

Finite Elements for the Quasi-Geostrophic Equations of the Ocean

Erich L Foster

Dissertation submitted to the Faculty of the
Virginia Polytechnica Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Traian Iliescu, Chair
Slimane Adjerid
John A. Burns
Anne E. Staples

17 April 2013
Blacksburg, Virginia

Keywords: Quasi-Geostrophic equations, finite element method, Argyris element,
wind-driven ocean currents.

Copyright 2015, Erich L Foster

Finite Elements for the Quasi-Geostrophic Equations of the Ocean

Erich L Foster

(ABSTRACT)

The quasi-geostrophic equations (QGE) are usually discretized in space by the finite difference method. The finite element (FE) method, however, offers several advantages over the finite difference method, such as the easy treatment of complex boundaries and a natural treatment of boundary conditions [61]. Despite these advantages, there are relatively few papers that consider the FE method applied to the QGE.

Most FE discretizations of the QGE have been developed for the streamfunction-vorticity formulation. The reason is simple: The streamfunction-vorticity formulation yields a second order *partial differential equation (PDE)*, whereas the streamfunction formulation yields a fourth order PDE. Thus, although the streamfunction-vorticity formulation has two variables (q and ψ) and the streamfunction formulation has just one (ψ), the former is the preferred formulation used in practical computations, since its conforming FE discretization requires low-order (C^0) elements, whereas the latter requires a high-order (C^1) FE discretization.

We present a conforming FE discretization of the QGE based on the Argyris element and we present a two-level FE discretization of the Stationary QGE (SQGE) based on the same conforming FE discretization using the Argyris element. We also, for the first time, develop optimal error estimates for the FE discretization QGE. Numerical tests for the FE discretization and the two-level FE discretization of the QGE are presented and theoretical error estimates are verified. By benchmarking the numerical results against those in the published literature, we conclude that our FE discretization is accurate. Furthermore, the numerical results have the same convergence rates as those predicted by the theoretical error estimates.

To my daughter Juno: You can accomplish anything if you believe.

Acknowledgments

No words can really express the gratitude I feel towards my wife, Chantelle, for doing all the hard work while I worked long hours. Without your help and patience I wouldn't have been able to accomplish so much.

Special thanks to both my academic brothers, David Wells and Zhu Wang. Without both of your help, I would still be figuring things out.

Thanks to my advisor, Traian Iliescu, for always being available. I also would like to say thanks for his patience in guiding me through multiple revisions.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	A Brief History of Wind-Driven Ocean Theory	2
1.3	Problem Overview	5
1.4	Existing Results	6
1.5	Outline	9
2	Governing Equations	10
2.1	The Quasi-Geostrophic Equations	10
2.2	The Stationary Quasi-Geostrophic Equations	14
3	Mathematical Framework	16
3.1	QGE Strong Formulation	17
3.2	SQGE Weak Formulation	18
3.3	Well-Posedness of SQGE	19
4	Finite Element Discretization	21
4.1	QGE Finite Element Discretization	21
4.2	SQGE Finite Element Discretization	22
4.3	Argyris Element	23
4.3.1	Basis Functions	23
4.3.2	Transformation	28
4.3.3	Normal Derivatives	33
4.3.4	Interpolation Error	34
4.3.5	Boundary Conditions	35
4.4	Two-Level Method	37
4.4.1	Two-Level Algorithm	37
5	Error Analysis	38
5.1	SQGE	39
5.2	Two-Level Method	46
5.3	QGE	49

6	Numerical Experiments	57
6.1	SQGE	57
6.1.1	Linear Stommel Model	58
6.1.2	Linear Stommel-Munk Model	64
6.1.3	SQGE	66
6.2	Two-Level SQGE	73
6.2.1	Practical Considerations	73
6.2.2	Computational Efficiency	74
6.2.3	Rates of Convergence	75
6.3	QGE	78
6.3.1	Rates of Convergence	78
7	Conclusions	83
	Bibliography	85

List of Tables

4.1	Constraints for Argyris triangle [1, 13, 20, 27], where i, j, k correspond to the various degrees of freedom of the Argyris element.	25
4.2	Coefficients for the Argyris basis matrix, M	27
6.1	Errors and rate of convergence for the linear Stommel Model (6.1), Test 1 [81].	59
6.2	Errors and rate of convergence for the linear Stommel Model (6.1), Test 1b [81].	61
6.3	Errors and rate of convergence for the linear Stommel Model (6.1), Test 2 [61].	62
6.4	Errors and rate of convergence for the linear Stommel-Munk Model (6.7), Test 3 [15].	64
6.5	Errors and rate of convergence for the linear Stommel-Munk Model (6.7), Test 4 [15].	66
6.6	Errors and rate of convergence for the SQGE (2.29), Test 5.	67
6.7	Errors and rate of convergence for the SQGE (2.29), Test 6.	69
6.8	Table of parameter values used for the simulations of the Mediterranean Sea [8], where $A, \theta_0, \omega, H, L, r_e, \rho$ are the eddy viscosity, reference angle for the β -plane approximation, angular velocity of the Earth, domain height, domain length, radius of the Earth, and density of seawater, respectively.	71
6.9	Observed rates of convergence for SQGE applied to the Mediterranean Sea with forcing function $F = \sin(\frac{\pi}{2}y)$ and true solution obtained from fine mesh with $h = \frac{1}{640}$	72
6.10	Comparison of one-level and two-level methods: the L^2 -norm of the error (e_{L^2}), the H^2 -norm of the error (e_{H^2}) and simulation times.	74
6.11	Observed order of convergence in the H^2 norm for the coarse mesh, H , in the two-level method applied to (2.26) with exact solution (6.13).	75
6.12	Observed order of convergence in the H^2 norm for the fine mesh, h , in the two-Level method applied to (2.26) with the exact solution (6.13).	77
6.13	Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.16).	78
6.14	Observed spatial orders of convergence for Argyris applied to (2.26) with the exact solution (6.16) using implicit Euler for time discretization. Note the observed orders of convergence nearly matches the theoretical error estimates developed in Section 5.3. The L^2 order, however, drops off for the last spatial discretization due to nearing the machine epsilon.	79

6.15	Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.17). Note the observed order of convergence matches the theoretical error estimates developed in Section 5.3.	81
6.16	Observed order of convergence for Argyris applied to (2.26) with the exact solution (6.17). Note the observed order of convergence matches the theoretical error estimates developed in Section 5.3.	82

List of Figures

1.1	Map of the Gulf Stream by Franklin-Folger, 1769. Image courtesy of NOAA.gov.	2
1.2	The large scale surface currents of the ocean. Image courtesy of NOAA.gov.	3
1.3	The <i>Fram</i> caught in ice in March 1894 [62].	3
1.4	QGE is not NSE: Time averaged streamfunctions, time interval $0 \leq t \leq 10$ time step $dt = 1 \times 10^{-3}$, Reynolds number $Re = 200$, forcing term $F = \sin(\pi y)$.	8
4.1	Argyris element with its 21 degrees of freedom.	23
4.2	The reference triangle \widehat{K} for the Argyris element.	24
4.3	Illustration of continuity issue in normal derivatives for C^1 finite elements. .	34
4.4	Eliminated boundary nodes in C^0 finite elements.	36
4.5	Boundary decomposed into vertical and horizontal boundaries.	36
6.1	Linear Stommel model (6.1), Test 1a [81]: Observed errors versus expected rate of convergence.	60
6.2	Linear Stommel model (6.1), Test 1a [81]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs.	60
6.3	Linear Stommel model (6.1), Test 1b [81]: Observed errors versus expected rate of convergence.	61
6.4	Linear Stommel model (6.1), Test 1b [81]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs with $\epsilon_s = 1$	62
6.5	Linear Stommel model (6.1), Test 2 [61]: Observed errors versus expected rate of convergence.	63
6.6	Linear Stommel model (6.1), Test 2 [61]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs.	63
6.7	Linear Stommel-Munk model (6.7), Test 3 [15]: Observed errors versus ex- pected rate of convergence.	65
6.8	Linear Stommel-Munk model (6.7), Test 3 [15]: Streamlines of the approxi- mation, ψ^h , on a mesh size, $h = \frac{1}{32}$, and 28550 DoFs.	65
6.9	Linear Stommel-Munk model (6.7), Test 4 [15]: Observed errors versus ex- pected rate of convergence.	66
6.10	Linear Stommel-Munk model (6.7), Test 4 [15]: Streamlines of the approxi- mation, ψ^h , $h = \frac{1}{32}$, and 28550 DoFs.	67
6.11	SQGE (2.29), Test 5: Observed errors versus expected rate of convergence. .	68
6.12	SQGE (2.29), Test 5: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 36150 DoFs.	68

6.13	SQGE (2.29), Test 6: Observed errors versus expected rate of convergence. .	69
6.14	SQGE (2.29), Test 6: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 36150 DoFs.	70
6.15	Mesh of the Mediterranean Sea created using GMSH [38]. The mesh size corresponds to $h = \frac{1}{320}$ with $DoF = 240,342$	70
6.16	Comparison of large scale structures visible in the numerical results of the SQGE applied to the Mediterranean Sea and the observed large scale structures of the Mediterranean Sea.	72
6.17	The simulation times of the one-level method (green) and of the two-level method (blue) are displayed for all the pairs (h, H) in Table 6.10.	75
6.18	Observed order of convergence in the H^2 norm for the coarse mesh, H , in the two-Level method applied to (2.26) with the exact solution (6.13).	76
6.19	Observed order of convergence in the H^2 norm for the fine mesh, h , in the two-Level method applied to (2.26) with the exact solution (6.13).	77
6.20	Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.16).	79
6.21	Observed orders of convergence in space for Argyris applied to (2.26) with the exact solution (6.16) using implicit Euler for the time discretization.	80
6.22	Observed orders of convergence in time for Implicit-Euler applied to (2.26) with the exact solution (6.17).	81
6.23	Observed orders of convergence in space for Argyris applied to (2.26) with the exact solution (6.17) using implicit Euler for the time discretization.	82

Chapter 1

Introduction

1.1 Motivation

Promise is given that the purely mechanical difficulties connected with the handling of great quantities of data can be overcome so that the computational time factor will eventually cease to be the unsurmountable obstacle to the practical realizations of a program of numerical forecasting. –Jule Charney

Ocean currents have influenced humanity since the beginning of time, both through the effects ocean currents have on climate, but also through the effects they have on human travel. Vikings and Polynesians both made use of ocean currents in their exploration [25, 46]. The Vikings used the ocean currents to travel between Europe and the Americas [46]. In fact, the Vikings were able to sail all along the Greenland coast, because of a lack of ice bergs caused by a more northerly reaching Gulf Stream, which receded a century later [59].

The first documentation of ocean currents were made by Christopher Columbus (1492-1494), Vasco da Gama (1497-1499), and Ferdinand Magellan (1519-1522) [25, 81]. In fact, the Gulf Stream was first mapped by Benjamin Franklin in 1769 (Figure 1.1) [25, 81]. The first instrumental observations of ocean currents on a global scale were begun by the HMS *Challenger* in 1872 when she set sail for over three years around the world [75].

In more recent times we read of plans for shipping through the Northwestern Passage, which until recently had been closed by sea ice. Numerical climate models predicted that the passage would eventually open, however the passage has opened much earlier than anticipated [69].

The climate system is driven by energy from the sun. This solar energy is transferred from the low latitudes to the high latitudes via radiative processes, and oceanic and atmospheric circulations. The oceans have two important roles in the climate system: 1) They store and release heat seasonally, and 2) they transport heat around the globe by their large scale currents [85]. The oceans make up approximately 71% of the Earth's surface, therefore the absorption of solar energy is dominated by the oceans. Thus, climate variability is, to a large extent, an ocean-related phenomenon [75]. In fact, the oceans transport 3.2 petawatts from the tropics [80]. According to Winton [85] the lack of this northward transport of heat would result in a planet blanketed by a sea of ice.



Figure 1.1: Map of the Gulf Stream by Franklin-Folger, 1769. Image courtesy of NOAA.gov.

Additionally, we see that ocean circulation changes are strongly linked to changes in climate [59, 75]. In fact, the cold periods in Western Europe during the time periods 1440-1460, 1687-1703, and 1808-1821 can all be linked to a southward deflection of the Gulf Stream and a southward penetration of Arctic cold water [59].

The large scale ocean flows are characterized by three major features: the wind forcing, the stratification, and the effects of rotation [55, 81]. Annual mean wind patterns, where winds are westward near the equator and eastward at the midlatitudes [25], drive the subtropical and subpolar gyres, which correspond to the strong, persistent, sub-tropical and sub-polar western boundary currents: in the North Atlantic (the Gulf Stream and the Labrador Current), in the South Atlantic (the Brazil and the Falkland Currents), in the North Pacific (the Kuroshio and the Oyashio Currents), in the South Pacific (East Australia Current) and in the Indian Ocean (the Agulhas Current) [25, 81]. These major ocean currents are depicted in Figure 1.2 [44]. One of the common features of these gyres is that they display strong western boundary currents, weaker interior flows, and weak eastern boundary currents. It is these wind-driven ocean circulations that play a significant role in climate dynamics [7, 24, 28, 39].

1.2 A Brief History of Wind-Driven Ocean Theory

The influence of the Coriolis force on the currents of the ocean were first noted by Hadley, Coriolis, and Ferrel. The influence of these forces, however, were considered too small and therefore didn't show in the theory of the time [28]. Ekman's adviser Bjerknes was the first to indicate the importance of the influence of the Earth's rotation in the theory of motions of the ocean [28]. The true importance of the influence of the Coriolis force was first noted by the Norwegian scientist Fridtjof Nansen [7, 28]. Nansen designed a vessel, named the *Fram*, with the intent of allowing it to freeze in the polar ice, and in 1898 Nansen observed the drift of the *Fram* from its original location. As the vessel drifted Nansen noticed that the drift was always $20^\circ - 40^\circ$ to the right of the wind current [7].

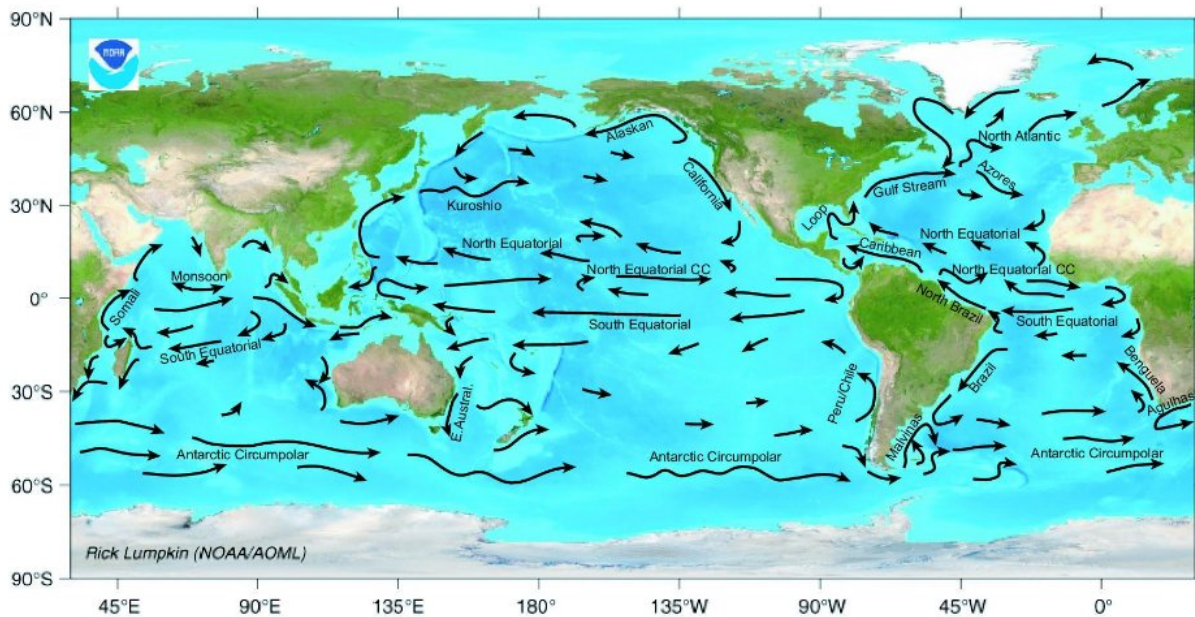


Figure 1.2: The large scale surface currents of the ocean. Image courtesy of NOAA.gov.



Figure 1.3: The *Fram* caught in ice in March 1894 [62].

In 1905 Ekman [28] took the observations made by Nansen and developed what is considered to be the origin of modern theories for wind-driven ocean circulation and their effects on ocean currents [67]. According to Ekman theory momentum from the wind is transferred at the surface of the ocean to the water, as it blows across the ocean surface, via wind stress. Then the Earth's rotation imparts a Coriolis acceleration on the moving water which causes a deflection of the transport to the right of the surface wind stress in the Northern Hemisphere and to the left of the surface wind stress in the Southern Hemisphere [7].

Rossby in his 1936 paper [70] introduced a force which he says had been “completely disregarded by theoretical investigators although its existence has been admitted implicitly by practically everyone who has approached physical oceanography from the descriptive side.” The force that Rossby introduced was the frictional force resulting from large-scale horizontal mixing. According to Rossby, “The introduction of this force permits us to see how motion generated in the surface layers may be diffused and finally dissipated without recourse to doubtful frictional forces at the bottom of the ocean.” Additionally, a significant contribution of Rossby was that he noticed one could represent the Coriolis force by [48]

$$f = f_0 + \beta y, \quad (1.1)$$

where f is the Coriolis parameter, f_0 is the reference Coriolis parameter, β is the β -plane parameter, and y is the y -coordinate (oriented northward).

Then Sverdrup introduced the idea of a variable *Coriolis parameter* [36]. By allowing for the variation of the Coriolis parameter in the north-south direction Sverdrup [79] introduced, for the first time, a north-south asymmetry of the problem domain. The term he introduced, $\nabla \cdot (\mathbf{x}\psi)$ is known as the β -term because of the typical notation for the derivative of the Coriolis parameter [36]. This asymmetry in the north-south direction accounted for the heretofore unexplained equatorial countercurrents.

Ekman's theory explains the deflection of ocean currents from the direction of the wind stress and Sverdrup's allowed for an asymmetry in the north-south direction, thereby resolving the matter of equatorial countercurrents. The combination of the two, however, was still not able to explain the existence of strong western boundary currents [36]. Next, Stommel [78] noticed that taking the gradient of the Coriolis term introduced an asymmetry in not only the north-south direction, but also in the east-west direction. This new term, $\frac{\partial\psi}{\partial x}$, which involves the streamfunction ψ introduces an asymmetry in the east-west direction. This can be seen by substituting $-x$ for x : clearly the sign of the β -term changes. This β -term is a convective term resulting from the rotation of the Earth. The model developed by Stommel was a simple model which included the β -term, a bottom friction term ($-\delta_S \nabla^2 \psi$), and a wind stress forcing term (F). The resultant equation is [36, 78, 81]

$$\frac{\partial\psi}{\partial x} = F - \delta_S \nabla^2 \psi. \quad (1.2)$$

Stommel noted that without the β -term the solution would be symmetric in the east-west direction and therefore no western boundary current would exist [78].

Soon after Stommel's 1948 paper, Munk [60] published a paper on the westward intensification of ocean currents. In his paper, Munk explains that “[i]n Ekman's and Stommel's model the ocean is assumed homogeneous, a case which the currents extend to the very bottom.” Munk points out that this is “in contrast with observations” and resulted in “very difficult” analysis for Ekman while Stommel was forced “to resort to a rather arbitrary frictional force along the bottom.” To address this Munk introduces lateral friction with a constant viscosity. Thus, the Munk problem is [36, 60, 81]

$$\frac{\partial\psi}{\partial x} = \delta_M^3 \Delta^2 \psi + F, \quad (1.3)$$

where δ_m is the Munk scale. In addition to the introduction of lateral viscosity, the Munk problem introduced a fourth-order operator (the biharmonic operator), which required an extra boundary condition. In Munk's 1950 paper [60], he used the same boundary conditions that we will consider, i.e. $\frac{\partial\psi}{\partial \mathbf{n}} = 0$, where \mathbf{n} is the outward unit normal vector.

In 1948 Charney [18] introduced a method by which he “filtered out the noise” from meteorological equations. By filtering out the so-called noise of the meteorological equations Charney was able to greatly simplify the equations of motion for the ocean and atmosphere, thus allowing for numerical weather prediction, which seemed to previously suffer from unsurmountable complexities. This method of filtering is known as the quasi-geostrophic ap-

proximation, and relies on the assumption that the Rossby number

$$Ro = \frac{U}{fL}, \quad (1.4)$$

where U , L are the characteristic velocity and length, respectively, is much much smaller than one. Later in his 1949 paper [19] Charney introduced a method “for reducing the three-dimensional forecast problem to a two-dimensional one.” It is the equations Charney introduced in this 1949 paper that has now come to be called the quasi-geostrophic equations and is the main focus of this thesis.

1.3 Problem Overview

With the continuous increase in computational power, complex mathematical models are becoming more and more popular in the numerical simulation of oceanic and atmospheric flows. For some geophysical flows in which computational efficiency is of paramount importance, however, simplified mathematical models are central. For example, the *quasi-geostrophic equations* (QGE), a standard mathematical model for large scale oceanic and atmospheric flows [23, 55, 64, 81], are often used in climate modelling [24].

The QGE are usually discretized in space by using the *finite difference method* (FDM) [72]. The *finite element method* (FEM), however, offers several advantages over the popular FDM, as outlined in [61]: (i) an easy treatment of complex boundaries, such as those of continents for the ocean, or mountains for the atmosphere; (ii) an easy grid refinement to achieve a high resolution in regions of interest [15]; (iii) a natural treatment of boundary conditions; and (iv) a straightforward approach for the treatment of multiply connected domains [61]. Despite these advantages, there are relatively few papers that consider the FEM applied to the QGE (see for example: [15, 32, 53, 61, 77]).

To our knowledge, *all* the FE discretizations of the QGE have been developed for the streamfunction-vorticity formulation, none using the streamfunction formulation. The reason is simple: The streamfunction-vorticity formulation yields a second order *PDE*, whereas the streamfunction formulation yields a fourth order PDE. Thus, although the streamfunction-vorticity formulation has two variables (q and ψ) and the streamfunction formulation has just one (ψ), the former is the preferred formulation used in practical computations, since its conforming FE discretization requires low-order (C^0) elements, whereas the latter requires a high-order (C^1) FE discretization.

The streamfunction formulation is, from both mathematical and computational points of view, completely different from the vorticity-streamfunction formulation. Indeed, the FE discretization of the streamfunction formulation generally requires the use of C^1 elements (for a conforming discretization), which makes their implementation challenging. From a mathematical point of view, however, the streamfunction formulation has the following significant advantage over the vorticity-streamfunction formulation: there are optimal error estimates for the FE discretization of the streamfunction formulation (see the error estimate (13.5) and Table 13.1 in [43]), whereas the available error estimates for the vorticity-streamfunction formulation are suboptimal.

Despite the simplification made in formulating the QGE from the full-fledged equations of ocean, the numerical simulation of the QGE is still computationally challenging when integrating over long time periods, as is the case in climate modelling. Therefore, it is necessary to reduce the computational cost of QGE simulations. We will consider a two-level method for reducing this computational cost. To our knowledge this will be the first time this approach has been applied to the streamfunction formulation of the QGE.

Two-level FE discretizations are very promising approaches for FE discretizations of nonlinear PDEs [30, 51]. Two-level FE discretizations aim to solve a particular nonlinear elliptic equation by first solving the system on a coarse mesh and then using the coarse mesh solution to solve the linearized elliptic equation on a fine mesh. The attraction of such a method is that one need only solve the nonlinear equations on a coarse mesh and then use this solution to solve on a fine mesh, thereby reducing computational time without sacrificing much in the way of solution accuracy. The development of the two-level finite element discretization was originally used by Xu in [87]. Later algorithms for the Navier-Stokes Equations (NSE) were developed by Layton [51], Fairag [30, 31], and Shao [74].

The goals of this paper are three-fold. First, we use a C^1 finite element (the Argyris element) to discretize the streamfunction formulation of the QGE. To the best of our knowledge, this is the *first* time that a C^1 finite element has been used in the numerical discretization of the QGE. Second, we derive optimal error estimates for the FE discretization of the QGE and present supporting numerical experiments. To the best of our knowledge, this is the first time that optimal error estimates for the QGE have been derived. Third, we present a two-level algorithm for solving the streamfunction formulation of the QGE and present the error analysis associated with this algorithm. To the best of our knowledge, this is the *first* time that such an algorithm has been applied to the streamfunction formulation of the QGE.

1.4 Existing Results

Although the FE discretization of the QGE is relatively scarce, the corresponding error analysis seems to be even more scarce. To our knowledge, *all* the error analysis for the FE discretization of the QGE has been done for the vorticity-streamfunction formulation, none being done for the streamfunction formulation. Furthermore, to the best of our knowledge, all the available error estimates for the FE discretization of the QGE are *suboptimal*. The first FE error analysis for the FE discretization of the QGE was carried out by Fix [32], in which suboptimal error estimates for the vorticity-streamfunction formulation were proved. Indeed, relationships (4.7) and (4.8) (and the discussion above these) in [32] show that the FE approximations for *both* the potential vorticity (denoted by ζ) and streamfunction (denoted by ψ) consist of piecewise polynomials of degree $k - 1$. At the top of page 381, the author concludes that the error analysis yields the following estimates:

$$\|\psi - \psi^h\|_1 = O(h^{k-1}), \quad (1.5)$$

$$\|\zeta - \zeta^h\|_0 = O(h^{k-1}). \quad (1.6)$$

Although the streamfunction error estimate (1.5) appears to be optimal, the potential vorticity error estimate (1.6) is clearly suboptimal. Indeed, using piecewise polynomials of degree

$k - 1$ for the FE approximation of the vorticity, one would expect an $O(h^k)$ error estimate in the L^2 norm. Medjo [57, 58] used a FE discretization of the vorticity-streamfunction formulation and proved error estimates for the time discretization, but no error estimates for the spatial discretization. Finally, Cascon et al. [15] proved both *a priori* and *a posteriori* error estimates for the FE discretization of the *linear Stommel-Munk* model (see Section 6 for more details). This model, while similar to the QGE, has one significant difference: the linear Stommel-Munk model is linear, whereas the QGE is nonlinear.

We note that the state-of-the-art in the FE error analysis for the QGE seems to reflect the FE error analysis for the *two-dimensional Navier-Stokes equations (2D NSE)*, to which the QGE are similar in form. Indeed, as carefully discussed in [43], the 2D NSE in streamfunction-vorticity formulation are easy to implement (only C^0 elements are needed for a conforming discretization), but the available error estimates are suboptimal (see Section 11.6 in [43]).

Remark 1. *It must be noted that although QGE and NSE look quite similar in their streamfunction forms, they are quite different. A significant difference lies in the asymmetry introduced by the β -term, $\frac{\partial \psi}{\partial x}$. This β -term introduces the Coriolis effect and differentiates the western boundary from the eastern boundary. As the Rossby number, Ro decreases the effect of the Coriolis becomes more and more significant. As displayed in Figure 1.4 the stronger the Coriolis force (smaller Ro) the narrower the western boundary layer becomes. Additionally, we see that in the two gyre forcing seen in Figure 1.4 as Ro decreases the distance between the gyres decreases, creating an internal boundary layer. Both these facts lend to precautions that must be taken into account when implementing the QGE, which are not necessary in the implementation of the NSE.*

Next, we summarize the discussion in [43], since we believe it sheds light on the QGE setting. For C^0 piecewise polynomial of degree k FE approximation for *both* the vorticity (denoted by ω) and streamfunction (denoted by ψ), the error estimates given in [41] are (see (1.26) in [43]):

$$|\psi - \psi^h|_1 + \|\omega - \omega^h\|_0 \leq C h^{k-1/2} |\ln h|^\sigma, \quad (1.7)$$

where $\sigma = 1$ for $k = 1$ and $\sigma = 0$ for $k > 1$. It is noted in [43] that the error estimate in (1.7) is not optimal: one may lose a half power in h for the derivatives of the streamfunction (i.e., for the velocity), and three-halves power for the vorticity. It is also noted that there is computational and theoretical evidence that (1.7) is not sharp with respect to the streamfunction error. Furthermore, in [33] it was shown that, for the *linear* Stokes equations, the derivatives of the streamfunction are essentially optimally approximated (see (11.27) in [43]):

$$|\psi - \psi^h|_1 \leq C h^{k-\varepsilon}, \quad (1.8)$$

where $\varepsilon = 0$ for $k > 1$ and $\varepsilon > 0$ is arbitrary for $k = 1$. That being said, it is then noted in [43] that (1.7) seems to be sharp for the vorticity error and thus vorticity approximations are, in general, very poor.

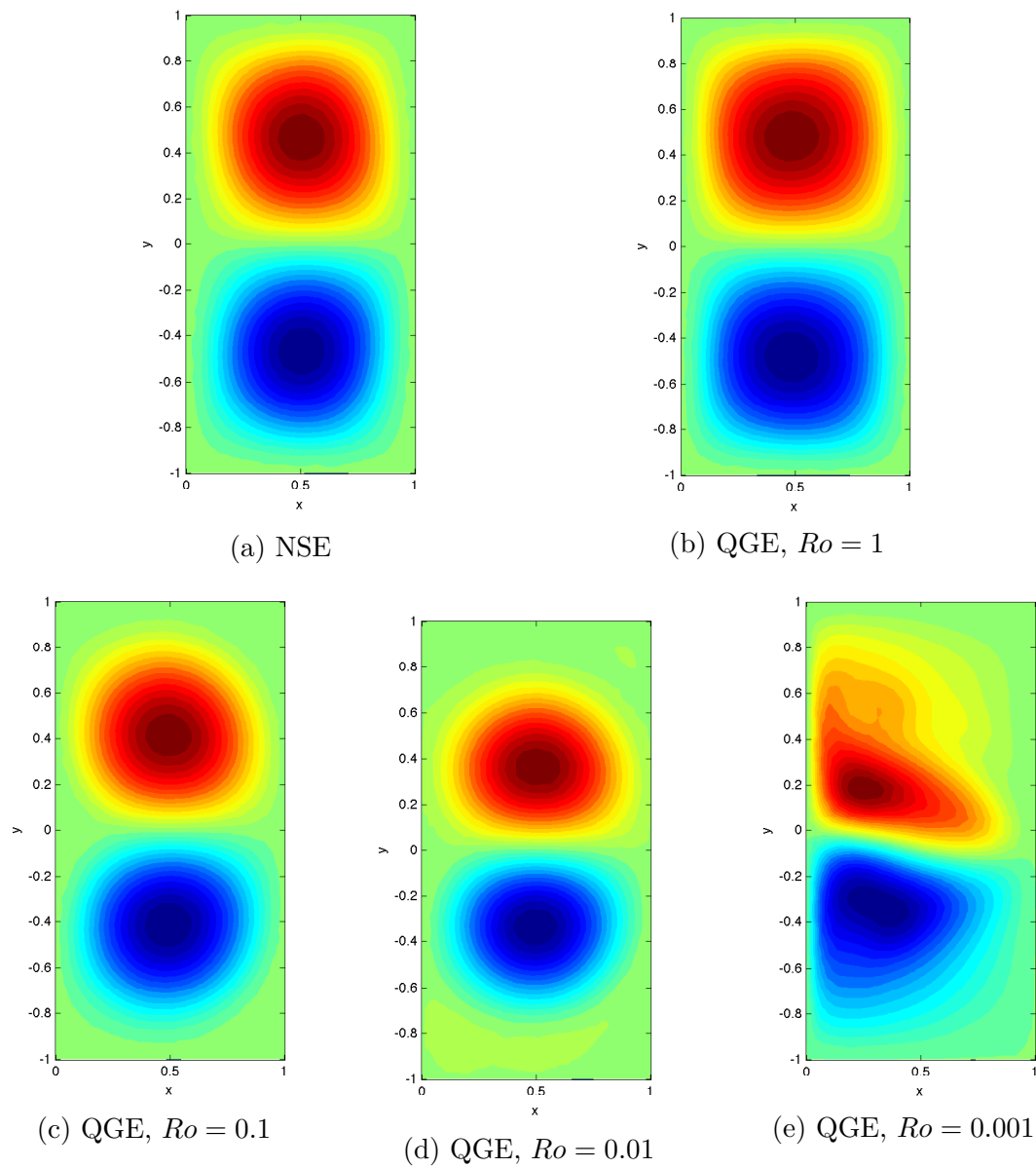


Figure 1.4: QGE is not NSE: Time averaged streamfunctions, time interval $0 \leq t \leq 10$ time step $dt = 1 \times 10^{-3}$, Reynolds number $Re = 200$, forcing term $F = \sin(\pi y)$.

1.5 Outline

The rest of the paper is organized as follows: Section 2.1 presents the QGE, and Section 2.2 presents the stationary QGE. Then Chapter 3 discusses the mathematical framework of both the QGE and SQGE. In Chapter 4 we outline the FE discretization for the QGE (Section 4.1), and SQGE (Section 4.2), with special emphasis placed on the Argyris element in Section 4.3. Additionally, a discussion of the two-level method is presented in Section 4.4. Rigorous error estimates for the FE discretization of the SQGE, the two-level method applied to the SQGE, and the QGE are derived in Section 5.1, Section 5.2, and Section 5.3, respectively. Several numerical experiments supporting the theoretical results for the QGE, SQGE and the two-level method are presented in Section 6.3, Section 6.1, and Section 6.2. These numerical experiments also tackle geophysical flows in realistic complex geometries with realistic parameter values. Finally, conclusions are found in Chapter 7.

Chapter 2

Governing Equations

One of the most popular mathematical models used in the study of large scale wind-driven ocean circulation is the QGE [23, 81]. The QGE represent a simplified model of the full-fledged equations (e.g., the Boussinesq equations), which allows efficient numerical simulations while preserving many of the essential features of the underlying large scale ocean flows. The assumptions used in the derivation of the QGE include the hydrostatic balance, the β -plane approximation, the geostrophic balance, and the eddy viscosity parametrization. Details of the derivation of the QGE and the approximations used along the way can be found in standard textbooks on geophysical fluid dynamics, such as [23, 55, 54, 56, 64, 81].

In the *one-layer QGE*, sometimes called the barotropic vorticity equation, the flow is assumed to be homogenous in the vertical direction. Thus, stratification effects are ignored in this model. The practical advantages of such a choice are obvious: the computations are two-dimensional, and, thus, the corresponding numerical simulation have a low computational cost. To include stratification effects, QGE models of increasing complexity have been devised by increasing the number of layers in the model (e.g., the two-layer QGE and the N -layer QGE [81]). As a first step, in this report we use the one-layer QGE (referred to as “the QGE” in what follows) to study the wind-driven circulation in an enclosed, mid-latitude rectangular basin, which is a standard problem, studied extensively by modelers [23, 52, 54, 55, 56, 81, 64].

2.1 The Quasi-Geostrophic Equations

The QGE are usually written as follows (see, e.g., equation (14.57) in [81], equation (1.1) in [55], equation (1.1) in [83], and equation (1) in [42]):

$$\frac{\partial q}{\partial t} + J(\psi, q) = A \Delta q + F \quad (2.1)$$

$$q = \Delta \psi + \beta y, \quad (2.2)$$

where q is the potential vorticity, ψ is the velocity streamfunction, β is the coefficient multiplying the y coordinate (which is oriented northward) in the β -plane approximation (2.4), F is the forcing, A is the eddy viscosity parametrization, and $J(\cdot, \cdot)$ is the Jacobian operator

given by

$$J(\psi, q) := \frac{\partial \psi}{\partial x} \frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial q}{\partial x}. \quad (2.3)$$

The β -plane approximation reads

$$f = f_0 + \beta y, \quad (2.4)$$

where f is the Coriolis parameter and f_0 is the reference Coriolis parameter (see, e.g., the discussion on page 84 in [22] or Section 2.3.2 in [81]).

