf{g}, & \textit{on } \Gamma_D, \\ \boldsymbol{\sigma}(\mathbf{u}, p)\,\mathbf{n} = \mathbf{0}, & \textit{on } \Gamma_N, \end{cases} \tag{$\mathrm{S}^{\mathrm{mix}}_{\mathrm{strong}}$}$$

where the Cauchy stress tensor is defined by

$$\boldsymbol{\sigma}(\mathbf{u}, p) := 2\mu\,\boldsymbol{\varepsilon}(\mathbf{u}) - p\,\mathbf{I}, \qquad \boldsymbol{\varepsilon}(\mathbf{u}) := \tfrac{1}{2}\big(\nabla\mathbf{u} + (\nabla\mathbf{u})^\top\big),$$

and $\mathbf{I}$ denotes the identity tensor.

The divergence operator acting on a vector field $\mathbf{v} = (v_1, v_2)$ is defined by

$$\nabla \cdot \mathbf{v} := \partial_i v_i,$$

and, for a tensor field $\boldsymbol{\tau} = (\tau_{ij})_{i,j=1}^2$, its divergence is defined row–wise by

$$(\nabla \cdot \boldsymbol{\tau})_i := \partial_j \tau_{ij}, \qquad i = 1, 2,$$

where repeated indices are summed (Einstein convention).

A pair $(\mathbf{u}, p)$ satisfying ($\mathrm{S}^{\mathrm{mix}}_{\mathrm{strong}}$) is called a strong solution *of the stationary Stokes problem with mixed boundary conditions.*

Definition 2.2.2 (Variational formulation of the Stokes problem with mixed boundary conditions). *Let $\Omega \subset \mathbb{R}^2$ be a bounded, connected, open set with Lipschitz boundary. Assume that the boundary is decomposed as*

$$\partial\Omega = \Gamma_D \cup \Gamma_N, \qquad \Gamma_D \cap \Gamma_N = \varnothing,$$

with $|\Gamma_D| > 0$. Let $\mu > 0$ be the dynamic viscosity and $\rho > 0$ the density. Let $\mathbf{f} \in L^2(\Omega)^2$ be a given body force per unit mass. Let $\mathbf{g} \in C^1(\Gamma_D)^2$ be prescribed Dirichlet data.

We define the homogeneous test space

$$V_0 := \left\{ \mathbf{v} \in H^1(\Omega)^2 \;\middle|\; \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D \right\},$$

the affine solution space

$$V_{\mathbf{g}} := \left\{ \mathbf{v} \in H^1(\Omega)^2 \;\middle|\; \mathbf{v} = \mathbf{g} \text{ on } \Gamma_D \right\},$$

where the boundary condition is understood in the sense of traces, and the (mean–zero) pressure space

$$Q := \left\{ q \in L^2(\Omega) \;\middle|\; \int_\Omega q\,dx = 0 \right\}.$$

The stationary incompressible Stokes problem with mixed boundary conditions consists of finding $(\mathbf{u}, p) \in V_{\mathbf{g}} \times Q$ *such that*

$$\begin{cases} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = F(\mathbf{v}), & \forall \mathbf{v} \in V_0, \\ b(\mathbf{u}, q) = 0, & \forall q \in Q, \end{cases} \tag{$\mathrm{S}^{\mathrm{mix}}_{\mathrm{weak}}$}$$

where

$$a(\mathbf{u}, \mathbf{v}) := \int_\Omega 2\mu\,\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,dx,$$

$$b(\mathbf{v}, q) := -\int_\Omega q\,\nabla \cdot \mathbf{v}\,dx,$$

$$F(\mathbf{v}) := \int_\Omega \rho\,\mathbf{f} \cdot \mathbf{v}\,dx.$$

The symmetric gradient is defined by

$$\boldsymbol{\varepsilon}(\mathbf{w}) := \tfrac{1}{2}\big(\nabla \mathbf{w} + (\nabla \mathbf{w})^\top\big).$$

The gradient of a vector field $\mathbf{w} = (w_1, w_2)$ *is defined in index notation by*

$$(\nabla \mathbf{w})_{ij} := \frac{\partial w_i}{\partial x_j}, \qquad i, j = 1, 2.$$

The divergence of a vector field $\mathbf{v} = (v_1, v_2)$ *is defined by*

$$\nabla \cdot \mathbf{v} := \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2}.$$

For two matrices $A = (a_{ij})$ *and* $B = (b_{ij})$ *in* $\mathbb{R}^{2\times 2}$*, their Frobenius inner product (double contraction) is defined by*

$$A : B := \sum_{i,j=1}^{2} a_{ij}\,b_{ij}.$$

Alll derivatives are understood in the sense of weak derivatives.

A pair $(\mathbf{u}, p)$ *satisfying* ($\mathrm{S}^{\mathrm{mix}}_{\mathrm{weak}}$) *is called a* weak solution *of the stationary incompressible Stokes problem with mixed boundary conditions.*

Proposition 2.2.3. *Under the assumptions of the strong problem stated in Definition 2.2.1, let*

$$(\mathbf{u}, p) \in C^2(\overline{\Omega})^2 \times C^1(\overline{\Omega})$$

be a strong solution of $(\mathrm{S}^{\mathrm{mix}}_{\mathrm{strong}})$*. Then there exists a constant* $c \in \mathbb{R}$ *such that* $(\mathbf{u}, \widetilde{p})$*, with* $\widetilde{p} := p + c$*, is a weak solution of* $(\mathrm{S}^{\mathrm{mix}}_{\mathrm{weak}})$ *in the sense of Definition 2.2.2.*

Proof. **1. Functional setting**

Since Ω is bounded, $|\Omega| < \infty$. Any function in $C^0(\overline{\Omega})$ is bounded on Ω, and therefore belongs to $L^2(\Omega)$. In particular,

$$\mathbf{f} \in L^2(\Omega)^2, \qquad \mathbf{u} \in L^2(\Omega)^2, \qquad p \in L^2(\Omega).$$

Moreover, from $\mathbf{u} \in C^2(\overline{\Omega})^2$ it follows that all first–order partial derivatives $\partial_j u_i$ are continuous on $\overline{\Omega}$ and hence bounded. Consequently, $\nabla \mathbf{u} \in L^2(\Omega)^{2\times 2}$ and therefore $\mathbf{u} \in H^1(\Omega)^2$.

Since $\mathbf{u} = \mathbf{g}$ on Γ_D in the classical sense and both $\mathbf{u}$ and $\mathbf{g}$ are continuous on $\overline{\Omega}$, their traces coincide on Γ_D. Thus $\mathbf{u} \in V_{\mathbf{g}}$.

Finally, the pressure is determined only up to an additive constant. Defining

$$c := -\frac{1}{|\Omega|} \int_\Omega p \, dx, \qquad \widetilde{p} := p + c,$$

we obtain $\int_\Omega \widetilde{p} \, dx = 0$, and hence $\widetilde{p} \in Q$.

2. Weak momentum equation

Let $\mathbf{v} \in V_0$. Since $(\mathbf{u}, p)$ is a strong solution, we have

$$-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p) = \rho \mathbf{f} \qquad \text{in } \Omega,$$

with $\boldsymbol{\sigma}(\mathbf{u}, p) \in C^1(\overline{\Omega})^{2\times 2}$. Multiplying by $\mathbf{v}$ and integrating over Ω yields

$$\int_\Omega \big(-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p)\big) \cdot \mathbf{v} \, dx = \int_\Omega \rho \, \mathbf{f} \cdot \mathbf{v} \, dx.$$

Using index notation and integrating by parts componentwise,

$$\int_\Omega \big(-\partial_j \sigma_{ij}(\mathbf{u}, p)\big) v_i \, dx = \int_\Omega \sigma_{ij}(\mathbf{u}, p) \, \partial_j v_i \, dx - \int_{\partial\Omega} \sigma_{ij}(\mathbf{u}, p) \, v_i \, n_j \, ds.$$

In tensor notation, this reads

$$\int_\Omega \big(-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p)\big) \cdot \mathbf{v} \, dx = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}, p) : \nabla \mathbf{v} \, dx - \int_{\partial\Omega} \big(\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}\big) \cdot \mathbf{v} \, ds.$$

We decompose the boundary term as

$$\int_{\partial\Omega} (\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}) \cdot \mathbf{v} \, ds = \int_{\Gamma_D} (\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}) \cdot \mathbf{v} \, ds + \int_{\Gamma_N} (\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n}) \cdot \mathbf{v} \, ds.$$

Since $\mathbf{v} \in V_0$, its trace vanishes on Γ_D, hence the first integral is zero. On Γ_N, the strong boundary condition gives $\boldsymbol{\sigma}(\mathbf{u}, p)\mathbf{n} = \mathbf{0}$, so the second integral is also zero. Therefore,

$$\int_\Omega \big(-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, p)\big) \cdot \mathbf{v}\, dx = \int_\Omega \boldsymbol{\sigma}(\mathbf{u}, p) : \nabla \mathbf{v}\, dx.$$

Using $\boldsymbol{\sigma}(\mathbf{u}, p) = 2\mu\, \boldsymbol{\varepsilon}(\mathbf{u}) - p\mathbf{I}$ and $\mathbf{I} : \nabla \mathbf{v} = \nabla \cdot \mathbf{v}$, we obtain

$$\int_\Omega \boldsymbol{\sigma}(\mathbf{u}, p) : \nabla \mathbf{v}\, dx = \int_\Omega 2\mu\, \boldsymbol{\varepsilon}(\mathbf{u}) : \nabla \mathbf{v}\, dx - \int_\Omega p\, \nabla \cdot \mathbf{v}\, dx.$$

Moreover, since $\boldsymbol{\varepsilon}(\mathbf{u})$ is symmetric, we have $\boldsymbol{\varepsilon}(\mathbf{u}) : \nabla \mathbf{v} = \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})$. Indeed, in indices,

$$\boldsymbol{\varepsilon}(\mathbf{u}) : \nabla \mathbf{v} = \frac{1}{2}(\partial_j u_i + \partial_i u_j)\, \partial_j v_i, \qquad \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v}) = \frac{1}{4}(\partial_j u_i + \partial_i u_j)(\partial_j v_i + \partial_i v_j),$$

and the two expressions coincide because

$$(\partial_j u_i + \partial_i u_j)\, \partial_i v_j = (\partial_i u_j + \partial_j u_i)\, \partial_j v_i$$

after swapping the dummy indices $i \leftrightarrow j$ in the left-hand side. Hence,

$$\int_\Omega 2\mu\, \boldsymbol{\varepsilon}(\mathbf{u}) : \nabla \mathbf{v}\, dx = \int_\Omega 2\mu\, \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\, dx.$$

Putting everything together, we conclude that for all $\mathbf{v} \in V_0$,

$$\int_\Omega 2\mu\, \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\, dx - \int_\Omega p\, \nabla \cdot \mathbf{v}\, dx = \int_\Omega \rho\, \mathbf{f} \cdot \mathbf{v}\, dx,$$

that is,

$$a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = F(\mathbf{v}).$$

The right–hand side is well defined, since $\mathbf{f} \in L^2(\Omega)^2$ and $\mathbf{v} \in V_0 \subset H^1(\Omega)^2 \subset L^2(\Omega)^2$, so that $\mathbf{f} \cdot \mathbf{v} \in L^1(\Omega)$.

3. Weak incompressibility constraint

Since $(\mathbf{u}, p)$ is a strong solution, we have

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{in } \Omega.$$

Let $q \in Q$. Because $\mathbf{u} \in H^1(\Omega)^2$, its divergence $\nabla \cdot \mathbf{u}$ belongs to $L^2(\Omega)$. Hence the integral

$$\int_\Omega q\, \nabla \cdot \mathbf{u}\, dx$$

is well defined, as $q \in L^2(\Omega)$.

Using the strong incompressibility condition pointwise in Ω, we obtain

$$b(\mathbf{u}, q) = -\int_\Omega q\, \nabla \cdot \mathbf{u}\, dx = 0 \qquad \forall q \in Q.$$

Combining Steps 1–3, we conclude that $(\mathbf{u}, \widetilde{p}) \in V_{\mathbf{g}} \times Q$ satisfies *($S^{\text{mix}}_{\text{weak}}$)*, and therefore it is a weak solution. □

For the subsequent analysis of well–posedness, since the velocity space $V_{\mathbf{g}}$ is affine rather than linear, it is convenient to reduce the affine formulation *($S^{\text{mix}}_{\text{weak}}$)*, introduced in Definition 2.2.2, to an equivalent homogeneous problem.

By definition of $V_{\mathbf{g}}$, any function $\mathbf{u} \in V_{\mathbf{g}}$ can be written as

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_g, \qquad \mathbf{u}_0 \in V_0,$$

where $\mathbf{u}_g \in V_{\mathbf{g}}$ is a fixed lifting of the Dirichlet data, that is, any function in $H^1(\Omega)^2$ whose trace coincides with $\mathbf{g}$ on Γ_D. The existence of such a lifting is ensured by the regularity assumptions on Ω and on the boundary data $\mathbf{g}$. In particular, since $\mathbf{g}$ is sufficiently regular on Γ_D, it can be extended to a function in $H^1(\Omega)^2$ whose trace coincides with $\mathbf{g}$ on Γ_D. Moreover, the lifting is not unique: if $\mathbf{w} \in V_0$, then $\mathbf{u}_g + \mathbf{w}$ is also an admissible lifting. Different choices of $\mathbf{u}_g$ lead to equivalent homogeneous formulations and do not affect the final solution $\mathbf{u}$. Substituting $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_g$ into *($S^{\text{mix}}_{\text{weak}}$)* yields the equivalent problem:

Definition 2.2.4 (Reduced Stokes variational problem)**.** *Find* $(\mathbf{u}_0, p) \in V_0 \times Q$ *such that*

$$\begin{cases} a(\mathbf{u}_0, \mathbf{v}) + b(\mathbf{v}, p) = F(\mathbf{v}) - a(\mathbf{u}_g, \mathbf{v}), & \forall \mathbf{v} \in V_0, \\ b(\mathbf{u}_0, q) = -b(\mathbf{u}_g, q), & \forall q \in Q. \end{cases} \tag{$\mathrm{S}^0_{\text{weak}}$}$$

For the precise definitions of the spaces V_0 *and* Q*, as well as of the bilinear forms* $a(\cdot,\cdot)$ *and* $b(\cdot,\cdot)$ *and the linear functional* $F(\cdot)$*, we refer to Definition 2.2.2.*

Chapter 3

Finite Element Method

In this chapter, the finite element method (FEM) is presented from a rigorous theoretical foundation. It begins by defining a finite element as the abstract triple (K, P, N) (element domain, shape function space, and degrees of freedom), establishing the key unisolvence property that guarantees unique reconstruction from nodal data, and constructing the associated nodal (shape-function) basis together with its matrix representation. These tools are then used to define the local finite element interpolant I_K, proving its linearity, nodal reproduction, and projection (idempotence) properties, and to extend the construction to a global setting. The chapter further illustrates the abstract framework through concrete triangular Lagrange elements (P_1 and P_2), the reference triangle and affine mappings, and briefly discusses other important element families. Finally, these ingredients are applied to a Galerkin discretization of the Stokes problem using the Taylor–Hood P_2/P_1 pair and outlining the elementwise computation and assembly of matrix entries via reference-element transformations and Gauss quadrature.

3.1 Formal definition of finite element

In this section, a rigorous and formal mathematical definition of a finite element is presented following the classical abstract framework introduced by Ciarlet [4] and further developed in standard references such as [1]. This definition provides a unified theoretical foundation for the finite element method, allowing us to identify precisely the mathematical objects involved and the structural properties they must satisfy.

Definition 3.1.1 (Finite Element). *A finite element is determined by the triple (K, P, N), where*

- *$K \subseteq \mathbb{R}^n$ is a closed and bounded set with nonempty interior and a piecewise smooth boundary. K is called the element domain.*
- *P is a finite-dimensional vector subspace of the ambient function space*

  $$\mathcal{F}(K) := \{f \mid f : K \to \mathbb{R}\},$$

  considered as a vector space over $\mathbb{R}$. P is called the shape function space.

- $N = \{N_i\}_{i=1}^k$ *is a basis of the dual space* P'*, called the set of degrees of freedom.*

The previous definition is deliberately general. No specific assumptions are made regarding the geometry of the set K, the precise nature of the function space P, or the form of the linear functionals defining the degrees of freedom. This level of abstraction is essential, as it allows a wide variety of finite elements used in practice to be treated within a unified theoretical framework.

In typical applications, K in 1D is usually a segment $[a, b]$, in 2D a triangle or a quadrilateral, and in 3D a tetrahedron or another polyhedral cell. The space P, consisting of functions $p : K \to \mathbb{R}$, is in practice usually chosen as a polynomial space, due to its simplicity and good properties, although other variants also exist. The functionals in N, $N_i : P \to \mathbb{R}$, in most cases correspond to evaluations at certain key points of K, called nodes, or to flux integrals over parts of the boundary ∂K, although here too a wide variety of choices exists.

Once the three fundamental components of a finite element have been identified, the next natural step is to understand the purpose of this construction. In the context of the finite element method, the finite element itself is not an end, but a tool that enables the definition of approximation operators mapping functions from infinite-dimensional spaces into computable finite-dimensional subspaces.

The ultimate goal of this definition is to construct an interpolant, that is, a mapping between a suitable function space and its finite-dimensional subspace P. As will be seen later in this section, this mapping is a projector and provides a fundamental approximation tool in the FEM.

In order to construct the interpolant, every element $p \in P$ must be uniquely determined by the values $\{N_i(p)\}_{i=1}^k \subset \mathbb{R}$. This property is guaranteed by the fact that N is a basis of P'. See Proposition 3.1.2. That proposition also presents a practical criterion for determining when a subset of P' is indeed a basis, since this verification is the most laborious of the three checks required to verify that a triple (K, P, N) defines a finite element.

Proposition 3.1.2 (Unisolvence Characterization of Dual Bases)**.** *Let P be a finite-dimensional vector space over $\mathbb{R}$ of dimension d, and let $N = \{N_1, \ldots, N_d\}$ be a subset of P'. The following statements are equivalent:*

1. *N is a basis of P'.*
2. ***(Unisolvence)** Let $p \in P$. If $N_i(p) = 0$ for all $i \in \{1, \ldots, d\}$, then $p = 0_P$.*
3. *The linear mapping between $\mathbb{R}$-vector spaces*

$$J : P \longrightarrow \mathbb{R}^d, \qquad J(p) := \big(N_1(p), \ldots, N_d(p)\big),$$

is an isomorphism.

Proof. (1)⇒(2). Assume that N is a basis of P'. Let $p \in P$ such that $N_i(p) = 0$ for all $i \in \{1, \ldots, d\}$. Since N spans P', for every $F \in P'$ there exist coefficients $\alpha_i \in \mathbb{R}$ such that

$F = \sum_{i=1}^{d} \alpha_i N_i$. Evaluating at p yields

$$F(p) = \sum_{i=1}^{d} \alpha_i N_i(p) = 0,$$

hence $F(p) = 0$ for all $F \in P'$. By contradiction, suppose $p \neq 0_P$. By Theorem A.1.7, the set $\{p\}$ can be extended to a basis $\{p, p_2, \ldots, p_d\}$ of P. Let $\{\Phi_1, \ldots, \Phi_d\} \subset P'$ be the associated dual basis, whose existence and definition follows from Theorem A.1.6, characterized by $\Phi_i(p_j) = \delta_{ij}$. By definition, $\Phi_1(p) = 1 \neq 0$, which is a contradiction. Therefore, necessarily $p = 0_P$, and unisolvence is proved.

(2)⇒(3). The mapping J is linear, since each $N_i \in P'$ is linear and the coordinate concatenation preserves linearity; hence J is a linear homomorphism. Since $\dim P = \dim \mathbb{R}^d = d$, it suffices to prove that J is injective. This follows directly from the unisolvence hypothesis: if $p \in \ker J$, then $J(p) = 0_{\mathbb{R}^d}$ and, by unisolvence, $p = 0_P$. Therefore, $\ker J = \{0\}$.

