

# DISCONTINUOUS GALERKIN FINITE ELEMENT METHOD FOR THE STOKES PROBLEM USING RECTANGULAR MESHES

Dissertation Submitted To

UNIVERSITY OF KWAZULU-NATAL

in partial fulfilment of the requirements  
of the award of Master of Science Degree

in

APPLIED MATHEMATICS

by

WONDER KUDZO EKPE

COLLEGE OF AGRICULTURE, ENGINEERING AND SCIENCE  
School of Mathematics Statistics and Computer Science



September 2020

Supervisor: Dr Kanagaratnam Arunakirinathar

## Abstract

In this paper, we present the primal formulation of the Discontinuous Galerkin Method for the Stokes problem using the primitive variables - velocity and pressure. In this formulation, non-overlapping conforming elements are used such that both velocity and pressure fluxes are considered to be discontinuous over the inter-element interfaces. However, to enforce *artificial* continuity, we introduce a jump term, equipped with a penalty constant  $\gamma > 0$ , which penalises the jump of numerical fluxes across elements. For the domain of the problem set, we demand that the velocity  $u \in [H_0^1(\Omega)]^d$  and for the pressure,  $p \in L_0^2(\Omega)$ . Bilinear forms for both velocity and pressure are defined such that symmetry is ensured in terms of the numerical fluxes  $u$  and  $v$ , thus giving rise to our type of formulation known in literature as Symmetric Interior Penalty Galerkin (SIPG) Method. Using quadrilateral meshes, we state and prove the inf-sup condition and velocity and pressure error estimates. A simple case of lid-driven cavity flow in a square domain is studied for our formulation and the results are compared with results reported by others in the literature. The study also establishes a simple way of estimating the convergence rate of our scheme, which we found to be optimal in the mesh size.

## Acknowledgement

I thank God for his love and protection for my life throughout this studies. This piece of project would not have been completed without supports from few individuals. I would like to thank my supervisor — for his advice and directions towards the success of this work. I also thank — for their support in diverse ways. God bless you all.

# Dedication

I dedicate this project to my family - Ekpe and Abofra.

# Contents

<b>Abstract</b>	<b>i</b>
<b>List of Tables</b>	<b>v</b>
<b>List of Figures</b>	<b>vi</b>
<b>1 Chapter One: Introduction</b>	<b>1</b>
1.1 Background and Literature Review . . . . .	1
1.2 Application of the DG Method . . . . .	9
1.3 Aims and Objectives of the Study . . . . .	10
1.4 Overview . . . . .	10
<b>2 Chapter Two: Mathematical Preliminaries</b>	<b>12</b>
2.1 The Navier-Stokes Equation . . . . .	12
2.2 Basic Functional Analysis . . . . .	13
2.3 The Galerkin Finite Element Method . . . . .	18
2.4 Weak Formulation of Stokes Equations . . . . .	19
2.5 Formulation of Discontinuous Galerkin FEM . . . . .	25
2.5.1 Discretization . . . . .	26
2.5.2 Mesh Assumptions . . . . .	30
2.5.3 The Discrete Problem . . . . .	31
2.5.4 Multiplicative trace and Inverse Inequalities . . . . .	36
<b>3 Chapter Three: Numerical and Error Analysis</b>	<b>40</b>
3.1 Tools for Computing Integral Terms . . . . .	40
3.1.1 Reference Element . . . . .	40
3.2 Numerical Quadrature for a Rectangle . . . . .	45
3.2.1 Basis Functions . . . . .	46
3.2.2 Effect of the Size of Penalty Parameter $\gamma$ . . . . .	49
3.3 Error Analysis of the Discontinuous Galerkin Method . . . . .	50
3.3.1 A-priori error estimates . . . . .	56
3.4 Numerical Examples . . . . .	62
<b>4 Chapter Four: Conclusions and Future Work</b>	<b>69</b>
<b>Bibliography</b>	<b>71</b>

## **List of Tables**

1	Comparison of convergence of different errors at various iterations . . . . .	67
---	---	----

## List of Figures

2.1	A Two-dimensional Domain showing Boundaries . . . . .	13
2.2	Discretization of the Solution Domain . . . . .	27
3.1	Bilinear Elements in Reference Coordinates (Left) and in Physical Coordinates (Right). . . . .	41
3.2	Basis bilinear functions defined over a rectangle . . . . .	48
3.3	A Two-D Lid-driven Cavity problem . . . . .	62
3.4	Streamline plots for Stokes solutions showing conner singularities . . . . .	63
3.5	Streamline plot of velocity field (left) and mesh plot of pressure field (right)	64
3.6	Contour (left) and mesh plots (right) of the estimated energy error . . . . .	64
3.7	Plots of upper error bound (red *- line), approximation error (solid black line) and residual error (blue -o- line) . . . . .	66
3.8	Error plot of $\ln \  u - u_h \ $ against $\ln h$ . . . . .	68

# Chapter One: Introduction

## 1.1 Background and Literature Review

Since the dawn of computers, the use of numerical approach in solving mathematical problems became much easier. The use of these numerical methods become necessary for problems to which there exists no (stable or unique) analytical solution. Unfortunately, most real-life phenomena such as fluid flows rarely do have analytical solutions, except for those considered in simple flow domains. The difficulties in obtaining analytical solutions are due to the fact that most of these real-life situations are modelled by equations that are non-linear in nature. In the study of fluid flows, the well-known Navier-Stokes Equations (NSEs), which are a couple of partial differential equations (PDEs), are used to model flows in different domains and media. These equations are nonlinear in nature and hence are difficult to solve analytically.

To overcome these difficulties, numerical methods are often employed to approximate solutions to these non-linear equations. The basic principle underlying all the numeral methods is the idea of discretizing (partitioning) of the problem domain and finding the solution in each of these discretized parts. There are a number of these methods. Amongst these are, Finite Difference Method (FDM), Finite Volume Method (FVM) and Finite Element Method (FEM). In the FDM, the domain under consideration is partitioned into grids of discrete points and the governing equation is approximated by substituting occurring derivatives by difference quotients between these gridpoints [67]. This results into system of algebraic equations that are then solved using appropriate approach. Though the finite

difference schemes are easy to implement, it is more difficult in implementing in complex domains. Unlike the FDM, the domain in the FVM is divided into subdomains called volumes. The physical conservation laws such as mass, momentum and energy conservation are applied to each subdomain, and as in the case of the FDM, to large systems of equations which have to be solved in a similar fashion. One advantage of this method is that, it caters for problems with discontinuities [29]. It is however not very reliable with respect to stability and convergence analyses [43]. Lastly, the FEM, which will be employed in this work, has a similar formulation as the FVM. Its basic idea is to divide the geometry of the domain into subdomains (called elements) and to establish an approximation equation for each of the elements. The algebraic equations are then assembled and solved together to estimate the solution to the governing equation. [65]

Over the years researchers have used one or the other of these numerical methods in approximating solutions to the Navier-Stokes equation. The FDM has been used by a number of researchers to model fluid flows. These include Chorin [15], Yongho, *et al* [14]. In most of these works, the authors have been able to use the Artificial Compressibility (AC) method to deal with the incompressibility condition imposed in the NSEs, equation (2.3) [43]. Other researchers also have been able to use pressure based iteration schemes, in which marker-and-cell (MAC); and Semi-Implicit Method for Pressure-Linked Equations (SIMPLE/R) methods are used (see for example [15]). Projection method have been used by other writers such as, Yongho *et al* [14], Motoyoshi [44] and Nahid [60]. Marchi *et al* [57], Mohammad *et al* [33] and Aikkara & Aboobacker [1], all employed Fi-

nite Volume method to tackle different flow problems. In the implementation of the Finite Element Method, the following researchers resorted to the general Galerkin formulation - Montlaur [23], Ducasse [29], Omari [62] and Jiajian [43]. When the governing equation is discretized using the general Galerkin method, the resulting system of algebraic linear equations pose additional complications. This is due to the absence of the pressure term in the incompressibility equation (2.3) (see for example [71]). To overcome this problem, some authors such as Sharma [72], make use of the Penalty Galerkin Method, which segregates the computation of the velocity and pressure. Galerkin weighted residual method of finite element was employed by [80]. Before we move our attention from the generalities and into more specific review of more related literature, it suffices to mention that, in all of the above-mentioned methods, different formulations of the Navier-Stokes equations were used. These include, the primitive variable formulation (velocity-pressure), velocity-vorticity and stream-function formulations. Each of these formulations have their respective strengths and weaknesses during implementation stage. (see for example [60], [36], [79], [34] and [53]).

The DGM exhibits better stability properties around sharp gradients such as boundary or interior layers and also cater for discontinuities which are often present in the analytical solution of convection/transport dominated PDE problems [35].

Since its introduction by Reed and Hill in 1973, a number of papers have been published on the DGM. Some of the modern works are those by Cockburn *et al* [20]. An extensive review of the basic idea, formulation and implementation was included in this review pa-

per. Another work by Xu and Shu [85] also explained applications of the DGM to partial differential equations (PDEs) with higher order spatial derivatives. Again, our NSEs fall under these PDEs.

The DGM exhibits better stability properties around sharp gradients such as boundary or interior layers and also cater for discontinuities which are often present in the analytical solution of convection/transport dominated PDE problems [35].

Since its introduction by Reed and Hill in 1973, a number of papers have been published on the DGM. Some of the modern works are those by Cockburn *et al* [20]. An extensive review of the basic idea, formulation and implementation was included in this review paper. Another review paper by Xu and Shu [85] also explained application of the DGM to partial differential equations (PDEs) with higher order spatial derivatives. Again our NSEs fall under these PDEs.

The development of the DG method has been evolutionary. Different authors have come out with varying formulations, mostly based on the physical problem under consideration. Some of these techniques include Runge-Kutta Discontinuous Galerkin (RKDG) by Cockburn and Shu [20], the Local Discontinuous Galerkin (LDG), also by Cockburn and Shu [18] and Compact Discontinuous Galerkin (CDG) by Peraire and Persson [63]. In 1977, Baker studied a different variant of the CDG method known as the Interior Penalty (IP) method. The IP was also studied extensively by Arnold in 1982 [27]. Others are the Embedded DG (by Güzey *et al.*, [40]) and more recently the Hybridizable Discontinuous

Galerkin (HDG), by Cockburn and Gopalakrishnan in 2005. In the following paragraph, we discuss the characteristics of some of these different formulations.

The RKDG was developed to handle the earlier difficulties of efficiency and stability with the DG method. This was done by restricting the Courant number [73]. It was introduced originally for second order systems, mainly schemes in time and space. However, The RKDG was improved and applied to high-order PDEs in solid dynamics and fluid mechanics. In the LDG formulation, the primary variable, with its derivative are used as the independent unknown to be solved. By this, the governing equation is reduced to a number of first-order PDEs, which are then solved individually by applying the standard DG method. The LDG formulation in fluid mechanics involves vorticity, velocity and pressure. Later the vorticity was replaced by introduction of the lifting operators, which leads to the velocity-pressure formulation [23]. This coupled formulation has been applied, especially to the Stokes and Oseen problems, and also to the Navier-Stokes equations (see, for e.g., [17]). Even though the LDG has the benefit of good conservative schemes, one major drawback is loss of compactness due to large degree of freedoms, as a result of the introduction of additional unknowns. To circumvent this loss of compactness, CDG method was introduced. The CDG method is similar to the LDG, especially, in lower-order schemes. The difference between the two becomes more pronounced in higher dimensions. The CDG uses local lifting operators to eliminate coupling between degrees of freedom of non-neighbouring elements [58]. As a result, the schemes are more compact and produce sparser matrix systems, resulting in a better stability than that from the LDG [63]. The

IP method also shares commonalities with the CDG. They both produce symmetric and coercive bilinear forms by inducing compact formulations. However, unlike the CDG, the IP method does not make use of lifting operators. This makes it more straight-forward to implement and hence cutting down on computational cost. In this work, we make use of the IP method formulation due to its simplicity and ease of implementation.

Most frequently, all works published on the variant DG formulations by authors such as Baker in 1977, Wheeler in 1978 and by Arnold in 1982, are all referred to as Interior Penalty Galerkin (IPG) method, albeit, their formulations do not depend on the development of the standard DG method [27]. The main stand-out feature of the IPG is that, in its formulation, the continuity requirement for the conforming finite element solution in the DG is replaced by an interior penalty. This is to ensure continuity of numeric fluxes over inter-elements boundaries.

The DG methods however also do exhibit certain weakness: There is always a possibility of introduction of a large number of degrees of freedom, which leads to large stencils. They cause reduced sparsity of the global system due to a rather intense coupling across element interfaces. There are however, few remedies proposed to circumvent these disadvantage. One of such technique is by hybridization as used in published works including [19] and [11]. In the hybridization approach, additional unknowns are introduced to the inter-element faces, thereby reducing the coupling across these elements [see, e.g., [81]].

After few years of almost no activity since its introduction, the 1990s saw more works done on the IPG method. In 1999, works by Rivi re *et al* where they introduce an in-

terior and boundary penalty to an earlier method that had been studied by [9] and [61]. Rivi re *et al* published further findings on this, where they proposed the Nonsymmetric Interior Penalty Galerkin method (NIPG). According to their work, optimality of the error estimates for this method is not guaranteed in the  $L^2(\Omega)$ -norm, using the Aubin-Nitsche approach. However, works published by Mats *et al* [51] and [22] showed ways by which optimality could be achieved using the Aubin-Nitsche method. In their paper, Mats *et al* proved that the NIPG is inf-sup stable with respect to a mesh-dependent energy norm. This was applied to quadratic and higher order polynomials on a general unstructured grids in two dimensions. A positive feature of this NIPG is that, coercivity of the bilinear forms is achieved for any value of the penalization parameter  $\gamma > 0$  [66]. We also mention, briefly, a similar but slightly different approach known in the literature as the Incomplete Interior Penalty Galerkin (IIPG) method. One major difference between the SIPG and IIPG is that, to achieve coercivity, one needs a relatively larger values of  $\gamma$  in the case of IIPG.

In this work, we consider the primal method, where the Symmetric Interior Penalty approach is used. Largely, we follow closely, the approach used by Dolejvı and Feistauer in [27]. This method has many advantages over the NIPG and IIPG. One of such advantages is that, an error estimate in the  $L^2(\Omega)$ -norm can be achieved by using the Aubin-Nitsche tricks. Many authors have published a lot of work on the SIPG method. These include Arnold in [3], Wheeler in [82], Girault *et al* [38], Rivierer *et al* [66], Dolejvı *et al* [27] and Houston *et al* [42], to list a few. The formulation of these three variants - the NIPG, the

IIPG and the SIPG - are all similar, except for a multiplicative factor that determines the sign between the inter-element fluxes. This factor is assigned values  $-1, 0$  or  $1$ . In the case of the SIGP, this value is taken to be  $-1$  (see section 2.5).

The Symmetric Interior Penalty Galerkin method has been applied to many different problems. Carrier *et al* [12] applied the SIPG method in studying the stability of the Stokes flow of arbitrary order. In their work, they modeled a heterogeneous, incompressible Stokes flow, using elements of order  $Q_k^2 - Q_{k-1}$  with hierarchical Legendre basis polynomials. Their results showed that a careful choice of the penalty parameter, which is based on local values of the viscosity, greatly affects the convergence of their discretisation and iterative methods. Some other papers in which the SIPG was employed are by Baker *et al* [6], and Karakasian and Jureidini [45]. In both of these works, the velocity field was approximated using discontinuous solenoidal piecewise functions across the inter-element boundaries, with continuous pressure. The only difference is that, whereas the former applied this method to Stokes problem, the stationery Navier-Stokes equation was the problem of interest in the latter. In their conclusion, they both found out that, with only requiring that meshes meet local quasi-uniformity assumptions, Karakasian and Jureidini achieved results that have optimal rates of convergence. Another work done by Epshteyn and Rivière [31], is where stability and convergence of the SIPG method was studied. In their work, they derived computable bounds of the penalty parameter which are needed to guarantee a stability and convergence of the system. Their result showed an improved coercivity, where there has been mesh-dependence of the constant on quantities such as

the polynomial order and the angles of the triangles formed in the 2-D domain.

Another aspect considered in this work is the elemental implementation challenges. We take a look at an appropriate geometric transformation of physical elements to a reference or master element. There is a greater freedom of choice of element types that can be used to estimate our solution in the DGM. This is possible, especially, when we relax the constraint of  $C^0$  continuity condition of the solution [55]. Owing to the property of the DGM not requiring continuity of the solution, it is possible to employ quadrilateral basis functions in the approximating the solution. These basis functions can be defined such that, the usual cross-terms of the variables, that are common in the standard CDG are excluded. Example of such element is a linear quadrilateral element. In her thesis, Maggi found out that there is an evidence of a considerable benefit in the use of quadrilateral elements in terms of accuracy and computational time. One of the purposes of this work is to test the performance of these quadrilateral elements applied to the Stokes problem. Our motivation for this part is from works published by Wirasaet *et al* in [84] and Maggi in [55].