As noted in Chapter 10.7.2 in [81] (see also [73]), the eddy viscosity parameter A in (2.1) is usually several orders of magnitude higher than the molecular viscosity. This choice allows the use of a coarse mesh in numerical simulations. The horizontal velocity \mathbf{u} can be recovered from ψ and q by using the following formula:

$$\mathbf{u} := \nabla^\perp \psi = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right). \quad (2.5)$$

The computational domain considered in this section (and in most of this thesis) is standard [42], a rectangular, closed basin on a β -plane with the y coordinate increasing northward and the x coordinate eastward. The center of the basin is at $y = 0$, the northern and southern boundaries are at $y = \pm L$, respectively, and the western and eastern boundaries are at $x = 0$ and $x = L$ (see Figure 1 in [42]).

We are now ready to nondimensionalize the QGE (2.1)-(2.2). There are several ways of nondimensionalizing the QGE, based on different scalings and involving different parameters (see standard textbooks on geophysical fluid dynamics, such as [23, 55, 64, 81]). Since the FE error analysis in this report is based on a precise relationship among the nondimensional parameters of the QGE (see, e.g., (5.13)), we present below a careful nondimensionalization of the QGE.

We first need to choose a length scale and a velocity scale. The length scale we choose is L , the width of the computational domain. The velocity scale is the Sverdrup velocity

$$U := \frac{\pi \tau_0}{\rho H \beta L}, \quad (2.6)$$

where τ_0 is the amplitude of the wind stress, ρ is the density of the fluid, and H is the height of the fluid. The Sverdrup balance, which was used in the derivation of (2.6), expresses the balance between the two dominant effects in the system: the β -effect and the curl of the divergence of the wind stress (see, e.g., Section 14.1.3 in [81]). Once the length and velocity scales are chosen, the variables in the QGE (2.1)-(2.2) can be nondimensionalized as follows:

$$x^* = \frac{x}{L}, \quad y^* = \frac{y}{L}, \quad t^* = \frac{t}{L/U}, \quad q^* = \frac{q}{\beta L}, \quad \psi^* = \frac{\psi}{U L}, \quad (2.7)$$

where a $*$ superscript denotes the nondimensional variable. We first nondimensionalize (2.2). Using (2.7), (2.2) becomes

$$\beta L q^* = \frac{1}{L^2} \Delta^*(U L \psi^*) + \beta (L y^*). \quad (2.8)$$

Dividing (2.8) by βL , we get:

$$q^* = \left(\frac{U}{\beta L^2} \right) \Delta^* \psi^* + y^*. \quad (2.9)$$

Defining the *Rossby number* Ro as follows

$$Ro := \frac{U}{\beta L^2}, \quad (2.10)$$

equation (2.9) becomes

$$q^* = Ro \Delta^* \psi^* + y^*. \quad (2.11)$$

We note that all the nondimensionalizations in (2.7) are naturally based on the velocity scale U and the length scale L , except q^* . Indeed, a nondimensionalization of the form

$$\tilde{q} = \frac{U}{L}, \quad (2.12)$$

would probably be more natural. Note that the alternative nondimensionalization in (2.12) is indeed correct, i.e., the variable \tilde{q} is nondimensional. The main reason for which the nondimensionalization in (2.7) is used instead the one in (2.12) is that the former yields just one constant (the Rossby number Ro) in (2.11), whereas the latter would yield two constants.

Next, we nondimensionalize (2.1). We start with the left-hand side:

$$\frac{\partial q}{\partial t} = (\beta U) \frac{\partial q^*}{\partial t^*}, \quad (2.13)$$

$$\begin{aligned} J(\psi, q) &= \frac{\partial \psi}{\partial x} \frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y} \frac{\partial q}{\partial x} = U \frac{\partial \psi^*}{\partial x^*} \beta \frac{\partial q^*}{\partial y^*} - U \frac{\partial \psi^*}{\partial y^*} \beta \frac{\partial q^*}{\partial x^*} \\ &= (\beta U) J^*(\psi^*, q^*). \end{aligned} \quad (2.14)$$

Next, we nondimensionalize the right-hand side of (2.1). The first term can be nondimensionalized as follows:

$$\begin{aligned} \frac{\partial q}{\partial t} A \Delta q &= A \left(\frac{\partial^2 q}{\partial x^2} + \frac{\partial^2 q}{\partial y^2} \right) = A \left(\frac{1}{L^2} \frac{\partial^2}{\partial x^{*2}} (\beta L q^*) + \frac{1}{L^2} \frac{\partial^2}{\partial y^{*2}} (\beta L q^*) \right) \\ &= A \frac{\beta}{L} \Delta^* q^*. \end{aligned} \quad (2.15)$$

Thus, inserting (2.13)-(2.15) in (2.1), we get:

$$(\beta U) \frac{\partial q^*}{\partial t^*} + (\beta U) J^*(\psi^*, q^*) = A \frac{\beta}{L} \Delta^* q^* + F. \quad (2.16)$$

Dividing by βU , we get:

$$\frac{\partial q^*}{\partial t^*} + J^*(\psi^*, q^*) = \left(\frac{A}{U L} \right) \Delta^* q^* + \frac{F}{\beta U}. \quad (2.17)$$

Defining the *Reynolds number* Re as follows

$$Re := \frac{U L}{A}, \quad (2.18)$$

equation (2.17) becomes

$$\frac{\partial q^*}{\partial t^*} + J^*(\psi^*, q^*) = Re^{-1} \Delta^* q^* + \frac{F}{\beta U}. \quad (2.19)$$

As noted above, the Sverdrup balance expresses the balance between the two dominant effects in the system: the β -effect and the curl of the divergence of the wind stress (see, e.g., Section 14.1.3 in [81]). The velocity scale U in (2.6) was chosen according to the Sverdrup balance. Thus, the last term on the right-hand side of (2.19) has the following units:

$$\left[\frac{F}{\beta U} \right] \sim \left[\frac{\nabla \times (\nabla \cdot \tau_0)}{\rho} \right] \stackrel{(2.6)}{\sim} \left[\frac{\nabla \times (\nabla \cdot \tau_0)}{\frac{\rho}{\frac{\pi \tau_0}{\rho H L}}} \right], \quad (2.20)$$

which, after the obvious simplification, is clearly nondimensional. Thus, (2.20) clearly shows that the last term on the right-hand side of (2.19) is nondimensional, so (2.19) becomes:

$$\frac{\partial q^*}{\partial t^*} + J^*(\psi^*, q^*) = Re^{-1} \Delta^* q^* + F^*. \quad (2.21)$$

Dropping the $*$ superscript in (2.21), we obtain the nondimensional *vorticity-streamfunction formulation* of the *one-layer quasi-geostrophic equations*

$$\frac{\partial q}{\partial t} + J(\psi, q) = Re^{-1} \Delta q + F \quad (2.22)$$

$$q = Ro \Delta \psi + y, \quad (2.23)$$

where Re and Ro are the Reynolds and Rossby numbers, respectively.

Substituting (2.23) in (2.22) and dividing by Ro , we get the *streamfunction formulation* of the *one-layer quasi-geostrophic equations*

$$\frac{\partial [\Delta \psi]}{\partial t} - Re^{-1} \Delta^2 \psi + J(\psi, \Delta \psi) + Ro^{-1} \frac{\partial \psi}{\partial x} = Ro^{-1} F. \quad (2.24)$$

Equations (2.22)-(2.23) and (2.24) are the usual formulations of the one-layer QGE in streamfunction-vorticity and streamfunction formulations, respectively. We note that the streamfunction-vorticity formulation has two unknowns (q and ψ), whereas the streamfunction formulation has only one unknown (ψ). The streamfunction-vorticity formulation, however, is more popular than the streamfunction formulation, since the former is a second-order PDE, whereas the latter is a fourth-order PDE.

We also note that (2.23)-(2.22) and (2.24) are similar in form to the 2D NSE written in the streamfunction-vorticity and streamfunction formulations, respectively. Indeed, (2.22)-(2.23) and (2.24) are almost the same as (11.3)-(11.4) and (13.1) in [43], which are obtained by first writing

$$\mathbf{u} = \begin{pmatrix} \frac{\partial \psi}{\partial y} \\ -\frac{\partial \psi}{\partial x} \end{pmatrix} \quad (2.25)$$

and then taking the curl of the 2D NSE.

There are, however, several significant differences between the QGE and the 2D NSE. First, we note that the term y in (2.23) and the corresponding term $\frac{\partial \psi}{\partial x}$ in (2.24), which model the *rotation effects* in the QGE, do not have counterparts in the 2D NSE. Furthermore, the Rossby number, Ro , in the QGE, which is a measure of the rotation effects, does not appear in the 2D NSE. However, apart from these two significant differences, the streamfunction-vorticity and streamfunction formulations of the QGE and the 2D NSE in the streamfunction form are quite similar.

Thus, for notation consistency (i.e., to ensure that the velocity and the streamfunction are related by (2.25)), we will consider the QGE (2.24) with ψ replaced with $-\psi$:

$$-\frac{\partial [\Delta \psi]}{\partial t} + Re^{-1} \Delta^2 \psi + J(\psi, \Delta \psi) - Ro^{-1} \frac{\partial \psi}{\partial x} = Ro^{-1} F. \quad (2.26)$$

At this point, let us comment on the significance of the two parameters in (2.26), the Reynolds number Re and the Rossby number Ro . As in the 2D NSE case, Re is the coefficient of the diffusion term $-\Delta q = \Delta^2 \psi$. The higher Re , the smaller the magnitude of the diffusion term as compared with the nonlinear convective term $-J(\psi, \Delta \psi) = (\nabla \psi^\perp \cdot \nabla) q = (\mathbf{u} \cdot \nabla) q$. Since Ro , quantifying the rotation effects in the QGE, does not appear in the 2D NSE, its significance deserves a special attention. We first note that, for small Ro , which corresponds to large rotation effects, the forcing term $Ro^{-1} F$ becomes large as compared with the other terms. But probably the most interesting term in (2.26) is $Ro^{-1} \frac{\partial \psi}{\partial x}$, which could be interpreted as a convection type term with respect to ψ , not to $q = -\Delta \psi$. When Ro is small, $Ro^{-1} \frac{\partial \psi}{\partial x}$ becomes large. In conclusion, the physically relevant cases for large scale oceanic flows, in which Re is large and Ro is small (i.e., small diffusion and high rotation, respectively) translate mathematically into a *convection-dominated* PDE with *large forcing*. Thus, from a mathematical point of view, we expect the restrictive conditions used to prove the well-posedness of the 2D NSE [40, 41, 43] will be even more restrictive in the QGE setting, due to the rotation effects. We will later see that this is indeed the case.

To completely specify the equations in (2.26), we need to impose boundary conditions. The question of appropriate boundary conditions for the QGE is a thorny one, especially for the vorticity-streamfunction formulation (see [81, 21] for a careful discussion of this issue). In this report, we consider

$$\psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0 \quad \text{on } \partial \Omega \quad (2.27)$$

$$\psi(0) = \psi_0 \quad \text{when } t = 0 \quad (2.28)$$

which are also the boundary conditions used in [43] for the streamfunction formulation of the 2D NSE.

2.2 The Stationary Quasi-Geostrophic Equations

When testing finite element code it is useful to simplify the problem even though in the real world stationary flows don't exist, the *stationary* QGE (SQGE) are useful, e.g. in testing

code. Additionally, the time dependence of the QGE, (2.24), adds additional complexity to the finite element error analysis. Finite error analysis for time dependent problems usually can be split into two parts; analysis of the spatial discretization that arises through the application of finite elements, and the discretization error in time that arises from the application of the method of lines. Thus, a FE error analysis of the SQGE is a good push off point for analysis of the time-dependent QGE and is therefore the main motivation for presenting the SQGE. It is not only the theory that motivates the study of the SQGE. In practice, the time discretization for the QGE is built around that of the SQGE.

The SQGE are obtained by setting $\frac{\partial q}{\partial t}$ to 0 in (2.22) and therefore we get the *streamfunction formulation* of the *one-layer stationary quasi-geostrophic equations*

$$Re^{-1} \Delta^2 \psi + J(\psi, \Delta \psi) - Ro^{-1} \frac{\partial \psi}{\partial x} = Ro^{-1} F. \quad (2.29)$$

Equations (2.22)-(2.23), (2.24), and (2.29) are the usual formulations of the one-layer QGE in streamfunction-vorticity formulation, the one-layer QGE in streamfunction formulation, and the steady one-layer QGE in streamfunction formulation, respectively.

To completely specify the equations in (2.29), we need to impose boundary conditions. For consistency and with the QGE (2.24) we consider

$$\psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0 \quad \text{on } \partial \Omega,$$

which, as stated previously, are also the boundary conditions used in [43] for the streamfunction formulation of the 2D NSE.

Chapter 3

Mathematical Framework

In this chapter we build the mathematical framework needed to apply the FEM. To this end, we introduce the functional spaces required for the formulation of the QGE, SQGE, and we also discuss the well-posedness of SQGE. All of which are prerequisites for both apply the FE discretization, as is done in Chapter 4, and the error estimates proven in Chapter 5.

For completeness we will first define an inner product, L^p spaces, and Sobolev spaces along with their associated norms.

Definition 3.1. Inner Product [50]:

Let f and g be square integrable functions such that $f, g : \Omega \rightarrow \mathbb{R}$. Then their inner product is defined as

$$(f, g) := \int_{\Omega} f g \, d\mathbf{x} \quad (3.1)$$

Definition 3.2. L^p -norm [50]:

Let $1 \leq p < \infty$ and $f \in L^p(\Omega)$ and X be a normed-space, over Ω , with the associated norm

$$\|f\|_{L^p(\Omega)} := \left(\int_{\Omega} |f|^p \, d\mathbf{x} \right)^{\frac{1}{p}}. \quad (3.2)$$

Then the completion of the norm-space X , denoted $L^p(\Omega)$ is called an L^p -space and we denote the norm associated with $L^p(\Omega)$ as $\|\cdot\|_{L^p(\Omega)}$ and is defined as in (3.2).

We will denote the typical L^2 -norm by $\|\cdot\|$.

Definition 3.3. Sobolev space [29]:

Let $1 \leq p < \infty$, $f \in L^p(\Omega)$ be a real valued function, s be non-negative, and α be a multi-index with $|\alpha| \leq s$. Then the Sobolev space $W^{s,p}(\Omega)$ is defined as

$$W^{s,p}(\Omega) := \{f \in L^p(\Omega) : D^{\alpha} f \in L^p(\Omega) \quad \forall |\alpha| \leq s\}, \quad (3.3)$$

where $D^{\alpha} f$ is the order α derivative of f in the weak sense. The associated norm is given by

$$\|f\|_{W^{s,p}(\Omega)}^p := \sum_{|\alpha| \leq s} \|D^{\alpha} f\|_{L^p(\Omega)}^p. \quad (3.4)$$

In what follows we denote $W^{s,2}(\Omega)$ as $H^s(\Omega)$ and $\|f\|_{H^s(\Omega)}$ as $\|f\|_s$, while the semi-norm will be denoted as $|f|_s := \|D^s f\|$.

3.1 QGE Strong Formulation

We are now ready to derive the strong formulation of the QGE in streamfunction form, (2.26). To this end, it is necessary to introduce the appropriate functional setting. Let ψ be a unique solution [52] of the QGE, (2.26), then

$$\psi \in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^2(\Omega)) \quad (3.5)$$

where

$$\begin{aligned} L^2(0, T; H_0^2(\Omega)) &:= \left\{ \psi(t, \mathbf{x}) : [0, T] \rightarrow H_0^2(\Omega) : \int_0^T \|\Delta \psi\| dt < \infty \right\} \\ L^\infty(0, T; L^2(\Omega)) &:= \left\{ \psi(t, \mathbf{x}) : [0, T] \rightarrow L^2(\Omega) : \operatorname{ess\,sup}_{0 < t < T} \|\psi\| < \infty \right\}. \end{aligned}$$

With the functional setting in place we can derive the strong formulation for the QGE, (2.26).

Additionally, let

$$X := H_0^2(\Omega) = \left\{ \psi \in H^2(\Omega) : \psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega \right\}.$$

Then multiplying (2.26) by a test function $\chi \in X$ and then transferring some of the derivatives from ψ to χ by use of the divergence theorem, we get in a standard way the *strong formulation* of the QGE in streamfunction formulation:

$$\int_\Omega \left(-\frac{\partial [\Delta \psi]}{\partial t} + Re^{-1} \Delta^2 \psi + J(\psi, \Delta \psi) - Ro^{-1} \frac{\partial \psi}{\partial x} \right) \chi d\mathbf{x} = Ro^{-1} \int_\Omega F \chi d\mathbf{x}.$$

We first rewrite the Jacobian term, from (2.26) in a more useful form

$$\begin{aligned} -\frac{\partial}{\partial t} \int_\Omega \Delta \psi \chi d\mathbf{x} + Re^{-1} \int_\Omega \Delta^2 \psi \chi d\mathbf{x} - \int_\Omega \left(\frac{\partial [\Delta \psi]}{\partial x} \frac{\partial \psi}{\partial y} - \frac{\partial [\Delta \psi]}{\partial y} \frac{\partial \psi}{\partial x} \right) \chi d\mathbf{x} \\ - Ro^{-1} \int_\Omega \frac{\partial \psi}{\partial x} \chi d\mathbf{x} = Ro^{-1} \int_\Omega F \chi d\mathbf{x}, \end{aligned}$$

which yields the following equation:

$$-\frac{\partial}{\partial t} \int_\Omega \Delta \psi \chi d\mathbf{x} + Re^{-1} \int_\Omega \Delta^2 \psi \chi d\mathbf{x} - \int_\Omega (\nabla [\Delta \psi] \cdot \nabla \times \psi) \chi d\mathbf{x} - Ro^{-1} \int_\Omega \frac{\partial \psi}{\partial x} \chi d\mathbf{x} = Ro^{-1} \int_\Omega F \chi d\mathbf{x}.$$

For the remaining portion of the strong formulation we shall deal with each summand individually (from left to right) and so

$$\begin{aligned} -\frac{\partial}{\partial t} \int_\Omega \Delta \psi \chi d\mathbf{x} &= -\frac{\partial}{\partial t} \int_\Omega \nabla (\nabla \psi \chi) - \nabla \psi \cdot \nabla \chi d\mathbf{x} \\ &= -\frac{\partial}{\partial t} \int_{\partial\Omega} \nabla \psi \chi \cdot \mathbf{n} dS + \frac{\partial}{\partial t} \int_\Omega \nabla \psi \cdot \nabla \chi d\mathbf{x}. \end{aligned}$$

The next summand gives

$$\begin{aligned}
Re^{-1} \int_{\Omega} \Delta^2 \psi \chi \, d\mathbf{x} &= Re^{-1} \int_{\Omega} \nabla [\nabla^3 \psi \chi] + \nabla^3 \psi \nabla \chi \, d\mathbf{x} \\
&= Re^{-1} \int_{\partial\Omega} \nabla^3 \psi \chi \cdot \mathbf{n} \, dS \overset{0}{=} Re^{-1} \int_{\Omega} \nabla [\Delta \psi \nabla \chi] - \Delta \psi \Delta \chi \, d\mathbf{x} \\
&= -Re^{-1} \int_{\partial\Omega} \Delta \psi \nabla \chi \cdot \mathbf{n} \, dS \overset{0}{=} Re^{-1} \int_{\Omega} \Delta \psi \Delta \chi \, d\mathbf{x}.
\end{aligned}$$

For the third and final summand we have

$$\begin{aligned}
- \int_{\Omega} \nabla [\Delta \psi] \cdot \nabla \times \psi \chi \, d\mathbf{x} &= - \int_{\Omega} \nabla [\Delta \psi] \cdot \nabla \times \psi \chi \, d\mathbf{x} \\
&= - \int_{\Omega} \nabla [\Delta \psi \cdot \nabla \times \psi \chi] - \Delta \psi \nabla [\nabla \times \psi \chi] \overset{0}{=} \\
&\quad - \Delta \psi \nabla \times \psi \cdot \nabla \chi \, d\mathbf{x} \\
&= - \int_{\partial\Omega} \Delta \psi \nabla \times \psi \chi \cdot \mathbf{n} \, dS \overset{0}{=} \\
&\quad + \int_{\Omega} \Delta \psi \nabla \times \psi \cdot \nabla \chi \, d\mathbf{x} \\
&= \int_{\Omega} \Delta \psi (\psi_y \chi_x - \psi_x \chi_y) \, d\mathbf{x}.
\end{aligned}$$

Thus, the weak form of the QGE in its streamfunction form is

$$\begin{aligned}
\frac{\partial}{\partial t} \int_{\Omega} \nabla \psi \nabla \chi \, d\mathbf{x} + Re^{-1} \int_{\Omega} \Delta \psi \Delta \chi \, d\mathbf{x} + \int_{\Omega} \Delta \zeta (\psi_y \chi_x - \psi_x \chi_y) \, d\mathbf{x} - Ro^{-1} \int_{\Omega} \psi_x \chi \, d\mathbf{x} \\
= Ro^{-1} \int_{\Omega} F \chi \, d\mathbf{x} \quad \forall \chi \in X.
\end{aligned} \tag{3.6}$$

Therefore, taking (\cdot, \cdot) to be the L^2 inner product and letting

$$b(\psi; \psi, \chi) = \int_{\Omega} \Delta \psi (\psi_y \chi_x - \psi_x \chi_y) \, d\mathbf{x} \tag{3.7}$$

gives the following strong formulation of the QGE in streamfunction formulation:

Find $\psi \in X$ such that

$$\frac{\partial}{\partial t} (\nabla \psi, \nabla \chi) + Re^{-1} (\Delta \psi, \Delta \chi) + b(\psi; \psi, \chi) - Ro^{-1} (\psi_x, \chi) = Ro^{-1} (F, \chi), \quad \forall \chi \in X. \tag{3.8}$$

3.2 SQGE Weak Formulation

We are now ready to derive the weak formulation of the SQGE in streamfunction formulation (2.29). To this end, we first introduce the appropriate functional setting. As in Section 3.1, let

$$X := H_0^2(\Omega) = \left\{ \psi \in H^2(\Omega) : \psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0 \text{ on } \partial\Omega \right\}.$$

The difference in the functional setting for the QGE and the SQGE stems from the lack of time derivative in the SQGE. However, in the spatial domain everything remains the same.

Thus, the derivation of the weak formulation for the SQGE follows immediately from the strong formulation for the QGE with the exception that we set the bilinear form $(\nabla\psi, \nabla\chi)$ to zero. Therefore, using the notation presented in Section 3.1, the weak form of the SQGE in streamfunction formulation is:

$$\begin{aligned} &\text{Find } \psi \in X \text{ such that} \\ &Re^{-1}(\Delta\psi, \Delta\chi) + b(\psi; \psi, \chi) - Ro^{-1}(\psi_x, \chi) = Ro^{-1}(F, \chi), \quad \forall \chi \in X. \end{aligned} \quad (3.9)$$

The linear form (F, χ) , the bilinear forms $(\Delta\psi, \Delta\chi)$, (ψ_x, χ) and the trilinear form $b(\psi; \psi, \chi)$ are still continuous [16] and have the same bounds as those for the QGE, i.e. (5.7), (5.8), (5.9), and (5.10).

3.3 Well-Posedness of SQGE

For small enough data, one can use the same type of arguments as in [40, 41] to prove that the QGE in streamfunction formulation (3.9) are well-posed [6, 47, 86]. In what follows, we will always assume that the small data condition involving Re , Ro and F , is satisfied and, thus, that there exists a unique solution ψ to (3.9).

Using a standard argument (see Theorem 2.1 in [16]), one can also establish the following stability estimate:

Theorem 1. *The solution ψ of (3.9) satisfies the following stability estimate:*

$$|\psi|_2 \leq Re Ro^{-1} \|F\|_{-2}. \quad (3.10)$$

Proof. Setting $\chi = \psi$ in (3.9), we get:

$$Re^{-1}(\Delta\psi, \Delta\psi) + b(\psi; \psi, \psi) - Ro^{-1}(\psi_x, \psi) = Ro^{-1}(F, \psi). \quad (3.11)$$

From the trilinear form

$$b(\zeta; \psi, \chi) = \int_{\Omega} \Delta\zeta (\psi_y \chi_x - \psi_x \chi_y) d\mathbf{x},$$

we see that $b(\zeta; \psi, \psi) = 0$ for all $\psi \in X$. Therefore, we have

$$b(\psi; \psi, \psi) = 0. \quad (3.12)$$

We also note that, applying Green's theorem, we have

$$\begin{aligned} (\psi_x, \psi) &= \iint_{\Omega} \frac{\partial\psi}{\partial x} dx dy = \frac{1}{2} \iint_{\Omega} \frac{\partial}{\partial x} (\psi^2) dx dy \\ &= \frac{1}{2} \iint_{\Omega} \left(\frac{\partial}{\partial x} (\psi^2) - \frac{\partial}{\partial x} (0) \right) dx dy = \frac{1}{2} \int_{\partial\Omega} 0 dx + \psi^2 dy \\ &= 0, \end{aligned} \quad (3.13)$$

where in the last equality in (3.13) we used that $\psi = 0$ on $\partial\Omega$ (since $\psi \in H_0^2(\Omega)$). Substituting (3.13) and (3.12) in (3.11) and using the Cauchy-Schwarz inequality, we get:

$$\begin{aligned} Re^{-1}(\Delta\psi, \Delta\psi) &= Ro^{-1}(F, \psi) \\ Re^{-1}|\psi|_2^2 &= Ro^{-1}(F, \psi), \end{aligned} \tag{3.14}$$

which is equivalent to

$$\begin{aligned} |\psi|_2 &\leq Re Ro^{-1} \sup_{\psi \in X} \frac{(F, \psi)}{|\psi|_2} \\ |\psi|_2 &\leq Re Ro^{-1} \|F\|_{-2}. \end{aligned} \tag{3.15}$$

and thus we have proven (3.11). □

Chapter 4

Finite Element Discretization

In this chapter we introduce the conforming FE discretization of the QGE (Section 4.1), and the SQGE (Section 4.2). Additionally, an in depth discussion of the Argyris element and the transformation developed by Dominguez et. al. [27] is presented in Section 4.3. Finally, we introduce the two-level method and apply it to the SQGE in Section 4.4.

4.1 QGE Finite Element Discretization

In this section, we present the functional setting and some auxiliary results for the FE discretization of the streamfunction formulation of the QGE (3.8). Let \mathcal{T}^h denote a finite element triangulation of Ω with mesh size (maximum triangle diameter) h . We consider a *conforming* FE discretization of (3.8), i.e., let X^h be piecewise polynomials such that $X^h \subset X = H_0^2(\Omega)$.

The FE discretization of the streamfunction formulation of the QGE (3.8) reads:

$$\begin{aligned} &\text{Find } \psi^h \in X^h \text{ such that} \\ &\frac{\partial}{\partial t}(\nabla \psi^h, \nabla \chi^h) + Re^{-1}(\Delta \psi^h, \Delta \chi^h) + b(\psi^h, \psi^h, \chi^h) - Ro^{-1}(\psi_x^h, \chi^h) = Ro^{-1}(F, \chi^h), \quad \forall \chi^h \in X^h. \end{aligned} \tag{4.1}$$

Using standard arguments [40, 41], one can prove that, if the small data condition used in proving the well-posedness result for the continuous case holds, then (4.1) has a unique solution ψ^h (see Theorem 2.1 in [16]).

As noted in Section 6.1 in [20] (see also Section 13.2 in [43], Section 3.1 in [49], and Theorem 5.2 in [11]), in order to develop a conforming FE discretization for the QGE (3.8), we are faced with the problem of constructing subspaces of the space $H_0^2(\Omega)$. Since the standard, piecewise polynomial FE spaces are locally regular, this construction amounts in practice to finding FE spaces X^h that satisfy the inclusion $X^h \subset C^1(\overline{\Omega})$, i.e., finding C^1 finite elements.

Several finite elements meet this requirement (see, e.g., Section 6.1 in [20], Section 13.2 in [43], and Section 5 in [11]): the Argyris triangular element, the Bell triangular element, the Hsieh-Clough-Tocher triangular element (a macroelement), and the Bogner-Fox-Schmit rectangular element).

A description of the particular finite element we are using (the Argyris triangle Figure 4.1) and its transformation will be thoroughly discussed in Section 4.3). Additionally, we note that (4.1) is only a semi-discretization, since the formulation is still continuous in time, but discretized in space. We apply the *method of lines* (MoL) in the time domain, i.e. use a *finite difference* approximation for the time derivative.

4.2 SQGE Finite Element Discretization

In this section, we present the functional setting and some auxiliary results for the FE discretization of the streamfunction formulation of the SQGE (3.9). Let \mathcal{T}^h denote a finite element triangulation of Ω with mesh size (maximum triangle diameter) h . We consider a *conforming* FE discretization of (3.9), i.e., $X^h \subset X = H_0^2(\Omega)$.

The FE discretization of the streamfunction formulation of the SQGE (3.9) reads:

$$\begin{aligned} &\text{Find } \psi^h \in X^h \text{ such that} \\ &Re^{-1}(\Delta\psi^h, \Delta\chi^h) + b(\psi^h, \psi^h, \chi^h) - Ro^{-1}(\psi_x^h, \chi^h) = Ro^{-1}(F, \chi^h), \quad \forall \chi^h \in X^h. \end{aligned} \quad (4.2)$$

Using standard arguments [40, 41], one can prove that, if the small data condition used in proving the well-posedness result for the continuous case holds, then (4.2) has a unique solution ψ^h (see Theorem 2.1 in [16]). Furthermore, one can prove the following stability result for ψ^h using the same arguments as those used in the proof of (1) for the continuous setting.

Theorem 2. *The solution ψ^h of (4.2) satisfies the following stability estimate:*

$$|\psi^h|_2 \leq Re Ro^{-1} \|F\|_{-2}. \quad (4.3)$$

Proof. The proof is almost identical to the proof of Theorem 1, but is given here for completeness.

Let $\chi^h = \psi^h$ in (4.2) which gives

$$Re^{-1}(\Delta\psi^h, \Delta\psi^h) + b(\psi^h, \psi^h, \psi^h) - Ro^{-1}(\psi_x^h, \psi^h) = Ro^{-1}(F, \psi^h) \quad \forall \psi^h \in X^h.$$

Since, $b(\psi^h, \psi^h, \psi^h) = 0$ and $(\psi_x^h, \psi^h) = 0$ we have

$$\begin{aligned} Re^{-1}(\Delta\psi^h, \Delta\psi^h) &= Ro^{-1}(F, \psi^h) \\ Re^{-1} \|\psi^h\|_2^2 &= Ro^{-1}(F, \psi^h) \\ \|\psi^h\|_2 &\leq Re Ro^{-1} \sup_{\psi^h \in X^h} \frac{(F, \psi^h)}{|\psi^h|_2} \\ \|\psi^h\|_2 &\leq Re Ro^{-1} \|F\|_{-2}. \end{aligned}$$

Thus, the proof is complete. □

Again, we point out that as noted in Section 6.1 in [20] (see also Section 13.2 in [43], Section 3.1 in [49], and Theorem 5.2 in [11]), in order to develop a conforming FE discretization for the SQGE (3.9), we are faced with the problem of constructing subspaces of the space $H_0^2(\Omega)$. As was discussed previously the Argyris finite element is an element of class C^1 and therefore will be the FE used in this thesis for the discretization of the SQGE.

4.3 Argyris Element

If we are to require conforming finite elements, Lagrange finite elements are not enough to guarantee continuity in the first derivative, and so to ensure continuity in the first derivative [49] a C^1 FE will be required. In this thesis we focus on the Argyris element (depicted in Figure 4.1), although the FE error analysis presented in Section 5.1, Section 5.2, and Section 5.3 only requires the use of a C^1 FE. The Argyris element is probably the best known of all C^1 finite elements [1, 27], but appears to be rarely implemented.

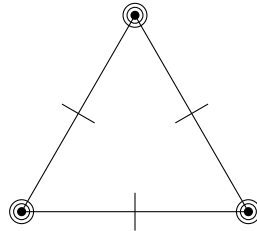


Figure 4.1: Argyris element with its 21 degrees of freedom.

The Argyris element does, in fact, ensure C^1 continuity [27, 63], but at a cost of twenty-one degrees of freedom. However, these twenty-one degrees of freedom give basis polynomials of degree five and therefore the FE discretization with the Argyris element has a high rate of convergence [27]. These degrees of freedom include the value at each vertex, the value of the first derivatives at each vertex, the value of the second derivatives at each vertex, the value of the mixed derivative at each vertex, and finally the value of the normal derivatives at each of the edge midpoints. Here in lies the main difficulty in implementing the Argyris triangle, the normal derivatives. Not only do we now have 21 degrees of freedom that we must worry about, but the added complexity of a transformation that maintains the direction of the normal derivatives is required. Since working on the reference element is the most common way of working with finite elements and normal derivatives are not respected by affine transformation a more complicated transformation will have to be employed. This is unlike a standard Lagrange element where only a simple affine transformation is required [27]. Dominguez *et al.* developed such a transformation. This transformation, which is a 21×21 matrix, C , allows for all calculations to be done on a reference element [27], vastly simplifying the calculations for the various matrices and load vector required for the QGE finite elements calculations and allowing for faster running code. Since the transformation developed by Dominguez is so new, and is not the standard for implementation of the Argyris element, his transformation will be discussed at length in Subsection 4.3.2 for completeness.

4.3.1 Basis Functions

In most references, which present the Argyris element, only the constraints found in Table 4.1 are presented. Thus, for completeness, we derive and explicitly state the the basis functions for the Argyris FE in this subsection. The Argyris triangle has 21 degrees of freedom and therefore has 21 basis functions per triangle. Additionally, the basis for Argyris in reference

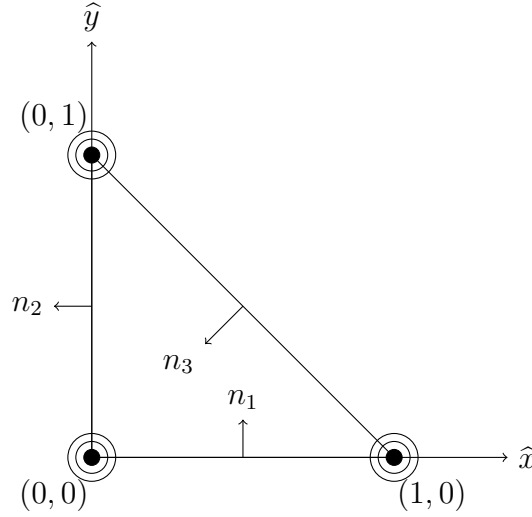


Figure 4.2: The reference triangle \widehat{K} for the Argyris element.

coordinates (depicted in Figure 4.2) belong to the space \mathbb{P}_5 , which has the standard monomial basis

$$\{1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, x^3y, x^2y^2, xy^3, y^4, x^5, x^4y, x^3y^2, x^2y^3, xy^4, y^5\}.$$

Thus, the i^{th} basis function for the Argyris triangle can be written as

$$\begin{aligned} \widehat{\varphi}_i(\widehat{x}, \widehat{y}) = & m_1^i + m_2^i \widehat{x} + m_3^i \widehat{x}^2 + m_4^i \widehat{x}^3 + m_5^i \widehat{x}^4 + m_6^i \widehat{x}^5 + m_7^i \widehat{y} + m_8^i \widehat{y}^2 + m_9^i \widehat{y}^3 \\ & + m_{10}^i \widehat{y}^4 + m_{11}^i \widehat{y}^5 + m_{12}^i \widehat{x}\widehat{y} + m_{13}^i \widehat{x}\widehat{y}^2 + m_{14}^i \widehat{x}\widehat{y}^3 + m_{15}^i \widehat{x}\widehat{y}^4 + m_{16}^i \widehat{x}^2\widehat{y} \\ & + m_{17}^i \widehat{x}^2\widehat{y}^2 + m_{18}^i \widehat{x}^2\widehat{y}^3 + m_{19}^i \widehat{x}^3\widehat{y} + m_{20}^i \widehat{x}^3\widehat{y}^2 + m_{21}^i \widehat{x}^4\widehat{y}. \end{aligned} \quad (4.4)$$

Now, consider the reference triangle \widehat{K} in Figure 4.2 with vertices numbered counterclockwise 1, 2, and 3, i.e. $(\widehat{x}_1, \widehat{y}_1) = (0, 0)$, $(\widehat{x}_2, \widehat{y}_2) = (1, 0)$, and $(\widehat{x}_3, \widehat{y}_3) = (0, 1)$. Additionally, let the vector v_i represent the i^{th} edge with

$$v_1 = [\widehat{x}_2 - \widehat{x}_1, \widehat{y}_2 - \widehat{y}_1]^T, \quad v_2 = [\widehat{x}_3 - \widehat{x}_1, \widehat{y}_3 - \widehat{y}_1]^T \text{ and } v_3 = [\widehat{x}_3 - \widehat{x}_2, \widehat{y}_3 - \widehat{y}_2]^T.$$

For book keeping purposes, we will use the convention that the i^{th} normal vector is the rotation of the v_i counter-clockwise 90° . This is the same convention used by Dominguez et. al. [27]. Therefore, the i^{th} basis function can be found using the restriction in Table 4.1.