(3)⇒(1). Since $J : P \to \mathbb{R}^d$ is a linear isomorphism, by Theorem A.1.9 its pullback

$$J^* : (\mathbb{R}^d)' \longrightarrow P', \qquad J^*(\varphi) := \varphi \circ J,$$

is also a linear isomorphism. Let $\{e_1^*, \ldots, e_d^*\}$ be the canonical basis of $(\mathbb{R}^d)'$, defined by $e_i^*(x_1, \ldots, x_d) = x_i$. Then, for $p \in P$,

$$\big(J^*(e_i^*)\big)(p) = e_i^*(J(p)) = e_i^*\big(N_1(p), \ldots, N_d(p)\big) = N_i(p),$$

hence $J^*(e_i^*) = N_i$ as a map for all i. Therefore, $\{N_1, \ldots, N_d\} = J^*(\{e_1^*, \ldots, e_d^*\})$ is the image of a basis under an isomorphism, and hence, by Theorem A.1.4, is a basis of P'. □

Beyond guaranteeing uniqueness of the nodal representation, the unisolvence property also allows the explicit construction of a basis of the space P adapted to the degrees of freedom. This basis is designed to interpolate nodal values exactly and can be regarded as the functional analogue of the canonical basis in finite-dimensional vector spaces.

Definition 3.1.3 (Nodal Basis)**.** *Let (K, P, N) be a finite element, and* $\dim P = d$. *The nodal basis of P associated with the element is a d-tuple $\Phi = (\phi_1, \ldots, \phi_d) \subset P$ such that*

$$N_i(\phi_j) = \delta_{ij} \qquad \forall\ i, j = 1, \ldots, d.$$

Remark 3.1.4. *The functions $\{\phi_1, \ldots, \phi_d\}$ of the nodal basis introduced in Definition 3.1.3 are commonly referred to as shape functions in the finite element literature. This terminology reflects their role in describing the local "shape" of the finite element approximation within the element domain K.*

Although the definition of the nodal basis is conceptually natural, it is not immediately obvious that such a collection of functions forms a basis of the space P. This fact must be established rigorously and relies fundamentally on the unisolvence property proved earlier.

Proposition 3.1.5 (Nodal Basis)**.** *Let (K, P, N) be a finite element with* $\dim P = d$. *Then the nodal basis defined in Definition 3.1.3 is indeed a basis of P.*

Proof. By Definition 3.1.1, N is a basis of P', and by Proposition 3.1.2, the linear map

$$J : P \longrightarrow \mathbb{R}^d, \qquad J(p) = (N_1(p), \ldots, N_d(p)),$$

is an isomorphism. By the definition of the nodal basis, the image of $\Phi = (\phi_1, \ldots, \phi_d) \subset P$ under J, that is $J(\Phi)$, coincides with the canonical basis of $\mathbb{R}^d$. Therefore, based on Theorem A.1.4 Φ is a basis of P. □

From a practical point of view, it is not sufficient to know the abstract existence of the nodal basis; one must also understand its specific representation. In particular, for both analysis and implementation purposes, it is essential to express each nodal basis function in terms of a fixed and explicitly known basis of the space P, typically the canonical basis. This link provides a direct connection between the abstract definition of degrees of freedom and their specific realization as shape functions. The following result formalizes this construction by showing how the coefficients of the nodal basis functions arise as the solution of a linear system.

Proposition 3.1.6 (Matrix Representation of the Nodal Basis)**.** *Let (K, P, N) be a finite element with $\dim P = d$, and let $B = \{b_1, \ldots, b_d\}$ be the canonical basis of P. Let $\Phi = \{\phi_1, \ldots, \phi_d\}$ be the nodal basis associated with $N = \{N_1, \ldots, N_d\} \subset P'$. Then, for each $j \in \{1, \ldots, d\}$, the coordinate vector of ϕ_j in the canonical basis of P, denoted by $(\phi_j)_B \in \mathbb{R}^d$, is obtained as the solution of the linear system*

$$M(\phi_j)_B = e_j,$$

where e_j is the j-th vector of the canonical basis of $\mathbb{R}^d$ and

$$M_{ij} := N_i(b_j).$$

In particular, if $C = [(\phi_1)_B \ \cdots \ (\phi_d)_B]$, then $C = M^{-1}$.

Proof. By definition of the nodal basis, $N_i(\phi_j) = \delta_{ij}$ for all i, j. Write ϕ_j in the canonical basis B as

$$\phi_j = \sum_{\ell=1}^{d} c_\ell^{(j)} b_\ell, \qquad (\phi_j)_B = \left(c_1^{(j)}, \ldots, c_d^{(j)}\right)^\top.$$

Evaluating N_i and using linearity, we obtain

$$\delta_{ij} = N_i(\phi_j) = \sum_{\ell=1}^{d} c_\ell^{(j)} N_i(b_\ell) = \sum_{\ell=1}^{d} M_{i\ell}\, c_\ell^{(j)}.$$

This is precisely the i-th equation of the linear system

$$M(\phi_j)_B = e_j.$$

The matrix M is invertible because it is the matrix (with respect to the bases B and the canonical basis of $\mathbb{R}^d$) of the isomorphism J from Proposition 3.1.2. Therefore, the system admits a unique solution.

The columns of C are exactly the desired coordinate vectors, and hence $C = M^{-1}$. □

Up to this point, the degrees of freedom have been defined only on the space P. However, in order for the interpolant to act as an approximation operator, it is necessary to extend the action of these functionals to a larger class of functions defined on the element domain. This motivates the introduction of a functional space in which all nodal variables are well defined, without requiring the function itself to belong to the finite-dimensional space P.

Definition 3.1.7 (Admissible Function Space for Nodal Variables). *Let (K, P, N) be a finite element with $\dim P = d$. The set of functions for which the nodal variables are well defined is given by*

$$A_K := \{ v : K \to \mathbb{R} \mid N_i(v) \in \mathbb{R} \text{ for all } i \in \{1, \dots, d\} \}.$$

At this point, all the necessary tools are available to construct the finite element interpolant.

Definition 3.1.8 (Local Finite Element Interpolant). *Let (K, P, N) be a finite element with $\dim P = d$. The local interpolant is defined as*

$$I_K : A_K \longrightarrow P, \qquad I_K(v) := \sum_{i=1}^{d} N_i(v)\, \phi_i \in P.$$

Proposition 3.1.9 (Properties of the Local Interpolant). *Let (K, P, N) be a finite element. The local interpolant defined above satisfies:*

1. ***Linearity.***

$$I_K(\alpha u + \beta v) = \alpha I_K u + \beta I_K v \quad \text{for all } u, v \in A_K \text{ and } \alpha, \beta \in \mathbb{R}.$$

2. ***Reproduction of nodal values.*** *For every $v \in A_K$ and every i,*

$$N_i(I_K(v)) = N_i(v).$$

In particular, $I_K v$ is the unique function in P that shares with v the same nodal values $\{N_i(v)\}_{i=1}^d$.

3. ***Idempotence (projection property).***

$$I_K^2 = I_K.$$

In particular, the restriction of I_K to P is the identity: if $p \in P$, then $I_K p = p$.

Proof. **(1)** Let $u, v \in A_K$ and $\alpha, \beta \in \mathbb{R}$. Using the linearity of each N_j,

$$I_K(\alpha u + \beta v) = \sum_{j=1}^{d} N_j(\alpha u + \beta v)\, \phi_j = \sum_{j=1}^{d} \big(\alpha N_j(u) + \beta N_j(v)\big)\phi_j = \alpha I_K u + \beta I_K v.$$

(2) For $v \in A_K$ and any $i \in \{1, \dots, d\}$,

$$N_i(I_K v) = N_i\Big(\sum_{j=1}^{d} N_j(v)\phi_j\Big) = \sum_{j=1}^{d} N_j(v)\, N_i(\phi_j) = \sum_{j=1}^{d} N_j(v)\, \delta_{ij} = N_i(v).$$

Uniqueness follows from unisolvence: if $p \in P$ satisfies $N_i(p) = N_i(v)$ for all i, then $N_i(p - I_K v) = 0$ for all i, and by part (2) of Proposition 3.1.2 we conclude $p - I_K v = 0$, i.e. $p = I_K v$.

(3) By part (2), for $v \in A_K$ and all i,

$$N_i\big(I_K(I_K v)\big) = N_i(I_K v) = N_i(v).$$

By unisolvence again, $I_K(I_K v) = I_K v$, that is, $I_K^2 = I_K$. In particular, if $p \in P$, then $N_i(I_K p) = N_i(p)$ for all i; by unisolvence, $I_K p = p$. □

The local interpolant is one of the central objects of the finite element method, since it provides the connection between the continuous model and the discrete approximation. That is, it takes a function from the space containing the continuous solution of a differential problem with boundary conditions and maps it into the finite-dimensional subspace P, while exactly preserving the degrees of freedom. In this way, a computable representation is obtained that retains the essential information and allows for the local assembly of matrices and vectors, leading, as will be discussed later, to a global linear system.

Moreover, its linear and idempotent nature makes it a stable projector, which is a key ingredient in error analysis. From approximation estimates of the form $\|v - I_K v\|$, a priori bounds for the method are derived (via, e.g., Céa's lemma), and the choice of discrete spaces is designed to ensure stability and consistency. In summary, all the effort devoted to constructing I_K—and to guaranteeing unisolvence, stability, and good approximation properties—results in a reliable bridge between the continuous problem and its effective numerical solution.

3.2 Geometric setting and global approximation

In the previous section, the notion of a finite element was introduced in an abstract and local setting, together with the associated degrees of freedom, nodal basis functions, and local interpolation operator. These constructions were defined on a single element and provide the building blocks of the finite element method. In order to obtain a global approximation over a physical domain, it is now necessary to specify the geometry of the domain under consideration, its discretization into finite elements, and the way in which local interpolants are combined to define a global interpolation operator.

Definition 3.2.1 (Physical domain). *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with boundary $\partial\Omega$. The set Ω is called the physical domain of the problem. Functions of interest, such as solutions or data, are defined on Ω, and boundary conditions are prescribed on $\partial\Omega$.*

Definition 3.2.2 (Subdivision of the physical domain). *Let $\Omega \subset \mathbb{R}^n$ be the physical domain introduced in Definition 3.2.1. A subdivision of Ω is a finite collection of subsets of $\mathbb{R}^n$*

$$\mathcal{T}_h = \{K_i \subset \mathbb{R}^n\}_{i \in I},$$

where the index set I is finite, such that:

1. *each $K_i \in \mathcal{T}_h$ is a closed set with nonempty interior, called an element;*

2. *the union of the elements covers the closure of the domain,*

$$\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K;$$

3. *for any two distinct elements $K_i, K_j \in \mathcal{T}_h$, the intersection $K_i \cap K_j$ is either empty or contained in the intersection of their boundaries,*

$$K_i \cap K_j \subset \partial K_i \cap \partial K_j.$$

Once the physical domain has been discretized into a subdivision, the local constructions introduced in the previous section can be extended to the whole domain. In particular, each element $K \in \mathcal{T}_h$ is endowed with the structure of a finite element, with its own local space of shape functions, degrees of freedom, and associated local interpolant. The remaining task is to combine these local interpolation operators in a coherent way in order to define a single approximation operator acting on functions defined over the entire domain Ω. This leads naturally to the definition of the global finite element interpolant.

Definition 3.2.3 (Global finite element interpolant)**.** *Let $\mathcal{T}_h$ be a subdivision of the physical domain Ω in the sense of Definition 3.2.2. Assume that each element $K \in \mathcal{T}_h$ is endowed with the structure of a finite element (K, P, N), and let $I_K : A_K \to P$ denote the associated local interpolant defined in Definition 3.1.8.*

The global finite element interpolant is the operator

$$I_h : \{\, v : \Omega \to \mathbb{R} \mid v|_K \in A_K \text{ for all } K \in \mathcal{T}_h \,\} \longrightarrow \{\, w : \Omega \to \mathbb{R} \,\}$$

defined by

$$(I_h v)|_K := I_K(v|_K), \qquad \forall K \in \mathcal{T}_h.$$

By construction, the natural domain of the global interpolant I_h is the set of functions for which all local degrees of freedom are well defined. Consequently, the effective domain of I_h is determined solely by the nature of the degrees of freedom of the chosen finite element. Moreover, and this is critical, the image of the global interpolant satisfies

$$\mathrm{Im}(I_h) = V_h,$$

where V_h denotes the global finite element space associated with the mesh $\mathcal{T}_h$, which is required to be a finite-dimensional subspace of the functional space in which the variational problem to be approximated is posed. Therefore, it is the global interpolator—determined by the domain subdivision and by the chosen finite element structure for every subdivision element—that builds the conforming Galerkin approximation to a fixed variational problem. This observation captures one of the central ideas of the finite element method under the Galerkin perspective.

3.3 Specific finite elements used

In this section, several specific examples of finite elements are presented in order to illustrate the abstract theory introduced previously. The focus is placed first on the type of elements that will be used in the applications discussed later in this work, while other relevant families of finite elements are briefly mentioned afterwards.

From this point on, we restrict our attention to two-dimensional finite elements defined on triangular domains. More precisely, the element domain $K \subset \mathbb{R}^2$ is always assumed to be a closed triangular region in $\mathbb{R}^2$, so it has nonempty interior and piecewise smooth boundary. Consequently, K satisfies all the geometric requirements of the general definition of a finite element. The space P will be a finite-dimensional polynomial space defined on K, whose dimension depends on the degree of the polynomials considered. Finally, the set of degrees of freedom N will consist of either evaluations at convenient points or suitably chosen integral functionals over edges, depending on the specific element under consideration. Further details are given in each specific example.

3.3.1 Common reference triangle and conforming triangular mesh

As detailed in the previous sections, the construction of the finite element method requires working with a collection of finite elements that discretize the physical domain under consideration, commonly referred to as a mesh. Although all elements in the mesh share the same finite element type, it is neither convenient nor necessary to treat each physical element independently.

A fundamental idea of the finite element method is to define all finite element objects on a single, fixed domain—called the reference element and then transfer these objects to each physical element of the mesh by means of suitable bijective mappings.

Definition 3.3.1 (Reference Triangle)**.** *In the two-dimensional case considered here, and under the choice of triangular elements as the basic geometric entities, the reference element is defined as the triangular domain $\widehat{K} \subset \mathbb{R}^2$ with vertices*

$$\widehat{\mathbf{v}}_1 = (0,0), \qquad \widehat{\mathbf{v}}_2 = (1,0), \qquad \widehat{\mathbf{v}}_3 = (0,1).$$

This domain is called the reference (or master) triangle.

Definition 3.3.2 (Affine Mapping from the Reference Triangle)**.** *Let $K \subset \mathbb{R}^2$ be a triangular element with vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ that are not collinear. The affine mapping from the reference triangle $\widehat{K}$ onto the physical element K is the function*

$$F_K : \widehat{K} \longrightarrow K,$$

defined by

$$F_K(\widehat{\mathbf{x}}) = \mathbf{v}_1 + B_K \widehat{\mathbf{x}},$$

where the matrix $B_K \in \mathbb{R}^{2\times 2}$ is given by

$$B_K = \begin{pmatrix} \mathbf{v}_2 - \mathbf{v}_1 & \mathbf{v}_3 - \mathbf{v}_1 \end{pmatrix}.$$

Since the vertices of K are not collinear, the matrix B_K is invertible. Therefore, the mapping F_K is bijective and admits an affine inverse $F_K^{-1} : K \to \widehat{K}$.

The affine mapping F_K provides a systematic way to transport functions, basis functions, and degrees of freedom between the reference element and any physical element of the mesh. In particular, shape functions defined on K are obtained as compositions of reference shape functions on $\widehat{K}$ with the inverse mapping F_K^{-1}, a process commonly referred to as the pullback functions.

The reference triangle (Definition 3.3.1) and the affine mapping (Definition 3.3.2) play a fundamental role in the theoretical analysis of the finite element method. Key properties such as unisolvence, approximation estimates and stability bounds are first established on the reference element and then transferred uniformly to all physical elements through the affine mapping.

It is also necessary to specify how the physical domain Ω is decomposed into a finite collection of triangular elements to which the reference element can be mapped. This leads to the notion of a conforming triangular mesh, which provides the geometric framework required to assemble local finite element spaces and interpolants into globally well-defined and conforming finite element spaces.

Definition 3.3.3 (Conforming triangular mesh)**.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain. A* conforming triangular mesh *of Ω is a subdivision $\mathcal{T}_h$ of $\overline{\Omega}$ in the sense of Definition 3.2.2 such that all elements $K \in \mathcal{T}_h$ are nondegenerate triangles and, in addition, the following conformity condition holds:*

(4) For any two distinct elements $K, K' \in \mathcal{T}_h$, the intersection $K \cap K'$ is either empty, a common vertex, or a common edge. In particular, no vertex of any triangle lies in the interior of an edge of another triangle.

The mesh size *of the triangulation is defined by*

$$h := \max_{K \in \mathcal{T}_h} \operatorname{diam}(K).$$

3.3.2 Lagrange P_1 finite element

We begin with the simplest and most widely used finite element: the linear Lagrange element defined on triangular domains.

Definition 3.3.4 (Lagrange P_1 Finite Element)**.** *Let $K \subset \mathbb{R}^2$ be a triangular domain with vertices*

$$\mathbf{v}_1 = (x_1, y_1), \quad \mathbf{v}_2 = (x_2, y_2), \quad \mathbf{v}_3 = (x_3, y_3),$$

which are not collinear. The Lagrange P_1 finite element on K is defined by the triple (K, P, N), where

- *the polynomial space is*

$$P := \mathbb{P}_1(K) = \{\, p(x, y) = \alpha + \beta x + \gamma y \mid \alpha, \beta, \gamma \in \mathbb{R},\ (x, y) \in K \,\},$$

- *and the set of degrees of freedom is*

$$N = \{N_1, N_2, N_3\}, \qquad N_i(p) := p(x_i, y_i), \quad i = 1, 2, 3.$$

It is immediate that $\dim P = 3$.

Proposition 3.3.5 (Unisolvence of the Lagrange P_1 Element)**.** *The degrees of freedom $N = \{N_1, N_2, N_3\}$ defined in Definition 3.3.4 form a basis of the dual space P' and hence the triple defined there, is effectely a finite element.*

Proof. Since $\dim P = 3$, by Proposition 3.1.2 it suffices to verify that the only function $p \in P$ satisfying

$$p(x_i, y_i) = 0 \quad \text{for } i = 1, 2, 3$$

is the zero function.

Let $p(x, y) = \alpha + \beta x + \gamma y$. The above conditions lead to the linear system

$$\begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The determinant of the coefficient matrix is equal (up to a multiplicative constant) to twice the area of the triangle K. Since the vertices of K are not collinear, the triangle has nonempty interior and this determinant is nonzero. Therefore, the only solution is the trivial one, and the unisolvence property holds.