### 1.2 Application of the DG Method

In today's fast growing demands for simulation and automation, the application and importance of DGM cannot be overstated. It is applied in gas dynamics, magneto-hydrodynamics, electro-magnetism, granular flows, meteorology, modelling of shallow water, oceanography, aeroacoustic, oil recovery simulation, semiconductor device simulation, transport of contaminant in porous media, turbo-machinery, turbulent flows, visco-elastic

flows and weather forecasting among others. It is evident from the above-mentioned works and reviews that the DGM emerges as a better choice for tackling problems where the classical FEM are not applicable to and FVM produce typically low order approximations. It is for these and many other reasons we decide on the use of the Discontinuous Galerkin Finite Element Method to solve the incompressible cavity flow problem.

### **1.3 Aims and Objectives of the Study**

The main aim of this work is to study the primal formulation of the Discontinuous Galerkin Finite Element Method. Due to the inter-elements discontinuities, our formulation focuses on the Interior Penalty (IP) approach. For the sake of stability and coercivity of the derived bilinear forms, we will employ the symmetric form of the formulation. We also will explore the approximation properties of the elements to be used to approximate the solutions. Regarding the error analysis, our interest borders on derivation of energy norms and adaptive algorithms based on residual-based a priori error estimation in space. Ultimately, this work seeks to study the convergence and inf-sup stability of the discretization for the Stokes problem.

### **1.4 Overview**

The remainder of this work is structured as follows. Chapter 2 contains the mathematical formulation of Navier-Stokes equations. We start with some preliminary discussion on functional analysis, where the functional spaces of our intended solutions are defined. The Stokes equations is derived form the formulated Navier-Stokes equation. Following this,

the weak formulation of the Stokes problem is done, based on the Galerkin Finite Element Method. This chapter concludes with a detailed formulation of the Symmetric Interior Penalty Method of the Discontinuous Galerkin Method. In chapter 3, we look at the numerical and error analysis of the discretized system. The first part consists of tools for numerical computations, including the quadrature of the chosen element type, while the second part of the chapter deals with the error analysis. A couple of numerical examples are presented and solved in the second section of this chapter. The dissertation finally ends with conclusion and suggestions for future work in chapter 4.

## Chapter Two: Mathematical Preliminaries

In this Chapter, we state the Navier-Stokes equations and formulate the Discontinuous Galerkin Finite Element method. We also include the formulation of the test problem at the last section of this chapter.

### 2.1 The Navier-Stokes Equation

The purpose of this section is to state and describe the incompressible Navier-Stokes equations, which govern the motion of Newtonian fluid. The detailed formulation however, is left out from this work.

Modelling and propagation of fluid flow is governed by the Navier-Stokes Equations (NSEs). These set of equations are derived from the principles of conservation of mass, momentum and energy. Additionally, the Reynolds Transport Theorem is used with the continuity equation to derive these equations. Here we state the conservation form of Navier-Stokes Equations, without the detailed derivation:

$$\nabla \cdot \mathbf{u} = 0 \tag{2.1}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} + f, \tag{2.2}$$

where,  $\nu = \frac{\mu}{\rho}$  is the kinematic viscosity. Written in their components form, the two-dimensional incompressible Navier-Stokes equations are:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad (2.3)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + f_x, \quad (2.4)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial y} + \nu \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) + f_y, \quad (2.5)$$

where  $u$  and  $v$  are the velocity components in the  $x$  and  $y$  directions respectively.  $\nu$  is the fluid kinematic viscosity,  $f$  is the total force, and  $\rho$  is the mass density of the fluid.

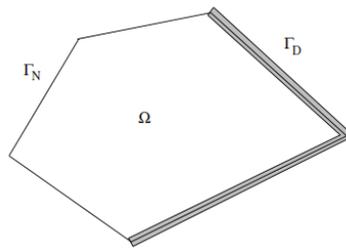


Figure 2.1: A Two-dimensional Domain showing Boundaries

Next, we present the formulation of the general Galerkin Finite Element method.

## 2.2 Basic Functional Analysis

In this section, we review some of the preliminary functional spaces, within which our solutions will be treated.

### Piecewise Polynomial Functions

For a given linear function  $v$ , we construct a vector space  $\mathbb{P}_1(I)$  on an interval  $I = [x_0, x_1]$  on the real axis as:

$$\mathbb{P}_1(I) = \{v : v(x) = c_0 + c_1x, x \in I\}, \quad (2.6)$$

where  $c_0, c_1 \in \mathbb{R}$  are constants. Basically, we can determine any  $v \in \mathbb{P}_1(I)$  uniquely by specifying values of the constants  $c_0$  and  $c_1$ . Alternatively, we can define  $\alpha_0 = v(x_0)$  and  $\alpha_1 = v(x_1)$ , and by a simple linear matrix system, these values again determine  $v \in \mathbb{P}_1(I)$  uniquely. Since a linear combination of linear polynomials results in a linear polynomial, equation (2.6) defines a vector space.

The relevance of this discussion is that, if we consider the points  $x_0$  and  $x_1$  to be nodes of some physical discretized domain with nodal values  $\alpha_0$  and  $\alpha_1$ , then we can define a new set of nodal basis  $\{\lambda_0, \lambda_1\}$  such that:

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

This therefore implies that, for 1-dimensional domain structure, any linear basis function defined, has a value of 1 at node  $x_j$  and 0 at all other nodes.

Similarly, we can define the space of continuous piecewise linear function by noting that, if one divides domain of a function  $v$  into subintervals, then it is possible to assign a linear function to each of these subintervals. This subdivision will probably introduce

discontinuities over the different subintervals. However, continuity of the function  $v$  can be ensured by using the degrees of freedom at the inter-nodal boundaries. This will be treated, in details, in the later part of this work.

From the subdivisions and using the subintervals, the space  $V_h$  for continuous piecewise linear polynomials is defined as:

$$V_h = \{v : v \in C^0(I), v|_{I_i} \in \mathbb{P}_1(I_i)\}. \quad (2.7)$$

Here,  $C^0(I)$  is the space of the continuous piecewise function defined over subintervals,  $I_i = [x_{i-1}, x_i]$ , for  $i = 1, 2, 3, \dots, n$ . The space defined in equation (2.7) ensures that all functions in it are continuous over the interval  $I$  and are linear on the subintervals  $I_i$ .

**Lemma 1.** *Any function  $v \in V_h$  can be defined uniquely by using nodal values of that function, thus:*

$$\{v(x_i)\}_{i=0}^n.$$

*The converse immediately follows that there exists a function  $v \in V_h$  for any given set of nodal values,  $\{\alpha_i\}_{i=0}^n$ .*

From lemma 1, we can introduce a set of basis functions  $\{\varphi_j\}_{j=0}^n$  such that

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

Next, we need to define interpolants, that will approximate any given function in spaces defined above. In the case of the linear polynomials, we define an interpolation function  $\Pi f \in \mathbb{P}_1(I)$  to a continuous function  $f$  as:

$$\Pi f(x) = f(x_0)\varphi_0 + f(x_1)\varphi_1.$$

The main function of the interpolation is to approximate our solution by using a pre-defined interpolant,  $\Pi f$ . The use of this approach introduces an error, which is a measure of the difference between the true solution and the approximated value ( $f - \Pi f$ ). In our analysis, we employ the  $L^2$  error norm, defined as

$$\|v\|_{L^2(I)} = \left( \int_I v^2 dx \right)^{1/2}.$$

In this section we looked at basic polynomial spaces and functional interpolations in one-dimension. The higher dimensions, with their applications to our problem will be extended to our two-dimensional case later in equation (2.22).

### Formal Definitions

Given the preliminaries in the preceding section, we now move to state formal definitions of the functional spaces and tools that will be needed for our analyses. Sketches of proofs of some of these tools will be given in the relevant sections.

**Definition 2.1.** *Consider solutions  $u, v$  and  $q$  belonging to our defined domain (including*

boundaries). We define the following basic functional spaces within which we expect to find these solutions, such that  $u, v \in V$  and  $q \in Q$ .

$$L^2(\Omega) = \{v : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} v^2 < \infty\},$$

$$H^1(\Omega) = \{v : v \in L^2(\Omega) \mid \nabla v \in L^2(\Omega)\},$$

$$H_0^1(\Omega) = \{v : v \in H^1(\Omega) \mid v = 0 \text{ on } \partial\Omega\},$$

$$L_0^2(\Omega) = \{q : q \in L^2(\Omega) \mid \int_{\Omega} q \, dx = 0\},$$

where, for  $1 \leq s < \infty$ ,  $L^s(\Omega)$ , called the Lebesgue space, is a linear space of all functions measurable on  $\Omega$  such that

$$\int_{\Omega} |u|^s \, dx < +\infty,$$

equipped with the norm

$$\|u\|_{L^s(\Omega)} = \left( \int_{\Omega} |u|^s \, dx \right)^{1/s}.$$

The function spaces  $V$  and  $Q$  are Hilbert spaces defined as:

$$V = (H_0^1(\Omega))^d, \quad d \in \{2, 3\} \text{ and } Q = L_0^2(\Omega).$$

These possess the following properties:

$$\text{inner product : } (\mathbf{u}, \mathbf{w})_V = \int_{\Omega} (\nabla \mathbf{v}, \nabla \mathbf{w}) \, dx,$$

$$\text{induced norm : } \|\mathbf{v}\|_V = \|\nabla \mathbf{v}\|_{L^2(\Omega)} = (\mathbf{v}, \mathbf{v})_V^{1/2},$$

for  $V$ , and for  $Q$ , we have:

$$\text{inner product : } (q, r)_Q = \int_{\Omega} (qr) \, dx,$$

$$\text{induced norm : } \|q\|_Q = \|q\|_{L^2(\Omega)} = (q, q)_Q^{1/2}.$$

Lastly, one other tool needed for analysis is the Cauchy-Schwarz inequality, stated as:

**Lemma 2** (Discrete Cauchy-Schwarz Inequality). *Let  $\{a_i\}_{i=1}^n$  and  $\{b_i\}_{i=1}^n$  be two sequences of real numbers. Then*

$$\left| \sum_{i=1}^n a_i b_i \right| \leq \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \left( \sum_{i=1}^n b_i^2 \right)^{1/2}. \quad (2.8)$$

### 2.3 The Galerkin Finite Element Method

The Finite Element Method (FEM) involves dividing the region of interest (domain) into a number of smaller regions called elements. The differential equations (governing equations) under consideration are then reduced to simple algebraic equations by the use of approximation/test functions, which are assigned to each element. These functions are then solved with appropriate boundary conditions and the solutions thus attained are assembled to obtain an approximate solution to the original problem [77], [16].

## 2.4 Weak Formulation of Stokes Equations

In this section, we derive the weak formulation of a model problem. To help us do this, we consider the following model problem for the Stokes equations. The Stokes equations are just same as the NSEs equations (2.1) and (2.2), but without the nonlinear convective term  $(\mathbf{u} \cdot \nabla)\mathbf{u}$ . This is a special case of the momentum equation with large viscosity. Let  $\Omega \subset \mathbb{R}^d$ , with  $d = 2, 3$ , be an open and bounded, with Lipschitz boundary. The Stokes problem is stated as: Find  $(\mathbf{u}, p) : \Omega \times \Omega \rightarrow \mathbb{R}^d \times \mathbb{R}$ , such that

$$\begin{cases} -\nu \Delta \mathbf{u} + \nabla p &= \mathbf{f} \text{ in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 \text{ in } \Omega. \end{cases} \quad (2.9)$$

The process of weak formulation involves two main steps - one, multiply the momentum equation with a test function,  $v \in V$  and the continuity equation by  $q \in Q$ ; and, secondly, integrating the system over the domain  $\Omega$  using integration by parts [70].

One last assumption on our solution  $(\mathbf{u}, p)$  is that,  $\mathbf{u} \in [C^2(\Omega) \cap C^1(\bar{\Omega})]^d$  and  $p \in [C^1(\Omega) \cap C(\bar{\Omega})]$ . Here, by  $C^1$  and  $C^2$  we mean the functions satisfy some continuity requirements at the boundaries, (i.e., at the intersections of the domain and its boundaries). If however, we represent the Dirichlet and Neumann boundary conditions (which we will discuss in the next session) by  $\Gamma_D$  and  $\Gamma_N$  respectively, then for our test functions  $\mathbf{v}$  and  $q$ , we have that  $\mathbf{v} \in (C_{\Gamma_D}^\infty(\Omega))^d$ , where,

$$C_{\Gamma_D}^\infty(\Omega) = \{v \in C^\infty(\Omega) \cap H^1(\Omega)\},$$

such that  $\exists U \subset \mathbb{R}^d$ , and  $v(x) = 0 \forall x \in U \cap \Omega$ ; and  $q \in C^\infty(\Omega)$ .

The above is necessary since we know that  $C^\infty(\Omega)$  is not a subset of  $H^1(\Omega)$ , [see [37] and [70]]. We now return to the weak formulation. We proceed as:

$$-\int_{\Omega} \nu \Delta \mathbf{u} \cdot \mathbf{v} + \int_{\Omega} (\nabla p) \cdot \mathbf{v} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, \quad (2.10)$$

$$\int_{\Omega} q \nabla \cdot \mathbf{u} = 0. \quad (2.11)$$

Using integration by parts, the first term on LHS of (2.10) gives :

$$\begin{aligned} -\int_{\Omega} \nu \Delta \mathbf{u} \cdot \mathbf{v} &= -\int_{\Omega} \nu \sum_{i=1}^d \Delta u_i v_i \\ &= \int_{\partial\Omega} \sum_{i=1}^d \nu (\nabla u_i \cdot \mathbf{n}) v_i + \int_{\Omega} \nu \sum_{i=1}^d \nabla u_i \nabla v_i \\ &= \int_{\partial\Omega} \nu (\nabla \mathbf{u} \cdot \mathbf{v}) \cdot \mathbf{n} + \int_{\Omega} (\nu \nabla \mathbf{u}) : (\nabla \mathbf{v}). \end{aligned}$$

Applying the condition  $v = 0$  on  $\partial\Omega$ , the above equation simplifies to:

$$-\int_{\Omega} \nu \Delta \mathbf{u} \cdot \mathbf{v} = \int_{\Omega} (\nu \nabla \mathbf{u}) : (\nabla \mathbf{v}), \quad (2.12)$$

where, the product  $A : B$  represents the dyadic product defined as

$$A : B = \sum_{i=1}^d \sum_{j=1}^d A_{ij} \cdot B_{ij}.$$

Next, the second term in equation (2.10) is transformed as follows:

$$\begin{aligned}\int_{\Omega} (\nabla p) \cdot \mathbf{v} &= \int_{\Omega} \sum_{i=1}^d \frac{\partial p}{\partial x_i} v_i \\ &= \int_{\Omega} \sum_{i=1}^d \frac{\partial}{\partial x_i} (p v_i) - p \frac{\partial v_i}{\partial x_i}.\end{aligned}$$

Using the product rule, the above equation gives

$$\int_{\Omega} (\nabla p) \cdot \mathbf{v} = \int_{\Omega} \nabla \cdot (p \mathbf{v}) - p \nabla \cdot \mathbf{v},$$

and finally, applying the Gaussian theorem, we have:

$$\int_{\Omega} (\nabla p) \cdot \mathbf{v} = \int_{\partial \Omega} (p \mathbf{v}) - \int_{\Omega} p \nabla \cdot \mathbf{v}.$$

Again, we observe that, the first integral of the RHS vanishes on the boundary. Thus, the pressure term in equation (2.10) becomes:

$$\int_{\Omega} (\nabla p) \cdot \mathbf{v} = - \int_{\Omega} p \nabla \cdot \mathbf{v}. \quad (2.13)$$

Substituting equations (2.12) and (2.13) into equation (2.10), we have:

$$\int_{\Omega} (\nu \nabla \mathbf{u}) : (\nabla \mathbf{v}) - \int_{\Omega} p \nabla \cdot \mathbf{v} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}.$$

Our standard weak formulation for FEM then reads: find  $(\mathbf{u}, p) \in (V \times Q)$  such that

$$\int_{\Omega} (\nu \nabla \mathbf{u}) : (\nabla \mathbf{v}) - \int_{\Omega} p \nabla \cdot \mathbf{v} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in V, \quad (2.14)$$

$$\int_{\Omega} q \nabla \cdot \mathbf{u} = 0, \quad \forall q \in Q. \quad (2.15)$$

Solutions to the weak forms equations (2.14)-(2.15) are obtained in smaller subspaces (elements) and on the boundaries as compared to what is expected in the case of the strong form, equations (2.9), which seeks to find solution for the whole domain  $\Omega$ .