Now, let z be the vector containing the monomial basis for \mathbb{P}_5 , i.e.

$$z = [1, \widehat{x}, \widehat{y}, \widehat{x}^2, \widehat{x}\widehat{y}, \widehat{y}^2, \widehat{x}^3, \widehat{x}^2\widehat{y}, \widehat{x}\widehat{y}^2, \widehat{y}^3, \widehat{x}^4, \widehat{x}^3\widehat{y}, \widehat{x}^2\widehat{y}^2, \widehat{x}\widehat{y}^3, \widehat{y}^4, \widehat{x}^5, \widehat{x}^4\widehat{y}, \widehat{x}^3\widehat{y}^2, \widehat{x}^2\widehat{y}^3, \widehat{x}\widehat{y}^4, \widehat{y}^5]^T.$$

Then the i^{th} Argyris basis function on \widehat{K} is given by

$$\widehat{\varphi}_i = M_i z,$$

$i, j, k = 1, 2, 3$	$i = 4, 5, 6$ and $j, k = 1, 2, 3$	$i = 7, 8, 9$ and $j, k = 1, 2, 3$	$i = 10, 11, 12$ and $j, k = 1, 2, 3$
$\varphi_i(x_j, y_j) = \delta_{i,j}$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$	$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = \delta_{i,j+3}$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$	$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = \delta_{i,j+6}$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$	$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = \delta_{i,j+9}$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$
$i = 13, 14, 15$ and $j, k = 1, 2, 3$	$i = 16, 17, 18$ and $j, k = 1, 2, 3$	$i = 19, 20, 21$ and $k = 1, 2, 3$	
$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$	$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = \delta_{i,j+15}$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = 0$	$\varphi_i(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial x}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x^2}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial x \partial y}(x_j, y_j) = 0$ $\frac{\partial^2 \varphi_i}{\partial y^2}(x_j, y_j) = 0$ $\frac{\partial \varphi_i}{\partial n_k} = \delta_{i,k-18}$	

Table 4.1: Constraints for Argyris triangle [1, 13, 20, 27], where i, j, k correspond to the various degrees of freedom of the Argyris element.

where M_i is the i^{th} row of the matrix M . Therefore, the evaluation of $\widehat{\varphi}_i(\widehat{x}, \widehat{y})$ comes down to the matrix-vector multiplication

$$\widehat{\varphi}_i(\widehat{x}, \widehat{y}) = M_i z(\widehat{x}, \widehat{y}).$$

To determine the matrix M we must solve the linear system

$$ZM^T = I_{21}$$

that results from the constraints in Table 4.1. Here

$$\begin{aligned} Z = [& z(0, 0), z(0, 1), z(1, 0), z_{\widehat{x}}(0, 0), z_{\widehat{y}}(0, 0), z_{\widehat{x}}(1, 0), z_{\widehat{y}}(1, 0), z_{\widehat{x}}(0, 1), z_{\widehat{y}}(0, 1), \\ & z_{\widehat{xx}}(0, 0), z_{\widehat{xy}}(0, 0), z_{\widehat{yy}}(0, 0), z_{\widehat{xx}}(1, 0), z_{\widehat{xy}}(1, 0), z_{\widehat{yy}}(1, 0), \\ & z_{\widehat{xx}}(0, 1), z_{\widehat{xy}}(0, 1), z_{\widehat{yy}}(0, 1), z_{\widehat{y}}^{(1/2, 0)}, -z_{\widehat{x}}(0, 1/2), \\ & -\frac{1}{\sqrt{2}}(z_{\widehat{x}}^{(1/2, 1/2)} + z_{\widehat{y}}^{(1/2, 1/2)})]^T. \end{aligned}$$

I_{21} is a 21×21 identity matrix and M is the matrix containing the coefficients for all 21 basis functions. Therefore solving this system will result in a matrix, M , that contains the coefficients for the Argyris basis functions for the reference triangle as in (4.4). Thus, the matrix M can be seen in Table 4.2.

1	0	0	0	0	0	-10	0	0	-10	15	0	-30	0	15	-6	0	30	0	-6
0	0	0	0	0	10	0	0	0	-15	0	15	0	0	0	6	0	-15	0	0
0	0	0	0	0	0	0	0	0	10	0	15	0	-15	0	0	0	-15	0	6
0	1	0	0	0	-6	0	-11	0	8	0	10	18	0	0	-3	0	1	-10	0
0	0	1	0	0	0	-11	0	-6	0	0	18	10	0	8	0	-8	-10	1	-3
0	0	0	0	0	-4	0	0	0	7	0	7	$-\frac{7}{2}$	0	0	-3	0	$\frac{7}{2}$	0	0
0	0	0	0	0	0	-5	0	0	0	14	$\frac{37}{2}$	0	0	0	0	-8	$-\frac{37}{2}$	0	0
0	0	0	0	0	0	0	-5	0	0	0	$\frac{37}{2}$	14	0	0	0	0	$-\frac{37}{2}$	-8	0
0	0	0	0	0	0	0	0	-4	0	0	$-\frac{7}{2}$	7	0	0	0	0	$\frac{7}{2}$	0	-3
0	0	0	$\frac{1}{2}$	0	$-\frac{3}{2}$	0	0	0	$\frac{3}{2}$	0	$-\frac{3}{2}$	0	0	0	$-\frac{1}{2}$	0	$\frac{3}{2}$	1	0
0	0	0	0	1	0	-4	0	-4	0	0	5	10	5	0	0	-2	-6	-2	0
0	0	0	0	0	$\frac{1}{2}$	0	0	0	$-\frac{3}{2}$	0	0	$-\frac{3}{2}$	0	$\frac{3}{2}$	0	0	1	0	$-\frac{1}{2}$
0	0	0	0	0	0	$\frac{1}{2}$	0	0	0	-1	0	$\frac{1}{4}$	0	0	$\frac{1}{2}$	0	$-\frac{1}{4}$	0	0
0	0	0	0	0	0	0	1	0	0	0	-3	$-\frac{7}{2}$	0	0	0	2	$\frac{7}{2}$	0	0
0	0	0	0	0	0	0	0	0	0	0	0	$\frac{5}{4}$	0	0	0	0	$-\frac{3}{4}$	0	0
0	0	0	0	0	0	0	0	0	0	0	0	$\frac{5}{4}$	0	0	0	0	$-\frac{5}{4}$	0	0
0	0	0	0	0	0	0	1	0	0	0	0	$-\frac{7}{2}$	-3	0	0	0	$-\frac{3}{4}$	0	0
0	0	0	0	0	0	0	0	0	$\frac{1}{2}$	0	0	$\frac{1}{4}$	0	-1	0	0	$\frac{5}{2}$	2	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$-\frac{1}{4}$	0	$\frac{1}{2}$
0	0	0	0	0	0	16	0	0	0	-32	-32	0	0	0	0	16	32	0	0
0	0	0	0	0	0	0	-16	0	0	0	32	32	0	0	0	0	-16	-32	-16
0	0	0	0	0	0	0	0	0	0	0	0	$8\sqrt{2}$	0	0	0	0	$-8\sqrt{2}$	0	0

Table 4.2: Coefficients for the Argyris basis matrix, M .

Therefore, the Argyris basis functions, on the reference element \widehat{K} , are

$$\begin{aligned}
\widehat{\varphi}_1(\widehat{x}, \widehat{y}) &= 1 - 10\widehat{x}^3 - 10\widehat{y}^3 + 15\widehat{x}^4 - 30\widehat{x}^2\widehat{y}^2 + 15\widehat{y}^4 - 6\widehat{x}^5 + 30\widehat{x}^3\widehat{y}^2 + 30\widehat{x}^2\widehat{y}^3 - 6\widehat{y}^5 \\
\widehat{\varphi}_2(\widehat{x}, \widehat{y}) &= 10\widehat{x}^3 - 15\widehat{x}^4 + 15\widehat{x}^2\widehat{y}^2 + 6\widehat{x}^5 - 15\widehat{x}^3\widehat{y}^2 - 15\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_3(\widehat{x}, \widehat{y}) &= 10\widehat{y}^3 + 15\widehat{x}^2\widehat{y}^2 - 15\widehat{y}^4 - 15\widehat{x}^3\widehat{y}^2 - 15\widehat{x}^2\widehat{y}^3 + 6\widehat{y}^5 \\
\widehat{\varphi}_4(\widehat{x}, \widehat{y}) &= \widehat{x} - 6\widehat{x}^3 - 11\widehat{x}\widehat{y}^2 + 8\widehat{x}^4 + 10\widehat{x}^2\widehat{y}^2 + 18\widehat{x}\widehat{y}^3 - 3\widehat{x}^5 + \widehat{x}^3\widehat{y}^2 - 10\widehat{x}^2\widehat{y}^3 - 8\widehat{x}\widehat{y}^4 \\
\widehat{\varphi}_5(\widehat{x}, \widehat{y}) &= \widehat{y} - 11\widehat{x}^2\widehat{y} - 6\widehat{y}^3 + 18\widehat{x}^3\widehat{y} + 10\widehat{x}^2\widehat{y}^2 + 8\widehat{y}^4 - 8\widehat{x}^4\widehat{y} - 10\widehat{x}^3\widehat{y}^2 + \widehat{x}^2\widehat{y}^3 - 3\widehat{y}^5 \\
\widehat{\varphi}_6(\widehat{x}, \widehat{y}) &= -4\widehat{x}^3 + 7\widehat{x}^4 - \frac{7}{2}\widehat{x}^2\widehat{y}^2 - 3\widehat{x}^5 + \frac{7}{2}\widehat{x}^3\widehat{y}^2 + \frac{7}{2}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_7(\widehat{x}, \widehat{y}) &= -5\widehat{x}^2\widehat{y} + 14\widehat{x}^3\widehat{y} + \frac{37}{2}\widehat{x}^2\widehat{y}^2 - 8\widehat{x}^4\widehat{y} - \frac{37}{2}\widehat{x}^3\widehat{y}^2 - \frac{27}{2}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_8(\widehat{x}, \widehat{y}) &= -5\widehat{x}\widehat{y}^2 + \frac{37}{2}\widehat{x}^2\widehat{y}^2 + 14\widehat{x}\widehat{y}^3 - \frac{27}{2}\widehat{x}^3\widehat{y}^2 - \frac{37}{2}\widehat{x}^3\widehat{y}^2 - 8\widehat{x}\widehat{y}^4 \\
\widehat{\varphi}_9(\widehat{x}, \widehat{y}) &= -4\widehat{y}^3 - \frac{7}{2}\widehat{x}^3 + 7\widehat{y}^4 + \frac{7}{2}\widehat{x}^3\widehat{y}^2 + \frac{7}{2}\widehat{x}^2\widehat{y}^3 - 3\widehat{y}^5 \\
\widehat{\varphi}_{10}(\widehat{x}, \widehat{y}) &= \frac{1}{2}\widehat{x}^2 - \frac{3}{2}\widehat{x}^3 + \frac{3}{2}\widehat{x}^4 - \frac{3}{2}\widehat{x}^2\widehat{y}^2 - \frac{1}{2}\widehat{x}^5 + \frac{3}{2}\widehat{x}^3\widehat{y}^2 + \widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{11}(\widehat{x}, \widehat{y}) &= \widehat{x}\widehat{y} - 4\widehat{x}^2\widehat{y} - 4\widehat{x}\widehat{y}^2 + 5\widehat{x}^3\widehat{y} + 10\widehat{x}^2\widehat{y}^2 + 5\widehat{x}\widehat{y}^3 - 2\widehat{x}^4\widehat{y} - 6\widehat{x}^3\widehat{y}^2 - 6\widehat{x}^2\widehat{y}^3 - 2\widehat{x}\widehat{y}^4 \\
\widehat{\varphi}_{12}(\widehat{x}, \widehat{y}) &= \frac{1}{2}\widehat{y}^2 - \frac{3}{2}\widehat{y}^3 - \frac{3}{2}\widehat{x}^2\widehat{y}^2 + \frac{3}{2}\widehat{y}^4 + \widehat{x}^3\widehat{y}^2 + \frac{3}{2}\widehat{x}^2\widehat{y}^3 - \frac{1}{2}\widehat{y}^5 \\
\widehat{\varphi}_{13}(\widehat{x}, \widehat{y}) &= \frac{1}{2}\widehat{x}^3 - \widehat{x}^4 + \frac{1}{4}\widehat{x}^2\widehat{y}^2 + \frac{1}{2}\widehat{x}^5 - \frac{1}{4}\widehat{x}^3\widehat{y}^2 - \frac{1}{4}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{14}(\widehat{x}, \widehat{y}) &= \widehat{x}^2\widehat{y} - 3\widehat{x}^3\widehat{y} - \frac{7}{2}\widehat{x}^2\widehat{y}^2 + 2\widehat{x}^4\widehat{y} + \frac{7}{2}\widehat{x}^3\widehat{y}^2 + \frac{5}{2}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{15}(\widehat{x}, \widehat{y}) &= \frac{5}{4}\widehat{x}^2\widehat{y}^2 - \frac{3}{4}\widehat{x}^3\widehat{y}^2 - \frac{5}{4}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{16}(\widehat{x}, \widehat{y}) &= \frac{5}{4}\widehat{x}^2\widehat{y}^2 - \frac{5}{4}\widehat{x}^3\widehat{y}^2 - \frac{3}{4}\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{17}(\widehat{x}, \widehat{y}) &= \widehat{x}\widehat{y}^2 - \frac{7}{2}\widehat{x}^2\widehat{y}^2 - 3\widehat{x}\widehat{y}^3 + \frac{5}{2}\widehat{x}^3\widehat{y}^2 + \frac{7}{2}\widehat{x}^2\widehat{y}^3 + 2\widehat{x}\widehat{y}^4 \\
\widehat{\varphi}_{18}(\widehat{x}, \widehat{y}) &= \frac{1}{2}\widehat{y}^3 + \frac{1}{4}\widehat{x}^2\widehat{y}^2 - \widehat{y}^4 - \frac{1}{4}\widehat{x}^3\widehat{y}^2 - \frac{1}{4}\widehat{x}^2\widehat{y}^3 + \frac{1}{2}\widehat{y}^5 \\
\widehat{\varphi}_{19}(\widehat{x}, \widehat{y}) &= 16\widehat{x}^2\widehat{y} - 32\widehat{x}^3\widehat{y} - 32\widehat{x}^2\widehat{y}^2 + 16\widehat{x}^4\widehat{y} + 32\widehat{x}^3\widehat{y}^2 + 16\widehat{x}^2\widehat{y}^3 \\
\widehat{\varphi}_{20}(\widehat{x}, \widehat{y}) &= -16\widehat{x}\widehat{y}^2 + 32\widehat{x}^2\widehat{y}^2 + 32\widehat{x}\widehat{y}^3 - 16\widehat{x}^3\widehat{y}^2 - 32\widehat{x}^2\widehat{y}^3 - 16\widehat{x}\widehat{y}^4 \\
\widehat{\varphi}_{21}(\widehat{x}, \widehat{y}) &= \sqrt{2} \left(8\widehat{x}^2\widehat{y}^2 - 8\widehat{x}^3\widehat{y}^2 - 8\widehat{x}^2\widehat{y}^3 \right).
\end{aligned} \tag{4.5}$$

Remark 2. Throughout our search for an explicit statement of the Argyris basis functions we only came across two reference, [84] and [68], both of which are online references. However, this is not to mean that a reference containining the explicit statement of the Argyris basis functions doesn't exist, but it certainly eluded the author of this thesis.

4.3.2 Transformation

For completeness we present the transformation developed by Dominguez et. al. in this section. All notations and conventions are the same as those developed by Dominguez, but we present the derivation of the transformation here, because of the novelty and importance of this transformation to this thesis. For the original derivation one should see [26] and [27]. Now that we have the Argyris

basis functions on the reference triangle we need to relate those to the Argyris basis functions on a general triangle. First, consider the affine transformation from \widehat{K} to K , i.e. $F : \widehat{K} \rightarrow K$ such that

$$F(\widehat{\mathbf{x}}) = B\widehat{\mathbf{x}} + \mathbf{b} := \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{bmatrix} \begin{bmatrix} \widehat{x} \\ \widehat{y} \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}. \quad (4.6)$$

Also denote the vectors representing the edge of the triangle as

$$\mathbf{v}_1 = \mathbf{x}_2 - \mathbf{x}_1 \quad \mathbf{v}_2 = \mathbf{x}_3 - \mathbf{x}_1 \quad \mathbf{v}_3 = \mathbf{x}_3 - \mathbf{x}_2.$$

Let \mathbf{n}_i be the unit normal vector corresponding to the i^{th} side obtained by rotating the vector \mathbf{v}_i by $\pi/2$ in the positive direction. Additionally, let \mathbf{m}_i be the midpoint to the i^{th} side. Now consider the linear functionals corresponding to the vertices of the triangle K denoted as

$$\begin{aligned} \mathcal{L}_i(\varphi) &:= \varphi(\mathbf{x}_i), \\ \mathcal{L}_i^\circ(\varphi) &:= \partial_\circ \varphi(\mathbf{x}_i) \quad \text{if } \circ \in \{x, y\}, \\ \mathcal{L}_i^\circ(\varphi) &:= \partial_\circ \varphi(\mathbf{x}_i) \quad \text{if } \circ \in \{xx, xy, yy\} \end{aligned}$$

for $i \in \{1, 2, 3\}$. For the functionals corresponding to the sides define

$$\mathcal{L}_i^n(\varphi) := \nabla_{\mathbf{x}} \phi(\mathbf{m}_i) \cdot \mathbf{n}_i \quad i \in \{1, 2, 3\}.$$

Now renumber the linear functionals as \mathcal{L}_j for $j \in \{1, \dots, 21\}$ as they are listed below

$$\begin{aligned} &\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \\ &\mathcal{L}_1^x, \mathcal{L}_1^y, \mathcal{L}_2^x, \mathcal{L}_2^y, \mathcal{L}_3^x, \mathcal{L}_3^y, \\ &\mathcal{L}_1^{xx}, \mathcal{L}_1^{xy}, \mathcal{L}_1^{yy}, \mathcal{L}_2^{xx}, \mathcal{L}_2^{xy}, \mathcal{L}_2^{yy}, \mathcal{L}_3^{xx}, \mathcal{L}_3^{xy}, \mathcal{L}_3^{yy}, \\ &\mathcal{L}_1^n, \mathcal{L}_2^n, \mathcal{L}_3^n, \end{aligned}$$

where the linear functionals $\widehat{\mathcal{L}}_i$ and $\widehat{\mathcal{L}}_i^n$ are the corresponding functionals on the reference triangle \widehat{K} . The basis functions of the Argyris triangle are fifth degree polynomials in $\mathbb{P}_5(K)$ that satisfy

$$\mathcal{L}_i(\varphi_j) = \delta_{ij} \quad i, j \in \{1, \dots, 21\}$$

and similarly on the reference triangle

$$\widehat{\mathcal{L}}_i(\widehat{\varphi}_j) = \delta_{ij} \quad i, j \in \{1, \dots, 21\}.$$

Now define a new set of functionals by

$$\widetilde{\mathcal{L}}_i(\varphi) := \widehat{\mathcal{L}}_i(\varphi \circ F). \quad (4.7)$$

Since $\{\widetilde{\mathcal{L}}_i\}$ and $\{\widehat{\mathcal{L}}_i\}$ are both basis for the dual space of $\mathbb{P}_5(K)$ there is a nonsingular matrix C such that

$$\widetilde{\mathcal{L}}_i = \sum_{j=1}^{21} c_{ij} \mathcal{L}_j \quad i \in \{1, \dots, 21\}. \quad (4.8)$$

Therefore, it follows that (see [27])

$$\varphi_i \circ F = \sum_{j=1}^{21} c_{ij} \widehat{\varphi}_j \quad i \in \{1, \dots, 21\}. \quad (4.9)$$

Thus we have shown there exists a C which transforms any triangle K to a reference triangle \widehat{K} .

Now, we must determine a simple expression for matrix C . To do this we introduce a new set of linear functionals

$$\mathcal{L}_i^* \quad i \in \{1, \dots, 24\}.$$

We define these functionals in the following way: first we let $\mathcal{L}_i^* := \mathcal{L}_i$ $i = \{1, \dots, 18\}$, and then introduce the new functionals

$$\mathcal{L}_i^\circ \quad \circ \in \{\perp, \parallel\} \quad i \in \{1, 2, 3\}.$$

These functionals will be defined using the following relationship:

$$\mathcal{L}_i^\perp(\varphi) := \nabla_{\mathbf{x}}\varphi(\mathbf{m}_i) \cdot R\mathbf{v}_i, \quad \mathcal{L}_i^\parallel(\varphi) := \nabla_{\mathbf{x}}\varphi(\mathbf{m}_i) \cdot \mathbf{v}_i,$$

where

$$R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

is the matrix that rotates a vector $\pi/2$ counter-clockwise. These relationships will define a relationship of the edge and firsts derivatives at a midpoint to the directional derivatives, i.e. the normal derivative and the derivative pointing in the direction of the corresponding edge.

Now, order the linear functionals as $\mathcal{L}_1^\perp, \mathcal{L}_2^\perp, \mathcal{L}_3^\perp, \mathcal{L}_1^\parallel, \mathcal{L}_2^\parallel, \mathcal{L}_3^\parallel$. We can write each $\tilde{\mathcal{L}}_i$ as a linear combination of \mathcal{L}_j^* and therefore we have the relationship

$$\tilde{\mathcal{L}}_i = \sum_{j=1}^{24} d_{ij} \mathcal{L}_j^*, \quad i \in \{1, \dots, 21\}, \quad (4.10)$$

where d_{ij} are constants.

Now from (4.6) we see that

$$\begin{aligned} \frac{\partial x}{\partial \widehat{x}} &= B_{11} & \frac{\partial x}{\partial \widehat{y}} &= B_{12} \\ \frac{\partial y}{\partial \widehat{x}} &= B_{21} & \frac{\partial y}{\partial \widehat{y}} &= B_{22}. \end{aligned}$$

With this we can determine the relationship between the gradient of $\widehat{\varphi}$ on the triangle \widehat{K} to the gradient of φ on the triangle K

$$\begin{aligned} \nabla_{\widehat{\mathbf{x}}}(\varphi \circ F) &= \begin{bmatrix} \frac{\partial \varphi \circ F}{\partial \widehat{x}} \\ \frac{\partial \varphi \circ F}{\partial \widehat{y}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial \varphi \circ F}{\partial x} \frac{\partial x}{\partial \widehat{x}} + \frac{\partial \varphi \circ F}{\partial y} \frac{\partial y}{\partial \widehat{x}} \\ \frac{\partial \varphi \circ F}{\partial x} \frac{\partial x}{\partial \widehat{y}} + \frac{\partial \varphi \circ F}{\partial y} \frac{\partial y}{\partial \widehat{y}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial x}{\partial \widehat{x}} & \frac{\partial y}{\partial \widehat{x}} \\ \frac{\partial x}{\partial \widehat{y}} & \frac{\partial y}{\partial \widehat{y}} \end{bmatrix} \begin{bmatrix} \frac{\partial \varphi \circ F}{\partial x} \\ \frac{\partial \varphi \circ F}{\partial y} \end{bmatrix} \end{aligned}$$

thus we have

$$\nabla_{\widehat{\mathbf{x}}}(\varphi \circ F) = B^T \nabla_{\mathbf{x}} \varphi \circ F \quad (4.11)$$

If we define the condensed Hessian, in the same way as Dominguez et. al. [27], as $H_x(\varphi) = [\varphi_{xx}, \varphi_{xy}, \varphi_{yy}]^T$ then using (4.6) we can determine the relationship between the condensed Hessian of $\widehat{\varphi}$ on the triangle \widehat{K} to the condensed Hessian of φ on the triangle K

$$\begin{aligned} H_{\widehat{\mathbf{x}}}(\varphi \circ F) &= \begin{bmatrix} \frac{\partial^2 \varphi \circ F}{\partial \widehat{x}^2} \\ \frac{\partial^2 \varphi \circ F}{\partial \widehat{x} \partial \widehat{y}} \\ \frac{\partial^2 \varphi \circ F}{\partial \widehat{y}^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial \widehat{x}} \left(\frac{\partial \varphi \circ F}{\partial x} \frac{\partial x}{\partial \widehat{x}} + \frac{\partial \varphi \circ F}{\partial y} \frac{\partial y}{\partial \widehat{x}} \right) \\ \frac{\partial}{\partial \widehat{y}} \left(\frac{\partial \varphi \circ F}{\partial x} \frac{\partial x}{\partial \widehat{x}} + \frac{\partial \varphi \circ F}{\partial y} \frac{\partial y}{\partial \widehat{x}} \right) \\ \frac{\partial}{\partial \widehat{y}} \left(\frac{\partial \varphi \circ F}{\partial x} \frac{\partial x}{\partial \widehat{y}} + \frac{\partial \varphi \circ F}{\partial y} \frac{\partial y}{\partial \widehat{y}} \right) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial^2 \varphi \circ F}{\partial x^2} \left(\frac{\partial x}{\partial \widehat{x}} \right)^2 + 2 \frac{\partial^2 \varphi \circ F}{\partial x \partial y} \frac{\partial x}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{x}} + \frac{\partial^2 \varphi \circ F}{\partial y^2} \left(\frac{\partial y}{\partial \widehat{x}} \right)^2 \\ \frac{\partial^2 \varphi \circ F}{\partial x^2} \frac{\partial x}{\partial \widehat{x}} \frac{\partial x}{\partial \widehat{y}} + \left(\frac{\partial x}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{y}} + \frac{\partial x}{\partial \widehat{y}} \frac{\partial y}{\partial \widehat{x}} \right) \frac{\partial^2 \varphi \circ F}{\partial x \partial y} + \frac{\partial^2 \varphi \circ F}{\partial y^2} \frac{\partial y}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{y}} \\ \frac{\partial^2 \varphi \circ F}{\partial x^2} \left(\frac{\partial x}{\partial \widehat{y}} \right)^2 + 2 \frac{\partial^2 \varphi \circ F}{\partial x \partial y} \frac{\partial x}{\partial \widehat{y}} \frac{\partial y}{\partial \widehat{y}} + \frac{\partial^2 \varphi \circ F}{\partial y^2} \left(\frac{\partial y}{\partial \widehat{y}} \right)^2 \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{\partial x}{\partial \widehat{x}} \right)^2 & 2 \frac{\partial x}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{x}} & \left(\frac{\partial y}{\partial \widehat{x}} \right)^2 \\ \frac{\partial x}{\partial \widehat{x}} \frac{\partial x}{\partial \widehat{y}} & \left(\frac{\partial x}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{y}} + \frac{\partial x}{\partial \widehat{y}} \frac{\partial y}{\partial \widehat{x}} \right) & \frac{\partial y}{\partial \widehat{x}} \frac{\partial y}{\partial \widehat{y}} \\ \left(\frac{\partial x}{\partial \widehat{y}} \right)^2 & 2 \frac{\partial x}{\partial \widehat{y}} \frac{\partial y}{\partial \widehat{y}} & \left(\frac{\partial y}{\partial \widehat{y}} \right)^2 \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \varphi \circ F}{\partial x^2} \\ \frac{\partial^2 \varphi \circ F}{\partial x \partial y} \\ \frac{\partial^2 \varphi \circ F}{\partial y^2} \end{bmatrix} \end{aligned}$$

and so

$$H_{\widehat{\mathbf{x}}}(\varphi \circ F) = \Theta H_{\mathbf{x}}(\varphi) \circ F, \quad (4.12)$$

where

$$\Theta = \begin{bmatrix} B_{11}^2 & 2B_{11}B_{21} & B_{21}^2 \\ B_{12}B_{11} & B_{12}B_{21} + B_{11}B_{22} & B_{21}B_{22} \\ B_{12}^2 & 2B_{22}B_{12} & B_{22}^2 \end{bmatrix}.$$

Therefore we see that

$$\widetilde{\mathcal{L}}_i = \mathcal{L}_i, \quad \begin{bmatrix} \widetilde{\mathcal{L}}_i^x \\ \widetilde{\mathcal{L}}_i^y \end{bmatrix} = B^T \begin{bmatrix} \mathcal{L}_i^x \\ \mathcal{L}_i^y \end{bmatrix}, \quad \begin{bmatrix} \widetilde{\mathcal{L}}_i^{xx} \\ \widetilde{\mathcal{L}}_i^{xy} \\ \widetilde{\mathcal{L}}_i^{yy} \end{bmatrix} = \Theta \begin{bmatrix} \mathcal{L}_i^{xx} \\ \mathcal{L}_i^{xy} \\ \mathcal{L}_i^{yy} \end{bmatrix}.$$

Additionally, notice $\mathbf{v}_i = B\widehat{\mathbf{v}}_i$ for $i = \{1, 2, 3\}$ where $\widehat{\mathbf{v}}_i$ are defined on the reference triangle \widehat{K} . It should also be noted that

$$\begin{bmatrix} \mathcal{L}_i^\perp \\ \mathcal{L}_i^\parallel \end{bmatrix} = \begin{bmatrix} -v_i^y & v_i^x \\ v_i^x & v_i^y \end{bmatrix} \begin{bmatrix} \mathcal{L}_i^x \\ \mathcal{L}_i^y \end{bmatrix}$$

and therefore

$$\begin{bmatrix} \mathcal{L}_i^x \\ \mathcal{L}_i^y \end{bmatrix} = \frac{1}{|v_i|^2} \begin{bmatrix} -v_i^y & v_i^x \\ v_i^x & v_i^y \end{bmatrix} \begin{bmatrix} \mathcal{L}_i^\perp \\ \mathcal{L}_i^\parallel \end{bmatrix}.$$

Then

$$\begin{aligned} \tilde{\mathcal{L}}_i^n(\varphi) &= \widehat{\mathcal{L}}_i^n(\varphi \circ F) \\ &= \frac{1}{|\widehat{\mathbf{v}}_i|} R\widehat{\mathbf{v}}_i \cdot \nabla_{\widehat{\mathbf{x}}}(\varphi \circ F)(\widehat{\mathbf{m}}_i) \\ &= \frac{1}{|\widehat{\mathbf{v}}_i|} R\widehat{\mathbf{v}}_i \cdot B^T \nabla_{\mathbf{x}}\varphi(\mathbf{m}_i) \\ &= \frac{1}{|\widehat{\mathbf{v}}_i|} R\widehat{\mathbf{v}}_i \cdot \ell_i^{-2} B^T \begin{bmatrix} -v_i^y & v_i^x \\ v_i^x & v_i^y \end{bmatrix} \begin{bmatrix} \mathcal{L}_i^\perp \\ \mathcal{L}_i^\parallel \end{bmatrix}, \end{aligned}$$

where ℓ_i is the length of the i^{th} side of the triangle K . This can be written in the form

$$\tilde{\mathcal{L}}_i^n = f_i \mathcal{L}_i^\perp + g_i \mathcal{L}_i^\parallel,$$

where

$$f_i = \frac{1}{\ell_i^2 |\widehat{\mathbf{v}}_i|} R\widehat{\mathbf{v}}_i \cdot B^T R\mathbf{v}_i \quad g_i = \frac{1}{\ell_i^2 |\widehat{\mathbf{v}}_i|} R\widehat{\mathbf{v}}_i \cdot B^T \mathbf{v}_i.$$

Therefore, we can construct a 21×24 the matrix D in block diagonal form in the following way:

$$D = \text{diag}[I_3, B^T, B^T, B^T, \Theta, \Theta, \Theta, Q],$$

where

$$Q = \left[\begin{array}{cc|cc} f_1 & & g_1 & \\ & f_2 & & g_2 \\ & & f_3 & \\ & & & g_3 \end{array} \right].$$

It also holds that

$$\mathcal{L}_i^* = \sum_{j=1}^{21} e_{ij} \mathcal{L}_j \quad i \in \{1, \dots, 24\}. \quad (4.13)$$

For the degrees of freedom corresponding to the vertices, i.e. function values, first derivative values, and second/mixed derivative values, it should be clear that $\mathcal{L}_i^* = \mathcal{L}_i$ for $i = 1, \dots, 18$. For the degrees of freedom corresponding to the normal derivatives and tangential derivatives at the midpoints we see that since $\ell_i \mathbf{n}_i = R\mathbf{v}_i$ and $\mathcal{L}_i^\perp = \ell_i \mathcal{L}_i^n$. Now let ϕ be an arbitrary polynomial in $\mathbb{P}_5(K)$ and define

$$\psi(t) := \varphi(t\mathbf{x}_\beta + (1-t)\mathbf{x}_\alpha) \in \mathbb{P}_5(t) \quad \alpha < \beta \text{ and } \alpha, \beta \in \{1, 2, 3\}.$$

Then, we get

$$\psi'(1/2) = \frac{15}{8}(\psi(1) - \psi(0)) - \frac{7}{16}(\psi'(1) + \psi'(0)) + \frac{1}{32}(\psi''(1) - \psi''(0)). \quad (4.14)$$

Now take γ to be the index corresponding to $\mathbf{v}_\gamma = \mathbf{x}_\beta - \mathbf{x}_\alpha$ and since

$$\begin{aligned}\psi'(1/2) &= \mathcal{L}_\gamma(\varphi) \\ \psi(0) &= \mathcal{L}_\alpha(\varphi) \\ \psi(1) &= \mathcal{L}_\beta(\varphi) \\ \psi'(0) &= v_\gamma^x \mathcal{L}_\alpha^x(\varphi) + v_\gamma^y \mathcal{L}_\alpha^y(\varphi) \\ \psi'(1) &= v_\gamma^x \mathcal{L}_\beta^x(\varphi) + v_\gamma^y \mathcal{L}_\beta^y(\varphi) \\ \psi''(0) &= (v_\gamma^x)^2 \mathcal{L}_\alpha^{xx}(\varphi) + 2v_\gamma^x v_\gamma^y \mathcal{L}_\alpha^{xy}(\varphi) + (v_\gamma^y)^2 \mathcal{L}_\alpha^{yy}(\varphi) \\ \psi''(1) &= (v_\gamma^x)^2 \mathcal{L}_\beta^{xx}(\varphi) + 2v_\gamma^x v_\gamma^y \mathcal{L}_\beta^{xy}(\varphi) + (v_\gamma^y)^2 \mathcal{L}_\beta^{yy}(\varphi),\end{aligned}$$

we see that by applying this to (4.14) we get the following expression

$$\begin{aligned}\mathcal{L}_\gamma^{\parallel} &= \frac{15}{8}(-\mathcal{L}_\alpha + \mathcal{L}_\beta) - \frac{7}{16}(v_\gamma^x \mathcal{L}_\alpha^x(\varphi) + v_\gamma^y \mathcal{L}_\alpha^y(\varphi) + v_\gamma^x \mathcal{L}_\beta^x(\varphi) + v_\gamma^y \mathcal{L}_\beta^y(\varphi)) \\ &\quad + \frac{1}{32}(-(v_\gamma^x)^2 \mathcal{L}_\alpha^{xx}(\varphi) - 2v_\gamma^x v_\gamma^y \mathcal{L}_\alpha^{xy}(\varphi) - (v_\gamma^y)^2 \mathcal{L}_\alpha^{yy}(\varphi) \\ &\quad + (v_\gamma^x)^2 \mathcal{L}_\beta^{xx}(\varphi) + 2v_\gamma^x v_\gamma^y \mathcal{L}_\beta^{xy}(\varphi) + (v_\gamma^y)^2 \mathcal{L}_\beta^{yy}(\varphi)).\end{aligned}$$

Therefore, we have the 24×21 matrix

$$E = \begin{bmatrix} I_{18} & \mathbf{0} \\ \mathbf{0} & L \\ T & \mathbf{0} \end{bmatrix}.$$

We then can define

$$L = \text{diag}[\ell_1, \ell_2, \ell_3]$$

and the (3×18) block matrix T as being composed of three sub blocks $(3 \times 3, 3 \times 6, \text{ and } 3 \times 9)$ respectively)

$$\frac{15}{8} \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}, \quad -\frac{7}{16} \begin{bmatrix} \mathbf{v}_1^T & \mathbf{v}_1^T & \mathbf{0} \\ \mathbf{v}_2^T & \mathbf{0} & \mathbf{v}_2^T \\ \mathbf{0} & \mathbf{v}_3^T & \mathbf{v}_3^T \end{bmatrix}, \quad \frac{1}{32} \begin{bmatrix} -\mathbf{w}_1^T & \mathbf{w}_1^T & \mathbf{0} \\ -\mathbf{w}_2^T & \mathbf{0} & \mathbf{w}_2^T \\ \mathbf{0} & -\mathbf{w}_3^T & \mathbf{w}_3^T \end{bmatrix},$$

where $\mathbf{w}_i^T = [(v_i^x)^2, 2v_i^x v_i^y, (v_i^y)^2]$.