Consequently, N is a basis of P' and (K, P, N) defines a finite element for any nondegenerate triangular domain K. □

Proposition 3.3.6 (Nodal basis of the P_1 Lagrange element on the reference triangle)**.** *Let $\widehat{K}$ be the reference triangle introduced in Definition 3.3.1 and let $P = \mathbb{P}_1(\widehat{K})$. The nodal basis of P in the sense of Definition 3.1.3 is given by*

$$\widehat{\phi}_1(\xi, \eta) = 1 - \xi - \eta, \qquad \widehat{\phi}_2(\xi, \eta) = \xi, \qquad \widehat{\phi}_3(\xi, \eta) = \eta, \qquad (\xi, \eta) \in \widehat{K}.$$

Proof. The functions $\widehat{\phi}_1, \widehat{\phi}_2, \widehat{\phi}_3$ belong to $\mathbb{P}_1(\widehat{K})$ and satisfy

$$\widehat{\phi}_j(\widehat{\mathbf{v}}_i) = \delta_{ij}, \qquad i, j = 1, 2, 3,$$

where $\{\widehat{\mathbf{v}}_i\}_{i=1}^3$ are the vertices of $\widehat{K}$. Therefore

$$\widehat{N}_i(\widehat{\phi}_j) = \delta_{ij}, \qquad i, j = 1, 2, 3,$$

where $\widehat{N}_i$ denotes the nodal functional associated with the vertex $\widehat{\mathbf{v}}_i$. Therefore, the functions $\widehat{\phi}_1, \widehat{\phi}_2, \widehat{\phi}_3$ satisfy the defining conditions of the nodal basis given in Definition 3.1.3. By unisolvence, the nodal basis is unique, and the result follows. □

Proposition 3.3.7 (Nodal basis of the P_1 Lagrange element on a physical triangle)**.** *Let K be a physical triangle and let $F_K : \widehat{K} \to K$ be the affine mapping introduced in Definition 3.3.2. Let $\{\widehat{\phi}_1, \widehat{\phi}_2, \widehat{\phi}_3\}$ be the nodal basis of $\mathbb{P}_1(\widehat{K})$ given in Proposition 3.3.6. Then the nodal basis $\{\phi_1, \phi_2, \phi_3\}$ of $\mathbb{P}_1(K)$ in the sense of Definition 3.1.3 is given by*

$$\phi_i := \widehat{\phi}_i \circ F_K^{-1}, \qquad i = 1, 2, 3.$$

Proof. Since F_K is bijective and affine, each function ϕ_i belongs to $\mathbb{P}_1(K)$. Moreover, for each vertex $\mathbf{v}_j$ of K,

$$N_j(\phi_i) = \phi_i(\mathbf{v}_j) = \widehat{\phi}_i\big(F_K^{-1}(\mathbf{v}_j)\big) = \widehat{\phi}_i(\widehat{\mathbf{v}}_j) = \delta_{ij}.$$

Therefore, $\{\phi_1, \phi_2, \phi_3\}$ satisfies the defining conditions of the nodal basis given in Definition 3.1.3. By unisolvence, the nodal basis on K is unique. □

Definition 3.3.8 (Local interpolant for the Lagrange P_1 element)**.** *Let $(K, \mathbb{P}_1(K), \{N_i\}_{i=1}^3)$ be the Lagrange P_1 finite element from Definition 3.3.4, and let $\{\phi_1, \phi_2, \phi_3\}$ be its nodal basis given in Proposition 3.3.7. Consider the admissible space A_K of Definition 3.1.7. Then the associated local finite element interpolant*

$$I_K : A_K \longrightarrow \mathbb{P}_1(K)$$

is given, for every $v \in A_K$, by

$$I_K(v) := \sum_{i=1}^{3} N_i(v)\, \phi_i, \qquad \text{that is,} \qquad I_K(v)(x) = \sum_{i=1}^{3} v(\mathbf{v}_i)\, \phi_i(x) \quad \forall x \in K.$$

The local interpolant I_K associates to a generic function v defined on the element K a unique affine function in $\mathbb{P}_1(K)$ that coincides with v at the vertices of K. In other words, $I_K v$ is the piecewise linear approximation of v over K obtained by prescribing the nodal values $v(\mathbf{v}_i)$ and extending them linearly inside the element through the nodal basis functions. This construction replaces the original function by its simplest approximation on K, while preserving the nodal values and providing a locally linear representation.

3.3.3 Lagrange P_2 finite element

We now introduce the quadratic Lagrange finite element on triangular domains, which provides a higher–order approximation while remaining C^0–conforming.

Definition 3.3.9 (Lagrange P_2 Finite Element)**.** *Let $K \subset \mathbb{R}^2$ be a nondegenerate triangular domain with vertices $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and let $\mathbf{m}_{12}, \mathbf{m}_{23}, \mathbf{m}_{31}$ denote the midpoints of its edges. The Lagrange P_2 finite element on K is defined by the triple (K, P, N), where*

- *the polynomial space is*

  $$P := \mathbb{P}_2(K) = \left\{\, p(x,y) = \alpha + \beta x + \gamma y + \delta x^2 + \varepsilon xy + \zeta y^2 \;\middle|\; \alpha, \beta, \gamma, \delta, \varepsilon, \zeta \in \mathbb{R},\ (x,y) \in K \,\right\},$$

- *and the set of degrees of freedom is*

  $$N = \{N_1, \dots, N_6\}, \qquad N_i(p) := p(\mathbf{a}_i), \quad i = 1, \dots, 6,$$

  where the nodal points are given by

  $$\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}, \qquad \{\mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_6\} = \{\mathbf{m}_{12}, \mathbf{m}_{23}, \mathbf{m}_{31}\}.$$

It is immediate that $\dim P = 6$.

Proposition 3.3.10 (Unisolvence of the Lagrange P_2 Element)**.** *The degrees of freedom $N = \{N_i\}_{i=1}^6$ defined in Definition 3.3.9 form a basis of the dual space P'. Consequently, the triple (K, P, N) introduced there is consistent with the definition of a finite element.*

Proof. We first particularize the argument to the reference triangle $\widehat{K}$ introduced in Definition 3.3.1, and we will later extend it to a generic physical triangle $K \subset \mathbb{R}^2$.

Let $\widehat{p} \in \mathbb{P}_2(\widehat{K})$ satisfy $\widehat{N}_i(\widehat{p}) = 0$ for all $i = 1, \ldots, 6$, where the nodal points are the three vertices and the three edge midpoints of $\widehat{K}$. Writing $\widehat{p}$ in the monomial basis,

$$\widehat{p}(\xi, \eta) = \alpha + \beta\,\xi + \gamma\,\eta + \delta\,\xi^2 + \varepsilon\,\xi\eta + \zeta\,\eta^2,$$

the nodal conditions yield the homogeneous linear system

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0.5 & 0 & 0.25 & 0 & 0 \\ 1 & 0.5 & 0.5 & 0.25 & 0.25 & 0.25 \\ 1 & 0 & 0.5 & 0 & 0 & 0.25 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The determinant of this matrix is $\det = \frac{1}{64} \neq 0$, hence the only solution is $\alpha = \beta = \gamma = \delta = \varepsilon = \zeta = 0$, i.e. $\widehat{p} \equiv 0$ on $\widehat{K}$. Therefore, the unisolvence property holds on the reference triangle $\widehat{K}$ and based on Proposition 3.1.2, this set of degrees of freedom is a basis of P'.

To show that the unisolvence property also holds on a generic physical triangle $K \subset \mathbb{R}^2$, we make use of item (iii) in Proposition 3.1.2.

To this end, we define the linear mapping

$$E_K : \mathbb{P}_2(K) \longrightarrow \mathbb{R}^6, \qquad E_K(p) := \big(p(\mathbf{v}_1),\, p(\mathbf{v}_2),\, p(\mathbf{v}_3),\, p(\mathbf{m}_{12}),\, p(\mathbf{m}_{23}),\, p(\mathbf{m}_{31})\big),$$

which, in the notation of Proposition 3.1.2, corresponds to the operator J associated with the set of degrees of freedom N. The goal is to show that E_K is an isomorphism.

Let $F_K : \widehat{K} \to K$ be the affine isomorphism introduced in Definition 3.3.2. This mapping induces an affine pullback operator on polynomials,

$$\Phi : \mathbb{P}_2(\widehat{K}) \longrightarrow \mathbb{P}_2(K), \qquad \widehat{p} \longmapsto \widehat{p} \circ F_K^{-1}.$$

The operator Φ is an isomorphism, trivially linear with explicit inverse given by

$$\Phi^{-1}(p) = p \circ F_K.$$

We now justify the identity

$$E_K = E_{\widehat{K}} \circ \Phi^{-1}$$

in a componentwise manner. Let $p \in \mathbb{P}_2(K)$ be arbitrary, by definition of $E_{\widehat{K}}$,

$$E_{\widehat{K}} \circ \Phi^{-1}(p) = E_{\widehat{K}}(p \circ F_K) = \big((p \circ F_K)(\widehat{\mathbf{a}}_1),\, \ldots,\, (p \circ F_K)(\widehat{\mathbf{a}}_6)\big),$$

where $\{\widehat{\mathbf{a}}_i\}_{i=1}^6$ are the corresponding nodal points of the reference triangle $\widehat{K}$.

Since the affine mapping $F_K : \widehat{K} \to K$ maps vertices to vertices and edge midpoints to edge midpoints, we have

$$F_K(\widehat{\mathbf{a}}_i) = \mathbf{a}_i, \qquad i = 1, \ldots, 6.$$

Therefore, for each $i = 1, \ldots, 6$,

$$(p \circ F_K)(\widehat{\mathbf{a}}_i) = p\big(F_K(\widehat{\mathbf{a}}_i)\big) = p(\mathbf{a}_i),$$

and hence

$$E_{\widehat{K}}(p \circ F_K) = E_K(p).$$

This proves the claimed identity. Since $E_{\widehat{K}}$ is an isomorphism—by the unisolvence result established above on the reference triangle $\widehat{K}$—and Φ^{-1} is also an isomorphism, it follows that E_K is an isomorphism as the composition of two isomorphisms.

Therefore, the unisolvence property holds on any physical triangle K, which completes the proof. □

Proposition 3.3.11 (Nodal basis of the P_2 Lagrange element on the reference triangle)**.** *Let $\widehat{K}$ be the reference triangle introduced in Definition 3.3.1 and let $P = \mathbb{P}_2(\widehat{K})$. The nodal basis of P in the sense of Definition 3.1.3 is given by Proposition 3.1.6 and it can be written as*

$$\widehat{\phi}_1(\xi,\eta) = \lambda_1(2\lambda_1 - 1), \qquad \widehat{\phi}_2(\xi,\eta) = \lambda_2(2\lambda_2 - 1), \qquad \widehat{\phi}_3(\xi,\eta) = \lambda_3(2\lambda_3 - 1),$$

$$\widehat{\phi}_4(\xi,\eta) = 4\lambda_1\lambda_2, \qquad \widehat{\phi}_5(\xi,\eta) = 4\lambda_2\lambda_3, \qquad \widehat{\phi}_6(\xi,\eta) = 4\lambda_3\lambda_1,$$

where $\lambda_1, \lambda_2, \lambda_3$ denote the barycentric coordinates on $\widehat{K}$, given by

$$\lambda_1(\xi,\eta) = 1 - \xi - \eta, \qquad \lambda_2(\xi,\eta) = \xi, \qquad \lambda_3(\xi,\eta) = \eta.$$

Proof. By construction, the functions $\widehat{\phi}_1, \ldots, \widehat{\phi}_6$ are polynomials of degree at most 2, hence $\widehat{\phi}_i \in \mathbb{P}_2(\widehat{K})$ for all $i = 1, \ldots, 6$. Moreover, using the barycentric representation and the definition of the nodal points (vertices and edge midpoints), one checks that

$$\widehat{\phi}_j(\widehat{\mathbf{a}}_i) = \delta_{ij}, \qquad i, j = 1, \ldots, 6,$$

and therefore

$$\widehat{N}_i(\widehat{\phi}_j) = \delta_{ij}, \qquad i, j = 1, \ldots, 6.$$

Thus $\{\widehat{\phi}_1, \ldots, \widehat{\phi}_6\}$ satisfies the defining conditions of the nodal basis given in Definition 3.1.3. By unisolvence, the nodal basis is unique, and the result follows. □

Remark 3.3.12. *The nodal basis of the Lagrange P_2 element on a physical triangle K is obtained by composition with the inverse of the affine mapping F_K, where F_K is the isomorphism introduced in Definition 3.3.2. This construction is completely analogous to the P_1 case discussed in Proposition 3.3.7.*

Definition 3.3.13 (Local interpolant for the Lagrange P_2 element)**.** *Let $(K, \mathbb{P}_2(K), \{N_i\}_{i=1}^6)$ be the Lagrange P_2 finite element from Definition 3.3.9, and let $\{\phi_1, \ldots, \phi_6\}$ be its nodal basis given in Remark 3.3.12. Since the degrees of freedom are point evaluations, one typically takes $A_K = C^0(K)$. The associated local finite element interpolant*

$$I_K : C^0(K) \longrightarrow \mathbb{P}_2(K)$$

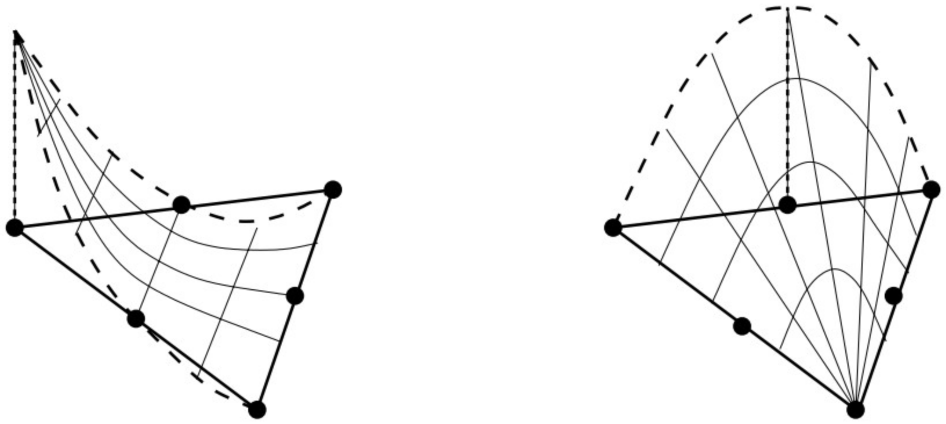


Figure 3.1: Illustration of two nodal basis of the Lagrange P_2 element over a physic triangular element. The nodal values are prescribed at the three vertices and the three edge midpoints, yielding a piecewise quadratic interpolant that is continuous across the element. Taken from [6].

is given, for every $v \in C^0(K)$, by

$$I_K(v) := \sum_{i=1}^{6} N_i(v)\,\phi_i, \qquad \textit{that is,} \qquad I_K(v)(x) = \sum_{i=1}^{6} v(\mathbf{a}_i)\,\phi_i(x) \quad \forall x \in K.$$

The local interpolant I_K associates to a continuous function v a unique quadratic polynomial in $\mathbb{P}_2(K)$ that coincides with v at the vertices and edge midpoints of K. Compared with the P_1 case, this construction provides a higher–order local approximation while preserving the nodal values and the C^0 conformity of the finite element space as we will see below.

3.3.4 Taylor-Hood global interpolant

Definition 3.3.14 (Global Lagrange interpolant)**.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded open polygonal domain and let $\mathcal{T}_h$ be a conforming triangular mesh of Ω in the sense of Definition 3.3.3. For each element $K \in \mathcal{T}_h$ we consider either the P_1 finite element $(K, \mathbb{P}_1(K), N_K^{(1)})$ or the P_2 finite element $(K, \mathbb{P}_2(K), N_K^{(2)})$. In both cases, the associated nodal bases, local interpolants, and related constructions are those introduced in Sections 3.3.2 and 3.3.3.*

For $k \in \{1, 2\}$, let

$$I_K^{(k)} : A_K \longrightarrow \mathbb{P}_k(K)$$

denote the corresponding local Lagrange interpolant, where A_K is the admissible function space for the nodal degrees of freedom (Definition 3.1.7).

The global Lagrange interpolant of degree k is the operator

$$I_h^{(k)} : \mathcal{A}(\Omega) \longrightarrow \{\, w : \Omega \to \mathbb{R} \,\},$$

defined elementwise by

$$(I_h^{(k)}w)|_K := I_K^{(k)}(w|_K), \qquad \forall K \in \mathcal{T}_h, \qquad \forall w \in \mathcal{A}(\Omega),$$

where

$$\mathcal{A}(\Omega) := \left\{ w : \Omega \to \mathbb{R} \;\middle|\; w|_K \in A_K \textit{ for all } K \in \mathcal{T}_h \right\}.$$

Proposition 3.3.15 (Characterization of the image of the global Lagrange interpolant)**.** *Let $k \in \{1,2\}$ and let $I_h^{(k)} : \mathcal{A}(\Omega) \to \{ w : \Omega \to \mathbb{R} \}$ be the global Lagrange interpolant defined in Definition 3.3.14. Then*

$$\operatorname{Im}\Big(I_h^{(k)}\Big) = \left\{ v_h \in C^0(\overline{\Omega}) \;\middle|\; v_h|_K \in \mathbb{P}_k(K) \textit{ for all } K \in \mathcal{T}_h \right\}.$$

Proof. Let $v_h \in \operatorname{Im}(I_h^{(k)})$. Then $v_h = I_h^{(k)}w$ for some $w \in \mathcal{A}(\Omega)$. By the elementwise definition of $I_h^{(k)}$ we have $v_h|_K = I_K^{(k)}(w|_K) \in \mathbb{P}_k(K)$ for every $K \in \mathcal{T}_h$, hence v_h is piecewise polynomial of degree at most k. Moreover, since $\mathcal{T}_h$ is conforming, any two adjacent elements K^+ and K^- share either a full edge or a vertex. Let $E = K^+ \cap K^-$ be an interior edge. By construction of the Lagrange interpolant, the restrictions $v_h|_{K^+}$ and $v_h|_{K^-}$ coincide at all Lagrange nodes lying on E (the two endpoints for $k = 1$, and the endpoints together with the midpoint for $k = 2$). The traces of $v_h|_{K^+}$ and $v_h|_{K^-}$ on E are polynomials of degree at most k in one variable; hence, since two polynomials of degree $\leq k$ that agree at $k + 1$ distinct points must coincide identically, their traces agree at every point of E. Therefore v_h is continuous across every interior edge, and consequently $v_h \in C^0(\overline{\Omega})$. This proves

$$\operatorname{Im}(I_h^{(k)}) \subset \left\{ v_h \in C^0(\overline{\Omega}) \;\middle|\; v_h|_K \in \mathbb{P}_k(K) \text{ for all } K \in \mathcal{T}_h \right\}.$$

Conversely, let $v_h \in \left\{ v_h \in C^0(\overline{\Omega}) \;\middle|\; v_h|_K \in \mathbb{P}_k(K) \text{ for all } K \in \mathcal{T}_h \right\}$ and set $w := v_h$. Then $w \in \mathcal{A}(\Omega)$ and, for each $K \in \mathcal{T}_h$, polynomial reproduction of the local interpolant yields $I_K^{(k)}(w|_K) = w|_K = v_h|_K$ by Proposition 3.1.9. Hence $(I_h^{(k)}w)|_K = v_h|_K$ for all K, and therefore $I_h^{(k)}w = v_h$ in Ω. Thus $v_h \in \operatorname{Im}(I_h^{(k)})$, proving

$$\left\{ v_h \in C^0(\overline{\Omega}) \;\middle|\; v_h|_K \in \mathbb{P}_k(K) \text{ for all } K \in \mathcal{T}_h \right\} \subset \operatorname{Im}(I_h^{(k)}).$$

□

Proposition 3.3.16 (Finite dimensionality and H^1–conformity of the image of the global interpolant)**.** *Let $k \in \{1,2\}$ and let $I_h^{(k)}$ be the global Lagrange interpolant defined in Definition 3.3.14. Then $\operatorname{Im}(I_h^{(k)})$ is a finite–dimensional vector subspace of $H^1(\Omega)$. Moreover,*

$$\dim\big(\operatorname{Im}(I_h^{(k)})\big) = N_h^{(k)},$$

where $N_h^{(k)}$ denotes the total number of global Lagrange nodes of degree k associated with the mesh $\mathcal{T}_h$. To simplify the notation, we write $N_h^{(k)} = d_k$ from now on.

Proof. By Proposition 3.3.15, any $v_h \in \operatorname{Im}(I_h^{(k)})$ satisfies $v_h|_K \in \mathbb{P}_k(K)$ for all $K \in \mathcal{T}_h$ and $v_h \in C^0(\overline{\Omega})$. Hence, for each element $K \in \mathcal{T}_h$, the restriction $v_h|_K$ coincides with a polynomial function on the closed set K, and is therefore classically differentiable on its interior K°, with

$$\nabla(v_h|_{K^\circ}) \in \mathbb{P}_{k-1}(K^\circ)^2.$$

We define a piecewise vector field $\mathbf{g}_h$ by

$$\mathbf{g}_h(x) := \nabla(v_h|_{K^\circ})\,(x) \qquad \text{for } x \in K^\circ,\ K \in \mathcal{T}_h.$$

This definition is unambiguous almost everywhere in Ω, since the union of all element interiors covers Ω up to a set of Lebesgue measure zero. For each $K \in \mathcal{T}_h$, the restriction $\mathbf{g}_h|_{K^\circ}$ is a polynomial and thus belongs to $L^2(K)^2$, which implies $\mathbf{g}_h \in L^2(\Omega)^2$.

To identify $\mathbf{g}_h$ as the weak gradient of v_h, let $\boldsymbol{\varphi} \in C_c^\infty(\Omega)^2$. Since the set

$$\Omega \setminus \bigcup_{K\in\mathcal{T}_h} K^\circ$$

has Lebesgue measure zero, integrals over Ω coincide with integrals over the union of the element interiors. Hence,

$$\int_\Omega v_h\,\nabla\cdot\boldsymbol{\varphi}\,dx = \int_{\bigcup_{K\in\mathcal{T}_h} K^\circ} v_h\,\nabla\cdot\boldsymbol{\varphi}\,dx = -\int_{\bigcup_{K\in\mathcal{T}_h} K^\circ} \mathbf{g}_h\cdot\boldsymbol{\varphi}\,dx = -\int_\Omega \mathbf{g}_h\cdot\boldsymbol{\varphi}\,dx,$$

which shows that $\mathbf{g}_h$ is the weak gradient of v_h, that means the gradient with weak derivatives.