However, one can simplify the weak forms further, by defining continuous bilinear forms as:

$$a(\mathbf{u}, \mathbf{v}) = \int_{\Omega} (\nu \nabla \mathbf{u}) : (\nabla \mathbf{v}),$$

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

where  $a$  and  $b$  are defined as  $a : V \times V \rightarrow \mathbb{R}$ ,  $b : V \times Q \rightarrow \mathbb{R}$ . These bilinear forms are equipped with the following norms:

$$\|a\| := \sup_{\mathbf{u}, \mathbf{v} \in V \setminus \{0\}} \frac{a(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\|_V \|\mathbf{v}\|_V},$$

$$\|b\| := \sup_{\substack{\mathbf{v} \in V \setminus \{0\} \\ p \in Q \setminus \{0\}}} \frac{b(\mathbf{v}, p)}{\|\mathbf{v}\|_V \|p\|_Q}.$$

So, by using the above definitions, the weak formulation of the Stoke problem can be

re-written as: Find  $(\mathbf{u}, p) \in V \times Q$  such that

$$\begin{cases} a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = f(\mathbf{v}), & \forall \mathbf{v} \in V, \\ b(\mathbf{u}, q) = 0, & \forall q \in Q. \end{cases} \quad (2.16)$$

To summarize the process of the weak formulation in this section, we observe that the weak solution to the system (2.16) need not necessarily be a classical solution. If however, this solution is classical, then one can say that (2.9) and (2.16) are equivalent. This is possible, especially when we require the use of homogeneous boundary conditions [2]. To see that there exist unique solution to the weak Stoke problem, we state the following theorems (without proofs).

**Theorem 2.1** (Lax-Milgram Theorem). *Let  $V$  be a Hilbert space equipped with norm defined in this section. Let  $f : V \rightarrow \mathbb{R}$  be a continuous linear functional on  $V$ , and let  $a : V \times V \rightarrow \mathbb{R}$  be a bilinear form on  $V \times V$ , such that there exists a constant  $c_1 > 0$ :*

$$a(u, u) \geq c_1 \|u\|^2, \quad \forall u \in V, \quad \text{Coercivity}$$

*and continuous, and, hence, there exists a constant  $c_2 > 0$  such that:*

$$|a(u, v)| \leq c_2 \|u\| \|v\|, \quad \forall u, v \in V. \quad \text{Continuity}$$

Then there exists a unique solution  $u \in V$  such that

$$a(u, v) = f(v), \quad \forall v \in V.$$

Using the theorem above, we can still refine our effort of finding our solution in the domain by choosing a finite dimension space  $V_h \subset V$ , (where  $h$  is discretization parameter), such that:

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in V_h, \tag{2.17a}$$

$$a(u, v_h) = f(v_h), \quad \forall v_h \in V_h. \tag{2.17b}$$

Subtracting equation (2.17b) from equation (2.17a), we have:

$$a(u_h, v_h) - a(u, v_h) = 0,$$

$$a(u_h - u, v_h) = 0.$$

If we define  $u_h - u = e_h$  as the error, then

$$a(e_h, v_h) = 0, \quad \forall v_h \in V_h,$$

which is the so-called Galerkin orthogonality [cf., e.g., [27]]. The approximate solution  $u_h$

converges to the exact solution  $u$  as  $h$  approaches zero. That is:

$$\lim_{h \rightarrow 0} \|u - u_h\| = 0. \quad \square$$

## 2.5 Formulation of Discontinuous Galerkin FEM

In previous section, the formulation of the weak form of the Stokes problem was based on the assumption that, the solution and the test functions are linear. However, this is not always possible in real-life problems. Most often, we deal with domains that are not that “smooth” and are always not continuous. That is to say, the two-sided limits, at a point  $a$ , of such functions are not equal or do not exist. To handle this discontinuities, the Discontinuous Galerkin Method is designed to introduce continuity across the inter-elements in the domain. This is made possible by defining piecewise polynomials representing the approximation to the solution sort for. In this work, we focus on he variant of formulation known as Symmetric Interior Galerkin Method (SIPG).

The formulation involves integrating the defined piecewise polynomials by parts on each mesh element, and summing them over all elements. To ensure stability and continuity, two stabilization terms are added to the system. The first is a symmetrizing term which corresponds to fluxes obtained after integration by part, and the second, is a penalization term which imposes a weak continuity of the numerical solution. This process generates different bilinear forms, involving the primitive variables as well as the introduced terms. For the scheme thus developed to be coercive, there is a known threshold of this penal-

ization term that has to be used. This is discussed briefly in section 3.2.2. Stability and convergence are also covered by these terms [31]. In the following subsections, we go through the process of the formulation of the DGFEM.

### 2.5.1 Discretization

To proceed, we define the following. Suppose we partition the closure ( $\bar{\Omega}$ ) of the solution domain ( $\Omega$ ), into finite number of closed  $d$ -dimensional elements  $K$ , with mutually disjoint interiors that are mutually disjoint. The system of triangulation can either be conforming (regular) or nonconforming (irregular). If the intersection of two adjacent elements is empty or, if they share a common face, an edge or a vertex, the triangulation (mesh) is said to be conforming. On the other hand, if they do not share a common face, etc., then we have the nonconforming mesh, and is said to contain hanging nodes [78]. In this work, we shall assume irregular meshes; however, we shall consider the elements sharing common face. Let  $\mathcal{T}_h$  represent this partition called triangulation of  $\Omega$ , where  $h > 0$  is a parameter. Then we define

$$\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K.$$

Since our problem is based on two dimensional domain, we consider  $d = 2$ , meaning our domain is a polygon and hence,  $K_i \in \mathcal{T}_h$  are triangles. Furthermore, the following notations will be employed in our work. For every element  $K \in \mathcal{T}_h$ , we denote the boundary as  $\partial K$  and define  $\mathbf{n}_i$  as the unit outward normal vector on the boundary  $\partial K_i$ . Next, consider  $K$

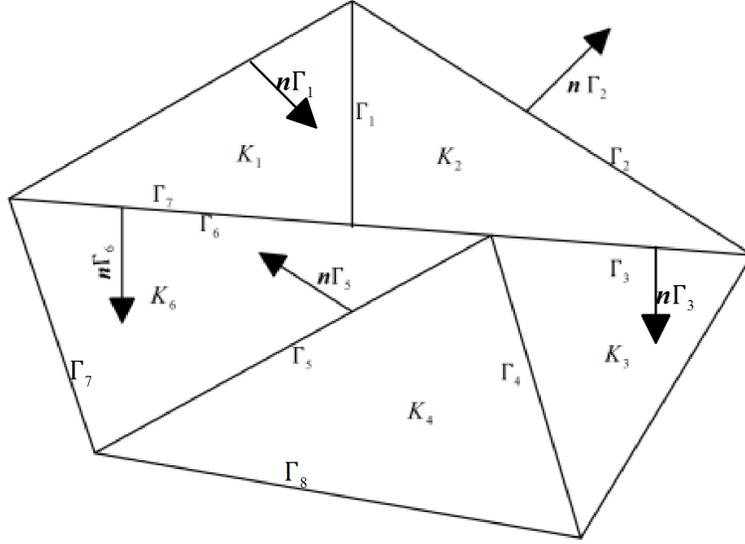


Figure 2.2: Discretization of the Solution Domain

and  $K' \in \mathcal{T}_h$ , neighbouring elements such that  $\partial K \cap \partial K'$  has a positive  $(d-1)$ -dimensional Lebesgue measure [27] and  $\Gamma_i \subset K_i$  a face of element  $K_i$  in the domain with  $|\Gamma_i|$  being the length of face  $\Gamma_i$ . So, as example, from figure 2.2 , we have,

$$\partial K_4 = \Gamma_4 \cup \Gamma_5 \cup \Gamma_8.$$

For the boundary conditions, the following boundary faces can be defined:

$$\mathcal{F}_h^B = \{\Gamma \in \mathcal{F}_h; \Gamma \subset \partial\Omega\},$$

where,  $\mathcal{F}_h$  denotes the system of all faces such that

$$\mathcal{F}_h^D = \{\Gamma \in \mathcal{F}_h; \Gamma \subset \partial\Omega_D\}, \quad (2.18)$$

$$\mathcal{F}_h^N = \{\Gamma \in \mathcal{F}_h; \Gamma \subset \partial\Omega_N\}, \quad (2.19)$$

$$\mathcal{F}_h^I = \mathcal{F}_h \setminus \mathcal{F}_h^B. \quad (2.20)$$

Here,  $\mathcal{F}_h^D$ ,  $\mathcal{F}_h^N$  and  $\mathcal{F}_h^I$  represent the set of all Dirichlet boundary faces, Neumann boundary faces and the set of all inner faces respectively. We will also use the notation  $\mathcal{F}_h^{ID} = \mathcal{F}_h^I \cup \mathcal{F}_h^D$  to denote union set of all faces belonging to the both boundary types. For each internal face  $\mathcal{F}_h^I$ , each pair of neighbouring elements  $K_\Omega^{(L)}, K_\Omega^{(R)} \in \mathcal{T}_h$ , are such that  $\Gamma \subset \partial K_\Omega^{(L)} \cap \partial K_\Omega^{(R)}$ , implying that these two elements are adjacent and share a common face in  $\Omega$ . In these neighbouring elements,  $K_\Omega^{(L)}, K_\Omega^{(R)}$  refer to elements on the left and on the right of the boundary under consideration, respectively

Unlike the solutions derived for the problem in equation (2.9), the DGM seeks solutions in each element  $K_i \in \mathcal{T}_h$  by weakly imposing the Dirichlet boundary condition on the boundary  $\partial K_i$ . This is achieved by using velocity on the neighbouring/adjacent elements, which can be obtained by taking average of these numerical fluxes [78]. Hence our search for the solution is limited to a functional space called "Broken" Sobolev space, defined as:

$$H^k(\Omega, \mathcal{T}_h) = \{v \in L^2(\Omega); \quad v|_K \in H^k(K), \forall K \in \mathcal{T}_h\}. \quad (2.21)$$

Here,  $k \in \mathbb{N}$  is an integer and this space imposes a restriction on  $K$  to be in the Sobolev space  $H^k(K)$ . And for our approximating polynomials, we define the space of piecewise polynomial functions:

$$S_{hp} = \{v \in L^2(\Omega); v|_K \in P_p(K), \forall K \in \mathcal{T}_h\}, \quad (2.22a)$$

$$Q_h = \{q \in L^2(\Omega); q|_K \in P_p(K), \forall K \in \mathcal{T}_h\} \quad (2.22b)$$

where  $P_p(K)$  represents the space of all approximation polynomials of degree  $d \leq p$  on  $K$ ,  $p > 0$  is maximum of the degree of polynomial approximation. In the  $hp$  formulation, the degree,  $p$  in  $S_{hp}$  is greater than that of the pressure space  $Q_h$ .

For these spaces with  $v \in H^k(\Omega, \mathcal{T}_h)$ , we define norm and semi-norms as

$$\|v\|_{H^k(\Omega, \mathcal{T}_h)} = \left( \sum_{K \in \mathcal{T}_h} \|v\|_{H^k(K)}^2 \right)^{1/2} \quad \text{and}, \quad (2.23)$$

$$|v|_{H^k(\Omega, \mathcal{T}_h)} = \left( \sum_{K \in \mathcal{T}_h} |v|_{H^k(K)}^2 \right)^{1/2}. \quad (2.24)$$

Finally, the following notations will be used for any  $\Gamma \in \mathcal{F}_h^I$  and for  $v \in H^1(\Omega, \mathcal{T}_h)$ :

$$v|_{\Gamma}^{(L)} = \text{the trace of } v|_{K_{\Gamma}^{(L)}} \text{ on } \Gamma, \quad v|_{\Gamma}^{(R)} = \text{the trace of } v|_{K_{\Gamma}^{(R)}} \text{ on } \Gamma,$$

$$\langle v \rangle_{\Gamma} = \frac{1}{2} \left( v|_{\Gamma}^{(L)} + v|_{\Gamma}^{(R)} \right) \text{ [average flux]}, \quad [v]_{\Gamma} = v|_{\Gamma}^{(L)} - v|_{\Gamma}^{(R)} \text{ [jump of flux accross faces]}$$

. If however, we consider solution  $v$  on the boundaries and we let  $\Gamma \in \mathcal{F}_h^B$  and  $K_{\Gamma}^{(L)}$ , then

we have that  $\Gamma \subset \partial K_\Gamma^{(L)} \cap \partial\Omega$ . So for  $v \in H^1(\Omega, \mathcal{T}_h)$ , we set

$$v_\Gamma = v|_\Gamma^{(L)} = v|_\Gamma^{(R)} = \text{the trace of } v|_{K_\Gamma^{(L)}} \text{ on } \Gamma.$$

### 2.5.2 Mesh Assumptions

For the sake of coherence and consistency of our presentation, we state and follow the following assumptions. These will be useful in the error analysis of our formulation.

- The mesh triangulation  $\{\mathcal{T}_h\}_{h \in (0, \bar{h})}$ ,  $h > 0$  is shape regular, such that, for some constant  $C_1$ , we have

$$\frac{h_K}{\rho_K} \leq C_1 \quad \forall K \in \mathcal{T}_h; h \in (0, \bar{h}). \quad (2.25)$$

- If we let  $h_\Gamma > 0$  to be the size of the face  $\Gamma$ , then there exist some constants  $C_2, C_3 > 0$ , independent of  $h, K$  and  $\Gamma$  such that

$$C_2 \leq \frac{h_\Gamma}{h_K} \leq C_3. \quad (2.26)$$

In the above, we define  $h_K$  to be the diameter  $h_K = \text{diam}(K)$  of element  $K$ . If we inscribe a number of balls in  $K$ , then we define  $\rho_K$  as the radius of the largest  $d$ -dimensional ball inscribed into  $K$ . Next, we state the following lemmas, without proof, which are useful in the analysis of theorems [28].

**Lemma 3.** *Following the above assumptions, we have that, for every  $v \in H^1(\Omega, \mathcal{T}_h)$*

$$\sum_{\Gamma \in \mathcal{F}_h^{ID}} h_\Gamma^{-1} \int_\Gamma [v]^2 ds \leq \frac{2}{C_4} \sum_{K \in \mathcal{T}_h} h_K^{-1} \int_{\partial K} |v|^2 ds, \quad (2.27)$$

$$\sum_{\Gamma \in \mathcal{F}_h^{ID}} h_\Gamma \int_\Gamma \langle v \rangle^2 ds \leq C_3 \sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |v|^2 ds. \quad (2.28)$$

Now that the underlying principles have been discussed, we turn our attention to finding solution of the discrete form of our problem (2.9).

### 2.5.3 The Discrete Problem

In this section, we derive the discrete problem, using equation (2.9) with the necessary boundary conditions as follows: Find  $(u, p) : \Omega \times \Omega \rightarrow \mathbb{R}^d \times \mathbb{R}$  such that

$$-\nu \Delta u + \nabla p = f \quad \text{in } \Omega, \quad (2.29a)$$

$$u = u_D \quad \text{on } \partial\Omega_D, \quad (2.29b)$$

$$\mathbf{n} \cdot \nabla u = g_N \quad \text{on } \partial\Omega_N, \quad (2.29c)$$

where,  $u_D$  and  $u_N$  are given functions on parts of the boundaries (Dirichlet and Neumann respectively)  $\partial\Omega$  such that  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ ,  $\partial\Omega_D \cap \partial\Omega_N = \emptyset$  and  $\partial\Omega_D \neq \emptyset$ .  $\mathbf{n} \cdot \nabla u = \frac{\partial u}{\partial \mathbf{n}}$  is the derivative of the function  $u$  in the direction of the outer unit normal to the boundary. We now proceed, as before, by multiplying the first equation in (2.29a) by a test function  $v \in H^2$ , integrating (by part) over the domain  $\Omega$ . For the sake of simplicity, we take  $\nu = 1$  and also treat the term containing the pressure separately. Thus, by using the Green's

theorem, the first term of our equation gives:

$$- \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx. \quad (2.30)$$

Next, we apply the boundary conditions and sum over the different faces:

$$\sum_{K \in \mathcal{T}_h} \int_K (\nabla u \cdot \nabla v) \, dx - \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\mathbf{n}_K \cdot \nabla u) v \, ds = \int_{\Omega} f v \, dx. \quad (2.31)$$

We observe that the second term, which represents sums over the boundary can be split into the different faces as:

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\mathbf{n}_K \cdot \nabla u) v \, ds &= \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} (\mathbf{n}_{\Gamma} \cdot \nabla u) v \, ds + \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} (\mathbf{n}_{\Gamma} \cdot \nabla u) v \, ds \\ &\quad + \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \mathbf{n}_{\Gamma} \cdot [(\nabla u_{\Gamma}^{(L)})v_{\Gamma}^{(L)} - (\nabla u_{\Gamma}^{(R)})v_{\Gamma}^{(R)}] \, ds. \end{aligned} \quad (2.32)$$

But from equation (2.29c), we observe that

$$\sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} (\mathbf{n}_{\Gamma} \cdot \nabla u) v \, ds = \int_{\partial \Omega_N} g_N v \, ds,$$

and similarly, the integrand in the last term of equation (2.32) can be re-written as

$$\mathbf{n}_{\Gamma} \cdot (\nabla u_{\Gamma}^{(L)})v_{\Gamma}^{(L)} - \mathbf{n}_{\Gamma} \cdot (\nabla u_{\Gamma}^{(R)})v_{\Gamma}^{(R)} = \mathbf{n}_{\Gamma} \cdot \langle \nabla u \rangle_{\Gamma} [v]_{\Gamma}.$$

Substituting the above two equations and equation (2.32) back into equation (2.31), we obtain

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_K (\nabla u \cdot \nabla v) \, dx - \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \mathbf{n} \cdot \nabla uv \, ds - \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \mathbf{n} \cdot \langle \nabla u \rangle [v] \, ds \\ = \int_{\Omega} f v \, dx + \int_{\partial\Omega_N} g_N v \, ds, \quad v \in H^1(\Omega, \mathcal{T}_h), \end{aligned}$$

upon dropping the subscript  $\Gamma$ .

