INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeib Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
RICE UNIVERSITY

A Priori Error Estimates of Finite Element Models of Systems of Shallow Water Equations

by

Mónica L. Martínez

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Mary F. Wheeler, Co-Chair
Ernest and Virginia Cockrell Chair in Aerospace Engineering and Engineering Mechanics
The University of Texas, Austin, TX

Clint N. Dawson, Co-Chair
Associate Professor of Aerospace Engineering and Engineering Mechanics
The University of Texas, Austin, TX

Richard A. Tapia, Co-Chair
Noah Harding Professor of Computational and Applied Mathematics

John Edward Akin
Professor of Mechanical Engineering and Computational and Applied Mathematics

Houston, TX
April, 1998
Abstract

*A Priori* Error Estimates of Finite Element Models of Systems of Shallow Water Equations

by

Mónica L. Martínez

In recent years, there has been much interest in the numerical solution of shallow water equations. The numerical procedure used to solve the shallow water equations must resolve the physics of the problem without introducing spurious oscillations or excessive numerical diffusion. Staggered-grid finite difference methods have been used extensively in modeling surface flow without introducing spurious oscillations. Finite element methods, permitting a high degree of grid flexibility for complex geometries and facilitating grid refinement near land boundaries to resolve important processes, have become much more prevalent. However, early finite element simulations of shallow water systems were plagued by spurious oscillations and the various methods introduced to eliminate these oscillations through artificial diffusion were generally unsuccessful due to excessive damping of physical components of the solution.

Here, we give a brief overview on some finite element models of the shallow water equations, with particular attention given to the *wave* and *characteristic* formulations. In the literature, standard analysis, based on Fourier decompositions of these methods, has always neglected contributions from the nonlinear terms.
We derive $L^\infty ((0, T); L^2(\Omega))$ and $L^2 ((0, T); H^1(\Omega))$ a priori error estimates for both the continuous-time and discrete-time Galerkin approximation to the nonlinear wave model, finding these to be optimal in $H^1(\Omega)$.

Finally, we derive $L^\infty ((0, T); L^2(\Omega))$ and $L^2 ((0, T); H^1(\Omega))$ a priori error estimates for our proposed Characteristic-Galerkin approximation to the nonlinear primitive model. We find these estimates to be optimal in $H^1(\Omega)$ but with less restrictive time-step constraints when compared to the Galerkin estimates for the wave model.
To my parents

Francisco Martínez Domínguez
María Remedios Martínez Llaca
Acknowledgments

I would like to express my sincere thanks and deepest gratitude to those persons integral in helping me complete the Ph.D. program.

First, I am grateful to my committee for their support and for agreeing to listen to my defense during the summer.

My thesis advisors Clint N. Dawson and Mary F. Wheeler have supported and encouraged me throughout this long process. Their professionalism, scholarship, and leadership have become paradigms to follow. And their extensive scientific knowledge has set new standards for me to learn as much. I thank them for welcoming me into the Center for Subsurface Modeling family.

I am also indebted to Richard A. Tapia for his mentorship and friendship. From my first meeting with him in San Antonio at a SACNAS conference, I have benefited from his insights and commentary into teaching, research and outreach. His serendipitous use of “kick in the butt” and “pat on the back” have helped me appreciate the highs and survive the lows associated with graduate school. He and his family have also become a second family to me, as we watched Rockets basketball games or educational videos, as we made cookies, practiced our merengue dancing at fiestas, or enjoyed some good food (thanks to Jean Tapia).

Srinivas Chippada has taught me basic engineering axioms, suffered through countless practice talks, and patiently listened to my explanations of error estimates and their proofs.

My seventh-grade algebra teacher Mr. Martini and my undergraduate-senior mathematics professor H. Royden, both now in heaven, nurtured my love for mathematics.
And finally, Miguel Canales Jr., my partner in all things non-mathematical, never stopped believing in me when I lost belief in myself, always challenged me to live up to my potential when I thought there was no more to give, was always my friend when I needed him to be so, always provided inspiration when I could find none, always made me laugh amidst frustration, always held my hand when it needed holding, was always with me in spirit even though this part of our life journeys forced us a thousand miles apart, and continues to love me even though I’m a nerdy mathematician.
Special Acknowledgments

As in all graduate experiences, mine was interlaced with many peaks and valleys. As we are more than the sum of our parts, which includes the influence of those touching our lives, I would like to acknowledge the following people who influenced my graduate career in profound ways.

The Rice CAAM administrative staff Daria Lawrence, Fran Moshiri, and Linda Neyra, who are competent beyond words and friends to all those in need, served as inspirational models of self-determination in their pursuit of their own academic degrees despite the many obstacles that lay before them.

Rice graduate student peers Martha M. Carey, Pamela J. Williams, Cassandra Moore McZeeal, Donald Williams, Leticia Velazquez, Miguel Argaez, Cristina Villalobos, and John Rodriguez made my days a little bit more interesting and a little bit more sane. Rice CAAM professors Amr El-Bakry, Mahmoud El-Alem, and Danny Sorensen, were all supportive and encouraging.

Rice CAAM Systems Administrator Michael Pearlmam came to my rescue everytime I accidentally killed my computer. CRPC Manager of Outreach Programs Theresa Chatman, Exxon’s Ruth Gonzalez and College of William and Mary’s Virginia Torczon have all been counsel and friend.

Sandia National Laboratories’ Juan C. Meza helped me weather the sometimes-difficult phases of the graduate experience. I am also indebted to him for his mentorship and friendship that began when I first came to him as a summer intern.

The Texas Institute for Computational and Applied Mathematics (TICAM) provided extraordinary support to those of us Rice students who made the move to Austin in 1995 with the Subsurface Modeling Group. In particular, I thank TICAM administrative assistants Connie Baxter, Sheri Brice and Ruth Hengst, TICAM
systems administrators Larry Thorne and Chris McCraw, CAM Professors Graham Carey, Leszek Demkowicz, and finally, J. Tinsley Oden, Director of TICAM, without whose blessing, my time at UT would not have been memorable, let alone possible.

Finally, I am indebted to the SACNAS organizations for being the vehicle, through its conferences and mentorship programs, by which I have met role-models that looked like me and allowed me to be a role-model in turn.
# Contents

Abstract  
Acknowledgments  
List of Illustrations

## 1 Introduction

## 2 Shallow Water Equations

2.1. Primitive Formulation, P-SWE.  
2.1.1. An Overview of the FEM for the P-SWE.  
2.2. Wave Formulation, W-SWE.  
2.2.1. Harmonic Wave SWE.  
2.2.2. Historical Development of the W-SWE.  
2.3. Characteristic Formulation, C-SWE.  
2.3.1. The Method of Characteristics.  
2.3.2. Historical Development of the C-SWE.  
2.4. Well-Posedness and Boundary Conditions.

## 3 Notation, Definitions and Mathematical Notions

3.1. Mathematical Inequalities.  
3.2. A Priori Error Estimate Tools.  
3.3. Approximation Theory Notions.  
3.3.1. The $L^2$ Projection.  
3.3.2. One Final Comment on Notation
4 Finite Element Model of W-SWE.

4.1. Equivalence of Solutions. ........................................ 30

4.2. Continuous-Time Nonlinear Model. .................................. 31
   4.2.1. Continuous-Time Variational Formulation. ...................... 31
   4.2.2. The Continuous-Time Galerkin Approximation. .................. 33
   4.2.3. Error Equations. ............................................. 34
   4.2.4. Choice of Projections. ....................................... 35
   4.2.5. Choice of Test Functions. .................................... 35
   4.2.6. Bounding the GWCE and CME Error Equation. ................... 37
   4.2.7. A Priori Error Estimate. .................................... 46

4.3. Discrete-Time Nonlinear Model. ................................... 49
   4.3.1. ADCIRC Temporal Discretization. .............................. 50
   4.3.2. The Discrete-Time Galerkin Approximation. ..................... 52
   4.3.3. Error Equations. ............................................. 53
   4.3.4. Choice of Test Functions. .................................... 56
   4.3.5. Bounding the first time-step of the GWCE Error Equation.... 56
   4.3.6. Bounding the GWCE-Error Equations. ........................... 62
   4.3.7. Bounding the CME-Error Equation. ................................ 67
   4.3.8. Bounding the Sum of the Error Equations. ..................... 69
   4.3.9. A Priori Error Estimate ..................................... 73

4.4. Recent Applications of Error Estimates. .......................... 75

5 Finite Element Model of C-SWE ........................................... 76

5.1. Weak Formulation. ............................................... 77

5.2. Characteristic-Galerkin. ......................................... 78
   5.2.1. CE-NCME Error Equations and Choice of Projections. ........ 79
5.2.2. Properties of Characteristic
5.2.3. Bounding the CE Error Equation
5.2.4. Bounding the NCME Error Equation
5.2.5. Bounding the Sum of the Error Equations

6 Extensions and Future Directions

Bibliography
Illustrations

1.1 Computational Grid: Galveston Bay .......................... 3
1.2 Computational Grid: Gulf of Mexico ........................... 4
1.3 Computational Grid: West Atlantic Ocean and Gulf of Mexico .... 5

2.1 Underwater topography, vertical scale and horizontal scale. ........ 8
2.2 Vertical cross-section depicting elevation and bathymetry. ........ 10
2.3 Land and River/Sea Boundaries .................................. 21
Chapter 1

Introduction

In recent years, there has been much interest in the numerical solution of shallow water equations. Simulation of shallow water systems can serve numerous purposes. First, it can serve as a means for modeling tidal fluctuations for those interested in capturing tidal energy for commercial purposes or safe-guarding against potentially damaging tidal forces acting on stilted objects, like oil-well platforms in the far ocean or piers near-shore. Second, these simulations can be used to determine long-range circulation patterns in environmentally-sensitive water bodies. Third, these simulations can be used to compute tidal ranges and surges caused by extreme geologic and atmospheric events such as tsunamis and hurricanes. This information can be used in the development planning of coastal areas. Finally, the shallow water hydrodynamic model can be coupled to a transport model in considering flow and transport phenomena, thus making it possible to study remediation options for polluted bays and estuaries, to predict the impact of commercial projects on fisheries, to model freshwater-saltwater interactions, and to study allocation of allowable discharges by municipalities and by industries in meeting water quality controls.

The numerical procedure used to solve the shallow water equations must resolve the physics of the problem without introducing spurious oscillations or too much numerical diffusion. Staggered-grid finite difference methods have been used extensively in modeling surface flow without introducing spurious oscillations. Examples of hydrodynamic simulators based on finite-difference models include the 3-dimensional Coastal Hydraulics and Sediment Transport (CH3D-SED) simulator used by the U.S. Army Corps of Engineers’ Waterways Experiment Station (CE-WES) and the Computer
Hardware. Advanced Mathematics and Model Physics (CHAMMP) global ocean model used by the Department of Energy (DOE).

Finite element methods, permitting a high degree of grid flexibility for complex geometries and facilitating grid refinement near land boundaries to resolve important processes, have become much more prevalent. Some sample unstructured triangular finite element grids of water bodies are depicted in Figures 1.1 - 1.3. Notice the refinement along land boundaries and the coarser mesh along ocean boundaries. The grid in Figure 1.1 is courtesy of J. Matsumoto at the Texas Water Development Board (TWDB). The grids in Figures 1.2-1.3 are courtesy of J. J. Westerink at the University of Notre Dame.

Early finite element simulations of shallow water systems were plagued by spurious oscillations and the various methods introduced to eliminate these oscillations through artificial diffusion were generally unsuccessful due to excessive damping of physical components of the solution.

In the late 1970s and early 1980s, the governing equations that constituted the shallow water equations were reformulated into what is is now called the wave formulation. This alternative formulation was proven to eliminate spurious oscillations while matching tidal data accurately. A hydrodynamic simulator Advanced Circulation Model (ADCIRC), based on this wave formulation is currently being used by the TWDB, the CE-WES, and the U.S. Navy. Still, this wave formulation also has some major drawbacks. First, the wave formulation contains a tuning parameter that is difficult to adjust in practice. Second, the wave formulation doesn't handle very well advective flow that dominates in long narrow rivers or channels. For example, such flow exists in the Houston Ship Channel seen in Figure 1.1

In the mid-1980s and 1990s, the governing equations were again reformulated from the viewpoint of characteristic methods. These formulations have been shown
to handle advective flow while permitting larger time-steps to be taken. The split-characteristic finite element model of Zienkiewicz and Ortiz[64, 65] is one such example.

For the purpose of this dissertation, we will exclude finite difference models from our study and concentrate solely on a few finite element models, all based on the more robust triangular elements. We will also exclude stability estimates of these models since the corresponding analysis follows easily from the techniques developed here.
In Chapter 2, we give an overview on some finite element models of the primitive formulation of the shallow water equations. We will give introductory information to the *wave* formulation as well as to the *characteristic* formulation of the shallow water equations. Then we will discuss well-posedness which includes issues concerning boundary conditions although, for clarity, we will, omit the corresponding analysis here. In Chapter 3, we will introduce mathematical notions that will be used throughout the paper.
In Chapter 4, we will develop $L^\infty((0,T); L^2(\Omega))$ and $L^2((0,T); H^1(\Omega))$ a priori error estimates, optimal in $H^1(\Omega)$, for the wave formulation. The importance of these estimates is clear when one considers that the literature is devoid of such estimates for the full nonlinear shallow water equations. We will detail the error estimate derivation for a continuous Galerkin approximation with homogeneous Dirichlet boundary conditions. We will then extend these estimates to the fully discrete case. Moreover, we
will give a very brief application of these estimates in work recently done by Wheeler et al [18].

In Chapter 5, we develop a Characteristic-Galerkin method for solving the system of shallow water equations. In that section, we derive $\mathcal{L}^\infty((0,T);\mathcal{L}^2(\Omega))$ and $\mathcal{L}^2((0,T);\mathcal{H}^1(\Omega))$ a priori error estimates which are optimal in $\mathcal{H}^1(\Omega)$ and require a less restrictive time-step constraint than does the estimate for the Galerkin case. For the sake of completeness, we discuss in Chapter 6 some exciting directions that we hope to pursue in the near-future.
Chapter 2

Shallow Water Equations

Shallow water equations can be used to study flow in fluid domains whose

- vertical length scale $H$ is much smaller than the characteristic (horizontal direction) length scale $L$;

- and, in which, the underwater topography doesn't change too fast, that is, for $\alpha$ the degree of inclination of the bottom sea bed, $\tan \alpha \approx \sin \alpha$;

see Weiyan [59], see Figure 2.1.

The 2-dimensional shallow water equations (hereafter, referred to as SWE) are obtained by depth averaging the 3-dimensional incompressible Navier-Stokes equations using appropriate free-surface and boundary conditions along with a hydrostatic pressure assumption.

We can see how this paradigm results in the SWE that we will study by first considering the 3-dimensional incompressible Navier-Stokes equations. Suppose

$$\mathbf{v}(x, y, z, t) = \begin{pmatrix} \dot{u}(x, y, z, t) \\ \dot{v}(x, y, z, t) \\ \dot{w}(x, y, z, t) \end{pmatrix}$$

is the fluid velocity,

- $p(x, y, z, t)$ is pressure,

- $\rho(x, y, z, t)$ is density,

- $0 < \mu$ is viscosity,

- $\mathbf{F}$ are external forces which include a Newtonian tide potential and a Coriolis force,
Figure 2.1 Underwater topography, vertical scale and horizontal scale.

- and $S(x, y, z, t)$ is a stress tensor$^*$. 

Then, the Navier-Stokes equations can be described by

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

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + \frac{1}{\rho} \nabla p - \frac{\mu}{\rho} \Delta \mathbf{v} = \mathbf{F} + \frac{1}{\rho} \nabla \cdot \mathbf{S} \tag{2.2}$$

with free surface boundary conditions at $z = \xi$:

$$\dot{\mathbf{w}} = \frac{\partial \xi}{\partial t} + \frac{\partial \xi}{\partial x} \mathbf{u} + \frac{\partial \xi}{\partial y} \mathbf{v},$$

$$\sigma_{31} = \tau_{ws-x},$$

$$\sigma_{32} = \tau_{ws-y},$$

$$p = p_{a}.$$  

$^*$Recall that the stress tensor has the following structure.

$$\mathbf{S} = \begin{pmatrix}
\sigma_{11} & \sigma_{12} & \sigma_{13} \\
\sigma_{21} & \sigma_{22} & \sigma_{23} \\
\sigma_{31} & \sigma_{32} & \sigma_{33}
\end{pmatrix}$$
and bottom surface boundary conditions at \( z = -h_b \):

\[
\begin{align*}
\mathbf{v} &= 0, \\
\sigma_{31} &= \tau_b u, \\
\sigma_{32} &= \tau_b v,
\end{align*}
\]

where,

- \( \xi(x, y, t) \) is the free surface elevation above a reference plane,

- \( h_b(x, y) \) is the bathymetric depth under that reference plane (see Figure 2.2),

- \( \boldsymbol{\tau}_{ws}(x, y, t) \) is the applied free surface wind stress with \( \boldsymbol{\tau}_{ws} = \begin{pmatrix} \tau_{ws-x} \\ \tau_{ws-y} \end{pmatrix} \),

- \( p_a(x, y, t) \) is the atmospheric pressure at the free surface,

- and \( \tau_b(x, y, t) \) is the bottom friction function.

As an exercise, we will derive the continuity equation for the SWE by integrating (2.1) over the water-column:

\[
0 = \int_{-h_b}^{\xi} \nabla \cdot \mathbf{v} \, dz = \int_{-h_b}^{\xi} \left( \frac{\partial \hat{u}}{\partial x} + \frac{\partial \hat{v}}{\partial y} \right) \, dz + \int_{-h_b}^{\xi} \frac{\partial \hat{w}}{\partial z} \, dz. \quad (2.3)
\]

Now, using

\[
\frac{\partial}{\partial s} \int_{a(s)}^{b(s)} f(s, z) \, dz = \frac{\partial b(s)}{\partial s} f(s, z) \bigg|_{z=b} - \frac{\partial a(s)}{\partial s} f(s, z) \bigg|_{z=a} + \int_{a(s)}^{b(s)} \frac{\partial f(s, z)}{\partial s} \, dz,
\]

we can rewrite (2.3) as

\[
0 = \frac{\partial}{\partial x} \int_{-h_b}^{\xi} \hat{u} \, dz + \frac{\partial}{\partial y} \int_{-h_b}^{\xi} \hat{v} \, dz \\
- \frac{\partial \xi}{\partial x} \bigg|_{z=\xi} - \frac{\partial \xi}{\partial y} \bigg|_{z=\xi} + \frac{\partial (-h_b)}{\partial x} \bigg|_{z=-h_b} + \frac{\partial (-h_b)}{\partial y} \bigg|_{z=-h_b} + \int_{-h_b}^{\xi} \frac{\partial \hat{w}}{\partial z} \, dz.
\]

Applying the prescribed boundary conditions yields

\[
0 = \nabla \cdot (\mathbf{u} H) + \frac{\partial \xi}{\partial t}.
\]
where \( \mathbf{u} = [u(x, t), v(x, t)] \) and

\[
\begin{align*}
u & = \frac{1}{H} \int_{-h_b}^{\xi} \hat{u} \, dz, \\
v & = \frac{1}{H} \int_{-h_b}^{\xi} \hat{v} \, dz
\end{align*}
\]

are the depth-averaged horizontal velocities.

Deriving the equations of motion for the SWE is more involved. To do so, we need to first assume that pressure is hydrostatic, that is, the fluid pressure is assumed to have a linear distribution over the water-column (as long as the fluid is continuous). This assumption can be written as the “state equation” for pressure

\[
\frac{\partial p}{\partial z} = -\rho g.
\]

Throughout, \( g \) is acceleration due to gravity (\( g = 32 ft/s^2 \) or \( g = 9.8 m/s^2 \)). Integrating the state equation over the depth of the water body, we get the equation for pressure

\[
p = p_a + \rho g (\xi + h_b)
\]

so we can remove pressure as a variable in (2.2). Then, we depth-integrate only the \( x \)- and \( y \)- (horizontal) direction momentum equations in (2.2). Thus, vertical velocity is assumed to be negligible.
We should point out that in some of the literature, the shallow water equations are derived from the turbulent Reynolds equation (with Boussinesq’s approximation for relating Reynolds stress to velocity) which are themselves derived from the Navier-Stokes equations. We refer the reader to [42, 9, 59, 51]. To summarize, the shallow water equations have the main properties that the flow is compressible, that the horizontal length scale is much greater than the vertical length scale, and that the underwater topography doesn’t change too fast. It should be noted that the latter two properties implies that $\nabla h_b$ is small in the sense that $\left|\frac{\partial h_b}{\partial x}\right|, \left|\frac{\partial h_b}{\partial y}\right| \approx \frac{d|h_b|}{d|L|} < \frac{d|H|}{d|L|} << 1$.

2.1. Primitive Formulation, P-SWE.

Letting $U = uH$, the 2-dimensional governing equations, in operator form [38], are the continuity equation (CE)

$$\text{CE}(\xi, u; h_b) \equiv \frac{\partial \xi}{\partial t} + \nabla \cdot U = 0 \quad (2.4)$$

and the non-conservative momentum equations (NCME), as derived by Westerink et al [62],

$$\text{NCME}(\xi, u; \Phi) = \frac{\partial u}{\partial t} + (u \cdot \nabla) u + g \nabla \xi - \frac{\mu}{H} \Delta U + \tau_b u + f_c k \times u + \mathcal{F} = 0, \quad (2.5)$$

where $\Phi = (h_b, g, \mu, f_c, \tau_b, \mathcal{F})$. Here, $k$ is a unit vector in the vertical direction, and $\mathcal{F} = (-\frac{1}{H} \tau_{ws} + \nabla p_a - g \nabla \mathcal{N})$, where now $\tau_{ws}$ is the applied free surface wind stress relative to the reference density of water, $p_a(x, t)$ is the atmospheric pressure at the free surface relative to the reference density of water, and $\mathcal{N}(x, t)$ is the Newtonian equilibrium tide potential relative to the effective Earth elasticity factor. As defined in [42], the bottom friction function is given by

$$\tau_b(\xi, u) = c_f \frac{\sqrt{u^2 + v^2}}{H}, \quad (2.6)$$
where \( c_f \) is a friction coefficient. And, the Coriolis function is given by 
\[ f_c(x) = 2 \omega \sin \alpha, \]
where \( \omega \) is the angular velocity of the earth in its daily rotation \((7.29 \times 10^{-5} \text{ s}^{-1})\) and \( \alpha \) is the degrees latitude. It should be noted that the final form of the viscosity term is a point of contention in the literature - other forms of it are \( \mu \Delta \mathbf{u} \) and \( \frac{\mu}{H} \nabla \cdot H \nabla \mathbf{u} \); see [28, 9].

The conservative momentum equations (CME)\(^\dagger\) are derived from the (NCME) as
\[
\text{CME} \equiv H(\text{NCME}) + \mathbf{u}(\text{CE}) = 0.
\]

We shall refer to (CE)-(NCME) and (CE)-(CME) as the primitive shallow water equations (P-SWE) to be consistent with literature terminology.

The numerical procedure used to solve the shallow water equations must resolve the physics of the problem without introducing spurious oscillations, numerical noise or excessive numerical diffusion.

Moreover, Westerink et al [62] note a need for greater grid refinement near land boundaries to resolve important processes there and to prevent high frequency (small wavelength) components of the solution from aliasing (polluting low frequencies). Permitting a high degree of grid flexibility, the finite element method is a good candidate.

2.1.1. An Overview of the FEM for the P-SWE.

There has been a substantial effort over the past two decades in applying finite element methods to P-SWE. Early finite element simulations of shallow water systems were

\(^\dagger\)Recall that the conservative form of the momentum equations is obtained by multiplying the NCME by the primary variable \( \phi \) of the continuity equation, expanding the modified advective term \( \phi \mathbf{u} \cdot \nabla \mathbf{u} = \nabla \cdot \mathbf{u} (\mathbf{u} \phi) - \mathbf{u} \nabla (\mathbf{u} \phi) \) to incorporate the definition of the continuity equation, and finally combining like terms. In incompressible flow, \( \phi \) is constant.
plagued by numerical noise, including spurious oscillations. Various methods were introduced to eliminate the spurious oscillations through artificial diffusion. For example, Partridge and Brebbia [49, 12, 13], employing quadratic triangular elements, use excessive bottom friction. King and Norton [37, 36], employing mixed-interpolation triangular elements with linear polynomials approximating elevation and quadratic polynomials approximating velocity, use excessive viscosity in the momentum equation. Conner and Wang [58], employing linear triangular elements, also use artificial viscosity in the momentum equation. These methods tended to damp-out physical components of the solution.

A few years later, Gresho and Lee [32] comment that numerical noise in finite element models was partly due to inadequate mesh refinement in regions of steep gradients. Then, Walters and Carey [56, 53] showed that the choice of elements also played an equally important role. For instance, equal-order linear and quadratic approximations to the variables of the primitive formulation are prone to spurious oscillations of wavelength $2h$, where $h$ corresponds to the mesh size.

Additionally, Walters and Carey proved that mixed-interpolation spaces, consisting of piecewise constant polynomials to approximate elevation and linear polynomials to approximate velocity did not have convergent modes on triangles in two dimensions, unlike the one-dimensional case in which there are no spurious oscillations. And, mixed-interpolation spaces consisting of quadratic polynomials to approximate elevation and linear polynomials to approximate velocity were shown to have spurious modes in both one and two dimensions.

However, Walters and Carey [56] also proved (and Walters [53] later numerically confirmed) that the use of a mixed-interpolation space, consisting of linear polynomials
to approximate elevation and quadratic polynomials to approximate velocity. removed
the spurious modes from the elevation variable but not from the velocity variable\textsuperscript{1}.

It is thus interesting to note that the Hydraulics Laboratory at CE-WES has
had some success with a mixed-interpolation finite-element model of the primitive
formulation, based on the work of King and Norton, in their simulators RMA-2
(2D-model), RMA-4(2D-model), RMA-10(3-D on an unstructured mesh) in model-
ing the Atchafalaya Bay Delta, Charleston Estuary, Chesapeake Bay, Galveston Bay
Navigation Channels, Mississippi River at Redeye Crossing, and San Francisco Bay.
Still, the viscosity parameter might be driven too high to stabilize the numerical
method.