Finally, since $v_h \in C^0(\overline{\Omega})$ and $\overline{\Omega}$ is compact, v_h is bounded and therefore belongs to $L^2(\Omega)$. Together with $\mathbf{g}_h \in L^2(\Omega)^2$, this shows that $v_h \in H^1(\Omega)$, that is,

$$\mathrm{Im}(I_h^{(k)}) \subset H^1(\Omega).$$

$\mathrm{Im}(I_h^{(k)})$ closure under addition and scalar multiplication follows directly from the linearity of $I_h^{(k)}$, see Proposition 3.1.9, hence $\mathrm{Im}(I_h^{(k)})$ is a vector subspace of $H^1(\Omega)$.

To prove finite dimensionality, let $\{a_i\}_{i=1}^{d_k} \subset \mathbb{R}^2$ denote the set of all global Lagrange nodes of degree k, defined as the union of the local Lagrange nodes of degree k associated with each element $K \in \mathcal{T}_h$, where nodes shared by adjacent elements are identified due to the conformity of the mesh. Let $\{\varphi_i\}_{i=1}^{d_k}$ be the corresponding global nodal basis functions, obtained by assembling the local nodal basis functions in such a way that each φ_i coincides locally with the corresponding shape function on the elements containing the node a_i, and vanishes outside their union. See Figure 3.2 for an illustration of the global Lagrange P_1 nodal basis function associated with an interior vertex.

By definition of the global nodal basis, each φ_i satisfies $\varphi_i(a_j) = \delta_{ij}$ and has support contained in the union of the elements sharing the node a_i. In particular, for any fixed point $x \in \Omega$, only a finite number of basis functions φ_i are nonzero at x, namely those associated with the Lagrange nodes of the element $K \in \mathcal{T}_h$ such that $x \in K$. Consequently, the sum

$$v_h(x) = \sum_{i=1}^{d_k} v_h(a_i)\,\varphi_i(x)$$

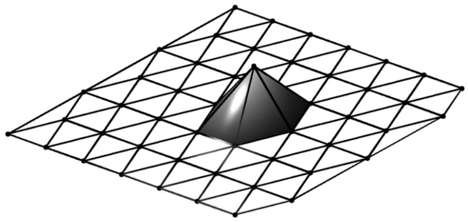


Figure 3.2: Global Lagrange P_1 nodal basis function associated with an interior vertex of a conforming triangular mesh. The function is piecewise affine on each triangle, globally continuous, and supported only on the elements sharing the corresponding node, forming the characteristic "hat" or pyramid shape. Image reproduced from [14].

reduces, for this fixed point x, to a finite sum over the Lagrange nodes of the element K,

$$v_h(x) = \sum_{a_i \in \mathcal{N}(K)} v_h(a_i)\, \varphi_i(x),$$

where $\mathcal{N}(K)$ denotes the set of Lagrange nodes associated with K. In particular, $\#\mathcal{N}(K) = 3$ for P_1 elements and $\#\mathcal{N}(K) = 6$ for P_2 elements. This shows that every function in $\mathrm{Im}(I_h^{(k)})$ admits the global nodal expansion

$$v_h = \sum_{i=1}^{d_k} v_h(a_i)\, \varphi_i.$$

Since the nodal basis functions are linearly independent and span $\mathrm{Im}(I_h^{(k)})$, it follows that

$$\dim\big(\mathrm{Im}(I_h^{(k)})\big) = d_k < \infty.$$

Finally, any finite–dimensional subspace of a Hilbert space is closed. Since $\mathrm{Im}(I_h^{(k)}) \subset H^1(\Omega)$, we conclude that $\mathrm{Im}(I_h^{(k)})$ is a closed subspace of $H^1(\Omega)$ and therefore a Hilbert space in its own right when endowed with the $H^1(\Omega)$ inner product. □

This construction clarifies the role of the global interpolant: it provides a canonical mapping from generic Ω real functions into a subspace finite-dimensional of $H^1(\Omega)$.

3.3.5 Notes on more exotic but important elements

As will be shown below, the Lagrange family introduced above provides, through its associated global interpolant, an H^1–conforming finite element space on conforming meshes. In this context, H^k–conformity means that the range of the global interpolant is a finite–dimensional subspace of the Sobolev space in which the variational formulation is posed. In particular, for second–order boundary value problems whose weak formulation lives in $H^1(\Omega)$, the global interpolant associated with Lagrange elements produces a discrete space $V_h \subset H^1(\Omega)$, making the approximation compatible with the variational setting. Although these spaces also consist of globally continuous functions, their essential property is precisely this H^1–conformity rather than classical continuity itself.

In contrast, certain boundary value problems—notably fourth–order models such as the biharmonic equation—admit natural variational formulations posed in $H^2(\Omega)$ and therefore require H^2–conforming discrete spaces, that is, finite–dimensional spaces $W_h \subset H^2(\Omega)$. Standard Lagrange elements cannot achieve this property, since their construction only guarantees interelement continuity of function values and does not ensure the regularity required for membership in $H^2(\Omega)$. Instead, specialized finite elements such as Hermite or Argyris elements are employed. For instance, the Argyris triangle uses the local polynomial space $\mathbb{P}_5(K)$ (dimension 21) and prescribes at each vertex the function value, first derivatives, and second derivatives, together with suitable edge derivative data, thereby producing a global discrete space contained in $H^2(\Omega)$.

Because such H^2–conforming elements employ derivative evaluations as degrees of freedom, the admissible domain of the associated interpolant must consist of sufficiently smooth functions. Accordingly, the global interpolant takes the form

$$I_h : \mathcal{A}(\Omega) \longrightarrow W_h,$$

where $\mathcal{A}(\Omega)$ is a function space with enough regularity to make all derivative degrees of freedom well defined (typically $C^1(\overline{\Omega})$ or $C^2(\overline{\Omega})$), and where $W_h \subset H^2(\Omega)$ is the resulting finite element space [1].

For incompressible flow models (Stokes/Navier–Stokes), the constraint $\nabla\cdot\mathbf{u} = 0$ motivates the use of stable finite element pairs and, in some settings, exactly mass–conservative discretizations. The Taylor–Hood family (e.g. P_2/P_1 on triangles) is a classical stable choice that typically provides high accuracy for both velocity and pressure; however, the discrete velocity field is only *weakly* divergence–free, i.e. the constraint is satisfied in the variational sense but not pointwise. When exact local mass conservation is required, it is natural to employ velocity spaces conforming to $H(\mathrm{div};\Omega)$, such as Raviart–Thomas (RT) or Brezzi–Douglas–Marini (BDM) elements. These spaces enforce continuity of the normal component of the velocity across interelement edges, which yields exact flux balance at the element level.

In the Raviart–Thomas family on a triangular mesh, the local space of order k is

$$\mathbf{RT}_k(K) = \mathbb{P}_k(K)^2 + \mathbf{x}\,\mathbb{P}_k(K),$$

and its degrees of freedom are given by: (i) moments of the normal component on each edge,

$$\int_e (\mathbf{v}\cdot\mathbf{n})\, q\, ds \quad \forall\, q \in \mathbb{P}_k(e),$$

which ensure continuity of normal fluxes, and (ii) internal moments

$$\int_K \mathbf{v}\cdot\mathbf{q}\, dx \quad \forall\, \mathbf{q} \in \mathbb{P}_{k-1}(K)^2,$$

which fix the interior behaviour. With a compatible pressure space, these elements yield discrete velocities that are exactly divergence–free in the appropriate polynomial sense, a property particularly relevant in long–time simulations and transport–dominated regimes.

In electromagnetic and magnetohydrodynamic (MHD) models, unknowns such as the electric or magnetic field naturally belong to function spaces involving the curl operator, leading to the Sobolev space $H(\text{curl};\Omega)$ and motivating the use of *edge elements*. The Nédélec family is specifically designed so that tangential components are continuous across interelement edges, which is precisely the conformity requirement for $H(\text{curl})$ problems and ensures the correct transmission of electromagnetic quantities across interfaces.

On a triangular mesh, the first Nédélec family of order k is locally defined by

$$\mathbf{N}_k(K) = \mathbb{P}_k(K)^2 + \mathbf{x}^\perp \mathbb{P}_k(K),$$

where $\mathbf{x}^\perp = (-y, x)$. Its degrees of freedom are given by: (i) edge moments of the tangential component,

$$\int_e (\mathbf{v}\cdot\mathbf{t})\, q\, ds \quad \forall q \in \mathbb{P}_k(e),$$

which enforce tangential continuity across edges, and (ii) internal moments

$$\int_K \mathbf{v}\cdot\mathbf{q}\, dx \quad \forall \mathbf{q} \in \mathbb{P}_{k-1}(K)^2,$$

which determine the interior behaviour. These properties make Nédélec elements particularly suitable for Maxwell equations and related curl–dominated systems.

This brief overview emphasizes a general principle: the choice of a finite element family is dictated not only by the desired polynomial degree (accuracy), but also by the Sobolev space in which the variatonal problem to solve is posed and by the conservation or structural properties one wishes to preserve at the discrete level.

3.4 Stokes problem approximated by Galerkin

In this section the Stokes variational problem with mixed boundary conditions is approximated by a Galerkin formulation based on the Taylor–Hood P_2/P_1 finite elements over a conforming triangular mesh.

The finite element method is obtained by applying the Galerkin principle to the homogeneous formulation (S^0_{weak}), defined in Definition 2.2.4 in Section 2.2.1 and shortened here for context:

Find $(\mathbf{u}_0, p) \in V_0 \times Q$ such that

$$\begin{cases} a(\mathbf{u}_0,\mathbf{v}) + b(\mathbf{v},p) = F(\mathbf{v}) - a(\mathbf{u}_g,\mathbf{v}), & \forall \mathbf{v} \in V_0, \\ b(\mathbf{u}_0,q) = -b(\mathbf{u}_g,q), & \forall q \in Q. \end{cases} \tag{$\mathrm{S}^0_{\text{weak}}$}$$

We therefore introduce finite-dimensional subspaces

$$V_{0,h} \subset V_0 = \left\{ \mathbf{v} \in H^1(\Omega)^2 \;\middle|\; \mathbf{v} = \mathbf{0} \text{ on } \Gamma_D \right\}, \qquad Q_h \subset Q = L^2_0(\Omega),$$

which will be constructed using the global Lagrange interpolant.

Definition 3.4.1 (Discrete homogeneous velocity space)**.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain and let $\mathcal{T}_h$ be a conforming triangular mesh of Ω.*

Recalling that the continuous homogeneous velocity space is

$$V_0 := \left\{ \mathbf{v} \in H^1(\Omega)^2 \;\middle|\; \mathbf{v} = \mathbf{0} \ \textit{on } \Gamma_D \right\},$$

and motivated by Proposition 3.3.16 with $k = 2$*, we define the discrete homogeneous velocity space as*

$$V_{0,h} := \left\{ \mathbf{v}_h \in \operatorname{Im}(I_h^{(2)})^2 \;\middle|\; \mathbf{v}_h = \mathbf{0} \ \textit{on } \Gamma_D \right\}.$$

Then $V_{0,h}$ *is a finite–dimensional subspace of* V_0*. Moreover, if* $N_{h,D}^{(2)}$ *denotes the number of global* P_2 *Lagrange nodes lying on* Γ_D *and* $N_h^{(2)}$ *the total number of global* P_2 *Lagrange nodes in the mesh, then*

$$\dim(V_{0,h}) = 2\big(N_h^{(2)} - N_{h,D}^{(2)}\big),$$

since the homogeneous Dirichlet condition fixes both velocity components at all P_2 *nodes on* Γ_D*.*

To approximate the non-homogeneous Dirichlet condition, we define a discrete lifting of the boundary data by means of the global interpolant:

$$\mathbf{u}_{g,h} := I_h^{(2)}\mathbf{g}.$$

This function belongs to $\operatorname{Im}(I_h^{(2)})^2$ and matches the Dirichlet data at the Lagrange nodes lying on Γ_D. The full discrete velocity space is then given by

$$V_{\mathbf{g},h} := \mathbf{u}_{g,h} + V_{0,h}.$$

Definition 3.4.2 (Discrete pressure space)**.** *Recalling that the continuous pressure space is* $Q = L_0^2(\Omega)$*, and observing that* $\operatorname{Im}(I_h^{(1)}) \subset H^1(\Omega) \subset L^2(\Omega)$*, we define the discrete pressure space as*

$$Q_h := \left\{ q_h \in \operatorname{Im}(I_h^{(1)}) \;\middle|\; \int_\Omega q_h \, dx = 0 \right\}.$$

Then Q_h *is a finite–dimensional subspace of* $L_0^2(\Omega)$*. Moreover, its dimension is equal to the total number of global Lagrange nodes of degree* 1 *associated with the mesh* $\mathcal{T}_h$*, minus one, since the zero–mean condition* $\int_\Omega q_h \, dx = 0$ *defines a nontrivial linear constraint on* $\operatorname{Im}(I_h^{(1)})$*.*

The pair $(V_{0,h}, Q_h)$ together with the discrete lifting $\mathbf{u}_{g,h}$ yields the classical Taylor–Hood P_2/P_1 finite element approximation of the Stokes problem with mixed boundary conditions.

Definition 3.4.3 (Galerkin discretization of the mixed Stokes problem)**.** *The Galerkin approximation of* (S^0_{weak}) *consists of finding*

$$(\mathbf{u}_{0,h}, p_h) \in V_{0,h} \times Q_h$$

such that

$$\begin{cases} a(\mathbf{u}_{0,h}, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = F(\mathbf{v}_h) - a(\mathbf{u}_{g,h}, \mathbf{v}_h), & \forall \mathbf{v}_h \in V_{0,h}, \\ b(\mathbf{u}_{0,h}, q_h) = -b(\mathbf{u}_{g,h}, q_h), & \forall q_h \in Q_h. \end{cases} \tag{S^0_h}$$

The discrete velocity approximation of the original mixed boundary value problem is then defined by

$$\mathbf{u}_h := \mathbf{u}_{g,h} + \mathbf{u}_{0,h}.$$

Accordingly, the pair $(\mathbf{u}_h, p_h) \in V_{\mathbf{g},h} \times Q_h$ constitutes the Taylor–Hood finite element approximation of the Stokes problem with mixed boundary conditions. It is well known that the Taylor–Hood finite element pair satisfies the discrete inf–sup (Ladyzhenskaya–Babuška–Brezzi) condition, and therefore yields a stable and convergent mixed approximation; see, e.g., [3, 11, 7].

3.4.1 Algebraic linear system associated with Taylor–Hood P_2/P_1

Since $V_{0,h}$ and Q_h are finite–dimensional spaces by Definitions 3.4.1 and 3.4.2, their bases are induced by the global Lagrange nodal bases associated with the interpolants $I_h^{(2)}$ and $I_h^{(1)}$, as characterized in Proposition 3.3.15. We describe these bases explicitly. Let

$$N_2 := N_h^{(2)}, \qquad N_1 := N_h^{(1)},$$

denote respectively the total number of global Lagrange nodes of degree 2 and degree 1 associated with the mesh $\mathcal{T}_h$. Let

$$\{a_i^{(2)}\}_{i=1}^{N_2}, \qquad \{a_j^{(1)}\}_{j=1}^{N_1},$$

denote respectively the global Lagrange nodes of degree 2 and degree 1 associated with the mesh $\mathcal{T}_h$. Let

$$\{\varphi_i^{(2)}\}_{i=1}^{N_2}$$

be the corresponding global scalar nodal basis of $\mathrm{Im}(I_h^{(2)})$, characterized by

$$\varphi_i^{(2)}(a_j^{(2)}) = \delta_{ij}, \qquad \mathrm{supp}(\varphi_i^{(2)}) \subset \bigcup \{K \in \mathcal{T}_h \mid a_i^{(2)} \in K\}.$$

For each element $K \in \mathcal{T}_h$, the restriction $\varphi_i^{(2)}|_K$ coincides either with zero or with one of the local quadratic Lagrange shape functions defined on K, as constructed on the reference triangle in Proposition 3.3.11 and transported to physical elements according to Remark 3.3.12. The global functions are obtained by assembling these local shape functions consistently across the mesh.

Each $\varphi_i^{(2)}$ is globally continuous, piecewise polynomial of degree 2 on every element of $\mathcal{T}_h$, and vanishes outside the patch of elements sharing the node $a_i^{(2)}$. Representative examples corresponding to interior vertex nodes, interior edge midpoint nodes, and boundary vertex nodes are shown in Figure 3.3.

Recalling that

$$V_{0,h} = \{\mathbf{v}_h \in \mathrm{Im}(I_h^{(2)})^2 \mid \mathbf{v}_h = 0 \text{ on } \Gamma_D\},$$

only those P_2 nodes not lying on Γ_D generate velocity degrees of freedom. Let

$$\mathcal{N}_{h,0}^{(2)} := \{\, a_i^{(2)} \mid a_i^{(2)} \notin \Gamma_D \,\}.$$

For each $a_i^{(2)} \in \mathcal{N}_{h,0}^{(2)}$ we define the two vector-valued functions

$$\boldsymbol{\varphi}_{i,1} = \begin{pmatrix} \varphi_i^{(2)} \\ 0 \end{pmatrix}, \qquad \boldsymbol{\varphi}_{i,2} = \begin{pmatrix} 0 \\ \varphi_i^{(2)} \end{pmatrix}.$$

These functions satisfy

$$\boldsymbol{\varphi}_{i,\alpha}(a_j^{(2)}) = \delta_{ij}\,\mathbf{e}_\alpha, \qquad \alpha \in \{1,2\},$$

where $\{\mathbf{e}_1, \mathbf{e}_2\}$ denotes the canonical basis of $\mathbb{R}^2$, and they vanish at all Dirichlet nodes by construction. The collection of all such functions forms a basis of $V_{0,h}$. Its cardinality is

$$N_u = 2\,\#\mathcal{N}_{h,0}^{(2)} = \dim(V_{0,h}),$$

since two velocity degrees of freedom are associated with each unconstrained P_2 node.

Regarding the base of Q_h, let

$$\{\psi_j^{(1)}\}_{j=1}^{N_1}$$

be the corresponding global scalar nodal basis of $\mathrm{Im}(I_h^{(1)})$, characterized by

$$\psi_j^{(1)}(a_k^{(1)}) = \delta_{jk}, \qquad \mathrm{supp}(\psi_j^{(1)}) \subset \bigcup\{K \in \mathcal{T}_h \mid a_j^{(1)} \in K\}.$$

For each element $K \in \mathcal{T}_h$, the restriction $\psi_j^{(1)}|_K$ coincides either with zero or with one of the local linear Lagrange shape functions defined on K in Proposition 3.3.7, constructed on the reference triangle and transported to physical elements by affine pullback, exactly as described for the quadratic case. The global functions are obtained by assembling these local shape functions consistently across the mesh.

Each $\psi_j^{(1)}$ is globally continuous, piecewise polynomial of degree 1 on every element of $\mathcal{T}_h$, and vanishes outside the patch of elements sharing the node $a_j^{(1)}$. A example corresponding to interior vertex nodes is shown in Figure 3.2.

Recalling that

$$Q_h = \left\{ q_h \in \mathrm{Im}(I_h^{(1)}) \;\middle|\; \int_\Omega q_h\,dx = 0 \right\},$$

the zero–mean condition defines a single nontrivial linear constraint on $\mathrm{Im}(I_h^{(1)})$. Hence Q_h is a codimension–one subspace of $\mathrm{Im}(I_h^{(1)})$.