Finally, notice that by combining (4.10) and (4.13) we get the matrix in (4.8) and

$$C = DE. \tag{4.15}$$

Thus, we have the transformation for the Argyris triangle K into a reference triangle \widehat{K} .

4.3.3 Normal Derivatives

During the discussion above we have been concerned with transforming a basis on a local triangle to that of a basis on a reference triangle. However, we have not considered the implications that results from the assumption of always rotating the normal derivative $\frac{\pi}{2}$ counterclockwise. If we always rotate $\frac{\pi}{2}$ counterclockwise there will be a jump discontinuity in the basis function along that edge. This results from assuming $\frac{\partial \varphi_i}{\partial \mathbf{n}_i} = 1$ on triangle I while at the same time assuming

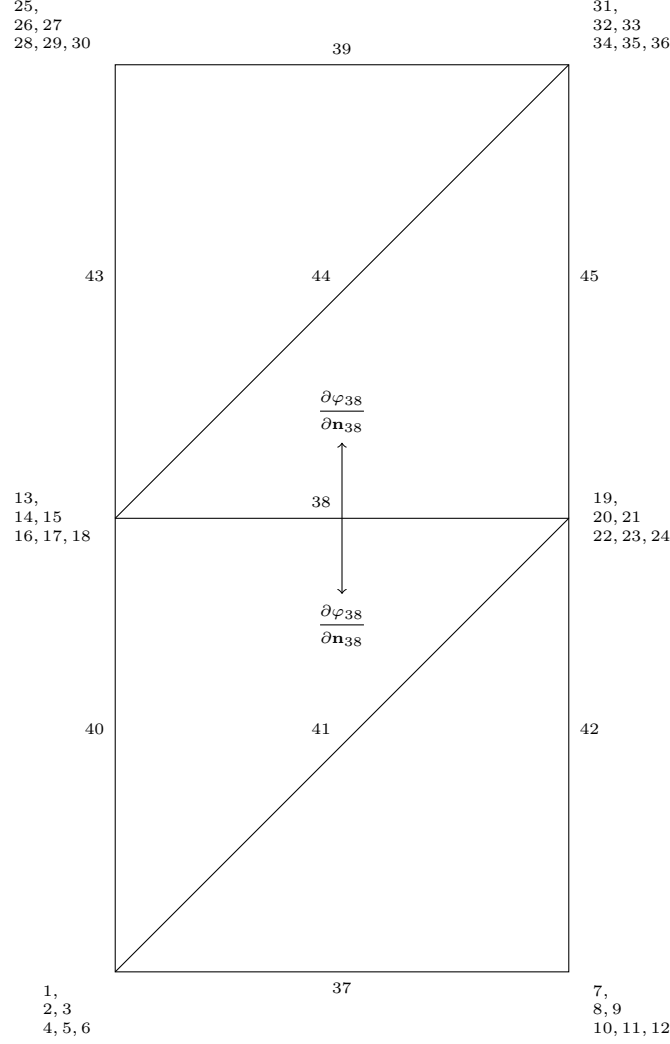


Figure 4.3: Illustration of continuity issue in normal derivatives for C^1 finite elements.

$\frac{\partial \varphi_i}{\partial \mathbf{n}_i} = -1$ on the adjacent triangle J . The potential issue can be seen graphically in Figure 4.3. This is what causes the discontinuity along shared edges. To address this one has two options: one can multiply the value of φ_i by negative one on the triangle J , without changing φ_i on triangle I , or one can avoid the issue all together by numbering the nodes in such a way that this mismatching of normal derivatives doesn't occur.

4.3.4 Interpolation Error

By using Theorem 6.1.1 and inequality (6.1.5) in [20], i.e.

$$\|\psi - I^h \psi\|_{m,q} \leq Ch^{k+1-m} |\psi|_{k+1,p}, \quad (4.16)$$

we obtain the following two approximation properties for the Argyris FE space X^h :

$$\forall \chi \in H^6(\Omega) \cap H_0^2(\Omega), \exists \chi^h \in X^h \quad \text{such that} \quad \|\chi - \chi^h\|_2 \leq C h^4 |\chi|_6, \quad (4.17)$$

$$\forall \chi \in H^4(\Omega) \cap H_0^2(\Omega), \exists \chi^h \in X^h \quad \text{such that} \quad \|\chi - \chi^h\|_2 \leq C h^2 |\chi|_4, \quad (4.18)$$

$$\forall \chi \in H^3(\Omega) \cap H_0^2(\Omega), \exists \chi^h \in X^h \quad \text{such that} \quad \|\chi - \chi^h\|_2 \leq C h |\chi|_3, \quad (4.19)$$

where C is a generic constant that can depend on the data, but not on the mesh size h . Approximation property (4.17) follows from inequality (6.1.5) in [20] with $q = 2$, $p = 2$, $m = 2$ and $k + 1 = 6$. Approximation property (4.18) follows from inequality (6.1.5) in [20] with $q = 2$, $p = 2$, $m = 2$ and $k + 1 = 4$. Approximation property (4.19) follows from inequality (6.1.5) in [20] with $q = 2$, $p = 2$, $m = 2$ and $k + 1 = 3$.

4.3.5 Boundary Conditions

Unlike Lagrange FEs *boundary conditions* (BCs) for C^1 FEs can be challenging to implement. We will demonstrate the difficulty of implementing BCs for C^1 FEs through an example. The example we will use is the *Poisson problem*, which would normally have a C^0 FE discretization. The motivation for using this problem is that it should be a very familiar implementation in C^0 and thus a useful demonstration of the difficulties of implementing C^1 FEs. Therefore, consider the following *Poisson problem*

$$\begin{aligned} -\Delta u &= f & \text{on } \Omega \\ u &= 0 & \text{on } \partial\Omega. \end{aligned} \quad (4.20)$$

The weak form of (4.20) is given by

$$\begin{aligned} \text{Find } u \in X &:= H_0^1(\Omega) \text{ such that} \\ (\nabla u, \nabla v) &= (f, v) \quad \forall v \in X. \end{aligned} \quad (4.21)$$

Thus, for conforming FEs the FE formulation is given by

$$\begin{aligned} \text{Find } u^h \in X^h &\subset X \text{ such that} \\ (\nabla u^h, \nabla v^h) &= (f, v^h) \quad \forall v^h \in X^h. \end{aligned} \quad (4.22)$$

Since $u \in H_0^1(\Omega)$ and $u^h \in X^h \subset H_0^1(\Omega)$ only C^0 Lagrange FEs need be used. However, for demonstration purposes we assume $u^h \in X^h \subset X \subset C^1(\Omega)$. Therefore, Lagrange FEs are no longer an appropriate choice for our FE discretization.

With Lagrange FEs the most common way to deal with BCs is to enforce them by explicitly setting the degrees of freedom (DoFs) on the boundary to zero and reduce the linear system to contain only the internal nodes. This can be seen graphically in Figure 4.4. One significant benefit of treating BCs in this way is that one reduces the size of the linear system that must be solved, and therefore one reduce computational time.

This method, however, can be problematic when using C^1 FEs. Consider the boundary $\partial\Omega = \Gamma_1 \cup \Gamma_2$, where Γ_1 and Γ_2 correspond to the vertical and horizontal boundaries, respectively (see Figure 4.5). On Γ_1 we know $u = 0$ which implies $u_y = u_{yy} = 0$ on Γ_1 . On Γ_2 we know $u = 0$ which implies $u_x = u_{xx} = 0$ on Γ_2 . Thus, simply setting $u^h = 0$ on the boundary results in the DoFs corresponding to u_x , u_y , u_{xx} , and u_{yy} not necessarily being set to zero. Therefore, one must force

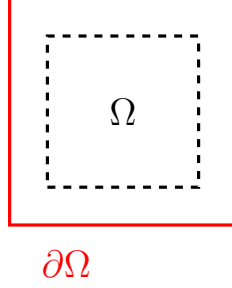


Figure 4.4: Eliminated boundary nodes in C^0 finite elements.

these DoFs to be zero too or another scheme must be implemented to deal with this case. However, setting the appropriate DoFs to zero in this manner requires some analysis of the problem domain before implementation, which is problematic when one wants to implement a generalized code. To read more about implementation of boundary conditions for Hermite type FEs see Remark 2.3.8 in Ciarlet [20]. To avoid analyzing boundary conditions for each domain one can implement a Lagrange Multiplier scheme [3, 4, 5, 12, 65, 66].

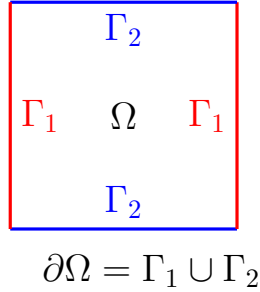


Figure 4.5: Boundary decomposed into vertical and horizontal boundaries.

We begin by rewriting the FE formulation, (4.22), in the following way

$$\begin{aligned} Au &= \ell \\ \Lambda u &= b, \end{aligned} \tag{4.23}$$

where $\Lambda u = b$ is the constraint equation describing the boundary conditions and $Au = \ell$ is the FE formulation of the Poisson equation. Then the *variational form* of (4.23) is

$$L(u, \lambda) = \frac{1}{2} u^T A u - u^T \ell + \lambda^T (\Lambda u - b), \tag{4.24}$$

where λ is the Lagrange multiplier. Using the first-order optimality condition results in

$$\begin{aligned} \frac{dL}{du} &= Au - \ell + \Lambda^T \lambda = 0 \\ \frac{dL}{d\lambda} &= \Lambda u - b = 0. \end{aligned} \tag{4.25}$$

Thus, given $b = \mathbf{0}$ the new system of equations using Lagrange multipliers for boundary constraints is given by

$$\begin{bmatrix} A & \Lambda^T \\ \Lambda & \mathbf{0} \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} \ell \\ \mathbf{0} \end{bmatrix}. \quad (4.26)$$

Since the homogeneous boundary conditions were not enforced in A , we note that A is singular. To avoid poor conditioning of the system (4.26), we add a small perturbation matrix, $\varepsilon \mathbf{I}$, in the lower right corner of the matrix of the linear system (4.26). We note that a similar approach (the penalty method) is used in the Navier-Stokes equations context (see, e.g., [43]).

4.4 Two-Level Method

Although the streamfunction formulation of the QGE contains only one flow variable, ψ , it still suffers from having to solve a large nonlinear system of equations. This is usually done by using a nonlinear solver, such as Newton's method. These nonlinear solvers typically require solving large linear systems multiple times to obtain the solution to the nonlinear system. Solving these large linear systems multiple times can be time consuming. Thus, a two-level algorithm can improve solve times greatly over the standard nonlinear solver, since we need only solve the nonlinear system on a coarse mesh and then use that solution to solve a linear system on a fine mesh. To this end, this section discusses the two-level finite element discretization for the pure streamfunction formulation of the SQGE and QGE.

4.4.1 Two-Level Algorithm

We consider an approximate solution to (3.9) by a two-level finite element procedure [30, 51]. Let $X^h, X^H \subset H_0^2(\Omega)$ denote two conforming finite element spaces with $H \gg h$. We compute an approximate solution ψ^h in the finite element space X^h by solving a linear system for the degrees of freedom in X^h . This linear system requires: first, the computation of the approximate solution ψ^H to the nonlinear system in the finite element space X^H , where the mesh is very coarse, i.e. $H \gg h$, and then using this solution, ψ^H , to linearize the problem on the fine mesh. This procedure is as follows:

Algorithm 1

Step 1: Solve the following nonlinear system on a coarse mesh for $\psi^H \in X^H$:

$$Re^{-1}(\Delta\psi^H, \Delta\chi^H) + b(\psi^H; \psi^H, \chi^H) - Ro^{-1}(\psi_x^H, \chi^H) = Ro^{-1}(F, \chi^H), \quad \text{for all } \chi^H \in X^H. \quad (4.27)$$

Step 2: Solve the following linear system on a fine mesh for $\psi^h \in X^h$:

$$Re^{-1}(\Delta\psi^h, \Delta\chi^h) + b(\psi^H; \psi^h, \chi^h) - Ro^{-1}(\psi_x^h, \chi^h) = Ro^{-1}(F, \chi^h), \quad \text{for all } \chi^h \in X^h. \quad (4.28)$$

Chapter 5

Error Analysis

The main goal of this Chapter is to develop rigorous numerical analysis for the conforming FE discretization of the QGE (2.26). In particular we focus on the use of the Argyris element, but the majority of the results apply to any conforming FE discretization of QGE (2.26), SQGE (2.29), and the two-level method applied to SQGE (2.29).

In the sections that follow we rely heavily on the Young inequality, Hölder inequality, Cauchy-Schwarz inequality, Poincaré-Friedrichs inequality, and Ladyzhenskaya inequality, so for completeness we will state them here.

Definition 5.1. Young inequality [71]:

For $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, and any two positive numbers a and b ,

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (5.1)$$

Definition 5.2. Hölder inequality [71]:

Let Ω be a measurable set, $1 \leq p, q < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then their product $f g$ is integrable over Ω and

$$\int_{\Omega} |f g| d\mathbf{x} \leq \|f\|_p \|g\|_q. \quad (5.2)$$

Definition 5.3. Cauchy-Schwarz inequality [71]:

Let Ω be a measurable set and f and g be square-integrable over Ω . Then their product $f g$ is also over Ω and

$$\int_{\Omega} |f g| d\mathbf{x} \leq \sqrt{\int_{\Omega} f^2 d\mathbf{x}} \sqrt{\int_{\Omega} g^2 d\mathbf{x}}. \quad (5.3)$$

Definition 5.4. Poincaré-Friedrichs inequality [52]:

Let Ω be a measurable set and $u \in H_0^1(\Omega)$. Then there is a positive constant $C_P = C_P(\Omega)$ such that

$$\|u\| \leq C_P \|\nabla u\| \quad \forall u \in X. \quad (5.4)$$

Definition 5.5. Ladyzhenskaya inequality [52]:

For any vector function $u : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with compact support and with the indicated finite L^4 and L^2 norms,

$$\|u\|_{L^4(\mathbb{R}^2)} \leq 2^{1/4} \|u\|_{L^2(\mathbb{R}^2)}^{1/2} \|\nabla u\|_{L^2(\mathbb{R}^2)}^{1/2}. \quad (5.5)$$

With these definitions we can now proceed with the error analysis.

5.1 SQGE

The main goal of this section is to develop a rigorous error analysis for the FE discretization of the SQGE (2.29) by using conforming FEs. In particular, we develop error bounds for the FE discretization of (2.29) with Argyris element in Theorem 4. First, in Theorem 3 we prove error estimates in the H^2 norm by using an approach similar to that used in [16]. Second, in Theorem 4, we prove error estimates in the L^2 and H^1 norms by using a duality argument.

The following lemma will introduce some useful bounds for the forms introduced in Section 3.1.

Lemma 1. *The linear form (F, χ) , the bilinear forms $(\nabla\psi, \nabla\chi)$, $(\Delta\psi, \Delta\chi)$, (ψ_x, χ) and the trilinear form $b(\psi; \psi, \chi)$ are continuous [16]: There exist $\Gamma_0, \Gamma_2 > 0$ such that for all $\psi, \chi, \varphi \in X$*

$$(\nabla\psi, \nabla\chi) \leq \Gamma_0 |\psi|_2 |\chi|_2, \quad (5.6)$$

$$(\Delta\psi, \Delta\chi) \leq |\psi|_2 |\chi|_2, \quad (5.7)$$

$$b(\psi; \varphi, \chi) \leq \Gamma_1 \|\psi\|_2 \|\varphi\|_2 \|\chi\|_2, \quad (5.8)$$

$$(\psi_x, \chi) \leq \Gamma_2 |\psi|_2 |\chi|_2, \quad (5.9)$$

$$(F, \chi) \leq \|F\|_{-2} |\chi|_2. \quad (5.10)$$

Proof. We first establish the most straightforward estimate (5.7). This follows directly from the Hölder inequality, (5.2), i.e.

$$(\Delta\psi, \Delta\chi) \leq \|\Delta\psi\|_{L^p} \|\Delta\chi\|_{L^r}, \quad \frac{1}{p} + \frac{1}{r} = 1$$

with $p = r = 2$.

The bound (5.10) can be proven by first noting that, $\forall \chi \in X$,

$$\frac{(F, \chi)}{|\chi|_2} \leq \sup_{\xi \in X} \frac{(F, \xi)}{|\xi|_2} = \|F\|_{-2}.$$

Multiplying both sides of the above inequality by $|\chi|_2$ gives the desired result:

$$(F, \chi) \leq \|F\|_{-2} |\chi|_2.$$

Next, we prove (5.6) by first applying Hölder inequality, (5.2) with $p = r = 2$. Therefore

$$(\nabla\psi, \nabla\chi) \leq \|\nabla\psi\| \|\nabla\chi\|.$$

Now, we apply the Poincaré inequality, (5.4) to get the desired result:

$$(\nabla\psi, \nabla\chi) \leq \Gamma_0 |\psi|_2 |\chi|_2,$$

where Γ_0 is the square of the Poincaré constant.

Next, we will obtain (5.9) by first applying the Hölder inequality to get

$$(\psi_x, \chi) \leq \|\psi_x\| \|\chi\|.$$

Noting that $\|\psi_x\| \leq \|\nabla\psi\|$ and applying the Poincaré inequality once to the term involving ψ and twice to the term involving χ gives the desired result

$$(\psi_x, \chi) \leq \Gamma_2 |\psi|_2 |\chi|_2,$$

where Γ_2 is the cube of the Poincaré constant.

Finally, for (5.8) we start by applying the Hölder inequality, (5.2)

$$b(\psi; \varphi, \chi) \leq \|\Delta\psi\|_{L^p} \|\nabla\varphi\|_{L^q} \|\nabla\chi\|_{L^r}, \quad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1, \quad (5.11)$$

where $1 \leq p, q, r \leq \infty$. Thus, setting $p = 2$, and $q = r = 4$ gives

$$b(\psi; \varphi, \chi) \leq \|\Delta\psi\| \|\nabla\varphi\|_{L^4} \|\nabla\chi\|_{L^4}.$$

Then applying the Ladyzhenskaya inequality, (5.5) results in

$$b(\psi; \varphi, \chi) \leq \Gamma \|\Delta\psi\| \|\nabla\varphi\|^{1/2} \|\Delta\varphi\|^{1/2} \|\nabla\chi\|^{1/2} \|\Delta\chi\|^{1/2}, \quad (5.12)$$

where Γ is the square Ladyzhenskaya constant, from (5.5). Using the Poincaré inequality, (5.4) gives

$$b(\psi; \varphi, \chi) \leq \Gamma_1 \|\Delta\psi\| \|\Delta\varphi\| \|\Delta\chi\|,$$

where Γ_1 is the square of the Poincaré constant time Γ from (5.12). \square

Theorem 3. *Let ψ be the solution of (3.9) and ψ^h be the solution of (4.2). Furthermore, assume that the following small data condition is satisfied:*

$$Re^{-2} Ro \geq \Gamma_1 \|F\|_{-2}, \quad (5.13)$$

where Re is the Reynolds number defined in (2.18), Ro is the Rossby number defined in (2.10), Γ_1 is the continuity constant of the trilinear form b in (5.8), and F is the forcing term. Then the following error estimate holds:

$$|\psi - \psi^h|_2 \leq C(Re, Ro, \Gamma_1, \Gamma_2, F) \inf_{\chi^h \in X^h} |\psi - \chi^h|_2, \quad (5.14)$$

where

$$C(Re, Ro, \Gamma_1, \Gamma_2, F) := \frac{Ro^{-1}\Gamma_2 + 2Re^{-1} + \Gamma_1 Re Ro^{-1} \|F\|_{-2}}{Re^{-1} - \Gamma_1 Re Ro^{-1} \|F\|_{-2}} \quad (5.15)$$

is a generic constant that can depend on Re , Ro , Γ_1 , Γ_2 , F , but not on the meshsize h .

Remark 3. *Note that the small data condition in Theorem 3 involves both the Reynolds number and the Rossby number, the latter quantifying the rotation effects in the QGE.*

Furthermore, note that the standard small data condition $Re^{-2} \geq \Gamma_1 \|F\|_{-2}$ used to prove the uniqueness for the steady-state 2D NSE [40, 41, 52] is significantly more restrictive for the QGE, since (5.13) has the Rossby number (which is small when rotation effects are significant) on the left-hand side. This is somewhat counterintuitive, since in general rotation effects are expected to help in proving the well-posedness of the system. We think that the explanation for this puzzling situation is the following: Rotation effects do make the mathematical analysis of 3D flows more amenable by giving them a 2D character. We, however, are concerned with 2D flows (the QGE). In this case, the small data condition (5.13) (needed in proving the uniqueness of the solution) indicates that rotation effects make the mathematical analysis of the (2D) QGE more complicated than that of the 2D NSE.

Proof. Since $X^h \subset X$, (3.9) holds for all $\chi = \chi^h \in X^h$. Subtracting (4.2) from (3.9) with $\chi = \chi^h \in X^h$ gives

$$\begin{aligned} Re^{-1} \left(\Delta [\psi - \psi^h], \Delta \chi^h \right) + b(\psi; \psi, \chi^h) - b(\psi^h; \psi^h, \chi^h) \\ - Ro^{-1} \left([\psi - \psi^h]_x, \chi^h \right) = 0 \quad \forall \chi^h \in X^h. \end{aligned} \quad (5.16)$$

Next, adding and subtracting $b(\psi^h; \psi, \chi^h)$ to (5.16), we get:

$$\begin{aligned} Re^{-1}(\Delta [\psi - \psi^h], \Delta \chi^h) + b(\psi; \psi, \chi^h) - b(\psi^h; \psi, \chi^h) + b(\psi^h; \psi, \chi^h) - b(\psi^h; \psi^h, \chi^h) \\ - Ro^{-1}([\psi - \psi^h]_x, \chi^h) = 0 \quad \forall \chi^h \in X^h. \end{aligned} \quad (5.17)$$

The error e can be decomposed as

$$e := \psi - \psi^h = (\psi - \lambda^h) + (\lambda^h - \psi^h) := \eta + \Phi^h, \quad (5.18)$$

where $\lambda^h \in X^h$ is arbitrary. Thus, equation (5.17) can be rewritten as

$$\begin{aligned} Re^{-1}(\Delta \eta + \Delta \Phi^h, \Delta \chi^h) + b(\eta + \Phi^h, \psi, \chi^h) + b(\psi^h; \eta + \Phi^h, \chi^h) \\ + Ro^{-1}(\eta_x + \Phi_x^h, \chi^h) = 0 \quad \forall \chi^h \in X^h. \end{aligned} \quad (5.19)$$

Letting $\chi^h := \Phi^h$ in (5.19), we obtain

$$\begin{aligned} Re^{-1}(\Delta \Phi^h, \Delta \Phi^h) - Ro^{-1}(\Phi_x^h, \Phi^h) = -Re^{-1}(\Delta \eta, \Delta \Phi^h) - b(\eta; \psi, \Phi^h) - b(\psi^h; \psi, \Phi^h) \\ - b(\psi^h; \eta, \Phi^h) - b(\psi^h; \Phi^h, \Phi^h) + Ro^{-1}(\eta, \Phi^h). \end{aligned} \quad (5.20)$$

Note that, since $(\Phi_x^h, \Phi^h) = -(\Phi^h, \Phi_x^h) \forall \Phi^h \in X^h \subset X = H_0^2$, it follows that

$$(\Phi_x^h, \Phi^h) = 0. \quad (5.21)$$

Also, it follows immediately from (3.7) that

$$b(\psi^h; \Phi^h, \Phi^h) = 0. \quad (5.22)$$

Combining (5.22), (5.21), and (5.20), we get:

$$\begin{aligned} Re^{-1}(\Delta \Phi^h, \Delta \Phi^h) = -Re^{-1}(\Delta \eta, \Delta \Phi^h) - b(\Phi^h; \psi, \Phi^h) - b(\eta; \psi, \Phi^h) \\ - b(\psi^h; \eta, \Phi^h) + Ro^{-1}(\eta_x, \Phi^h). \end{aligned} \quad (5.23)$$

Using

$$(\Delta \Phi^h, \Delta \Phi^h) = |\Phi^h|_2^2$$

and inequalities (5.7) – (5.9) in equation (5.23) gives

$$\begin{aligned} Re^{-1} |\Phi^h|_2^2 \leq Re^{-1} |\eta|_2 |\Phi^h|_2 + \Gamma_1 \left(|\eta|_2 |\psi|_2 |\Phi^h|_2 + |\psi^h|_2 |\eta|_2 |\Phi^h|_2 \right) \\ + \Gamma_1 |\Phi^h|_2^2 |\psi|_2 + Ro^{-1} \Gamma_2 |\eta|_2 |\Phi^h|_2. \end{aligned} \quad (5.24)$$

Simplifying and rearranging terms in (5.24) gives

$$|\Phi^h|_2 \leq \left(Re^{-1} - \Gamma_1 |\psi|_2 \right)^{-1} \left(Re^{-1} + \Gamma_1 |\psi|_2 + \Gamma_1 |\psi^h|_2 + Ro^{-1} \Gamma_2 \right) |\eta|_2. \quad (5.25)$$

Using (5.25) and the triangle inequality along with the stability estimates (3.10) and (4.3) gives

$$\begin{aligned}
|e|_2 &\leq |\eta|_2 + |\Phi^h|_2 \\
&\leq \left[1 + \frac{Re^{-1} + \Gamma_1 |\psi|_2 + \Gamma_1 |\psi^h|_2 + Ro^{-1} \Gamma_2}{Re^{-1} - \Gamma_1 |\psi|_2} \right] |\eta|_2 \\
&\leq \left[1 + \frac{Re^{-1} + \Gamma_1 (Re Ro^{-1} \|F\|_{-2}) + \Gamma_1 (Re Ro^{-1} \|F\|_{-2}) + Ro^{-1} \Gamma_2}{Re^{-1} - \Gamma_1 (Re Ro^{-1} \|F\|_{-2})} \right] |\eta|_2 \\
&= \left[\frac{Ro^{-1} \Gamma_2 + 2 Re^{-1} + \Gamma_1 Re Ro^{-1} \|F\|_{-2}}{Re^{-1} - \Gamma_1 Re Ro^{-1} \|F\|_{-2}} \right] |\psi - \lambda^h|_2,
\end{aligned} \tag{5.26}$$

where $\lambda^h \in X^h$ is arbitrary. Taking the infimum over $\lambda^h \in X^h$ in (5.26) proves the error estimate (5.14). \square

In Theorem 3, we proved an error estimate in the H^2 norm. In Theorem 4, we will prove error estimates in the L^2 and H^1 norms by using a duality argument. To this end, we first notice that the QGE (2.26) can be written as

$$\mathcal{N} \psi = Ro^{-1} F, \tag{5.27}$$

where the nonlinear operator \mathcal{N} is defined as

$$\mathcal{N} \psi := Re^{-1} \Delta^2 \psi + J(\psi, \Delta \psi) - Ro^{-1} \frac{\partial \psi}{\partial x}. \tag{5.28}$$

The linearization of \mathcal{N} around ψ , a solution of (2.26), yields the following *linear* operator:

$$\mathcal{L} \chi := Re^{-1} \Delta^2 \chi + J(\chi, \Delta \psi) + J(\psi, \Delta \chi) - Ro^{-1} \frac{\partial \chi}{\partial x}. \tag{5.29}$$

To find the dual problem associated with the QGE (5.27), we first define the *dual operator* \mathcal{L}^* of \mathcal{L} :

$$(\mathcal{L} \chi, \psi^*) = (\chi, \mathcal{L}^* \psi^*) \quad \forall \psi^* \in X. \tag{5.30}$$

To find \mathcal{L}^* , we use the standard procedure: In (5.30), we use the definition of \mathcal{L} given in (5.29) and we “integrate by parts” (i.e., use Green’s theorem):

$$\begin{aligned}
(\mathcal{L} \chi, \psi^*) &= \left(Re^{-1} \Delta^2 \chi + J(\chi, \Delta \psi) + J(\psi, \Delta \chi) - Ro^{-1} \frac{\partial \chi}{\partial x}, \psi^* \right) \\
&= \left(\chi, Re^{-1} \Delta^2 \psi^* - J(\psi, \Delta \psi^*) + Ro^{-1} \frac{\partial \psi^*}{\partial x} \right) + \left(J(\chi, \Delta \psi), \psi^* \right),
\end{aligned} \tag{5.31}$$

where to get the first term on the right-hand side of (5.31) we used the skew-symmetry of the trilinear form b in the last two variables and Green’s theorem (just as we did in the proof of Theorem 1). Next, we apply Green’s theorem to the second term on the right-hand side of (5.31):

$$\begin{aligned}
\left(J(\chi, \Delta \psi), \psi^* \right) &= \int_{\Omega} \chi_x \Delta \psi_y \psi^* - \chi_y \Delta \psi_x \psi^* d\mathbf{x} \\
&\stackrel{Green}{=} \int_{\partial \Omega} (\chi \Delta \psi_{xy} \psi^* - \chi \Delta \psi_{yx} \psi^*) \cdot \mathbf{n} dS + \int_{\Omega} \chi \Delta \psi_x \psi_y^* - \chi \Delta \psi_y \psi_x^* d\mathbf{x} \\
&= \left(\chi, J(\Delta \psi, \psi^*) \right).
\end{aligned} \tag{5.32}$$

Equations (5.31)-(5.32) imply:

$$\begin{aligned} (\mathcal{L} \chi, \psi^*) &= \left(\chi, Re^{-1} \Delta^2 \psi^* - J(\psi, \Delta \psi^*) + Ro^{-1} \frac{\partial \psi^*}{\partial x} \right) + \left(\chi, J(\Delta \psi, \psi^*) \right) \\ &= (\chi, \mathcal{L}^* \psi^*). \end{aligned} \quad (5.33)$$

Thus, the *dual operator* \mathcal{L}^* is given by

$$\mathcal{L}^* \psi^* = Re^{-1} \Delta^2 \psi^* - J(\psi, \Delta \psi^*) + J(\Delta \psi, \psi^*) + Ro^{-1} \frac{\partial \psi^*}{\partial x}. \quad (5.34)$$

For any given $g \in L^2(\Omega)$, the weak formulation of the *dual problem* is:

$$(\mathcal{L}^* \psi^*, \chi) = (g, \chi) \quad \forall \chi \in X = H_0^2(\Omega). \quad (5.35)$$

We assume that ψ^* , the solution of (5.35), satisfies the following elliptic regularity estimates:

$$\psi^* \in H^4(\Omega) \cap H_0^2(\Omega), \quad (5.36)$$

$$\|\psi^*\|_4 \leq C \|g\|_0, \quad (5.37)$$

$$\|\psi^*\|_3 \leq C \|g\|_{-1}, \quad (5.38)$$

where C is a generic constant that can depend on the data, but not on the mesh size h .

Remark 4. We note that this type of elliptic regularity was also assumed in [16] for the streamfunction formulation of the 2D NSE. In that report, it was also noted that, for a polygonal domain with maximum interior vertex angle $\theta < 126^\circ$, the assumed elliptic regularity was actually proved by Blum and Rannacher [10]. We note that the theory developed in [10] carries over to our case as well. In Section 5 in [10] it is proved that, for weakly nonlinear problems that involve the biharmonic operator as linear main part and that satisfy certain growth restrictions, each weak solution satisfies elliptic regularity results of the form (5.36)-(5.38). Assuming that Ω is a bounded polygonal domain with inner angle ω at each boundary corner satisfying $\omega < 126.283696\dots^\circ$, Theorem 7 in [10] with $k = 0$ and $k = 1$ implies (5.36)-(5.38). Using an argument similar to that used in Section 6(b) in [10] to prove that the streamfunction formulation of the 2D NSE satisfies the restrictions in Theorem 7, we can prove that ψ^* , the solution of our dual problem (5.35), satisfies the elliptic regularity results in (5.36)-(5.38). Indeed, the main point in Section 6(b) in [10] is that the corner singularities arising in flows around sharp corners are essentially determined by the linear main part Δ^2 in the streamfunction formulation of the 2D NSE, which is the linear main part of our dual problem (5.35) as well.

We emphasize that although, the error estimates in the L^2 and H^1 norms proven in Theorem 4 are derived for the particular space $X^h \subset H_0^2(\Omega)$ consisting of Argyris elements, the same results can be derived for other conforming C^1 finite element spaces.

Theorem 4. Let ψ be the solution of (3.9) and ψ^h be the solution of (4.2). Assume that the same small data condition as in Theorem 3 is satisfied:

$$Re^{-2} Ro \geq \Gamma_1 \|F\|_{-2}. \quad (5.39)$$

Furthermore, assume that $\psi \in H^6(\Omega) \cap H_0^2(\Omega)$. Then there exist positive constants C_0, C_1 and C_2 that can depend on $Re, Ro, \Gamma_1, \Gamma_2, F$, but not on the meshsize h , such that

$$|\psi - \psi^h|_2 \leq C_2 h^4 |\psi|_6, \quad (5.40)$$

$$|\psi - \psi^h|_1 \leq C_1 h^5 |\psi|_6, \quad (5.41)$$

$$\|\psi - \psi^h\|_0 \leq C_0 h^6 |\psi|_6. \quad (5.42)$$

Proof. Estimate (5.40) follows immediately from (4.17) and Theorem 3. Estimates (5.42) and (5.41) follow from a duality argument.

The error in the primal problem (3.9) and the interpolation error in the dual problem (5.35) are denoted as

$$e := \psi - \psi^h \quad e^* := \psi^* - \psi^{*h}, \quad (5.43)$$

respectively.