In the beginning of the 1990s, Bernardi and Pirroneau [8, 9] theoretically analyze a
spatial finite element approximation and implicit finite difference temporal scheme of
four formulations of the primitive equations at low Reynolds numbers, which excludes
the full non-linear equations under study here.

Recently, Agoshkov et al [2, 3, 5, 4] have investigated a spatial finite element ap-
proximation, where the velocity field is approximated by piecewise linear polynomials
and the elevation is approximated by some embellished linear polynomials, examples
of which are the bubble functions. They have studied the effects of various admissible
boundary conditions, and proven stability of various temporal finite difference schemes
for P-SWE system under limiting conditions. However, these authors defer both the
issue of spurious waves and of error estimates.

\textsuperscript{1}Of side interest is the fact that this mixed-interpolation space eliminates spurious oscillations in
the incompressible Navier-Stokes equations because the transient term is absent from the continuity
equation and so it cannot couple oscillations between elevation and velocity.
2.2. Wave Formulation, W-SWE.

The wave formulations of the shallow water equations are based on a reformulation of the continuity equation into a wave (or harmonic) equation. Computational and experimental evidence in the literature suggest that these formulations lead to approximate solutions with reduced oscillations. Moreover, these approximate solutions have accurately matched tidal data.

2.2.1. Harmonic Wave SWE.

Before we review the wave formulation of Gray and Lynch, we briefly review independent work done on another wave formulation. Walters [53, 54, 55, 57] published a series of articles on what he called a harmonic wave model (or spectral form) of the SWE. In this model, the primary variables $\xi$ and $u$ are expanded into finite Fourier series of a time-average component plus a sum of periodic components with angular frequencies $\omega_n$:

$$
\xi = \xi_0(x) + \frac{1}{2} \sum_{n=-N}^{N} \xi_n(x)e^{-i\omega_n t},
$$

$$
u = u_0(x) + \frac{1}{2} \sum_{n=-N}^{N} u_n(x)e^{-i\omega_n t}.
$$

These expansions are substituted into the CE and the NCME. The resulting equations are then multiplied by $e^{i\omega_n t}$ and averaged over time (by taking integrals from $-T$ to $T$ and dividing by $2T$) followed by taking the limit in time as $T \rightarrow \infty$. Finally, velocity modes are eliminated from the harmonic CE, using the harmonic momentum equation, to arrive at the final harmonic CE. The harmonic CE together with the harmonic momentum equation constitute the harmonic wave model. An obvious advantage of this form is that time-stepping is completely circumvented.

Finally, Walters showed that his harmonic decomposition-in-time wave model of the shallow water equations was not prone to spurious oscillations and matched tidal
data better than any of the standard methods (equal-order quadratic interpolation, equal-order linear interpolation, and mixed-order interpolation) applied to the primitive formulation. Still this model is efficient only for problems in which forces are periodic and for which there exist a small number of frequency components, [53].

2.2.2. Historical Development of the W-SWE.

In 1979, Lynch and Gray [45] derived the wave continuity equation (WCE), from the mass and momentum conservation equations, as a means to eliminate oscillations without resorting to numerical damping and to model motion in which forces may be aperiodic. The WCE is called the wave form of the CE because it is a second-order wave equation in $\xi$ when linearized. The WCE can be described from

$$\text{WCE}(\xi, u; \Phi) \equiv \frac{\partial(\text{CE})}{\partial t} - \nabla \cdot (\text{CME}) + \tau(\text{CE}) = 0,$$

where $\tau(x, t)$ is a non-linear friction coefficient that balances the good mass-conservative properties of the CE and the good spurious-mode suppression of the second-order hyperbolic wave equation.

In this shallow water formulation, the WCE is then coupled to either the CME or to the NCME. The equivalence of this model to the more standard one based on the CE is discussed by Kinnmark[38].

This formulation has led to the development of robust finite element algorithms for depth-integrated coastal circulation models. The WCE approach has motivated a substantial computational and analytical effort [23, 26, 45, 53, 40]. Using Fourier phase/space analysis of the linearized WCE-NCME and WCE-CME system of equations, Foreman [26] and Kinnmark [38] prove that the WCE formulation suppresses spurious oscillations of the numerical solution, and is capable of capturing $2h$ waves (on a uniform grid [38]) that would otherwise pollute longer wavelengths. The WCE formulation has also motivated substantial field applications; see [30], [27], [29], [46],
These studies have demonstrated the advantage of the WCE formulation for finite element applications in terms of achieving both a high level of computational accuracy and efficiency.

The generalized wave continuity equation (GWCE) [38] is essentially the same as the WCE except that multiplication of the continuity equation by \( \tau \) is replaced with multiplication by some general function, \( G \), that may be independent of time. Luettich et al [42] chose to replace \( \tau \) by a time-independent positive constant \( \tau_0 \). Their version of the GWCE is given by

\[
\text{GWCE} \equiv \frac{\partial^2 \xi}{\partial t^2} + \tau_0 \frac{\partial \xi}{\partial t} - \nabla \cdot \left[ \nabla \cdot \left( \frac{U^2}{H} \right) + gH \nabla \xi + \mu \nabla \frac{\partial \xi}{\partial t} \right] + (\tau_b - \tau_0)U + f_k \times U + H\mathbf{F} = 0, \tag{2.8}
\]

where it was assumed that \( \nabla \cdot \nabla U \approx \nabla (\nabla \cdot U) \).

The choice of \( G = \tau_0 \) yields a system of time-independent matrices when the GWCE is discretized in time using a three-level implicit scheme for linear terms and treating non-linear terms explicitly. Although \( \tau_0 \) can be chosen arbitrarily, Kinnmark [38] suggests that \( \frac{\tau_0}{\tau_b} \in [0, 2] \), and Kolar et al [39] alternatively suggest that \( \frac{\tau_0}{\max(\tau_b)} \) be on the order 1-10 for best numerical results. Still, this tuning-parameter \( \tau_0 \) appears to be the basis of some controversy within the shallow water community.

The GWCE can be coupled to the CME, given by

\[
\text{CME} \equiv \frac{\partial U}{\partial t} + \nabla \cdot \left( \frac{U^2}{H} \right) + gH \nabla \xi - \mu \Delta U + \tau_i U + f_k \times U + H\mathbf{F} = 0, \tag{2.9}
\]

or to the NCME. A finite element simulator, \textit{ADCIRC}, based on the GWCE-NCME system of equations using equal-order interpolation, namely linear polynomials, to approximate elevation and velocity unknowns, has been developed by Luettich, et al. In [42], it was demonstrated that the approximations generated by this simulator

\[\text{Recall, that } \nabla \cdot \nabla U \equiv \nabla (\nabla \cdot U) \text{ in incompressible flow.}\]
accurately matched tidal data taken from the English Channel and southern North Sea. We shall refer to the GWCE-NCME and the GWCE-CME as the wave shallow water equations (W-SWE). Note that in 1981, Lynch [43] showed that the W-SWE converges to the harmonic wave SWE as $\Delta t \to 0$.

2.3. Characteristic Formulation, C-SWE.

The characteristic formulation of the SWE is based on manipulating the governing equations into a form in which the time derivative and the advective term have been absorbed into a directional derivative.

Denote by

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \cdot \nabla$$

as the substantial or material or directional derivative. We can write the continuity equation (2.4) in characteristic form as

$$\frac{D H}{Dt} + H \nabla \cdot u = 0,$$

and the non-conservative momentum equation (2.5) in characteristic form as

$$\frac{Du}{Dt} + g \nabla \xi - \frac{\mu}{H} \Delta u H + \tau_{y} u + f_{z} k \times u + F = 0.$$

2.3.1. The Method of Characteristics.

A typical hyperbolic equation

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = f$$

can be solved by the method of characteristics (MOC).

Now rewrite the above equation in characteristic form as

$$\frac{Du}{Dt} = f.$$

(2.10)
A typical finite difference approximation for solving (2.10) is
\[
\frac{D\mathbf{u}}{Dt} \approx \frac{\mathbf{u}^{n+1} - \mathbf{u}^*}{\Delta t} = f^{n+1}.
\] (2.11)

We first take a Lagrangian step by tracking each particle to obtain \(\mathbf{u}^*\). Upon determining \(\mathbf{u}^*\), we revert back to an Eulerian approach to solve for \(\mathbf{u}^{n+1}\). This type of process is called Eulerian-Lagrangian. The name Eulerian-Lagrangian comes from two contrasting viewpoints of fluid flow in which, in the *Eulerian* viewpoint, the flow field is studied at a given spatial region (fixed finite control volume), and in which, in the *Lagrangian* viewpoint, the motion of a given fluid particle in a fixed space region is studied. The Eulerian-Lagrangian approach is also called the modified method of characteristics in the mathematics literature.

The method of characteristics (MOC) is a well-known technique for solving first-order hyperbolic equations, usually when there is advection-dominated flow. The modified method of characteristics (MMOC) is simply a time-stepping procedure along the characteristic "lines" (determined by the MOC) in combination with any spatial discretization. The MMOC together with the Galerkin FEM constitute the Characteristic-Galerkin (CG) method.

### 2.3.2. Historical Development of the C-SWE.

In the mid-1990s, Zienkiewicz and Ortiz [64, 65] have proposed a Characteristic-Galerkin approximation to P-SWE with promising numerical results. Their method relies on a Chorin-type projection[19] with fractional-stepping along the characteristics (of the NCME) which are themselves approximated using a Taylor expansion. However, their particular Taylor expansion assumes that the foot of the characteristic is very close to the nodal point around which the expansion was taken. Moreover, these authors don’t provide any theoretical framework for their observations.
2.4. Well-Posedness and Boundary Conditions.

Theoretically, a mathematical problem is well-posed\cite{33} if

1. the solution exists.

2. the solution is unique.

3. the solution depends continuously on the initial and boundary conditions (stability).

However, for some problems, such as the shallow water equations discussed here, analytical proofs of these properties may be intractable if not impossible without oversimplifying the problem. Still various attempts have been made at addressing these well-posedness properties, the last of which is the easiest to prove if the initial and boundary conditions are understood, the latter of which is discussed often in the literature.

In the context of surface flow, the boundary of $\Omega$ consists of a land boundary, a river boundary and an ocean boundary as in Figure 2.3. The conditions expressed on these boundaries are usually complicated and cannot be easily classified as either essential or natural, although Agoshkov et al \cite{5} have recently undertaken this classification with one caveat - they have excluded the cases for which they could not conclude the system was stable or not. Drolet and Gray \cite{23} are only able to determine the number of boundary conditions to fulfill well-posedness for the P-SWE and the W-SWE.

For the P-SWE and GWCE-CME models (in subcritical flow), two boundary conditions must be specified for inflow, and one each for outflow and no-flow boundaries. For the GWCE-NCME model (in subcritical flow) three boundary conditions must be specified for inflow and one each for outflow and no-flow boundaries.

Based on some physical understanding, the following conditions are usually used in the shallow water community \cite{59}. Let $\nu, \tau$ correspond to the normal and tangential
Figure 2.3 Land and River/Sea Boundaries

directions to the domain boundary. Then, \( u_\nu = u \cdot \nu \) and \( u_\tau = u \cdot \tau \) correspond to the normal and tangential velocities, respectively. Hence,

- Fixed land boundary: \( \partial \Omega_L \). This no-flow boundary condition can be described by
  \[
  u \cdot \nu = 0, \quad \nabla u_\tau \cdot \nu = 0, \quad \text{and} \quad \frac{\partial \xi}{\partial \nu} = 0.
  \]

- River boundary: \( \partial \Omega_R \). This river discharge may be looked upon as an inflow open boundary condition.
  \[
  u \cdot \nu = \dot{u} \cdot \nu, \quad \mu \frac{\partial u}{\partial \nu} \cdot \tau = 0 \quad \text{and} \quad \xi = \dot{\xi}.
  \]

- Open ocean boundary: \( \partial \Omega_O \). The open ocean boundary conditions are generally difficult to determine in practice, namely because velocity information is difficult to accurately measure. The ocean boundaries are therefore chosen sufficiently far away from land boundaries to minimize their effect on those regions. To that end, we may choose,
  \[
  \xi = \dot{\xi} \quad \text{and} \quad \frac{\partial u}{\partial \nu} = 0.
  \]

Notice that we can write \( \frac{\partial u}{\partial \nu} = 0 \) as \( \nabla u_\nu \cdot \nu = \nabla u_\tau \cdot \nu = 0 \).

We will write \( \partial \Omega = \partial \Omega_L \cup \partial \Omega_R \cup \partial \Omega_O \).
Chapter 3

Notation, Definitions and Mathematical Notions

For the purpose of clarity, we now define some notation used hereafter. Let \( \Omega \) be a polyhedral domain in \( \mathbb{R}^2 \). Let \( \bar{\Omega} = \Omega \cup \partial \Omega \) where \( \partial \Omega \) is the boundary of \( \Omega \subset \mathbb{R}^2 \). Let \( \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \).

For \( 1 \leq p \leq \infty \), let

\[
L^p(\Omega) \equiv \left\{ \int_{\Omega} |\varphi|^p \, dx < \infty \right\}, \quad \text{if } 1 \leq p < \infty.
\]

\[
L^\infty(\Omega) \equiv \{ |\varphi(\mathbf{x})| \leq K \text{ a.e. on } \Omega, \text{ for a constant } K \}
\]

be the class of all Lebesgue measurable functions with norms

\[
||\varphi||_{L^p(\Omega)} = \left\{ \int_{\Omega} |\varphi|^p \, dx \right\}^{1/p}, \quad \text{if } 1 \leq p < \infty,
\]

\[
||\varphi||_{L^\infty(\Omega)} = \text{ess sup}_{\mathbf{x} \in \Omega} |\varphi(\mathbf{x})|.
\]

Moreover, the \( L^2 \) inner product is denoted by

\[
(\varphi, \varsigma) = \int_{\Omega} \varphi \circ \varsigma \, dx, \quad \varphi, \varsigma \in [L^2(\Omega)]^n,
\]

where "\( \circ \)" here refers to either multiplication, dot product, or double dot product as appropriate and \( n = 1, 2, \ldots \) refers to the space dimension. The \( L^2 \) norm is given by

\[
||\varphi|| = ||\varphi||_{L^2(\Omega)} = (\varphi, \varphi)^{1/2}.
\]

In \( \mathbb{R}^n \), \( \alpha = (\alpha_1, \ldots, \alpha_n) \) is an \( n \)-tuple with nonnegative integer components,

\[
D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}
\]

and \( |\alpha| = \sum_{i=1}^n \alpha_i \).

For \( \ell \) any nonnegative integer and \( 1 \leq p \leq \infty \), let

\[
\mathcal{W}^\ell_p \equiv \{ \varphi \in L^p(\Omega) \mid D^\alpha \varphi \in L^p(\Omega) \text{ for } |\alpha| \leq \ell \}
\]
be the Sobolev space with norm
\[\|\varphi\|_{W^s_p(\Omega)} = \left(\sum_{|\alpha| \leq s} \|D^\alpha \varphi\|_{L^p(\Omega)}^p\right)^{1/p}, \quad \text{if } 1 \leq p < \infty,\]
\[\|\varphi\|_{W^s_\infty(\Omega)} = \max_{|\alpha| \leq s} \|D^\alpha \varphi\|_{L^\infty(\Omega)}.\]

Moreover, the usual Sobolev space \(H^s\) is defined as \(H^s = W^s_2\). Additionally, \(H^1_0(\Omega)\) denotes the subspace of \(H^1(\Omega)\) obtained by completing, with respect to the norm \(\|\cdot\|_{H^1(\Omega)}\), the set of infinitely differentiable functions with compact support in \(\Omega\), \(C^\infty(\Omega)\). For relevant properties of these spaces, please refer to [1].

Now observe, for instance, that \(H^s\) are spaces of \(\mathbb{R}\)-valued functions. Spaces of \(\mathbb{R}^n\)-valued functions will be denoted in boldface type, but their norms will not be distinguished. Thus, \(L^2(\Omega) = [L^2(\Omega)]^n\) has norm \(\|\varphi\|^2 = \sum_{i=1}^n \|\varphi_i\|^2\); \(H^1(\Omega) = [H^1(\Omega)]^n\) has norm \(\|\varphi\|_{H^1(\Omega)}^2 = \sum_{i=1}^n \sum_{|\alpha| \leq 1} \|D^\alpha \varphi_i\|^2\); etc.

Moreover, we define a discrete temporal subdomain of \([0,T]\) by \(J_{\Delta t} = \{t^k | t^k \in [0,T], t^k = k\Delta t, k = 0, \ldots, N, N\Delta t = T, \Delta t \geq 0\}\), and denote \(J^0_{\Delta t} = J_{\Delta t} - \{0\}\).

For \(X\), a normed space with norm \(\|\cdot\|_X\) and a map \(\varphi : [0,T] \to X\), define \(\varphi^k = \varphi^k(x) = \varphi(x,t^k)\) and
\[\|\varphi\|_{L^2([0,T];X)}^2 = \int_0^T \|\varphi\|_X^2 dt,\]
\[\|\varphi\|_{L^\infty([0,T];X)} = \sup_{0 \leq t \leq T} \|\varphi\|_X,\]
\[\|\varphi\|_{L^2_{\Delta t}([0,T];X)}^2 = \sum_{k=0}^N \|\varphi^k\|_X^2 \Delta t,\]
\[\|\varphi\|_{L^\infty_{\Delta t}([0,T];X)} = \sup_{0 \leq k \leq N} \|\varphi^k\|_X.\]

3.1. Mathematical Inequalities.

There will be occasion to use the following inequalities.

Let \(0 \leq a, b \in \mathbb{R}\), then
\[2(a - b)a = a^2 - b^2 + (a - b)^2 \geq a^2 - b^2.\]
Let $0 \leq a, b \in \mathbb{R}$, then

$$(a + b)^p \leq 2^{p-1} (a^p + b^p), \quad 1 \leq p < \infty;$$

$$(a + b)^p \leq a^p + b^p. \quad 0 \leq p < 1.$$  

**Arithmetic-Geometric Mean Inequality (AGMI):** Let $a, b, \epsilon \in \mathbb{R}$, and $\epsilon > 0$, then

$$ab \leq \frac{1}{4\epsilon} a^2 + \epsilon b^2.$$  

**Hölder Inequality (HI):** Let $f \in L^p(\Omega), g \in L^q(\Omega)$ where $p, q \geq 1$, and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_\Omega |f(x)g(x)| \, dx \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}.$$  

Recall that the Cauchy-Schwartz Inequality is just a special case of the HI when $p = q = 2$.  

**Generalized Hölder Inequality (GHI):** Let $f \in L^p(\Omega), g \in L^q(\Omega), h \in L^r(\Omega)$ where $p, q, r \geq 1$, and $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, then

$$\int_\Omega |f(x)g(x)h(x)| \, dx \leq \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} \|h\|_{L^r(\Omega)}.$$  

Typical values of $p, q, r$ are $p = q = 2$ and $r = \infty$.  

**Integral Gronwall's Lemma (IGL):** If $f, g, h$ are piecewise continuous, non-negative functions, $h$ is non-decreasing and $\gamma, I$ are positive, such that

$$f(\tau) + g(\tau) \leq h(\tau) + \gamma \int_\alpha^\tau f(s) \, ds + I, \quad \forall \tau \in [\alpha, \beta],$$

then,

$$f(\tau) + g(\tau) \leq e^{\gamma(\tau-\alpha)} (h(\tau) + I).$$
Generalized Discrete Gronwall’s Lemma (GDGL): [Heywood and Rannacher [34]] Let $\Delta t, I$ and $f_k, g_k, h_k$, for integers $k \geq 0$, be nonnegative numbers such that

$$f^N + \sum_{k=0}^{N} g^k \Delta t \leq \sum_{k=0}^{N} h^k \Delta t + \sum_{k=0}^{N} \gamma^k f^k \Delta t + I, \quad N \geq 0.$$  

Suppose that $\gamma^k \Delta t \leq 1 \ \forall k$. Set $\sigma^k = \left(1 - \gamma^k \Delta t\right)^{-1}$. Then,

$$f^N + \sum_{k=0}^{N} g^k \Delta t \leq \exp \left( \sum_{k=0}^{N} \gamma^k \sigma^k \Delta t \right) \left\{ \sum_{k=0}^{N} h^k \Delta t + I \right\}, \quad N \geq 0.$$

The usual discrete Gronwall’s Lemma follows when the summation in term (*) extends only up to $N - 1$, in which case the result of the lemma follows with $\sigma^k = 1$. In particular, we have

$$f^N + \sum_{k=0}^{N} g^k \Delta t \leq \exp \left( \sum_{k=0}^{N} \gamma^k \Delta t \right) \left\{ \sum_{k=0}^{N} h^k \Delta t + I \right\}, \quad N \geq 0$$

$$\leq \exp \left( K \sum_{k=0}^{N} \Delta t \right) \left\{ \sum_{k=0}^{N} h^k \Delta t + I \right\}, \quad N \geq 0, \ K = \max_k \gamma^k$$

$$= e^{K(N \Delta t)} \left\{ \sum_{k=0}^{N} h^k \Delta t + I \right\}.$$

which is the result proved by Lees[41] in 1960.

There will also be occasion to use the following standard relations, given $f$:

$$\left( \frac{d}{dt} f, f \right) = \frac{1}{2} \frac{d}{dt} \left( ||f||^2 \right),$$

and, for $\alpha \in \mathbb{R}$,

$$\frac{1}{2} \frac{d}{dt} \left( e^{\alpha t} ||f||^2 \right) = \frac{1}{2} e^{\alpha t} \frac{d}{dt} \left( ||f||^2 \right) + \text{sgn}(\alpha) \frac{\alpha}{2} e^{\alpha t} ||f||^2.$$
3.2. *A Priori Error Estimate Tools.*

Since we will derive *a priori* error estimates, we first outline the standard paradigm for these derivations.

**EE1.** Derive the weak form of the error equations by subtracting the finite-dimensional weak (or variational) form from the equations satisfied by the finite element approximation.

**EE2.** Because the discretization error $E_\phi$ between the weak solution $\phi$ and the finite element approximation $\phi_h$ is difficult to estimate directly, we separate the discretization error into the sum of an approximation (or projection) error $\tilde{E}_\phi = \phi - \tilde{\phi}$ and an affine error $\tilde{E}_\phi = \phi_h - \tilde{\phi}$. This splitting is accomplished by taking $E_\phi = \phi - \phi_h$ and adding and subtracting a comparison function $\tilde{\phi}$:

$$\phi - \phi_h = \phi - \tilde{\phi} + \tilde{\phi} - \phi_h.$$

Then use the triangle inequality to separate and bound the errors:

$$||E_\phi|| \leq ||\tilde{E}_\phi|| + ||\tilde{E}_\phi||.$$

The approximation error is either known before-hand or can be easily determined. The affine error is not known and must therefore be estimated.

**EE3.** Choose the comparison function $\tilde{\phi}$, so that the approximation error can be estimated from approximation theory or from an auxiliary problem. In choosing a comparison function, one should consider the sharpness of the approximation error in relation to the sharpness of the affine error. Standard choices of comparison functions are the $L^2$ projection, an elliptic or $H^1$ projection [63, 35], a parabolic projection [48], a cutoff-function [14, 52], or the solution to the linearized model.
EE4. Choose the test function in the weak form of the error equations to facilitate the derivation of an estimate of the affine error.

EE5. Finally, knowing the affine error, add this to the approximation error, to get an upper bound estimate of the discretization error.

3.3. Approximation Theory Notions.

Let $\mathcal{T}$ be a quasi-uniform triangulation of $\Omega$ into elements $\omega_i$, $i = \{1, \ldots, n\}$, with $\text{diam}(\omega_i) = h_i$ and $h = \max_i h_i$.

Let $\mathcal{S}_h^0$ denote a finite dimensional subspace of $\mathcal{H}_0^1(\Omega)$ defined on this triangulation consisting of piecewise polynomials of degree less than $s_1$, and satisfying the standard approximation property

$$
\inf_{\varsigma \in \mathcal{S}_h^0} \|\phi - \varsigma\|_{\mathcal{H}_0^1(\Omega)} \leq K_0 h^{\ell-s_0} \|\phi\|_{\mathcal{H}_0^1(\Omega)}, \quad \phi \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}_0^\ell(\Omega),
$$

for integers $s_0, \ell$ and $0 \leq s_0 \leq \ell \leq s_1$ and where $K_0$ is a constant independent of $h$ and $\phi$.

Let $\mathcal{S}_h$ denote a finite dimensional subspace of $\mathcal{H}^1(\Omega)$, similarly defined as is $\mathcal{S}_h^0$, also satisfying (3.1) with $\mathcal{S}_h^0$ replaced by $\mathcal{S}_h$ and with $\mathcal{H}_0^1(\Omega)$ replaced by $\mathcal{H}^1(\Omega)$.

Moreover, we have the following standard inverse estimate.

**Lemma 3.1 (Inverse Estimate)** (See Brenner and Scott [14].) Let $h \in (0, 1]$ and $\mathcal{S}_h \subset \mathcal{W}_p^r(\Omega) \cap \mathcal{W}_q^m(\Omega)$, where $\Omega$ is a polyhedral domain in $\mathbb{R}^n$, $1 \leq p \leq \infty$, $1 \leq q \leq \infty$, and $0 \leq m \leq r$, then there exists a $K_0 = K_0(r, p, q)$ such that $\forall v \in \mathcal{S}_h$, we have

$$
\|v\|_{\mathcal{W}_p^r(\Omega)} \leq K_0 h^{m-r+\min(0, \frac{n}{p}-\frac{n}{q})} \|v\|_{\mathcal{W}_q^m(\Omega)}.
$$

Note that approximating spaces $\mathcal{S}_h^0$ also satisfies Lemma 3.1.
Finally, it is of some importance to distinguish what is meant as an optimal estimate of an error. We shall say that, for $\phi \in H^\ell(\Omega)$ and $\tilde{\phi} \in S^0_h$ (or $S_h$),

$$\|E_\phi\|_{L^2(\Omega)} \leq K_0 h^\ell \|\phi\|_{H^\ell(\Omega)}$$

and

$$\|E_\phi\|_{H^1(\Omega)} \leq K_0 h^{\ell-1} \|\phi\|_{H^\ell(\Omega)}$$

correspond to optimal estimates in $L^2$ and in $H^1$, respectively.

### 3.3.1. The $L^2$ Projection.

We define the $L^2$ projection of $\phi \in H^\ell_0(\Omega)$ into $S^0_h$ by $\tilde{\phi}$ satisfying

$$((\phi - \tilde{\phi})(\cdot, t), v) = 0, \quad \forall v \in S^0_h, t \geq 0. \quad (3.2)$$

The following results are standard.