Using the definition of the faces  $\mathcal{F}^D$  and  $\mathcal{F}^I$ , we can further simplify the above equation thus:

$$\sum_{K \in \mathcal{T}_h} \int_K (\nabla u \cdot \nabla v) \, dx - \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \mathbf{n} \cdot \langle \nabla u \rangle [v] \, ds = \int_{\Omega} f v \, dx + \int_{\partial\Omega_N} g_N v \, ds, \quad v \in H^1(\Omega, \mathcal{T}_h). \quad (2.33)$$

This means that on each face,  $\Gamma_i$ , the velocity fluxes,  $\nabla u$  and  $\nabla v$  contribute a piecewise gradient across each element  $K_i$  adjacent to a corresponding face.

At this point, we turn to the treatment of the inter-element discontinuities. To (weakly) assume continuity of numerical fluxes across the neighbouring elements, we define the so-called interior and boundary penalty bilinear forms:

$$\begin{aligned} J_h^\gamma(u, v) &= \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \gamma [u][v] \, ds + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \gamma u v \, ds. \\ &= \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma [u][v] \, ds, \quad u, v \in H^1(\Omega, \mathcal{T}_h), \end{aligned} \quad (2.34)$$

and on the boundary, this penalty term is associated with the linear form as:

$$J_D^\gamma(v) = \sum_{\Gamma \in \mathcal{F}^D} \int_{\Gamma} \gamma u_D v \, ds. \quad (2.35)$$

The constant  $\gamma$  is the discontinuity-penalization parameter, whose values influence the nature of flux flow across the interior faces and it is defined as

$$\gamma_{\Gamma} = \frac{\beta}{h_{\Gamma}}, \quad \Gamma \in \mathcal{F}_h^{ID}. \quad (2.36)$$

Here,  $\beta > 0$  is the penalization constant and  $h_{\Gamma}$  is the diameter of the interior face  $\Gamma$ .

Hence, the jump terms (2.34) and (2.35) can be written as:

$$J_h^\gamma(u, v) = \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \frac{\beta}{h_{\Gamma}} [u][v] \, ds \quad \text{and} \quad J_D^\gamma(v) = \sum_{\Gamma \in \mathcal{F}^D} \int_{\Gamma} \frac{\beta}{h_{\Gamma}} u_D v \, ds. \quad (2.37)$$

Now turning back to (2.33), we notice that the left-hand side is non-symmetric in terms of  $u$  and  $v$ . To rectify this, we note that a function  $u \in H^1(\Omega) \cap H^2(\Omega, \mathcal{T}_h)$ , satisfies the boundary condition (2.29b) and

$$\sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \mathbf{n} \cdot \langle \nabla v \rangle [u] \, ds = \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \mathbf{n} \cdot \nabla v u_D \, ds, \quad \forall v \in H^2(\Omega, \mathcal{T}_h). \quad (2.38)$$

This is valid, since by our definition,  $[u]$  is zero on the interior faces, and both  $[u]_{\Gamma} = u_{\Gamma} = u_D$  and  $\langle \nabla v \rangle_{\Gamma} = \nabla v_{\Gamma}$  on the Dirichlet boundary. Putting equations (2.33) and (2.38) together, different Discontinuous Galerkin Method can be formulated by summing (2.33)

using multiples of factors -1, 0 or 1 [cf. [28]]. However, in our work, for  $u, v \in H^2(\Omega, \mathcal{T}_h)$ , we use the bilinear diffusion form defined as

$$a_h^n(u, v) = \sum_{K \in \mathcal{T}_h} \int_K (\nabla u \cdot \nabla v) \, dx - \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} (\mathbf{n} \cdot \langle \nabla u \rangle [v] - \mathbf{n} \cdot \langle \nabla v \rangle [u]) \, ds, \quad (2.39)$$

for the left-hand side and on the right-hand side, we have

$$f_h^n(v) = \int_{\Omega} f v \, dx + \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\partial\Omega_N} g_N v; \, ds + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \mathbf{n} \cdot \nabla v u_D \, ds. \quad (2.40)$$

Now, to factor in our penalty weight  $\gamma$ , we define, for  $u, v \in H^2(\Omega, \mathcal{T}_h)$  the following bilinear form, by adding equation (2.37) to (2.39) and (2.40) as:

$$A_h^{n,\gamma}(u, v) = a_h^n(u, v) + J_h^\gamma(u, v), \quad (2.41)$$

$$F_h^{n,\gamma}(v) = f_h^n(v) + J_D^\gamma(v). \quad (2.42)$$

**Remark 2.1.** *We remark here that  $J_h^\gamma$  and  $J_D^\gamma$ , which denote the interior and the boundary penalty forms respectively replace the continuity of the conforming finite element approximate solutions representing Dirichlet boundary conditions (2.29b).*

For the remainder of this work, we will drop the superscripts,  $n$  and  $\gamma$  on simplicity grounds.

### 2.5.4 Multiplicative trace and Inverse Inequalities

Since in the formulation of the forms (2.39) and (2.37), velocity fluxes are summed up across different faces, by means of integrals, we need to define norms that will estimate norms over each element [27]. One major analytical tool used to obtain these estimates is the multiplicative trace inequality. Also, to be able to complete the error analysis, there is the need to the  $H^1$ -semi-norm by applying what is known in literature as the inverse inequality. It used to bound the *Broken* (in our case) Sobolev norm of a function using a lower-indexed norm [81]. These two inequalities are given in the following lemmas. The proofs of these lemmas can be found in [27].

**Lemma 4.** (*Multiplicative trace inequality*) *Assuming the assumption (2.25) is satisfied, then there exists a constant  $C_5 > 0$  which is independent of  $v, h$  and  $K$  such that*

$$\|v\|_{L^2(\partial K)}^2 \leq C_5 \left( \|v\|_{L^2(K)} \|v\|_{H^1(K)} + \frac{\|v\|_{L^2(K)}^2}{h_K} \right), \quad (2.43)$$

$$K \in \mathcal{T}_h, v \in H^1(K), h \in (0, \bar{h}).$$

**Lemma 5.** (*Inverse inequality*) *Let the assumption (2.25) be satisfied. There exists a constant  $C_6 > 0$  which is independent of  $v, h$  and  $K$  such that*

$$\|v\|_{H^1(K)} \leq \frac{C_6}{h_K} \|v\|_{L^2(K)} \quad \forall v \in P_p(K), K \in \mathcal{T}_h, h \in (0, \bar{h}). \quad (2.44)$$

Before we proceed any further, let's turn our attention to the pressure term from equation

(2.29a). To discretise the pressure term, we follow the same steps as before, multiplying by a test function  $v$  and summing over each element in the domain:

$$- \sum_{K \in \mathcal{T}_h} \int_K p \nabla \cdot v + \sum_{\Gamma \in \mathcal{F}^I} \int_{\Gamma} \langle p \rangle [v] \cdot \mathbf{n} + \sum_{\Gamma \in \mathcal{F}^B} \int_{\partial\Omega} p^* [v] \cdot \mathbf{n},$$

where  $p^*$  represents the trace of the pressure on either of the two boundary faces of elements  $\Gamma_i \in \mathcal{F}^B$ .

Hence, we define the bilinear form for the pressure as:

$$b(\mathbf{v}, p) = - \sum_{K \in \mathcal{T}_h} \int_K p \nabla \cdot v + \sum_{\Gamma \in \mathcal{F}^I} \int_{\Gamma} \langle p \rangle [v] \cdot \mathbf{n} + \sum_{\Gamma \in \mathcal{F}^B} \int_{\partial\Omega} p^* [v] \cdot \mathbf{n}. \quad (2.45)$$

Putting all together, the Discontinuous Galerkin formulation for the discrete Stokes problem is defined as: Find  $(\mathbf{u}_h, p_h) \in S_{hp} \times Q_h$  such that:

$$A_h(\mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) = F_h(v), \quad \forall \mathbf{v}_h \in S_{hp}, \quad (2.46)$$

$$b(\mathbf{u}_h, q_h) = 0, \quad \forall q_h \in Q_h, \quad (2.47)$$

where,  $S_{hp}$  and  $Q_h$  are given by (2.22).

If we let our approximation solution  $u_h = \sum_{i=1}^n u_i N_i$ ,  $p_h = \sum_{i=1}^m p_i L_i$ ,  $v_h = N_j$  and  $q_h = L_j$ , then the systems (2.46) and (2.47) can be represented in the form

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}. \quad (2.48)$$

Here,  $A$  and  $B$  are  $n \times n$  and  $m \times n$  matrices respectively, and while  $n$  represents the degrees of freedom of the velocity,  $m$  represents that for the pressure field.

Before we can attempt solving for  $u$  from (2.48), we need to be sure that the matrix  $A$  is invertible. At this point we assume this to be the case, which is normally true for the Stokes problem [54].

From above, we have

$$Au + B^T p = f \text{ and } Bu = 0.$$

By using block-wise Gauss elimination, we have

$$u = A^{-1}f - A^{-1}B^T p$$

substituting into the second equation gives

$$BA^{-1}f - BA^{-1}B^T p = 0.$$

We note that we have eliminated  $u$  from the system, leaving us only with  $p$  in the following relation :  $BA^{-1}B^T p = BA^{-1}f$ , which is known as the Schur complement. However, for a steady incompressible flow, we seek a solution for the pressure only up to a constant. Nevertheless, it can be determined to a reference point or pre-defined boundary conditions. For the *invertibility* of the matrix  $BB^T$ , a necessary condition is that, if  $B$  is  $m \times n$  matrix then  $\text{kernel}(B) = 0$ . It is worth mentioning that for this condition to hold,  $m \geq n$ . This

implies

$$\max_u \frac{(u, B^T p)}{\|u\|} \geq \beta \|p\|, \quad \forall p.$$

When applied to our Stokes problem, we obtain

$$\sup_{u \in S_{hp}} \frac{(p, \nabla \cdot u)}{\|u\|_1} \geq \beta \|p\|_0 > 0, \quad \forall p \in Q_h.$$

This relation will be discussed further to obtain convergence rate of the scheme.

In the following chapter, we discuss how the integrals in equations (2.39), (2.40) and (2.45) can be evaluated.

## Chapter Three: Numerical and Error Analysis

This chapter is dedicated to the numerical aspects of our formulation, as applied to the steady Stokes problem in particular. The first part is concerned with the implementation of the DGM formulation using rectangular element and definition of relevant basis functions. Lastly, we take a look at some error analysis that are associated with, not only the Finite element formulations but specifically with the Discontinuous Galerkin Method.

### 3.1 Tools for Computing Integral Terms

In this section, we explain some useful terms and discuss their applications to our formulation.

#### 3.1.1 Reference Element

Our choice of quadrilateral elements is motivated by the fact that, with the same size of elements, triangular meshes will have as twice as many elements as quadrilateral meshes. Obviously, this is possible if we note that a quadrilateral can be formed by merging two adjacent triangular elements in a domain. The story is not different for the case of edges - quadrilateral meshes have less edges than that in triangular elements. This should generally results in an improved computational cost, especially, in the DG schemes. Works published by [55], [84] show a better performance of quadrilateral elements over the triangular counterparts.

The task of approximating a function on a domain involves associating a finite element to

each element of our triangulation. However, it is well-known that evaluating the functions, using integrals on the physical elements is costly and almost impossible, especially for higher dimensions. To avoid this difficulty, we introduce a reference or master finite element, where the integrals for each element are computed on this reference element defined in the coordinate system:  $(\xi, \eta)$  and then moved or transferred to the physical elements  $(x, y)$  [59], [55] and [49].

Using a unit square as our reference element  $\hat{K}$ , we assume to locate it across the four quadrants and assign the following vertices;  $\hat{A}_1 = (-1, -1)$ ,  $\hat{A}_2 = (1, -1)$ ,  $\hat{A}_3 = (1, 1)$  and  $\hat{A}_4 = (-1, 1)$ . We define a physical element  $K$  with vertices;  $A_i(x_i, y_i)$  for  $i = 1, 2, 3, 4$ . These two bilinear coordinates are shown in 3.1. In order to compute the integrals on the

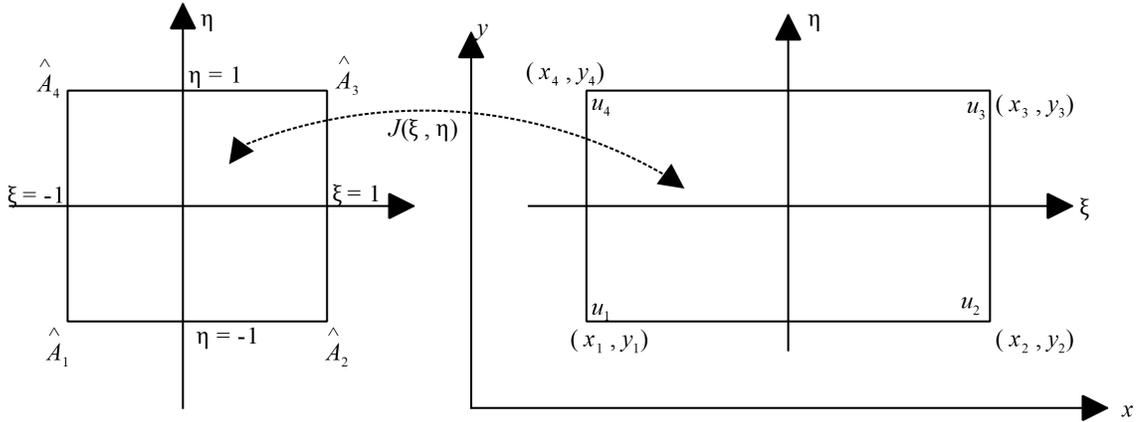


Figure 3.1: Bilinear Elements in Reference Coordinates (Left) and in Physical Coordinates (Right).

physical elements, we define the invertible affine mapping  $F_K : \hat{K} \mapsto K$  as:

$$F_K(\xi, \eta)^T = (x, y)^T, \quad (3.1)$$

where the coordinate  $x$  and  $y$  are given as

$$x = \sum_{i=1}^4 \varphi_i(\xi, \eta) x_i, \quad y = \sum_{i=1}^4 \varphi_i(\xi, \eta) y_i.$$

Hence, a variable  $u$  in the physical element is mapped as:

$$u = \sum_{i=1}^4 \varphi_i(\xi, \eta) u_i,$$

where  $u_i$  is the nodal variable at node  $i$ . For our quadrilateral elements, the shape functions

$\varphi_i$  are given as

$$\begin{aligned} \varphi_1(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 - \eta), & \varphi_2(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 - \eta), \\ \varphi_3(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 + \eta), & \varphi_4(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 + \eta) \end{aligned} \tag{3.2}$$

So the mapping defined in (3.1) can be written (generally) as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \varphi_1 & \varphi_2 & \varphi_3 & \varphi_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varphi_1 & \varphi_2 & \varphi_3 & \varphi_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix},$$

giving us (for our specific case of rectangular meshes),

$$\begin{aligned} x &= \frac{1}{2} [(1 - \xi)x_1 + (1 + \xi)x_2], \\ y &= \frac{1}{2} [(1 - \eta)y_1 + (1 + \eta)y_4], \end{aligned} \tag{3.3}$$

using our coordinate systems.

$$\begin{bmatrix} x \\ y \end{bmatrix} = F_K \begin{bmatrix} \xi \\ \eta \end{bmatrix} = A_K \begin{bmatrix} \xi \\ \eta \end{bmatrix} + b_k,$$

where  $A_K$ , which is a non-singular matrix, defined as

$$A_K = \begin{bmatrix} a_{11}^K & a_{12}^K \\ a_{21}^K & a_{22}^K \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_2 - x_1 & 0 \\ 0 & y_4 - y_1 \end{bmatrix}, \quad b_K = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}. \quad (3.4)$$

The matrix  $A_K$  is as a result of computing derivatives of mapping functions (3.3) *w.r.t* the reference coordinates  $(\xi, \eta)$ , defined as

$$J_F = \frac{\partial(x, y)}{\partial(\xi, \eta)} = A_K,$$

with determinant given as

$$\det J_F = |J_F| = \frac{\mathcal{A}_Q}{4},$$

where  $\mathcal{A}_Q = (x_2 - x_1)(y_4 - y_1)$  is the area of our quadrilateral. From above, we can define the inverse  $F_K^{-1}$  of our mapping. Thus,

$$F_K^{-1} : K \mapsto \hat{K} : \quad F^{-1}(x) = A_K^{-1}(x - b_K) = \xi. \quad (3.5)$$

So our inverse mapping in equation (3.5) is:

$$\xi(x) = \frac{1}{\Delta x} (2x - x_1 - x_2) \quad \text{and} \quad \eta(y) = \frac{1}{\Delta y} (2y - y_1 - y_4),$$

where  $\Delta x = x_2 - x_1$  and  $\Delta y = y_4 - y_1$ .