We start by proving the L^2 norm estimate (5.42).

$$\begin{aligned} |e|^2 &= (e, e) = (e, \mathcal{L}^* \psi^*) = (\mathcal{L} e, \psi^*) \\ &= (e, \mathcal{L}^* e^*) + (e, \mathcal{L}^* \psi^{*h}) = (\mathcal{L} e, e^*) + (\mathcal{L} e, \psi^{*h}). \end{aligned} \quad (5.44)$$

The last term on the right-hand side of (5.44) is given by

$$(\mathcal{L} e, \psi^{*h}) = \left(Re^{-1} \Delta^2 e + J(e, \Delta \psi) + J(\psi, \Delta e) - Ro^{-1} \frac{\partial e}{\partial x}, \psi^{*h} \right). \quad (5.45)$$

To estimate this term, we consider the error equation obtained by subtracting (4.2) (with $\psi^h = \psi^{*h}$) from (3.9) (with $\chi = \psi^{*h}$):

$$\left(Re^{-1} \Delta^2 e - Ro^{-1} \frac{\partial e}{\partial x}, \psi^{*h} \right) + \left(J(\psi, \Delta \psi) - J(\psi^h, \Delta \psi^h), \psi^{*h} \right) = 0. \quad (5.46)$$

Using (5.46), equation (5.45) can be written as follows:

$$(\mathcal{L} e, \psi^{*h}) = \left(J(e, \Delta \psi) + J(\psi, \Delta e) - J(\psi, \Delta \psi) + J(\psi^h, \Delta \psi^h), \psi^{*h} \right). \quad (5.47)$$

Thus, by using (5.47) equation (5.44) becomes:

$$\begin{aligned} |e|^2 &= (\mathcal{L} e, e^*) + (\mathcal{L} e, \psi^{*h}) \\ &= Re^{-1} (\Delta e, \Delta e^*) - Ro^{-1} \left(\frac{\partial e}{\partial x}, e^* \right) + (J(e, \Delta \psi) + J(\psi, \Delta e), e^*) \\ &\quad + \left(J(e, \Delta \psi) + J(\psi, \Delta e) - J(\psi, \Delta \psi) + J(\psi^h, \Delta \psi^h), \psi^{*h} \right) \\ &= Re^{-1} (\Delta e, \Delta e^*) - Ro^{-1} (e_x, e^*) + b(e, \psi, e^*) + b(\psi, e, e^*) + b(e, \psi, \psi^{*h}) \\ &\quad - b(\psi, \psi, e^*) + b(\psi^h, \psi^h, e^*) \\ &= Re^{-1} (\Delta, \Delta e^*) - Ro^{-1} (e_x, e^*) + b(e, \psi, e^*) + b(\psi, e, e^*) \\ &\quad - b(e, \psi, e^*) + b(e, \psi^h, e^*) + b(e, e, \psi^*). \end{aligned} \quad (5.48)$$

Using the bounds in (5.7)-(5.9), (5.48) yields:

$$\begin{aligned} |e|^2 &= Re^{-1} (\Delta e, \Delta e^*) - Ro^{-1} (e_x, e^*) + b(e, \psi, e^*) + b(\psi, e, e^*) \\ &\quad - b(e, \psi, e^*) + b(e, \psi^h, e^*) + b(e, e, \psi^*) \\ &\leq Re^{-1} |e|_2 |e^*|_2 + Ro^{-1} \Gamma_2 |e|_2 |e^*|_2 + \Gamma_1 |e|_2 |\psi|_2 |e^*|_2 + \Gamma_1 |\psi|_2 |e|_2 |e^*|_2 \\ &\quad + \Gamma_1 |e|_2 |\psi|_2 |e^*|_2 + \Gamma_1 |e|_2 |\psi^h|_2 |e^*|_2 + \Gamma_1 |e|_2 |e|_2 |\psi^*|_2 \\ &= |e|_2 |e^*|_2 \left(Re^{-1} + Ro^{-1} \Gamma_2 + \Gamma_1 |\psi|_2 + \Gamma_1 |\psi|_2 + \Gamma_1 |\psi|_2 + \Gamma_1 |\psi^h|_2 \right) \\ &\quad + |e|_2^2 (\Gamma_1 |\psi^*|_2). \end{aligned} \quad (5.49)$$

We start bounding the terms on the right-hand side of (5.49). First, we note that using the stability estimates (3.10) for ψ and (4.3) for ψ^h , the right-hand side of (5.49) can be bounded as follows:

$$|e|^2 \leq C |e|_2 |e^*|_2 + |e|_2^2 (\Gamma_1 |\psi^*|_2), \quad (5.50)$$

where C is a generic constant that can depend on Re , Ro , Γ_1 , Γ_2 , F , but *not* on the mesh size h . By using the approximation results (4.18), we get:

$$|e^*|_2 \leq C h^2 |\psi^*|_4. \quad (5.51)$$

By using (5.36) and (5.37), the elliptic regularity results of the dual problem (5.35) with $g := e$, we also get:

$$|\psi^*|_4 \leq C |e|, \quad (5.52)$$

which obviously implies

$$|\psi^*|_2 \leq C |e|. \quad (5.53)$$

Inequalities (5.51)-(5.52) imply:

$$|e^*|_2 \leq C h^2 |e|. \quad (5.54)$$

Inserting (5.53) and (5.54) in (5.50), we get:

$$|e|^2 \leq C h^2 |e|_2 |e| + C |e|_2^2 |e|. \quad (5.55)$$

Using the obvious simplifications and the H^2 error estimate (5.40) in (5.55) yields:

$$|e| \leq C h^2 |e|_2 + C |e|_2^2 \leq C h^6 |\psi|_6 + C h^8 |\psi|_6^2 = C_0 h^6 |\psi|_6, \quad (5.56)$$

which proves the L^2 error estimate (5.42).

Next, we prove the H^1 norm estimate (5.41). Since the duality approach we use is similar to that we used in proving the L^2 norm estimate (5.42), we only highlight the main differences. We start again by writing the H^1 norm of the error in terms of the dual operator \mathcal{L}^* :

$$\begin{aligned} |e|_1^2 &= (\nabla e, \nabla e) = (e, -\Delta e) = (e, \mathcal{L}^* \psi^*) = (\mathcal{L} e, \psi^*) \\ &= (e, \mathcal{L}^* e^*) + (e, \mathcal{L}^* \psi^{*h}) = (\mathcal{L} e, e^*) + (\mathcal{L} e, \psi^{*h}). \end{aligned} \quad (5.57)$$

Thus, the second and fourth equalities in (5.57) clearly indicate that in the dual problem (5.35), one should choose $g = -\Delta e$, and not $g = e$, as we did in (5.52), when we proved the L^2 error estimate (5.42). Using (5.36) and (5.38), the elliptic regularity results of the dual problem (5.35) with $g := -\Delta e$, we also get:

$$|\psi^*|_3 \leq C |-\Delta e|_{-1} \leq C |e|_1, \quad (5.58)$$

where in the last inequality we used the fact that $e \in H_0^2(\Omega)$. Inequality (5.58) obviously implies

$$|\psi^*|_2 \leq C |e|_1. \quad (5.59)$$

All the results in (5.45)-(5.49) carry over to our setting. Thus, we get:

$$|e|_1^2 \leq C |e|_2 |e^*|_2 + C |e|_2^2 |\psi^*|_2, \quad (5.60)$$

where C is a generic constant that can depend on Re , Ro , Γ_1 , Γ_2 , F , but *not* on the mesh size h . By using the approximation result (4.19), we get:

$$|e^*|_2 \leq C h |\psi^*|_3. \quad (5.61)$$

Inequalities (5.61)-(5.58) imply:

$$|e^*|_2 \leq C h |e|_1. \quad (5.62)$$

Inserting (5.59) and (5.62) in (5.60), we get:

$$|e|_1^2 \leq C h |e|_2 |e|_1 + C |e|_2^2 |e|_1. \quad (5.63)$$

Using the obvious simplifications and the H^2 error estimate (5.40) in (5.63) yields:

$$|e|_1 \leq C h |e|_2 + C |e|_2^2 \leq C h^5 |\psi|_6 + C h^8 |\psi|_6 = C_1 h^5 |\psi|_6, \quad (5.64)$$

which proves the H^1 error estimate (5.41). \square

5.2 Two-Level Method

The main goal of this section is to develop a rigorous numerical analysis for Algorithm 1, the two-level algorithm introduced in Section 4.4. The proof for the error bounds follows a pattern that is similar to that presented in [30].

To this end, we first introduce an improved bound on the trilinear form $b(\zeta; \xi, \chi)$ using the discrete Sobolev inequality [20, 30]:

$$\|\nabla \varphi^h\|_{L^\infty} \leq c \sqrt{|\ln(h)|} |\varphi^h|_2.$$

The following lemma follows from the above inequality and (5.8) and was proven in [30]:

Lemma 2. *For any $\chi^h \in X^h$, the following inequalities hold:*

$$\begin{aligned} |b(\psi; \chi^h, \xi)| &\leq C \sqrt{|\ln(h)|} |\psi|_2 |\xi|_1 |\chi^h|_2, \\ |b(\psi; \xi, \chi^h)| &\leq C \sqrt{|\ln(h)|} |\psi|_2 |\xi|_1 |\chi^h|_2. \end{aligned}$$

The following lemma will prove useful for proving the error bounds for Algorithm 1:

Lemma 3. *For $\psi, \xi, \chi \in H_0^2(\Omega)$, we have*

$$b(\psi; \xi, \chi) = b^*(\xi; \chi, \psi) - b^*(\chi; \xi, \psi), \quad (5.65)$$

where

$$b^*(\xi; \chi, \psi) = \int_{\Omega} (\chi_y \xi_{xy} - \xi_x \chi_{yy}) \psi_y - (\xi_y \chi_{yx} - \xi_y \chi_{xx}) \psi_x \, d\mathbf{x}. \quad (5.66)$$

For a proof see the corresponding Lemma 5.6 in [30].

Before proving any error bounds we first prove that the continuous problem linearized about ψ^H has a unique solution.

Lemma 4. *Given a solution ψ^H of (4.27), then the solution to the following problem exists uniquely:*

$$\begin{aligned} \text{Find } \hat{\psi} \in H_0^2(\Omega) \text{ such that, for all } \chi \in H_0^2(\Omega), \\ Re^{-1}(\Delta \hat{\psi}, \Delta \chi) + b(\psi^H; \hat{\psi}, \chi) - Ro^{-1}(\hat{\psi}_x, \chi) = Ro^{-1}(F, \chi), \end{aligned} \quad (5.67)$$

and satisfies $\|\hat{\psi}\|_2 \leq Re Ro^{-1} \|F\|_{-2}$.

Proof. First we introduce a new continuous bilinear form $B : H_0^2(\Omega) \times H_0^2(\Omega) \rightarrow \mathbb{R}$ given by

$$B(\psi, \chi) = Re^{-1}(\Delta\psi, \Delta\chi) + b(\psi^H; \psi, \chi) - Ro^{-1}(\psi_x, \chi).$$

B is continuous and coercive and therefore $\hat{\psi}$ exists and is unique. Now setting $\chi = \hat{\psi}$ in (5.67) and noting that $(\psi_x, \chi) = -(\chi_x, \psi)$ which implies that $(\hat{\psi}_x, \hat{\psi}) = 0$ gives

$$\begin{aligned} Re^{-1}\|\hat{\psi}\|_2^2 &= Ro^{-1}(F, \hat{\psi}) \\ \|\hat{\psi}\|_2 &= Re Ro^{-1} \frac{(F, \hat{\psi})}{\|\hat{\psi}\|_2}. \end{aligned}$$

Taking the supremum over all $\hat{\psi} \in X$ leads to

$$\|\hat{\psi}\|_2 \leq Re Ro^{-1} \sup_{\hat{\psi} \in X} \frac{(F, \hat{\psi})}{\|\hat{\psi}\|_2}.$$

Therefore, by definition it follows that $\|\hat{\psi}\|_2 \leq Re Ro^{-1}\|F\|_{-2}$. \square

The following lemma shows that the solution to the fine mesh problem, (4.28) exists and has a stability bound dependent on Re , Ro , and the forcing function F .

Lemma 5. *The solution to (4.28) exists and satisfies*

$$\|\psi^h\|_2 \leq Re Ro^{-1}\|F\|_{-2}.$$

Proof. The bilinear form B is continuous and coercive on X^h and so ψ^h exists and is unique. Setting $\chi^h = \psi^h$ in (4.28) and again noting that $(\psi_x^h, \psi^h) = 0$ and using (5.10) gives

$$\begin{aligned} Re^{-1}\|\psi^h\|_2^2 &= Ro^{-1}(F, \psi^h) \\ \|\psi^h\|_2 &= Re Ro^{-1} \frac{(F, \psi^h)}{\|\psi^h\|_2}, \end{aligned}$$

which implies

$$\|\psi^h\| \leq Re Ro^{-1}\|F\|_{-2}.$$

\square

The following theorem gives the error bound after Step 2 and is the main result of this section. The proof of this theorem is similar to the proof for a similar theorem in [30].

Theorem 5. *Let $X^h, X^H \subset H_0^2(\Omega)$ be two finite element spaces. Let ψ be the solution to (3.9) and ψ^h the solution to (4.28). Then ψ^h satisfies*

$$|\psi - \psi^h|_2 \leq C_1 \inf_{\lambda^h \in X^h} |\psi - \lambda^h|_2 + C_2 \sqrt{|\ln h|} |\psi - \psi^H|_1, \quad (5.68)$$

where $C_1 = 2 + Re Ro^{-1} + Re^2 Ro^{-1} \Gamma_1 \|F\|_{-2}$ and $C_2 = 2Re^2 Ro^{-1} \Gamma_1 C \|F\|_{-2}$.

Proof. Subtracting (4.28) from (3.9) and letting $\chi = \chi^h$ yields

$$Re^{-1}(\Delta [\psi - \psi^h], \Delta \chi^h) + b(\psi; \psi, \chi^h) - b(\psi^H; \psi^h, \chi^h) - Ro^{-1}([\psi - \psi^h]_x, \chi^h) = 0, \quad \forall \chi^h \in X^h.$$

Using Lemma 3 gives

$$\begin{aligned} Re^{-1}(\Delta [\psi - \psi^h], \Delta \chi^h) + b^*(\psi; \chi^h, \psi) - b^*(\chi^h; \psi, \psi) \\ - b^*(\psi^h; \chi^h, \psi^H) + b^*(\chi^h; \psi^h, \psi^H) - Ro^{-1}([\psi - \psi^h]_x, \chi^h) = 0, \quad \forall \chi^h \in X^h. \end{aligned}$$

Now, adding the terms

$$-b^*(\psi^h; \chi^h, \psi) + b^*(\chi^h; \psi^h, \psi) + b^*(\psi^h; \chi^h, \psi) - b^*(\chi^h; \psi^h, \psi)$$

gives

$$\begin{aligned} Re^{-1}(\Delta [\psi - \psi^h], \Delta \chi^h) + b^*(\psi - \psi^h; \chi^h, \psi) + b^*(\chi^h; \psi^h - \psi, \psi) \\ + b^*(\psi^h; \chi^h, \psi - \psi^H) + b^*(\chi^h; \psi^h, \psi^H - \psi) - Ro^{-1}([\psi - \psi^h]_x, \chi^h) = 0, \quad \forall \chi^h \in X^h. \end{aligned}$$

Take $\lambda^h \in H_0^2(\Omega)$ arbitrary and define $e := \psi - \psi^h = \eta - \Phi^h$, where $\Phi^h = \psi^h - \lambda^h$ and $\eta = \psi - \lambda^h$. We have

$$\begin{aligned} Re^{-1}(\Delta \Phi^h, \Delta \chi^h) + b^*(\Phi^h; \chi^h, \psi) - b^*(\chi^h; \Phi^h, \psi) - Ro^{-1}(\Phi_x^h, \chi^h) \\ = Re^{-1}(\eta, \chi^h) + b^*(\eta; \chi^h, \psi) - b^*(\chi^h; \eta, \psi) \\ + b^*(\psi^h; \chi^h, \psi - \psi^H) + b^*(\chi^h; \psi^h, \psi^H - \psi) - Ro^{-1}(\eta_x, \chi^h), \quad \forall \chi^h \in X^h. \end{aligned}$$

Since this holds for any $\chi^h \in H_0^2(\Omega)$, it holds in particular for $\chi^h = \Phi^h \in H_0^2(\Omega)$, which implies

$$\begin{aligned} Re^{-1}(\Delta \Phi^h, \Delta \Phi^h) - Ro^{-1}(\Phi_x^h, \Phi^h) = Re^{-1}(\Delta \eta, \Delta \Phi^h) + b^*(\eta; \Phi^h, \psi) - b^*(\Phi^h; \eta, \psi) \\ + b^*(\psi^h; \Phi^h, \psi - \psi^H) + b^*(\Phi^h; \psi^h, \psi^H - \psi) - Ro^{-1}(\eta_x, \Phi^h). \end{aligned}$$

Note that $(\Phi_x, \Phi) = -(\Phi, \Phi_x)$ and so it follows that $(\Phi_x^h, \Phi^h) = 0$. This combined with Lemma 3 implies

$$\begin{aligned} Re^{-1}(\Delta \Phi^h, \Delta \Phi^h) = Re^{-1}(\Delta \eta, \Delta \Phi^h) + b(\psi; \eta, \Phi^h) \\ + b^*(\psi^h; \Phi^h, \psi - \psi^H) + b^*(\Phi^h; \psi^h, \psi^H - \psi) - Ro^{-1}(\eta_x, \Phi^h). \end{aligned}$$

Using the error bounds given in (5.7), (5.8), (5.9), (5.10), Lemma 2, Theorem 1, Theorem 2 and Lemma 4 gives

$$\begin{aligned} Re^{-1}|\Phi^h|_2^2 &\leq Re^{-1}|\eta|_2 |\Phi^h|_2 + \Gamma_1 |\psi|_2 |\eta|_2 |\Phi^h|_2 \\ &\quad + 2\Gamma_1 C |\psi^H|_2 |\Phi^h|_2 |\psi - \psi^H|_1 \sqrt{|\ln(h)|} + Ro^{-1} \Gamma_2 |\eta|_2 |\Phi^h|_2 \\ &= \left(Ro^{-1} \Gamma_2 + Re^{-1} + \Gamma_1 |\psi|_2 \right) |\eta|_2 |\Phi^h|_2 + 2\Gamma_1 C |\psi^H|_2 |\Phi^h|_2 |\psi - \psi^H|_1 \sqrt{|\ln(h)|} \\ |\Phi^h|_2 &\leq \left(1 + Re Ro^{-1} \Gamma_2 + Re^2 Ro^{-1} \Gamma_1 \|F\|_{-2} \right) |\eta|_2 \\ &\quad + 2Re^2 Ro^{-1} \Gamma_1 C \|F\|_{-2} |\psi - \psi^H|_1 \sqrt{|\ln(h)|} \end{aligned}$$

Adding $|\eta|_2$ to both sides and using the triangle inequality ($|\psi - \psi^h|_2 \leq |\Phi^h|_2 + |\eta|_2$) gives

$$\begin{aligned} |\Phi^h|_2 &\leq \left(2 + Re Ro^{-1} \Gamma_2 + Re^2 Ro^{-1} \Gamma_1 \|F\|_{-2}\right) |\eta|_2 \\ &\quad + 2Re^2 Ro^{-1} \Gamma_1 C \|F\|_{-2} |\psi - \psi^H|_1 \sqrt{|\ln(h)|} \end{aligned}$$

Thus, we have the following estimate for the error bounds:

$$|\psi - \psi^h|_2 \leq C_1 \inf_{\lambda^h \in X^h} |\psi - \lambda^h|_2 + C_2 \sqrt{|\ln h|} |\psi - \psi^H|_1$$

where $C_1 = 2 + Re Ro^{-1} \Gamma_2 + Re^2 Ro^{-1} \Gamma_1 \|F\|_{-2}$ and $C_2 = 2Re^2 Ro^{-1} \Gamma_1 C \|F\|_{-2}$. \square

As an example, consider the case of the Argyris triangle. For this element we have the following inequalities, which follow from approximation theory [9] and Theorem 6.1.1 [20]:

$$|\psi - \psi^h|_j \leq C h^{6-j} |\psi|_6 \quad (5.69)$$

$$|\psi - \psi^H|_j \leq C H^{6-j} |\psi|_6 \quad (5.70)$$

Corollary 5.6. *Let $X^h, X^H \in H_0^2(\Omega)$ be the Argyris finite elements. Then, ψ^h satisfies*

$$|\psi - \psi^h|_2 \leq C_1 h^4 |\psi|_6 + C_2 \sqrt{|\ln(h)|} H^5 |\psi|_6. \quad (5.71)$$

Proof. This follows directly by substituting (5.69) and (5.70) into (5.68). \square

5.3 QGE

In this section we present the convergence, and the error analysis associated with the semi-discretization. We will rely heavily on the Section 5.1 which contains the finite error analysis for the SQGE.

Proposition 1. *The solution to (4.1), ψ^h , is stable. For any $t > 0$ the following inequality holds:*

$$\frac{1}{2} \|\nabla \psi^h(t)\|^2 + \frac{Re^{-1}}{2} \int_0^t \|\Delta \psi^h(t')\|^2 dt' \leq \frac{1}{2} \|\nabla \psi_0^h\|^2 + \frac{Re Ro^{-2}}{2} \int_0^t \|F(t')\|_{-1}^2 dt'. \quad (5.72)$$

Proof. Take $\chi^h = \psi^h$ in (4.1) and note that $b(\psi^h; \psi^h, \psi^h) = 0$. Since $(\psi_x^h, \psi^h) = -(\psi^h, \psi_x^h)$, we get $(\psi_x^h, \psi^h) = 0$. Thus, we have

$$\frac{1}{2} \frac{d}{dt} \|\nabla \psi^h\|^2 + Re^{-1} \|\Delta \psi^h\|^2 = Ro^{-1} (F, \psi^h) \leq Ro^{-1} \|F\|_{-1} \|\nabla \chi\|.$$

Using Young inequality gives

$$\frac{1}{2} \frac{d}{dt} \|\nabla \psi^h\|^2 + Re^{-1} \|\Delta \psi^h\|^2 = \frac{Ro^{-2}}{2\epsilon} \|F\|_{-1}^2 + \frac{\epsilon}{2} \|\nabla \psi^h\|^2. \quad (5.73)$$

Now taking $\epsilon = Re^{-1}$ in (5.73) results in

$$\frac{1}{2} \frac{d}{dt} \|\nabla \psi^h\|^2 + \frac{Re^{-1}}{2} \|\Delta \psi^h\|^2 = \frac{Re Ro^{-2}}{2} \|F\|_{-1}^2.$$

Assuming $\|\Delta \psi^h\| \in L^1(0, T)$ and integrating over $(0, t)$ gives the final result. \square

Lemma 6. *There are finite constants $\Gamma_3, \Gamma_4 > 0$ such that, for all $\psi, \varphi, \chi \in X$, the following inequalities hold:*

$$b(\psi; \varphi, \chi) \leq \Gamma_3 \|\Delta\psi\| \|\Delta\varphi\| \left(\|\nabla\chi\|^{1/2} \|\Delta\chi\|^{1/2} \right) \quad (5.74)$$

$$b(\psi; \varphi, \chi) \leq \Gamma_4 \left(\|\nabla\psi\|^{1/2} \|\Delta\psi\|^{1/2} \right) \|\Delta\varphi\| \|\Delta\chi\|. \quad (5.75)$$

Proof. We will begin with the proof of (5.74) and then prove (5.75) last. For (5.74) we begin in the same way as we did for estimate (5.8) in Lemma 1. That is, we apply the Hölder inequality, (5.2) to $b(\psi; \varphi, \chi)$ and set $p = 2$, while setting $q = r = 4$, giving

$$b(\psi; \varphi, \chi) \leq \|\Delta\psi\| \|\nabla\varphi\|_{L^4} \|\nabla\chi\|_{L^4}.$$

Now, applying the Ladyzhenskaya inequality (5.5), we get

$$b(\psi; \varphi, \chi) \leq \Gamma \|\Delta\psi\| \|\nabla\varphi\|^{1/2} \|\Delta\varphi\|^{1/2} \|\nabla\chi\|^{1/2} \|\Delta\chi\|^{1/2}.$$

Next using that the semi-norm, $|\cdot|_2$, and the norm, $\|\cdot\|_2$, are equivalent in $H_0^2(\Omega)$ (see (1.2.8) in Ciarlet [20]) we have

$$b(\psi; \varphi, \chi) \leq \Gamma_3 \|\Delta\psi\| \|\Delta\varphi\| \left(\|\nabla\chi\|^{1/2} \|\Delta\chi\|^{1/2} \right),$$

which proves estimate (5.74).

For estimate (5.75) we apply Lemma 3, thus using (5.66), Hölder inequality, (5.2) reads

$$b(\psi; \varphi, \chi) \leq \|\nabla\psi\|_{L^p} \|\Delta\varphi\|_{L^q} \|\nabla\chi\|_{L^r}, \quad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1. \quad (5.76)$$

Letting $p = r = 4$, and $q = 2$ in (5.76) gives

$$b(\psi; \varphi, \chi) \leq \|\nabla\psi\|_{L^4} \|\Delta\varphi\| \|\nabla\chi\|_{L^4}.$$

Then applying the Ladyzhenskaya inequality, (5.5) we get

$$b(\psi; \varphi, \chi) \leq \Gamma \|\Delta\varphi\| \|\nabla\psi\|^{1/2} \|\Delta\psi\|^{1/2} \|\nabla\chi\|^{1/2} \|\Delta\chi\|^{1/2}.$$

Finally, using that the semi-norm, $|\cdot|_2$, and the norm, $\|\cdot\|_2$, are equivalent in $H_0^2(\Omega)$ gives the desired result

$$b(\psi; \varphi, \chi) \leq \Gamma_4 \left(\|\nabla\psi\|^{1/2} \|\Delta\psi\|^{1/2} \right) \|\Delta\varphi\| \|\Delta\chi\|.$$

□

The next theorem proves the convergence of the FE approximation ψ^h to the exact solution, ψ , to QGE (2.26). The proof is similar to the proof for Theorem 22 in [52].

Theorem 6. *Let ψ be a unique solution to the QGE (2.26). Then there exists constants $C_1(T, Re, \Gamma_3)$, $C_2(T, Re, \Gamma_3)$, $C_3(Re, Ro, \Gamma_2)$, and $C_4(T, F, \psi_0, \Gamma_1, \Gamma_2, Re, Ro, \|\Delta\psi\|_{L^4(0,T;L^2)})$ such that for all $t \in [0, T]$*

$$\begin{aligned} & \|\nabla(\psi - \psi^h)\|^2 + Re^{-1} \int_0^T \|\Delta(\psi - \psi^h)\|^2 dt \leq C_1(T, Re, \Gamma_3) \|\nabla(\psi_0 - \psi^h(0))\|^2 \\ & + \inf_{\lambda^h(t) \in X^h} \left\{ C_2(T, Re, \Gamma_3) \int_0^T \|\nabla(\psi - \lambda^h)_t\|_{-1}^2 + C_3(Re, Ro, \Gamma_2) \|\Delta(\psi - \lambda^h)\|^2 dt \right. \\ & + C_4(T, Re, Ro, F, \Gamma_1, \Gamma_2, \Gamma_4, \|\Delta\psi\|_{L^4(0,T;L^2)}) \|\Delta(\psi - \lambda^h)\|_{L^4(0,T;L^2)} \\ & \left. + 2 \|\nabla(\psi - \lambda^h)\|^2 \right\}, \end{aligned} \quad (5.77)$$

where

$$C_1(T, Re, \Gamma_3) := 4 \exp \left(\frac{Re^{-3} \Gamma_3^4}{432} \int_0^T \|\Delta \psi\|^4 dt \right), \quad (5.78)$$

$$C_2(T, Re, \Gamma_3) := \frac{9}{4} Re C_1(T, Re, \Gamma_3), \quad (5.79)$$

$$C_3(Re, Ro, \Gamma_2) := 2 + 2 \left(Re^{-2} + Ro^{-2} \Gamma_2^2 \right), \quad (5.80)$$

$$C_4(T, F, \psi_0, Re, Ro, \Gamma_1, \Gamma_3, \Gamma_4, \int_0^T \|\Delta \psi\|^4 dt) := \frac{C_2(T, Re, \Gamma_3)}{2} \quad (5.81)$$

$$\left(\Gamma_1^2 \|\Delta \psi\|_{L^4(0,T;L^2)}^4 + \Gamma_4 C^*(F, \psi_0, Re, Ro) \right) \quad (5.82)$$

and $C^*(F, \psi_0, Re, Ro)$, Γ_1 , Γ_2 , Γ_3 , Γ_4 are the stability bound from Proposition 1, and the constants from Lemma 1 and Lemma 6, respectively.

Proof. Let $\chi = \chi^h \in X^h$ and subtract (4.1) from (3.8) then let $e := \psi - \psi^h$ and we obtain

$$(\nabla e_t, \nabla \chi^h) + [b(\psi; \psi, \chi^h) - b(\psi^h; \psi^h, \chi^h)] + Re^{-1}(\Delta e, \Delta \chi^h) - Ro^{-1}(e_x, \chi^h) = 0 \quad \forall \chi^h \in X^h \subset X.$$

Now adding and subtracting $b(\psi^h; \psi, \chi^h)$ gives

$$(\nabla e_t, \nabla \chi^h) + [b(e; \psi, \chi^h) + b(\psi^h; e, \chi^h)] + Re^{-1}(\Delta e, \Delta \chi^h) - Ro^{-1}(e_x, \chi^h) = 0 \quad \forall \chi^h \in X^h \subset X.$$

Take $\lambda^h : [0, T] \rightarrow X^h$ arbitrary and decomposing the error $e = \eta - \Phi^h$, where $\eta := \psi - \lambda^h$ and $\Phi^h := \psi^h - \lambda^h$, results in

$$\begin{aligned} (\nabla \Phi_t^h, \nabla \chi^h) + Re^{-1}(\Delta \Phi^h, \Delta \chi^h) &= (\nabla \eta_t, \nabla \chi^h) + Re^{-1}(\Delta \eta, \Delta \chi^h) - Ro^{-1}[(\eta_x, \chi^h) - (\Phi_x^h, \chi^h)] \\ &\quad + [b(\eta; \psi, \chi^h) - b(\Phi^h; \psi, \chi^h) + b(\psi^h; \eta, \chi^h) - b(\psi^h; \Phi^h, \chi^h)]. \end{aligned}$$

Let $\chi^h = \Phi^h$; then noting $b(\psi^h; \Phi^h, \Phi^h) = 0$ and $(\Phi_x^h, \Phi^h) = -(\Phi^h, \Phi_x^h)$ implies $(\Phi_x^h, \Phi^h) = 0$, we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla \Phi^h\|^2 + Re^{-1} \|\Delta \Phi^h\|^2 &= (\nabla \eta_t, \nabla \Phi^h) + Re^{-1}(\Delta \eta, \Delta \Phi^h) - Ro^{-1}(\eta_x, \Phi^h) \\ &\quad + [b(\eta; \psi, \Phi^h) - b(\Phi^h; \psi, \Phi^h) + b(\psi^h; \eta, \Phi^h)]. \end{aligned}$$

Using the Hölder inequality, (5.2), the bounds for (ψ_x, χ) given in Lemma 1, and the bounds for b from estimate (5.8) in Lemma 1, and Lemma 6 we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla \Phi^h\|^2 + Re^{-1} \|\Delta \Phi^h\|^2 &\leq \|\nabla \eta_t\|_{-1} \|\Delta \Phi^h\| + Re^{-1} \|\Delta \eta\| \|\Delta \Phi^h\| + Ro^{-1} \Gamma_2 \|\Delta \eta\| \|\Delta \Phi^h\| \\ &\quad + [b(\eta; \psi, \Phi^h) - b(\Phi^h; \psi, \Phi^h) + b(\psi^h; \eta, \Phi^h)]. \end{aligned} \quad (5.83)$$

Then using the Young inequality, (5.1) we have, for all $\epsilon > 0$

$$\|\nabla \eta_t\|_{-1} \|\Delta \Phi^h\| \leq \frac{\epsilon}{2} \|\Delta \Phi^h\|^2 + \frac{1}{2\epsilon} \|\eta_t\|_{-1}^2 \quad (5.84)$$

$$Re^{-1} \|\Delta \eta\| \|\Delta \Phi^h\| \leq \frac{\epsilon}{2} \|\Delta \Phi^h\|^2 + \frac{Re^{-2}}{2\epsilon} \|\Delta \eta\|^2 \quad (5.85)$$

$$Ro^{-1} \Gamma_2 \|\Delta \eta\| \|\Delta \Phi^h\| \leq \frac{\epsilon}{2} \|\Delta \Phi^h\|^2 + \frac{Ro^{-2} \Gamma_2^2}{2\epsilon} \|\Delta \eta\|^2. \quad (5.86)$$

Using the Young inequality, (5.1) with $\varepsilon > 0$ and estimate (5.8) in Lemma 1

$$\begin{aligned} b(\eta; \psi, \Phi^h) &\leq \Gamma_1 \|\Delta\eta\| \|\Delta\psi\| \|\Delta\Phi^h\| \\ &\leq \frac{\varepsilon}{2} \|\Delta\Phi^h\|^2 + \frac{\Gamma_1^2}{2\varepsilon} \|\Delta\eta\|^2 \|\Delta\psi\|^2, \end{aligned}$$

and letting $\varepsilon = 2\epsilon$ we have

$$b(\eta; \psi, \Phi^h) \leq \epsilon \|\Delta\Phi^h\|^2 + \frac{\Gamma_1^2}{4\epsilon} \|\Delta\eta\|^2 \|\Delta\psi\|^2. \quad (5.87)$$

Combing (5.84) - (5.87) and (5.83) we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla\Phi^h\|^2 + \frac{1}{2} (2Re^{-1} - 5\epsilon) \|\Delta\Phi^h\|^2 &\leq \frac{1}{2\epsilon} \left[\|\nabla\eta_t\|_{-1}^2 + (Re^{-2} + Ro^{-2}\Gamma_2^2) \|\Delta\eta\|^2 \right] \\ &\quad + \frac{\Gamma_1^2}{4\epsilon} \|\Delta\eta\|^2 \|\Delta\psi\|^2 - [b(\Phi^h; \psi, \Phi^h) - b(\psi^h; \eta, \Phi^h)]. \end{aligned} \quad (5.88)$$

For the term $b(\Phi^h; \psi, \Phi^h)$ we use Lemma 6 and the following version of the Young inequality, (5.1) (from [52] Equation (1.1.4)): given $a, b > 0$, for any $\epsilon > 0$, $1 \leq p \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, and

$$C(\epsilon, p, q) = \frac{(p\epsilon)^{-q/p}}{q} \quad ab \leq \epsilon a^p + C(\epsilon, p, q) b^q$$

with $p = 4/3$, $q = 4$ to obtain

$$\begin{aligned} b(\Phi^h; \psi, \Phi^h) &\leq \Gamma_3 \|\Delta\Phi^h\|^{3/2} (\|\Delta\psi\| \|\nabla\Phi^h\|^{1/2}) \\ &\leq \epsilon \|\Delta\Phi^h\|^2 + C_1^*(\Gamma_3, \epsilon) \|\Delta\psi\|^4 \|\nabla\Phi^h\|^2, \end{aligned} \quad (5.89)$$

where $C_1^*(\Gamma_3, \epsilon) = 27/256 \Gamma_3^4 \epsilon^{-3}$. Combining (5.88) and (5.89) gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla\Phi^h\|^2 + \frac{1}{2} (2Re^{-1} - 7\epsilon) \|\Delta\Phi^h\|^2 &\leq \frac{1}{2\epsilon} \left[\|\nabla\eta_t\|_{-1}^2 + (Re^{-2} + Ro^{-2}\Gamma_2^2) \|\Delta\eta\|^2 \right] \\ &\quad + \frac{\Gamma_1^2}{4\epsilon} \|\Delta\eta\|^2 \|\Delta\psi\|^2 + C_1^*(\Gamma_3, \epsilon) \|\Delta\psi\|^4 \|\nabla\Phi^h\|^2 + b(\psi^h; \eta, \Phi^h). \end{aligned} \quad (5.90)$$

For the final term involving $b(\psi^h; \eta, \Phi^h)$ we use inequality (5.75) and the Young inequality, (5.1) with $\varepsilon = 2\epsilon$, i.e.

$$\begin{aligned} b(\psi^h; \eta, \Phi^h) &\leq \Gamma_4 (\|\nabla\psi^h\|^{1/2} \|\Delta\psi^h\|^{1/2}) \|\Delta\eta\| \|\Delta\Phi^h\| \\ &\leq \epsilon \|\Delta\Phi^h\|^2 + \frac{\Gamma_4^2}{4\epsilon} \|\nabla\psi^h\| \|\Delta\eta\|^2. \end{aligned} \quad (5.91)$$

By the stability estimate in Proposition 1, we have

$$\|\nabla\psi^h\| \leq C_2^*(F, \psi_0, Re, Ro). \quad (5.92)$$

Using (5.92), estimate (5.91) becomes

$$b(\psi^h; \eta, \Phi^h) \leq \epsilon \|\Delta\Phi^h\|^2 + \frac{\Gamma_4^2}{4\epsilon} C_2^{*2}(F, \psi_0, Re, Ro) \|\Delta\psi^h\| \|\Delta\eta\|^2. \quad (5.93)$$