**Lemma 3.2** Let $0 \leq q \leq \ell \leq s_1$. Let $\phi \in L^2((0, T); H^\ell_0(\Omega) \cap H^s(\Omega))$ and let $\tilde{\phi}$ be the corresponding $L^2$ projection of $\phi$ into $S^0_h$. And let $\theta = \phi - \tilde{\phi}$. If for some integer $j \geq 0$, $\left(\frac{\partial}{\partial t}\right)^j \phi \in L^2((0, T); H^j(\Omega))$, then $\left(\frac{\partial}{\partial t}\right)^j \tilde{\phi} \in L^2((0, T); S^0_h)$ and

$$\left\|\left(\frac{\partial}{\partial t}\right)^j \theta\right\|_{L^2((0, T); H^j(\Omega))} \leq K_0 h^{\ell-q} \left\|\left(\frac{\partial}{\partial t}\right)^j \phi\right\|_{L^2((0, T); H^s(\Omega))}, \quad (3.3)$$

for some constant $K_0$ independent of $\phi, q, h, \ell$, where $s = \min(\ell, s_1)$.

Finally, there will be occasion to employ the following lemma whose proof can be found in Brenner and Scott [14] in Corollary 4.8.9.

**Lemma 3.3** Let $\phi \in W^4_\infty(\Omega)$. Then, with $\tilde{\phi}$ defined via (3.2), its first-order spatial derivatives are bounded above in $L^\infty((0, T); L^\infty(\Omega))$ by a positive constant $K^*$. 
3.3.2. One Final Comment on Notation

We will also occasionally use the abbreviations

IBP(T) to mean integration by parts over time, and

IBP(Ω) to mean integration by parts over spatial domain Ω.

Finally, we let $K, K_i, (i = 0, 1, 2, ..)$ and $\epsilon$ be generic constants, not necessarily the same at every occurrence.
Chapter 4

Finite Element Model of W-SWE.

4.1. Equivalence of Solutions.

Gray and Lynch [31] have stated that the numerical solution to the GWCE-NCME is essentially the same as that obtained from the GWCE-CME. However, Kinnmark [38] proves that

1. the solution to the GWCE-NCME, plus initial condition \( CE\big|_{t=0} = 0 \), is analytically equivalent to the solution of CE-NCME;

2. the solution to the GWCE-CME, plus initial condition \( CE\big|_{t=0} = 0 \), is analytically equivalent to the solution of CE-NCME.

But, if \( CE\big|_{t=0} \neq 0 \), then

1. the solution to the GWCE-NCME, with the initial condition \( CE\big|_{t=0} \neq 0 \), is analytically equivalent, as \( t \uparrow \), to the solution of the CE-NCME because \( CE \to 0 \) as \( t \uparrow \) when \( \tau_o > 0 \);

2. the solution to the GWCE-CME, with the initial condition \( CE\big|_{t=0} \neq 0 \), is analytically equivalent, as \( t \uparrow \), to the solution of the CE-NCME because \( CE \to 0 \) as \( t \uparrow \) when \( (\tau_o - \nabla \cdot u) > 0 \).

Because of the simpler form of the NCME, it was this form that was actually implemented in ADCIRC. However, because of the natural symmetry between terms in the GWCE and in the CME formulations, we derive an \textit{a priori} error estimate for this system instead. Another alternative is to rewrite the advective term in the GWCE
in non-conservative form to conform to the NCME. Of course, one can always invoke the equivalence of solution of this system and that for the GWCE-NCME.

4.2. Continuous-Time Nonlinear Model.

4.2.1. Continuous-Time Variational Formulation.

Let \( T = g \xi \nabla \xi + f_c k_x U + H \mathcal{F} \). Consider the coupled system given by the GWCE-CME

\[
\frac{\partial^2 \xi}{\partial t^2} + \tau_o \frac{\partial \xi}{\partial t} - \nabla \cdot \left[ \nabla \cdot \left( \frac{U^2}{H} \right) + gh_b \nabla \xi + \mu \nabla \frac{\partial \xi}{\partial t} + (\tau_o - \tau_v)U + T \right] = 0. \tag{4.1}
\]

\[
\frac{\partial U}{\partial t} + \nabla \cdot \left( \frac{U^2}{H} \right) + gh_b \nabla \xi - \mu \Delta U + \tau_v U + T = 0. \tag{4.2}
\]

with the initial conditions

\[
\begin{align*}
\xi(x, 0) &= \xi_0(x), \\
\frac{\partial \xi}{\partial t}(x, 0) &= \xi_1(x), \\
u(x, 0) &= u_0(x),
\end{align*}
\tag{4.3}
\]

and, for simplicity, consider homogeneous Dirichlet boundary conditions.

As noted in Kinnmark [38], the condition necessary for the solution of the GWCE-CME system of equations to be the same as the solution of the primitive form is that

\( CE \bigg|_{t=0} = 0 \), that is, \( \xi_1(x) = -\nabla \cdot U_0 \), where \( U_0 = u_0 H_0 = u_0 (h_b + \xi_0) \).

Before proceeding, we need to list some assumptions. For \((x, t) \in \bar{\Omega} \times (0, T)\),

**M1.** \( \exists \) positive constants \( H_- \) and \( H^* \) such that \( H_- \leq H(x, t) \leq H^* \),

**M2.** \( \tau_o \) is positive,

**M3.** \( \mu \) is positive,

**M4.** \( f_c, \nabla p_a(x, t), \nabla N(x, t), \tau_{ws}(x, t) \) are bounded.
A weak form of this system is as follows. For \( t \in (0, T] \), find \( \xi(x, t) \in \mathcal{H}_0^1(\Omega) \) and \( U(x, t) \in \mathcal{H}_0^1(\Omega) \) such that

\[
\left( \frac{\partial^2 \xi}{\partial t^2}, v \right) + \tau_0 \left( \frac{\partial \xi}{\partial t}, v \right) + \left( \nabla \cdot \left( \frac{U^2}{H} \right) , \nabla v \right) + (gh_b \nabla \xi, \nabla v) + \mu \left( \nabla \frac{\partial \xi}{\partial t}, \nabla v \right) \\
+ ((\tau_0 - \tau_0) U, \nabla v) + (I, \nabla v) = 0, \quad \forall v \in \mathcal{H}_0^1(\Omega), \quad t > 0. \tag{4.4}
\]

\[
\left( \frac{\partial U}{\partial t}, w \right) + \left( \nabla \cdot \left( \frac{U^2}{H} \right) , w \right) + (gh_b \nabla \xi, w) + \mu (\nabla U, \nabla w) \\
+ (\tau_0 U, w) + (I, w) = 0, \quad \forall w \in \mathcal{H}_0^1(\Omega), \quad t > 0, \tag{4.5}
\]

with initial conditions

\[
\begin{align*}
(\xi(x, 0), v) &= (\xi_0(x), v), \quad \forall v \in \mathcal{H}_0^1(\Omega), \\
(\frac{\partial \xi}{\partial t}(x, 0), v) &= (\xi_1(x), v), \quad \forall v \in \mathcal{H}_0^1(\Omega), \\
(U(x, 0), w) &= (U_0(x), w), \quad \forall w \in \mathcal{H}_0^1(\Omega),
\end{align*}
\tag{4.6}
\]

Until recently, there were no a priori error estimates for the Galerkin finite element model of the nonlinear wave formulation. In 1996, Chippada, Dawson, Martínez, and Wheeler [16] were able to derive a priori error estimates for the nonlinear W-SWE. We detail these results here.

Let us assume for \((x, t) \in \tilde{\Omega} \times (0, T]\)

\textbf{M5.} the solutions \((\xi, U)\) to (4.4)-(4.6) exist and are unique.

\textbf{M6.} \(\xi_0(x), \xi_1(x) \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega),\)

\textbf{M7.} \(U_0(x) \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega),\)

\textbf{M8.} \(H(x, \cdot) \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega) \cap \mathcal{W}_\infty^1(\Omega), \quad t \in (0, T),\)

\textbf{M9.} \(U(x, \cdot) \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega) \cap \mathcal{W}_\infty^1(\Omega), \quad t \in (0, T).\)

Moreover, from definition, we have
D1. ∃ positive constants $\gamma_*, \gamma^*$ such that $0 < \gamma_* \leq gh_b(x) \leq \gamma^*$,

and, together with assumptions M1, M9, we have

D2. ∃ non-negative constants $\tau_{b*}, \tau_{b^*}$ such that $0 \leq \tau_{b*} \leq \tau_{b(x,t)} \leq \tau_{b^*}$.

4.2.2. The Continuous-Time Galerkin Approximation.

Let $H_h(x,t) = \xi_h(x,t) + h_b(x)$. We define the continuous-time Galerkin approximations to $\xi, U$ (satisfying (4.4)-(4.5), (4.6)) to be the mappings $\xi_h(x,t) \in S^0_h, U_h(x,t) \in S^0_h$ for each $t > 0$ satisfying

$$
\left( \frac{\partial^2 \xi_h}{\partial t^2}, v \right) + \tau_o \left( \frac{\partial \xi_h}{\partial t}, v \right) + \left( \nabla \cdot \left( \frac{U_h^2}{H_h} \right), \nabla v \right) + (gh_b \nabla \xi_h, \nabla v) \\
+ \mu \left( \nabla \frac{\partial \xi_h}{\partial t}, \nabla v \right) + ((\tau_{b*} - \tau_o)U_h, \nabla v) + (T_h, \nabla v) = 0, \quad \forall v \in S^0_h; \tag{4.7}
$$

$$
\left( \frac{\partial U_h}{\partial t}, w \right) + \left( \nabla \cdot \left( \frac{U_h^2}{H_h} \right), w \right) + (gh_b \nabla \xi_h, w) + \mu (\nabla U_h, \nabla w) \\
+ (\tau_{b*}U_h, w) + (T_h, w) = 0, \quad \forall w \in S^0_h, \tag{4.8}
$$

with initial conditions

$$
\begin{align*}
(\xi_h(x,0), v) &= (\xi_0(x), v), \quad \forall v \in S^0_h, \\
(\frac{\partial \xi_h(x,0)}{\partial t}, v) &= (\xi_1(x), v), \quad \forall v \in S^0_h, \\
(U_h(x,0), w) &= (U_0(x), w), \quad \forall w \in S^0_h,
\end{align*}
\tag{4.9}
$$

For simplicity, $\mu$ and the forcing terms in $T, T_h$ are taken to be given data.

We suppose ∃ positive constants $K_{**} \leq \frac{H_s}{2}$ and $K^{**} \geq 2K^*$ such that

N1. $K_{**} \leq H_h(x,t) \leq K^{**},$

N2. $\|U_h\|_{C^\infty((0,T); C^\infty(\Omega))} \leq K^{**},$

N3. for $\tau_{b*}(x,t) = c_f \frac{\sqrt{u_{h*}^2 + v_{h*}^2}}{H_h},$ ∃ non-negative constants $\tau_{b*, \tau_{b^*}}$ such that

$$
0 \leq \tau_{b*} \leq \tau_{b*(x,t)} \leq \tau_{b*}.
$$
4.2.3. Error Equations.

Given projections $\hat{\xi} \in S^0_h$ and $\hat{U} \in S^0_h$, we denote the affine errors in elevation and velocity as

$$\psi_\xi = (\xi - \hat{\xi}) \quad \text{and} \quad \psi_U = (U - \hat{U}),$$

respectively; and we also denote the approximation errors in elevation and velocity as

$$\theta_\xi = (\xi - \hat{\xi}) \quad \text{and} \quad \theta_U = (U - \hat{U}),$$

respectively.

Subtracting (4.4)–(4.6) from (4.7)–(4.9) yields the error equations as follows.

$$
\left( \frac{\partial^2 \psi_\xi}{\partial t^2}, v \right) + \tau_0 \left( \frac{\partial \psi_\xi}{\partial t}, v \right) + (gh_b \nabla \psi_\xi, \nabla v) + \mu \left( \nabla \cdot \left( \nabla \psi_\xi \right), \nabla v \right) \\
= \left( \frac{\partial^2 \theta_\xi}{\partial t^2}, v \right) + \tau_0 \left( \frac{\partial \theta_\xi}{\partial t}, v \right) + \left( \nabla \cdot \left( \left( \frac{U^2}{H} \right) - \left( \frac{U^2_h}{H_h} \right) \right), \nabla v \right) \\
+ (gh_b \nabla \theta_\xi, \nabla v) + \mu \left( \nabla \cdot \left( \nabla \theta_\xi \right), \nabla v \right) + ((\tau_b U - \tau_h U_h) - \tau_0 (U - U_h), \nabla v) \\
+ (T - T_h, \nabla v), \quad \forall v \in S^0_h.
$$

and

$$\psi_\xi(\cdot, 0) = \frac{\partial \psi_\xi}{\partial t}(\cdot, 0) = 0. \tag{4.10}$$

Additionally,

$$
\left( \frac{\partial \psi_U}{\partial t}, w \right) + \mu \left( \nabla \psi_U, \nabla w \right) + (\tau_h \psi_U, w) \\
= \left( \frac{\partial \theta_U}{\partial t}, w \right) + \left( \nabla \cdot \left( \left( \frac{U^2}{H} \right) - \left( \frac{U^2_h}{H_h} \right) \right), w \right) + (gh_b (\theta_\xi - \psi_\xi), w) \\
+ \mu \left( \nabla \theta_U, \nabla w \right) + (\tau_b U - \tau_h \hat{U}, w) + (T - T_h, w), \quad \forall w \in S^0_h, \tag{4.12}
$$

and

$$\psi_U(\cdot, 0) = 0. \tag{4.13}$$
4.2.4. Choice of Projections.

It turns out that an elliptic projection does not improve the estimate we derive (see [16]) because of the non-linear terms in the coupled system of equations, although this should be obvious from close examination of the error equations (4.10)-(4.12). Therefore, we shall use, as comparison functions, the $L^2$ projections $\hat{\xi} \in \mathcal{S}_h^0$ and $\hat{U} \in \mathcal{S}_h^0$ defined as in (3.2) with properties summarized in Lemma 3.2. Notice then that the terms $\left( \left( \frac{\partial}{\partial t} \right)^j \theta_\xi, v \right), \quad j = 1, 2$ in (4.10) and the term $\left( \frac{\partial \theta U}{\partial t}, w \right)$ in (4.12) vanish due to the definition of the projections.

4.2.5. Choice of Test Functions.

Let $r$ be a positive constant to be chosen. Let $v = v_1 + v_2$ be the test function in (4.10), where

$$v_1(\cdot, t) = \int_t^T e^{-rt} \psi_\xi(\cdot, t) \, dt \quad \text{and} \quad v_2(\cdot, t) = \int_t^T e^{-rt} \frac{\partial}{\partial t} \psi_\xi(\cdot, t) \, dt.$$  

These test functions are not standard. However, a version of these test functions (without the exponential decay weight) has been used in the derivation of error estimates for finite element methods for second-order hyperbolic by Baker [7] and Cowsar, Dupont, Wheeler [20]. We will also choose $w = \psi_{U}$ as the test function in (4.12).

First, we will investigate the use of $v_1$ and $v_2$ as the test functions in (4.10). Note that $v_1(\cdot, T) = 0$, $v_2(\cdot, T) = 0$,

$$\frac{\partial v_1}{\partial t} = -e^{-rt} \psi_\xi(\cdot, t), \quad \text{and} \quad \frac{\partial v_2}{\partial t} = -e^{-rt} \frac{\partial}{\partial t} \psi_\xi(\cdot, t).$$

Now, consider the first term of (4.10). When $v = v_1$, we obtain, upon integrating over $(0, T]$ and recalling equations (4.11),(4.13),

$$\int_0^T \left( \frac{\partial^2 \psi_\xi}{\partial t^2}, v_1 \right) dt \quad \stackrel{\text{IBP(T)}}{=} \quad -\int_0^T \left( \frac{\partial \psi_\xi}{\partial t}, \frac{\partial v_1}{\partial t} \right) dt + \left. \left( \frac{\partial \psi_\xi}{\partial t}, v_1 \right) \right|_0^T$$
\[
= \int_0^T e^{-rt} \left( \frac{\partial \psi_x}{\partial t}, \psi_x \right) dt \\
= \frac{1}{2} \int_0^T \frac{\partial}{\partial t} \left( e^{-rt} \|\psi_x\|^2 \right) dt + \frac{r}{2} \int_0^T e^{-rt} \|\psi_x\|^2 dt \\
= \frac{1}{2} e^{-rT} \|\psi_x(\cdot, T)\|^2 + \frac{r}{2} \int_0^T e^{-rt} \|\psi_x\|^2 dt.
\]

Similarly, using \( v = v_2 \), we obtain, upon integrating over \((0, T)\),
\[
\int_0^T \left( \frac{\partial^2 \psi_x}{\partial t^2}, v_2 \right) dt = -\int_0^T \left( \frac{\partial \psi_x}{\partial t}, \frac{\partial v_2}{\partial t} \right) dt = \int_0^T e^{-rt} \left\| \frac{\partial \psi_x}{\partial t} \right\|^2 dt.
\]

Now, consider the second term of (4.10). When \( v = v_1 \), we obtain, upon integrating over \((0, T)\)
\[
\tau_\circ \int_0^T \left( \frac{\partial \psi_x}{\partial t}, v_1 \right) dt = \tau_\circ \int_0^T \left( \psi_x, \frac{\partial v_1}{\partial t} \right) dt = \tau_\circ \int_0^T e^{-rt} \|\psi_x\|^2 dt.
\]

When \( v = v_2 \), we obtain, upon integrating over \((0, T)\)
\[
\tau_\circ \int_0^T \left( \frac{\partial \psi_x}{\partial t}, v_2 \right) dt = \tau_\circ \int_0^T \left( \psi_x, \frac{\partial v_2}{\partial t} \right) dt = \tau_\circ \int_0^T e^{-rt} \left( \psi_x, \frac{\partial \psi_x}{\partial t} \right) dt \\
= \frac{\tau_\circ}{2} e^{-rT} \|\psi_x(\cdot, T)\|^2 + \tau_\circ \frac{r}{2} \int_0^T e^{-rt} \|\psi_x\|^2 dt.
\]

Now, consider the diffusion term in (4.10). Integrating in time over \((0, T)\) yields, when \( v = v_2 \),
\[
\mu \int_0^T \left( \nabla \frac{\partial \psi_x}{\partial t}, \nabla v_1 \right) dt = \mu \int_0^T e^{-rt} \|\nabla \psi_x\|^2 dt.
\]
and when \( v = v_2 \),
\[
\mu \int_0^T \left( \nabla \frac{\partial \psi_x}{\partial t}, \nabla v_2 \right) dt = -\frac{\mu}{2} \int_0^T e^{-rt} \frac{d}{dt} (\nabla v_2, \nabla v_2) dt \\
= -\frac{\mu}{2} \int_0^T \frac{d}{dt} \left( e^{-rt} \|\nabla v_2\|^2 \right) dt + \frac{r \mu}{2} \int_0^T e^{-rt} \|\nabla v_2\|^2 dt \\
= \frac{\mu}{2} \|\nabla v_2(\cdot, 0)\|^2 + \frac{r \mu}{2} \int_0^T e^{-rt} \|\nabla v_2\|^2 dt.
\]

Finally, we are also able to manipulate the third term when \( v = v_1 \):
\[
\int_0^T (g h_b \nabla \psi_x, \nabla v_1) dt = -\frac{1}{2} \int_0^T e^{-rt} \frac{d}{dt} (g h_b \nabla \psi_x, \nabla v_1) dt
\]
\[
\begin{align*}
&= -\frac{1}{2} \int_0^T \frac{d}{dt} \left( e^{rt} \left\| \sqrt{gh_b} \nabla v_1 \right\|^2 \right) dt + \frac{r}{2} \int_0^T e^{rt} \left\| \sqrt{gh_b} \nabla v_1 \right\|^2 dt \\
&\geq \frac{\tau_s}{2} \left\| \nabla v_1(\cdot, 0) \right\|^2 + \frac{r\tau_s}{2} \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt.
\end{align*}
\]

The temporal integration of terms \( \left( \frac{\partial \psi_U}{\partial t}, \psi_U \right), (\tau_b \psi_U, \psi_U), \mu (\nabla \psi_U, \nabla \psi_U) \) in (4.12) are straightforward.

### 4.2.6. Bounding the GWCE and CME Error Equation.

Using \( v = v_1 + v_2 \) in (4.10) and \( w = \psi_U \) in (4.12), integrating in time over \( (0, T) \), and summing the resulting equations yields

\[
\begin{align*}
&\left( \frac{\tau_0 + 1}{2} \right) e^{-rt} \left\| \psi_\xi(\cdot, T) \right\|^2 + \left( \frac{r(\tau_0 + 1)}{2} + \tau_0 \right) \int_0^T e^{-rt} \left\| \psi_\xi \right\|^2 dt \\
&+ \int_0^T e^{-rt} \left\| \frac{\partial \psi_\xi}{\partial t} \right\|^2 dt + \mu \int_0^T e^{-rt} \left\| \nabla \psi_\xi \right\|^2 dt \\
&+ \frac{\tau_s}{2} \left\| \nabla v_1(\cdot, 0) \right\|^2 + \frac{r\tau_s}{2} \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt \\
&+ \frac{\mu}{2} \left\| \nabla v_2(\cdot, 0) \right\|^2 + \frac{r\mu}{2} \int_0^T e^{rt} \left\| \nabla v_2 \right\|^2 dt \\
&+ \frac{1}{2} \left\| \psi_U(\cdot, T) \right\|^2 + \mu \left\| \nabla \psi_U \right\|^2_{C^2((0,T);C^2(\Omega))} + \left\| \sqrt{\tau_b} \psi_U \right\|^2_{C^2((0,T);C^2(\Omega))} \\
&\leq \int_0^T \left( \nabla \cdot \left( \left( \frac{U^2}{H} \right) - \left( \frac{U_h^2}{H_h} \right) \right), \nabla (v_1 + v_2) + \psi_U \right) dt \\
&+ \int_0^T (gh_b \nabla \theta_\xi, \nabla (v_1 + v_2) + \psi_U) dt \\
&- \int_0^T (gh_b \nabla \psi_\xi, \nabla v_2 + \psi_U) dt \\
&+ \mu \int_0^T \left( \nabla \frac{\partial \xi}{\partial t}, \nabla (v_1 + v_2) \right) dt \\
&+ \mu \int_0^T (\nabla \theta_U, \nabla \psi_U) dt \\
&+ \int_0^T ((\tau_b - \tau_b U_h) - \tau_b (U - U_h), \nabla (v_1 + v_2)) dt \\
&+ \int_0^T (\tau_b U - \tau_b \tilde{U}, \psi_U) dt \\
&+ \int_0^T (T - T_h, \nabla (v_1 + v_2) + \psi_U) dt.
\end{align*}
\]
\[ = \mathcal{E}_1 + \cdots + \mathcal{E}_8. \quad (4.14) \]

It will be implicitly understood that we use either the the Hölder Inequality or the Arithmetic Geometric Mean Inequality or both when we bound the terms \( \mathcal{E}_i, i = 1, \ldots, 8 \). However, we will explicitly mention additional justifications in the derivation of bounds for these terms.

Term \( \mathcal{E}_1 \) is a bit complicated. To bound it above, we will employ the following expansion of the advective term

\[
\nabla \cdot \left( \frac{U^2}{H} \right) = \left( \frac{U}{H} \cdot \nabla U \right) + (\nabla \cdot U) \frac{U}{H} - (\nabla H \cdot U) \frac{U}{H^2}
\]

so that

\[
\left( \frac{U}{H} \cdot \nabla U \right) - \left( \frac{U_h}{H_h} \cdot \nabla U_h \right)
\]

\[
= \left( \frac{U}{H} \cdot \nabla (U - \bar{U}) \right) - \left( \frac{U_h}{H_h} \cdot \nabla (U_h - \bar{U}) \right) + \left[ \frac{U}{H} - \frac{U_h}{H_h} \right] \cdot \nabla \bar{U};
\]

\[
(\nabla \cdot U) \frac{U}{H} - (\nabla \cdot U_h) \frac{U_h}{H_h}
\]

\[
= \left( \nabla \cdot (U - \bar{U}) \right) \frac{U}{H} - \left( \nabla \cdot (U_h - \bar{U}) \right) \frac{U_h}{H_h} + \left( \nabla \cdot \bar{U} \right) \left[ \frac{U}{H} - \frac{U_h}{H_h} \right];
\]

and

\[
(\nabla \xi \cdot U) \frac{U}{H^2} - (\nabla \xi_h \cdot U_h) \frac{U_h}{H_h^2}
\]

\[
= \nabla (\xi - \bar{\xi}) \cdot U \frac{U}{H^2} - \nabla (\xi_h - \bar{\xi}_h) \cdot U_h \frac{U_h}{H_h^2} + \nabla \bar{\xi} \cdot U \frac{U}{H^2} - \nabla \bar{\xi} \cdot U_h \frac{U_h}{H_h^2}.
\]

Therefore,

\[
\mathcal{E}_1 = \int_0^T \left( \nabla \cdot \left( \left( \frac{U^2}{H} \right) - \left( \frac{U_h^2}{H_h} \right) \right), \nabla (v_1 + v_2) + \psi_U \right) \, dt
\]

\[
= \int_0^T \left( \frac{U}{H} \cdot \nabla \theta U, \nabla (v_1 + v_2) + \psi_U \right) \, dt
\]

\[
- \int_0^T \left( \frac{U_h}{H_h} \cdot \nabla \psi U, \nabla (v_1 + v_2) + \psi_U \right) \, dt
\]
\[ + \int_0^T \left( \left[ \frac{U}{H} - \frac{U_h}{H_h} \right] \cdot \nabla \tilde{U}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
+ \int_0^T \left( \nabla \cdot \theta U \right) \frac{U}{H}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
- \int_0^T \left( \nabla \cdot \psi U \right) \frac{U_h}{H_h}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
+ \int_0^T \left( \nabla \cdot \tilde{U} \right) \left[ \frac{U}{H} - \frac{U_h}{H_h} \right], \nabla (v_1 + v_2) + \psi_U \right) dt \\
- \int_0^T \left( \nabla \theta \xi \cdot U \right) \frac{U}{H^2}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
+ \int_0^T \left( \nabla \psi \xi \cdot U_h \right) \frac{U_h}{H_h^2}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
- \int_0^T \nabla \xi \cdot \left\{ \left[ \frac{U}{H} \right]^2 - \left[ \frac{U_h}{H_h} \right]^2 \right\}, \nabla (v_1 + v_2) + \psi_U \right) dt \\
- \int_0^T \nabla h \cdot \left\{ \left[ \frac{U}{H} \right]^2 - \left[ \frac{U_h}{H_h} \right]^2 \right\}, \nabla (v_1 + v_2) + \psi_U \right) dt \]

= A_1 + \cdots + A_{10}.