What we have been trying to show in this section is the relationship between a function

on the reference element and that on the physical element: from physical element,

$$\hat{f}(\xi, \eta) = f(x, y) \quad \text{and} \quad \nabla \hat{f}(\xi, \eta) = A_K^T \nabla f(x, y).$$

as opposed to (3.1.1). For us to be able to compute the integral transforms, we needed to find the derivatives of equation (3.1) and (3.5). The derivative of the mapping in (3.1) gives us the Jacobian and that of (3.5) is the inverse Jacobian. This in effect helps us to compute the elemental mass matrix [84].

**Remark 3.1.** *We note that the existence of this inverse mapping is dependent on the assumption that the Jacobian is constant on the reference element  $\hat{K}$ .*

### 3.2 Numerical Quadrature for a Rectangle

The construction mass matrices helps in computing the  $L_2$  projection, which is necessary in approximating our solution. One way to compute this projection is by use of quadrature or numerical integration.

In the previous section, our concept of reference element is based on the Fubini theorem:

**Theorem 3.1** (Fubini Theorem). *Let  $K$  be the image of  $F_K(\hat{K})$  where  $\hat{K}$  is the reference element and  $F_K$  is continuously differentiable mapping. Let  $f(\mathbf{x}) : K \rightarrow \mathbb{R}$  be an integrable function defined on  $K$ . We define the function  $\hat{f} : \hat{K} \rightarrow \mathbb{R}$  by  $\hat{f}(\hat{\mathbf{x}}) = f(\mathbf{x})$  where  $\mathbf{x} = F_K(\hat{\mathbf{x}})$ . Then*

$$\int_K f(\mathbf{x}) d\mathbf{x} = \int_{\hat{K}} \hat{f}(\hat{\mathbf{x}}) | \det J_F(\hat{\mathbf{x}}) | d\hat{\mathbf{x}},$$

where  $J_F$  is the Jacobian matrix in equation (3.4).

Theorem 3.1 serves as a tool for evaluating the integrals of our Discontinuous Galerkin Method. This enables us to integrate the functions over the reference element, which is easier than integrating over the physical element. Using numerical quadrature rule, we can approximate the integrals on the reference element using a form, similar to that described in this theorem thus:

$$\int_{\hat{K}} \hat{f} \approx \sum_{j=1}^{N_p} w_j \hat{f}(n_{x,j}, n_{y,j}),$$

where  $N_p$  and  $w_j$  represent the number of integration points and quadrature weights respectively.  $(n_{x,j}, n_{y,j})$  are the quadrature nodes inside the reference element. Using our transformation, the corresponding integrals on the physical element can be computed as:

$$\begin{aligned} \int_K f &= \frac{\mathcal{A}_Q}{4} \int_{\hat{K}} \hat{f} \approx \frac{\mathcal{A}_Q}{4} \sum_{j=1}^{N_p} w_j \hat{f}(n_{x,j}, n_{y,j}), \\ \int_K \nabla f \cdot w &\approx \frac{\mathcal{A}_Q}{4} \sum_{j=1}^{N_p} w_j (A_K^T)^{-1} \hat{\nabla} \hat{f}(n_{x,j}, n_{y,j}) \cdot \hat{w}(n_{x,j}, n_{y,j}), \\ \int_K \nabla f \cdot \nabla w &\approx \frac{\mathcal{A}_Q}{4} \sum_{j=1}^{N_p} w_j (A_K^T)^{-1} \hat{\nabla} \hat{f}(n_{x,j}, n_{y,j}) \cdot (A_K^T)^{-1} \hat{\nabla} \hat{w}(n_{x,j}, n_{y,j}). \end{aligned} \quad (3.6)$$

### 3.2.1 Basis Functions

The formulation of the DG method is such that the functions in the DG solution space  $S_{hp}$  do not have to meet the requirement of continuity across the elements, as in the case of continuous formulation. Due to this flexibility, we are able to construct quadrilateral basis functions that do not include the cross-terms that we see in the continuous formu-

lation. One such case is the use of a linear rectangular element instead of general bilinear quadrilateral element [55]. In this case, the space  $S_{hp}$  will have only one support element  $K_i$  on which the functions are defined, but vanish outside  $K_i$ . By this assumption, we define

$$S_{hp} = \text{span}\{\varphi_i^K : 1 \leq i \leq N_{loc}, K \in \mathcal{T}\},$$

where the global basis functions  $\varphi_i$  is defined as

$$\varphi_i^K(x) = \begin{cases} \hat{\varphi}_i \circ F_K(x), & \text{if } x \in K, \\ 0, & \text{if } x \notin K, \end{cases} \quad (3.7)$$

with the functions  $\{\hat{\varphi}_i\}$ 's represent the local bases defined on our reference element  $\hat{K}$  and  $N_{loc}$  denoting the local dimension of order based on  $K$ .

Employing the Legendre polynomials (which are a subset of the Jacobi polynomials), with weights  $\alpha$  and  $\beta$ , then the global basis functions defined in equation (3.7) can be written as

$$\varphi_i^{(1)} = P_i^{(0,0)}(\xi) \quad \text{and} \quad \varphi_j^{(2)} = P_j^{(0,0)}(\eta),$$

where  $P_p^{(\alpha,\beta)}$  are the  $n$ -th order Jacobian polynomials, with  $\alpha = \beta = 0$  in this case. For the rectangular elements, these basis functions are given as

$$\tilde{\phi}_{ij}(\xi, \eta) = \varphi_i^{(1)}(\xi)\varphi_j^{(2)}(\eta). \quad (3.8)$$

We see from these functions that they are constructed as tensor product of the global 1-D

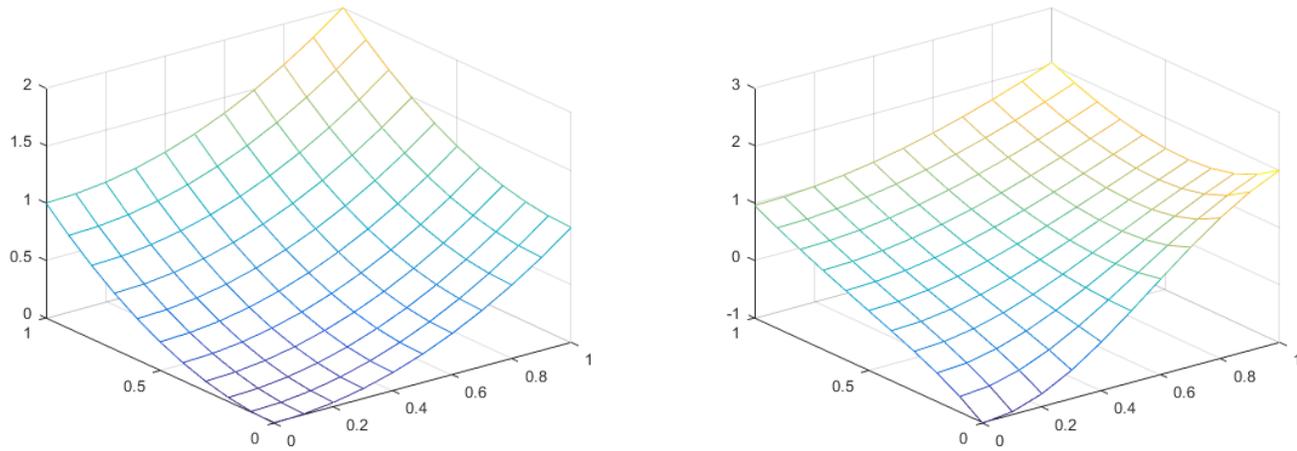


Figure 3.2: Basis bilinear functions defined over a rectangle

bases.

One important property of this orthogonality of the Legendre polynomials is that, it easily leads to a diagonal mass matrix of our formulation. In the particular case of rectangular elements, the area integral is shown to be given as, using the orthogonality property,

$$\int_{\bar{\Omega}} \tilde{\phi}_{ij} \tilde{\phi}_{kl} d\hat{A} = \begin{cases} \frac{4}{(2i+1)(2j+1)}, & \text{if } i = k \text{ and } j = l, \\ 0, & \text{otherwise.} \end{cases} \quad (3.9)$$

Evaluating and ordering (3.9) hierarchically, we have

$$\begin{aligned}
 \phi_1^{\mathcal{A}_Q} &= \bar{\phi}_{00}, & \phi_2^{\mathcal{A}_Q} &= \bar{\phi}_{01}, & \phi_3^{\mathcal{A}_Q} &= \bar{\phi}_{10}, \\
 \phi_4^{\mathcal{A}_Q} &= \bar{\phi}_{02}, & \phi_5^{\mathcal{A}_Q} &= \bar{\phi}_{11}, & \phi_6^{\mathcal{A}_Q} &= \bar{\phi}_{20}, \\
 & \dots & & \dots & & \dots \\
 \phi_{(k+1)(k+2)/2-k}^{\mathcal{A}_Q} &= \bar{\phi}_{0k}, & \dots, & & \phi_{(k+1)(k+2)/2}^{\mathcal{A}_Q} &= \bar{\phi}_{k0}.
 \end{aligned} \tag{3.10}$$

### 3.2.2 Effect of the Size of Penalty Parameter $\gamma$

In our DG formulation, we introduced a parameter  $\gamma$  to penalize the discontinuities across the boundaries and inter-element faces. To ensure coercivity of our bilinear forms (2.41) and (2.42), we must carefully choose the size of  $\gamma$  in a way that it will not be too small or not too large. This coercivity is needed for stability and convergence of our formulation (SIPG)[59]. Different approaches and choices of the size of  $\gamma$  have been used in the literature. For example, Dobrev *et al* in [26], selected the parameter based on the diffusion coefficient  $\nu$ . Murat [59] examined this effect using the Poisson problem with  $\nu = 1$ . From this, their choice was  $\gamma = 3k(k+1)$ . For our formulation, we choose an element dependent parameter, whose size is determined for each iteration. The size, as given in equation (2.36), is determined dynamically during the process of iteration.

### 3.3 Error Analysis of the Discontinuous Galerkin Method

The use of numerical schemes to approximate problems introduces some error elements in the final solution. To minimize such errors, it has always been our aim to reach the following goals:

- i. the approximation solution  $u_h$  of our problem (2.29a) exists and it is unique;
- ii. the approximation solution  $u_h$  converges to the exact solution  $u$  in the  $\|\cdot\|_{S_{hp}}$  - norm as  $h \rightarrow 0$ , i.e.,

$$\lim_{h \rightarrow 0} \|u - u_h\|_{S_{hp}} = 0,$$

and

- iii. we seek  $\alpha > 0$  (order of convergence), independent of  $h$  such that

$$\|u - u_h\|_{S_{hp}} \leq Ch^\alpha, \quad h \in (0, \bar{h}), \quad (3.11)$$

where  $C > 0$  is a constant, independent of  $h$ . In other words, we seek *a priori* error estimate to our problem.

In this section, we use the residual-based method to study the error estimates. We state here that, the jump term  $J^\gamma$  plays no major part in the proof.

1. *Existence and Uniqueness.* The proof of the existence and uniqueness of the velocity  $\mathbf{u}_h$  in our bilinear form  $a_h$  is based on its coercivity, given by the Lax-Milgram theorem 2.1. Existence of the pressure on the other hand, can be proven by employing the inf-sup

condition [64]:

$$\beta \|q\|_{L^2(\Omega)} \leq \sup_{\mathbf{v} \in V, \mathbf{v} \neq 0} \frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_V} \quad \forall q \in Q.$$

We start by proving the continuity of the bilinear form (2.39).

**Lemma 6.** *(Continuity) There exists an inequality*

$$|a_h^{n,\gamma}(u, v)| \leq C \|u\|_{1,\gamma} \|v\|_{1,\gamma}, \quad \forall u, v \in H^2(\Omega, \mathcal{T}_h),$$

where

$$\begin{aligned} \|v\|_{1,\gamma}^2 &= \|v\|^2 + \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla v \rangle)^2 ds \\ &= |v|_{H^1(\Omega, \mathcal{T}_h)}^2 + J_h^\gamma(v, v) + \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla v \rangle)^2 ds, \end{aligned}$$

which is satisfied by any bilinear form  $a_h$  and  $C > 0$  is a constant independent of  $u$ ,  $h$  and

$p$ .

Before we state the proof of the above, we define the norm  $\|\cdot\|$  as

$$\|v\|^2 = \sum_{K \in \mathcal{T}_h} \left( \|\nabla v\|^2 + \|v\|^2 \right) + \sum_{\Gamma \in \mathcal{F}_h^{ID}} \gamma_{\Gamma} \| [v]^2 \|^2,$$

where the two norms used in the first summation are defined with respect to the  $L^2(K)$ ,

while that defined in the last term on the right is with respect to the  $L^2(\Gamma)$ .

*Proof.* To prove lemma 6 if we denote the various terms in bilinear form 2.39 by

$$\begin{aligned} T_1 &= \sum_{K \in \mathcal{T}_h} \int_K |\nabla u \cdot \nabla v| \, dx, \\ T_2 &= \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_K |\mathbf{n} \cdot \langle \nabla u \rangle [v]| \, ds, \\ T_3 &= \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_K |\mathbf{n} \cdot \langle \nabla v \rangle [u]| \, ds. \end{aligned}$$

It follows then that

$$|a_h(u, v)| \leq T_1 + T_2 + T_3.$$

Now, from the norm and semi-norm defined in (2.23), we see that

$$T_1 \leq \sum_{K \in \mathcal{T}_h} |u|_{H^1(K)} |v|_{H^1(K)} \leq |u|_{H^1(\Omega, \mathcal{T})} |v|_{H^1(\Omega, \mathcal{T})}. \quad (3.12)$$

Next, from the Cauchy inequality (2.8), we have

$$\begin{aligned} T_2 &\leq \sum_{\Gamma \in \mathcal{F}_h^{ID}} \left( \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla u \rangle)^2 \, ds \right)^{1/2} \left( \int_{\Gamma} \gamma [v]^2 \, ds \right)^{1/2} \\ &\leq \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla u \rangle)^2 \, ds \right)^{1/2} \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma [v]^2 \, ds \right)^{1/2} \end{aligned} \quad (3.13)$$

and

$$T_3 \leq \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla v \rangle)^2 \, ds \right)^{1/2} \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma [u]^2 \, ds \right)^{1/2}. \quad (3.14)$$

Next, we derive the bound from the inequalities (3.12)-(3.14) as:

$$\begin{aligned}
 |a_h(u, v)| &\leq |u|_{H^1(\Omega, \mathcal{T}_h)} |v|_{H^1(\Omega, \mathcal{T}_h)} \\
 &+ \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla u \rangle)^2 ds \right)^{1/2} \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma [v]^2 ds \right)^{1/2} \\
 &+ \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla v \rangle)^2 ds \right)^{1/2} \left( \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma [u]^2 ds \right)^{1/2} \\
 &\leq \left( |u|_{H^1(\Omega, \mathcal{T}_h)}^2 + \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla u \rangle)^2 ds + J_h^\gamma(u, u) \right)^{1/2} \\
 &\times \left( |v|_{H^1(\Omega, \mathcal{T}_h)}^2 + \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \gamma^{-1} (\mathbf{n} \cdot \langle \nabla v \rangle)^2 ds + J_h^\gamma(v, v) \right)^{1/2} \\
 &= \|u\|_{1, \gamma} \|v\|_{1, \gamma},
 \end{aligned} \tag{3.15}$$

as required.  $\square$

Next, we prove the coercivity of the bilinear form  $a_h^n$ , as defined by (2.41).

**Lemma 7.** (Coercivity) *Let the assumptions (2.25) and (2.26) be valid, the penalty term  $\gamma_\Gamma$  be defined as in (2.36). Further, let*

$$\beta \geq 4C_T, \tag{3.16}$$

where  $C_T = C_3 C_5 (1 + C_6)$ , with the constants  $C_i$  as defined in the respective lemmas above.

Then for all  $\Gamma \in \mathcal{T}_h^{ID}$ ,

$$A(v, v)_h \geq \frac{1}{2} \|v\|_h^2, \quad \forall v \in S_{hp}, \tag{3.17}$$

where,  $A(v, v)_h$  is given in (2.41) and

$$\|v_h\|_h^2 = |v_h|_{H^1(\Omega, \mathcal{T}_h)}^2 + J_h^\gamma(v_h, v_h),$$

with  $|v_h|$  being the semi-norm defined in (2.23).