A basis of Q_h is obtained by restricting the nodal basis $\{\psi_j^{(1)}\}_{j=1}^{N_1}$ to the subspace satisfying the mean–zero condition. The collection of the resulting functions forms a basis of Q_h. Its cardinality is

$$N_p = N_1 - 1 = \dim(Q_h),$$

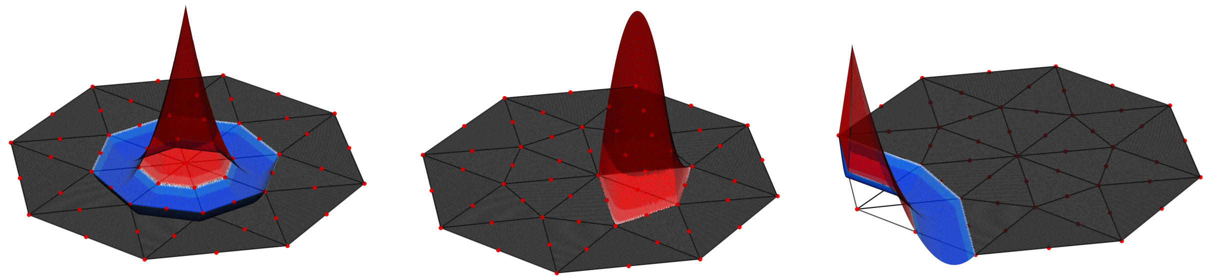
since the zero–mean condition removes exactly one independent degree of freedom.

Any discrete solution $(\mathbf{u}_{0,h}, p_h) \in V_{0,h} \times Q_h$ admits the expansions

$$\mathbf{u}_{0,h} = \sum_{i=1}^{N_u} U_i\,\boldsymbol{\varphi}_i, \qquad p_h = \sum_{j=1}^{N_p} P_j\,\psi_j,$$

for uniquely determined coefficient vectors

$$\mathbf{U} \in \mathbb{R}^{N_u}, \qquad \mathbf{P} \in \mathbb{R}^{N_p}.$$



(a) Interior vertex node (b) Interior edge midpoint node (c) Boundary vertex node

Figure 3.3: Global Lagrange P_2 nodal basis functions associated with three different types of nodes in a conforming triangular mesh. Each function is globally continuous, piecewise quadratic on every element, and supported only on the union of the elements sharing the corresponding node. From left to right: basis function associated with an interior vertex node, with an interior edge midpoint node, and with a boundary vertex node.

Proposition 3.4.4 (Reduction to basis test functions)**.** *Let V be a finite–dimensional vector space with basis $\{\varphi_k\}_{k=1}^N$, and let $\mathcal{L} : V \to \mathbb{R}$ be a linear functional. If*

$$\mathcal{L}(\varphi_k) = 0 \quad \textit{for all } k = 1, \dots, N,$$

then $\mathcal{L}(v) = 0$ for all $v \in V$.

Proof. Let $v \in V$ be arbitrary. Since $\{\varphi_k\}$ is a basis of V, there exist unique coefficients $\{v_k\}_{k=1}^N$ such that

$$v = \sum_{k=1}^{N} v_k \, \varphi_k.$$

By linearity,

$$\mathcal{L}(v) = \sum_{k=1}^{N} v_k \, \mathcal{L}(\varphi_k).$$

If $\mathcal{L}(\varphi_k) = 0$ for all k, then $\mathcal{L}(v) = 0$. □

In *(S_h^0)*, the mappings

$$\mathbf{v}_h \longmapsto a(\mathbf{u}_{0,h}, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) - F(\mathbf{v}_h) + a(\mathbf{u}_{g,h}, \mathbf{v}_h)$$

and

$$q_h \longmapsto b(\mathbf{u}_{0,h}, q_h) + b(\mathbf{u}_{g,h}, q_h)$$

are linear in the test arguments. Since $V_{0,h}$ and Q_h are finite–dimensional and $\{\boldsymbol{\varphi}_k\}_{k=1}^{N_u}$, $\{\psi_\ell\}_{\ell=1}^{N_p}$ are bases of these spaces, Proposition 3.4.4 shows that it is sufficient to impose the equalities in *(S_h^0)* only for the basis functions.

Consequently, one obtains exactly N_u independent equations from the velocity test space and N_p independent equations from the pressure test space. Since the unknown coefficient vectors $\mathbf{U} \in \mathbb{R}^{N_u}$ and $\mathbf{P} \in \mathbb{R}^{N_p}$ contain $N_u + N_p$ scalar unknowns, the resulting algebraic system is square.

Substituting the expansions of $\mathbf{u}_{0,h}$ and p_h into (S_h^0) and testing with $\boldsymbol{\varphi}_k$ and ψ_ℓ yields

$$\sum_{i=1}^{N_u} U_i\, a(\boldsymbol{\varphi}_i, \boldsymbol{\varphi}_k) + \sum_{j=1}^{N_p} P_j\, b(\boldsymbol{\varphi}_k, \psi_j) = F(\boldsymbol{\varphi}_k) - a(\mathbf{u}_{g,h}, \boldsymbol{\varphi}_k), \qquad k = 1, \dots, N_u,$$

$$\sum_{i=1}^{N_u} U_i\, b(\boldsymbol{\varphi}_i, \psi_\ell) = -b(\mathbf{u}_{g,h}, \psi_\ell), \qquad \ell = 1, \dots, N_p.$$

We now rewrite the algebraic relations above in matrix form by expanding the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ according to their definitions in the mixed variational Stokes formulation (Definition 2.2.2).

Recall that, for $\mathbf{w}, \mathbf{v} \in H^1(\Omega)^2$ and $q \in L^2(\Omega)$,

$$a(\mathbf{w}, \mathbf{v}) = \int_\Omega 2\mu\, \varepsilon(\mathbf{w}) : \varepsilon(\mathbf{v})\, dx, \qquad b(\mathbf{v}, q) = -\int_\Omega q\, \nabla\cdot\mathbf{v}\, dx, \qquad F(\mathbf{v}) = \int_\Omega \mathbf{f}\cdot\mathbf{v}\, dx + \int_{\Gamma_N} \mathbf{t}\cdot\mathbf{v}\, ds,$$

where $\varepsilon(\mathbf{w}) = \frac{1}{2}(\nabla\mathbf{w} + \nabla\mathbf{w}^\top)$.

Let $\{\boldsymbol{\varphi}_k\}_{k=1}^{N_u}$ be the basis of $V_{0,h}$ and $\{\psi_\ell\}_{\ell=1}^{N_p}$ the basis of Q_h introduced above. For $k, i \in \{1, \dots, N_u\}$ and $\ell, j \in \{1, \dots, N_p\}$ we define the matrices

$$A \in \mathbb{R}^{N_u \times N_u}, \qquad B \in \mathbb{R}^{N_u \times N_p},$$

by the entries

$$A_{ki} := a(\boldsymbol{\varphi}_i, \boldsymbol{\varphi}_k) = \int_\Omega 2\mu\, \varepsilon(\boldsymbol{\varphi}_i) : \varepsilon(\boldsymbol{\varphi}_k)\, dx,$$

$$B_{kj} := b(\boldsymbol{\varphi}_k, \psi_j) = -\int_\Omega \psi_j\, \nabla\cdot\boldsymbol{\varphi}_k\, dx.$$

We also define the vectors $\mathbf{F} \in \mathbb{R}^{N_u}$ and $\mathbf{G} \in \mathbb{R}^{N_p}$ by

$$F_k := F(\boldsymbol{\varphi}_k) - a(\mathbf{u}_{g,h}, \boldsymbol{\varphi}_k) = \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}_k\, dx + \int_{\Gamma_N} \mathbf{t}\cdot\boldsymbol{\varphi}_k\, ds - \int_\Omega 2\mu\, \varepsilon(\mathbf{u}_{g,h}) : \varepsilon(\boldsymbol{\varphi}_k)\, dx,$$

$$G_\ell := -b(\mathbf{u}_{g,h}, \psi_\ell) = \int_\Omega \psi_\ell\, \nabla\cdot\mathbf{u}_{g,h}\, dx.$$

With these definitions, the relations obtained from testing with the basis functions become

$$\sum_{i=1}^{N_u} A_{ki}\, U_i + \sum_{j=1}^{N_p} B_{kj}\, P_j = F_k, \qquad k = 1, \dots, N_u,$$

$$\sum_{i=1}^{N_u} B_{i\ell}\, U_i = G_\ell, \qquad \ell = 1, \dots, N_p,$$

which can be written compactly as the block linear system

$$\begin{pmatrix} A & B \\ B^\top & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix}. \tag{S$_h^{\mathrm{alg}}$}$$

The matrix A is symmetric by symmetry of $a(\cdot,\cdot)$, and the global system has the classical saddle–point structure of mixed formulations. Moreover, by the local support of the basis functions $\boldsymbol{\varphi}_i$ and ψ_j, the matrices A and B are sparse.

3.4.1.1 Computation of matrix elements

The algebraic system involves the matrices

$$A_{ki} = \int_\Omega 2\mu\, \varepsilon(\boldsymbol{\varphi}_i) : \varepsilon(\boldsymbol{\varphi}_k)\, dx, \qquad B_{kj} = -\int_\Omega \psi_j\, \nabla \cdot \boldsymbol{\varphi}_k\, dx,$$

and the vectors

$$F_k = \int_\Omega \mathbf{f} \cdot \boldsymbol{\varphi}_k\, dx + \int_{\Gamma_N} \mathbf{t} \cdot \boldsymbol{\varphi}_k\, ds - \int_\Omega 2\mu\, \varepsilon(\mathbf{u}_{g,h}) : \varepsilon(\boldsymbol{\varphi}_k)\, dx,$$

$$G_\ell = \int_\Omega \psi_\ell\, \nabla \cdot \mathbf{u}_{g,h}\, dx.$$

Since $\mathcal{T}_h$ is a conforming triangulation of Ω, all volume integrals decompose elementwise:

$$A_{ki} = \sum_{K \in \mathcal{T}_h} A_{ki}^{(K)}, \qquad B_{kj} = \sum_{K \in \mathcal{T}_h} B_{kj}^{(K)},$$

$$F_k = \sum_{K \in \mathcal{T}_h} F_k^{(K)} + \int_{\Gamma_N} \mathbf{t} \cdot \boldsymbol{\varphi}_k\, ds, \qquad G_\ell = \sum_{K \in \mathcal{T}_h} G_\ell^{(K)},$$

where

$$A_{ki}^{(K)} = \int_K 2\mu\, \varepsilon(\boldsymbol{\varphi}_i) : \varepsilon(\boldsymbol{\varphi}_k)\, dx,$$

$$B_{kj}^{(K)} = -\int_K \psi_j\, \nabla \cdot \boldsymbol{\varphi}_k\, dx,$$

$$F_k^{(K)} = \int_K \mathbf{f} \cdot \boldsymbol{\varphi}_k\, dx - \int_K 2\mu\, \varepsilon(\mathbf{u}_{g,h}) : \varepsilon(\boldsymbol{\varphi}_k)\, dx,$$

$$G_\ell^{(K)} = \int_K \psi_\ell\, \nabla \cdot \mathbf{u}_{g,h}\, dx.$$

Let $F_K : \widehat{K} \to K$ be the affine mapping with Jacobian B_K and determinant $J_K = \det B_K$. For any integrable function g on K,

$$\int_K g(x)\, dx = \int_{\widehat{K}} g(F_K(\widehat{x}))\, |J_K|\, d\widehat{x}.$$

If $\boldsymbol{\varphi}_i = \widehat{\boldsymbol{\varphi}}_i \circ F_K^{-1}$, then

$$\nabla \boldsymbol{\varphi}_i = (\nabla_{\widehat{x}} \widehat{\boldsymbol{\varphi}}_i)\, B_K^{-1}, \qquad \varepsilon(\boldsymbol{\varphi}_i) = \frac{1}{2}\Big((\nabla_{\widehat{x}} \widehat{\boldsymbol{\varphi}}_i) B_K^{-1} + B_K^{-T} (\nabla_{\widehat{x}} \widehat{\boldsymbol{\varphi}}_i)^\top \Big).$$

After change of variables, the elemental integrals become

$$A_{ki}^{(K)} = \int_{\widehat{K}} 2\mu\, \varepsilon(\widehat{\boldsymbol{\varphi}}_i) : \varepsilon(\widehat{\boldsymbol{\varphi}}_k)\, |J_K|\, d\widehat{x},$$

$$B_{kj}^{(K)} = -\int_{\widehat{K}} \widehat{\psi}_j\, (\nabla \cdot \widehat{\boldsymbol{\varphi}}_k)\, |J_K|\, d\widehat{x},$$

$$F_k^{(K)} = \int_{\widehat{K}} \mathbf{f}(F_K(\widehat{x})) \cdot \widehat{\boldsymbol{\varphi}}_k \, |J_K| \, d\widehat{x} - \int_{\widehat{K}} 2\mu \, \varepsilon(\widehat{\mathbf{u}}_{g,h}) : \varepsilon(\widehat{\boldsymbol{\varphi}}_k) \, |J_K| \, d\widehat{x},$$

$$G_\ell^{(K)} = \int_{\widehat{K}} \widehat{\psi}_\ell \, (\nabla \cdot \widehat{\mathbf{u}}_{g,h}) \, |J_K| \, d\widehat{x}.$$

All integrals over $\widehat{K}$ are evaluated numerically by Gauss quadrature:

$$\int_{\widehat{K}} g(\widehat{x}) \, d\widehat{x} \approx \sum_{q=1}^{N_q} w_q \, g(\widehat{x}_q).$$

Thus, for example,

$$A_{ki}^{(K)} \approx \sum_{q=1}^{N_q} w_q \, 2\mu \, \varepsilon(\widehat{\boldsymbol{\varphi}}_i(\widehat{x}_q)) : \varepsilon(\widehat{\boldsymbol{\varphi}}_k(\widehat{x}_q)) \, |J_K|,$$

$$F_k^{(K)} \approx \sum_{q=1}^{N_q} w_q \, \mathbf{f}(F_K(\widehat{x}_q)) \cdot \widehat{\boldsymbol{\varphi}}_k(\widehat{x}_q) \, |J_K| - \sum_{q=1}^{N_q} w_q \, 2\mu \, \varepsilon(\widehat{\mathbf{u}}_{g,h}(\widehat{x}_q)) : \varepsilon(\widehat{\boldsymbol{\varphi}}_k(\widehat{x}_q)) \, |J_K|.$$

The global matrices and right–hand side vectors are obtained by assembling the elemental contributions over all $K \in \mathcal{T}_h$.

Chapter 4

Numerical results

This chapter presents the numerical results obtained for the two–dimensional stationary incompressible Stokes problem introduced in Section 2.2.1. The computations are performed on a flow domain containing an airfoil section, and the discrete solutions are obtained by solving the algebraic system derived in Section 3.4.1 within the Taylor–Hood finite element framework.

4.1 Domain and boundary conditions specification

In order to pursue the aim of simulating the airflow around the two–dimensional section of a wind turbine blade, a computational domain is constructed so as to contain the corresponding airfoil geometry.

The blade section considered throughout this work corresponds to the NACA 633418 profile, a laminar airfoil belonging to the NACA 6–series, originally designed to achieve extended regions of low drag. This particular profile is representative of blade sections commonly employed in wind energy applications due to its favorable lift–to–drag characteristics and its structural thickness, which makes it especially suitable for turbine blades.

The fluid domain is defined as a square region $\Omega_\square \subset \mathbb{R}^2$. The airfoil section is embedded inside this square and removed from it, so that the final computational domain is given by

$$\Omega = \Omega_\square \setminus \overline{\Omega_{\text{blade}}},$$

where Ω_{blade} denotes the interior of the NACA 633418 profile.

The outer square boundary is chosen sufficiently distant from the blade in all directions in order to approximate an external flow configuration and to minimize the influence of artificial boundaries on the velocity and pressure fields near the airfoil surface. This configuration allows us to model the aerodynamic behavior of the blade section under controlled inflow conditions while preserving a mathematically well–posed boundary value problem.

Although real wind turbine blades operate in a fully three–dimensional environment, the present study focuses on a two–dimensional cross section, which provides valuable insight into the local aerodynamic behavior while significantly reducing computational complexity.

The boundary $\partial\Omega$ is decomposed into the outer boundary of the square and the blade boundary:

$$\partial\Omega = \partial\Omega_{\square} \;\cup\; \partial\Omega_{\text{blade}}.$$

We introduce a Cartesian coordinate system (x, y) with the blade placed inside the box as shown in Figure 4.1. The outward unit normal to $\partial\Omega$ is denoted by $\mathbf{n}$. The outer boundary $\partial\Omega_{\square}$ is split into an *inflow–driven* Dirichlet part and a *natural* (do–nothing) part:

$$\partial\Omega_{\square} = \Gamma_D \cup \Gamma_N, \qquad \Gamma_D \cap \Gamma_N = \emptyset,$$

where, in this setup,

$$\Gamma_D := \Gamma_{\text{left}} \;\cup\; \Gamma_{\text{bottom}}, \qquad \Gamma_N := \Gamma_{\text{right}} \;\cup\; \Gamma_{\text{top}}.$$

That is, the left vertical side and the bottom horizontal side of the box carry Dirichlet data, while the top and right sides use the natural (do–nothing) boundary condition associated with the variational formulation. Finally, on the blade boundary we enforce a no–slip condition:

$$\mathbf{u} = \mathbf{0} \quad \text{on } \Gamma_{\text{blade}}.$$

On Γ_D we prescribe a uniform velocity vector whose magnitude is a controllable parameter and whose direction encodes the angle of attack. Specifically, let $U_\infty > 0$ be the chosen inflow magnitude and let $\alpha \in \mathbb{R}$ be the angle of attack (measured counterclockwise from the positive x–axis). The imposed Dirichlet datum is

$$\mathbf{g}(U_\infty, \alpha) := U_\infty \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}.$$

The boundary condition on Γ_D is therefore

$$\mathbf{u} = \mathbf{g}(U_\infty, \alpha) \quad \text{on } \Gamma_D,$$

to be understood in the trace sense (as in the mixed variational framework in Definition 2.2.2 of the thesis).

On Γ_N we impose the natural boundary condition associated with the Stokes stress, i.e. a traction–free condition

$$\sigma(\mathbf{u}, p)\,\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_N, \qquad \sigma(\mathbf{u}, p) = 2\mu\,\varepsilon(\mathbf{u}) - p\,I.$$

This is the standard *do–nothing* boundary condition: it arises directly from the integration by parts in the weak formulation and avoids prescribing pressure values on the outflow boundary.

4.2 Mesh

Figures 4.1–4.2 show two views of the unstructured triangular mesh used in the computations. The mesh is intentionally *graded*: it is coarser far from the airfoil, while it is strongly refined in a neighborhood of Γ_{blade} and in the near–wake region downstream of the trailing edge. This refinement strategy is motivated by the fact that the largest velocity gradients (hence the dominant contributions to viscous stresses) concentrate around the solid boundary and in the region where the flow reorganizes after passing the trailing edge. A fine boundary resolution is therefore essential to obtain accurate wall tractions and pressure distributions, which are the primary outputs of interest in the present study.

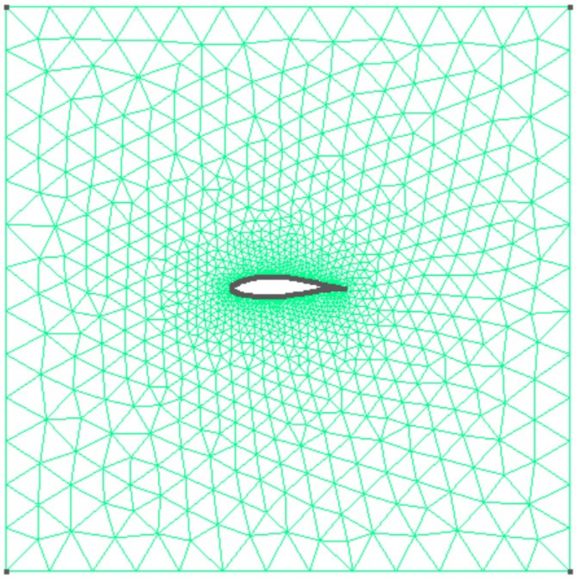


Figure 4.1: Overview of the computational domain Ω and the conforming unstructured triangular discretization used in this work. The outer boundary is the rectangle $\Omega_\square = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$ with $(x_{\min}, x_{\max}, y_{\min}, y_{\max}) = (-2.0,\ 3.0,\ -2.5,\ 2.5)$, in which the NACA 633418 airfoil is embedded and removed, so that $\Omega = \Omega_\square \setminus \overline{\Omega_{\text{blade}}}$. The mesh statistics are: 4799 nodes and 2321 triangular elements. Using the Taylor–Hood P_2/P_1 discretization, the resulting mixed Stokes problem leads to a sparse linear system of size 10837×10837, assembled and stored in BCOO format with 186288 nonzero entries.