Now lets look at each $A_i, i = 1, \ldots, 10$ individually.

From assumptions $\textbf{M1, M9}$, we obtain the following upper bound on the affine error terms

\[ A_1 \leq K \left\| \nabla \theta U \right\|^2_{L^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt \\
+ K \int_0^T e^{rt} \left\| \nabla v_2 \right\|^2 dt + K \left\| \psi U \right\|^2_{L^2((0,T);L^2(\Omega))}; \]

\[ A_4 \leq K \left\| \nabla \theta \xi \right\|^2_{L^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt \\
+ K \int_0^T e^{rt} \left\| \nabla v_2 \right\|^2 dt + K \left\| \psi U \right\|^2_{L^2((0,T);L^2(\Omega))}; \]

\[ A_7 \leq K \left\| \nabla \xi \right\|^2_{L^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt \\
+ K \int_0^T e^{rt} \left\| \nabla v_2 \right\|^2 dt + K \left\| \psi U \right\|^2_{L^2((0,T);L^2(\Omega))}. \]

Using assumptions $\textbf{M3, N1, N2}$, yields

\[ A_2 \leq c \left\| \nabla \psi U \right\|^2_{L^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \left\| \nabla v_1 \right\|^2 dt \]
\[ A_5 \leq \epsilon \| \nabla U \|_{L^2((0,T);C^2(\Omega))}^2 + K \int_0^T e^{-rt} \| \nabla \psi_2 \|_{L^2((0,T);C^2(\Omega))}^2 dt + K \int_0^T e^{-rt} \| \nabla \psi_1 \|_{L^2((0,T);C^2(\Omega))}^2 dt \]

\[ A_6 \leq \epsilon \int_0^T e^{-rt} \| \nabla U \|_{L^2((0,T);C^2(\Omega))}^2 dt + K \int_0^T e^{-rt} \| \nabla \psi_1 \|_{L^2((0,T);C^2(\Omega))}^2 dt \]

To bound the remaining \( A_i \) terms, we will need the following lemma.

**Lemma 4.1** Let assumptions M1, M9, N1 hold. Then, there exists constants \( K_1(H_*, K_*) \), \( K_2(K_*) \) such that

\[
\left\| \frac{U}{H} - \frac{U_h}{H_h} \right\| \leq K_1 \left( \| \theta \| + \| \psi \| \right) + K_2 \left( \| \theta U \| + \| \psi U \| \right).
\]

**Proof:**

\[
\left\| \frac{U}{H} - \frac{U_h}{H_h} \right\| = \left\| \frac{U (H_h - H) + (U - U_h) H}{H H_h} \right\|
\leq \left\| \frac{U}{H H_h} \right\|_{C^\infty(\Omega)} \| H_h - H \| + \left\| \frac{1}{H_h} \right\|_{C^\infty(\Omega)} \| U - U_h \|
= K_1 \| \xi - \xi \| + K_2 \| U - U_h \|.
\]

Assumptions M1, M9, N1 are used to get the first part of the inequality and assumption N1 is used to get the second part of the inequality. \( \square \)

From Lemmas 4.1 and 3.3, assumptions M1, M2, M9, N1, we have

\[ A_3 \leq \epsilon \int_0^T e^{-rt} \| \psi \|_{L^2((0,T);C^2(\Omega))}^2 dt + K \| \theta \|_{L^2((0,T);C^2(\Omega))}^2 + K \| \theta U \|_{L^2((0,T);C^2(\Omega))}^2 + K \| \psi \|_{L^2((0,T);C^2(\Omega))}^2 \]

\[ A_6 \leq \epsilon \int_0^T e^{-rt} \| \psi \|_{L^2((0,T);C^2(\Omega))}^2 dt + K \| \theta \|_{L^2((0,T);C^2(\Omega))}^2 + K \| \theta U \|_{L^2((0,T);C^2(\Omega))}^2 + K \| \psi U \|_{L^2((0,T);C^2(\Omega))}^2 \]
From Lemmas 4.1 and 3.3, and assumptions M1, M2, M9, N1, N2,

\[
A_9 = - \int_0^T \left( \nabla \xi \cdot \left\{ \left[ \frac{U}{H} - \frac{U_h}{H_h} \right] \left[ \frac{U}{H} + \frac{U_h}{H_h} \right] + \left[ \frac{U_h U - U U_h}{H H_h} \right] \right\} \nabla (v_1 + v_2) + \psi U \right) dt \\
\leq \epsilon \int_0^T e^{-rt} ||\psi_\xi||^2dt + K ||\theta_\xi||_{L^2((0,T);L^2(\Omega))}^2 + K ||\theta U||_{L^2((0,T);L^2(\Omega))}^2 \\
+ K \int_0^T e^{rt} ||\nabla v_1||^2dt + K \int_0^T e^{rt} ||\nabla v_2||^2dt + K ||\psi U||_{L^2((0,T);L^2(\Omega))}^2.
\]

Finally, from Lemmas 4.1 and 3.3 and assumptions M1, M2, M8, M9, N1, N2, we have

\[
A_{10} = - \int_0^T \left( \nabla h_b \cdot \left\{ \left[ \frac{U}{H} - \frac{U_h}{H_h} \right] \left[ \frac{U}{H} + \frac{U_h}{H_h} \right] + \left[ \frac{U_h U - U U_h}{H H_h} \right] \right\} \nabla (v_1 + v_2) + \psi U \right) dt \\
\leq \epsilon \int_0^T e^{-rt} ||\psi_\xi||^2dt + K ||\theta_\xi||_{L^2((0,T);L^2(\Omega))}^2 + K ||\theta U||_{L^2((0,T);L^2(\Omega))}^2 \\
+ K \int_0^T e^{rt} ||\nabla v_1||^2dt + K \int_0^T e^{rt} ||\nabla v_2||^2dt + K ||\psi U||_{L^2((0,T);L^2(\Omega))}^2.
\]

Therefore, after substituting the upper bounds of \( A_i, i = \{1, \ldots, 10\} \) back into \( E_1 \), we get

\[
E_1 \leq \epsilon \int_0^T e^{-rt} ||\psi_\xi||^2dt + \epsilon \int_0^T e^{-rt} ||\nabla \psi_\xi||^2dt + \epsilon ||\nabla \psi U||_{L^2((0,T);L^2(\Omega))}^2 \\
+ K ||\theta_\xi||_{L^2((0,T);H^1(\Omega))}^2 + K ||\theta U||_{L^2((0,T);H^1(\Omega))}^2 \\
+ K \int_0^T e^{rt} ||\nabla v_1||^2dt + K \int_0^T e^{rt} ||\nabla v_2||^2dt + K ||\psi U||_{L^2((0,T);L^2(\Omega))}^2.
\]

The bound on \( E_2 \) is straightforward. From definition D1,

\[
E_2 = \int_0^T (gh_b \nabla \theta_\xi, \nabla (v_1 + v_2) + \psi U) dt \\
\leq K ||\nabla \theta_\xi||_{L^2((0,T);L^2(\Omega))}^2 + K \int_0^T e^{rt} ||\nabla v_1||^2dt \\
+ K \int_0^T e^{rt} ||\nabla v_2||^2dt + K ||\psi U||_{L^2((0,T);L^2(\Omega))}^2.
\]

The bound on \( E_3 \) is also straightforward. Using definition D1 and assumption M3,

\[
E_3 = - \int_0^T (gh_b \nabla \psi_\xi, \nabla v_2 + \psi U) dt
\]
\[
\leq \epsilon \int_0^T e^{-rt} \|\nabla \psi_1\|^2 dt + K \int_0^T e^{rt} \|\nabla v_2\|^2 dt + K \|\psi_U\|^2_{C^2((0,T);L^2(\Omega))}.
\]

The bound on \(\mathcal{E}_4\) is straightforward.

\[
\mathcal{E}_4 = \mu \int_0^T \left( \nabla \frac{\partial \theta_{\xi}}{\partial t}, \nabla (v_1 + v_2) \right) dt \leq K \left\| \nabla \frac{\partial \theta_{\xi}}{\partial t} \right\|^2_{L^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \|\nabla v_1\|^2 dt + K \int_0^T e^{rt} \|\nabla v_2\|^2 dt.
\]

From assumption M3, we have

\[
\mathcal{E}_5 = \mu \int_0^T (\nabla \theta_U, \nabla \psi_U) dt \leq \epsilon \|\nabla \psi_U\|^2_{C^2((0,T);L^2(\Omega))} + K \|\theta_U\|^2_{C^2((0,T);L^2(\Omega))}.
\]

To bound \(\mathcal{E}_6 + \mathcal{E}_7\), first observe that

\[
\tau_b U - \tau_{b_h} U_h = \tau_b(U - U) - \tau_{b_h}(U_h - U) + (\tau_b - \tau_{b_h})\tilde{U}
\]

and

\[
\tau_b U - \tau_{b_h} \tilde{U} = \tau_b(U - \tilde{U}) + (\tau_b - \tau_{b_h})\tilde{U}.
\]

Therefore, we can write \(\mathcal{E}_6 + \mathcal{E}_7\) as

\[
\mathcal{E}_6 + \mathcal{E}_7 = \int_0^T ((\tau_b U - \tau_{b_h} U_h) - \tau_o (U - U_h), \nabla (v_1 + v_2)) dt + \int_0^T (\tau_b U - \tau_{b_h} \tilde{U}, \psi_U) dt
\]

\[
= \int_0^T (\tau_b \theta_U, \nabla (v_1 + v_2) + \psi_U) dt - \int_0^T (\tau_o (\theta_U - \psi_U), \nabla (v_1 + v_2)) dt
\]

\[
- \int_0^T (\tau_{b_h} \psi_U, \nabla (v_1 + v_2)) dt + \int_0^T ((\tau_b - \tau_{b_h})\tilde{U}, \nabla (v_1 + v_2) + \psi_U) dt
\]

\[
= B_1 + B_2 + B_3 + B_4.
\]

The bound on \(B_1\) follows from definition D2,

\[
B_1 \leq K \|\theta_U\|^2_{C^2((0,T);L^2(\Omega))} + K \int_0^T e^{rt} \|\nabla v_1\|^2 dt + K \|\psi_U\|^2_{C^2((0,T);L^2(\Omega))}.
\]
The bound on \( B_2 \) is straightforward.

\[
B_2 \leq K \| \theta U \|_{L^2((0,T);C^2(\Omega))}^2 + K \int_0^T e^{rt} \| \nabla v_1 \|^2 dt + K \int_0^T e^{rt} \| \nabla v_2 \|^2 dt + K \| \psi U \|_{L^2((0,T);C^2(\Omega))}^2.
\]

The bound on \( B_3 \) follows from assumption N3.

\[
B_3 \leq K \int_0^T e^{rt} \| \nabla v_1 \|^2 dt + K \int_0^T e^{rt} \| \nabla v_2 \|^2 dt + K \| \psi U \|_{L^2((0,T);C^2(\Omega))}^2.
\]

And, using Lemmas 4.1 and 3.3 the definitions of \( \tau_b \) and \( \tau_{bh} \) as well as assumptions M1, M2, M9, N1, N2,

\[
B_4 = \left( (\tau_b - \tau_{bh}) U, \nabla (v_1 + v_2) + \psi U \right) dt
\]

\[
= c_f \left( \left[ \frac{||U/H||_{L^2}}{H} - \frac{||U_h/H_h||_{L^2}}{H_h} \right] U, \nabla (v_1 + v_2) + \psi U \right) dt
\]

\[
= c_f \left( \left[ \frac{H_h ||U/H||_{L^2} - H ||U_h/H_h||_{L^2}}{H H_h} \right] U, \nabla (v_1 + v_2) + \psi U \right) dt
\]

\[
= c_f \left( \left[ \frac{H_h (||U/H||_{L^2} - ||U_h/H_h||_{L^2})}{H H_h} \right] U, \nabla (v_1 + v_2) + \psi U \right) dt
\]

\[
\leq c_f \left( \left[ \frac{(\psi - \theta \xi) ||U||_{L^2}}{H^2 H_h} + \frac{|\theta U||_{L^2} + ||\psi U||_{L^2}}{H H_h} \right] \nabla (v_1 + v_2) + \psi U \right) dt
\]

\[
\leq C \int_0^T e^{-rt} ||\psi||_2^2 dt + K ||\phi||_{L^2((0,T);C^2(\Omega))}^2 + K ||\theta U||_{L^2((0,T);C^2(\Omega))}^2
\]

\[
+ K \int_0^T e^{rt} \| \nabla v_1 \|^2 dt + K \int_0^T e^{rt} \| \nabla v_2 \|^2 dt + K \| \psi U \|_{L^2((0,T);C^2(\Omega))}^2.
\]

Therefore, after substituting the upper bounds of \( B_i, i = 1, \ldots, 4 \), back into \( \mathcal{E}_6 + \mathcal{E}_7 \), we obtain

\[
\mathcal{E}_6 + \mathcal{E}_7 \leq \epsilon \int_0^T e^{-rt} ||\psi||_2^2 dt + K ||\phi||_{L^2((0,T);C^2(\Omega))}^2 + K ||\theta U||_{L^2((0,T);C^2(\Omega))}^2
\]

\[
+ K \int_0^T e^{rt} \| \nabla v_1 \|^2 dt + K \int_0^T e^{rt} \| \nabla v_2 \|^2 dt + K \| \psi U \|_{L^2((0,T);C^2(\Omega))}^2.
\]
Continuing, we have

\[ \mathcal{E}_s = \int_0^T (\mathbf{T} - \mathbf{T}_h, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ = \int_0^T (g \xi \nabla \xi - g \xi_h \nabla \xi_h, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ + \int_0^T (f_c \mathbf{k} \times (\mathbf{U} - \mathbf{U}_h), \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ + \int_0^T (H \mathcal{F} - H_h \mathcal{F}_h, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ = C_1 + C_2 + C_3. \]

Now, from assumptions \textbf{M1}, \textbf{M2}, \textbf{M3}, \textbf{N1}, and Lemma 3.3,

\[ C_1 = \int_0^T (g \xi \nabla \xi - g \xi_h \nabla \xi_h, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ = \int_0^T (g \xi \nabla (\xi - \tilde{\xi}) - g \xi_h \nabla (\xi_h - \tilde{\xi}) + g (\xi - \xi_h) \nabla \tilde{\xi}, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ = \int_0^T (g \xi \nabla \theta \xi - g \xi_h \nabla \psi \xi + g \theta \xi \nabla \tilde{\xi} - g \psi \xi \nabla \tilde{\xi}, \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ \leq \epsilon \int_0^T e^{-rt} ||\psi \xi||^2 \, dt + \epsilon \int_0^T e^{-rt} ||\nabla \psi \xi||^2 \, dt + K' \left( ||\theta \xi||_{L^2((0,T); H^1(\Omega))}^2 + K' \left( ||\psi \xi||_{L^2((0,T); H^1(\Omega))}^2 \right) \right). \]

From definition \textbf{D4},

\[ C_2 = \int_0^T (f_c \mathbf{k} \times (\mathbf{U} - \mathbf{U}_h), \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ \leq K' \left( ||\theta \mathbf{U}||_{L^2((0,T); H^1(\Omega))}^2 + K \left( \int_0^T e^{rt} ||\nabla v_1||^2 \, dt \right) \right) + K' \left( \int_0^T e^{rt} ||\nabla v_2||^2 \, dt + K' \left( ||\psi \mathbf{U}||_{L^2((0,T); H^1(\Omega))}^2 \right) \right). \]

From assumptions \textbf{M1}, \textbf{M2}, \textbf{M4},

\[ C_3 = \int_0^T (H - H_h) \mathcal{F} + H_h(\mathcal{F} - \mathcal{F}_h), \nabla (v_1 + v_2) + \psi_U) \, dt \]
\[ = \int_0^T \left( (\theta - \psi \xi) + (\theta \xi - \psi \xi) \frac{\tau ws}{H}, \nabla (v_1 + v_2) + \psi_U \right) \, dt \]
\[ \leq \epsilon \int_0^T e^{-rt} ||\psi \xi||^2 \, dt + K' \left( ||\theta \xi||_{L^2((0,T); H^1(\Omega))}^2 \right) \]
\[ + K' \left( \int_0^T e^{rt} ||\nabla v_1||^2 \, dt + K' \left( \int_0^T e^{rt} ||\nabla v_2||^2 \, dt + K' \left( ||\theta \mathbf{U}||_{L^2((0,T); H^1(\Omega))}^2 \right) \right) \right). \]
After substituting the upper bounds of $C_i, i = \{1, \ldots, 3\}$ back into $E_8$, we get

$$E_8 \leq \epsilon \int_0^T e^{-rt} \|\psi_\xi\|^2 dt + \epsilon \int_0^T e^{-rt} \|\nabla \psi_\xi\|^2 dt$$

$$+ K \|\theta_\xi\|^2_{L^2((0,T); H^1(\Omega))} + K \|\theta_U\|^2_{L^2((0,T); C^2(\Omega))}$$

$$+ K \int_0^T e^{rt} \|\nabla v_1\|^2 dt + K \int_0^T e^{rt} \|\nabla v_2\|^2 dt + K \|\psi_U\|^2_{L^2((0,T); C^2(\Omega))}.$$

Substituting the upper bounds on $E_i$ back into (4.14) and absorbing $\epsilon$ terms into the LHS yields

$$\left(\frac{\tau_0 + i}{2}\right) e^{-rT} \|\psi_\xi(\cdot,T)\|^2 + \left(\frac{r(\tau_0 + 1)}{2}\right) \int_0^T e^{-rt} \|\psi_\xi\|^2 dt$$

$$+ \int_0^T e^{-rt} \left\|\frac{\partial \psi_\xi}{\partial t}\right\|^2 dt + \frac{\mu}{2} \int_0^T e^{-rt} \|\nabla \psi_\xi\|^2 dt$$

$$+ \frac{\gamma_*}{2} \|\nabla v_1(\cdot,0)\|^2 + r \frac{\gamma_*}{2} \int_0^T e^{rt} \|\nabla v_1\|^2 dt$$

$$+ \frac{\mu}{2} \|\nabla v_2(\cdot,0)\|^2 + r \frac{\mu}{2} \int_0^T e^{rt} \|\nabla v_2\|^2 dt$$

$$+ \frac{1}{2} \|\psi_U(\cdot,T)\|^2 + \|\sqrt{\tau_h} \psi_U\|^2_{L^2((0,T); C^2(\Omega))} + \frac{\mu}{2} \|\nabla \psi_U\|^2_{L^2((0,T); C^2(\Omega))}$$

$$\leq K \|\theta_\xi\|^2_{L^2((0,T); H^1(\Omega))} + K \left\|\frac{\partial \theta_\xi}{\partial t}\right\|^2_{L^2((0,T); C^2(\Omega))}$$

$$+ K \|\theta_U\|^2_{L^2((0,T); H^1(\Omega))} + K v_1 \int_0^T e^{rt} \|\nabla v_1\|^2 dt$$

$$+ K v_2 \int_0^T e^{rt} \|\nabla v_2\|^2 dt + K \|\psi_U\|^2_{L^2((0,T); C^2(\Omega))}. \quad (4.15)$$

Let $r = \max \{2K v_1/\gamma_*, 2K v_2/\mu\}$, such that $r_1 = \left(r \frac{\gamma_*}{2} - K v_1\right) \geq 0$ and $r_2 = \left(r \frac{\mu}{2} - K v_2\right) \geq 0$. Then, we can rewrite (4.15) as

$$\left(\frac{\tau_0 + 1}{2}\right) e^{-rT} \|\psi_\xi(\cdot,T)\|^2 + \left(\frac{r(\tau_0 + 1)}{2}\right) \int_0^T e^{-rt} \|\psi_\xi\|^2 dt$$

$$+ \int_0^T e^{-rt} \left\|\frac{\partial \psi_\xi}{\partial t}\right\|^2 dt + \frac{\mu}{2} \int_0^T e^{-rt} \|\nabla \psi_\xi\|^2 dt$$

$$+ \frac{\gamma_*}{2} \|\nabla v_1(\cdot,0)\|^2 \underbrace{+ r_1 \int_0^T e^{rt} \|\nabla v_1\|^2 dt}_{A}$$

$$+ \frac{\gamma_*}{2} \|\nabla v_1(\cdot,0)\|^2 \underbrace{+ r_1 \int_0^T e^{rt} \|\nabla v_1\|^2 dt}_{B}$$
\[ + \frac{\mu}{2} \frac{\| \nabla v_2(\cdot, 0) \|^2 + r_2 \int_0^T e^{ \tau t} \| \nabla v_2 \|^2 dt}{C} \]
\[ + \frac{1}{2} \| \psi_U(\cdot, T) \|^2 + \frac{1}{\sqrt{\pi h}} \frac{\| \psi_U \|^2}{L^2((0,T); L^2(\Omega))} + \frac{\mu}{2} \| \nabla \psi_U \|^2_{L^2((0,T); L^2(\Omega))} \]
\[ \leq K \| \theta_\xi \|^2_{L^2((0,T); H^1(\Omega))} + K \frac{\| \nabla \theta_\xi \|^2}{L^2((0,T); L^2(\Omega))} + K \frac{\| \psi_U \|^2}{L^2((0,T); L^2(\Omega))} + K \| \theta_U \|^2_{L^2((0,T); H^1(\Omega))} \cdot (4.16) \]

Let \( \Lambda = \| \psi_U \|^2 \) in (4.16). Using IGL (with \( f = \Lambda \)), and observing that terms \( A, B, C, D \) are all non-negative yields
\[ \left( \frac{\tau_0 + 1}{2} \right) e^{-rT} \| \psi_\xi(\cdot, T) \|^2 + \left( \frac{r(\tau_0 + 1)}{2} \right) \int_0^T e^{-r t} \| \psi_\xi \|^2 dt \]
\[ + \int_0^T e^{-r t} \| \frac{\partial \psi_\xi}{\partial t} \|^2 dt + \frac{\mu}{2} \int_0^T e^{-r t} \| \nabla \psi_\xi \|^2 dt \]
\[ + \frac{1}{2} \| \psi_U(\cdot, T) \|^2 + \frac{1}{\sqrt{\pi h}} \frac{\| \psi_U \|^2}{L^2((0,T); L^2(\Omega))} + \frac{\mu}{2} \| \nabla \psi_U \|^2_{L^2((0,T); L^2(\Omega))} \]
\[ \leq e^{KT} \left( K \frac{\| \theta_\xi \|^2}{L^2((0,T); H^1(\Omega))} + K \frac{\| \nabla \theta_\xi \|^2}{L^2((0,T); L^2(\Omega))} \right) + K \| \theta_U \|^2_{L^2((0,T); H^1(\Omega))} \cdot (4.17) \]

4.2.7. A Priori Error Estimate.

It should be pointed out that the proof for the a priori error estimate is based on an argument similar to that made by Ewing, Wheeler [25] to handle nonlinearities, but using an \( L^2 \) projection instead of an elliptic projection as the comparison function. Using Ewing and Wheeler's argument, we first assume that the Galerkin approximations are bounded by some constant (via assumptions N1 and N2). But then we show that for sufficiently small \( h \), that we can remove the estimates' dependence on the assumed bound of the approximations, being dependent instead on a bound on the comparison functions and on \( H_* \). This will be the case here, but only for polynomials of degree at least two (\( s_1 \geq 3 \)). We now finish the estimate.
Use the approximation result stated in Lemma 3.2, 
\[
\|\theta_\varepsilon\|_{L^2((0,T);H^1(\Omega))}, \left\| \frac{\partial \theta_\varepsilon}{\partial t} \right\|_{L^2((0,T);C^2(\Omega))}, \|\theta_\varepsilon\|_{L^2((0,T);H^1(\Omega))} \leq K_0 h^{\ell-1},
\]
to obtain
\[
\|\psi_\varepsilon(\cdot, T)\| + \left\| \frac{\partial \psi_\varepsilon}{\partial t} \right\|_{L^2((0,T);C^2(\Omega))} + \|\psi_\varepsilon\|_{L^2((0,T);H^1(\Omega))} \\
+ \|\psi_U(\cdot, T)\| + \left\| \sqrt{\tau_h} \psi_U \right\|_{L^2((0,T);C^2(\Omega))} \\
+ \|\nabla \psi_U\|_{L^2((0,T);C^2(\Omega))} \leq K h^{\ell-1}. \tag{4.18}
\]

Finally, applying the triangle inequality to the affine error and to the approximation error yields the following error estimate
\[
\left\| \frac{\partial}{\partial t}(\xi - \xi_h) \right\|_{L^2((0,T);C^2(\Omega))} + \|\xi - \xi_h\|_{L^2((0,T);H^1(\Omega))} \\
+ \|\xi - \xi_h(\cdot, T)\| + \left\| \sqrt{\tau_h}(\xi - \xi_h) \right\|_{L^2((0,T);C^2(\Omega))} \\
+ \|\nabla \xi - \nabla \xi_h\|_{L^2((0,T);C^2(\Omega))} \leq K h^{\ell-1}.
\]

The proof of the theorem is now complete in the case of linears since we assumed that N1, N2 hold for this case.

We now establish a result for the case of at least quadratic polynomials \((s_1 \geq 3)\). From the inverse assumptions, the boundedness of the \(L^2\) projection and estimate (4.18), we obtain
\[
\|U_h\|_{L^\infty((0,T);C^\infty(\Omega))} \leq \left\| U_h - \tilde{U} \right\|_{L^\infty((0,T);C^\infty(\Omega))} + \|\tilde{U}\|_{L^\infty((0,T);C^\infty(\Omega))} \\
\leq K_0 h^{-1} \left\| U_h - \tilde{U} \right\|_{L^\infty((0,T);C^2(\Omega))} + K^* \\
\leq K_0 h^{-1} K h^{\ell-1} + K^* \\
= K h^{\ell-2} + K^*.
\]

For \(h\) sufficiently small, viz, \(h^{\ell-2} < \frac{K^*}{K}\), we get
\[ ||U_h||_{L^\infty((0,T);L^\infty(\Omega))} < 2K^* \leq K^{**}. \]

The upper bound for \( H_h(x,t) \) is shown similarly.