*Proof.* From the definitions of equations (2.34), (2.36) and (2.39), if we let  $\epsilon > 0$  then from application of the Cauchy and Young's inequalities, we see that

$$\begin{aligned} a_h^s(v_h, v_h) &= |v_h|_{H^1(\Omega, \mathcal{T}_h)}^2 - 2 \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \mathbf{n} \cdot \langle \nabla v_h \rangle [v_h] \, ds \\ &\geq |v_h|_{H^1(\Omega, \mathcal{T}_h)}^2 - 2 \left( \frac{1}{\epsilon} \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} h_{\Gamma} (\mathbf{n} \cdot \langle \nabla v_h \rangle)^2 \, ds \right)^{1/2} \left( \epsilon \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} \frac{1}{h_{\Gamma}} [v_h]^2 \, ds \right)^{1/2} \\ &\geq |v_h|_{H^1(\Omega, \mathcal{T}_h)}^2 - \zeta - \frac{\epsilon}{\beta} J_h^\gamma(v_h, v_h), \end{aligned} \quad (3.18)$$

where we define  $\zeta$  as

$$\zeta = \frac{1}{\epsilon} \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} h_{\Gamma} |\langle \nabla v_h \rangle|^2 \, ds.$$

Next, we deduce from the combination of assumption (2.26), inequalities (2.43), (2.44) and lemma (2.28) that

$$\zeta \leq \frac{C_3}{\epsilon} \sum_{K \in \mathcal{T}_h} h_K \|\nabla v_h\|_{L^2(\partial K)}^2 \quad (3.19)$$

$$\begin{aligned} &\leq \frac{C_3 C_5}{\epsilon} \sum_{K \in \mathcal{T}_h} h_K \left( |v_h|_{H^1(K)} |\nabla v_h|_{H^1(K)} + h_K^{-1} |v_h|_{H^1(K)}^2 \right) \\ &\leq \frac{C_T}{\epsilon} |v_h|_{H^1(\Omega, \mathcal{T})}^2. \end{aligned} \quad (3.20)$$

But if we choose  $\epsilon = 2C_T$ , then from (3.16), (3.19) and (3.20), we see that

$$\begin{aligned} a_h^n(v_h, v_h) &\geq \frac{1}{2} \left( |v_h|_{H^1(\omega, \mathcal{T}_h)}^2 - \frac{C_T}{\beta} J_h^\gamma(v_h, v_h) \right) \\ &\geq \frac{1}{2} \left( |v_h|_{H^1(\omega, \mathcal{T}_h)}^2 - J_h^\gamma(v_h, v_h) \right). \end{aligned} \quad (3.21)$$

But we seek to establish coercivity for the bilinear form  $A_h^{n,\gamma}(v_h, v_h)$ , hence, as defined in (2.41), we write

$$\begin{aligned} A_h^{n,\gamma}(v_h, v_h) &= a_h^n(v_h, v_h) + J^\gamma(v_h, v_h) \\ &\geq \frac{1}{2} \left( |v_h|_{H^1(\omega, \mathcal{T}_h)}^2 + J_h^\gamma(v_h, v_h) \right) \\ &= \frac{1}{2} \|v_h\|^2, \end{aligned} \quad (3.22)$$

and this, on the account of the Lax-Milgram theorem 2.1, implies existence uniqueness of the approximate solution. □

**Remark 3.2.** *The choice and selection of the values of the constants  $C_1 - C_6$ , and their combination with  $C_T$  is taken, in part from [28].*

2. *Convergence.* Now, we study the convergence of the solution sequence  $(u_h, p_h)$  of the system (2.46) and (2.47) approaching the exact solution set  $(u, p)$  of the standard Stokes problem (2.29a). We proceed by using the following theorem:

### 3.3.1 A-priori error estimates

In this section an *a priori* error estimate is obtained. Most often, the exact solutions of the physical problems we seek to solve are not known. The error cannot therefore be calculated directly, hence the use of error estimates. In this sense, we employ *a priori* error estimates, since we need no information from the computed solution  $u_h$ .

**Basic Error:** We follow the approach employed in [83].

Let  $(u, p)$  be the exact solution to our problem (2.29a), and  $(u_h, p_h)$  represent the approximate solution to the discrete problem (2.46). Then, for some constant  $C > 0$ , independent of  $h$ , there hold

$$\| u - u_h \|_h \leq Ch \left( \inf_{w \in S_{hp}} \| u - w \|_h + \inf_{q \in Q_h} \| p - q \|_{L^2(\Omega)} + \sup_{v \in S_{hp}} \frac{|R_h(u, p; v)|}{\| v \|_h} \right), \quad (3.23)$$

$$\| p - p_h \|_{L^2(\Omega)} \leq Ch^2 \left( \inf_{q \in Q_h} \| p - q \|_{L^2(\Omega)} + \inf_{w \in S_{hp}} \| u - w \|_h + \sup_{v \in S_{hp}} \frac{|R_h(u, p; v)|}{\| v \|_h} \right), \quad (3.24)$$

where the residual term  $R(u, p, v)$  is defined as follows, with  $I$  being identity matrix in  $\mathbb{R}^{2 \times 2}$ ,

$$\begin{aligned} R_h(u, p, v) &= \sum_{K \in \mathcal{T}_h} \int_K (\nabla u - pI) : \nabla v \, dx - \sum_{K \in \mathcal{T}_h} \int_K \nabla u : \mathcal{L}(v) \, dx \\ &+ \int_{\Gamma} p \mathcal{M}(v) \, dx - \int_{\Omega} f \cdot v \, dx. \end{aligned} \quad (3.25)$$

In the above definition, we introduced the so-called lifting operators  $\mathcal{L} : S_{hp} \rightarrow s_{hp}$  and  $\mathcal{M} : S_{hp} \rightarrow Q_h$ , (with  $s_{hp}$  being a support space of  $S_{hp}$ ), given as

$$\begin{aligned} \int_{\Omega} \mathcal{L}(v) : \tau \, dx &= \int_{\Gamma} [v] : \langle \tau \rangle \, ds, \quad \forall \tau \in s_{hp}, \\ \int_{\Omega} \mathcal{M}(v)q \, dx &= \int_{\Gamma} [v] \langle q \rangle \, ds, \quad \forall q \in Q_h. \end{aligned}$$

These lifting operators consist of broken gradient and a correction term, which cater for the jumps between the inter-element faces. Generally speaking, they convert a scalar-valued function to a vector counterpart. The residual term (3.25) is bounded in the following way. The proof can be found in [83].

**Lemma 8.** *Assuming uniqueness of our solution  $(u, p)$  and for any  $(w, q) \in V_h \times Q_h$ , the following bounds are satisfied:*

$$\begin{aligned} |R_h(u, p; v)| &\leq C \|v\|_h \left( \|u - w\|_h + \|p - q\|_{L^2(\Omega)} \right) \\ &\quad + \left| \int_{\Gamma} \langle \nabla u - \nabla w \rangle : [v] \, ds - \int_{\Gamma} \langle p - q \rangle [v] \, ds \right|. \end{aligned}$$

Next, we state and proof the main result by combining the abstract bounds stated in (3.23) with the above lemma. This is shown in the following error estimates.

**Theorem 3.2** (Error Estimates). *Let  $(u, p)$  and  $(u_h, p_h)$  be the exact and approximate solutions to (2.29a) and (2.46) respectively with  $n \geq 2$ . If we replace the discrete pressure*

$p$  with a function  $\tilde{q}$  then for any  $(w, \tilde{q}) \in S_{sp} \times \tilde{Q}_h$ , there holds

$$\|u - u_h\|_h + \|p - p_h\|_{L^2(\Omega)} \leq Ch^3(E_1 + E_2 + E_3),$$

where

$$\begin{aligned} E_1^2 &= \sum_{K \in \mathcal{T}_h} \left( |u - w|_{H^1(K)}^2 + h_K^{-2} \|u - w\|_{L^2(K)}^2 + \|p - \tilde{q}\|_{L^2(K)}^2 \right), \\ E_2^2 &= \sum_{K \in \mathcal{T}_h \setminus K_0} h_K^2 \left( |u - w|_{H^2(K)}^2 + |p - \tilde{q}|_{H^1(K)}^2 \right), \\ E_3^2 &= \sum_{K \in K_0} h_K^2 \left( |u - w|_{H^2(K)}^2 + |p - \tilde{q}|_{H^1(K)}^2 \right). \end{aligned} \quad (3.26)$$

We use  $K_0$  to denote elements around a singular point.

*Proof.* Let  $w \in S_{hp}$ ,  $\tilde{q} \in Q_h$  be arbitrary functions and define  $q = \tilde{q} - |\Omega|^{-1} \int_{\Omega} \tilde{q} \, dx \in Q_h$ .

Then a combination of the bounds defined in (3.23) and Lemma (8) gives

$$\begin{aligned} &\|u - u_h\|_h + \|p - p_h\|_{L^2(\Omega)} \\ &\leq Ch^2 \left( \|u - u_h\|_h + \|p - q\|_{L^2(\Omega)} + \sup_{v \in S_{hp}} \frac{|E_h(u - w, p - q; v)|}{\|v\|_h} \right), \end{aligned} \quad (3.27)$$

where

$$E_h(u - w, p - q; v) = \int_{\Gamma} \langle \nabla u - \nabla w \rangle : [v] \, ds - \int_{\Gamma} \langle p - q \rangle [v] \, ds. \quad (3.28)$$

If we take a function  $\phi \in H^1(K)$ , then, using assumptions (2.25) and (2.26), with the

trace inequality (2.43) then the RHS of (3.27) can be estimated using

$$\| \phi \|_{L^2(\partial K)}^2 \leq C \left( h_K^{-1} \| \phi \|_{L^2(K)} + h_K |\phi|_{H^1(K)}^2 \right).$$

This helps us write

$$\begin{aligned} \| u - w \|_h^2 &= \sum_{K \in \mathcal{T}_h} |u - w|_{H^1(K)}^2 + \int_{\Gamma} \frac{p_{\Gamma}^2}{h_{\Gamma}} |[u - w]|^2 ds \\ &\leq Ch^2 \left( \sum_{K \in \mathcal{T}_h} |u - w|_{H^1(K)}^2 + C \sum_{K \in \mathcal{T}_h} h_K^{-1} \| u - w \|_{L^2(\partial K)}^2 \right) \quad (3.29) \\ &\leq Ch^2 \sum_{K \in \mathcal{T}_h} \left( h_K^{-2} \| u - w \|_{L^2(K)}^2 + |u - w|_{H^1(K)}^2 \right) \\ &\leq Ch^2 E_1^2. \end{aligned}$$

Here, we define  $h$  and  $p$  as

$$h_{\Gamma}(\mathbf{x}) := \begin{cases} \min\{h_K, h_{K'}\}, & \mathbf{x} \text{ in the interior of } \partial K \cap \partial K', \\ h_K, & \mathbf{x} \text{ in the interior of } \partial K \cap \partial \Omega \end{cases}$$

and

$$p_{\Gamma}(\mathbf{x}) := \begin{cases} \max\{p_K, p_{K'}\}, & \mathbf{x} \text{ in the interior of } \partial K \cap \partial K', \\ p_K, & \mathbf{x} \text{ in the interior of } \partial K \cap \partial \Omega. \end{cases}$$

For the pressure norm, we have that

$$\begin{aligned}
 \| p - q \|_{L^2(\Omega)} &= \| p - \tilde{q} - |\Omega|^{-1} \int_{\Omega} (p - \tilde{q}) \, dx \|_{L^2(\Omega)} \\
 &\leq \| p - \tilde{q} \|_{L^2(\Omega)} + |\Omega|^{1/2} \int_{\Omega} |p - \tilde{q}| \, dx \\
 &\leq 2 \| p - \tilde{q} \|_{L^2(\Omega)} \\
 &\leq 2E_1,
 \end{aligned} \tag{3.30}$$

noting that  $\int_{\Omega} p \, dx = \int_{\Omega} q \, dx = 0$ . But from (3.28), we see that

$$\begin{aligned}
 |E_h(u - w, p - q; v)| &\leq \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} (|\langle \nabla u - \nabla w \rangle : [v]| + |\langle p - q \rangle [v]|) \, ds \\
 &\leq \sum_{\Gamma \in \mathcal{F}_h^{ID}} \int_{\Gamma} (|\langle \nabla u - \nabla w \rangle| + |\langle p - q \rangle|) |[v]| \, ds \\
 &\leq \sum_{\Gamma \in \mathcal{F}_h^{ID}} \| [v] \|_{L^\infty(\Gamma)} \int_{\Gamma} (|\langle \nabla u - \nabla w \rangle| + |\langle p - q \rangle|) \, ds. \tag{3.31}
 \end{aligned}$$

The norm of the velocity jump  $\| [v] \|$  can be re-written in form of inequality, by employing the inverse inequality, as:

$$\| [v] \|_{L^\infty(\Gamma)} = \| |[v]|^2 \|_{L^\infty(\Gamma)}^{1/2} \leq C \frac{p_\Gamma}{\sqrt{h_\Gamma}} \| |[v]|^2 \|_{L^1(\Gamma)}^{1/2} \leq C \frac{p_\Gamma}{\sqrt{h_\Gamma}} \| [v] \|_{L^2(\Gamma)}.$$

Hence, applying the regularity assumption and substituting into (3.31), we have

$$\begin{aligned}
 |E_h(u - w, p - q; v)| &\leq C \sum_{\Gamma \in \mathcal{F}_h^{ID}} \|\text{pr}h_\Gamma^{-1/2}[v]\|_{L^2(\Gamma)} \int_\Gamma (|\langle \nabla u - \nabla w \rangle| + |\langle p - q \rangle|) ds \\
 &\leq C \left( \int_\Gamma \frac{p_\Gamma^2}{h_\Gamma} |[v]|^2 ds \right)^{1/2} \left( \sum_{K \in \mathcal{T}_h} \|\nabla u - \nabla w\|_{L^1(\partial K)}^2 + \|p - q\|_{L^1(\partial K)}^2 \right)^{1/2} \\
 &\leq C \|v\|_h \left( \sum_{K \in \mathcal{T}_h} \|\nabla u - \nabla w\|_{L^1(\partial K)}^2 + \|p - q\|_{L^1(\partial K)}^2 \right)^{1/2}.
 \end{aligned}$$

Then by the trace inequality, we have that

$$\begin{aligned}
 \frac{|E_h(u - w, p - q; v)|}{\|v\|_h} &\leq C \left( \sum_{K \in \mathcal{T}_h} (|u - w|_{H^1(K)}^2 + \|p - q\|_{L^2(K)}^2) \right. \\
 &\quad + \sum_{K \in \mathcal{T}_h \setminus K_0} h_K^2 (|u - w|_{H^2(K)}^2 + |p - q|_{H^1(K)}^2) \\
 &\quad \left. + \sum_{K \in K_0} h_K (|u - w|_{H_0^1(K)}^2 + |p - q|_{H_0^1(K)}^2) \right)^{1/2}.
 \end{aligned}$$

In our formulation, we set the pressure to be constant and hence  $\nabla(q - \tilde{q}) = 0$ . Hence, if

we apply (3.30) for all  $v \in S_{hp}$ , we have

$$\begin{aligned}
 \frac{|E_h(u - w, p - q; v)|}{\|v\|_h} &\leq C \left( E_1^2 + \sum_{K \in \mathcal{T}_h \setminus K_0} h_K^2 (|u - w|_{H^2(K)}^2 + |p - \tilde{q}|_{H^1(K)}^2) \right. \\
 &\quad \left. + \sum_{K \in K_0} h_K (|u - w|_{H_0^1(K)}^2 + |p - \tilde{q}|_{H_0^1(K)}^2) \right)^{1/2} \\
 &\leq C(E_1 + E_2 + E_3). \tag{3.32}
 \end{aligned}$$

The proof is completed by combing (3.29)-(3.32) and adding (3.27).  $\square$

### 3.4 Numerical Examples

In this section, we take a look at some numerical results for the Stokes problem, in the domain  $\Omega = [-1, 1]^2$ , discussed in this work. We will use a simple lid-driven cavity problem to numerically verify the error estimates developed in this section. This Stokes problem describes fluid flow in a square (or rectangular) domain. The setup of the test problem is such that all the boundaries except the top lid is set to zero. This is to say the no-slip Dirichlet boundary condition is applied to the two sides and the bottom of the cavity, while the top lid is set in a uniform motion (from left to right), see figure 3.3. As a result

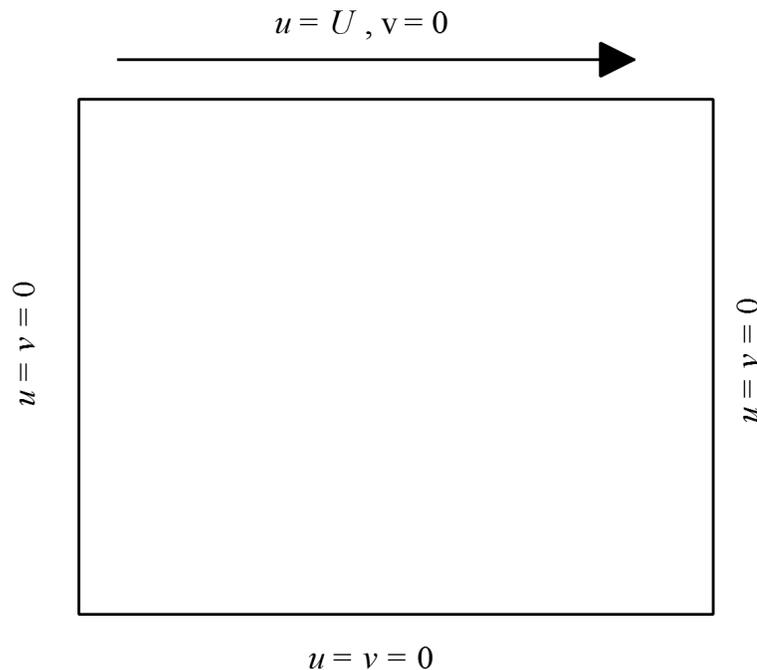


Figure 3.3: A Two-D Lid-driven Cavity problem

of the non-zero horizontal velocity at the top lid, we have different computations. These are the *leaky* cavity, a *water-tight* cavity or a *regularized* cavity. In our case, we consider

the leaky setup, where the velocity at the top lid is given as  $U = 1$ .