Combining (5.90) and (5.93) gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla \Phi^h\|^2 + \frac{1}{2} (2Re^{-1} - 9\epsilon) \|\Delta \Phi^h\|^2 &\leq \frac{1}{2\epsilon} \left[\|\nabla \eta_t\|_{-1}^2 + (Re^{-2} + Ro^{-2} \Gamma_2^2) \|\Delta \eta\|^2 \right] \\ &+ \frac{\Gamma_1^2}{4\epsilon} \|\Delta \psi\|^2 \|\Delta \eta\|^2 + \frac{\Gamma_4}{4\epsilon} C_2^*(F, \psi_0, Re, Ro) \|\Delta \psi^h\| \|\Delta \eta\|^2 + C_1^*(\Gamma_3, \epsilon) \|\Delta \psi\|^4 \|\nabla \Phi^h\|^2. \end{aligned} \quad (5.94)$$

Take $\epsilon = Re^{-1}/9$ in (5.94). Letting $C_3^*(F, \psi_0, Re, Ro, \Gamma_4) = \frac{\Gamma_4}{2} C_2^*(F, \psi_0, Re, Ro)$, $C_4^*(Re) = \frac{9}{2} Re$, $C_5^*(Re, \Gamma_3) = \frac{27}{256} 9^3 Re^3$, $C_6^*(Re, Ro, \Gamma_2) = Re^{-2} + Ro^{-2} \Gamma_2^2$, and $C_7^*(\Gamma_1) = \frac{\Gamma_1^2}{2}$, we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\nabla \Phi^h\|^2 + \frac{Re^{-1}}{2} \|\Delta \Phi^h\|^2 &\leq C_4^*(Re) \left[\|\nabla \eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta \eta\|^2 \right. \\ &\left. + C_7^*(\Gamma_1) \|\Delta \psi\|^2 \|\Delta \eta\|^2 + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta \psi^h\| \|\Delta \eta\|^2 \right] + C_5^*(Re, \Gamma_3) \|\Delta \psi\|^4 \|\nabla \Phi^h\|^2. \end{aligned} \quad (5.95)$$

Let $a(t) := 2 C_5^*(Re, \Gamma_3) \|\Delta \psi\|^4$ and

$$A(t) := \int_0^t a(t') dt' < \infty. \quad (5.96)$$

Multiplying (5.95) by the integrating factor $e^{-A(t)}$, we get

$$\begin{aligned} &\left\{ \frac{d}{dt} \left[\|\nabla \Phi^h\|^2 \right] - 2 C_5^*(Re, \Gamma_3) \|\Delta \psi\|^4 \|\nabla \Phi^h\|^2 \right\} e^{-A(t)} + Re^{-1} \|\Delta \Phi^h\|^2 e^{-A(t)} \\ &\leq 2 C_4^*(Re) \left[\|\nabla \eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta \eta\|^2 + C_7^*(\Gamma_1) \|\Delta \psi\|^2 \|\Delta \eta\|^2 \right. \\ &\quad \left. + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta \psi^h\| \|\Delta \eta\|^2 \right] e^{-A(t)}, \end{aligned}$$

which can also be written as

$$\begin{aligned} &\left\{ e^{-A(t)} \frac{d}{dt} \left[\|\nabla \Phi^h\|^2 \right] - \frac{d}{dt} [A(t)] e^{-A(t)} \|\nabla \Phi^h\|^2 \right\} + Re^{-1} \|\Delta \Phi^h\|^2 e^{-A(t)} \\ &\leq 2 C_4^*(Re) \left[\|\nabla \eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta \eta\|^2 + C_7^*(\Gamma_1) \|\Delta \psi\|^2 \|\Delta \eta\|^2 \right. \\ &\quad \left. + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta \psi^h\| \|\Delta \eta\|^2 \right] e^{-A(t)}, \end{aligned}$$

and simplifies to

$$\begin{aligned} &\frac{d}{dt} [e^{-A(t)} \|\nabla \Phi^h\|^2] + Re^{-1} \|\Delta \Phi^h\|^2 e^{-A(t)} \\ &\leq 2 C_4^*(Re) \left[\|\nabla \eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta \eta\|^2 + C_7^*(\Gamma_1) \|\Delta \psi\|^2 \|\Delta \eta\|^2 \right. \\ &\quad \left. + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta \psi^h\| \|\Delta \eta\|^2 \right] e^{-A(t)}, \end{aligned}$$

Now, integrating over $[0, T]$ and multiplying by $e^{A(T)}$ gives

$$\begin{aligned} \|\nabla\Phi^h(T)\|^2 + Re^{-1} \int_0^T \|\Delta\Phi^h\|^2 e^{A(T)-A(t)} dt &\leq e^{A(T)-A(0)} \|\nabla\Phi^h(0)\|^2 \\ &\quad + 2C_4^*(Re) \left[\int_0^T \|\nabla\eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta\eta\|^2 e^{A(T)-A(t)} dt \right. \\ &\quad \left. + \int_0^T \left(C_7^*(\Gamma_1) \|\Delta\psi\|^2 + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta\psi^h\| \right) \|\Delta\eta\|^2 e^{A(T)-A(t)} dt \right]. \end{aligned} \quad (5.97)$$

Noting that $e^{A(T)-A(t)} \geq 1$, $e^{A(T)-A(t)} \leq e^{A(T)}$, and $A(0) = 0$, (5.97) becomes

$$\begin{aligned} \|\nabla\Phi^h(T)\|^2 + Re^{-1} \int_0^T \|\Delta\Phi^h\|^2 dt &\leq C_8^*(T, Re, \Gamma_3) \|\nabla\Phi^h(0)\|^2 \\ &\quad + C_9^*(T, Re, \Gamma_3) \left[\int_0^T \|\nabla\eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta\eta\|^2 dt \right. \\ &\quad \left. + \int_0^T \left(C_7(\Gamma_1) \|\Delta\psi\|^2 + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) \|\Delta\psi^h\| \right) \|\Delta\eta\|^2 dt \right], \end{aligned} \quad (5.98)$$

where

$$C_8^*(T, Re, \Gamma_3) = \exp\left(2 \frac{27}{256} 9^3 Re^3 \Gamma_3^4 \int_0^T \|\Delta\psi\|^4 dt\right), \quad (5.99)$$

$$C_9^*(T, Re, \Gamma_3) = 9Re \exp\left(2 \frac{27}{256} 9^3 Re^3 \gamma_3^4 \int_0^T \|\Delta\psi\|^4 dt\right). \quad (5.100)$$

By the Hölder inequality, (5.2) we have

$$\int_0^T \|\Delta\psi^h\| \|\Delta\eta\|^2 dt \leq \|\Delta\psi^h\|_{L^2(0,T;L^2)}^2 \|\Delta\eta\|_{L^4(0,T;L^2)}^2, \quad (5.101)$$

$$\int_0^T \|\Delta\psi\|^2 \|\Delta\eta\|^2 dt \leq \|\Delta\psi\|_{L^4(0,T;L^2)}^2 \|\Delta\eta\|_{L^4(0,T;L^2)}^2. \quad (5.102)$$

Note that $\|\Delta\psi^h\|_{L^2(0,T;L^2)} \leq C_{10}^*(Re, Ro, F)$ from the stability bound (Proposition 1), while $\|\Delta\psi\|_{L^4(0,T;L^2)} \leq C_{11}^*(\|\Delta\psi\|_{L^4(0,T;L^2)})$ by hypothesis. Thus, (5.98) can be written as

$$\begin{aligned} \|\nabla\Phi^h\|^2 + Re^{-1} \int_0^T \|\Delta\Phi^h\|^2 dt &\leq C_8^*(T, Re, \Gamma_3) \|\nabla\Phi^h(0)\|^2 \\ &\quad + C_9^*(T, Re, \Gamma_3) \left[\int_0^T \|\nabla\eta_t\|_{-1}^2 + C_6^*(Re, Ro, \Gamma_2) \|\Delta\eta\|^2 dt \right. \\ &\quad + (C_7^*(\Gamma_1) C_{11}^*(\|\Delta\psi\|_{L^4(0,T;L^2)})) \\ &\quad \left. + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) C_{10}^*(Re, Ro, F) \right] \|\Delta\eta\|_{L^4(0,T;L^2)}^2. \end{aligned}$$

Remark 5. We note that the stability bound in Proposition 1 does not provide an estimate for $\|\Delta\psi^h\|_{L^4(0,T;L^2)}$. This was the reasoning for treating the nonlinear terms $b(\eta; \psi, \Phi^h)$ and $b(\psi^h; \eta, \Phi^h)$ in (5.83) differently.

Now adding $\|\nabla\eta(T)\|^2$ and $Re^{-1} \int_0^T \|\Delta\eta\|^2 dt$ to both sides and using the triangle inequality gives

$$\begin{aligned} \frac{1}{2} \|\nabla(\psi - \psi^h)(T)\|^2 + \frac{Re^{-1}}{2} \int_0^T \|\Delta(\psi - \psi^h)\|^2 dt &\leq C_8^*(T, Re, \Gamma_3) \|\nabla\Phi^h(0)\|^2 \\ &+ C_9^*(T, Re, \Gamma_3) \int_0^T \|\nabla(\psi - \lambda^h)\|_{-1}^2 dt + (1 + C_6^*(Re, Ro, \Gamma_2)) \|\Delta(\psi - \lambda^h)\|^2 dt \\ &+ (C_7^*(\Gamma_1) C_{11}^*(\|\Delta\psi\|_{L^4(0,T;L^2)})) \\ &+ C_3^*(F, \psi_0, Re, Ro, \Gamma_4) C_{10}^*(Re, Ro, F) \|\Delta(\psi - \lambda^h)\|_{L^4(0,T;L^2)}^2 \\ &+ \|\nabla(\psi - \lambda^h)(T)\|^2. \end{aligned} \quad (5.103)$$

By the triangle inequality we have $2\|\nabla e(0)\|^2 \geq \|\nabla\Phi^h(0)\|^2 + \|\nabla\eta(0)\|^2$, therefore adding $C_8^*(T, Re, \Gamma_3) \|\nabla\eta(0)\|^2$ to the right hand side of (5.103) and using the triangle inequality gives

$$\begin{aligned} \frac{1}{2} \|\nabla(\psi - \psi^h)(T)\|^2 + \frac{Re^{-1}}{2} \int_0^T \|\Delta(\psi - \psi^h)\|^2 dt &\leq 2C_8^*(T, Re, \Gamma_3) \|\nabla e(0)\|^2 \\ &+ C_9^*(T, Re, \Gamma_3) \int_0^T \|\nabla(\psi - \lambda^h)\|_{-1}^2 dt + (1 + C_6^*(Re, Ro, \Gamma_2)) \|\Delta(\psi - \lambda^h)\|^2 dt \\ &+ (C_7^*(\Gamma_1) C_{11}^*(\|\Delta\psi\|_{L^4(0,T;L^2)})) \\ &+ C_3^*(F, \psi_0, Re, Ro, \Gamma_4) C_{10}^*(Re, Ro, F) \|\Delta(\psi - \lambda^h)\|_{L^4(0,T;L^2)}^2 \\ &+ \|\nabla(\psi - \lambda^h)(T)\|^2. \end{aligned} \quad (5.104)$$

Finally, taking $\inf_{\lambda^h \in X^h}$ of both sides of (5.104) and letting

$$\begin{aligned} C_1(T, Re, \Gamma_3) &= 4C_8^*(T, Re, \Gamma_3) \\ C_2(T, Re, \Gamma_3) &= 2C_9^*(T, Re, \Gamma_3) \\ C_3(Re, Ro, \Gamma_2) &= 2 + 2C_6^*(Re, Ro, \Gamma_2) \\ C_4(T, F, \psi_0, Re, Ro, \Gamma_1, \Gamma_3, \Gamma_4, \|\Delta\psi\|_{L^4(0,T;L^2)}) &= 2C_9^*(T, Re, \Gamma_3) \\ &\quad \left(C_7^*(\Gamma_1) C_{11}^*(\|\Delta\psi\|_{L^4(0,T;L^2)}) + C_3^*(F, \psi_0, Re, Ro, \Gamma_4) C_{10}^*(F, \psi_0, Re, Ro) \right) \end{aligned}$$

gives

$$\begin{aligned} \|\nabla(\psi - \psi^h)(T)\|^2 + Re^{-1} \int_0^T \|\Delta(\psi - \psi^h)\|^2 dt &\leq C_1(T, Re, \Gamma_3) \left\| \nabla[\psi - \psi^h](0) \right\|^2 \\ &+ \inf_{\lambda^h \in X^h} \left\{ C_2(T, Re, \Gamma_3) \int_0^T \left\| \nabla(\psi - \lambda^h) \right\|_{-1}^2 dt + C_3(Re, Ro, \Gamma_2) \left\| \Delta(\psi - \lambda^h) \right\|^2 dt \right. \\ &+ C_4(T, F, \psi_0, Re, Ro, \Gamma_1, \Gamma_3, \Gamma_4, \|\Delta\psi\|_{L^4(0,T;L^2)}) \left\| \Delta(\psi - \lambda^h) \right\|_{L^4(0,T;L^2)}^2 \\ &\left. + 2\|\nabla(\psi - \lambda^h)(T)\|^2 \right\}, \end{aligned}$$

which is the desired result. \square

Next we determine the FE rates of convergence yielded by the error estimate (5.77) in Theorem 6. The following lemma will allow for the determination of FE convergence rates in the terms involving $\frac{\partial}{\partial t}$ in the error estimate (5.77).

Lemma 7. *Suppose the nodes N_j of the finite element mesh do not move. Let $X^h \subset X$, consisting of Argyris elements and I^h the associated \mathbb{P}^5 -interpolation operator (see Theorem 6.1.1 in [20]). Then, assuming that $\psi, \psi_t \in H^6$, we have*

$$\frac{\partial}{\partial t} (I^h \psi) = I^h \left(\frac{\partial \psi}{\partial t} \right), \quad (5.105)$$

$$\left\| \frac{\partial [\nabla \psi]}{\partial t} - \frac{\partial}{\partial t} \nabla (I^h \psi) \right\| \leq C h^5 \left| \frac{\partial \psi}{\partial t} \right|_6, \text{ and} \quad (5.106)$$

$$\left\| \frac{\partial [\nabla \psi]}{\partial t} - \frac{\partial}{\partial t} \nabla (I^h \psi) \right\|_{-1} \leq C h^5 \left| \frac{\partial \psi}{\partial t} \right|_6. \quad (5.107)$$

Remark 6. *Estimate (32) in Theorem 6 in Section 5.7 of [29] shows that $H^6 \hookrightarrow C^1$. Thus, the interpolation operator I^h can be applied to ψ and ψ_t .*

Proof. Equation (5.105) follows from the explicit formulas for the \mathbb{P}^5 interpolant, I^h (see Ciarlet [20]), estimate (5.106) follows from a combination of (5.105) and estimate (6.1.9) in Theorem 6.1.1 in [20] with $p = q = 2$ and $m = 1$. Finally, estimate (5.107) follows from the fact that $\|\cdot\|_{-1} \leq \|\cdot\|$. \square

Lemma 8. *Let X^h denote the FE space associated with the Argyris element. Suppose that the interpolation estimates from Lemma 7 hold and that $\psi, \psi_t \in H^6(\Omega)$. Then,*

$$\int_0^T \|\nabla (\psi - I^h \psi)_t\|_{-1}^2 + \|\Delta (\psi - I^h \psi)\|^2 dt \leq C h^8 \int_0^T h^2 |\psi_t|_6^2 + |\psi|_6^2 dt \quad (5.108)$$

and

$$\|\Delta (\psi - I^h \psi)\|_{L^4(0,T;L^2(\Omega))}^2 \leq C h^8 |\psi|_{L^4(0,T;H^6(\Omega))}^2. \quad (5.109)$$

Proof. At each time instance we see from inequality (6.1.9) in [20] that $\|\Delta (\psi - I^h \psi)\| \leq C h^4 |\psi|_6$. Squaring and integrating this and using the interpolation error bounds from Lemma 7 gives the first estimate. The second estimate follows analogously, i.e.

$$\begin{aligned} \|\Delta (\psi - I^h \psi)\|_{L^4(0,T;L^2(\Omega))} &= \left(\int_0^T \|\Delta (\psi - I^h \psi)\|^4 dt \right)^{\frac{1}{4}} \\ &\leq C h^4 \left(\int_0^T |\psi|_6^4 dt \right)^{\frac{1}{4}}, \end{aligned}$$

which leads to the desired estimate $\|\Delta (\psi - I^h \psi)\|_{L^4(0,T;L^2(\Omega))}^2 \leq C h^8 \|\psi\|_{L^4(0,T;H^6(\Omega))}^2$. \square

Theorem 7. *Let X^h be the FE space associated with the Argyris element and an I^h the associated \mathbb{P}^5 -interpolation operator (see Theorem 6.1.1 in [20]). Suppose the interpolation estimates from Lemma 7 hold and that $\psi, \psi_t \in H^6(\Omega)$. Suppose also that the assumptions of Theorem 6 hold. Then,*

$$\begin{aligned} \|\nabla (\psi - \psi^h)(T)\|^2 + Re^{-1} \int_0^T \|\Delta (\psi - \psi^h)\|^2 dt &\leq h^8 C \left\{ (C_1(T, Re, \Gamma_3) + 2) h^2 |\psi|_6^2 \right. \\ &\quad + C_2(T, Re, \Gamma_3) \left(h^2 \|\psi_t\|_{L^2(0,T;H^6(\Omega))}^2 + C_3(Re, Ro, \Gamma_2) \|\psi\|_{L^2(0,T;H^6(\Omega))}^2 \right) \\ &\quad \left. + C_4(T, F, \psi_0, Re, Ro, \Gamma_1, \Gamma_3, \Gamma_4, \|\psi\|_{L^4(0,T;L^2)}) \|\psi\|_{L^4(0,T;H^6(\Omega))}^2 \right\}. \end{aligned} \quad (5.110)$$

Proof. The proof follows from Theorem 6, Lemma 7, and Lemma 8. \square

Chapter 6

Numerical Experiments

In this section we aim to verify the rates of convergence, developed in Chapter 5, for both the QGE and SQGE. To this end we first use the *method of manufactured solutions* on a rectangular domain and calculate the L^2 , H^1 , H^2 errors from the difference of the exact solution and the numerical solution obtained from applying the Argyris element to the corresponding problem. Additionally, the rates of convergence for the two-level method applied to the SQGE is verified in a similar manner. Efficiency of the two-level method is also demonstrated by looking at the time required to simulate the SQGE in both the one-level method and the two-level method.

After verifying the rates of convergence obtained from the method of manufactured solutions, we then look to studying the rates of convergence on realistic domains with realistic parameter values. To this end we apply the FEM to the SQGE on both the Mediterranean and North Atlantic Ocean. We treat the finest mesh as our “true solution” and compare to the solutions obtained on coarser meshes.

6.1 SQGE

The main goal of this section is twofold. First, we show that the FE discretization of the streamfunction formulation of the SQGE (2.29) with the Argyris element produces accurate numerical approximations. To this end, we benchmark our numerical results against those in the published literature [81, 15, 61]. The second goal is to show that the numerical results follow the theoretical error estimates in Theorem 3 and Theorem 4, i.e., we compare the observed rates of convergence to the theoretical rates of convergence developed in Section 5.1.

Although the pure streamfunction formulation of the steady SQGE (2.29) is our main concern, we also test our Argyris FE discretization on two simplified settings: (i) the *linear Stommel* model; and (ii) *linear Stommel-Munk* model. The reason for using these two additional numerical tests is that they are standard test problems in the geophysical fluid dynamics literature (see, e.g., Chapter 13 in Vallis [81] as well as the reports of Myers *et al.* [61], and Cascon *et al.* [15]). This allows us to benchmark our numerical results against those in the published literature. Since both the linear Stommel and the linear Stommel-Munk models lack the nonlinearity present in the SQGE (2.29), they represent good stepping stones to verifying our FE discretization.

The first simplified model used in our numerical investigation is the *linear Stommel* model (equation (14.22) in [81] and equation (11) in [61])

$$\epsilon_s \Delta \psi + \frac{\partial \psi}{\partial x} = f, \quad (6.1)$$

where ϵ_s is the Stommel number. The Stommel number is given by (equation (10) in [61])

$$\epsilon_s = \frac{\gamma}{\beta L},$$

where γ , β , and L are the bottom friction decay, the measure of the β -plane effect, and characteristic length scale, respectively. We note that the linear Stommel Model (6.1) is just the linear Stommel-Munk model (6.2) in which the biharmonic term is dropped (i.e. $\epsilon_m = 0$). Thus it is a good starting point, since is less complex than both the SQGE and the linear Stommel-Munk model.

The *linear Stommel-Munk* model (see p. 587 in [81] and problem 2 in [15]) is given by

$$\epsilon_s \Delta \psi - \epsilon_m \Delta^2 \psi + \frac{\partial \psi}{\partial x} = f, \quad (6.2)$$

where ϵ_s and ϵ_m are the Stommel number and Munk scale, respectively. The Stommel number and Munk scale are given by (equation (10) in [61])

$$\epsilon_m = \frac{A}{\beta L^3},$$

where A is the eddy viscosity. The model is supplemented with appropriate boundary conditions, which will be described for each of the subsequent numerical tests.

We note that the linear Stommel-Munk model (6.2) is similar in form to the SQGE (2.29). Indeed, both models contain the biharmonic operator $\Delta^2 \psi$, the rotation term $\frac{\partial \psi}{\partial x}$, and the forcing terms f and F , respectively. The two main differences between the two models are the following: First, the SQGE are nonlinear, since they contain the Jacobian term $J(\psi, q)$; the Stommel-Munk model is linear, since it doesn't contain the Jacobian term. The second difference is that the linear Stommel-Munk model contains a Laplacian term $\Delta \psi$, whereas the SQGE does not.

We also note that the two models use different parameters: the Reynolds number Re and the Rossby number Ro in (2.29) and the Stommel number ϵ_s and the Munk scale ϵ_m in the linear Stommel-Munk model. These two sets of parameters, however, are related by the following relations:

$$\epsilon_m = Ro Re^{-1}, \quad (6.3)$$

$$\epsilon_s = Ro \frac{\gamma L}{U}. \quad (6.4)$$

The rest of the section is organized as follows: In Subsection 6.1.1 we present results for the linear Stommel model (6.1). In Subsection 6.1.2 we present results for the linear Stommel-Munk model (6.2). Finally, in Subsection 6.1.3 we present results for the nonlinear SQGE (2.29).

6.1.1 Linear Stommel Model

This subsection presents the results for the FE discretization of the linear Stommel model (6.1) by using the Argyris element. The computational domain is $\Omega = [0, 1] \times [0, 1]$. For completeness, we present results for two numerical tests. The first test, denoted by Test 1, corresponds to the exact solution used by Vallis (equation (14.38) in [81]), while the second test, denoted by Test 2, corresponds to the exact solution used by Myers *et al.* (Equations 15 and 16 in [61]).

Test 1a: In this test, we choose the same setting as that used by Vallis (equation (14.38) in [81]). In particular, the forcing term and the non-homogeneous Dirichlet boundary conditions are chosen to match those given by the exact solution

$$\psi(x, y) = (1 - x - e^{-x/\epsilon_s}) \pi \sin \pi y. \quad (6.5)$$

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	70	1.15×10^{-1}	—	1.81×10^0	—	8.37×10^1	—
$1/4$	206	1.02×10^{-2}	3.50	3.12×10^{-1}	2.54	2.55×10^1	1.72
$1/8$	694	4.46×10^{-4}	4.51	2.59×10^{-2}	3.59	3.90×10^0	2.71
$1/16$	2534	1.09×10^{-5}	5.36	1.22×10^{-3}	4.41	3.49×10^{-1}	3.48
$1/32$	9670	1.97×10^{-7}	5.79	4.35×10^{-5}	4.80	2.34×10^{-2}	3.90

Table 6.1: Errors and rate of convergence for the linear Stommel Model (6.1), Test 1 [81].

We choose the same Stommel number as that used by Vallis, i.e. $\epsilon_s = 0.04$. The exact solution (6.5) considered by Vallis satisfies $\psi \rightarrow 0$ as $x \rightarrow 0$, but does not vanish at $x = 1$. In our numerical tests, we used a standard lifting procedure to treat these non-homogeneous boundary conditions, i.e. for a problem of the form

$$\begin{aligned} L\psi &= f \text{ on } \Omega \\ \psi &= g \text{ on } \partial\Omega, \end{aligned}$$

we reformulate it to be

$$\begin{aligned} L\tilde{\psi} &= \tilde{f} \text{ on } \Omega \\ \tilde{\psi} &= 0 \text{ on } \partial\Omega \end{aligned}$$

where $L\tilde{\psi} = Lu - LS = f - g = \tilde{f}$. The solution ψ from the original problem can then be found by $\psi(x, y) = \tilde{\psi} + S$. The function $S(x, y)$ is assumed to have the form

$$S(x, y) = A(y)(1 - x) + B(y)x$$

and satisfies the boundary conditions given by (6.6). After some simple algebra we see that

$$S(x, y) = -xe^{-x/\epsilon_s} \pi \sin(\pi y).$$

The function \tilde{f} can be determined by applying the operator L corresponding to the *linear Stommel* problem to $u - S$.

Applying the finite element method to the *linear Stommel* problem with the new modified \tilde{f} , corresponding to the exact solution given by Vallis, and homogeneous boundary conditions using Argyris Finite Elements we get a solution that matches the solution presented by Vallis, as can be seen in Figure 6.2. Additionally, the table of errors Table 6.1 shows the order of convergence appears to be approaching the expected rates for L^2 , H^1 , and H^2 norms. Figure 6.2 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 9670 DoFs. Comparing Figure 6.2 with Figure 14.5 in [81], we notice that our approximation is close to his. Since the exact solution is available, we can compute the errors in various norms. Table 6.1 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h (the DoFs are also included).

We note that the errors in Table 6.1 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. The orders of convergence in Table 6.1 are close to the theoretical ones for the fine meshes, but not as close for the coarse meshes. We think that

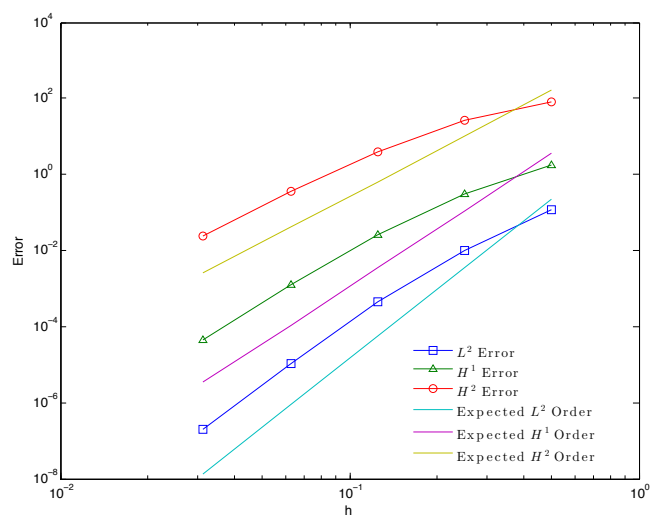


Figure 6.1: Linear Stommel model (6.1), Test 1a [81]: Observed errors versus expected rate of convergence.

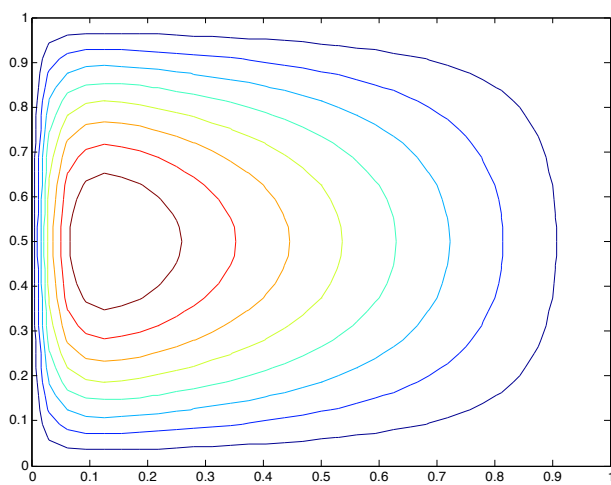


Figure 6.2: Linear Stommel model (6.1), Test 1a [81]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	70	1.69×10^{-5}	—	3.43×10^{-4}	—	8.72×10^{-3}	—
$1/4$	206	3.72×10^{-7}	5.50	1.34×10^{-5}	4.68	5.62×10^{-4}	3.96
$1/8$	694	4.89×10^{-9}	6.25	3.76×10^{-7}	5.16	3.25×10^{-5}	4.11
$1/16$	2534	7.08×10^{-11}	6.11	1.12×10^{-8}	5.07	1.96×10^{-6}	4.05
$1/32$	9670	1.08×10^{-12}	6.04	3.44×10^{-10}	5.02	1.21×10^{-7}	4.02

Table 6.2: Errors and rate of convergence for the linear Stommel Model (6.1), Test 1b [81].

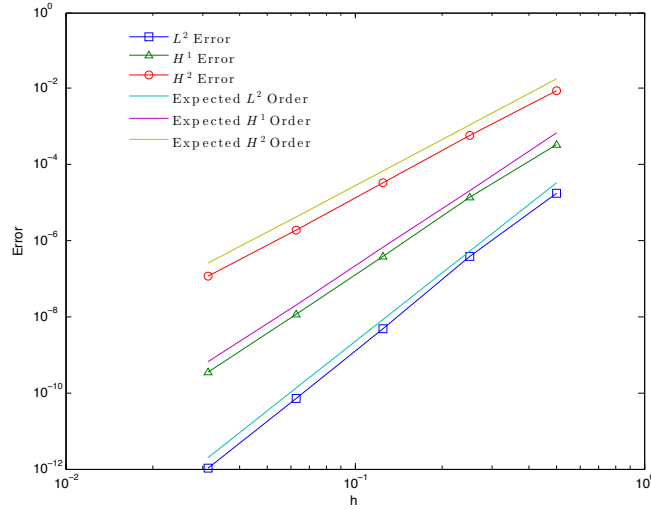


Figure 6.3: Linear Stommel model (6.1), Test 1b [81]: Observed errors versus expected rate of convergence.

the inaccuracies on the coarse meshes are due to their inability to capture the thin boundary layer on the left-hand side (i.e., at $x = 0$). The finer the mesh gets, the better this boundary layer is captured and the better the numerical accuracy becomes.

Test 1b: In the Second part of Test 1, we verify the hypothesis above, that is, whether the degrading accuracy of the approximation is indeed due to the thin western boundary layer. To this end, we change the Stommel number in Test 1a to be $\epsilon_s = 1$, which will result in a much thicker western boundary layer. We then run the same numerical test as before, but with the new Stommel number. As can be seen in Table 6.2, the rates of convergence are the expected theoretical orders of convergence. This shows that the reason for the inaccuracies in Table 6.1 were indeed due to the thin western boundary layer.

Test 2: For our second test we use the exact solution given by Myers (Equations 15 and 16 in [61]), i.e.

$$\psi(x, y) = \frac{\sin(\pi y)}{\pi(1 + 4\pi^2\epsilon_s^2)} \left\{ 2\pi\epsilon_s \sin(\pi x) + \cos(\pi x) + \frac{1}{e^{R_1} - e^{R_2}} [(1 + e^{R_2})e^{R_1 x} - (1 + e^{R_1})e^{R_2 x}] \right\}, \quad (6.6)$$

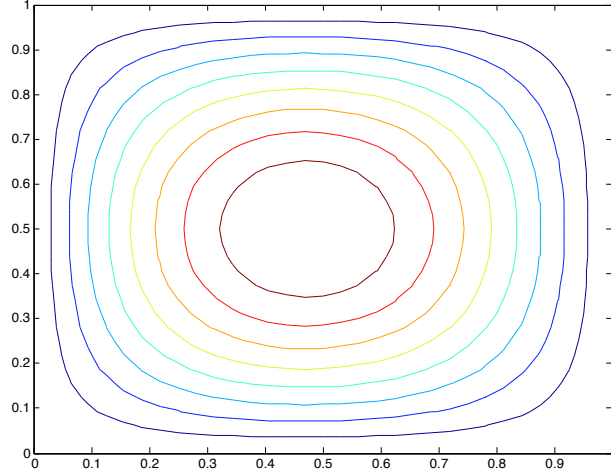


Figure 6.4: Linear Stommel model (6.1), Test 1b [81]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs with $\epsilon_s = 1$.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	70	5.65×10^{-3}	—	1.45×10^{-1}	—	6.60×10^0	—
$1/4$	206	4.28×10^{-4}	3.72	2.08×10^{-2}	2.80	1.63×10^0	2.02
$1/8$	694	1.46×10^{-5}	4.87	1.41×10^{-3}	3.89	2.07×10^{-1}	2.98
$1/16$	2534	2.95×10^{-7}	5.63	5.83×10^{-5}	4.59	1.65×10^{-2}	3.65
$1/32$	9670	4.97×10^{-9}	5.89	2.00×10^{-6}	4.87	1.07×10^{-3}	3.95

Table 6.3: Errors and rate of convergence for the linear Stommel Model (6.1), Test 2 [61].

where R_1 and R_2 are the positive and negative roots, respectively, of

$$R = \frac{-1 \pm \sqrt{1 + 4\pi^2 \epsilon_s^2}}{2\epsilon_s}.$$

The forcing term and the homogeneous Dirichlet boundary conditions are chosen to match those given by the exact solution (6.6). We choose the same Stommel number as that used by Myers, i.e. $\epsilon_s = 0.05$.

Figure 6.6 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 9670 DoFs. Comparing Figure 6.6 with Figure 2 in [61], we notice that our approximation is close to that in [61]. Since the exact solution is available, we can compute the errors in various norms. Table 6.3 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h .

We note that the errors in Table 6.3 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. Again, we see that the orders of convergence in Table 6.3 are close to the theoretical ones for the fine meshes, but not as close for the coarse meshes. We attribute this to the inaccuracies at the thin boundary layer on the left-hand side (i.e., at $x = 0$). The finer the mesh gets, the better this boundary layer is captured and the better the numerical accuracy becomes.

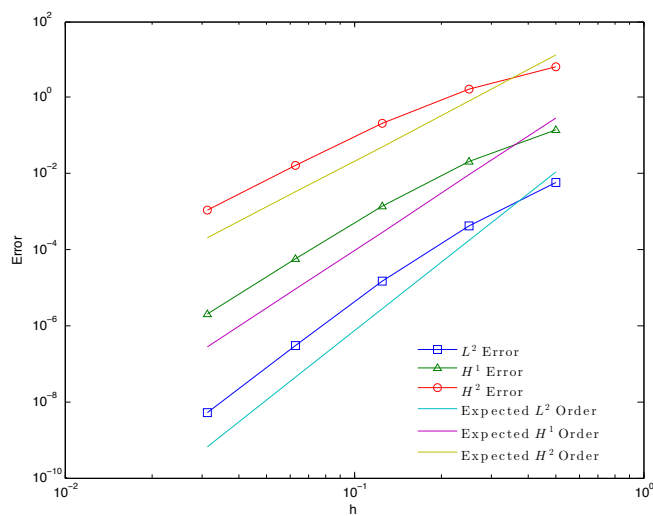


Figure 6.5: Linear Stommel model (6.1), Test 2 [61]: Observed errors versus expected rate of convergence.