Use assumption M1. inverse assumptions, and estimates on the approximation and affine errors to get the lower bound for \( H_h(x,t) \). that is,

\[
H_h = H - (H - H_h) = H - \theta_\xi + \psi_\xi \\
\geq H_* - K_h h^{t-2}.
\]

For \( h \) sufficiently small, viz. \( h^{t-2} < \frac{H_*}{2K} \), we get

\[
H_h > \frac{H_*}{2} \geq K_{**}.
\]

Thus for the case of quadratics and higher, there exists a \( K \) bounded above independent of \( K^{**}, K_{**} \). This completes the development of the \textit{a priori} estimate. We summarize our results below.
Theorem 4.1 (Chippada, Dawson, Martínez, Wheeler)

Let $0 \leq s_0 \leq \ell \leq s_1$. Let $(\xi, U)$ be the solution to (4.4)-(4.6) with homogeneous Dirichlet boundary conditions. Let $(\xi_h, U_h)$ be the Galerkin approximations to $(\xi, U)$. If $\xi(t) \in H^l_0(\Omega) \cap H^l(\Omega) \cap W^l_\infty(\Omega)$, $U(t) \in H^l_0(\Omega) \cap H^l(\Omega) \cap W^l_\infty(\Omega)$ for each $t$; if $\xi_h(t) \in S_h(\Omega)$, $U_h(t) \in S_h(\Omega)$ for each $t$; and assumptions M1-M7 and N1-N3 hold with definitions D1-D2; then there exists a constant $\bar{K} = \bar{K}(T, s_1, r, H_*, K_*, K^*, K_{**})$ such that

$$
\| (\xi - \xi_h)(\cdot, T) \| + \left\| \frac{\partial}{\partial t} (\xi - \xi_h) \right\|_{L^2((0,T); L^2(\Omega))} + \| \xi - \xi_h \|_{L^2((0,T); H^l(\Omega))}
+ \| (U - U_h)(\cdot, T) \| + \left\| \sqrt{r_h} (U - U_h) \right\|_{L^2((0,T); L^2(\Omega))} + \| \nabla U - \nabla U_h \|_{L^2((0,T); L^2(\Omega))} \leq \bar{K} h^{l-1}.
$$

Moreover, for $h$ sufficiently small and $s_1 \geq 3$ then

$$
\| \xi_h \|_{L^\infty((0,T); L^\infty(\Omega))} + \| U_h \|_{L^\infty((0,T); L^\infty(\Omega))} < 2K^* \leq K^{**},
$$

and

$$
\xi_h > \frac{H_*}{2} \geq K_{**}.
$$

Thus, the dependence of $\bar{K}$ on $K_{**}, K^{**}$ (via assumptions N1-N2) is removed.

4.3. Discrete-Time Nonlinear Model.

For completeness, we derive an a priori error estimate for a discrete-time Galerkin finite element approximation to the GWCE-CME system of equations.
Assume that M1–M9 hold. Additionally, assume

**M10.** \( \frac{\partial \xi}{\partial t}(\mathbf{x}, t), \frac{\partial^2 \xi}{\partial t^2}(\mathbf{x}, t) \in \mathcal{L}^\infty((0, \Delta t); \mathcal{H}^1(\Omega)); \frac{\partial^3 \xi}{\partial t^3}(\mathbf{x}, t) \in \mathcal{L}^\infty((0, \Delta t); \mathcal{L}^2(\Omega)), \)

**M11.** \( \xi(\mathbf{x}, t) \in \mathcal{H}^1((0, T); \mathcal{H}^1(\Omega)) \cap \mathcal{H}^2((0, T); \mathcal{H}^1(\Omega)) \cap \mathcal{H}^4((0, T); \mathcal{L}^2(\Omega)), \)

**M12.** \( \mathbf{U}(\mathbf{x}, t) \in \mathcal{H}^2((0, T); \mathcal{L}^2(\Omega)) \cap \mathcal{H}^2((0, T); \mathcal{H}^1(\Omega)) . \)

Then, the discrete-time variational form of (4.1)-(4.3) can be described by (4.4) at \( t = t^k \), and (4.5) at \( t = t^k \), with initial conditions

\[
(\xi^0, v) = (\xi_0, v), \quad \forall v \in \mathcal{H}_0^1(\Omega),
\]

\[
(\xi^{-1}, v) = (\xi^1, v) - 2\Delta t(\xi_1, v), \quad \forall v \in \mathcal{H}_0^1(\Omega),
\]

\[
(U^0, w) = (U_0, w), \quad \forall w \in \mathcal{H}_0^1(\Omega).
\]

Cowser et al [20] suggest the form of the second equation in (4.19) to specially handle truncation errors for the second-order time derivative approximation at time \( t = 0 \).

Now we define

\[
\partial_t(\xi^k) = (\Delta t)^{-1}(\xi^{k+1} - \xi^k); \quad \partial_t(\xi^k) = (2\Delta t)^{-1}(\xi^{k+1} - \xi^{k-1});
\]

\[
\partial_t(\mathbf{u}^k) = (\Delta t)^{-1}(\mathbf{u}^{k+1} - \mathbf{u}^k); \quad \partial_t^2(\mathbf{u}^k) = (\Delta t)^{-2}(\mathbf{u}^{k+1} - 2\mathbf{u}^k + \mathbf{u}^{k-1});
\]

and \( \xi^{k+\frac{1}{2}} = (\xi^{k+1} + \xi^k)/2 \).

### 4.3.1. **ADCIRC Temporal Discretization.**

In **ADCIRC**, the GWCE-NCME spatial discretization is based on linear triangular elements. The temporal discretization is straightforward. Claiming improved stability in advection-dominated flow, Luettich et al [42], first transform the advective terms in the GWCE which are in conservative form into the non-conservative form using the CE, that is,

\[
\nabla \cdot (\mathbf{uuH}) = \mathbf{uH} \cdot \nabla \mathbf{u} + \mathbf{u} \nabla \cdot (\mathbf{uuH}) = \mathbf{uH} \cdot \nabla \mathbf{u} - \mathbf{u} \frac{\partial \xi}{\partial t}.
\]
(One should notice that this transformation improves the symmetry of the advective terms between the GWCE and NCME equations.) The GWCE is then discretized in time using a k-centered three-time-level implicit scheme for linear terms and treating nonlinear terms explicitly. Thus, the Galerkin finite element approximations $\xi_h^k(x) = \xi_h(x, t = t^k)$ and $u_h^k(x) = u_h(x, t = t^k)$ satisfy

$$
\left( \frac{\partial^2 \xi_h^k}{\partial t^2}, v \right) + \tau_0 \left( \partial_t \xi_h^k, v \right) + \left( u_h^k H_h^{k^\text{H}} \cdot \nabla u_h^k - u_h^k \partial_\nu \xi_h^k, \nabla v \right) \\
+ \sum_{i=-1}^1 \alpha_i \left( gh_b \nabla \xi_h^{k+i}, \nabla v \right) + \mu \left( \partial_t \nabla \xi_h^k, \nabla v \right) + \left( (\tau_h^k - \tau_0) u_h^k H_h^{k}, \nabla v \right) \\
+ \left( T_h^k, \nabla v \right) = 0, \quad \forall v \in S_h^0, t > 0,
$$

where $\sum_{i=-1}^1 \alpha_i = 1$ and $\alpha_{-1} = \alpha_1$. A global time-independent linear system results at each time step which is solved either directly or iteratively (say by a Jacobi conjugate gradient method).

A Crank-Nicolson approach (centering at $k + \frac{1}{2}$) is taken for solving the NCME except for advective terms which are treated explicitly and for the diffusive term which is variably-weighted over two time levels. Thus, $\xi^k, u_h^k$ also satisfy

$$
\left( \partial_t u_h^k, w \right) + \left( u_h^k \cdot \nabla u_h^k, w \right) + \mu \sum_{i=0}^1 \beta_i \left( \nabla (u_h H_h)^{k+i}, \frac{1}{H_h^{k+i}} \nabla w \right) \\
+ \left( \tau_h^k u_h^{k+\frac{1}{2}}, w \right) + \left( T_h^{k+\frac{1}{2}}, w \right) = 0, \quad \forall w \in S_h^0, t > 0,
$$

where $\sum_{i=0}^1 \beta_i = 1$. By lumping terms, a diagonal time-dependent linear system results at each time step, which must be solved iteratively unless $\mu = 0$ or $\beta_1 = 0$.

The corresponding temporal discretization of the GWCE-CME is

$$
\left( \frac{\partial^2 \xi_h^k}{\partial t^2}, v \right) + \tau_0 \left( \partial_t \xi_h^k, v \right) + \left( \nabla \cdot \left( \frac{(U_h^k)^2}{H_h^k} \right), \nabla v \right) \\
+ \sum_{i=-1}^1 \alpha_i \left( gh_b \nabla \xi_h^{k+i}, \nabla v \right) + \mu \left( \partial_t \nabla \xi_h^k, \nabla v \right) \\
+ \left( (\tau_h^k - \tau_0) U_h^k, \nabla v \right) + \left( T_h^k, \nabla v \right) = 0, \quad \forall v \in S_h^0, k \geq 0,
$$
\( (\partial_t U_h^k, w) + \left( \nabla \cdot \left( \frac{(U_h^k)^2}{H_h^k} \right), w \right) + (gh_b \nabla \xi_h^{k+\frac{1}{2}}, w) + \mu \sum_{i=0}^{1} \beta_i \left( \nabla U_h^{k+i}, \nabla w \right) + \left( \tau_{b_h} U_h^{k+\frac{1}{2}}, w \right) + \left( T_h^{k+\frac{1}{2}}, w \right) = 0, \quad \forall w \in S_h^0, \ k \geq 0. \)

### 4.3.2. The Discrete-Time Galerkin Approximation.

It will be clear that because of the explicit treatment of advective terms (and forcing terms) in ADCIRC, eliminating tight coupling between the GWCE and the momentum equations, the best that can be achieved is a first-order-in-time scheme.

Thus, we look at a slightly different time discretization scheme in the GWCE. We define the discrete-time Galerkin approximations to \( \xi^k(x), U^k(x) \) to be the mappings \( \xi^k_h(x) \in S_h, U^k_h(x) \in S_h \) satisfying

\[
\frac{2}{\Delta t} \left( \partial_t \xi_h^1, v \right) + \tau_0 \left( \partial_t \xi_h^1, v \right) + \left( gh_b \nabla \xi_h^{\frac{1}{2}}, \nabla v \right) + \mu \left( \partial_t \nabla \xi_h^1, \nabla v \right) \\
= \frac{2}{\Delta t} (\xi_1, v) - \left( \nabla \cdot \left( \frac{(U_h^0)^2}{H_h^0} \right), \nabla v \right) \\
- \left( \left( \tau_{b_h} - \tau_0 \right) U_h^0, \nabla v \right) - \left( T_h^0, \nabla v \right), \quad \forall v \in S_h(\Omega); \quad (4.20)
\]

\[
\left( \partial_t \xi_h^k, v \right) + \tau_0 \left( \partial_t \xi_h^k, v \right) + \left( \nabla \cdot \left( \frac{(U_h^k)^2}{H_h^k} \right), \nabla v \right) \\
+ \left( gh_b \nabla \xi_h^{k+\frac{1}{2}}, \nabla v \right) + \mu \left( \partial_t \nabla \xi_h^k, \nabla v \right) \\
+ \left( \left( \tau_{b_h} - \tau_0 \right) U_h^k, \nabla v \right) + \left( T_h^k, \nabla v \right) = 0, \quad \forall v \in S_h(\Omega), \ k \geq 1; \quad (4.21)
\]

\[
\left( \partial_t U_h^{k+1}, w \right) + \left( \nabla \cdot \left( \frac{(U_h^k)^2}{H_h^k} \right), w \right) + \left( gh_b \nabla \xi_h^{k+\frac{1}{2}}, w \right) \\
+ \mu \sum_{i=0}^{1} \beta_i \left( \nabla U_h^{k+i}, \nabla w \right) + \left( T_h^{k+\frac{1}{2}}, w \right) = 0, \quad \forall w \in S_h(\Omega), \ k \geq 0 \quad (4.22)
\]

with initial conditions

\[
(\xi_h^0, v) = (\xi_0, v), \quad \forall v \in S_h(\Omega),
\]

\[
(\xi_h^{-1}, v) = (\xi_h^1, v) - 2\Delta t(\xi_1, v), \quad \forall v \in S_h(\Omega), \quad (4.23)
\]

\[
(U_h^0, w) = (U_0, w), \quad \forall w \in S_h(\Omega).
\]
Equation (4.20) arises from considering (4.21) at \( k = 0 \) and using the fictitious value \( \xi_h^{-1} \) defined in (4.23).

In contrast to the flexibility of the \textit{ADClRC} simulator, we only consider the cases when diffusion terms in the CME are treated explicitly \((\beta_1 = 0, \beta_0 = 1)\) or implicitly \((\beta_1 = 1, \beta_0 = 0)\).

\subsection{Error Equations.}

To obtain an estimate for discretization errors \((\xi^k - \xi_h^k)\) and \((U^k - U_h^k)\), again we must first obtain an estimate on the affine error terms \((\xi_h^k - \tilde{\xi}^k)\) and \((U_h^k - \tilde{U}^k)\).

Before proceeding, we again make certain assumptions about the Galerkin approximations. We employ an inductive argument similar to that made in [25] to handle nonlinearities. That is, one can show, for sufficiently small mesh size \(h\), that the bound of the Galerkin approximations depends on a smaller bound (of the comparison projections) rather than on the initially assumed larger bound.

In particular, after showing the base case \((t^0)\), we will assume that the Galerkin approximations (at \(t^k, k = 1, \ldots, N - 1\)) are bounded by some constant. As a consequence of the error estimate, in the case of polynomials of degree at least two \((s_1 \geq 3)\), it will be true that the Galerkin approximations (at \(k = N\)) are bounded. And, this concludes the inductive argument.

It should not escape the reader that the boundedness assumptions are essential to the error estimate in the case of linear polynomial approximations.

Let us proceed with the inductive argument. From, Lemma 3.3 and assumptions \textbf{M1}, \textbf{M8} and \textbf{M9} (interpreted at time \(t = t^k\)), \(\exists\) positive constants \(C_\ast, C'\) such that for \(k = 0, \ldots, N\),

\[
C_\ast \leq \left\{ h_b + \tilde{\xi}^k, \nabla (h_b + \tilde{\xi}^k) \right\} \leq C' ,
\]

(4.24)
and
\[ \left\{ \left| \mathbf{U}^k \right|, \left| \nabla \mathbf{U}^k \right| \right\} \leq C^*. \tag{4.25} \]
Moreover, with \( C_* = C^*/2 \), \( C^{**} = 2C^* \), we assume that for \( k = 0, \ldots, (N - 1) \), and \( h \) and \( \Delta t \) sufficiently small,

**N3.** \( C_* < H_h(x, t^k) < C^{**} \).

**N4.** \( |U^k_h| < C^{**} \).

We immediately have the base case for the inductive proof: \( C_* < H^0_h < C^{**} \) and \( |U^0_h| < C^{**} \). In the remainder of this paper, from the derivation of the \textit{a priori} error estimate, we prove the hypothesis for time \( t = t^N \).

Observe that **N3** and **N4** imply **M3** holds at \( t = t^k \).

Now, subtract (4.4) at \( t = t^0 \) from (4.20); subtract (4.4) at \( t = t^k, k \geq 1 \) from (4.21), and subtract (4.5) at \( t = t^k \) from (4.22). Use the fact that we can write, using Taylor's Theorem with integral remainder, the following truncation-in-time terms:

\[ \delta_0^0 = \frac{\partial^2 \xi^0}{\partial t^2} + \frac{2}{\Delta t} (\xi_1 - \partial_t \xi^1) = -\frac{1}{(\Delta t)^2} \int_0^{\Delta t} (\Delta t - s)^2 \frac{\partial^3 \xi}{\partial t^3} ds; \]

\[ \delta_0^k = \frac{\partial^2 \xi^k}{\partial t^2} - \partial_t^2 \xi^k \]

\[ = -\frac{1}{6(\Delta t)^2} \left\{ \int_{t^k}^{t^{k+1}} (s - t_{k+1})^3 \frac{\partial^4 \xi}{\partial t^4} ds + \int_{t^k}^{t^{k+1}} (t_{k+1} - s)^3 \frac{\partial^4 \xi}{\partial t^4} ds \right\}, \quad k \geq 1; \]

\[ \delta_1^k = \frac{\partial \xi^k}{\partial t} - \partial_t \xi^k = -\frac{1}{\Delta t} \int_{t^k}^{t^{k+1}} (t_{k+1} - s) \frac{\partial^2 \xi}{\partial t^2} ds; \]

\[ \delta_2^k = -(\xi^{k+1} - \xi^k) = -\int_{t^k}^{t^{k+1}} \frac{\partial \xi}{\partial t} ds; \]

\[ \epsilon_1^k = \frac{\partial U^k}{\partial t} - \partial_t U^k = -\frac{1}{\Delta t} \int_{t^k}^{t^{k+1}} (t_{k+1} - s) \frac{\partial^2 U}{\partial t^2} ds; \]

\[ \epsilon_2^k = -(U^{k+1} - U^k) = -\int_{t^k}^{t^{k+1}} \frac{\partial U}{\partial t} ds. \]

Consequently, we obtain the following GWCE-CME error equations:

\[
\left( \frac{2 + \tau_0 \Delta t}{\Delta t} \right) \left( \partial_t \psi \xi^1, v \right) + \left( gh \nabla \psi \frac{\Delta t}{2}, \nabla v \right) + \mu \left( \partial_t \nabla \psi \xi^1, \nabla v \right)
\]
\[ \begin{aligned}
= \left( \nabla \left[ \left( \frac{U^0}{H^0}\right)^2 - \left( \frac{U_h^0}{H_h^0}\right)^2 \right] \right) \cdot \nabla v + \left( \tau_b - \tau_0 \right) \theta_U^0 \cdot \nabla v \\
+ \left( \tau_b - \tau_0 \right) \theta_U^0 \cdot \nabla v + \left( g h_b \nabla \theta_x^0 \cdot \nabla v \right) + \mu \left( \nabla \theta_x^0 \cdot \nabla v \right) \\
+ \left( \nabla \left[ \left( \frac{U^k}{H^k}\right)^2 - \left( \frac{U_h^k}{H_h^k}\right)^2 \right] \right) \cdot \nabla v \\
+ \left( \tau_b - \tau_0 \right) \theta_U^k \cdot \nabla v + \left( g h_b \nabla \theta_x^k \cdot \nabla v \right) + \left( \nabla \theta_x^k \cdot \nabla v \right) + \left( \nabla \delta_{1} \cdot \nabla v \right) + \mu \left( \nabla \delta_{1} \cdot \nabla v \right).
\end{aligned} \tag{4.26}
\]

\( (\partial_{\xi}^2 \psi_{\xi}^k \cdot v) + \tau_0 \left( \partial_{\xi} \psi_{\xi}^k \cdot v \right) + \left( \tau_b - \tau_0 \right) \psi_U^k \cdot \nabla v \\
+ \left( g h_b \nabla \psi_{\xi}^{k+\frac{1}{2}} \cdot \nabla v \right) + \mu \left( \partial_{\xi} \nabla \psi_{\xi}^k \cdot \nabla v \right)
\]

\[ \begin{aligned}
= \left( \nabla \left[ \left( \frac{U^k}{H^k}\right)^2 - \left( \frac{U_h^k}{H_h^k}\right)^2 \right] \right) \cdot \nabla v \\
+ \left( \tau_b - \tau_0 \right) \theta_U^k \cdot \nabla v + \left( g h_b \nabla \theta_x^k \cdot \nabla v \right) + \left( \nabla \theta_x^k \cdot \nabla v \right) + \left( \nabla \delta_{1} \cdot \nabla v \right) + \mu \left( \nabla \delta_{1} \cdot \nabla v \right), \quad \forall v \in S_h^0(\Omega), \quad k \geq 1;
\end{aligned} \tag{4.27}
\]

\[ \begin{aligned}
(\partial_{\xi} \psi_{\xi}^{k+1} \cdot w) + \left( \tau_b \psi_{\xi}^{k+\frac{1}{2}} \cdot w \right) + \mu \sum_{i=0}^{1} \beta_i \left( \nabla \psi_{\xi}^{k+1} \cdot \nabla w \right)
\end{aligned} \]

\[ \begin{aligned}
= \left( \nabla \left[ \left( \frac{U^k}{H^k}\right)^2 - \left( \frac{U_h^k}{H_h^k}\right)^2 \right] \right) \cdot w + \left( \tau_b \theta_U^{k+\frac{1}{2}} \cdot w \right) \\
+ \left( \tau_b - \tau_0 \right) \theta_U^{k+\frac{1}{2}} \cdot w + \left( g h_b \nabla \theta_x^{k+\frac{1}{2}} \cdot w \right) - \left( g h_b \nabla \psi_{\xi}^{k+\frac{1}{2}} \cdot w \right) \\
+ \mu \sum_{i=0}^{1} \beta_i \left( \nabla \theta_U^{k+i} \cdot \nabla w \right) + \left( \nabla \theta_x^{k+i} \cdot \nabla w \right) + \left( T^k - T_h^k \right) \cdot w + \frac{1}{2} \left( g h_b \nabla \delta_x^k \cdot w \right) \\
+ \left( \psi_{\xi}^k \cdot w \right) + \frac{1}{2} \left( \tau_b \psi_{\xi}^k \cdot w \right) + \frac{1}{2} \beta \left( \nabla \psi_{\xi}^k \cdot \nabla w \right), \quad \forall w \in S_h^0(\Omega), \quad k \geq 0.
\end{aligned} \tag{4.28}
\]

We have used the fact that, \( \psi_{\xi}^0 = 0, \psi_{\xi}^{-1} = \psi_{\xi}^1 \), and \( \psi_U^0 = 0 \) as well as the projection definitions (3.3.1.).

Note that with the exception of the how the initial conditions are implemented, the extension of the continuous-time analysis to discrete-time has been straight-forward.
4.3.4. Choice of Test Functions.

We choose the special test functions described in the previous section, now in discrete form. Let \( v = v_a + v_b \) be the test function in (4.26) where \( v_a = \psi_\xi^1 \) and \( v_b = \partial_{i_b} \psi_\xi^1 \). Let \( r \) be a positive constant to be determined and let \( v = v_1^{k+1} + v_2^{k+1} \) be the test function in (4.27), where

\[
 v_1^{k+1} = \sum_{j=k+1}^N e^{-r j \Delta t} \psi_{\xi}^j \Delta t \quad \text{and} \quad v_2^{k+1} = \sum_{j=k+1}^N e^{-r j \Delta t} \partial_{i_b} \left( \psi_{\xi}^j \right) \Delta t.
\]

Finally, let \( w = \psi_{U_h}^{k+1} \) be the test function in (4.28).

4.3.5. Bounding the first time-step of the GWCE Error Equation.

First, manipulate (4.26) to derive upper bounds for \( \| \psi_\xi^1 \|_{\mathcal{H}^1(\Omega)} \) and for \( \| \partial_{i_b} \psi_\xi^1 \|_{\mathcal{H}^1(\Omega)} \).

We will look at the test functions separately. First, use \( v_a = \psi_\xi^1 \) in (4.26) to obtain

\[
 \left( \frac{2 + \tau_s \Delta t}{2 \Delta t^2} \right) \| \psi_\xi^1 \|^2 + \frac{\tau_s}{2} \| \nabla \psi_\xi^1 \|^2 + \frac{\mu}{2 \Delta t} \| \nabla \psi_\xi^1 \| \leq \mathcal{P}_1 + \cdots \mathcal{P}_{10}. \quad (4.29)
\]

Here, \( \mathcal{P}_i \) denotes an error term that must be bounded in some fashion. Terms \( \mathcal{P}_1, \mathcal{P}_2 \) and \( \mathcal{P}_4 - \mathcal{P}_6 \) are treated similarly to the corresponding terms in continuous time (using test function \( v_1 \) with \( r = 0 \)) investigated thoroughly in [16]. We now investigate the remaining terms.

To estimating \( \mathcal{P}_1 \), use the relations of the previous section, Lemma 4.2 and Lemma 4.3 in [16], and assumptions \textbf{M1, M9, M10, M12, N3, N4} to obtain

\[
 \mathcal{P}_1 = \left( \nabla \cdot \left[ \left( \frac{(U^0)^2}{H^0} \right) - \left( \frac{(U_h^0)^2}{H_h^0} \right) \right], \nabla \psi_\xi^1 \right) \leq \epsilon \| \nabla \psi_\xi^1 \|^2 + K \| \theta_\xi^0 \|^2_{\mathcal{H}^1(\Omega)} + K \| \theta_U^0 \|^2_{\mathcal{H}^1(\Omega)}.
\]

In estimating \( \mathcal{P}_3 \), use the relations of the previous section, Lemma 4.2 and Lemma 4.3 in [16] and assumptions \textbf{M1, M9, M10, M12, N3, N4} to obtain
\[ P_3 = (\tau^0_0 - \tau^0_h) \tilde{U}^0, \nabla \psi^1 \]
\[ \leq c_f \left( \left( \frac{\psi^0_0 - \theta^0_0}{H^0_0 (H^0)^2} \right) \left\| U^0 \right\|_E + \frac{\left\| \theta U^0 \right\|_E}{H^0} + \frac{\left\| \psi U^0 \right\|_E}{H^0} \right) + \frac{\left( \psi^0_0 - \theta^0_0 \right) \left\| U^0_h \right\|_E}{H^0_0 (H^0)^2 H^0} \left( \tilde{U}^0, \nabla \psi^1 \right) \]
\[ \leq \epsilon \left\| \nabla \psi^1 \right\|^2 + K \left\| \theta^0_0 \right\|^2 + K \left\| \theta U^0 \right\|^2. \]

In estimating \( P_7 \), consider
\[ \left\| \delta_0^0 \right\|^2 \leq \frac{\Delta t^2}{5} \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}, \tag{4.30} \]
so that,
\[ P_7 = (\delta_0^0, \psi^1) \leq \epsilon \left\| \psi^1 \right\|^2 + K(\Delta t)^2 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}. \]

In estimating \( P_8 \) and \( P_{10} \), consider, for instance,
\[ \left\| \delta_1^0 \right\|^2 \leq \frac{\Delta t^2}{3} \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}, \]
so that,
\[ P_8 = \tau_0 (\delta_1^0, \psi^1) \leq \epsilon \left\| \psi^1 \right\|^2 + K(\Delta t)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}. \]
\[ P_{10} = \mu (\nabla \delta_1^0, \nabla \psi^1) \leq \epsilon \left\| \nabla \psi^1 \right\|^2 + K(\Delta t)^2 \left\| \nabla \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}. \]

Finally, in estimating \( P_9 \), consider
\[ \left\| \nabla \delta_2^0 \right\|^2 \leq \Delta t^2 \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}, \]
so that,
\[ P_9 = \frac{1}{2} (gh_0 \nabla \delta_2^0, \psi^1) \leq \epsilon \left\| \psi^1 \right\|^2 + K(\Delta t)^2 \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))}. \]

Consequently, we observe that the right hand side of (4.29) is bounded above by
\[ \epsilon_0 \left\| \psi^1 \right\|^2 + \epsilon_1 \left\| \nabla \psi^1 \right\|^2 + K \left\| \theta^0_0 \right\|^2_{H^1(\Omega)} + K \left\| \partial_t \nabla \theta^0 \right\|^2 \]
\[ + K \left\| \theta U^0 \right\|^2_{H^1(\Omega)} + K(\Delta t)^2 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty((0,\Delta t); C^2(\Omega))} \]
\[ + K(\Delta t)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^\infty((0,\Delta t); H^1(\Omega))}. \]
Let $\epsilon_0 = \frac{3}{16} \sigma^2$ so that $\Delta t < \frac{8}{\sigma^2} = \kappa_1$. Let $\sigma_0 = \left(\frac{2 + \tau_0 \Delta t}{2 \Delta t}\right) - \epsilon_0 > 0$ and $\sigma_1 = \frac{\gamma_* - \epsilon_1}{2} > 0$.