In spite of its simple setup and formulation, one major difficulty of the lid-driven cavity problem is the presence of discontinuities at some of its corners, resulting in corner singularities for the solution. This property makes it difficult for the use of the standard Galerkin methods to approximate the boundary data. The two plots in figure 3.4 show these singularities. For different values of the penalization constant,  $\gamma$ , contour plots are used to show these singularities at the corners. In solving our problem, we made use of

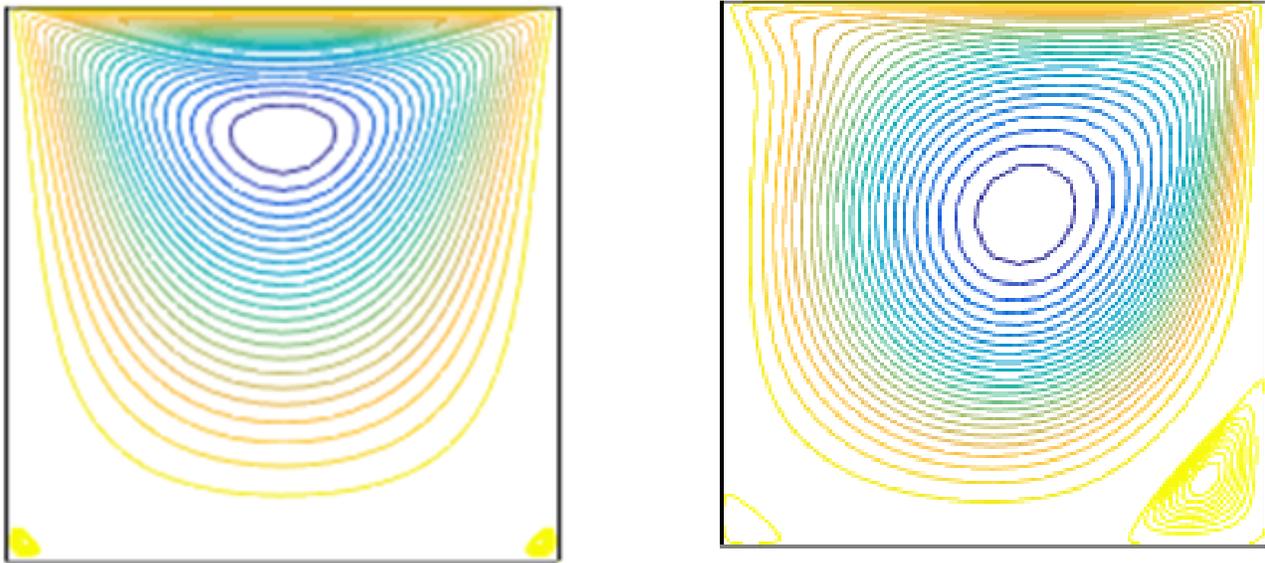


Figure 3.4: Streamline plots for Stokes solutions showing conner singularities

Incompressible Flow & Iterative Solver Software (IFISS)\*

The full solution setup and process is described thus: As stated earlier, the leaky cavity type, with  $16 \times 16$  uniform grids were employed. To ensure optimality, unequal element type of the  $Q_2 - P_1$  was used. The plots are shown in figure 3.5.

---

\*This is from a document by [75] Guide.

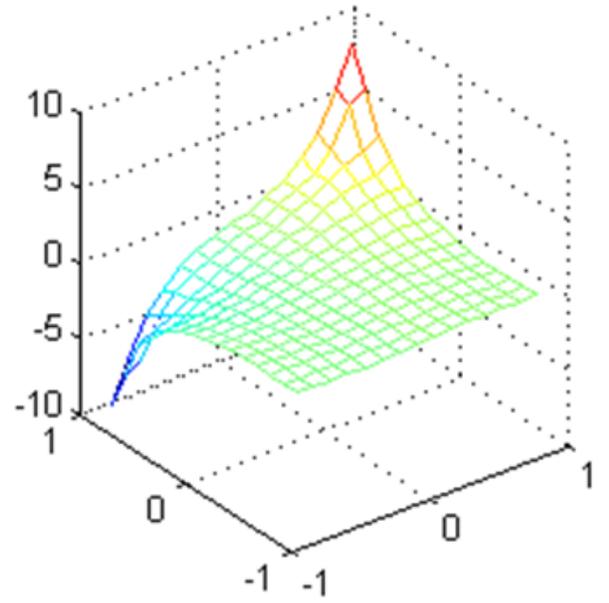
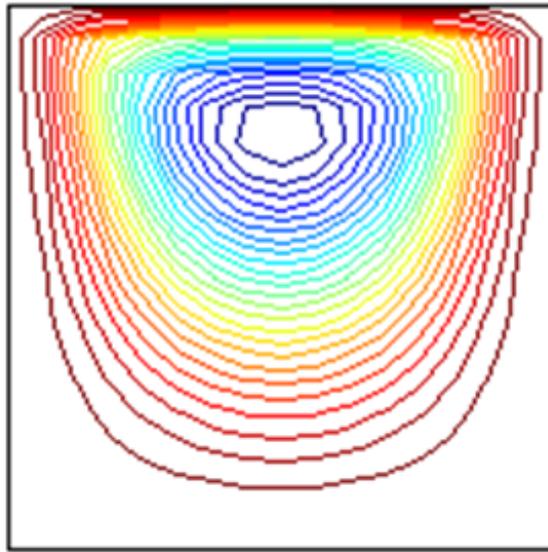


Figure 3.5: Streamline plot of velocity field (left) and mesh plot of pressure field (right)

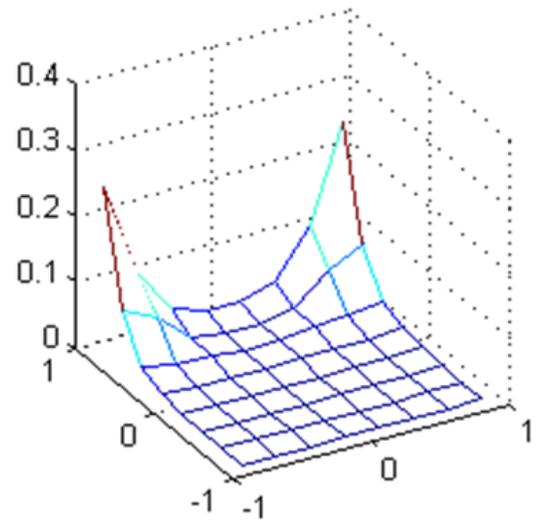
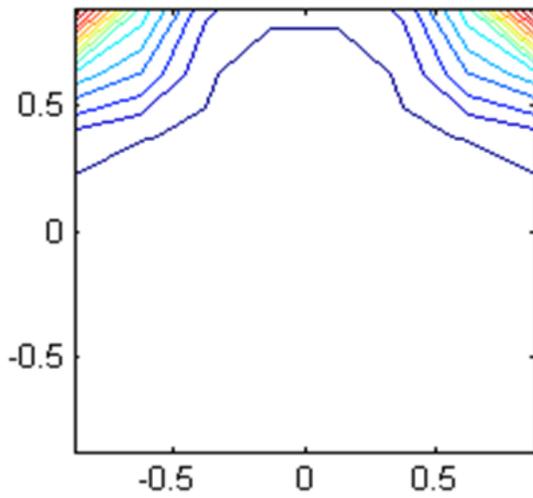


Figure 3.6: Contour (left) and mesh plots (right) of the estimated energy error

Once the problem is solved, some error analyses were conducted to observe the performance of the error bounds derived in section 3.3. In numerical approximations, different types of errors are introduced. These include, the (domain) discretization error, quadrature and finite arithmetic errors; and the approximation error [65]. From the matrix 2.48, we think of solving the system  $MKx = M\mathbf{b}$ , where,  $M$  is a symmetric and positive definite preconditioner. Two errors are associated with this system, namely, the residual error, given as  $\mathbf{r}^{(k)} = \mathbf{b} - K\mathbf{x}^{(k)}$ , and the computational (algebraic) error  $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$ , for the  $k$ th step. For our Stokes problem, we choose  $M$  as

$$M = \begin{bmatrix} A^{-1} & 0 \\ 0 & Q^{-1} \end{bmatrix},$$

where  $Q$  is the pressure matrix. As described by Elma *et al* [30], we define the iteration stopping tolerance as

$$\frac{\sqrt{2}}{\gamma^2} \|\mathbf{r}^{(k)}\|_{M_*} \leq \eta^{s(K)},$$

where,  $\gamma$  is the inf-sup constant,  $M_*$  is the approximation to the preconditioner  $M$  and  $\eta$  the *a priori* error estimate see [76]. In figure 3.7, the plots of the approximation error estimate  $\eta^{(k)}$ , the computational upper bound estimate  $\|\mathbf{e}^{(k)}\|_E$  and the preconditioned residual errors are plotted. These plots show the convergence of the estimated error bounds derived earlier in this work. The numerical error estimates for the different error types at different iteration numbers are shown in table 1. From this table, it is evidently clear that convergence is achieved in only nineteen iterations. This is in agreement with what

is observed in the error plots in figure 3.7.

In this chapter, we discussed the implementation of the approximating domain, where transformation between the physical element to a reference element is described. We also derived various error bounds for a-prior error estimates of the Stokes problem. This is achieved by introducing a residual term to cater for the broken gradients and jumps between the elements. Finally, numerical examples are used to test the theoretical results obtained.

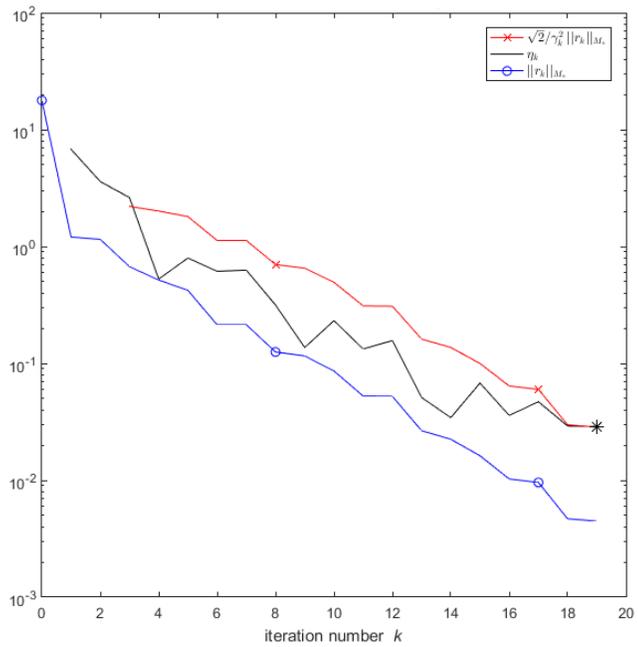


Figure 3.7: Plots of upper error bound (red -\*- line), approximation error (solid black line) and residual error (blue -o- line)

Itr. No.(k)	Err Est.	Appr. Err	Resd. Err.
1	6.7839e+00		1.2057e+00
2	3.5953e+00		1.1510e+005
3	2.6205e+00		6.7249e-01
6	6.1398e-01	1.1261e+00	2.1421e-01
8	3.1579e-01	7.0126e-01	1.2523e-01
10	2.3234e-01	4.9444e-01	8.6172e-02
12	1.5694e-01	3.0879e-01	2.6595e-02
18	2.9131e-02	2.9867e-02	4.6880e-03
19	2.8869e-02	2.8548e-02	4.4777e-03

Table 1: Comparison of convergence of different errors at various iterations

### Estimating the values of $C$ and $\alpha$

From the error equation (3.11), we would like to approximate the value of the constant  $\alpha$  by the following approach. Taking natural log of both sides of equation (3.11), we obtain, by using the equality sign

$$\begin{aligned} \ln \| u - u_h \| &= \ln (Ch^\alpha) \\ \implies \ln \| u - u_h \| &= \alpha C \ln h, \end{aligned} \quad (3.33)$$

where the term  $u - u_h$  represents our error,  $h$  is the mesh size and  $\alpha$  and  $C$  are constants. Now plotting  $\ln \| u - u_h \|$  against  $\ln h$ , we obtain a straight-line graph from which the slope equals the value of our constants  $\alpha$  and  $C$ . If we set the product  $\alpha C = m$ , then from the graph, we observe that the slope  $m$  represents the value of our constants  $\alpha C$ . This implies that, for any given mesh size,  $h$  and a constant  $C$ , the order of convergence  $\alpha$  can be estimated from the relationship (3.33). Finally, it is clear that the slope from

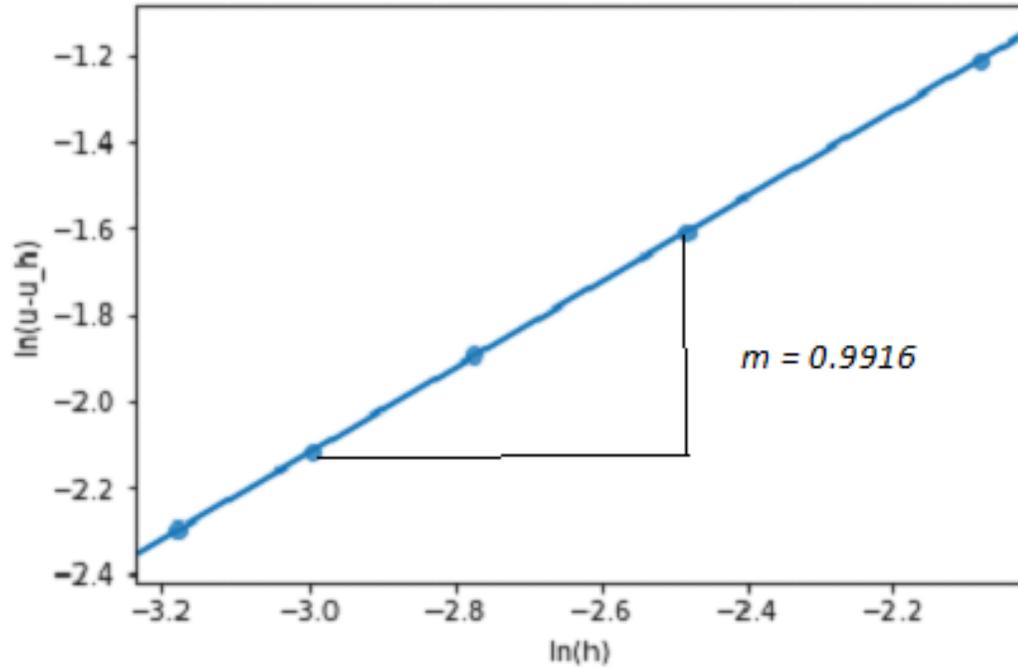


Figure 3.8: Error plot of  $\ln \| u - u_h \|$  against  $\ln h$

the graph is approximately equals to 1. And this tells us that, with a unit change in the natural log of the mesh size, the log value of the error also changes by same factor.

In this chapter, we discussed the implementation of the approximating domain, where transformation between the physical element to a reference element is described. We also derived various error bonds for a-prior error estimates of the Stokes problem. This is achieved by introducing a residual term to cater for the broken gradients and jumps between the elements. Numerical examples are used to test the theoretical results obtained. Finally, the value of convergence constant was estimated by plotting natural log of the error against natural log of the mesh size  $\ln \| u - u_h \|$  against  $\ln h$ .

## Chapter Four: Conclusions and Future Work

In this chapter, we summarise what we have been discussing thus far. We state and draw conclusion from the results obtained. Finally, the direction of future work is briefly described.

We presented formulation of a 2-dimensional Stokes problem using Symmetric Interior Penalty form of the Discontinuous Galerkin Method. We followed the standard approach, which resulted in a symmetric coercive and continuous bilinear forms in space variable for the diffusion terms of the steady-state Stokes problem. In this formulation, we introduced a penalty term, which is supposed to enforce continuity across the different elements of the solution domain. The approach is similar to methods used in [27] and [83].

Next, we considered the construction of elements that are particularly useful in the choice of basis functions. We constructed basis functions that are devoid of bi-polynomial functions. The bilinear functions thus defined in the standard continuous Galerkin FEM contain quadratic cross terms. These functions are needed to ensure continuity between the elements but introduce extra degrees of freedom, resulting in higher computational cost. Our basis functions are orthogonal and hierarchical. The rules for computing numerical quadrature, specifically, for quadrilaterals are defined.