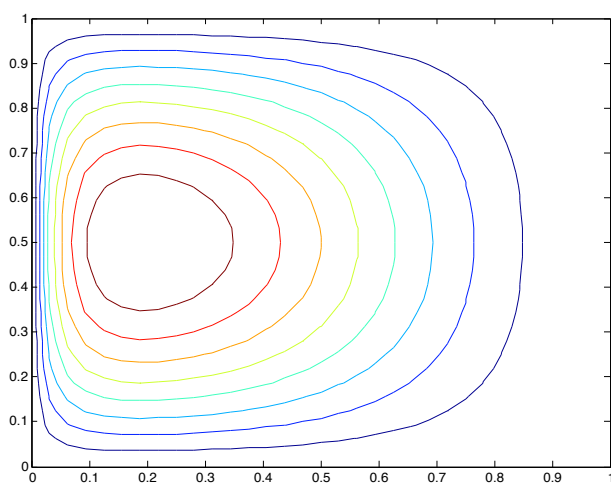


Figure 6.6: Linear Stommel model (6.1), Test 2 [61]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 9670 DoFs.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	170	2.99×10^{-3}	—	4.08×10^{-2}	—	7.62×10^{-1}	—
$1/4$	550	3.22×10^{-5}	6.54	1.03×10^{-3}	5.31	4.08×10^{-2}	4.23
$1/8$	1958	3.44×10^{-7}	6.55	2.49×10^{-5}	5.37	2.25×10^{-3}	4.18
$1/16$	7366	4.57×10^{-9}	6.23	7.03×10^{-7}	5.15	1.34×10^{-4}	4.07
$1/32$	28550	6.70×10^{-11}	6.09	2.11×10^{-8}	5.06	8.26×10^{-6}	4.02

Table 6.4: Errors and rate of convergence for the linear Stommel-Munk Model (6.7), Test 3 [15].

6.1.2 Linear Stommel-Munk Model

This subsection presents results for the FE discretization of the linear Stommel-Munk model (6.2) by using the Argyris element. Our computational setting is the same as that used by Cascon *et al.* [15]: The computational domain is $\Omega = [0, 3] \times [0, 1]$, the Munk scale is $\epsilon_m = 6 \times 10^{-5}$, the Stommel number is $\epsilon_s = 0.05$, and the boundary conditions are

$$\psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega \quad (6.7)$$

For completeness, we present results for two numerical tests, denoted by Test 3 and Test 4, both corresponding to Test 1 and Test 2 in [15], respectively.

Test 3: For our third test we use the exact solution given by Test 1 in [15], i.e.

$$\psi(x, y) = \sin^2 \frac{\pi x}{3} \sin^2 \pi y. \quad (6.8)$$

The forcing term is chosen to match that given by the exact solution (6.8).

For this third test we take F corresponding to applying the linear operator L associated with the *linear Stommel-Munk* model to the exact solution (6.8).

Figure 6.8 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 28550 DoFs. Comparing Figure 6.8 with Figure 7 in [15], we notice that our approximation is close to that in [61]. Since the exact solution is available, we can compute the errors in various norms. Table 6.4 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h .

We note that the errors in Table 6.4 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. This time, we see that the orders of convergence in Table 6.4 are close to the theoretical ones for the fine meshes, but are higher than expected for the coarse meshes. We attribute this to the fact that the exact solution (6.8) does not display any boundary layers that could be challenging to capture by the Argyris element on a coarse mesh.

Test 4: For our fourth test we use the exact solution given by Test 2 in [15], i.e.

$$\psi(x, y) = \left[\left(1 - \frac{x}{3} \right) (1 - e^{-20x}) \sin \pi y \right]^2. \quad (6.9)$$

Again we take the forcing term F corresponding the exact solution (6.9).

Figure 6.10 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 28550 DoFs. Comparing Figure 6.10 with Figure 10

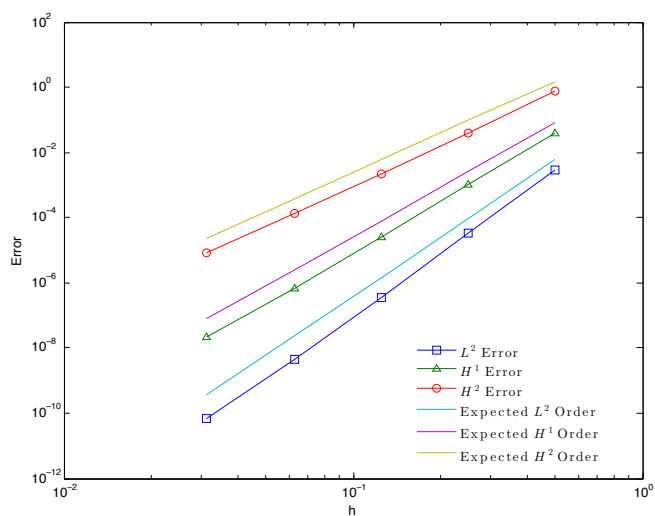


Figure 6.7: Linear Stommel-Munk model (6.7), Test 3 [15]: Observed errors versus expected rate of convergence.

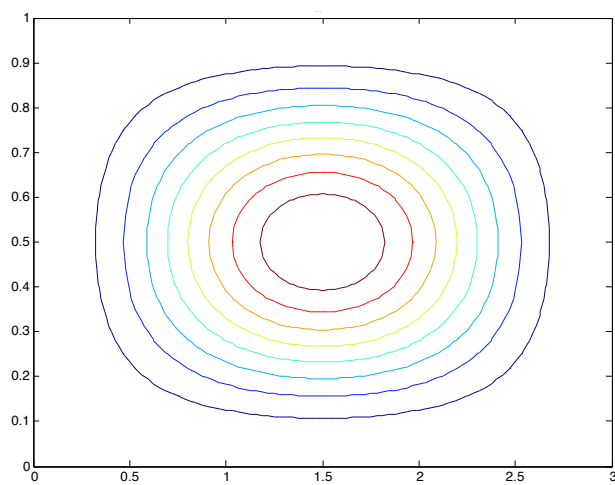


Figure 6.8: Linear Stommel-Munk model (6.7), Test 3 [15]: Streamlines of the approximation, ψ^h , on a mesh size, $h = \frac{1}{32}$, and 28550 DoFs.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	170	6.04×10^{-2}	—	1.16×10^0	—	3.90×10^1	—
$1/4$	550	1.13×10^{-2}	2.41	4.00×10^{-1}	1.54	2.14×10^1	0.866
$1/8$	1958	8.40×10^{-4}	3.75	5.91×10^{-2}	2.76	5.66×10^0	1.92
$1/16$	7366	2.82×10^{-5}	4.90	4.01×10^{-3}	3.88	7.38×10^{-1}	2.94
$1/32$	28550	5.59×10^{-7}	5.66	1.61×10^{-4}	4.64	5.97×10^{-2}	3.63

Table 6.5: Errors and rate of convergence for the linear Stommel-Munk Model (6.7), Test 4 [15].

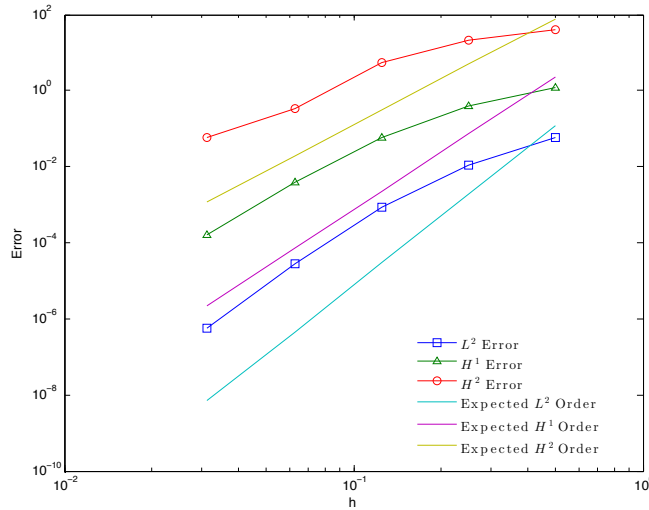


Figure 6.9: Linear Stommel-Munk model (6.7), Test 4 [15]: Observed errors versus expected rate of convergence.

in [61], we notice that our approximation is close to [61]. Since the exact solution is available, we can compute the errors in various norms. Table 6.5 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h .

We note that the errors in Table 6.5 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. Again, we see that the orders of convergence in Table 6.5 are close to the theoretical ones for the fine meshes, but not as close for the coarse meshes. As stated previously, we attribute this to the inaccuracies at the thin boundary layer on the left-hand side (i.e., at $x = 0$). The finer the mesh gets, the better this boundary layer is captured and the better the numerical accuracy becomes.

6.1.3 SQGE

This subsection presents results for the FE discretization of the streamfunction formulation of the SQGE (2.29) by using the Argyris element. Our computational domain is $\Omega = [0, 3] \times [0, 1]$, the Reynolds number is $Re = 1.667$, and the Rossby number is $Ro = 10^{-4}$. For completeness, we present results for two numerical tests, denoted by Test 5 and Test 6, both corresponding to the

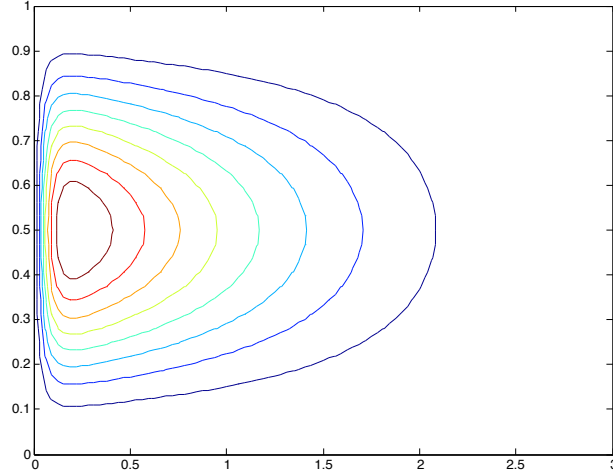


Figure 6.10: Linear Stommel-Munk model (6.7), Test 4 [15]: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 28550 DoFs.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	138	4.10×10^{-3}	—	4.91×10^{-2}	—	8.61×10^{-1}	—
$1/4$	536	2.23×10^{-5}	7.52	7.66×10^{-4}	6.00	2.96×10^{-2}	4.86
$1/8$	2349	2.35×10^{-7}	6.57	1.60×10^{-5}	5.58	1.31×10^{-3}	4.50
$1/16$	9152	3.03×10^{-9}	6.28	4.33×10^{-7}	5.21	7.53×10^{-5}	4.12
$1/32$	36150	4.64×10^{-11}	6.03	1.37×10^{-8}	4.99	4.84×10^{-6}	3.96
$1/64$	146090	9.86×10^{-13}	5.56	4.10×10^{-10}	5.06	2.92×10^{-7}	4.05

Table 6.6: Errors and rate of convergence for the SQGE (2.29), Test 5.

exact solutions given in Test 1 and Test 2 of [15], respectively.

Test 5: In this test, we take the same exact solution presented in *Test 3*, i.e.

$$\psi(x, y) = \sin^2 \frac{\pi x}{3} \sin^2 \pi y. \quad (6.10)$$

Again, the forcing term F and homogeneous boundary conditions, $\psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0$, correspond to the exact solution (6.10).

Figure 6.12 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 36150 DoFs. We note that the streamlines look as we expect and are similar to those given by Figure 7 in [61], which uses the same exact solution. Since the exact solution is available, we can compute the errors in various norms. Table 6.6 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h .

We note that the errors in Table 6.6 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. Again, since the exact solution (6.10) does not display any boundary layers, we see that the orders of convergence in Table 6.6 are close to the theoretical ones for the fine meshes, but are higher than expected for the coarse meshes. We also note

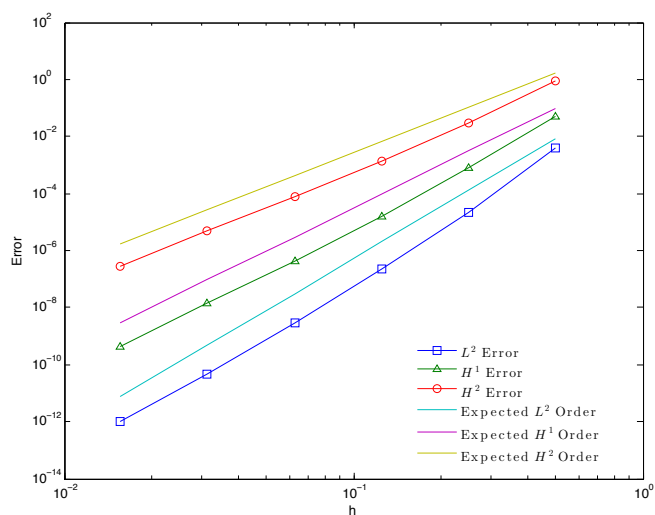


Figure 6.11: SQGE (2.29), Test 5: Observed errors versus expected rate of convergence.

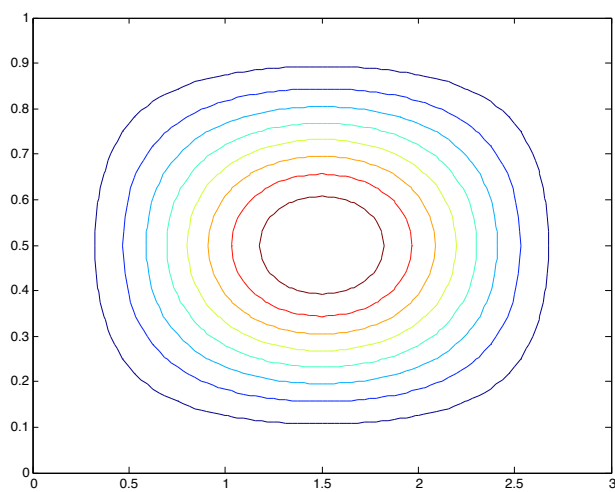


Figure 6.12: SQGE (2.29), Test 5: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 36150 DoFs.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/2$	138	4.36×10^{-1}	—	2.22×10^0	—	4.58×10^1	—
$1/4$	536	7.21×10^{-3}	5.92	2.86×10^{-1}	2.96	1.65×10^1	1.48
$1/8$	2349	1.92×10^{-3}	1.91	1.07×10^{-1}	1.42	8.06×10^0	1.03
$1/16$	9152	2.89×10^{-5}	6.05	3.94×10^{-3}	4.76	6.84×10^{-1}	3.56
$1/32$	36150	6.63×10^{-7}	5.45	1.75×10^{-4}	4.50	5.77×10^{-2}	3.57
$1/64$	146090	1.37×10^{-8}	5.60	6.91×10^{-6}	4.66	4.36×10^{-3}	3.73
$1/128$	585048	1.47×10^{-10}	6.55	1.63×10^{-7}	5.40	2.28×10^{-4}	4.26

Table 6.7: Errors and rate of convergence for the SQGE (2.29), Test 6.

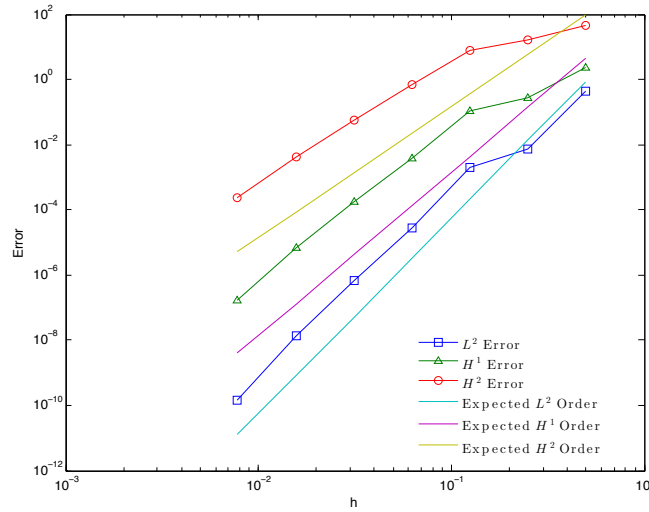


Figure 6.13: SQGE (2.29), Test 6: Observed errors versus expected rate of convergence.

the drop in the rate of convergence in the L^2 -norm for $1/64$ is due to reaching the machine epsilon and thus any rate of convergence for mesh size smaller than $1/64$ is meaningless.

Test 6: In this test, we take the same exact solution as used in *Test 4*, i.e.

$$\psi(x, y) = \left[\left(1 - \frac{x}{3} \right) (1 - e^{-20x}) \sin \pi y \right]^2. \quad (6.11)$$

The forcing term F and the homogeneous boundary conditions correspond to the exact solution (6.11).

Figure 6.14 presents the streamlines of the approximate solution obtained by using the Argyris Finite Element on a mesh with $h = \frac{1}{32}$ and 36150 DoFs. We note that the streamlines look as we expect and are similar to those given by Figure 10 in [61], which uses the same exact solution. Since the exact solution is available, we can compute the errors in various norms. Table 6.7 presents the errors e_0 , e_1 , and e_2 (i.e., the L^2 , H^1 , and H^2 errors, respectively) for various values of the mesh sizes, h .

We note that the errors in Table 6.7 follow the theoretical rates of convergence predicted by the estimates (5.40) - (5.42) in Theorem 4. Again for an exact solution which has a western boundary layer, we see that the orders of convergence in Table 6.7 are close to the theoretical ones for the

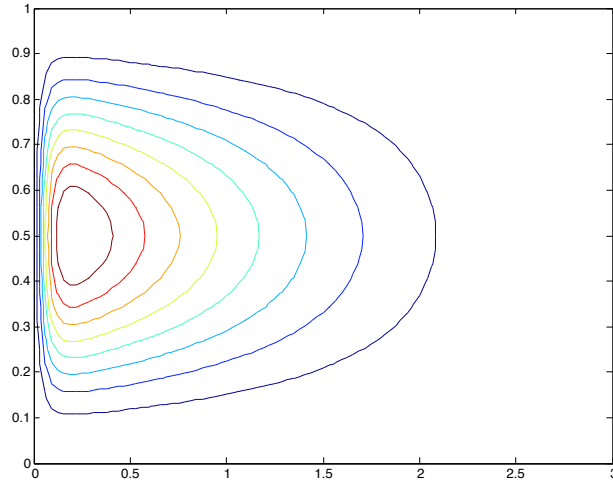


Figure 6.14: SQGE (2.29), Test 6: Streamlines of the approximation, ψ^h , $h = \frac{1}{32}$, and 36150 DoFs.

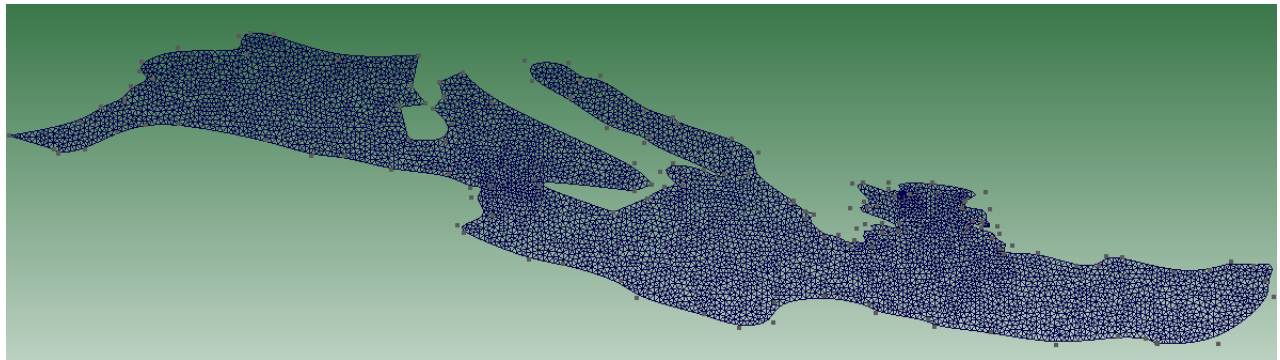


Figure 6.15: Mesh of the Mediterranean Sea created using GMSH [38]. The mesh size corresponds to $h = \frac{1}{320}$ with $DoF = 240,342$.

fine meshes, but not as close for the coarse meshes. We attribute this to the inaccuracies at the thin boundary layer on the left-hand side (i.e., at $x = 0$). The finer the mesh gets, the better this boundary layer is captured and the better the numerical accuracy becomes.

Mediterranean Sea

We have created a FE mesh of the Mediterranean Sea using GMSH [38]. The coastline data was obtained from GSHHS [17]. Major islands such as Corsica, Sardinia, and Sicily were connected to the nearest land mass in order to ensure a unique streamfunction (see the discussions in [43, 82]). Additionally, the Atlantic Ocean was closed off from the Mediterranean Sea at the Straits of Gibraltar, from the Red Sea at the Suez Canal, and the Sea of Marmara at the Dardanelles Strait, while the Gulf of Corinth was treated as land. The resultant FE mesh is displayed in Figure 6.15.

The following experiment used a forcing function given by $F = \sin(\frac{\pi}{4}y)$ which is the same forcing function given by Bryan [14]. Note: Bryan, in fact, used $\sin(\frac{\pi}{2}y)$, but on a domain with a

A	$2000 \text{ m}^2 \text{ s}^{-1}$
θ_0	40°
ω	$7.2526 \times 10^{-5} \text{ s}^{-1}$
H	1000 m
L	1000 km
r_e	6378.1 km
ρ	1024 kg/m^3

Table 6.8: Table of parameter values used for the simulations of the Mediterranean Sea [8], where $A, \theta_0, \omega, H, L, r_e, \rho$ are the eddy viscosity, reference angle for the β -plane approximation, angular velocity of the Earth, domain height, domain length, radius of the Earth, and density of seawater, respectively.

vertical extent which was of length two. In our case the vertical extent of the Northern Hemisphere is of length one. Thus, the appropriate forcing function becomes $\sin(\frac{\pi}{4}y)$. We choose similar parameters as those used in [8] and are summarized in Table 6.8.

Taking the derivative of relation (2.4) with respect to y gives the following relation for β to f and using the equation (2.80) from [81] gives

$$\beta = \frac{2\omega}{r_e} \cos \theta_0. \quad (6.12)$$

From the parameters given in Table 6.8 we see that

$$\beta \approx 1.742 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}.$$

Using this approximation for β and (2.6) with $\tau_0 = 0.6 \text{ dyne cm}^{-2}$ [45] gives the following approximation for the characteristic velocity:

$$U \approx 1.054 \times 10^{-2} \text{ m/s}.$$

Therefore by (2.10), the Rossby number is

$$Ro = 6.051 \times 10^{-4}$$

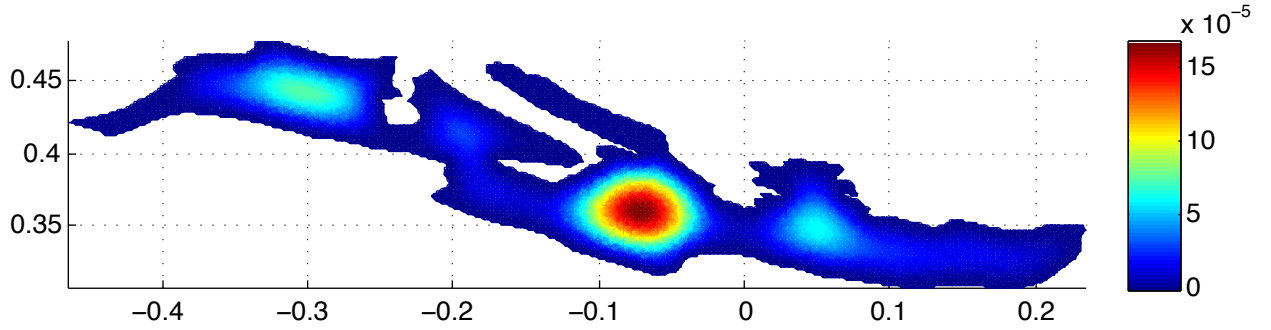
and by (2.18) the Reynolds number is

$$Re = 5.27.$$

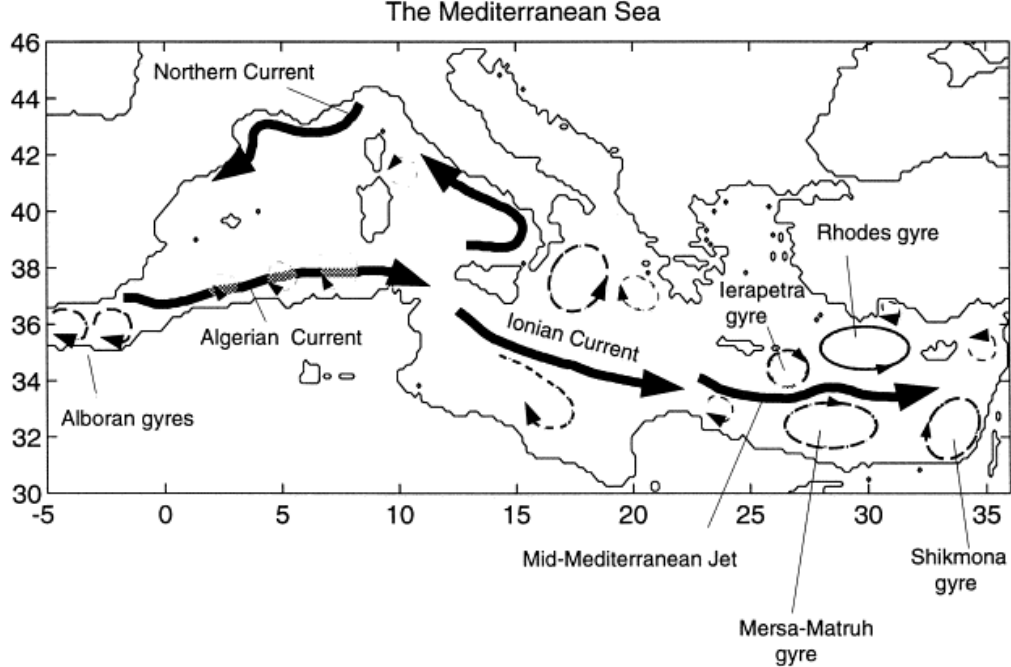
The solution obtained by applying the SQGE to the Mediterranean with parameters given in Table 6.8 and forcing function $F = \sin(\frac{\pi}{2}y)$ can be seen in Figure 6.16a. The shape of the streamfunction is very much in agreement with the solution obtained by Galan del Sastre (Figure 2.22 in [37]). Additionally, we note that the resultant large scale structures in Figure 6.16a are very similar to those in Figure 6.16b. It should be noted that some of the differences between our solution and the solution obtained by Galan del Sastre's and our solution and the observed large scale structures seen in Figure 6.16b are likely a result of his inclusion of more islands and not connected those islands to Europe or Africa. The rates of convergence are much lower than expected, however this likely is due to the fact that the domain is non-smooth which is likely causing the solution to be in a space which is less regular than the H^6 regularity [76] assumed by the error analysis in Section 5.1.

h	$DoFs$	e_0	L_2 order	e_1	H^1 order	e_2	H^2 order
$1/20$	1122	2.08×10^{-6}	—	1.95×10^{-4}	—	4.50×10^{-2}	—
$1/40$	4092	8.00×10^{-7}	1.38	6.68×10^{-5}	1.54	2.50×10^{-2}	0.850
$1/80$	15594	2.91×10^{-7}	1.46	2.47×10^{-5}	1.43	1.49×10^{-2}	0.741
$1/160$	60846	1.04×10^{-7}	1.49	9.05×10^{-6}	1.45	8.67×10^{-3}	0.785
$1/320$	240342	3.10×10^{-8}	1.75	2.75×10^{-6}	1.72	4.35×10^{-3}	0.994

Table 6.9: Observed rates of convergence for SQGE applied to the Mediterranean Sea with forcing function $F = \sin(\frac{\pi}{2}y)$ and true solution obtained from fine mesh with $h = \frac{1}{640}$.



(a) “True” solution of SQGE on the Mediterranean Sea with $Re = 5.27$, $Ro = 6.051 \times 10^{-4}$, $F = \sin(\frac{\pi}{2}y)$, and 955,302 DoFs corresponding to $h = \frac{1}{640}$.



(b) Large scale currents of the Mediterranean Sea [2].

Figure 6.16: Comparison of large scale structures visible in the numerical results of the SQGE applied to the Mediterranean Sea and the observed large scale structures of the Mediterranean Sea.

6.2 Two-Level SQGE

The goal of this section is two-fold: first, we illustrate the computational efficiency of the two-level method, and second, we verify the theoretical rates of convergence developed in Section 5.2. To illustrate the computational efficiency of the two-level method, we compare solution times for the full nonlinear one-level problem and for the two-level method applied to the SQGE. We choose coarse mesh/fine mesh pairs such that the ratio is $1/2$. To verify the theoretical rates of convergence, we compare the theoretical error estimates to the observed rates of convergence from our numerical tests. For the one-level problem we rely on our original code that was benchmarked in [34].

To this end we apply the two-level method to the SQGE (2.29) with $Re = Ro = 1$ and exact solution

$$\psi(x, y) = (\sin 4\pi x \cdot \sin 2\pi y)^2. \quad (6.13)$$

Additionally, the homogeneous boundary conditions are $\psi = \frac{\partial \psi}{\partial \mathbf{n}} = 0$ and the forcing function F corresponds to the exact solution (6.13). These boundary conditions and exact solution will be used in all the two-level tests that follow.

6.2.1 Practical Considerations

A key part of two-level algorithms is accessing a previous coarse mesh solution, i.e. finding the parent element given a child element. This step can negate any performance benefits if not implemented wisely. Indeed, let n be the number of elements in the FE discretization. For the unit square, a naïve search across every element takes $O(n/2)$ operations. This procedure may be improved with a binary search, which is summarized in Algorithm 2.

We note that every element on the fine mesh corresponds to exactly one element on the coarse mesh. However, a coarse mesh element may correspond to multiple elements on the fine mesh.

Algorithm 2 Given an element on the fine mesh determine the parent element on the coarse mesh.

Before examining the fine mesh, sort the coarse mesh elements by their centroid values.

Step 1: Select an element on the fine mesh and compute its centroid.

Step 2: Perform a binary search across the coarse mesh elements until the difference between the x -values of the fine mesh centroid and coarse mesh centroids is less than H , the coarse mesh step size. There should be many elements that fit this condition; save them as a list.

Step 3: Search through this list until we find the correct coarse mesh element (that is, the centroid of the fine-mesh element is an interior point of the correct coarse mesh element).

For the considered unit square, the binary search will examine on average $\log(n)$ elements, while the linear search component involves at most $\sqrt{n}/2$ elements. Therefore the search requires a $O(\sqrt{n}/2)$ number of element checks. Profiling results indicate that using Algorithm 2 to identify parent elements takes much less time than either setting up or solving the systems, so this approach is fast enough that lookup time does not contribute significantly to overall solution time.

H	h	DoFs, H	DoFs, h	e_{L^2}	e_{H^2}	time, s
—	0.05146	—	4362	4.29×10^{-8}	1.65×10^{-3}	3.328
0.1083	0.05146	1158	4362	1.09×10^{-7}	1.71×10^{-3}	2.372
—	0.02561	—	16926	5.75×10^{-10}	1.01×10^{-4}	19.92
0.05146	0.02561	4362	16926	7.69×10^{-10}	1.02×10^{-4}	11.82
—	0.01597	—	43074	4.75×10^{-11}	1.79×10^{-5}	55.69
0.03384	0.01597	10983	43074	5.27×10^{-11}	1.80×10^{-5}	33.19
—	0.01277	—	66678	8.66×10^{-12}	6.21×10^{-6}	102.4
0.02561	0.01277	16926	66678	9.69×10^{-12}	6.22×10^{-6}	59.03
—	0.009659	—	116614	3.88×10^{-12}	2.38×10^{-6}	161.7
0.02035	0.009659	29501	116614	6.84×10^{-12}	2.39×10^{-6}	95.93
—	0.007959	—	170598	4.79×10^{-12}	1.11×10^{-6}	325.1
0.01597	0.007959	43074	170598	9.09×10^{-12}	1.11×10^{-6}	172.3
—	0.006854	—	230574	1.79×10^{-11}	6.16×10^{-7}	401.7
0.01436	0.006854	58131	230574	1.30×10^{-11}	6.16×10^{-7}	219.5
—	0.006374	—	264678	3.91×10^{-11}	3.85×10^{-7}	559.7
0.01277	0.006374	66678	264678	2.31×10^{-11}	3.89×10^{-7}	291.9
—	0.005264	—	389994	3.85×10^{-11}	2.09×10^{-7}	753.4
0.01101	0.005264	98133	389994	6.50×10^{-11}	2.09×10^{-7}	397.7

Table 6.10: Comparison of one-level and two-level methods: the L^2 -norm of the error (e_{L^2}), the H^2 -norm of the error (e_{H^2}) and simulation times.

6.2.2 Computational Efficiency

To illustrate the computational efficiency of the two-level method, we compare the simulation time for the standard one-level method (i.e. the full nonlinear system, without the two-level method) with the simulation time for the two-level method.

In Table 6.10, the L^2 -norm of the error (e_{L^2}), the H^2 -norm of the error (e_{H^2}) and the simulation times are listed for various mesh sizes. For each fine mesh, we choose a coarse mesh that ensures the same order of magnitude for the errors in the one-level and two-level methods. For small values of the fine mesh size, h , the two-level method was significantly faster than the one-level method. The errors in the H^2 -norm were nearly identical, while the error in the L^2 -norm were generally of the same order of magnitude. We also note that the tolerance in Newton's method seems to cause a plateau in the L^2 -norm of the error. The results in Table 6.10 are illustrated graphically in Figure 6.17. In this figure the simulation times of the one-level method (green) and of the two-level method (blue) are displayed for all the pairs (h, H) in Table 6.10. Figure 6.17 clearly shows that as the number of degrees of freedom (DoFs) increases, the computational efficiency of the two-level method increases as well.

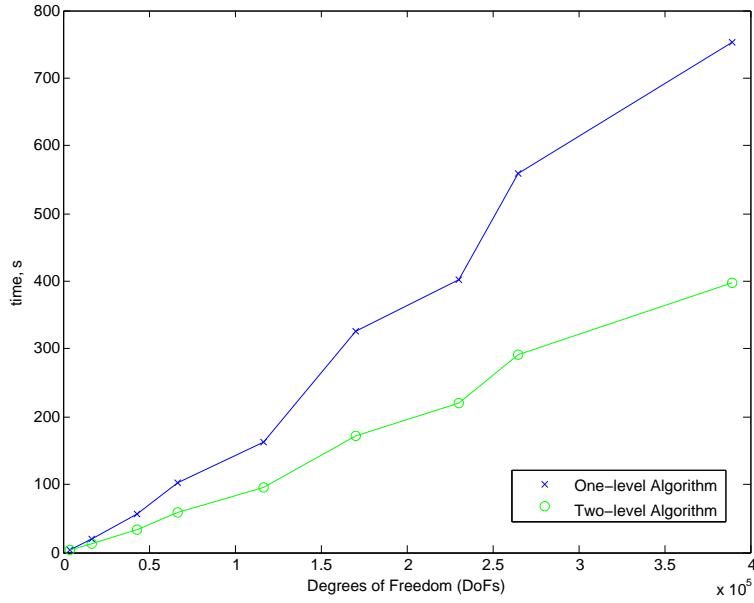


Figure 6.17: The simulation times of the one-level method (green) and of the two-level method (blue) are displayed for all the pairs (h, H) in Table 6.10.

H	h	DoFs, H	DoFs, h	e_{H^2}	H^2 order
0.43	0.0063	350	296710	4.54×10^{-1}	—
0.21	0.0063	1270	296710	1.75×10^{-2}	4.45
0.10	0.0063	4838	296710	4.92×10^{-4}	5.04
0.05	0.0063	18886	296710	1.32×10^{-5}	5.20
0.025	0.0063	74630	296710	6.02×10^{-7}	4.45

Table 6.11: Observed order of convergence in the H^2 norm for the coarse mesh, H , in the two-level method applied to (2.26) with exact solution (6.13).

6.2.3 Rates of Convergence

The goal of this subsection is to verify, numerically, the theoretical rates of convergence in (5.71) of Corollary 5.6. Unlike the theoretical error estimates for the one-level method developed in [34], for the two-level method we must verify rates of convergence for two different meshes: the fine mesh, h , and the coarse mesh, H .

To verify, numerically, the theoretical rate of convergence with respect to H , given in (5.71), we fix h to a small value and we vary H . Thus, the total error in (5.71) will be dominated by the H term, i.e. the total error will be of order $O(H^5)$. In Table 6.11, we fix $h = 0.0063$ and we vary H . The error in the L^2 -norm (e_{L^2}), the error in the H^2 -norm (e_{H^2}), and the rate of convergence with respect to H are listed in Table 6.11. The rate of convergence follows the theoretical rate predicted in (5.71) (i.e. fifth-order). For the last mesh pair, however, the rate of convergence appears to drop off. This occurs because, for small values of H , the total error in (5.71) is not dominated anymore by the H term.