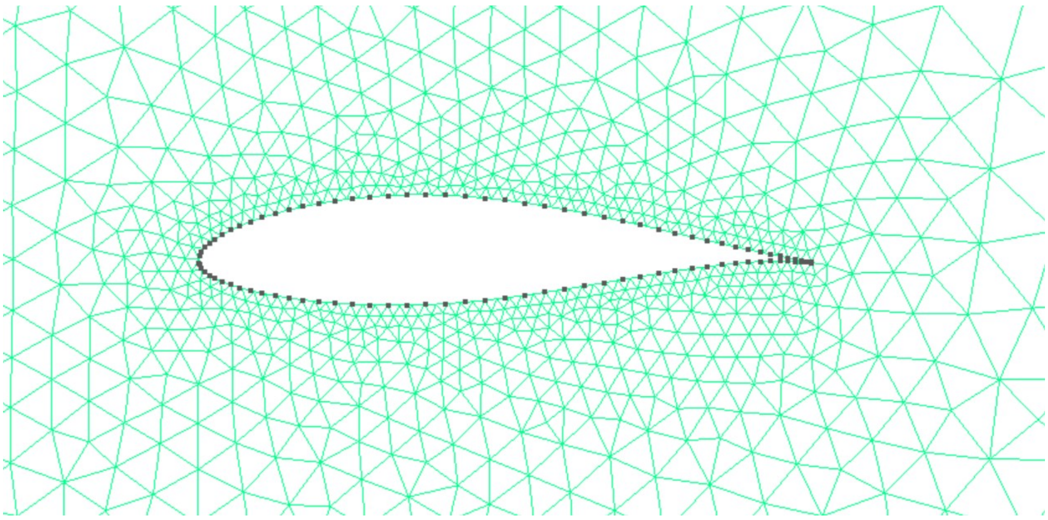


Figure 4.2: Detail of the mesh refinement near the blade boundary Γ_{blade}. The resolution is increased around the leading edge, along the surface, and near the trailing edge to better capture boundary stresses and pressure variations.

4.3 Field images and pressure coefficient along the blade

In this section we present the velocity and pressure fields obtained from the numerical solution. All simulations are performed on the graded unstructured mesh shown in Figure 4.1.

The simulation corresponds to a low Reynolds laminar regime. The imposed inflow magnitude is $U_\infty = 1$, the dynamic viscosity is $\mu = 1$, the angle of attack is $\alpha = 8°$, the density is set to $\rho = 1$ and the airfoil chord is normalized to $L = 1$. These values lead to a Reynolds number

$$Re = \frac{\rho U_\infty L}{\mu} = 1,$$

which places the flow in a strongly viscous regime where inertial effects are negligible and the Stokes approximation is appropriate.

Figure 4.3 shows the two components of the velocity field. The uniform inflow prescribed on the left and bottom boundaries is visible, while the no–slip condition on the blade surface produces a deceleration of the flow and a smooth deflection around the profile. Downstream of the trailing edge the wake reorganizes without separation, which is consistent with the very low Reynolds number considered.

Figure 4.4 displays the pressure distribution in the domain. A pressure increase appears near the leading edge stagnation region followed by a gradual decrease along the suction side of the airfoil. The pressure gradients remain smooth throughout the domain, reflecting the elliptic character of the Stokes equations.

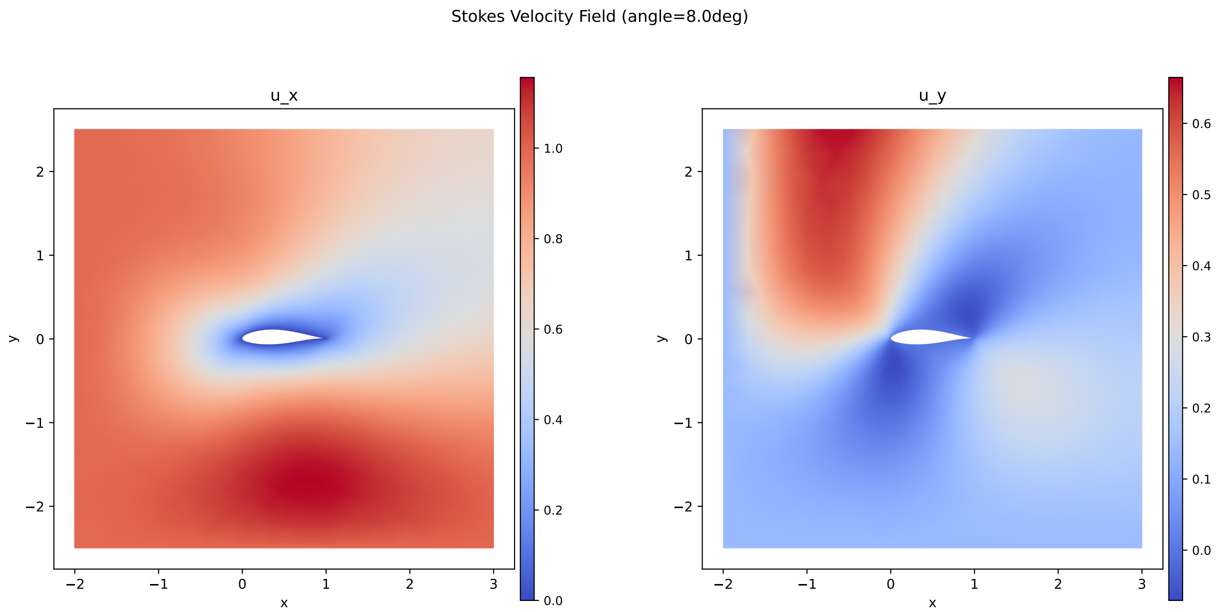


Figure 4.3: Velocity field components (u_x, u_y) for $U_\infty = 1$, $\mu = 1$ and $\alpha = 8°$ on the mesh shown in Figure 4.1.

4.4 Pressure coefficient through the blade

In order to quantify the pressure distribution along the airfoil surface, the nondimensional pressure coefficient is evaluated along both sides of the blade. The coefficient is defined as

$$C_p = \frac{p - p_\infty}{\frac{1}{2}\rho U_\infty^2},$$

where p denotes the local pressure on the airfoil boundary, p_∞ is the reference free–stream pressure, ρ is the fluid density and U_∞ is the imposed inflow magnitude.

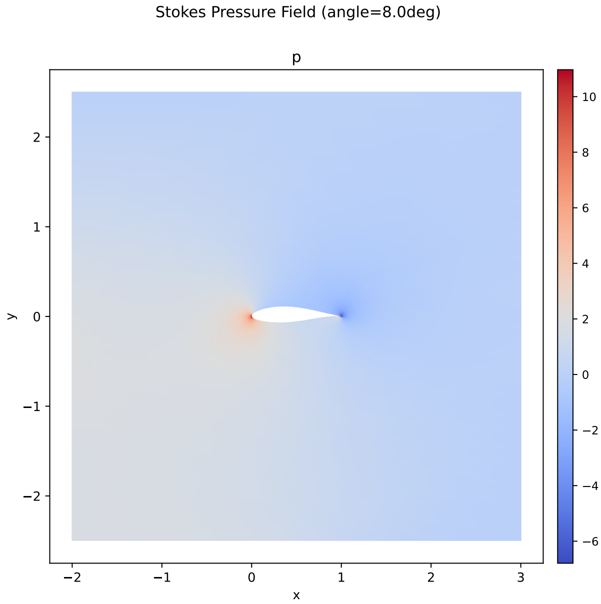


Figure 4.4: Pressure field for $U_\infty = 1$, $\mu = 1$ and $\alpha = 8°$ on the mesh shown in Figure 4.1.

The coefficient is sampled along the chordwise coordinate x/c, where c denotes the airfoil chord length. The evaluation is performed separately on the lower (pressure) side and on the upper (suction) side of the blade using the same simulation parameters described in the previous section, namely $U_\infty = 1$, $\mu = 1$, $\rho = 1$, $L = 1$ and $\alpha = 8°$, which correspond to a Reynolds number $Re = 1$. The underlying discrete solution is obtained on the graded mesh introduced in Figure 4.1.

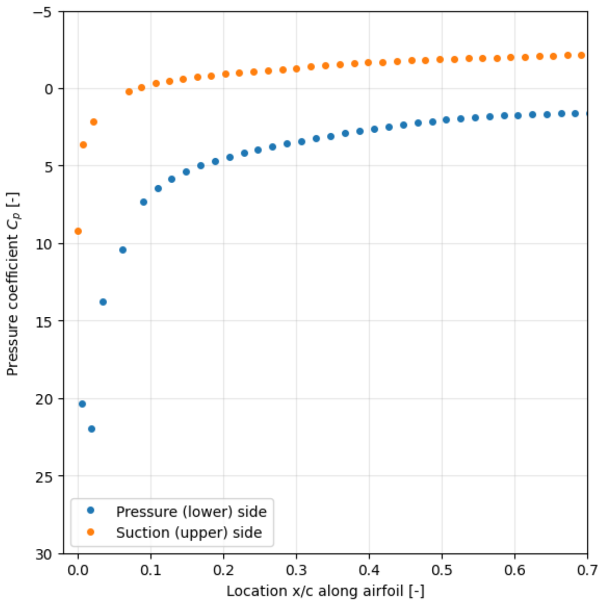


Figure 4.5: Pressure coefficient distribution C_p along the airfoil chord for the lower and upper surfaces obtained with $U_\infty = 1$, $\mu = 1$ and $\alpha = 8°$ on the mesh shown in Figure 4.1.

Figure 4.5 shows the resulting pressure coefficient distribution. A clear pressure peak appears near the leading edge on the lower surface, corresponding to the stagnation region. Moving towards the trailing edge, the pressure decreases smoothly along both sides of the airfoil. The suction side exhibits lower pressure values over most of the chord, while the pressure side remains comparatively higher. The absence of abrupt variations or oscillations is consistent with the strongly viscous regime and with the smooth character of the Stokes solution.

Chapter 5

Conclusions and future work

5.1 Conclusions

In this thesis, a rigorous mathematical formulation of the Finite Element Method has been developed, with application to the stationary two–dimensional Stokes problem. The main objective was to provide a clear theoretical basis for the numerical solver used in the computational framework.

Throughout the document, the variational theory was introduced as the analytical foundation of boundary value problems. Later, the formal definition of finite elements and the construction of conforming discretizations were presented. Finally, numerical simulations were performed in order to verify the correctness and stability of the proposed method.

The presented results are subject to the physical limitations of the stationary Stokes model. In particular, the Stokes equations are valid in the low Reynolds number regime, whereas the experimental conditions correspond to significantly higher Reynolds numbers. For this reason, a direct quantitative comparison with experimental measurements would not be physically meaningful.

Instead, the value of the present results lies in the verification of the mathematical and numerical framework, as well as in the successful adaptation of the solver to the airfoil geometry. The computed solutions exhibit physically consistent behavior within the assumptions of the model and provide a solid basis for future extensions to more realistic flow regimes.

In conclusion, this work establishes a solid bridge between mathematical rigor and numerical implementation, providing a reliable theoretical framework for the use of finite element methods in fluid mechanics applications.

5.2 Future work

In order to achieve a realistic comparison with experimental data (see Figure 5.1) and to contribute effectively to structural health monitoring of wind turbines, several extensions are required.

First, the stationary Stokes model must be extended to the Navier–Stokes equations in order to include inertial effects at relevant Reynolds numbers. For realistic operating conditions, appropriate stabilization techniques must be incorporated, such as SUPG-based formulations or suitable turbulence models (e.g., implicit LES or RANS approaches).

Second, transient simulations are required to capture time-dependent phenomena such as vortex shedding and dynamic aeroelastic effects. Only a time-dependent formulation allows a direct and physically consistent comparison with experimentally measured pressure signals.

Beyond these physical extensions, the differentiable structure provided by the JAX-based implementation must be systematically exploited. Automatic differentiation enables the computation of sensitivities of the pressure coefficient with respect to the angle of attack. This capability is essential for constructing efficient inverse formulations that relate pressure measurements to aerodynamic parameters.

Such reduced mappings between pressure distributions and angle of attack can then be integrated into physics-informed neural network (PINN) frameworks. This integration allows near real-time estimation of the angle of attack from measured pressure data. Combined with structural models, this approach enables the reconstruction of vibration characteristics and therefore provides indirect information about the structural state of the blade

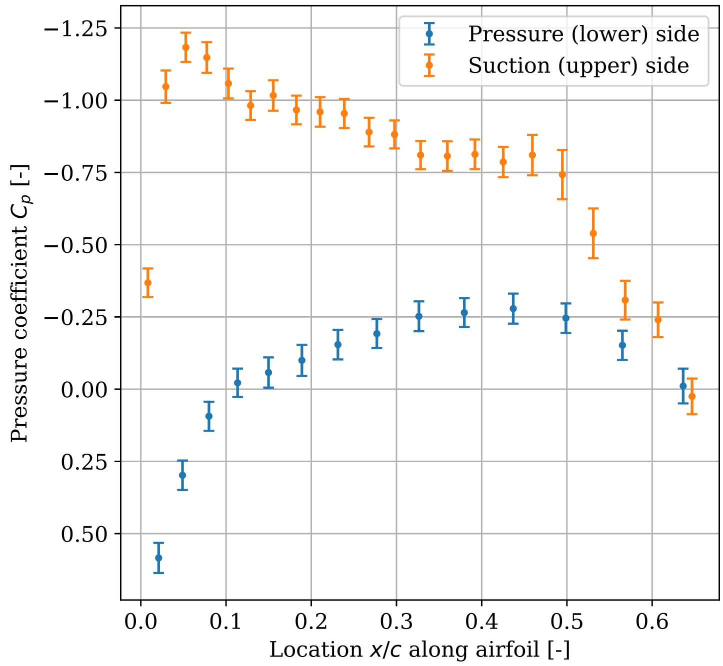


Figure 5.1: Pressure coefficient distribution along the NACA 633418 airfoil corresponding to the experimental case with the lowest Reynolds number considered in [10]. The wind tunnel operating conditions were: inflow velocity $U_\infty = 10\,\mathrm{m/s}$, air at 22°C and approximately 49% relative humidity, density $\rho = 1.225\,\mathrm{kg/m^3}$, and dynamic viscosity $\mu = 1.789 \times 10^{-5}\,\mathrm{Pa \cdot s}$. The airfoil chord length is $c = 0.16\,\mathrm{m}$ and the angle of attack is $\alpha = 8°$. With these parameters, $\mathrm{Re} \approx 1.1 \times 10^5$. Data extracted from [10].

Bibliography

[1] Brenner, S. C. and Scott, L. R. *The Mathematical Theory of Finite Element Methods.* 3rd. Vol. 15. Texts in Applied Mathematics. New York: Springer, 2008.

[2] Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* New York: Springer, 2011.

[3] Brezzi, F. and Fortin, M. *Mixed and Hybrid Finite Element Methods.* Vol. 15. Springer Series in Computational Mathematics. New York: Springer-Verlag, 1991.

[4] Ciarlet, P. G. *The Finite Element Method for Elliptic Problems.* Vol. 4. Studies in Mathematics and its Applications. Amsterdam: North-Holland, 1978.

[5] Diestel, J. *Sequences and Series in Banach Spaces.* Vol. 92. Graduate Texts in Mathematics. New York: Springer, 1984.

[6] Elman, H., Silvester, D., and Wathen, A. *Finite Elements and Fast Iterative Solvers. With Applications in Incompressible Fluid Dynamics.* Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press, 2014.

[7] Ern, A. and Guermond, J.-L. *Theory and Practice of Finite Elements.* Vol. 159. Applied Mathematical Sciences. New York: Springer, 2004.

[8] European Commission. *Wind energy in the European Union.* `https://energy.ec.europa.eu/topics/renewable-energy/eu-wind-energy_en` (accessed 2025-11-17). 2025.

[9] Evans, L. C. *Partial Differential Equations.* 2nd. Vol. 19. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2010.

[10] Franz, P. I. et al. "On the Potential of Aerodynamic Pressure Measurements for Structural Damage Detection". In: *Wind Energy Science Discussions [preprint]* (2025). DOI: `10.5194/wes-2025-26`.

[11] Girault, V. and Raviart, P.-A. *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms.* Vol. 5. Springer Series in Computational Mathematics. Berlin: Springer, 1986.

[12] Griese, F. et al. "Preconditioned FEM-based Neural Networks for Solving Incompressible Fluid Flows and Related Inverse Problems". In: *Journal of Computational and Applied Mathematics* **469** (2025), p. 116663. DOI: `10.1016/j.cam.2025.116663`.

[13] Royden, H. L. and Fitzpatrick, P. M. *Real Analysis. Third Edition.* Boston: Pearson, 2010.

[14] *Two–dimensional hat function on a triangular mesh.* `https://www.researchgate.net/figure/The-two-dimensional-hatfunction-is-illustrated-on-a-triangular-mesh-The-two-dimensional_fig8_360595481`. Accessed: 2025-11-07.

Appendix A

Preliminary definitions and basic results

A.1 Vector spaces and their duals

This section reviews the basic notions of vector spaces, linear mappings, and dual spaces that will be used throughout this work. Although these results are elementary from the point of view of linear algebra, they provide the foundation for the definition of the finite element structure.

Definition A.1.1 (Vector space)**.** *Let $\mathbb{K}$ be a field. A vector space (over $\mathbb{K}$) is a triple $(V, +, \cdot)$, where V is a set, $+ : V \times V \to V$ is a binary and internal operation such that $(V, +)$ is an abelian group, and $\cdot : \mathbb{K} \times V \to V$ is an action of the field $\mathbb{K}$ on V that defines a scalar multiplication on the group $(V, +)$.*

Definition A.1.2 (Basis)**.** *Let V be a vector space over a field $\mathbb{K}$. A subset $B = \{v_1, \dots, v_n\}$ is called a* basis *of V if:*

1. *For any scalars $\alpha_1, \dots, \alpha_n \in \mathbb{K}$,*

$$\sum_{i=1}^{n} \alpha_i v_i = 0_{\mathbb{K}} \quad \Longrightarrow \quad \alpha_1 = \cdots = \alpha_n = 0_{\mathbb{K}},$$

2. *For every $v \in V$, there exist scalars $\beta_1, \dots, \beta_n \in \mathbb{K}$ such that*

$$v = \sum_{i=1}^{n} \beta_i v_i.$$

If V admits a finite basis with n elements, then V is called finite-dimensional *and its dimension is $\dim V = n$.*

Definition A.1.3 (Linear mapping and Isomorphism)**.** *Let V, W be vector spaces over the same field $\mathbb{K}$. A mapping $T : V \to W$ is linear if for all $u, v \in V$ and $\alpha, \beta \in \mathbb{K}$ one has*

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v).$$

A linear map is called a isomorphism if it is bijective. In this case, V and W are said to be isomorphic*, and we write*

$$V \cong W.$$

Theorem A.1.4 (Basis Preservation under Isomorphism)**.** *Let V, W be vector spaces over a field $\mathbb{K}$ and let $T : V \to W$ be a vector space isomorphism. If $B = \{v_1, \ldots, v_n\}$ is a basis of V, then*

$$T(B) := \{T(v_1), \ldots, T(v_n)\}$$

is a basis of W.

Proof. We first show that $T(B)$ is linearly independent. Suppose that

$$\sum_{i=1}^{n} \alpha_i T(v_i) = 0.$$

By linearity of T, this implies

$$T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = 0.$$

Since T is injective, it follows that

$$\sum_{i=1}^{n} \alpha_i v_i = 0.$$

Because B is a basis, we obtain $\alpha_i = 0$ for all i, and hence $T(B)$ is linearly independent.

Next, we show that $T(B)$ spans W. Let $w \in W$. Since T is surjective, there exists $v \in V$ such that $T(v) = w$. Because B is a basis of V, there exist scalars $\beta_1, \ldots, \beta_n \in \mathbb{K}$ such that

$$v = \sum_{i=1}^{n} \beta_i v_i.$$

Applying T and using linearity,

$$w = T(v) = \sum_{i=1}^{n} \beta_i T(v_i),$$

which shows that w is a linear combination of elements of $T(B)$. Hence $T(B)$ spans W.

Therefore, $T(B)$ is linearly independent and spans W, and thus it is a basis of W. □

We denote by $\mathcal{L}(V, W)$ the set of all linear maps from V to W.