In terms of numerical analyses, we verified the existence and uniqueness of the solution of the symmetric Stokes scheme. Error bounds with the appropriate norms were derived

for both velocity and pressure. These helped in the derivation of a residual-based a priori error estimates.

In our numerical example, we used a free-licensed code (IFISS) to analyse the performance of the theoretical results obtained. Contour and mesh plots showed the distribution of both velocity and pressure fields in our lid-driven cavity example. Streamline plots revealed singularities as reported in the literature.

In attempt to test the theoretical results, we plotted different error types, namely, the upper error bound, the approximation error and the residual error to study the rate of convergence. We observe that these error converge just after in less than 20 iterations. This is shown in the values displayed in table 1, and agrees with the theoretical results obtained in chapter four.

Finally, a simple method of estimating the order of convergence is presented. It is established that, with a given constant, the rate is the slope of the straight line drawn using natural logs of the error and the mesh size  $h$ . The result is shown to be optimal in  $h$ .

The results and conclusions drawn from this work lay the foundation for future investigation and application of the DG method. I would like to study the performance of the DG schemes using equivalent parallelogram to approximate solutions for the general convection/reaction dominated stationary problems. On the error type, it will be of interest to look at the relatively new goal-oriented and hierarchical error estimates, instead of the residual-based discussed in this work.

## Bibliography

- [1] Aikkara, R. and Kadengal, A. (2010). Analysis of mixed convection flow characteristics in square cavity with uniformly heated bottom wall by finite volume method. *International Journal of Engineering Research & Technology (IJERT)*, 3(10):241–248.
- [2] Arnesen, A. J. (2010). Comparison of finite element methods for the navier-stokes equations. Masters thesis, University of Oslo.
- [3] Arnold, D. (1982). An interior penalty finite element method with discontinuous elements. *Society for Industrial and Applied Mathematics (SIAM) Journal of Numerical Analysis*, 19:742–760.
- [4] Baccouch, M. (2016). Analysis of a posteriori error estimates of the discontinuous galerkin method for nonlinear ordinary differential equations. *Journal of Applied Numerical Mathematics*, (106):129–153.
- [5] Badr, B. (2013). *Numerical Solutions to the Navier-Stokes Equations in Two and Three Dimensions*. Phd thesis, University of Manchester.
- [6] Baker, G. A., Jureidini, W. N., and Karakashian, O. A. (1990). Piecewise solenoidal vector fields and the stokes problem. *SIAM Journal on Numerical Analysis*, 27:1466 – 1485.
- [7] Basak, T., Roy, S., Sharma, P. K., and et al (2009a). Analysis of mixed convection

- flows within a square cavity with linearly heated side wall(s). *International Journal of Heat and Mass Transfer*, 52:2224–2242.
- [8] Basak, T., Roy, S., Sharma, P. K., and et al (2009b). Analysis of mixed convection flows within a square cavity with uniform and non-uniform heating of bottom wall. *International Journal of Thermal Sciences*, 48(5):891–912.
- [9] Baumann, C. E. (1997). An h-p adaptive discontinuous galerkin finite element method for computational fluid dynamics.
- [10] Blank, L. (2014). On divergence-free finite element methods for the stokes equations. Masters thesis, Freie University.
- [11] Brezzi, F. and Fortin, M. (1991). *Mixed and Hybrid Finite Element Methods*. Springer.
- [12] Carrier, D. E., May, D. A., and Schnepf, S. M. (2017). Symmetric interior penalty discontinuous galerkin discretisations and block preconditioning for heterogeneous stokes flow. preprint, <http://arxiv.org/abs/1607.03777v2>.
- [13] Causon, D. M., Mingham, C. G., and Qian, L. (2011). *Introductory Finite Volume Methods for PDEs*. bookboon.com.
- [14] Choi, Y., Jeong, D., Lee, S., and Kim, J. (2015). Numerical implementation of the two-dimensional incompressible navier–stokes equation. *KSIAM*, 19(2):103–121.
- [15] Chorin, A. J. (1997). A numerical method for solving incompressible viscous flow problems. *Computational Physics*, 135:118–125.

- [16] Chung, T. J. (2002). *Computational Fluid Dynamics*. Cambridge University Press.
- [17] Cockburn, B. and Kanschat, G. (2005). The local discontinuous galerkin methods for linearized incompressible fluid flow: A review. *Computers and Fluids*, 34.
- [18] Cockburn, B. and Shu, C.-W. (1998). Local discontinuous galerkin methods for the time-dependent reaction-diffusion systems. *SIAM Journal on Numerical Analysis*, 35(1).
- [19] Cockburn, B., Gopalakrishnan, J., and Lazarov, R. (2009). Unified hybridization of discontinuous galerkin, mixed and continuous galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47:1319 – 1365.
- [20] Cockburn, B. and Shu, C.-W. (2001). Runge–kutta discontinuous galerkin methods for convection-dominated problems. *Journal of Scientific Computing*, 16(3):173–261. <http://dx.doi.org/10.1023/A:1012873910884>.
- [21] Creusé, E., Kunert, G., and Nicaise, S. (2004). A posteriori error estimation for the stokes problem: Anisotropic and isotropic discretizations. *Math. Models Methods Appl. Sci.*, pages 1297–1341.
- [22] D, B. (2007). *Finite Elements. Theory, Fast Solvers and Applications in Solid Mechanics*. Cambridge University Press.
- [23] de Montlaur, A. (2009). *High-Order Discontinuous Galerkin Methods For Incompressible Flows*. Phd, Escola Politecnica Superior de Castelldefels.

- [24] DiPietro, D. A. and Ern, A. (2012). *Mathematical Aspects of Discontinuous Galerkin Methods*, volume 69.
- [25] Djoko, J. K. (2013). Discontinuous galerkin finite element discretization for steady stokes flows with threshold slip boundary condition. *Quaestiones Mathematicae*, 36(4):501–516.
- [26] Dobrev, V. A., Lazarov, R. D., and Zikatannov, L. T. (2008). Preconditioning of symmetric interior penalty discontinuous galerkin fem for elliptic problems.
- [27] Dolejší, V. and Feistauer, M. (2015). *Discontinuous Galerkin Method: Analysis and Applications to Compressible Flow*. Number 48 in Springer Series in Computational Mathematics. Springer International Publishing.
- [28] Dolejsi, V. and Feistauer, M. (2016). Discontinuous galerkin method. [http://www.karlin.mff.cuni.cz/~dolejsi/Vyuka/LectureNotes\\_DGM.pdf](http://www.karlin.mff.cuni.cz/~dolejsi/Vyuka/LectureNotes_DGM.pdf).
- [29] Ducasse, D. S. (2010). Analysis of mixed convection in an air filled square cavity. Masters thesis, University of KwaZulu-Natal.
- [30] Elman, H. C., Ramage, A., and Silvester, D. J. (2014). Ifiss: A computational laboratory for investigating incompressible flow problems. *Society for Industrial and Applied Mathematics (SIAM)*, 6(2):261–273.
- [31] Epshteyn, Y. and Rivière, B. (2007). Estimation of penalty parameters for symmetric interior penalty galerkin methods. *Journal of Computational and Applied Mathematics*, 206(2):843 – 872.

- [32] Ern, A. and Guermond, J.-L. (2004). *Theory and Practice of Finite Elements*. Springer Science and Business Media.
- [33] Esfe, M. H., Esforjani, S. S. M., Akbari, M., and Karimipour, A. (2014). Mixed-convection flow in a lid-driven square cavity filled with a nanofluid with variable properties: Effect of the nanoparticle diameter and of the position of a hot obstacle. *Heat Transfer Research*, 45(6):563–578.
- [34] Garvin, J., Lee, K. H., and Rehbolz, L. G. (2011). A numerical study for a velocity-vorticity-helicity formulation of the 3d time-dependent nse. *International Journal Of Numerical Analysis And Modeling*, 2(4):355–368.
- [35] Georgoulis, E. H., Iske, A., and Levesley, J., editors (2009). *Discontinuous Galerkin Methods for Linear Problems: An Introduction*. Springer. <http://www.springer.com/978-3-642-16875-8>.
- [36] Gera, B. and Singh, K. S. R. K. (2010). Cfd analysis of 2d unsteady flow around a square cylinder. *International Journal of Applied Engineering Research, Dindigul*, 1(3):602–610.
- [37] Girault, V. and Raviart, P. (1979). *Finite Element Approximation of the Navier-Stokes Equations*. Springer-Verlag. Google-Books-ID: 2sYZAQAAIAAJ.
- [38] Girault, V., Rivi re, B., and Wheeler, M. F. (2004). A discontinuous galerkin method with nonoverlapping domain decomposition for the stokes and navier-stokes problems. *Mathematics of Computation*, 74(249):53–84.

- [39] Gockenbach, M. S. (2006). *Understanding and Implementing the Finite Element Method*. Society for Industrial and Applied Mathematics (SIAM).
- [40] Güzey, S., Cockburn, B., and Stolarski, H. (2006). The embedded discontinuous galerkin method: Application to linear shell problems. *Int. J. Num. Meths in Engineering*, 70:1806–1824.
- [41] Harris, A., Harris, S., and Rauls, D. (2016). Numerical experiments using matlab: Superconvergence of nonconforming finite element approximation for second-order elliptic problems. *Journal of Applied Mathematics*, 7:2174 – 2182.
- [42] Houston, P., Schwab, C., and Süli, E. (2002). Discontinuous hp-finite element methods for advection–diffusion reaction problems. *SIAM Journal on Numerical Analysis*, 39(6):2133 – 2163.
- [43] Jiajan, W. (2010). Solution to incompressible navier stokes equations by using finite element method. Masters thesis, The University of Texas at Arlington.
- [44] Kalland, K. M. (2008). A navier-stokes solver for single- and two-phase flow. Masters thesis, University of Oslo.
- [45] Karakashian, O. A. and Jureidini, W. N. (1998). A nonconforming finite element method for the stationary navier-stokes equations. *SIAM Journal on Numerical Analysis*, 35(1):93 – 120.
- [46] Khanafer, K. M., Al-Amiri, A. M., and Pop, I. (2007). Numerical simulation of

- unsteady mixed convection in a driven cavity using an externally excited sliding lid. *European Journal of Mechanics B: Fluids*, 26:669–687.
- [47] Klein, B. (2015). *A high-order Discontinuous Galerkin solver for incompressible and low-Mach number flows*. Phd, Darmstadt University of Technology.
- [48] Kundu, P. K. and Cohen, I. M. (2008). *Fluid Mechanics*. Elsevier Inc., 4th edition.
- [49] Kwon, Y. K. and Bang, H. (1997). *The Finite Element Method Using MATLAB*. CRC Press LLC.
- [50] Larson, M. G. and Bengzon, F. (2010). *The Finite Element Method: Theory, Implementation, and Practice*. Springer.
- [51] Larson, M. G. and Niklasson, A. J. (2004). Analysis of a nonsymmetric discontinuous galerkin method for elliptic problems: Stability and energy error estimates. *SIAM Journal on Numerical Analysis*, 42(1):252 – 264.
- [52] Lewis, T. L. N. (2011). Finite element analysis: Mathematical theory and applications. Masters thesis, University of Illinois at Chicago.
- [53] Liu, J. and Weinan, E. (2000). Simple finite element method in vorticity formulation for incompressible flows. *Mathematics of Computation*, 70(234):579–593.
- [54] Logg, A. (2014). Lectures on the finite element method.
- [55] Maggi, A. L. (2011). Discontinuous galerkin finite element methods for shallow water

- flow: Developing a computational infrastructure for mixed element meshes. Masters thesis, The Ohio State University.
- [56] Makanda, G. (2015). *Numerical Study of Convective Fluid Flow in Porous and Non-porous Media*. Phd, University of KwaZulu-Natal.
- [57] Marchi, C. H., Suero, R., and Araki, L. K. (2009). The lid-driven square cavity flow: Numerical solution with a 1024 x 1024 grid. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 31(3):186–198.
- [58] Montlaur, A., Fernandez-Mendez, S., Peraire, J., and Huerta, A. (2007). Discontinuous galerkin methods for the navier-stokes equations using solenoidal approximations. *Int. J. Numer. Meth. Fluids*, 00.
- [59] Murat, U. (2014). *Adaptive Discontinuous Galerkin Methods for Non-linear Reactive Flows*. Phd, Middle East Technical University.
- [60] Nahid, D. J. and Chalhub, M. (2015). *Solution of The Incompressible Navier-Stokes Equations By Projection Methods Using the Integral Transform Technique*. Phd, Federal Fluminense University (UFF).
- [61] Oden, J. T. and Baumann, I. B. C. E. (1998). A discontinuous hp finite element method for diffusion problems. *Journal of Computational Physics*, 49:491–519.
- [62] Omari, R. (2013). Cfd simulations of lid driven cavity flow at moderate reynolds number. *European Scientific Journal*, 9(15):22–35.

- [63] Peraire, J. and Persson, P.-O. (2008). The compact discontinuous galerkin (cdg) method for elliptic problems. *SIAM J. Sci. Stat. Comput.*, 30(4):1806–1824.
- [64] Pietro, D. A. (2007). Analysis of a discontinuous galerkin approximation of the stokes problem based on an artificial compressibility flux. *International Journal for Numeric Methods in Fluid*, 55(55):793–813. [www.interscience.wiley.com](http://www.interscience.wiley.com).
- [65] Reddy, J. N. (2006). *An introduction to the finite felement method*. McGraw-Hill, 3 edition.
- [66] Rivière, B., Wheeler, M. F., and Girault, V. (1999). Improved energy estimates for interior penalty, constrained and discontinuous galerkin methods for elliptic problems part i. *Computational Geoscience*, 3.
- [67] Ruth, B. (2014). Implementation of test scenarios for incompressible flow using a divergence-free finite element approach. Bachelor thesis, Technische Universitat Munchen.
- [68] Safael, M. R., Goodarzl, M., and Mohammadi, M. (2011). Numerical modelling of turbulence mixed convection heat transfer in air filled enclosures by finite volume method. *International Journal of Multiphysics*, 5(4):307–323.
- [69] Sardar, S. (2012). Penalty-free discontinuous galerkin methods for the stokes and navier-stokes equations. Masters thesis, Rice University.
- [70] Schonknecht, N. (2015). On solvers for saddle point problems arising in finite element discretizations of incompressible flow problems. Masters thesis, Freie University.

- [71] Segal, A. I. (2015). Finite element methods for the incompressible navier-stokes equations. article.
- [72] Sharma, R. L. (2012). Viscous incompressible flow simulation using penalty finite element method. *EPJ Web of Conferences*, 25(5):1–12. <http://dx.doi.org/10.1051/epjconf/20122501085>.
- [73] Sheldon, J. P. (2016). *A Hybridizable Discontinuous Galerkin Method for Modelling Fluid-Structure Interactions*. Phd thesis, The Pennsylvania State University.
- [74] Shu, C. (2013). A brief survey on discontinuous galerkin methods in computation fluid dynamics. *Journal of Advances in Mechanics*, 43(6):541–553. [http://imechanica.org/files/13-059\\_revised.pdf](http://imechanica.org/files/13-059_revised.pdf).
- [75] Silvester, D., Elman, H., and Ramage, A. (2016). Incompressible Flow and Iterative Solver Software (IFISS), version 3.5. <http://www.manchester.ac.uk/ifiss/>.
- [76] Silvester, D. J. and Simoncini, V. (2010). An optimal iterative solver for symmetric indefinite systems stemming from mixed approximation,. *ACM. Trans. Math. Software*, 37(42).
- [77] Strang, G. and Fix, G. (2008). *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press.
- [78] Tosseli, A. (2002). hp discontinuous galerkin approximations for the stokes problem.
- [79] Uddin, K. M. S. and Saha, L. K. (2004). Solution to two-dimensional incompressible navier-stokes equations with simple, simpler and vorticity-stream function ap-

- proaches. *Computational Physics Section of Theoretical Physics*, 7:1–13. [arxiv.org/pdf/physics/0407002.pdf](https://arxiv.org/pdf/physics/0407002.pdf).
- [80] Uddin, K. M. S. and Saha, L. K. (2015). Numerical solutions of 2-d incompressible driven cavity flow with wavy bottom surface. *American Journal of Applied Mathematics*, 3(1):30–42.
- [81] Waluga, C. (2012). *Analysis of Hybrid Discontinuous Galerkin Methods for Incompressible Flow Problems*. Phd, RWTH Aachen University.
- [82] Wheeler, M. (1978). An elliptic collocation-finite element method with interior penalties. *SIAM Journal on Numerical Analysis*, 15(1):152 – 161.
- [83] Wihler, T. P. (2002). *Discontinuous Galerkin FEM for elliptic problems in polygonal domains*. Doctoral thesis, Swiss Federal Institute of Technology.
- [84] Wirasaet, D., Tanaka, S., Kubatko, E., Westerink, J., and Dawson, C. (2010). A performance comparison of nodal discontinuous galerkin methods on triangles and quadrilaterals. *Int. Journal Numer. Meth. Fluids*, 64:1336–1362.
- [85] Xu, Y. and Shu, C. (2010). Local discontinuous galerkin methods for high-order time-dependent partial differential equations. *Community of Computational Physics*, 7(1).