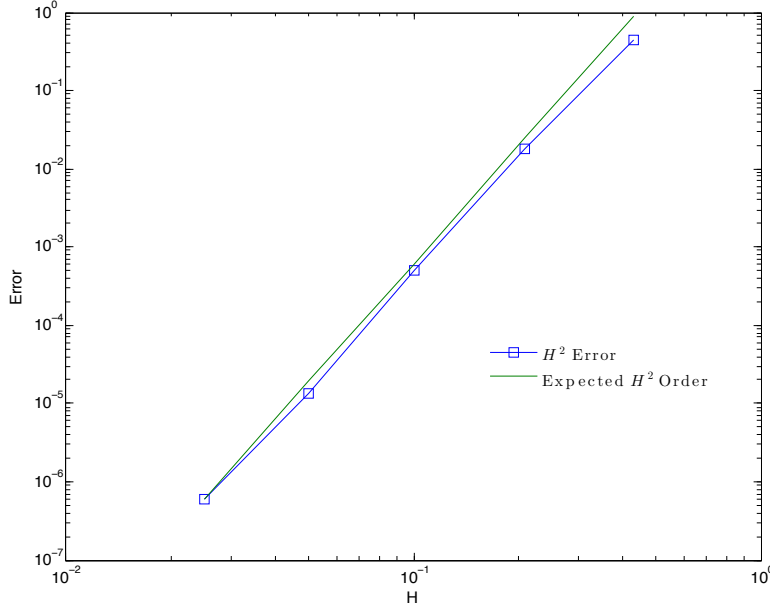


Figure 6.18: Observed order of convergence in the H^2 norm for the coarse mesh, H , in the two-Level method applied to (2.26) with the exact solution (6.13).

To verify, numerically, the theoretical rate of convergence with respect to h , given in (5.71), we must proceed with caution. The reason is that a straightforward approach would fix H and let h go to zero. This approach, however, would fail, since the H term would eventually dominate the total error. To avoid this, we consider the following scaling between the mesh sizes:

$$H = C h, \quad (6.14)$$

where $C > 1$. The scaling in (6.14) implies that the total error in (5.71) is of order $O(h^4)$. Indeed, the second term on the right hand side of (5.71) now scales as follows:

$$\begin{aligned} C_2 \sqrt{|\ln(h)|} H^5 &\approx C_2 C \sqrt{|\ln(h)|} h^5 \\ &\approx O(h^4), \end{aligned} \quad (6.15)$$

where in the last relation in (6.15) we used the fact that $\sqrt{|\ln(h)|} h \rightarrow 0$ when $h \rightarrow 0$ (which follows from l'Hospital's rule).

Remark 7. We emphasize that the scaling in (6.14) is not needed in the two-level algorithm. We only use it in this subsection to monitor the convergence rate with respect to h .

In this subsection, we consider $C = 2$ in (6.14). We note, however, that any other constant $C > 1$ could be used in (6.14). With this choice, we are now ready to verify, numerically, the theoretical rate of convergence with respect to h given in (5.71), which, as shown in (6.15), will be of order $O(h^4)$. In Table 6.12, for various mesh size pairs ($H = 2h, h$), we list the L^2 -norm of the error (e_{L^2}), the H^2 -norm of the error (e_{H^2}), and the rate of convergence. The rate of convergence follows the theoretical rate predicted in (5.71) (i.e. fourth-order).

H	h	DoFs, H	DoFs, h	e_{H^2}	H^2 order
1.1	0.43	38	106	5.83×10^0	—
0.43	0.21	106	350	6.61×10^{-1}	2.56
0.21	0.10	350	1270	3.65×10^{-2}	3.96
0.10	0.05	1270	4838	2.00×10^{-3}	4.10
0.050	0.025	4838	18886	1.20×10^{-4}	4.04
0.025	0.013	18886	74630	7.40×10^{-6}	4.02
0.013	0.0063	74630	296710	4.68×10^{-7}	3.99

Table 6.12: Observed order of convergence in the H^2 norm for the fine mesh, h , in the two-Level method applied to (2.26) with the exact solution (6.13).

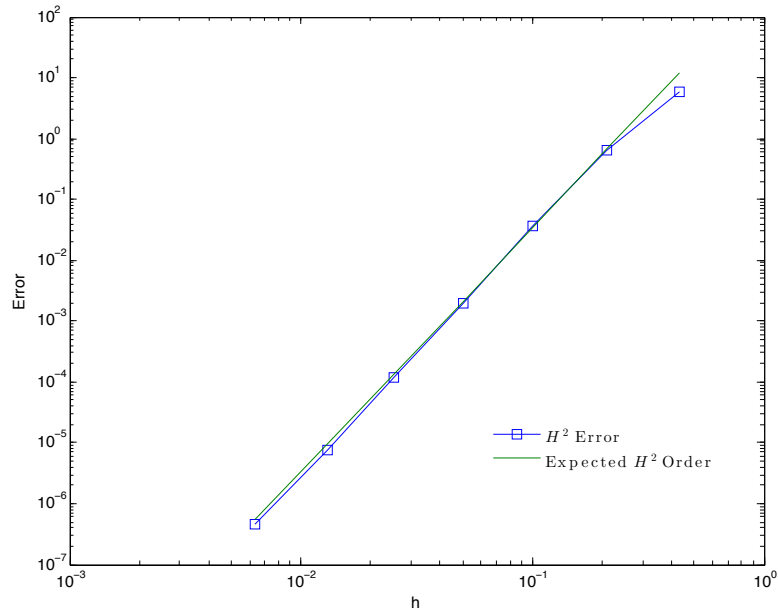


Figure 6.19: Observed order of convergence in the H^2 norm for the fine mesh, h , in the two-Level method applied to (2.26) with the exact solution (6.13).

k	h	DoFs	e_{L^2}	L^2 order
$1/2$	$1/16$	2853	5.65×10^{-4}	—
$1/4$	$1/16$	2853	3.49×10^{-4}	0.697
$1/8$	$1/16$	2853	1.916×10^{-4}	0.864
$1/16$	$1/16$	2853	1.001×10^{-4}	0.937
$1/32$	$1/16$	2853	5.11×10^{-5}	0.970
$1/64$	$1/16$	2853	2.58×10^{-5}	0.985
$1/128$	$1/16$	2853	1.30×10^{-5}	0.993
$1/256$	$1/16$	2853	6.51×10^{-6}	0.996
$1/512$	$1/16$	2853	3.26×10^{-6}	0.998
$1/1024$	$1/16$	2853	1.63×10^{-6}	0.999

Table 6.13: Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.16).

6.3 QGE

6.3.1 Rates of Convergence

In this subsection we verify the theoretical error estimates developed in Section 5.3. To this end, we apply the Argyris FE in space while applying implicit Euler in time to (2.26). To solve the resulting nonlinear system at each time step we apply Newton's method where the Newton's method is considered to have converged when the L^2 -norm of the difference in the current Newton iterate and the previous Newton iterate is less than 10^{-7} . Additionally, in each of the following computational tests we take $Re = 1$, $Ro = 1$, unless otherwise stated. In the tables that follow the variable k and h will refer to the time and spatial discretization, respectively. The tests which follow correspond to adding time dependence to the original tests used in our papers [34, 35] and the numerical tests for SQGE in both Section 6.1 and Section 6.2.

Test 1

For this test we take the exact solution to be

$$\psi(t; x, y) = (\sin \pi x \sin \pi y)^2 \sin t \quad (6.16)$$

which is very similar to **Test 6** in [34]. Here we add the time dependence in the sine term and let $\Omega = [0, 1]^2$. The time interval for integration is $t = [0, \frac{\pi}{2}]$. This test will have an intensifying western boundary layer as time increases.

Test 2

For this test we take the exact solution to be

$$\psi(t; x, y) = \left[\left(1 - \frac{x}{3}\right) \left(1 - e^{-20x}\right) \sin \pi y \right]^2 \sin t \quad (6.17)$$

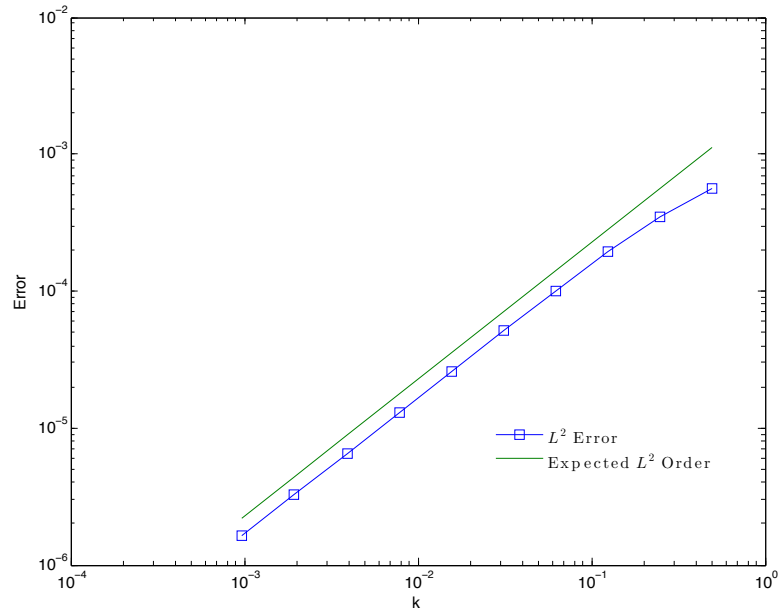


Figure 6.20: Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.16).

k	h	DoFs	e_{L^2}	L^2 order	e_{H^1}	H^1 order	e_{H^2}	H^2 order
$1/8192$	$1/2$	38	1.23×10^{-2}	—	1.18×10^{-1}	—	1.57×10^0	—
$1/8192$	$1/4$	174	2.12×10^{-5}	9.18	7.31×10^{-4}	7.34	2.79×10^{-2}	5.81
$1/8192$	$1/8$	662	7.88×10^{-7}	4.75	4.59×10^{-5}	3.99	3.04×10^{-3}	3.20
$1/8192$	$1/16$	2853	7.87×10^{-9}	6.65	9.05×10^{-7}	5.67	1.29×10^{-4}	4.56
$1/8192$	$1/32$	11690	6.97×10^{-11}	6.82	1.88×10^{-8}	5.59	5.98×10^{-6}	4.43
$1/8192$	$1/64$	47958	7.23×10^{-12}	3.27	5.26×10^{-10}	5.16	3.43×10^{-7}	4.12

Table 6.14: Observed spatial orders of convergence for Argyris applied to (2.26) with the exact solution (6.16) using implicit Euler for time discretization. Note the observed orders of convergence nearly matches the theoretical error estimates developed in Section 5.3. The L^2 order, however, drops off for the last spatial discretization due to nearing the machine epsilon.

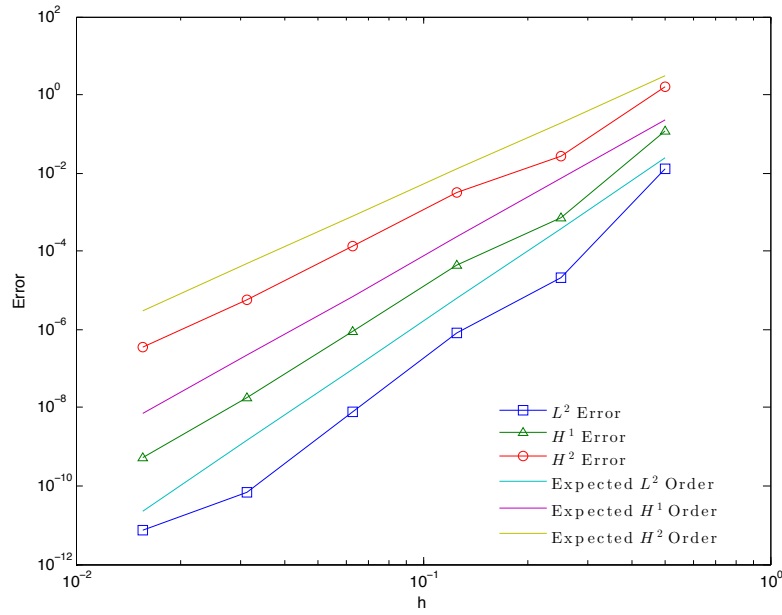


Figure 6.21: Observed orders of convergence in space for Argyris applied to (2.26) with the exact solution (6.16) using implicit Euler for the time discretization.

which is very similar to **Test 6** in [34]. Here we add the time dependence in the exponential term and let $\Omega = [0, 3] \times [0, 1]$. The time interval for integration is $t = [0, 0.5]$. This test will have an intensifying western boundary layer as time increases.

k	h	DoFs	e_{L^2}	L^2 order
$1/2$	$1/64$	47958	1.52×10^{-4}	—
$1/4$	$1/64$	47958	9.41×10^{-5}	0.695
$1/8$	$1/64$	47958	5.17×10^{-5}	0.865
$1/16$	$1/64$	47958	2.70×10^{-5}	0.937
$1/32$	$1/64$	47958	1.38×10^{-5}	0.970
$1/64$	$1/64$	47958	6.96×10^{-6}	0.985
$1/128$	$1/64$	47958	3.50×10^{-6}	0.993
$1/256$	$1/64$	47958	1.75×10^{-6}	0.996
$1/512$	$1/64$	47958	8.78×10^{-7}	0.998
$1/1024$	$1/64$	47958	4.40×10^{-7}	0.996

Table 6.15: Observed order of convergence for Implicit-Euler applied to (2.26) with the exact solution (6.17). Note the observed order of convergence matches the theoretical error estimates developed in Section 5.3.

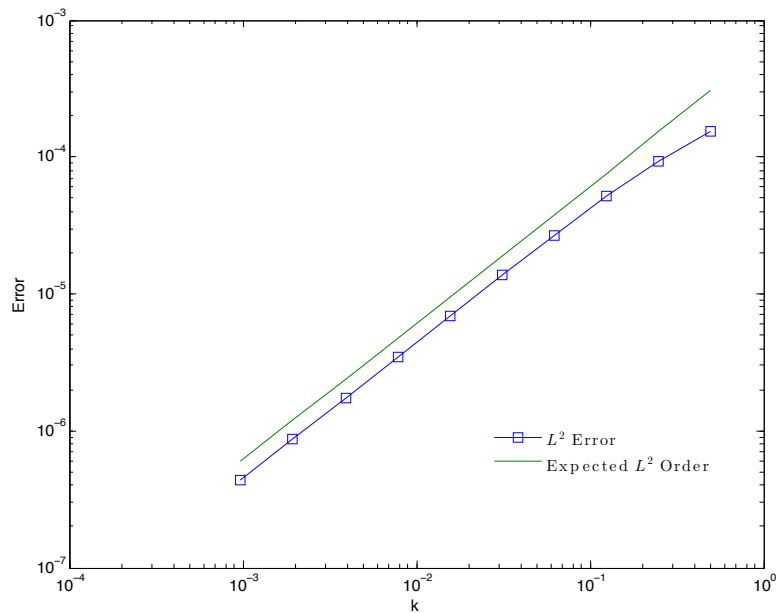


Figure 6.22: Observed orders of convergence in time for Implicit-Euler applied to (2.26) with the exact solution (6.17).

k	h	DoFs	e_{L^2}	L^2 order	e_{H^1}	H^1 order	e_{H^2}	H^2 order
$1/8192$	$1/2$	38	2.86×10^{-2}	—	5.16×10^{-1}	—	1.82×10^1	—
$1/8192$	$1/4$	174	4.79×10^{-3}	2.58	1.75×10^{-1}	1.56	9.28×10^0	0.973
$1/8192$	$1/8$	662	5.04×10^{-4}	3.25	3.38×10^{-2}	2.37	2.96×10^0	1.65
$1/8192$	$1/16$	2853	1.65×10^{-5}	4.94	2.17×10^{-3}	3.96	3.67×10^{-1}	3.01
$1/8192$	$1/32$	11690	4.17×10^{-7}	5.30	1.07×10^{-4}	4.34	3.47×10^{-2}	3.40
$1/8192$	$1/64$	47958	7.28×10^{-9}	5.84	3.70×10^{-6}	4.86	2.37×10^{-3}	3.87

Table 6.16: Observed order of convergence for Argyris applied to (2.26) with the exact solution (6.17). Note the observed order of convergence matches the theoretical error estimates developed in Section 5.3.

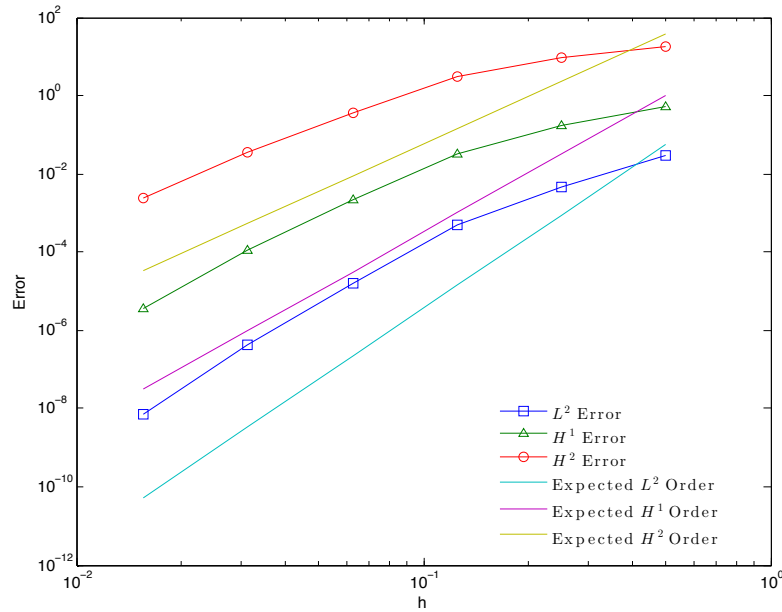


Figure 6.23: Observed orders of convergence in space for Argyris applied to (2.26) with the exact solution (6.17) using implicit Euler for the time discretization.

Chapter 7

Conclusions

In this thesis we studied the FE discretization of the pure streamfunction form of the QGE. The pure streamfunction form of the QGE requires C^1 FEs to be conforming, thus the implementation of the of the C^1 FE was discussed in length. In particular, the implementation of the Argyris element was discussed in Section 4.3, including an in depth discussion of the novel transformation developed by Dominguez et. el. [27] in Subsection 4.3.2. Additionally, we proposed a two-level FE discretization for the SQGE in Section 4.4 based on the conforming FE discretization used for the SQGE in Section 4.2. The two-level algorithm consists of two steps. In the first step, the nonlinear system is solved on a coarse mesh. In the second step, the nonlinear system is linearized around the approximation found in the first step, and the resulting linear system is solved on the fine mesh. This method significantly improved simulation times for the SQGE over that of the one-level method. Next, a conforming FE semi-discretization was developed for the time-dependent QGE based upon the FE discretization developed in Section 4.2.

For the conforming FE discretization of SQGE, in Section 5.1 we proved, for the first time, optimal error estimates in the H^2 , H^1 , and L^2 norms. The error analysis only relied on the FE discretization being conforming and was not specific to any particular element. Specifically, for the Argyris element we showed that the order of convergence in the H^2 , H^1 , L^2 norms were $O(h^4)$, $O(h^5)$, and $O(h^6)$, respectively. The error analysis in Section 5.1 was then used to develop new rigorous error estimates for the two-level FE discretization. These estimates are optimal in the following sense: for an appropriately chosen scaling between the coarse mesh, H , and the fine mesh, h , the error in the two-level method is of the same order as the error in the standard one-level method (i.e. solving the nonlinear system directly on the fine mesh). For the semi-discretization of the QGE we proved similar optimal error estimates, i.e. the order of convergence for the energy norm is of $O(h^4)$.

Numerical experiments for the SQGE (Section 6.1), two-level algorithm (Section 6.2), and QGE (Section 6.3) with the Argyris element, were also carried out. For the SQGE a few benchmark problems were used to verify our code. These benchmark problems included the Linear Stommel model [81, 61], and the Linear Stommel-Munk model [15]. With the agreement of our solutions to that of the literature our code was thus verified. We then verified numerically the theoretical error estimates developed in Section 6.1 and performed numerical experiments on non-rectangular domains, including a FE discretization of the SQGE applied to the Mediterranean. The test results for the Mediterranean appeared to be in agreement with observed surface currents and the numerical experiments developed by Galan del Sastre [37]. The code which was verified in Section 6.1 was then modified to allow for the use of the two-level method and numerical experiments were then carried out for the two-level method. These numerical experiments verified numerically the theoretical

error estimates developed in Section 5.2, both with respect to the coarse mesh size, H , and the fine mesh size, h . Furthermore, the numerical results showed that, for an appropriate scaling between the coarse and fine meshes, the two-level method significantly decreases the computational time of the standard one-level method. Next, the same code which was developed and verified for the SQGE was then modified to deal with time-dependence. We applied a implicit Euler scheme and verified numerically the theoretical spatial rates of convergence developed for the semi-discretization. Additionally, we observed expected rates of convergence in time for the implicit Euler scheme ($O(h)$).

The QGE has many unique challenges for numerical modelling. These challenges include, but are not limited to, unstable solutions, resulting from internal layers and western boundary layers, and high computational cost for large domains, such as the North Atlantic. To address these issues we plan to extend these studies in several directions, including stabilization methods such as *Petrov-Galerkin* stabilization, *adaptive mesh refinement*, model reduction using *Proper Orthogonal Decomposition*, and the incorporation of observed windstress data, which will include *parameter estimation*. Additionally, the streamfunction formulation suffers from non-unique streamfunctions for multiply connected domains, such as the ocean, seas, and large lakes with islands. To address this issue a method such as the one developed by van Gijzen et. al. [82] will be explored. Thus, a more realistic ocean mesh will be used and islands will no longer have to be connected to continents.

Bibliography

- [1] J. H. Argyris, I. Fried, and D. W. Scharpf. The TUBA family of plate elements for the matrix displacement method. *Aero. J.*, 72:701–709, 1968.
- [2] N. Ayoub, P.-Y. Le Traon, and P. De Mey. A description of the Mediterranean surface variable circulation from combined ERS-1 and TOPEX/POSEIDON altimetric data. *J. of Marine Sys.*, 18(1):3–40, 1998.
- [3] I. Babuška. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20(3):179–192, 1973.
- [4] H. Barbosa and T. Hughes. The finite element method with Lagrange multipliers on the boundary: circumventing the Babuška-Brezzi condition. *Comp. Meth. in Appl. Mech. and Eng.*, 85(1):109–128, 1991.
- [5] H. Barbosa and T. Hughes. Boundary Lagrange multipliers in finite element methods: error analysis in natural norms. *Numer. Math.*, 62(1):1–15, 1992.
- [6] V. Barcilon, P. Constantin, and E. S. Titi. Existence of solutions to the Stommel-Charney model of approximate deconvolution models of turbulence: Analysis, phenomenology and numerical analysis the Gulf Stream. *SIAM J. Math. Anal.*, 19(6):1355–1364, 1988.
- [7] D. Beesley, J. Olejarz, A. Tandon, and J. Marshall. A laboratory demonstration of Coriolis effects on wind-driven ocean currents. *Oceanography*, 21(2):72, 2008.
- [8] R. Bermejo and P. Galán del Sastre. Long-term behavior of the wind stress circulation of a numerical North Atlantic ocean circulation model. *ECCOMAS*, 2004.
- [9] M. Bernadou. Straight and curved finite elements of class C^1 and some applications to thin shell problems. In *Finite element methods (Jyväskylä, 1993)*, volume 164 of *Lecture Notes in Pure and Appl. Math.*, pages 63–77. Dekker, New York, 1994.
- [10] H. Blum, R. Rannacher, and R. Leis. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Meth. in the Appl. Sci.*, 2(4):556–581, 1980.
- [11] D. Braess. *Finite elements: Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, 2001.
- [12] J. H. Bramble. The Lagrange multiplier method for dirichlet’s problem. *Math. of Comp.*, 37(155):1–11, 1981.
- [13] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, third edition, 2008.

- [14] K. Bryan. A numerical investigation of a nonlinear model of wind-drive ocean. *J. of the Atmos. Sci.*, 20:594–606, 1963.
- [15] J. M. Cascon, G. C. Garcia, and R. Rodriguez. A priori and a posteriori error analysis for a large-scale ocean circulation finite element model. *Comp. Meth. Appl. Mech. Eng.*, 192(51-52):5305–5327, 2003.
- [16] M. E. Cayco and R. A. Nicolaides. Finite element technique for optimal pressure recovery from stream function formulation of viscous flows. *Math. of Comp.*, 46(174), 1986.
- [17] NOAA National Geophysical Data Center. Coastline extracted GSHHG .
<http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html>, January 2013.
- [18] J. G. Charney. On the scale of atmospheric motions. *Geofysiske Publikasjoner*, 17(2), 1948.
- [19] J. G. Charney. On a physical basis for numerical prediction of large-scale motions in the atmosphere. *J. of Meteorology*, 6(6):371–385, December 1949.
- [20] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, 1978.
- [21] P. F. Cummins. Inertial gyres in decaying and forced geostrophic turbulence. *J. Mar. Res.*, 50(4):545–566, 1992.
- [22] B. Cushman-Roisin. *Introduction to geophysical fluid dynamics*. Prentice Hall, Englewood Cliffs, New Jersey, 1994.
- [23] B. Cushman-Roisin and J. M. Beckers. *Introduction to geophysical fluid dynamics: Physical and numerical aspects*. International Geophysics. Elsevier Science & Technology, 2011.
- [24] H. E. Dijkstra. *Nonlinear physical oceanography: A dynamical systems approach to the large scale ocean circulation and el Nino*, volume 28. Springer Verlag, 2005.
- [25] H. E. Dijkstra. *Dynamical Oceanography*. Springer Verlag, 2008.
- [26] V. Dominguez and F. J. Sayas. A simple Matlab implementation of Argyris element. Technical Report 25, Universidad de Zaragoza, 2006.
- [27] V. Dominguez and F. J. Sayas. Algorithm 884: A simple Matlab implementation of the Argyris element. *ACM Transaction on Mathematical Software*, 35(2), 2008.
- [28] V. W. Ekman. On the influence of the earth’s rotation on ocean-currents. *Arkiv för matematik, astronomi och fysik*, 2(11), 1905.
- [29] L. C. Evans. *Partial Differential Equations*, volume 19. American Mathematical Society, Providence, 1989.
- [30] F. Fairag. A two-level finite-element discretization of the stream function form of the Navier-Stokes equations. *Comp. Math. Applic.*, 36(2):117–127, 1998.
- [31] F. Fairag. Numerical computations of viscous, incompressible flow problems using a two-level finite element method. *SIAM J. Sci. Comp.*, 24(6):1919–1929, 2003.
- [32] G. J. Fix. Finite element models for ocean circulation problems. *SIAM J. on Appl. Math.*, 29(3):371–387, 1975.

- [33] G. J. Fix, M. Gunzburger, R. Nicolaides, and J. Peterson. Mixed finite element approximations for the biharmonic equations. *Proc. 5th Internat. Sympos. on Finite Elements and Flow Problems (JT Oden, ed.)*, University of Texas, Austin, pages 281–286, 1984.
- [34] E. L. Foster, T. Iliescu, and Z. Wang. A finite element discretization of the streamfunction formulation of the stationary quasi-geostrophic equations of the ocean. *Comp. Meth. in Appl. Mech. and Eng.*, 261-262:105–117, 2013.
- [35] E. L. Foster, T. Iliescu, and D. Wells. A two-level finite element discretization of the streamfunction formulation of the stationary quasi-geostrophic equations of the ocean. *Comp. & Math. with Applic.*, 2013. In Revision.
- [36] B. Fox-Kemper. *Eddies and Friction: Removal of Vorticity from the Wind-Driven Gyre*. PhD thesis, MIT, June 2003.
- [37] P. Galán del Sastre. *Estudio Numérico del Atractor en Ecuaciones de Navier-Stokes Aplicadas a Modelos de Circulación del Océano*. PhD thesis, Universidad Complutense de Madrid, Madrid, 2004.
- [38] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *Inter. J. for Num. Meth. in Eng.*, 79(11):1309–1331, 2009.
- [39] M. Ghil, M. D. Chekroun, and E. Simonnet. Climate dynamics and fluid mechanics: Natural variability and related uncertainties. *Physica D: Nonlinear Phenomena*, 237:2111–2126, 2008.
- [40] V. Girault and P. A. Raviart. *Finite element approximation of the Navier-Stokes equations*. Volume 749 of Lecture Notes in Mathematics. Springer-Verlag, 1979.
- [41] V. Girault and P. A. Raviart. *Finite element methods for Navier-Stokes equations: theory and algorithms*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1986.
- [42] R. J. Greatbatch and B. T. Nadiga. Four-gyre circulation in a barotropic model with double-gyre wind forcing. *J. Phys. Oceanogr.*, 30:1461–1471, 200.
- [43] M. D. Gunzburger. *Finite element methods for viscous incompressible flows*. Computer Science and Scientific Computing. Academic Press Inc, 1989. A Guide to Theory, Practice, and Algorithms.
- [44] D. B. Haidvogel and A. Beckmann. *Numerical Ocean Circulation Modelling*. World Scientific Publishing Company, 4 edition, 1999.
- [45] S. Hellerman and M. Rosenstein. Normal monthly wind stress over the world ocean with error estimates. *J. Phys. Ocean.*, 13:1093–1104, 1983.
- [46] H. Ingstad and A. S. Instad. *The Viking discovery of America: The excavation of a Norse settlement in L’Anse aux Meadows, Newfoundland*. Breakwater Books, 2000.
- [47] V. M. Ipatova, V. I. Agoshkov, G. M. Kobelkov, and V. B. Zalesny. Theory of solvability of boundary value problems and data assimilation problems for ocean dynamics equations. *Russ. J. Numer. Anal. Math. Modelling*, 25(6):511–534, 2010.

- [48] I. N. James. Charney, J.G. 1947: The dynamics of long waves in a baroclinic westerly current. *Journal of Meteorology* 4, 135–62. *Progress in Physical Geography*, 33(5):718–723, August 2009.
- [49] C. Johnson. *Numerical solution of partial differential equations by the finite element method*, volume 32. Cambridge University Press, New York, 1987.
- [50] E. Kreyszig. *Introductory functional analysis with applications*. Wiley, New York, wiley classic library edition edition, 1989.
- [51] W. Layton. A two-level discretization method for the Navier-Stokes equations. *Comp. Math. Applic.*, 26(2):33–38, 1993.
- [52] W. J. Layton. *Introduction to the numerical analysis of incompressible viscous flows*, volume 6. Society for Industrial and Applied Mathematics (SIAM), 2008.
- [53] C. LeProvost, C. Bernier, and E. Blayo. A comparison of two numerical methods for integrating a quasi-geostrophic multilayer model of ocean circulations: finite element and finite difference methods. *J. Comput. Phys.*, 110:341–359, 1994.
- [54] A. J. Majda. *Introduction to PDEs and waves for the atmosphere and ocean*. AMS, New York, 2003.
- [55] A. J. Majda and X. Wang. *Non-linear dynamics and statistical theories for basic geophysical flows*. Cambridge University Press, 2006.
- [56] J. McWilliams. *Fundamentals of geophysical fluid dynamics*. Cambridge Univ Pr, 2006.
- [57] T. T. Medjo. Mixed formulation of the two-layer quasi-geostrophic equations of the ocean. *SIAM J. on Appl. Math.*, 15(4):489–502, 1999.
- [58] T. T. Medjo. Numerical simulations of a two-layer quasi-geostrophic equation of the ocean. *SIAM J. on Appl. Math.*, 37(6):2005–2022, 2000.
- [59] N. A. Mörner. Earth rotation, ocean circulation and paleoclimate. *GeoJournal*, pages 419–430, 1995.
- [60] W. H. Munk. On the wind-driven ocean circulation. *J. of Meteorology*, 7(2):79–93, April 1950.
- [61] P. G. Myers and A. J. Weaver. A diagnostic barotropic finite-element ocean circulation model. *J. Atmos. Oceanic Technol.*, 12:511, 1995.
- [62] F. Nansen. *Farthest North*. Harper & Borthers, New York, 1897.
- [63] M. Okabe. Full-explicit interpolation formulas for the Argyris triangle. *Comp. Meth. in Appl. Mech. and Eng.*, 106:381–394, 1993.
- [64] J. Pedlosky. *Geophysical fluid dynamics*. Springer, second edition, 1992.
- [65] J. Pitkäranta. Local stability conditions for the Babuska method of Lagrange multipliers. *Math. of Comp.*, 35(152):1113–1129, 1980.
- [66] J. Pitkäranta. The finite element method with Lagrange multipliers for domains with corners. *Mathematics of Computation*, 37(155):13–30, 1981.

- [67] J. F. Price, R. A. Weller, and R. R. Schudlich. Wind-driven ocean currents Ekman transport. *Science*, 238(4833):1534–1538, 1987.
- [68] Y. Renard and J. Pommier. Appendix a. finite element method list — GetFEM++. <http://download.gna.org/getfem/html/homepage/userdoc/appendixA.html#argyris-element>, March 2013.
- [69] J. Roach. Artic melt opens Northwest Passage. <http://news.nationalgeographic.com/news/2007/09/070917-northwest-passage.html>, September 2007.
- [70] C.-G. Rossby. Dynamics of steady ocean currents in the light of experimental fluid mechanics. *Pap. phys. Ocean. Meteor.*, 5(1), 1936.
- [71] H. Royden and P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010.
- [72] O. San, A. E. Staples, Z. Wang, and T. Iliescu. Approximate deconvolution large eddy simulation of a barotropic ocean circulation model. *Ocean Model.*, 2011. Submitted.
- [73] O. San, A. E. Staples, Z. Wang, and T. Iliescu. Approximate deconvolution large eddy simulation of a stratified two-layer quasigeostrophic ocean model. *Ocean Model.*, 2011. Submitted.
- [74] X. Shao and D. Han. A two-grid algorithm based on Newton iteration for the stream function form of the Navier-Stokes equations. *Appl. Math. J. Chinese Univ.*, 26(3):368–378, 2011.
- [75] G. Siedler, J. Church, and J. Gould. *Ocean circulation and climate: Observing and modelling the global ocean*, volume 77. Academic Press, 2001.
- [76] A. M. Soane, M. Sun, and R. Rostamian. The optimal convergence rate of a C^1 finite element method for non-smooth domains. *J. of Comp. and Appl. Math.*, 233(10):2711–2723, March 2010.
- [77] W. N. R. Stevens. Finite element, stream function-vorticity solution of steady laminar natural convection. *Int. J. Num. Meth. Fluids*, 2:381–394, 1982.
- [78] H. Stommel. The westward intensification of wind-driven ocean currents. 29(2):202–206, 1948.
- [79] H. U. Sverdrup. Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the Eastern Pacific. *Geophysics*, 33(11):318–326, 1947.
- [80] K. E. Trenberth and J. M. Caron. Estimates of meridional atmosphere and ocean heat transports. *J. of Climate*, 14(16):3433–3443, 2001.
- [81] G. K. Vallis. *Atmosphere and ocean fluid dynamics: Fundamentals and large-scale circulation*. Cambridge University Press, 2006.
- [82] M. B. van Gijzen, C. B. Vreugdenhil, and H. Oksuzoglu. The finite element discretization for stream-function problems on multiply connected domains. *J. of Comp. Phys.*, (140):30–46, 1998.
- [83] J. Wang and G. K. Vallis. Emergence of Fofonoff states in inviscid and viscous ocean circulation models. *J. Mar. Res.*, 52:83–127, 1994.
- [84] Wikipedia. Fonctions d’Argyris. http://fr.wikipedia.org/wiki/Fonctions_d’Argyris, March 2011.

- [85] M. Winton. On the climatic impact of ocean circulation. *J. of Climate*, 16(17):2875–2889, 2003.
- [86] G. Wolansky. Existence, uniqueness and stability of stationary barotropic flow with forcing and dissipation. *Comm. Pure Appl. Math.*, 41:19–46, 1988.
- [87] J. Xu. A novel two-grid method for semilinear elliptic equations. *SIAM J. on Sci. Comp.*, 15(1):231–237, 1994.