Definition A.1.5 (Dual space)**.** *Let V be a vector space over a field $\mathbb{K}$. The dual space of V is*

$$V' := \mathcal{L}(V, \mathbb{K}),$$

that is, the set of all linear functionals $F : V \to \mathbb{K}$. Addition and scalar multiplication are defined pointwise:

$$(F + G)(v) := F(v) + G(v), \qquad (\alpha F)(v) := \alpha F(v),$$

for $F, G \in V'$, $\alpha \in \mathbb{K}$, and $v \in V$, under which V' is again a vector space over $\mathbb{K}$.

Theorem A.1.6 (Dual basis). *Let V be a finite-dimensional vector space of dimension n over a field $\mathbb{K}$, and let $B = \{v_1, \ldots, v_n\}$ be a basis of V. Then $B^* = \{\varphi_1, \ldots, \varphi_n\}$, defined by*

$$\forall v \in V, \qquad \varphi_i(v) = \varphi_i\left(\sum_{j=1}^{n} x_j v_j\right) := x_i,$$

is a basis of V'.

Proof. First, we verify that for each i, $\varphi_i \in V'$. Given $a, b \in V$ and $\alpha, \beta \in \mathbb{K}$, and depicting a and b by their coordinates in the basis B,

$$\alpha a + \beta b = \sum_{j=1}^{n} (\alpha a_j + \beta b_j)\, v_j,$$

and therefore, by definition of φ_i,

$$\varphi_i(\alpha u + \beta v) = \alpha a_i + \beta b_i = \alpha \varphi_i(a) + \beta \varphi_i(b).$$

Hence $\{\varphi_1, \ldots, \varphi_n\} \subset V'$.

We now prove that this set is linearly independent and then that it is a spanning set. Let $\sum_{i=1}^{n} c_i \varphi_i = 0$. Evaluating at the basis vectors v_j, we obtain

$$0 = \left(\sum_{i=1}^{n} c_i \varphi_i\right)(v_j) = c_j,$$

hence $c_j = 0$ for all j, and the set is linearly independent.

On the other hand, let $F \in V'$ be arbitrary. Define the scalars $d_i := F(v_i)$ and consider

$$G := \sum_{i=1}^{n} d_i \varphi_i \in V'.$$

For an arbitrary vector $v = \sum_{i=1}^{n} x_i v_i$, we have

$$G(v) = \sum_{i=1}^{n} d_i \varphi_i(v) = \sum_{i=1}^{n} d_i x_i = \sum_{i=1}^{n} F(v_i) x_i = F\left(\sum_{i=1}^{n} x_i v_i\right) = F(v).$$

Thus G and F are equal as maps, and it follows that $\{\varphi_1, \ldots, \varphi_n\}$ spans V'. Therefore, $\{\varphi_1, \ldots, \varphi_n\}$ is a basis of V'. □

Theorem A.1.7 (Basis extension). *Let V be a finite-dimensional vector space of dimension n over a field $\mathbb{K}$, and let $S = \{v_1, \ldots, v_k\} \subset V$ be a linearly independent subset $(0 \le k \le n)$. Then S can be extended to a basis of V; that is, there exist vectors $w_{k+1}, \ldots, w_n \in V$ such that*

$$\mathcal{B} = \{v_1, \ldots, v_k, w_{k+1}, \ldots, w_n\}$$

is a basis of V.

Proof. If $k = n$, since S is linearly independent and contains n vectors in a space of dimension n, S is already a basis and there is nothing to prove. Assume $k < n$. Then $\langle S \rangle \neq V$, and therefore there exists $w_{k+1} \in V \setminus \langle S \rangle$. Hence $S \cup \{w_{k+1}\}$ is still linearly independent; otherwise $w_{k+1} \in \langle S \rangle$. By the dimension formula, we have $\dim\big(S \cup \{w_{k+1}\}\big) = k+1$. Repeating the argument $n-k$ times, we obtain a linearly independent set of n vectors; therefore, this set is a basis of V. □

Remark A.1.8. *The finiteness of the dimension is essential for the previous argument (termination in at most $n-k$ steps). In infinite-dimensional spaces, the statement remains true, but its standard proof relies on Zorn's Lemma (equivalent to the Axiom of Choice).*

Theorem A.1.9 (Pullback Isomorphism)**.** *Let V, W be vector spaces over a field $\mathbb{K}$ and let $J : V \to W$ be a vector space isomorphism. Its pullback, defined by*

$$J^* : W' \longrightarrow V', \qquad J^*(\phi) := \phi \circ J,$$

is also a vector space isomorphism between the dual spaces.

Proof. We must show that J^* is well-defined as a map, that it is linear, and that it is bijective.

(Well-definedness). Let $\phi \in W'$. Since $\phi : W \to \mathbb{K}$ is linear and $J : V \to W$ is linear, the composition

$$\phi \circ J : V \longrightarrow \mathbb{K}$$

is linear as a composition of linear maps. Therefore $\phi \circ J \in V'$, which implies that $J^*(\phi) \in V'$ for every $\phi \in W'$. Hence, J^* is well defined as a map from W' into V'.

(Linearity). Let $\phi, \psi \in W'$ and $\alpha, \beta \in \mathbb{K}$. Then

$$\begin{aligned} J^*(\alpha\phi + \beta\psi) &= (\alpha\phi + \beta\psi) \circ J \\ &= \alpha(\phi \circ J) + \beta(\psi \circ J) \\ &= \alpha J^*(\phi) + \beta J^*(\psi), \end{aligned}$$

which proves that J^* is linear.

(Bijectivity). Since J is an isomorphism, there exists an inverse isomorphism $J^{-1} : W \to V$. Consider its corresponding pullback

$$(J^{-1})^* : V' \longrightarrow W', \qquad (J^{-1})^*(\lambda) := \lambda \circ J^{-1}.$$

For every $\phi \in W'$, we have

$$\begin{aligned} \big((J^{-1})^* \circ J^*\big)(\phi) &= (J^{-1})^*(\phi \circ J) \\ &= (\phi \circ J) \circ J^{-1} \\ &= \phi \circ (J \circ J^{-1}) \\ &= \phi, \end{aligned}$$

and for every $\lambda \in V'$,

$$\begin{aligned}\big(J^* \circ (J^{-1})^*\big)(\lambda) &= J^*(\lambda \circ J^{-1}) \\ &= (\lambda \circ J^{-1}) \circ J \\ &= \lambda \circ (J^{-1} \circ J) \\ &= \lambda.\end{aligned}$$

Thus, $(J^{-1})^*$ is the inverse of J^*, and therefore J^* is bijective. Since it is also linear, J^* is a vector space isomorphism. □

A.2 Hilbert and Banach spaces

This section introduces the functional analytic framework based on Hilbert spaces, together with continuous linear functionals and bilinear forms. The notions of completeness, continuity, dual spaces, and orthogonality play a central role in the formulation and analysis of variational problems. These concepts are therefore critical for the variational theory to which the finite element method is applied.

Definition A.2.1 (Inner product space)**.** *Let V be a vector space over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. An inner product space is a pair $(V, (\cdot,\cdot)_V)$, where*

$$(\cdot,\cdot)_V : V \times V \to \mathbb{K}$$

is a mapping satisfying, for all $u, v, w \in V$ and $\alpha \in \mathbb{K}$:

1. $(u,u)_V \geq 0$, *with equality if and only if* $u = 0$,
2. $(u,v)_V = \overline{(v,u)_V}$,
3. $(\alpha u + v, w)_V = \alpha(u,w)_V + (v,w)_V$.

Definition A.2.2 (Normed vector space)**.** *Let V be a vector space over a field $\mathbb{K}$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A* norm *on V is a mapping $\|\cdot\|_V : V \to [0,\infty)$ such that for all $u, v \in V$ and $\alpha \in \mathbb{K}$:*

1. $\|v\|_V = 0$ *if and only if* $v = 0$,
2. $\|\alpha v\|_V = |\alpha|\, \|v\|_V$,
3. $\|u + v\|_V \leq \|u\|_V + \|v\|_V$.

The pair $(V, \|\cdot\|_V)$ is called a normed vector space*.*

Lemma A.2.3 (Cauchy–Schwarz inequality)**.** *Let $(V, (\cdot,\cdot)_V)$ be an inner product space over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then, for all $u, v \in V$,*

$$|(u,v)_V|^2 \leq (u,u)_V\, (v,v)_V.$$

Proof. If $u = 0$ or $v = 0$, then $(u, v)_V = 0$ and the inequality is immediate. Assume henceforth that $u \neq 0$ and $v \neq 0$. For any $\lambda \in \mathbb{K}$,

$$0 \leq (v - \lambda u,\, v - \lambda u)_V = (v, v)_V - \lambda(u, v)_V - \overline{\lambda}\,\overline{(u, v)_V} + |\lambda|^2(u, u)_V.$$

Now choose $\lambda := (u, v)_V/(u, u)_V$, well defined because $u \neq 0$, can be checked that the three last term in the right hand side are equal to $|(u, v)_V|^2/(u, u)_V$, therefore,

$$0 \leq (v, v)_V - \frac{|(u, v)_V|^2}{(u, u)_V}.$$

Multiplying by $(u, u)_V > 0$ we obtain

$$|(u, v)_V|^2 \leq (u, u)_V\,(v, v)_V,$$

which proves the claim. □

Proposition A.2.4 (Norm induced by an inner product)**.** *Let $(V, (\cdot, \cdot)_V)$ be an inner product space. Define*

$$\|v\|_V := \sqrt{(v, v)_V}, \qquad v \in V.$$

Then $\|\cdot\|_V$ is a norm on V. In particular, every inner product space is a normed vector space.

Proof. We verify the three norm axioms.

(Positivity and definiteness). By definition of an inner product, $(v, v)_V \geq 0$ for all $v \in V$, hence $\|v\|_V \geq 0$. Moreover, $\|v\|_V = 0$ if and only if $(v, v)_V = 0$, which holds if and only if $v = 0$.

(Homogeneity). Let $\alpha \in \mathbb{K}$ and $v \in V$. Using conjugate symmetry and linearity,

$$\|\alpha v\|_V^2 = (\alpha v, \alpha v)_V = \alpha(v, \alpha v)_V = \alpha\overline{\alpha}\,(v, v)_V = |\alpha|^2\|v\|_V^2,$$

hence $\|\alpha v\|_V = |\alpha|\,\|v\|_V$.

(Triangle inequality). Let $u, v \in V$. Expanding the inner product, we obtain

$$\|u + v\|_V^2 = (u + v, u + v)_V = \|u\|_V^2 + 2\mathrm{Re}(u, v)_V + \|v\|_V^2.$$

Since $\mathrm{Re}(u, v)_V \leq |(u, v)_V|$, and by the Cauchy–Schwarz inequality

$$|(u, v)_V| \leq \|u\|_V\,\|v\|_V,$$

it follows that

$$\|u + v\|_V^2 \leq \|u\|_V^2 + 2\|u\|_V\|v\|_V + \|v\|_V^2 = (\|u\|_V + \|v\|_V)^2.$$

Taking square roots yields $\|u + v\|_V \leq \|u\|_V + \|v\|_V$. □

Every normed vector space, and in particular every inner product space endowed with its induced norm, naturally defines a metric through

$$d(u,v) := \|u-v\|_V, \qquad u,v \in V.$$

This metric structure makes it possible to introduce topological notions such as convergence of sequences and completeness. In this context, completeness is a distinguished property, as it guarantees that Cauchy sequences converge within the space itself.

Definition A.2.5 (Cauchy sequence)**.** *Let $(V, \|\cdot\|_V)$ be a normed vector space. A sequence $(v_n) \subset V$ is called a Cauchy sequence if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that*

$$\|v_n - v_m\|_V < \varepsilon \qquad \text{for all } n, m \geq N.$$

Definition A.2.6 (Hilbert space)**.** *A Hilbert space is an inner product space $(V, (\cdot,\cdot)_V)$ that is complete with respect to the norm $\|\cdot\|_V$, that is, every Cauchy sequence in V converges to an element of V.*

Definition A.2.7 (Banach space)**.** *A Banach space is a normed vector space $(V, \|\cdot\|_V)$ that is complete with respect to the norm $\|\cdot\|_V$, that is, every Cauchy sequence in V converges to an element of V.*

Every Hilbert space is therefore a Banach space with the norm induced by its inner product.

Remark A.2.8. *Recall that a subset $M \subset V$ of a normed vector space $(V, \|\cdot\|_V)$ is called closed if it contains the limits of all convergent sequences in M, that is, whenever $(m_n) \subset M$ and $m_n \to m$ in V, one has $m \in M$. Notice that closedness is a purely topological property and does not imply any linear structure. In particular, a closed subset need not be a vector subspace. When a subset is both a vector subspace and closed, it is often referred to as a closed subspace; in the Hilbert setting, such a set is itself a Hilbert space when endowed with the induced inner product.*

Definition A.2.9 (Continuous linear functional)**.** *Let V be a normed vector space. A linear functional $F : V \to \mathbb{K}$ is said to be continuous if for every $v \in V$ and every sequence $(v_n) \subset V$ with $v_n \to v$ in V, one has*

$$F(v_n) \to F(v) \quad \text{in } \mathbb{K}.$$

Theorem A.2.10 (Characterization of continuity)**.** *Let V be a normed vector space and let $F : V \to \mathbb{K}$ be linear. The following statements are equivalent:*

1. *F is continuous on V;*
2. *F is continuous at 0_V;*
3. *there exists a constant $C > 0$ such that*

$$|F(v)| \leq C\|v\|_V \qquad \forall v \in V;$$

4. *F is bounded on the unit ball of V.*

Proof. We prove the chain of implications

$$(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (2) \Rightarrow (1),$$

which shows that all statements are equivalent.

(1) $\Rightarrow$ (2). If F is continuous on V, then it is continuous at 0.

(2) $\Rightarrow$ (3). Assume that F is continuous at 0. By continuity at 0, there exists $\delta > 0$ such that

$$\|v\|_V < \delta \quad \Longrightarrow \quad |F(v)| < 1.$$

Let $v \in V$ be arbitrary. If $v = 0$, the desired estimate is trivial. If $v \neq 0$, set

$$w := \frac{\delta}{2\|v\|_V}\, v,$$

so that $\|w\|_V = \delta/2 < \delta$, and hence $|F(w)| < 1$. By linearity,

$$|F(w)| = \left| \frac{\delta}{2\|v\|_V} F(v) \right| < 1,$$

which implies

$$|F(v)| < \frac{2}{\delta}\, \|v\|_V.$$

Therefore, there exists a constant $C > 0$ (for instance $C = 2/\delta$) such that

$$|F(v)| \leq C\|v\|_V \qquad \forall v \in V,$$

which proves (3).

(3) $\Rightarrow$ (4). Assume (3). If $\|v\|_V \leq 1$, then $|F(v)| \leq C$, hence F is bounded on the unit ball.

(4) $\Rightarrow$ (2). Assume that F is bounded on the unit ball, that is, there exists $M > 0$ such that

$$|F(v)| \leq M \qquad \forall v \in V \text{ with } \|v\|_V \leq 1.$$

Let $(v_n) \subset V$ be a sequence with $v_n \to 0$ in V. If $v_n = 0$, then $F(v_n) = 0$ by linearity, if $v_n \neq 0$, define

$$w_n := \frac{v_n}{\|v_n\|_V},$$

so that $\|w_n\|_V = 1$ and therefore $|F(w_n)| \leq M$. By linearity,

$$|F(v_n)| = \|v_n\|_V\, |F(w_n)| \leq M\|v_n\|_V \to 0,$$

since $\|v_n\|_V \to 0$. Hence $F(v_n) \to 0$ in $\mathbb{K}$, which shows that F is continuous at 0.

(2) $\Rightarrow$ (1). Assume that F is continuous at 0. Let $v \in V$ and let $(v_n) \subset V$ be a sequence with $v_n \to v$ in V. Then $v_n - v \to 0$ in V, and by linearity,

$$F(v_n) - F(v) = F(v_n - v) \to 0.$$

Therefore $F(v_n) \to F(v)$, which proves that F is continuous at v. Since v was arbitrary, F is continuous on V.

Therefore, all statements (1)–(4) are equivalent. □

Proposition A.2.11 (Closedness of the kernel of a bounded operator)**.** *Let H_1, H_2 be Hilbert spaces and let $T : H_1 \to H_2$ be a linear and continuous operator. Then* $\ker T$ *is a Hilbert space with the inner product inherited from H_1.*

Proof. First, $\ker T$ is a linear subspace of H_1. Indeed, if $x, y \in \ker T$ and $\alpha, \beta \in \mathbb{K}$, then

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) = 0,$$

hence $\alpha x + \beta y \in \ker T$.

We now show that $\ker T$ is closed in H_1. Let $(x_n) \subset \ker T$ be a sequence such that $x_n \to x$ in H_1. By continuity of T we have $T(x_n) \to T(x)$ in H_2. Since $T(x_n) = 0$ for all n, it follows that $T(x) = 0$, and therefore $x \in \ker T$. Thus $\ker T$ is closed.

Finally, every closed subspace of a Hilbert space is complete with the induced inner product. Hence $\ker T$, endowed with the restriction of the inner product of H_1, is itself a Hilbert space. □

Definition A.2.12 (Dual space of a normed space)**.** *Let V be a normed vector space. Its (topological) dual space is defined as*

$$V' := \{F : V \to \mathbb{K} \mid F \textit{ is linear and continuous}\}.$$

The dual space V' becomes a normed vector space when endowed with the operator norm

$$\|F\|_{V'} := \sup_{\|v\|_V \leq 1} |F(v)| = \sup_{v \neq 0} \frac{|F(v)|}{\|v\|_V}, \qquad F \in V'.$$

Moreover, the duality pairing *between V' and V is the map*

$$\langle \cdot, \cdot \rangle : V' \times V \to \mathbb{K}, \qquad \langle F, v \rangle := F(v).$$

Remark A.2.13. *If V is an infinite–dimensional normed vector space, then there exist linear functionals on V that are not continuous. Such functionals can be constructed using a Hamel basis of V and therefore rely on the axiom of choice. As a consequence, the algebraic dual of V is strictly larger than its topological dual. We refer to Diestel [5] for a detailed discussion of this distinction and related constructions.*

Definition A.2.14 (Bilinear forms and related properties)**.** *Let V and W be vector spaces over a field $\mathbb{K}$. A mapping*

$$a : V \times W \to \mathbb{K}$$

is called a bilinear form if it is linear in each argument separately. The bilinear form a is said to satisfy the following properties, when applicable:

1. *(Boundedness) If V and W are normed vector spaces, a is bounded if there exists a constant $M > 0$ such that*

$$|a(u, v)| \leq M \|u\|_V \|v\|_W \qquad \forall u \in V,\ \forall v \in W.$$

2. *(Symmetry) If $V = W$, the bilinear form a is symmetric if*

$$a(u, v) = a(v, u) \qquad \forall u, v \in V.$$

3. *(Coercivity) If $V = W$ is a normed vector space, the bilinear form a is coercive if there exists a constant $\alpha > 0$ such that*

$$a(v, v) \geq \alpha \|v\|_V^2 \qquad \forall v \in V.$$

A.3 Function spaces and weak derivatives

In order to formulate variational problems and to analyze finite element approximations, it is necessary to work within suitable spaces of functions that encode both regularity and integrability properties. Before introducing the notion of weak derivatives and Sobolev spaces, we begin by recalling some classical spaces of continuous and differentiable functions.These spaces provide the natural starting point for the functional framework of the finite element method and motivate the need for more general notions of differentiability. The exposition of the Chapter 1, [1] has been followed.

Definition A.3.1 (Lipschitz continuous function). *Let $\Omega \subset \mathbb{R}^2$ be a nonempty set. A function $f : \Omega \to \mathbb{R}$ is said to be* Lipschitz continuous *on Ω if there exists a constant $L \geq 0$ such that*

$$|f(x) - f(y)| \leq L\,|x - y| \qquad \forall\, x, y \in \Omega,$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^2$.

Definition A.3.2 (Boundary of a set). *Let $\Omega \subset \mathbb{R}^n$ be an open set. The boundary of Ω, denoted by $\partial\Omega$, is defined as*

$$\partial\Omega := \overline{\Omega} \setminus \Omega,$$

where $\overline{\Omega}$ denotes the closure of Ω in $\mathbb{R}^n$.