Thus,

$$
\sigma_0 \left\| \nabla \psi_1 \right\|^2 + \sigma_1 \left\| \nabla \psi_1 \right\|^2 + \frac{\mu}{\Delta t} \left\| \nabla \psi_1 \right\|^2 \\
\leq K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 + K \left\| \partial_t \nabla \theta_0 \right\|^2 + K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 \\
+ K' \left( \Delta t \right)^2 \left\| \frac{\partial^2 \theta}{\partial t^3} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \quad + K \left( \Delta t \right)^2 \left\| \frac{\nabla \psi_1}{\partial t} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \\
+ K \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|_{L^\infty((0, \Delta t); H^1(\Omega))}^2.
$$

Finally, we have

$$
\left\| \psi_1 \right\|_{H^1(\Omega)}^2 \leq K \left( h^{2(t-1)} + (\Delta t)^2 \right). \quad (4.31)
$$

Now, use $v_b = \partial_\mu \psi_1$ in (4.26) to obtain

$$
\left( \frac{2 + \tau_0 \Delta t}{\Delta t} \right) \left\| \partial_\mu \psi_1 \right\|^2 + \frac{\gamma_*}{2 \Delta t} \left\| \nabla \psi_1 \right\|^2 + \mu \left\| \partial_\mu \nabla \psi_1 \right\|^2 \\
\leq P_1 + \cdots P_{10}. \quad (4.32)
$$

After some work similar to that above, we determine that the right hand side of (4.32) is bounded above by

$$
\epsilon_2 \left\| \partial_\mu \psi_1 \right\|^2 + \epsilon_3 \left\| \partial_\mu \nabla \psi_1 \right\|^2 + K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 + K \left\| \partial_t \nabla \theta_0 \right\|^2 \\
+ K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 + K \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^3} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \quad + K \left( \Delta t \right)^2 \left\| \frac{\nabla \psi_1}{\partial t} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \\
+ K \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|_{L^\infty((0, \Delta t); H^1(\Omega))}^2.
$$

Let $\sigma_2 = \tau_0 - \epsilon_2 > 0$ and $\sigma_3 = \mu - \epsilon_3 > 0$. Then,

$$
\left( \frac{2 + \sigma_2 \Delta t}{\Delta t} \right) \left\| \partial_\mu \psi_1 \right\|^2 + \frac{\gamma_*}{2 \Delta t} \left\| \nabla \psi_1 \right\|^2 + \sigma_3 \left\| \partial_\mu \nabla \psi_1 \right\|^2 \\
\leq K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 + K \left\| \partial_\mu \nabla \theta_0 \right\|^2 + K \left\| \theta_0 \right\|_{H^1(\Omega)}^2 \\
+ K \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^3} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \quad + K \left( \Delta t \right)^2 \left\| \frac{\nabla \psi_1}{\partial t} \right\|_{L^\infty((0, \Delta t); L^2(\Omega))}^2 \\
+ K \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|_{L^\infty((0, \Delta t); H^1(\Omega))}^2.
$$
Finally, we have
\[ \left\| \partial_t \psi_t \right\|_{H^1(\Omega)}^2 + \left\| \nabla \psi_t \right\|_{L^2(\Omega)}^2 \leq K \left( h^{2(l-1)} + (\Delta t)^2 \right). \] (4.33)

Now, multiply (4.27) and (4.28) by \( \Delta t \) and sum from \( k = 0 \) to \( k = (N-1) \) using the test functions above. We will need to add \( \left\| \sqrt{g \theta_b} \right\|_{L^2(\Omega)} \) to both sides of the inequality obtained from (4.27), using \( \psi_t^{k+1} \), after multiplying by \( \Delta t \) and summing over \( k \).

Investigate the use of \( \psi_t^{k+1} \) and \( \psi_t^{k+1} \) as the test functions in (4.27) when we multiply by \( \Delta t \) and sum over \( k \). For a generic \( \varsigma \in H^1(\Omega) \), consider the following relations:

\[
\sum_{k=0}^{N-1} e^{\pm r \Delta t} \left( \partial_t \varsigma^k, \varsigma^{k+1} \right) = \frac{1}{2} \sum_{k=0}^{N-1} \left( e^{\pm r \Delta t} \left\| \varsigma^k \right\|^2 - e^{\pm r \Delta t} \left\| \varsigma^{k-1} \right\|^2 \right) \Delta t
\]
\[
+ \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left\| \varsigma^k \right\|^2 + \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left\| \varsigma^{k-1} \right\|^2 - \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left( \varsigma^{k-1}, \varsigma^k \right)
\]
\[
= \frac{1}{2} \sum_{k=0}^{N-1} \partial_t \varsigma^k \left( e^{\pm r \Delta t} \left\| \varsigma^k \right\|^2 \right) \Delta t + \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left( \varsigma^{k-1}, \varsigma^k \right) \Delta t
\]
\[
+ \frac{1}{2} \sum_{k=0}^{N-1} \left\| \partial_t \varsigma^k \right\|^2 (\Delta t)^2
\]
\[
= \frac{1}{2} \sum_{k=0}^{N-1} \partial_t \varsigma^k \left( e^{\pm r \Delta t} \left\| \varsigma^k \right\|^2 \right) \Delta t + \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} (1 - e^{\pm r \Delta t}) \left\| \varsigma^k \right\|^2
\]
\[
+ \frac{1}{2} e^{\pm r(N-1) \Delta t} (1 - e^{\pm r \Delta t}) \left\| \varsigma^{N-1} \right\|^2 - \frac{1}{2} e^{\pm r(N-1) \Delta t} (1 - e^{\pm r \Delta t}) \left\| \varsigma^{N-1} \right\|^2
\]
\[
+ \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left\| \partial_t \varsigma^k \right\|^2 (\Delta t)^2
\]
\[
= \frac{1}{2} e^{\pm r \Delta t} \left\| \varsigma^N \right\|^2 - \frac{1}{2} \left\| \varsigma^{-1} \right\|^2 + \frac{\theta}{2} e^{\theta \Delta t} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left\| \varsigma^k \right\|^2 \Delta t
\]
\[
+ \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r \Delta t} \left\| \partial_t \varsigma^k \right\|^2 (\Delta t)^2, \quad \theta \in (0, \pm r \Delta t), \quad \text{(A)}
\]

where the last equality results from an application of the Mean Value Theorem.

Similarly

\[
\sum_{k=0}^{N-1} e^{\pm r \Delta t} \left( \partial_t \varsigma^k, \varsigma^{k+1} \right) = \frac{1}{2} e^{\pm r \Delta t} \left\| \varsigma^N \right\|^2 - \frac{1}{2} \left\| \varsigma^0 \right\|^2
\]
\[ +\frac{r}{2} \sum_{k=0}^{N-1} e^{\pm r k \Delta t} \| \psi^{k+1} \|^2 \Delta t \]

\[ + \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r k \Delta t} \| \partial_{\xi} \psi^k \|^2 (\Delta t)^2, \quad \theta \in (0, \pm r \Delta t). \quad (B) \]

\[ \sum_{k=0}^{N-1} e^{\pm r k \Delta t} \left( \partial_{\xi} \psi^k, \psi^k \right) \Delta t = \frac{1}{2} \sum_{k=0}^{N-1} \partial_{\xi} \left( e^{\pm r k \Delta t} \| \psi^k \|^2 \right) \Delta t \]

\[ + \frac{r}{2} e^{\theta} \sum_{k=0}^{N-1} e^{\pm r k \Delta t} \| \psi^{k+1} \|^2 \Delta t \]

\[ - \frac{1}{2} \sum_{k=0}^{N-1} e^{\pm r k \Delta t} \| \partial_{\xi} \psi^k \|^2 (\Delta t)^2, \quad \theta \in (0, \pm r \Delta t). \quad (C) \]

We understand the following to be true: \( v^1_{1>N} = 0, \ v^2_{1>N} = 0. \)

Now, consider the first two terms of (4.27). When using the test function \( \nu^{k+1} \), we obtain

\[ \sum_{k=0}^{N-1} \left( \partial_{\xi}^2 \psi^k, \nu^{k+1} \right) \Delta t = - \sum_{k=0}^{N-1} \left( \partial_{\nu} \psi^k, \partial_{\xi} \nu^k \right) \Delta t - \left( \partial_{\nu} \psi^0, \nu^0 \right) + \left( \partial_{\nu} \psi^N, \nu^N \right) \]

\[ = \sum_{k=0}^{N} e^{-r k \Delta t} \left( \partial_{\nu} \psi^k, \psi^k \right) \Delta t - \left( \partial_{\nu} \psi^0, \nu^0 \right) \]

\[ = \frac{1}{2} e^{-r N \Delta t} \| \psi^{N-1} \|^2 - \frac{1}{2} \| \psi^{-1} \|^2 + \frac{r}{2} e^{\theta} \sum_{k=0}^{N} e^{-r k \Delta t} \| \psi^k \|^2 \Delta t \]

\[ + \frac{1}{2} \sum_{k=0}^{N} e^{-r k \Delta t} \| \partial_{\nu} \psi^k \|^2 (\Delta t)^2 - \left( \partial_{\nu} \psi^0, \nu^0 \right) \]

\[ \theta_0 \in (-r \Delta t, 0); \]

\[ \tau_0 \sum_{k=0}^{N-1} \left( \partial_{\xi} \psi^k, \nu^{k+1} \right) \Delta t = - \tau_0 \sum_{k=0}^{N-1} \left( \psi^k, \partial_{\xi} \nu^{k+1} \right) \Delta t - \tau_0 \left( \psi^0, \nu^0 \right) + \tau_0 \left( \psi^N, \nu^N \right) \]

\[ \equiv \tau_0 \sum_{k=0}^{N} e^{-r k \Delta t} \| \psi^k \|^2 \Delta t. \]

The first equalities above result from temporal summation by parts. We also have from the diffusion term upon summing by parts in time:

\[ \mu \sum_{k=0}^{N-1} \left( \partial_{\xi} \nabla \psi^k, \nabla v^{k+1} \right) \Delta t = \mu \sum_{k=0}^{N} e^{-r k \Delta t} \| \nabla \psi^k \|^2 \Delta t. \]
We are also able to manipulate the fourth term on the left side of (4.27) by using the definition of \( v_1^{k+1} \) as follows:

\[
\frac{1}{2} \sum_{k=0}^{N-1} \left( gh_b \nabla \psi_\xi^k, \nabla v_0^{k+1} \right) \Delta t \equiv \frac{1}{2} \sum_{k=0}^{N-1} e^{r_k \Delta t} \left( gh_b \partial_{\xi\psi} (\nabla v_0^{k+1}), \nabla v_1^{k+1} \right) \Delta t
\]

\[
\leq \frac{1}{4} \sum_{k=0}^{N-1} \partial_{\xi\psi} \left( e^{r_k \Delta t} \left\| \sqrt{gh_b} \nabla v_1^k \right\|^2 \right) \Delta t
\]

\[
+ \frac{r}{4} e^{r_k \Delta t} \left\| \sqrt{gh_b} \nabla v_0^{k+1} \right\|^2 \Delta t
\]

\[
- \frac{1}{4} \sum_{k=0}^{N-1} \left\| \partial_{\xi\psi} (\sqrt{gh_b} \nabla v_1^k) \right\|^2 \Delta t^2, \quad \theta \in (0, r \Delta t)
\]

\[
\geq \frac{\gamma_1}{4} \left\| \nabla v_0^0 \right\|^2 + \frac{r \gamma_1}{4} e^{r_k \Delta t} \left\| \nabla v_0^{k+1} \right\|^2 \Delta t
\]

\[
- \frac{1}{4} \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \sqrt{gh_b} \nabla \psi_\xi^k \right\|^2 (\Delta t)^2, \quad \theta \in (0, r \Delta t)
\]

and

\[
\frac{1}{2} \sum_{k=0}^{N-1} \left( gh_b \nabla \psi_\xi^{k+1}, \nabla v_1^{k+1} \right) \Delta t = \frac{1}{2} \sum_{k=0}^{N-1} \left( gh_b \nabla \psi_\xi^k, \nabla v_1^k \right) \Delta t + \frac{1}{2} \left( gh_b \nabla \psi_\xi^N, \nabla v_1^N \right) \Delta t
\]

\[
\equiv \frac{1}{2} \sum_{k=0}^{N} e^{r_k \Delta t} \left( gh_b \partial_{\xi\psi} (\nabla v_1^{k+1}), \nabla v_1^k \right) \Delta t
\]

\[
\leq \frac{1}{4} \sum_{k=0}^{N} \partial_{\xi\psi} \left( e^{r_k \Delta t} \left\| \sqrt{gh_b} \nabla v_1^k \right\|^2 \right) \Delta t
\]

\[
+ \frac{r}{4} e^{r_k \Delta t} \left\| \sqrt{gh_b} \nabla v_1^{k+1} \right\|^2 \Delta t
\]

\[
+ \frac{1}{4} \sum_{k=0}^{N} e^{r_k \Delta t} \left\| \partial_{\xi\psi} (\sqrt{gh_b} \nabla v_1^k) \right\|^2 \Delta t^2, \quad \theta \in (0, r \Delta t)
\]

\[
\geq \frac{\gamma_1}{4} \left\| \nabla v_0^0 \right\|^2 + \frac{r \gamma_1}{4} e^{r_k \Delta t} \left\| \nabla v_1^{k+1} \right\|^2 \Delta t
\]

\[
+ \frac{1}{4} \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \sqrt{gh_b} \nabla \psi_\xi^k \right\|^2 (\Delta t)^2, \quad \theta \in (0, r \Delta t)
\]

Similarly, using \( v_2^{k+1} \) as the test function in (4.27), we obtain

\[
\sum_{k=0}^{N-1} \left( \partial_{\xi\psi}^2 \psi_\xi^k, v_2^{k+1} \right) \Delta t = - \sum_{k=0}^{N-1} \left( \partial_{\xi\psi} \psi_\xi^k, \partial_{\xi\psi} v_2^k \right) \Delta t - \left( \partial_{\xi\psi} \psi_\xi^0, v_2^0 \right) + \left( \partial_{\xi\psi} \psi_\xi^N, v_2^N \right)
\]
where the first equality above results from summation by parts.

\[
\tau_0 \sum_{k=0}^{N-1} (\partial_t \psi^k, v^{k+1}_2) \Delta t \equiv \tau_0 \sum_{k=0}^{N-1} (\partial_t \psi^k, v^k_2) \Delta t + \tau_0 (\partial_t \psi^N, v^0_2) \Delta t - \tau_0 (\partial_t \psi^0, v^0_2) \Delta t
\]

\[
\equiv - \tau_0 \sum_{k=0}^{N} e^{r_k \Delta t} (\partial_t v^{k+1}_2, v^k_2) \Delta t - \tau_0 (\partial_t \psi^0, v^0_2) \Delta t
\]

\[
\equiv \frac{\tau_0}{2} \left\| v^0_2 \right\|^2 + \frac{r \tau_0}{2} e^\theta \sum_{k=0}^{N} e^{r_k \Delta t} \left\| v^{k+1}_2 \right\|^2 \Delta t + \tau_0 (\partial_t \psi^0, v^0_2) \Delta t
\]

\[
\frac{\tau_0}{2} \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \partial_t \psi^k \right\|^2 \left( \Delta t \right)^2, \ \theta \in (0, r \Delta t);
\]

\[
\mu \sum_{k=0}^{N-1} (\partial_t (\nabla \psi^k), \nabla v^{k+1}_2) \Delta t = \mu \left\| \nabla v^0_2 \right\|^2 + \frac{r \mu}{2} e^\theta \sum_{k=0}^{N} e^{r_k \Delta t} \left\| \nabla v^{k+1}_2 \right\|^2 \Delta t
\]

\[
+ \mu (\partial_t \nabla \psi^0, \nabla v^0_2) \Delta t
\]

\[
+ \frac{\mu}{2} \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \partial_t (\nabla \psi^k) \right\|^2 \left( \Delta t \right)^2, \ \theta \in (0, r \Delta t).
\]

The two terms above are manipulated using the definition of the test function followed by an application of (C).

4.3.6. Bounding the GWCE-Error Equations.

Recall that \( H^0 \) and \( |U^0| \) are bounded above by \( C^* \) and \( H^0 \) is bounded below by \( C_* \), and that we assumed that \( C_* < H^k < C^* \), and \( |U^k| < C^* \) for \( k = 0, \ldots, N - 1 \).

Now, using \( v^{k+1}_1 \) as the test function in (4.27), multiplying by \( \Delta t \) and summing from \( k = 0 \) to \( k = (N - 1) \), and adding \( \left\| \sqrt{h_b} v^0_1 \right\|^2 \) to both sides of the resulting inequality yields

\[
\frac{1}{2} e^{-r N \Delta t} \left\| \psi^N \right\|^2 + \left( \tau_0 + \frac{r}{2} e^\theta \right) \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \psi^k \right\|^2 \Delta t
\]
\[
+ \frac{\gamma^*}{2} \sum_{k=0}^{N-1} e^{r_k \Delta t} \left\| \nabla v_1^{k+1} \right\|^2 \Delta t + \mu \sum_{k=0}^{N} e^{-r_k \Delta t} \left\| \nabla \psi_1^k \right\|^2 \Delta t \\
+ \frac{1}{2} \sum_{k=0}^{N} e^{-r_k \Delta t} \left| \partial_x \psi_1^k \right|^2 (\Delta t)^2 + \frac{\gamma^*}{2} \left\| \nabla v_1^0 \right\|^2 + \gamma^* \left\| v_1^0 \right\|^2 \\
\leq \frac{1}{2} \left\| \psi_1^1 \right\|^2 - \left( \partial_x \psi_1^k, v_1^k \right) + \sum_{k=0}^{N-1} \left\{ \nabla \cdot \left[ \left( \frac{(U^k)^2}{H^k} \right) - \left( \frac{(U^H_h)^2}{H^H_h} \right) \right], \nabla v_1^{k+1} \right\} \\
- \left( (\tau_{\alpha} - \tau_0) \psi_1^k, \nabla v_1^{k+1} \right) + \left( (\tau_{\beta} - \tau_0) \theta U^k, \nabla v_1^{k+1} \right) \\
+ \left( (\tau_{\beta} - \tau_0) \delta U^k, \nabla v_1^{k+1} \right) + \left( \epsilon \delta_0, v_1^{k+1} \right) + \tau_0 \left( \delta_1, v_1^{k+1} \right) + \frac{1}{2} \left( \epsilon \delta_2, \nabla v_1^{k+1} \right) \\
+ \mu \left( \nabla \delta_1, \nabla v_1^{k+1} \right) \right\} \Delta t + \gamma^* \left\| v_1^0 \right\|^2 \\
= \tilde{S}_1 + \tilde{S}_2 + \left( \tilde{P}_1 + \cdots + \tilde{P}_{12} \right) .
\]

Here \( \tilde{S}_i \) denotes a term resulting from summation by parts and \( \tilde{P}_i \) corresponds to a term that must be bounded in some fashion. Terms \( \tilde{P}_1 - \tilde{P}_3 \), and \( \tilde{P}_5 - \tilde{P}_7 \) are treated similarly to their continuous-time analogs which are investigated thoroughly in [16].

Again, we investigate the remaining terms.

The treatment of terms \( \tilde{S}_1 \) and \( \tilde{S}_2 \) is straightforward. From (4.31) and (4.33),

\[
\tilde{S}_1 = \frac{1}{2} \left\| \psi_1^1 \right\|^2 \leq K \left( h^{2(l-1)} + \Delta t^2 \right) : \\
\tilde{S}_2 = - \left( \partial_x \psi_1^{k+1}, v_1^0 \right) \leq K \left( h^{2(l-1)} + \Delta t^2 \right) + \epsilon \left\| v_1^0 \right\|^2 .
\]

The treatment of \( \tilde{P}_1 \) is similar to that given in the previous section.

\[
\tilde{P}_1 = \sum_{k=0}^{N-1} \nabla \cdot \left[ \left( \frac{(U^k)^2}{H^k} \right) - \left( \frac{(U^H_h)^2}{H^H_h} \right) \right], \nabla v_1^{k+1} \right\} \Delta t \\
\leq \epsilon \sum_{k=0}^{N-1} e^{-r_k \Delta t} \left\| \psi_1^k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} e^{-r_k \Delta t} \left\| \nabla \psi_1^k \right\|^2 \Delta t \\
+ \epsilon \sum_{k=0}^{N-1} \left\| \theta_1^k \right\|_{H^1(\Omega)} \Delta t + K \sum_{k=0}^{N-1} \left\| \theta U^k \right\|_{H^1(\Omega)} \Delta t \\
+ K \sum_{k=0}^{N-1} e^{r_k \Delta t} \left\| \nabla v_1^{k+1} \right\|^2 \Delta t + K \sum_{k=0}^{N-1} \left\| \psi U^k \right\|^2 \Delta t .
\]

The treatment of \( \tilde{P}_4 \) is also similar to that given in the previous section:
\( \mathcal{P}_4 = \sum_{k=0}^{N-1} \left( (\tau_j^k - \tau_{h_j}^k) \bar{v}^k, \nabla v_{1}^{k+1} \right) \Delta t \)

\[ \leq \epsilon \sum_{k=0}^{N-1} e^{-r_k \Delta t} \| \psi^k \|^2 \Delta t + K \sum_{k=0}^{N-1} \| \theta^k \|^2 \Delta t + K \sum_{k=0}^{N-1} \| \psi^k_U \|^2 \Delta t \]

\[ + K \sum_{k=0}^{N-1} \| \theta^k_U \|^2 \Delta t + K \sum_{k=0}^{N-1} e^{-r_k \Delta t} \| \nabla v_{1}^{k+1} \|^2 \Delta t . \]

We recall that

\[ \mathcal{P}_7 = \sum_{k=0}^{N-1} \left( \mathbf{T}^k - \mathbf{T}_h^k, \nabla v_{1}^{k+1} \right) \Delta t \]

\[ \leq \epsilon \sum_{k=0}^{N-1} e^{-r_k \Delta t} \| \psi^k \|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} e^{-r_k \Delta t} \| \nabla \psi^k \|^2 \Delta t + K \sum_{k=0}^{N-1} \| \theta^k \|^2_{H^1(\Omega)} \Delta t \]

\[ + K \sum_{k=0}^{N-1} \| \psi^k_U \|^2 \Delta t + K \sum_{k=0}^{N-1} \| \theta^k_U \|^2 \Delta t + K \sum_{k=0}^{N-1} e^{-r_k \Delta t} \| \nabla v_{1}^{k+1} \|^2 \Delta t . \]

In estimating \( \mathcal{P}_8 \), consider

\[ \| \delta_0^k \|^2 \leq \frac{(\Delta t)^3}{136} \left\{ \int_{t^k}^{t^{k+1}} \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2 ds + \int_{t^{k-1}}^{t^k} \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2 ds \right\} , \quad k \geq 1. \]

Noting that \( v_{1}^{k+1} = v_{1}^0 - \sum_{j=0}^{k} e^{-r_j \Delta t} \psi^k \Delta t \) and using (4.30), we obtain

\[ \mathcal{P}_8 = \sum_{k=0}^{N-1} (\delta_0^k, v_{1}^{k+1}) \Delta t \leq \epsilon \| v_{1}^0 \|^2 + K \sum_{k=0}^{N-1} \sum_{j=0}^{k} e^{-r_j \Delta t} \| \psi^k \|^2 \Delta t, \]

\[ + K (\Delta t)^4 \int_{0}^{T} \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2 ds + K (\Delta t)^3 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{C^2(\Omega)}. \]

In estimating \( \mathcal{P}_9 \) and \( \mathcal{P}_{11} \), consider, for instance,

\[ \delta_1^k \leq \left| -\frac{1}{\Delta t} \left( \int_{t^k}^{t^{k+1}} (t^{k+1} - s)^2 ds \right) \right|^{\frac{1}{2}} \left( \int_{t^k}^{t^{k+1}} \left( \frac{\partial^2 \xi}{\partial t^2} \right)^2 ds \right)^{\frac{1}{2}} \]

\[ = (\frac{\Delta t}{3})^{1/2} \left( \int_{t^k}^{t^{k+1}} \left( \frac{\partial^2 \xi}{\partial t^2} \right)^2 ds \right)^{\frac{1}{2}}, \]

so that

\[ \| \delta_1^k \|^2 \leq \frac{\Delta t}{3} \int_{t^k}^{t^{k+1}} \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2 ds . \]
Thus,

\[ \mathcal{P}_9 = \tau_0 \sum_{k=0}^{N-1} (\delta_t^k, v_1^{k+1}) \Delta t \]

\[ \leq \epsilon \|v_1^0\|^2 + K \sum_{k=0}^{N-1} \left( \sum_{j=0}^{k} e^{-r_j \Delta t} \psi_k^j \Delta t \right)^2 \Delta t + K(\Delta t)^2 \int_0^T \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2 ds; \]

\[ \mathcal{P}_{11} = \mu \sum_{k=0}^{N-1} \left( \nabla \delta_t^k, \nabla v_1^{k+1} \right) \Delta t \]

\[ \leq \epsilon \|\nabla v_1^0\|^2 + K \sum_{k=0}^{N-1} \left( \sum_{j=0}^{k} e^{-r_j \Delta t} \nabla \psi_k^j \Delta t \right)^2 \Delta t + K(\Delta t)^2 \int_0^T \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2 ds. \]

The treatment of \( \mathcal{P}_{10} \) follows similarly

\[ \mathcal{P}_{10} = \frac{1}{2} \sum_{k=0}^{N-1} \left( gh_b \nabla \delta_2^k, \nabla v_1^{k+1} \right) \Delta t \]

\[ \leq \epsilon \|\nabla v_1^0\|^2 + K \sum_{k=0}^{N-1} \left( \sum_{j=0}^{k} e^{-r_j \Delta t} \nabla \psi_k^j \Delta t \right)^2 \Delta t \]

\[ + K(\Delta t)^2 \int_0^T \left\| \frac{\partial \xi}{\partial t} \right\|^2 ds. \]

Finally, algebraic manipulation yields

\[ \mathcal{P}_{12} = \gamma^* \|v_2^0\|^2 = \gamma^* \left( \sum_{k=0}^N e^{-r_k \Delta t} \psi_k^k \Delta t, \sum_{k=0}^N e^{-r_k \Delta t} \psi_k^k \Delta t \right) \]

\[ \leq \epsilon \left( \sum_{k=0}^N e^{-r_k \Delta t} \psi_k^k \Delta t \right)^2 \]

\[ + K \sum_{k=0}^N \left( \sum_{k=0}^N e^{-r_k \Delta t} \psi_k^k \right)^2 \Delta t \]

\[ = \epsilon \|v_2^0\|^2 + K \sum_{k=0}^N e^{-r_k \Delta t} \psi_k^k \Delta t. \]

Using \( v_2^{k+1} \) as the test function in (4.27), multiplying by \( \Delta t \) and summing from \( k = 0 \) to \( k = (N - 1) \), and using the relations above yields

\[ \sum_{k=0}^N e^{-r_k \Delta t} \left\| \partial_t \psi_k^k \right\|^2 \Delta t + \frac{\tau_0}{2} \sum_{k=0}^N e^{-r_k \Delta t} \left\| \partial_t \psi_k^k \right\|^2 \Delta t \]

\[ + \frac{\mu}{2} \sum_{k=0}^N e^{-r_k \Delta t} \left\| \nabla \psi_k^k \right\|^2 \Delta t \]

\[ + \frac{r \tau_0}{2} e^\theta \sum_{k=0}^N e^{-r_k \Delta t} \left\| v_2^{k+1} \right\|^2 \Delta t + \frac{\tau_0}{2} e^\theta \sum_{k=0}^N e^{-r_k \Delta t} \left\| \nabla v_2^{k+1} \right\|^2 \Delta t \]

\[ \leq - (\partial_t \psi_1^0, v_2^0) - \tau_0 \left( \partial_t \psi_1^1, v_2^0 \right) \Delta t - \mu \left( \partial_t \psi_1^1, \nabla v_2^0 \right) \Delta t \]
\[ \begin{align*}
- \sum_{k=0}^{N-1} \left\{ \nabla \cdot \left[ \left( \frac{(U^k)^2}{H^k} \right) - \left( \frac{(U_h^k)^2}{H_h^k} \right) \right], \nabla v_1^{k+1} \right\} \\
- \left( (\tau - \tau_h^k) \theta U^k, \nabla v_2^{k+1} \right) + \left( (\tau_h^k - \tau_h^{k+1}) U_h^k, \nabla v_2^{k+1} \right) - \left( gh_b \nabla \psi_k^{k+1}, \nabla v_2^{k+1} \right) \\
+ \left( gh_h \nabla \theta_k^{k+1}, \nabla v_2^{k+1} \right) + \mu \left( \partial_t \nabla \psi_k, \nabla v_2^{k+1} \right) + \left( T^k - T_h^k, \nabla v_2^{k+1} \right) \\
+ \left( \delta_0, v_2^{k+1} \right) + \tau_0 \left( \delta_1, v_2^{k+1} \right) + \frac{1}{2} \left( gh_h \nabla \delta_2, \nabla v_2^{k+1} \right) + \mu \left( \nabla \delta_1, \nabla v_2^{k+1} \right) \right\} \Delta t \\
= \dot{S}_1 + \dot{S}_2 + \dot{S}_3 + \left( \dot{P}_1 + \cdots + \dot{P}_{11} \right). \tag{4.35}
\end{align*} \]

All of these terms are treated analogously to those in (4.35) with the exception of \( \dot{S}_2, \dot{S}_3 \) and \( \dot{P}_8 - \dot{P}_{11} \) which we detail now.

From (4.31) and (4.33),
\[ \begin{align*}
\dot{S}_2 &= -\tau_0 \left( \partial_t \psi_k, v_2^0 \right) \Delta t \leq \epsilon \left\| v_2^0 \right\| + K \Delta t^2 \left( h^{2(t-1)} + \Delta t^2 \right), \\
\dot{S}_3 &= -\mu \left( \partial_t \psi_k, \nabla v_2^0 \right) \Delta t \leq \epsilon \left\| \nabla v_2^0 \right\| + K \Delta t^2 \left( h^{2(t-1)} + \Delta t^2 \right).
\end{align*} \]

Observe that the treatment of \( \dot{P}_8 - \dot{P}_{10} \) differs slightly from the treatment of related terms in (4.34) as follows:
\[ \begin{align*}
\dot{P}_8 &= \sum_{k=0}^{N-1} \left( \delta_0, v_2^{k+1} \right) \Delta t \leq K(\Delta t)^4 \int_0^T \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2 ds + K \sum_{k=0}^{N-1} e^{r \Delta t} \left\| v_2^{k+1} \right\|^2 \Delta t \\
&\quad + K(\Delta t)^3 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty(O, \Delta t; L^2(\Omega))} ;
\end{align*} \]

\[ \begin{align*}
\dot{P}_9 &= \tau_0 \sum_{k=0}^{N-1} \left( \delta_1, v_2^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2 ds + K \sum_{k=0}^{N-1} e^{r \Delta t} \left\| v_2^{k+1} \right\|^2 \Delta t;
\end{align*} \]

\[ \begin{align*}
\dot{P}_{10} &= \frac{1}{2} \sum_{k=0}^{N-1} \left( gh_h \nabla \delta_2, \nabla v_2^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left\| \frac{\partial \xi}{\partial t} \right\|^2 ds \\
&\quad + K \sum_{k=0}^{N} e^{r \Delta t} \left\| \nabla v_2^{k+1} \right\|^2 \Delta t;
\end{align*} \]

\[ \begin{align*}
\dot{P}_{11} &= \mu \sum_{k=0}^{N-1} \left( \nabla \delta_1, \nabla v_2^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2 ds + K \sum_{k=0}^{N} e^{r \Delta t} \left\| \nabla v_2^{k+1} \right\|^2 \Delta t.
\end{align*} \]
Finally, algebraic manipulation of the following terms from (4.28) yields:

\[
\sum_{k=0}^{N-1} \left( \partial_t \psi U^{k+1}, \psi U^{k+1} \right) \Delta t = \frac{1}{2} \sum_{k=0}^{N-1} \left( \| \psi U^{k+1} \|^{2} - \| \psi U^{k} \|^{2} + \| \partial_t \psi U^{k+1} \|^{2} \right) (\Delta t)^2
\]

\[
= \frac{1}{2} \| \psi U^{N} \|^{2} + \frac{1}{2} \sum_{k=0}^{N-1} \| \partial_t \psi U^{k+1} \|^{2} (\Delta t)^2;
\]

\[
\sum_{k=0}^{N-1} \left( \tau_{bh}^{k} \psi U^{k+\frac{1}{2}}, \psi U^{k+1} \right) \Delta t = \frac{1}{2} \sum_{k=0}^{N-1} \left( \| \sqrt{\tau_{bh}^{k}} \psi U^{k+1} \|^{2} \right) \Delta t
\]

\[
+ \frac{1}{2} \sum_{k=0}^{N-1} \left( \sqrt{\tau_{bh}^{k}} \psi U^{k+1} - (\psi U^{k+1} - \psi U^{k}) \right) \Delta t
\]

\[
\geq \frac{3}{4} \sum_{k=0}^{N-1} \left( \| \sqrt{\tau_{bh}^{k}} \psi U^{k+1} \|^{2} \right) \Delta t
\]

\[
- \frac{1}{4} \sum_{k=0}^{N-1} \left( \| \sqrt{\tau_{bh}^{k}} \partial_t \psi U^{k+1} \|^{2} \right) (\Delta t)^3;
\]

\[
\mu \sum_{k=0}^{N-1} \left( \beta_1 \nabla \psi U^{k+1} + \beta_0 \nabla \psi U^{k}, \nabla \psi U^{k+1} \right) \Delta t \geq \mu \left( \beta_1 + \beta_0 \frac{2}{2} \right) \sum_{k=0}^{N-1} \| \nabla \psi U^{k+1} \|^{2} \Delta t
\]

\[
- \frac{\mu \beta_0}{2} \sum_{k=0}^{N-1} \| \partial_t \psi U^{k+1} \|^{2} (\Delta t)^3.
\]

4.3.7. Bounding the CME-Error Equation.

Using \( w = \psi U^{k+1} \) as the test function in (4.28) followed by summation in time, yields

\[
\frac{1}{2} \| \psi U^{N} \|^{2} + \frac{1}{2} \sum_{k=0}^{N-1} \| \partial_t \psi U^{k+1} \|^{2} (\Delta t)^2
\]

\[
+ \frac{3}{4} \sum_{k=0}^{N-1} \| \sqrt{\tau_{bh}^{k}} \psi U^{k+1} \|^{2} \Delta t + \mu \left( \beta_1 + \beta_0 \frac{2}{2} \right) \sum_{k=0}^{N-1} \| \nabla \psi U^{k+1} \|^{2} \Delta t
\]

\[
\leq \frac{\tau_{bh}}{4} \sum_{k=0}^{N-1} \| \partial_t \psi U^{k+1} \|^{2} (\Delta t)^3 + \frac{\mu \beta_0}{2} \sum_{k=0}^{N-1} \| \partial_t \nabla \psi U^{k+1} \|^{2} (\Delta t)^3
\]

\[
+ \sum_{k=0}^{N-1} \left\{ \nabla \cdot \left[ \left( \frac{(U^{k})^2}{H^k} \right) - \left( \frac{(U_{bh}^{k})^2}{H_{bh}^k} \right) \right], \psi U^{k+1} \right\} + \left( \tau_{bh} \theta U^{k+\frac{1}{2}}, \psi U^{k+1} \right)
\]
\[ + \left( \tau_h^k - \tau_{h^k} \right) \tilde{U}^{k+\frac{1}{2}}, \psi_U^{k+1} \right) + \left( g h_b \nabla \theta \xi^k, \psi_U^{k+1} \right) - \left( g h_b \nabla \psi \xi^{k+\frac{1}{2}}, \psi_U^{k+1} \right) \\
+ \mu \sum_{i=0}^{m} \beta_i \left( \nabla \theta U^{k+i}, \nabla \psi U^{k+1} \right) + \left( T^k - T_h^k, \psi_U^{k+1} \right) + \frac{1}{2} \left( g h_b \nabla \psi \xi^k, \psi_U^{k+1} \right) \\
+ \left( \epsilon_1^k, \psi_U^{k+1} \right) + \frac{1}{2} \left( \tau_h^k \epsilon_2^k, \psi_U^{k+1} \right) + \mu \beta_1 \left( \nabla \epsilon_2^k, \nabla \psi \xi^{k+\frac{1}{2}} \right) \right) \Delta t \\
= \hat{S}_1 + \hat{S}_2 + \left( \hat{P}_1 + \cdots + \hat{P}_{11} \right). \tag{4.36} \]

Again, all the terms on the right-hand side are treated analogously to the terms in (4.35) with the exception of the following terms.

In estimating \( \hat{S}_2 \), we use the inverse assumption to obtain

\[ \hat{S}_2 = \frac{\mu}{2} \sum_{k=0}^{N-1} \left( \nabla \psi U^{k+1} \right)^2 \Delta t \leq \left( \frac{K_0 \mu}{2 h^2} \right) \sum_{k=0}^{N-1} \left( \nabla \psi U^{k+1} \right)^2 \Delta t. \]

And the truncation terms are estimated as follows

\[ \hat{P}_8 = \frac{1}{2} \sum_{k=0}^{N-1} \left( g h_b \nabla \psi \xi^k, \psi_U^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left( \nabla \psi \xi^k \right)^2 ds + K \sum_{k=0}^{N-1} \left( \psi U^{k+1} \right)^2 \Delta t; \]

\[ \hat{P}_9 = \sum_{k=0}^{N-1} \left( \epsilon_1^k, \psi_U^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left( \nabla \psi U^{k+1} \right)^2 ds + K \sum_{k=0}^{N-1} \left( \psi U^{k+1} \right)^2 \Delta t; \]

\[ \hat{P}_{10} = \frac{1}{2} \sum_{k=0}^{N-1} \left( \tau_h^k \epsilon_2^k, \psi_U^{k+1} \right) \Delta t \leq K(\Delta t)^2 \int_0^T \left( \nabla \psi \xi^{k+\frac{1}{2}} \right)^2 ds + K \sum_{k=0}^{N-1} \left( \psi U^{k+1} \right)^2 \Delta t; \]

\[ \hat{P}_{11} = \mu \beta_1 \sum_{k=0}^{N-1} \left( \nabla \psi U^{k+1} \right)^2 \Delta t \leq K(\Delta t)^2 \int_0^T \left( \nabla \psi \xi^{k+\frac{1}{2}} \right)^2 ds + \epsilon \sum_{k=0}^{N-1} \left( \psi U^{k+1} \right)^2 \Delta t. \]

Finally, after manipulating appropriate terms in (4.27) and (4.28), using the derived bounds on \( \left\| \psi \xi^k \right\| \left( H \Omega \right) \) and on \( \left\| \nabla \psi \xi^k \right\| \left( H \Omega \right) \), adding the resulting inequalities, applying a generalized discrete Gronwall’s inequality, and taking bounds above and below, we will obtain a relation giving an estimate of the affine error.

To apply Gronwall’s inequality, we will need to define a function

\[ \Lambda^k = e^{-r k \Delta t} \left\| \psi \xi^k \right\|^2 + \left\| \sum_{j=0}^{k} e^{-r j \Delta t} \psi \xi^k \Delta t \right\| \left( H \Omega \right) + \left\| \psi U^k \right\|^2. \]

We will then, after some work, obtain \( \Lambda^N \) in which \( \sum_{k=0}^{N} \Lambda^k \Delta t \) will be hidden using Gronwall’s Lemma.
4.3.8. Bounding the Sum of the Error Equations.

We observe that the right-hand-side of (4.34) can be bounded by

\[
\epsilon \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| \psi_k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| \nabla \psi_k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} \left\| \nabla \psi_{k+1} \right\|^2 \Delta t
\]

\[+ \epsilon \left\| v_1^0 \right\|^2 + \epsilon \left\| \nabla v_1^0 \right\|^2 + K \sum_{k=0}^{N-1} \left( \left\| \theta_k \right\|^2_{\mathcal{H}^1(\Omega)} + \left\| \nabla \theta_k \right\|^2 \right) \Delta t \]

\[+ K \sum_{k=0}^{N-1} \left\| \theta_k \right\|^2_{\mathcal{H}^1(\Omega)} \Delta t + K \left( h^{2(l-1)} + \Delta t^2 \right) \]

\[+ K' \left( \Delta t \right)^4 \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} + K' \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} \]

\[+ K' \left( \Delta t \right)^2 \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} + K' \left( \Delta t \right)^3 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty([0,\Delta t];C^2(\Omega))} \]

\[+ K_1 \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| v_1^{k+1} \right\|^2 \Delta t + K' \sum_{k=0}^{N-1} \left\| \psi_{k+1}^{k+1} \right\|^2 \Delta t + K \sum_{k=0}^{N-1} \left\| e^{-r k \Delta t} \psi_k \right\|^2 \Delta t \]

\[+ K' \sum_{k=0}^{N-1} \left\| \sum_{j=0}^{k} e^{-r j \Delta t} \psi_j \Delta t \right\|^2 + K' \sum_{k=0}^{N-1} \left\| \sum_{j=0}^{k} e^{-r j \Delta t} \nabla \psi_j \Delta t \right\|^2 \Delta t. \tag{4.37} \]

The right-hand-side of (4.35) can be bounded by

\[
\epsilon \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| \psi_k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| \nabla \psi_k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} \left\| \nabla \psi_{k+1} \right\|^2 \Delta t
\]

\[+ K' \sum_{k=0}^{N-1} \left( \left\| \theta_k \right\|^2_{\mathcal{H}^1(\Omega)} + \left\| \nabla \theta_k \right\|^2 \right) \Delta t + K' \sum_{k=0}^{N-1} \left\| \theta_k \right\|^2_{\mathcal{H}^1(\Omega)} \Delta t \]

\[+ K' \sum_{k=0}^{N-1} \left\| \psi_{k+1} \right\|^2 \Delta t + K' \left( h^{2(l-1)} + \Delta t^2 \right) \Delta t^2 \]

\[+ K' \left( \Delta t \right)^4 \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} + K' \left( \Delta t \right)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} \]

\[+ K' \left( \Delta t \right)^2 \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2_{L^2(J_{N+1};C^2(\Omega))} + K' \left( \Delta t \right)^3 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty([0,\Delta t];C^2(\Omega))} \]

\[+ K_2 \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| v_2^{k+1} \right\|^2 \Delta t + K_3 \sum_{k=0}^{N-1} e^{-r k \Delta t} \left\| \nabla v_2^{k+1} \right\|^2 \Delta t. \tag{4.38} \]

Finally, the right-hand-side of (4.36) can be bounded by
\[ \left( \frac{\hat{\tau} h^2 + 2K_0 \mu \theta_0}{4h^2} \right) \sum_{k=0}^{N-1} \left\| \partial_\xi \psi_U^{k+1} \right\|^2 (\Delta t)^3 + \epsilon \sum_{k=0}^{N-1} e^{-rk\Delta t} \left\| \psi_\xi^k \right\|^2 \Delta t + \epsilon \sum_{k=0}^{N-1} \left\| \nabla \psi_U^{k+1} \right\|^2 \Delta t + K \sum_{k=0}^{N-1} \left\| \theta^k \right\|_{H^1(\Omega)}^2 \Delta t + K \sum_{k=0}^{N-1} \left\| \theta_U^k \right\|_{H^1(\Omega)}^2 \Delta t + K \sum_{k=0}^{N-1} \left\| \theta_U^{k+1} \right\|_{H^1(\Omega)}^2 \Delta t + K \sum_{k=0}^{N-1} \left\| \psi_U^{k+1} \right\|^2 \Delta t + K(\Delta t)^2 \left\| \frac{\partial^2 U}{\partial t^2} \right\|^2_{L^2(J^1_4; L^2(\Omega))} + K(\Delta t)^2 \left\| \frac{\partial U}{\partial t} \right\|^2_{L^2(J^1_4; H^1(\Omega))} + K(\Delta t)^2 \left\| \nabla \frac{\partial \xi}{\partial t} \right\|^2_{L^2(J^1_4; H^1(\Omega))}. \] (4.39)

Choose
\[ r = \max \left\{ \frac{2K_1}{\gamma_*}, \frac{2K_2}{\tau_0}, \frac{2K_3}{\mu} \right\}, \]
such that,
\[ r_1 = r \frac{\gamma_*}{2} e^\theta - K_1 \geq 0, \]
\[ r_2 = r \frac{\tau_0}{2} e^\theta - K_2 \geq 0, \]
and
\[ r_3 = r \frac{\mu}{2} e^\theta - K_3 \geq 0. \]

Observe that
\[ \Lambda^k = e^{-rk\Delta t} \left\| \psi_\xi^k \right\|^2 + \left\| \psi_U^k \right\|^2 + \left\| \sum_{j=0}^{k} e^{-rj\Delta t} \psi_\xi^j \Delta t \right\|_{H^1(\Omega)}^2 \]
\[ \geq \left\| e^{-rk\Delta t} \psi_\xi^k \right\|^2 + \left\| \psi_U^k \right\|^2 + \left\| \sum_{j=0}^{k} e^{-rj\Delta t} \psi_\xi^j \Delta t \right\|_{H^1(\Omega)}^2. \]

Now, sum (4.34)–(4.36), using bounds (4.37)–(4.39) and the choice for \( r \) above to obtain
\[
\begin{align*}
&\frac{1}{2}e^{-rN\Delta t}\|\psi_{x}^{N}\|^{2} + \left(\tau_{0} + \frac{r^{2}}{2}\right)\sum_{k=0}^{N} e^{-rk\Delta t}\|\psi_{x}^{k}\|^{2}\Delta t
+ \sum_{k=0}^{N} e^{-rk\Delta t}\|\partial_{t}\psi_{x}^{k}\|^{2}\Delta t + \left(\frac{\tau_{0} + 1}{2}\right)\sum_{k=0}^{N} e^{-rk\Delta t}\|\partial_{\nu}\nabla_{x}^{k}\|^{2}(\Delta t)^{2} \tag{a}
\right)
+ \mu\sum_{k=0}^{N} e^{-rk\Delta t}\|\nabla\psi_{x}^{k}\|^{2}\Delta t + \frac{\mu}{2}\sum_{k=0}^{N} e^{-rk\Delta t}\|\partial_{\nu}\nabla_{x}^{k}\|^{2}(\Delta t)^{2} \tag{b}
\right)
\end{align*}
\]\[
\begin{align*}
+ \gamma_{v}\|v_{1}^{0}\|^{2} + \frac{\gamma_{v}}{2}\|\nabla v_{1}^{0}\|^{2} + \frac{\tau_{0}}{2}\|v_{2}^{0}\|^{2} + \frac{\mu}{2}\|\nabla v_{2}^{0}\|^{2} \tag{c}
\end{align*}
\]\[
\begin{align*}
+ r_{1} \sum_{k=0}^{N-1} e^{rk\Delta t}\|\nabla v_{1}^{k+1}\|^{2}\Delta t + r_{2} \sum_{k=0}^{N-1} e^{rk\Delta t}\|v_{2}^{k+1}\|^{2}\Delta t \tag{d}
\end{align*}
\]\[
\begin{align*}
+ r_{3} \sum_{k=0}^{N-1} e^{rk\Delta t}\|\nabla v_{2}^{k+1}\|^{2}\Delta t + \frac{1}{2}\|\psi_{U}^{N}\|^{2} + \frac{1}{2}\sum_{k=0}^{N-1}\|\partial_{\nu}\psi_{U}^{k+1}\|^{2}(\Delta t)^{2} \tag{e}
\right)
\end{align*}
\]\[
\begin{align*}
+ r_{3} \sum_{k=0}^{N-1} e^{rk\Delta t}\|\nabla v_{2}^{k+1}\|^{2}\Delta t + \frac{1}{2}\|\psi_{U}^{N}\|^{2} + \frac{1}{2}\sum_{k=0}^{N-1}\|\partial_{\nu}\psi_{U}^{k+1}\|^{2}(\Delta t)^{2} \tag{f}
\right)
\end{align*}
\]\[
\begin{align*}
+ \frac{1}{2}\sum_{k=0}^{N-1}\left\|\sqrt{\tau_{b}}\left( \psi_{U}^{k+1} \right) \right\|^{2}\Delta t + \mu\left( \beta_{1} + \frac{\beta_{0}}{2} \right) \sum_{k=0}^{N-1}\left\|\nabla\psi_{U}^{k+1}\right\|^{2}\Delta t \tag{g}
\right)
\end{align*}
\]\[
\begin{align*}
\leq \left( \frac{\tau_{0}^{2}h^{2} + 2K_{0}\mu\beta_{0}}{4\Delta t^{2}} \right) \sum_{k=0}^{N-1}\left\|\partial_{\nu}\psi_{U}^{k+1}\right\|^{2}(\Delta t)^{3} + \epsilon\left\|v_{1}^{0}\right\|^{2} + \epsilon\left\|\nabla v_{1}^{0}\right\|^{2} + \epsilon\sum_{k=0}^{N-1}\left\|\nabla\psi_{U}^{k+1}\right\|^{2}\Delta t
\right)
\end{align*}
\]\[
\begin{align*}
+ \epsilon\sum_{k=0}^{N-1}\left\|\psi_{x}^{k}\right\|^{2}\Delta t + \epsilon\sum_{k=0}^{N-1} e^{-rk\Delta t}\|\nabla\psi_{x}^{k}\|^{2}\Delta t + \epsilon\sum_{k=0}^{N-1}\left\|\nabla\psi_{U}^{k+1}\right\|^{2}\Delta t \tag{h}
\right)
\end{align*}
\]\[
\begin{align*}
+ K_{1} \sum_{k=0}^{N-1}\left\|\nabla v_{1}^{k+1}\right\|^{2}\Delta t + K_{2} \sum_{k=0}^{N-1} e^{rk\Delta t}\|v_{2}^{k+1}\|^{2}\Delta t
\right)
\end{align*}
\]\[
\begin{align*}
+ K_{3} \sum_{k=0}^{N-1}\left\|\nabla v_{2}^{k+1}\right\|^{2}\Delta t + K \sum_{k=0}^{N-1}\left( \|\theta_{x}^{k}\|^{2}_{H^{1}(\Omega)} + \|\partial_{t}\nabla\theta_{x}^{k}\|^{2} \right)\Delta t
\right)
\end{align*}
\]\[
\begin{align*}
+ K \sum_{k=0}^{N}\left( \|\theta_{U}^{k}\|^{2}_{H^{1}(\Omega)} + \|\theta_{U}^{k+1}\|^{2}_{H^{1}(\Omega)} \right)\Delta t
\right)
\end{align*}
\]\[
\begin{align*}
+ K(\Delta t)^{4}\left\|\frac{\partial\xi}{\partial t} \right\|_{L^{2}(J_{N+1};L^{2}(\Omega))} + K(\Delta t)^{2}\left\|\frac{\partial^{2}\xi}{\partial t^{2}} \right\|_{L^{2}(J_{N+1};H^{1}(\Omega))} \tag{i}
\right)
\end{align*}
\]\[
\begin{align*}
+ K(\Delta t)^{2}\left\|\nabla\frac{\partial\xi}{\partial t} \right\|_{L^{2}(J_{N+1};L^{2}(\Omega))} + K(\Delta t)^{3}\left\|\frac{\partial^{3}\xi}{\partial t^{3}} \right\|_{L^{\infty}(J_{N+1};L^{2}(\Omega))}
\right)
\end{align*}
\[ + K(\Delta t)^2 \left\| \frac{\partial^2 U}{\partial t^2} \right\|^2_{L^2(J_{N+1}; L^2(\Omega))} + K(\Delta t)^2 \left\| \frac{\partial U}{\partial t} \right\|^2_{L^2(J_{N+1}; H^1(\Omega))} + K \sum_{k=0}^{N} \Lambda^k \Delta t + K \left( h^{2(t-1)} + \Delta t^2 \right) + K \left( h^{2(t-1)} + \Delta t^2 \right) \Delta t^2. \]  