Definition A.3.3 (Lipschitz boundary in $\mathbb{R}^2$). *Let $\Omega \subset \mathbb{R}^2$ be an open set. We say that Ω has a Lipschitz boundary if for every $x_0 \in \partial\Omega$ there exist a radius $r > 0$, a rigid transformation*

$$\phi_{x_0} : \mathbb{R}^2 \to \mathbb{R}^2$$

(a translation composed with a rotation) such that $\phi_{x_0}(x_0) = 0$, and a Lipschitz continuous function

$$\varphi_{x_0} : \mathbb{R} \to \mathbb{R},$$

for which

$$\phi_{x_0}(\Omega) \cap B_r(0) = \{(x, y) \in B_r(0) : y > \varphi_{x_0}(x)\},$$

and

$$\phi_{x_0}(\partial\Omega) \cap B_r(0) = \{(x, y) \in B_r(0) : y = \varphi_{x_0}(x)\}.$$

Definition A.3.4 (Space of continuously differentiable functions). *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $k \in \mathbb{N}$. A function $v : \Omega \to \mathbb{R}$ is said to belong to $C^k(\Omega)$ if all its partial derivatives of order up to k exist and are continuous on Ω.*

More precisely, for a multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ with $|\alpha| := \alpha_1 + \cdots + \alpha_n \leq k$, the derivative

$$D^\alpha v := \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

exists and belongs to $C^0(\Omega)$. In the case $k = \infty$, the space $C^\infty(\Omega)$ is defined as the set of functions whose partial derivatives of all orders exist and are continuous on Ω.

Definition A.3.5 (Space $C^k(\overline{\Omega})$). *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $k \in \mathbb{N}$. We define*

$$C^k(\overline{\Omega}) := \left\{ v \in C^k(\Omega) \;:\; \text{for every multi-index } \alpha \text{ with } |\alpha| \leq k,\ D^\alpha v \text{ extends continuously to } \overline{\Omega} \right\}.$$

That is, a function v belongs to $C^k(\overline{\Omega})$ if $v \in C^k(\Omega)$ and for each multi-index α with $|\alpha| \leq k$ there exists a continuous function

$$g_\alpha : \overline{\Omega} \to \mathbb{R}$$

such that $g_\alpha = D^\alpha v$ in Ω.

The spaces $C^0(\Omega)$ and $C^k(\Omega)$ describe functions with increasing degrees of classical regularity inside the domain Ω. In many analytical constructions, however, it is also important to control the behavior of functions near the boundary of the domain or to restrict attention to functions that vanish outside a compact subset of Ω. This motivates the introduction of spaces of smooth functions with compact support, which will play a central role in the definition of weak derivatives and Sobolev spaces.

Definition A.3.6 (Compactly supported smooth functions)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $k \in \mathbb{N} \cup \{\infty\}$. The space $C_c^k(\Omega)$ is defined as*

$$C_c^k(\Omega) := \{\, v \in C^k(\Omega) \mid \operatorname{supp}(v) \subset \Omega \text{ is bounded}\},$$

where the support of v is defined as the closure in $\mathbb{R}^n$ of the set

$$\{x \in \Omega \mid v(x) \neq 0\}.$$

Functions in the space $C_c^k(\Omega)$ are classical C^k functions defined on the open set Ω whose nonzero values are confined to a strictly interior region of the domain. More precisely, a function $v \in C_c^k(\Omega)$ is k times continuously differentiable on Ω and, in addition, the closure of the set where v does not vanish remains entirely contained in Ω. As a consequence, such functions are identically zero in a neighbourhood of the boundary $\partial\Omega$, which makes them particularly well suited for integration by parts and for the definition of weak derivatives.

Proposition A.3.7 (Functions in $C_c^k(\Omega)$ vanish near the boundary)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $k \in \mathbb{N} \cup \{\infty\}$. If $v \in C_c^k(\Omega)$, then*

$$\operatorname{supp}(v) \cap \partial\Omega = \varnothing.$$

In particular, there exists an open set $U \subset \mathbb{R}^n$ satisfying $\partial\Omega \subset U$ such that

$$v = 0 \quad \text{on } U \cap \Omega.$$

Proof. Let $v \in C_c^k(\Omega)$. By Definition A.3.6,

$$\operatorname{supp}(v) = \overline{\{x \in \Omega \mid v(x) \neq 0\}} \subset \Omega.$$

Since $\partial\Omega = \overline{\Omega} \setminus \Omega$, it follows that $\Omega \cap \partial\Omega = \varnothing$, and hence

$$\operatorname{supp}(v) \cap \partial\Omega \subset \Omega \cap \partial\Omega = \varnothing,$$

which proves the first claim.

Because $\operatorname{supp}(v)$ is a closed set in $\mathbb{R}^n$, its complement $\mathbb{R}^n \setminus \operatorname{supp}(v)$ is open. By the previous step, $\partial\Omega \subset \mathbb{R}^n \setminus \operatorname{supp}(v)$, so we may define the open set

$$U := \mathbb{R}^n \setminus \operatorname{supp}(v),$$

which satisfies $\partial\Omega \subset U$. Moreover, by the definition of support, one has $v = 0$ on U, so also on $\Omega \cap U$, proving the existence of an open neighbourhood of the boundary where v vanishes. □

Therefore, functions in $C_c^k(\Omega)$ are not merely zero on $\partial\Omega$, but are identically zero in a whole neighbourhood of the boundary inside the domain.

The spaces $C^k(\Omega)$ introduced above impose strong pointwise regularity requirements. However, in many analytical and numerical contexts, it is sufficient to control functions in an integral sense rather than pointwise. This leads naturally to the introduction of Lebesgue spaces, which provide a framework for measuring the size of functions through integral norms and play a central role in modern analysis and partial differential equations.

Remark A.3.8. *The notions of measurability and Lebesgue integration used throughout this work are standard and can be found in classical references such as [13]. A detailed treatment of these topics is beyond the scope of this thesis. For the type of domains arising in finite element applications—namely open and bounded subsets of $\mathbb{R}^n$ with piecewise smooth boundaries—measurability is always satisfied, and the Lebesgue integral coincides with the Riemann integral whenever the latter is well defined. For this reason, no further attention will be paid to these aspects.*

Definition A.3.9 (Lebesgue L^p norm)**.** *Let $\Omega \subset \mathbb{R}^n$ be a measurable set and let $1 \leq p < \infty$. For a measurable function $v : \Omega \to \mathbb{R}$, the Lebesgue L^p norm is defined as*

$$\|v\|_{L^p(\Omega)} := \left(\int_\Omega |v(x)|^p \, \mathrm{d}x \right)^{1/p},$$

provided that the integral is finite.

Definition A.3.10 (Lebesgue spaces)**.** *Let $\Omega \subset \mathbb{R}^n$ be a measurable set and let $1 \leq p < \infty$. Consider the set*

$$\mathcal{L}^p(\Omega) := \left\{ v : \Omega \to \mathbb{R} \;\middle|\; v \text{ is measurable and } \|v\|_{L^p(\Omega)} < \infty \right\}.$$

Two functions $u, v \in \mathcal{L}^p(\Omega)$ are said to be equivalent if

$$\|u - v\|_{L^p(\Omega)} = 0.$$

The Lebesgue space $L^p(\Omega)$ is defined as the set of equivalence classes

$$L^p(\Omega) := \mathcal{L}^p(\Omega)/\sim,$$

where $\sim$ denotes this equivalence relation.

Remark A.3.11. *For $1 \leq p < \infty$, the mapping $\|\cdot\|_{L^p(\Omega)}$ defines a norm on $L^p(\Omega)$, and the resulting normed space is complete. Therefore, $L^p(\Omega)$ is a Banach space. A detailed proof of these results can be found, for instance, in Chapter 4 of [2].*

Remark A.3.12. *In the case $p = 1$, the equivalence relation $\|u - v\|_{L^1(\Omega)} = 0$ means that*

$$\int_\Omega |u(x) - v(x)| \, \mathrm{d}x = 0.$$

This implies that $u(x) = v(x)$ for all $x \in \Omega$ except possibly on a set of Lebesgue measure zero. For this reason, functions satisfying this property are said to be equal almost everywhere *in Ω from now on.*

Definition A.3.13 (Locally integrable functions). *Let $\Omega \subset \mathbb{R}^n$ be an open set. A function $v : \Omega \to \mathbb{R}$ is said to be locally integrable on Ω if $v \in L^1(K)$ for every compact set $K \subset \Omega$. In that case we write,*

$$v \in L^1_{\mathrm{loc}}(\Omega).$$

Proposition A.3.14 (Local integrability of L^2 functions). *Let $\Omega \subset \mathbb{R}^n$ be an open set. Then*

$$L^2(\Omega) \subset L^1_{\mathrm{loc}}(\Omega).$$

Proof. Let $v \in L^2(\Omega)$ and let $K \subset \Omega$ be compact. Since K is bounded, its Lebesgue measure $|K|$ is finite. Consider the constant function 1 on K. Then $1 \in L^2(K)$ and

$$\|1\|^2_{L^2(K)} = \int_K 1^2 \,\mathrm{d}x = |K|.$$

Applying the Cauchy–Schwarz inequality in $L^2(K)$ to the functions $|v|$ and 1, we obtain

$$\int_K |v(x)| \,\mathrm{d}x = \int_K |v(x)| \cdot 1 \,\mathrm{d}x \le \left(\int_K |v(x)|^2 \,\mathrm{d}x\right)^{1/2} \left(\int_K 1^2 \,\mathrm{d}x\right)^{1/2}.$$

Hence

$$\int_K |v(x)| \,\mathrm{d}x \le \|v\|_{L^2(K)} \, |K|^{1/2} < \infty,$$

because $v \in L^2(\Omega)$ implies $v \in L^2(K)$. Therefore $v \in L^1(K)$ for every compact $K \subset \Omega$, i.e. $v \in L^1_{\mathrm{loc}}(\Omega)$. □

At this point, with Lebesgue spaces already introduced, only one additional concept is needed in order to define Sobolev spaces: the notion of weak derivative. Weak differentiation extends the classical concept of differentiation to functions that may not be differentiable in the usual sense, while preserving the essential properties required for analysis.

Definition A.3.15 (Weak derivative). *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $v \in L^1_{\mathrm{loc}}(\Omega)$. A function $w \in L^1_{\mathrm{loc}}(\Omega)$ is called the weak derivative of v with respect to the variable x_i if*

$$\int_\Omega v \,\frac{\partial \varphi}{\partial x_i} \,\mathrm{d}x = -\int_\Omega w \,\varphi \,\mathrm{d}x \quad \textit{for all } \varphi \in C_c^\infty(\Omega).$$

In this case, w is denoted by $\partial_{x_i} v$.

This definition is motivated by the classical integration by parts formula. If $v \in C^1(\Omega)$, then for any test function $\varphi \in C_c^\infty(\Omega)$ one has

$$\int_\Omega v \,\frac{\partial \varphi}{\partial x_i} \,\mathrm{d}x = \int_{\partial\Omega} v \,\varphi \, n_i \,\mathrm{d}S - \int_\Omega \frac{\partial v}{\partial x_i} \,\varphi \,\mathrm{d}x,$$

where $n = (n_1, \ldots, n_n)$ denotes the outward unit normal vector to $\partial\Omega$. Since φ has compact support in Ω, it vanishes in a neighbourhood of $\partial\Omega$, and therefore the boundary term is zero. Hence,

$$\int_\Omega v \,\frac{\partial \varphi}{\partial x_i} \,\mathrm{d}x = -\int_\Omega \frac{\partial v}{\partial x_i} \,\varphi \,\mathrm{d}x.$$

The definition allows functions with limited regularity to possess derivatives that are meaningful in an integral sense. In particular, any function that is classically differentiable almost everywhere in Ω, and whose classical derivative is locally integrable, admits a weak derivative in the sense of the above definition. This fact is fundamental in the finite element method, since finite element approximations are typically only piecewise smooth and may fail to be differentiable at element interfaces, while still possessing well-defined weak derivatives.

In many applications, however, it is not sufficient to consider weak derivatives in isolation; one must also require that both the function and its derivatives satisfy suitable integrability properties. This leads naturally to the introduction of Sobolev spaces, which collect functions whose weak derivatives exist up to a given order and belong to Lebesgue spaces. These spaces provide the appropriate functional framework for the variational formulation of partial differential equations.

Definition A.3.16 (Sobolev space $H^1(\Omega)$)**.** *The Sobolev space $H^1(\Omega)$ is defined as*

$$H^1(\Omega) := \left\{ v \in L^2(\Omega) \;\middle|\; \partial_{x_i} v \in L^2(\Omega) \text{ for } i = 1, \ldots, n \right\},$$

where $\partial_{x_i} v$ denotes the weak derivative of v with respect to x_i.

Remark A.3.17. *By Proposition A.3.14, the assumptions $v \in L^2(\Omega)$ and $\partial_{x_i} v \in L^2(\Omega)$ in Definition A.3.16 imply $v, \partial_{x_i} v \in L^1_{\text{loc}}(\Omega)$, so that the definition A.3.15 of weak derivatives is consistent in this setting.*

Roughly speaking, functions in the Sobolev space $H^1(\Omega)$ are functions that are square integrable over Ω and whose first-order weak derivatives are also square integrable. In contrast with classical C^1 functions, elements of $H^1(\Omega)$ need not be continuous or differentiable pointwise everywhere.

Definition A.3.18 (Sobolev H^1 norm)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set. For a function $v \in H^1(\Omega)$, the Sobolev H^1 norm is defined as*

$$\|v\|_{H^1(\Omega)} := \left(\|v\|^2_{L^2(\Omega)} + \sum_{i=1}^{n} \|\partial_{x_i} v\|^2_{L^2(\Omega)} \right)^{1/2},$$

Remark A.3.19. *With the definition of $H^1(\Omega)$ given above, it can be shown that the Sobolev H^1 norm introduced in Definition A.3.18 is well defined and consistent with the weak differentiability framework. Moreover, endowed with this norm, $H^1(\Omega)$ is a complete normed space, and therefore a Banach space. A detailed proof of these results can be found, for instance, in Chapter 8 of [2].*

After introducing the Sobolev space $H^1(\Omega)$, the regularity requirements on the functions that may be considered as solutions of certain problems are relaxed. This allows a larger class of functions to be taken into account in variational formulations. On the other hand, this relaxation also implies a loss of pointwise control, in particular on the boundary of the domain, since pointwise evaluation is no longer meaningful in the classical sense for functions belonging only to $H^1(\Omega)$. As a result, boundary conditions cannot be imposed pointwise and must be interpreted in a weaker sense.

A classical result in Sobolev space theory (see, for instance, Chapter 8 of [2]) states that $C^\infty(\Omega)$ is dense in $H^1(\Omega)$. Consequently, every function $v \in H^1(\Omega)$ can be approximated in the H^1 norm by a sequence of smooth functions. For such smooth functions, boundary values are well defined in the classical sense. This observation provides a useful intuition to understand boundary behaviour for functions in $H^1(\Omega)$, as their boundary values may be interpreted as limits, in a suitable sense, of the boundary values of smooth approximations.

This idea is made precise through the notion of trace. For sufficiently regular domains, as physical problems have, there exists a linear and continuous operator, called the trace operator, which assigns to each function in $H^1(\Omega)$ a boundary value in a suitable function space.

Definition A.3.20 (Sobolev space $H_0^1(\Omega)$)**.** *The Sobolev space $H_0^1(\Omega)$ is defined as the closure of $C_c^\infty(\Omega)$ in $H^1(\Omega)$, that is,*

$$H_0^1(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{H^1(\Omega)}}.$$

Equivalently, a function $v \in H^1(\Omega)$ belongs to $H_0^1(\Omega)$ if and only if there exists a sequence $(v_k) \subset C_c^\infty(\Omega)$ such that

$$\|v_k - v\|_{H^1(\Omega)} \longrightarrow 0.$$

Remark A.3.21. *Since every function in $C_c^\infty(\Omega)$ is smooth and has compact support in Ω, it vanishes identically in a neighbourhood of the boundary $\partial\Omega$ and in a well defined way. Therefore, elements of $H_0^1(\Omega)$ can be approximated in the H^1 norm by smooth functions that are identically zero near the boundary. This provides an intuitive interpretation of $H_0^1(\Omega)$ as the subset of H^1 functions that satisfy homogeneous Dirichlet boundary conditions in a weak sense.*

Definition A.3.22 (Fractional Sobolev space $H^{1/2}(\partial\Omega)$ in $\mathbb{R}^2$)**.** *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with Lipschitz boundary. The fractional Sobolev space $H^{1/2}(\partial\Omega)$ is defined as*

$$H^{1/2}(\partial\Omega) := \left\{ u \in L^2(\partial\Omega) \;\middle|\; [u]_{H^{1/2}(\partial\Omega)} < \infty \right\},$$

where the Gagliardo seminorm is given by

$$[u]^2_{H^{1/2}(\partial\Omega)} := \int_{\partial\Omega}\int_{\partial\Omega} \frac{|u(x) - u(y)|^2}{|x-y|^2}\, dS_x\, dS_y.$$

The norm in $H^{1/2}(\partial\Omega)$ is defined by

$$\|u\|^2_{H^{1/2}(\partial\Omega)} := \|u\|^2_{L^2(\partial\Omega)} + [u]^2_{H^{1/2}(\partial\Omega)}.$$

Remark A.3.23 (Intuition behind $H^{1/2}(\partial\Omega)$)**.** *The space $H^{1/2}(\partial\Omega)$ can be interpreted as the natural trace space of $H^1(\Omega)$. Functions in $H^{1/2}(\partial\Omega)$ need not be classically differentiable along the boundary, but they possess a fractional regularity that measures how oscillations behave at small scales. The Gagliardo seminorm controls the average squared variation of u relative to the distance between points. Roughly speaking, this condition prevents jumps and excessive oscillations.*

Theorem A.3.24 (Trace Theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with Lipschitz boundary. Then there exists a bounded linear operator*

$$\gamma : H^1(\Omega) \longrightarrow H^{1/2}(\partial\Omega)$$

such that:

1. *For every* $u \in H^1(\Omega) \cap C^0(\overline{\Omega})$,
$$\gamma(u) = u|_{\partial\Omega},$$
that is, γ *coincides with the classical trace on continuous functions.*

2. *There exists a constant* $C > 0$*, depending only on* Ω*, such that*
$$\|\gamma(u)\|_{H^{1/2}(\partial\Omega)} \leq C \|u\|_{H^1(\Omega)} \qquad \forall u \in H^1(\Omega).$$

3. *The operator* γ *is surjective.*

The operator γ *is called the trace operator.*

Remark A.3.25. *A complete proof of the Trace Theorem can be found, for example, in Section 5.5 of [9].*

Remark A.3.26 (Intuition behind the Trace Theorem)**.** *The Trace Theorem states that functions in* $H^1(\Omega)$ *admit well–defined boundary values in the weaker fractional space* $H^{1/2}(\partial\Omega)$*. Intuitively, an* $H^1(\Omega)$ *function has one square–integrable derivative in the interior, which means that its oscillations are controlled in an averaged sense. Although such functions need not be continuous up to the boundary, their interior regularity is strong enough to prevent wild oscillations near* $\partial\Omega$*, so that a meaningful boundary value can still be assigned. However, this boundary value cannot in general belong to* $H^1(\partial\Omega)$*.*

Theorem A.3.27 (Poincaré inequality in $H^1_0(\Omega)$)**.** *Let* $\Omega \subset \mathbb{R}^n$ *be a bounded open set. Then there exists a constant* $C_P > 0$*, depending only on* Ω*, such that*
$$\|v\|_{L^2(\Omega)} \leq C_P \|\nabla v\|_{L^2(\Omega)} \qquad \forall v \in H^1_0(\Omega).$$

Reference. A proof can be found, for instance, in Chapter 9 of [2]. □

Lemma A.3.28 (Vanishing of a continuous function with zero action on test functions)**.** *Let* $\Omega \subset \mathbb{R}^n$ *be an open set and let* $g \in L^1_{loc}(\Omega)$ *satisfy*
$$\int_\Omega g(x)\,\varphi(x)\,dx = 0 \qquad \forall \varphi \in C^\infty_c(\Omega).$$
Then
$$g(x) = 0 \qquad \textit{for all } x \in \Omega.$$

Reference. A proof can be found, for instance, in Chapter 4 of [2]. □