(4.40)

Assume, in addition, that \( \Delta t \leq 2h^2 \left( \tilde{r}^* h^2 + 2K_0 \mu \beta_0 \right)^{-1} = \kappa \) so that

\[ \sigma_4 = \frac{1}{2} - \left( \frac{\tilde{r}^* h^2 + 2K_0 \mu \beta_0}{4h^2} \right) \Delta t \geq 0. \]

Using the above choice for \( \Delta t \), hiding all terms multiplied by \( \epsilon \), and observing that terms (a)-(h) are all non-negative, we can write (4.40) as follows

\[ \frac{1}{2} e^{-rN\Delta t} \left\| \psi_{\xi}^N \right\|^2 + \left( \frac{\sigma_0}{2} + \frac{\sigma_1}{2} e^{\sigma_0} \right) \sum_{k=0}^{N} e^{-r k \Delta t} \left\| \psi_{\xi}^k \right\|^2 \Delta t \]

\[ + \frac{1}{2} \left\| \psi_U^N \right\|^2 + \sigma_4 \sum_{k=0}^{N-1} \left\| \partial_{t^k} \psi_U^{k+1} \right\|^2 (\Delta t)^2 + \frac{\mu}{2} \left( \beta_1 + \frac{\beta_0}{2} \right) \sum_{k=0}^{N-1} \left\| \nabla \psi_U^{k+1} \right\|^2 \Delta t \]

\[ \leq K \left\| \theta_{\xi}^2 \right\|_{L^2(0, T-\Delta t; H^1(\Omega))}^2 + K \sum_{k=0}^{N-1} \left\| \partial_{t^k} \nabla \psi_{\xi}^k \right\|^2 \Delta t \]

\[ + K \left\| \theta_U \right\|^2_{L^2(0, T-\Delta t; H^1(\Omega))} + K \left\| \theta_U \right\|^2_{L^2(J_{N+1}; L^2(\Omega))} \]

\[ + K(\Delta t)^2 \left\| \frac{\partial^4 \xi}{\partial t^4} \right\|^2_{L^2(J_{N+1}; L^2(\Omega))} + K(\Delta t)^2 \left\| \frac{\partial^2 \xi}{\partial t^2} \right\|^2_{L^2(J_{N+1}; H^1(\Omega))} \]

\[ + K(\Delta t)^2 \left\| \frac{\nabla \xi}{\partial t} \right\|^2_{L^2(J_{N+1}; L^2(\Omega))} + K(\Delta t)^3 \left\| \frac{\partial^3 \xi}{\partial t^3} \right\|^2_{L^\infty((0,\Delta t); L^2(\Omega))} \]

\[ + K(\Delta t)^2 \left\| \frac{\partial^2 U}{\partial t^2} \right\|^2_{L^2(J_{N+1}; L^2(\Omega))} + K(\Delta t)^2 \left\| \frac{\partial U}{\partial t} \right\|^2_{L^2(J_{N+1}; H^1(\Omega))} \]

\[ + K \sum_{k=0}^{N} \Lambda^k \Delta t + K \left( h^{2(t-1)} + \Delta t^2 \right) + K \left( h^{2(t-1)} + \Delta t^2 \right) \Delta t^2. \]  

(4.41)

Recalling that

\[ \Lambda^N = e^{-rN\Delta t} \left\| \psi_{\xi}^N \right\|^2 + \left\| \psi_U^N \right\|^2 + \left\| \psi_0^N \right\|^2_{H^1(\Omega)}, \]

apply the generalized discrete Gronwall inequality and approximation properties of the \( L^2 \) projection to obtain
\[
\left\| \psi^N \right\|^2 + \left\| \psi_\xi \right\|^2_{L^2(J_N^N; H^1(\Omega))} + \left\| \psi_U^N \right\|^2 + \left\| \nabla \psi_U \right\|^2_{L^2(J_N^N; L^2(\Omega))} \\
\leq K \left\{ h^{2(t-1)} + \Delta t^2 \right\}.
\] (4.42)

The result of the theorem now follows by an application of the triangle inequality.

In the case \( s_1 \geq 3 \) and \( l > 2 \), we can now complete our induction argument.

\[
|U_h^N| \leq |\psi_U^N| + |\tilde{U}^N| \\
\leq K h^{-1}(h^{l-1} + \Delta t) + C^* \\
< C^{**}
\]

for \( h \) and \( \Delta t \) sufficiently small (with \( \Delta t = o(h) \)). A similar argument gives an upper bound for \( H_h^N \). For the lower bound on \( H_h^N \) we have

\[
H_h^N = \psi^N_\xi + \tilde{\xi}^N \\
\geq -K h^{-1}(h^{l-1} + \Delta t) + C_* \\
> C_{**},
\]

again for \( h \) and \( \Delta t \) sufficiently small (with \( \Delta t = o(h) \)).

**4.3.9. A Priori Error Estimate**

Following the paradigm set in the previous sections, we obtain the following result.
**Theorem 4.2** (Chippada, Dawson, Martínez, Wheeler - 12/96) Let $0 \leq s_0 \leq z$, $s_0 \leq \ell \leq s_1$, $0 \leq z < s_1$. Let $(\xi^k, U^k)$ be the solution to (4.4)-(4.5), (4.6), at time $t = t^k$. Let $(\xi_h^k, U_h^k)$ be the Galerkin approximations to $(\xi^k, U^k)$. If $\xi^k \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega) \cap \mathcal{W}_\infty^1(\Omega)$, $U^k \in \mathcal{H}_0^1(\Omega) \cap \mathcal{H}^l(\Omega) \cap \mathcal{W}_\infty^1(\Omega)$, $\xi_h^k \in \mathcal{S}_h(\Omega)$, $U_h^k \in \mathcal{S}_h(\Omega)$ for each $k$; with reasonable assumptions on surface and body forces, and with $\Delta t, h$ sufficiently small, then $\exists$ a constant $\bar{K} = \bar{K}(T, s_1, r, K_*, K^{**})$ such that

$$
\begin{align*}
\| \xi^N - \xi_h^N \| + \| \xi - \xi_h \|_{\mathcal{S}(J_N; \mathcal{H}^1(\Omega))} \\
+ \| U^N - U_h^N \| + \| \nabla U - \nabla U_h \|_{\mathcal{S}(J_{N-1}; \mathcal{L}^2(\Omega))} \\
\leq \bar{K} \left( h^{\ell - 1} + \Delta t \right).
\end{align*}
$$

where, $\Delta t \leq \min\{\sigma(h), \kappa_1, \kappa_2\}$, where $\kappa_1 = 8/\tau_0$, and $\kappa_2 = 2h^2 (\delta^{*} h^2 + 2K_0\mu\beta_0)^{-1}$. Moreover, for $h, \Delta t$ sufficiently small, $s_1 \geq 3$ and $l > 2$, then

$$
\bar{K} = \bar{K}(T, s_1, r, K_*, K^{**}).
$$

The proof of this theorem is similar to that used in proving Theorem 4.1, except that an induction argument is used to obtain the $K_*, K^{**}$ bounds on the approximations.

Observe that the time step restriction is less severe when the diffusion term in the momentum equation is completely implicit ($\beta_0 = 0$), whereby the time step dependence on the mesh size is nullified in $\kappa_2$. 

It is important to point out that for technical reasons, full handling of centering at $k$ in the GWCE and centering at $k + \frac{1}{2}$ in the CME was discarded in the derivation of Theorem 4.2.

4.4. Recent Applications of Error Estimates.

Recently, Wheeler et al [18], proposed a projection method for improving the local mass-conservation of a velocity field. Maintaining mass-conservation is important in transport problems in which some type of bio-mass is being tracked. Moreover, issues of mass-conservation become vital when one considers that, in hydrology, usually a hydrodynamic model is coupled to some water-quality model.

In the projection method, the mass-conserving velocity was the sum of a velocity obtained from some means plus a correction, $\Gamma$, obtained by a mixed finite element method.

In the application of this idea, the new velocity, $U_h^{new}$, was constructed from $\Gamma$ plus the ADCIRC velocity $U_h$. Moreover, it was easily shown that the task of determining the error in the new velocity was reduced to an approximation theory problem. In particular, if $\Pi_h U_h$ is the $\Pi_h$-projected velocity, then

$$||U_h^{new} - U|| \leq ||U_h - U|| + ||\Pi_h U_h - U||.$$ 

The $\Pi_h$-projection is well known in the MFEM literature; see, for example, [15] and in particular it satisfies

$$(\nabla \cdot (U - \Pi_h U_h), \zeta_h) = 0, \quad \zeta_h \in W_h,$$

where $\Pi_h U_h, \Gamma \in V_h$, and $(V_h, W_h)$ form the lowest-index Raviart-Thomas space. Moreover, it is known that

$$||\Pi_h U_h - U|| \leq Ch^{j+1},$$

where $0 \leq j$ corresponds to the index of the Raviart-Thomas spaces.
Chapter 5

Finite Element Model of C-SWE

Using a semi-implicit scheme in the ADCIRC (GWCE-NCME) hydrodynamic simulator forces the use of small time-steps on the order of 15-45 seconds to maintain stability. Clearly, either fully implicit time-stepping schemes or some other type of method must be used to achieve larger time steps. Still, ADCIRC can’t effectively handle highly-advective flow. In this application, advective acceleration usually occurs in narrow inlets, such as in the Houston Ship Channel in Galveston Bay.

To that end, we propose to use characteristics methods as a means to improve time-stepping and as a means to handle the advective terms. Moreover, we will derive a priori error estimates for the Characteristic-Galerkin finite element method for the primitive formulation. We first simplify the viscosity term in the (NCME) [9]. Moreover, to avoid technical difficulties associated with implementing boundary conditions along characteristics we shall assume that the solution is $\Omega$-periodic. Hereafter, we shall understand that each Sobolev space is a periodic Sobolev space on $\Omega$ and understand the meaning of the associated norms accordingly.

Let $\tau$ be a unit vector in the direction $(u,1)$ so that $\tau = \frac{1}{\alpha}(u,1)$, with $\alpha = \sqrt{|u|^2 + 1}$. Then, interpret

\[
\alpha \frac{\partial H}{\partial \tau} = \frac{\partial H}{\partial t} + u \cdot \nabla H,
\]

\[
\alpha \frac{\partial u}{\partial \tau} = \frac{\partial u}{\partial t} + u \cdot \nabla u
\]

as directional derivatives of $H$ and $u$ respectively, in the direction $\tau$ as similarly done in [24, 21]. It is well known that it is the much smaller norms of $\frac{\partial^2 H}{\partial \tau^2}$ and $\frac{\partial^2 u}{\partial \tau^2}$ compared to the norms of $\frac{\partial^2 H}{\partial t^2}$ and $\frac{\partial^2 u}{\partial t^2}$ obtained in standard time-stepping procedures, in the
presence of advection-dominated flow, that allow larger time-steps to be taken in MMOC.

Now, write (CE)-(NCME) as

$$\frac{\partial H}{\partial \tau} + H(\nabla \cdot u) = 0. \quad (5.3)$$
$$\alpha \frac{\partial u}{\partial \tau} + g\nabla (H - h_b) - \mu \Delta u + \tau_b u + f_c k \times u + \mathcal{F} = 0. \quad (5.4)$$

This form, in which characteristics are used for both the elevation and velocity solutions, seems to be consistent with a similar splitting shown in [4] to be stable under certain conditions.

5.1. Weak Formulation.

Approximate the directional derivatives with a backward difference by taking a first-order Taylor expansion of $H(\hat{x}, t^{k-1})$ and $u(\hat{x}, t^{k-1})$ about $(x, t^k)$:

$$\alpha \frac{\partial H^k}{\partial \tau} \approx \frac{H(x, t^k) - H(\hat{x}, t^{k-1})}{\Delta t}, \quad (5.5)$$
$$\alpha \frac{\partial u^k}{\partial \tau} \approx \frac{u(x, t^k) - u(\hat{x}, t^{k-1})}{\Delta t}, \quad (5.6)$$

and define the following truncation terms

$$\zeta^k = \alpha \frac{\partial H^k}{\partial \tau} - \frac{H(x, t^k) - H(\hat{x}, t^{k-1})}{\Delta t}, \quad (5.7)$$
$$\sigma^k = \alpha \frac{\partial u^k}{\partial \tau} - \frac{u(x, t^k) - u(\hat{x}, t^{k-1})}{\Delta t}. \quad (5.8)$$

For simplicity, let $\mathcal{f} = f(\hat{x})$. At time $t = t^k$, we can write (5.3)-(5.4) in one weak form as

$$\left( \left( \frac{H^k - \hat{H}^{k-1}}{\Delta t} \right), v \right) + \left( H^k (\nabla \cdot u^k), v \right) + \left( \zeta^k, v \right) = 0, \quad \forall v \in H^1(\Omega), \ k \geq 1, \quad (5.9)$$

$$\left( \left( \frac{u^k - \hat{u}^{k-1}}{\Delta t} \right), w \right) - \left( g(h^k - h_b), \nabla \cdot w \right) + \left( \mu (\nabla u^k, \nabla w) \right) + \left( \tau_b u^k, w \right) + \left( f_c k \times u^k, w \right) + \left( \mathcal{F}^k, w \right) + \left( \sigma^k, w \right) = 0, \quad \forall w \in H^1_0(\Omega), \ k \geq 1. \quad (5.10)$$
and let $H^k, u^k$ also satisfy the following initial conditions

$$H(x, 0) = H_0(x), \quad u(x, 0) = u_0(x).$$

Before proceeding, we need to list some assumptions. Suppose that assumptions M1, M3, M4 hold for $t = t^k$ and that for $(x, t) \in \bar{\Omega} \times J_N^k$,

R1. the solutions $(H, u)$ to (5.9)-(5.11) exist and are unique,

R2. $H_0(x) \in H^1_0(\Omega),$

R3. $u_0(x) \in H^1_0(\Omega),$

R4. $H(x, \cdot) \in H^1_0(\Omega) \cap H^t(\Omega) \cap W^1_{\infty}(\Omega), \quad t \in (0, T),$

R5. $u(x, \cdot) \in H^1_0(\Omega) \cap H^t(\Omega) \cap W^1_{\infty}(\Omega), \quad t \in (0, T)$

5.2. Characteristic-Galerkin.

Define the modified method of characteristics approximating (5.3)-(5.4) to be the maps $H_h : J_N \rightarrow S_h, u_h : J_N \rightarrow S_h$. Let the approximate characteristic be denoted by

$$\dot{x} = x - u_h^k(x) \Delta t.$$  \hfill (5.11)

Let $\tilde{f} = f(\dot{x})$. Then, $H_h^k, u_h^k$ satisfy

$$\left( \frac{H_h^k - \tilde{H}_h^{k-1}}{\Delta t}, v \right) + \left( H_h^{k-1}(\nabla \cdot u_h^{k-1}), v \right) = 0, \quad \forall v \in S_h, \quad k \geq 1, \quad (5.12)$$

$$\left( \frac{u_h^k - \tilde{u}_h^{k-1}}{\Delta t}, w \right) - \left( g(H_h^k - h_b), \nabla \cdot w \right) + \mu \left( \nabla u_h^k, \nabla w \right) + \left( \tau_h u_h^k, w \right)$$

$$+ \left( f_c k \times u_h^k \right) + \left( \mathcal{F}^{k-1}, w \right) = 0, \quad \forall w \in S_h, \quad k \geq 1, \quad (5.13)$$

with initial conditions

$$H_h^0 = \tilde{H}(x, 0), \quad u_h^0 = \tilde{u}(x, 0),$$

$$\hat{H}_h^0 = \tilde{H}(\hat{x}, 0), \quad \hat{u}_h^0 = \tilde{u}(\hat{x}, 0).$$
In the sections that follow, we will derive an *a priori* error estimate for the Characteristic-Galerkin method described here.

### 5.2.1. CE-NCME Error Equations and Choice of Projections.

First, define the additional truncation term

\[
\delta^k_i = H^k(\nabla \cdot u^k) - H^{k-1}(\nabla \cdot u^{k-1}) = \int_{t^{k-1}}^{t^k} \frac{\partial H(\nabla \cdot u)}{\partial t} \, dt.
\]

Given \( \mathcal{L}^2 \) projections \( \hat{H} \in \mathcal{S}_h \), and \( \hat{u} \in \mathcal{S}_h \), we denote the affine errors in elevation and velocity as

\[
\psi_H = (H_h - \hat{H}) \quad \text{and} \quad \psi_u = (u_h - \hat{u}).
\]

respectively; and we also denote the approximation errors in elevation and velocity as

\[
\theta_H = (H - \hat{H}) \quad \text{and} \quad \theta_u = (u - \hat{u}).
\]

Let us proceed, as in Section 4.2, with an inductive argument. Suppose that \( \hat{H}^k, \hat{u}^k \) (\( \forall t = t^k \)) are defined as in (3.2) \( \forall t^k \) and have the properties stated in Lemma 3.2. Next, we assume that for \( k = 0, \ldots, (N-1) \) M1 holds for each \( t = t^k \), and

**S1.** \[ |u^k_h| < K^{**}, |\nabla u^k_h| < K^{**}. \]

We immediately have the base case for the inductive proof: \( K^{**} < H^0_h < K^{**} \) and \(|\nabla u^0_h| < K^{**}\). Moreover, it follows that \(|\nabla \cdot u^0_h| < K^{**}\). In the remainder of this section, from the derivation of the *a priori* error estimate, we prove the hypothesis for time \( t = t^N \).

Now, write down the error equations resulting from subtracting (5.9)-(5.10) from (5.12)-(5.13), respectively, as

\[
\left( \left( \begin{array}{c} \psi_{H}^k - \psi_{H}^{k-1} \\ \Delta t \end{array} \right), v \right)
\]
\begin{align*}
&= \left( \left( \frac{\theta^k_{H} - \theta^k_{H}^{-1}}{\Delta t}, v \right), v \right) - \left( \left( \frac{\dot{H}^k - \dot{H}^{k-1}}{\Delta t}, v \right), v \right) \\
&\quad + \left( (H^k_{H} - H_{H}^{k-1}) \nabla \cdot u_{h}^{k-1}, v \right) + \left( H^{k-1}_{H} \nabla \cdot (u^{k-1} - u_{h}^{k-1}), v \right) \\
&\quad - \left( \zeta^k + \delta^k_{1}, v \right), \quad \forall v \in \mathcal{S}_{h}, \ k \geq 1,
\end{align*}

(5.14)

and

\begin{align*}
&\left( \left( \frac{\psi^k_{u} - \psi_{u}^{k-1}}{\Delta t}, w \right) + \mu \left( \nabla \psi^k_{u}, \nabla w \right) + \left( \tau_{bh}^k \psi^k_{u}, w \right) \right) \\
&\leq \left( \left( \frac{\theta^k_{u} - \theta^k_{u}^{-1}}{\Delta t}, w \right) - \left( \frac{\dot{u}^{k-1} - \dot{u}^{k-1}}{\Delta t}, w \right) \right) \\
&\quad - \left( g(H^k_{H} - H_{H}^k), \nabla w \right) + \mu \left( \nabla \theta^k_{u}, \nabla w \right) + \left( \tau_{bh}^k u^k - \tau_{bh}^k \dot{u}^k, w \right) \\
&\quad + \left( f_c (k \times (u^k - u_{h}^k), w \right) + \left( \sigma^k, w \right), \quad \forall w \in \mathcal{S}_{h}, \ k \geq 1,
\end{align*}

(5.15)

where \( \psi^0_{H} = \psi^0_{H} = 0 \) and \( \psi^{0}_u(x) = \psi^{0}_u = 0 \).

5.2.2. Properties of Characteristic

Before bounding terms, we will need to show that the characteristic equations have certain properties.

From differential geometry [22], we recall the following definition.

**Definition 5.1** Let \( \mathcal{G} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n \). Then, \( \mathcal{G} \) is a **homeomorphism** if

- \( \mathcal{G} \) is continuous and bijective (injective and surjective),

- the inverse of \( \mathcal{G} \), \( \mathcal{G}^{-1} : \mathcal{G}(A) \subset \mathbb{R}^n \rightarrow \mathbb{R}^n \), is continuous.

We will also need the following theorems.

**Theorem 5.1** Suppose \( u^k \in \mathcal{W}^1_{\infty}(\Omega) \). For fixed \( t \in J, t \geq \Delta t \), let \( \mathcal{H}_{\mathcal{K}}(x) \equiv x - u^k(x) \Delta t \). Then, \( \mathcal{H}_{\mathcal{K}} \) is a differentiable homeomorphism for \( \Delta t \) sufficiently small.
Proof: See [24] for a proof of this result.

Theorem 5.2 Let assumption S1 hold. For fixed \( t \in J, t \geq \Delta t \), let
\[
H_{\mathbf{x}}(x) \equiv x - u_h^k(x) \Delta t.
\]
Then, \( H_{\mathbf{x}} \) is a differentiable homeomorphism for \( \Delta t \) sufficiently small.

Proof: In [50], the result of the theorem is established for the case \( H_{\mathbf{x}} \) - but using an extrapolant of \( u_h^k \) (from the last two pressure time levels) instead of \( u_h^k \) proper. We will closely follow those arguments in establishing the result of the theorem for the case \( H_{\mathbf{x}} \).

The Jacobian matrix of this transformation is
\[
D H_{\mathbf{x}}(x) = \begin{pmatrix}
1 - \frac{\partial u_h^k}{\partial x_1} \Delta t & -\frac{\partial u_h^k}{\partial x_2} \Delta t \\
-\frac{\partial u_h^k}{\partial x_1} \Delta t & 1 - \frac{\partial u_h^k}{\partial x_2} \Delta t
\end{pmatrix}
\]
and the Jacobian is just the determinant of this transformation matrix
\[
J_{H_{\mathbf{x}}}(x) = \left( 1 - \frac{\partial u_h^k}{\partial x_1} \Delta t \right) \left( 1 - \frac{\partial v_h^k}{\partial x_2} \Delta t \right) - \left( \frac{\partial u_h^k}{\partial x_1} \right) \Delta t \left( \frac{\partial v_h^k}{\partial x_2} \right) \Delta t
= 1 - \frac{\partial u_h^k}{\partial x_1} \Delta t - \frac{\partial v_h^k}{\partial x_2} \Delta t + \left( \frac{\partial u_h^k}{\partial x_1} \right) \left( \frac{\partial v_h^k}{\partial x_2} \right) \Delta t^2 - \left( \frac{\partial u_h^k}{\partial x_2} \right) \left( \frac{\partial v_h^k}{\partial x_1} \right) \Delta t^2
= 1 - \nabla \cdot u_h^k(x) \Delta t + O(\Delta t^2).
\]
(5.16)
The last equality follows from S1.

Thus, by inductive assumptions
\[
|J_{H_{\mathbf{x}}}(x) - 1| \leq K \Delta t, \quad \forall x \in \Omega.
\]
(5.17)

Now, observe
\[
H_{\mathbf{x}}(x) - H_{\mathbf{x}}(x') = x - x' - (u_h^k(x) - u_h^k(x')) \Delta t
= \nabla u_h^k(\mathbf{x})(x - x') \Delta t
= (x - x') \left( 1 - \nabla u_h^k(\mathbf{x}) \Delta t \right).
\]
(5.18)
The first equality results from the definition of $\mathcal{H}_\Delta$. The second equality results from using the Mean Value Theorem. But, for $\Delta t$ sufficiently small,

$$
(1 - \nabla u_h^k(\bar{x}) \Delta t) \geq \left(1 - \|\nabla u_h^k\|_{L^\infty(\Omega)} \Delta t \right) > 0.
$$

(5.19)

Then, given $x, x' \in \Omega$. $\Delta t$ sufficiently small, taking absolute values in (5.18) and using (5.19), we obtain

$$
|\mathcal{H}_\Delta(x) - \mathcal{H}_\Delta(x')| = |x - x'| \left|1 - \nabla u_h^k(\bar{x}) \Delta t \right|
= |x - x'| \left(1 - \nabla u_h^k(\bar{x}) \Delta t \right)
\geq |x - x'| \left(1 - \|\nabla u_h^k\|_{L^\infty(\Omega)} \Delta t \right).
$$

(5.20)

Thus, from assumption $S1$, we have

$$
|\mathcal{H}_\Delta(x) - x| = |u_h^k(x) \Delta t| = O(\Delta t).
$$

(5.21)

From (5.17) and the inverse function theorem, $\mathcal{H}_\Delta$ is locally a differentiable homeomorphism onto its image. From (5.20), for $\Delta t$ sufficiently small, $\mathcal{H}_\Delta$ is globally one-to-one (injective).

To show that $\mathcal{H}_\Delta$ is onto (surjective), we again follow closely the arguments of Russell [50]. Let $\tilde{\Omega}$ be the union of $\Omega$ and its neighboring periodic copies. Since $\tilde{\Omega}$ is compact, $\mathcal{H}_\Delta$ is a closed one-to-one mapping of $\tilde{\Omega}$ and is therefore globally a homeomorphism of $\tilde{\Omega}$ onto its image, $\mathcal{H}_\Delta(\tilde{\Omega})$. Now, since $\tilde{\Omega}$ is simply-connected, so is $\mathcal{H}_\Delta(\tilde{\Omega})$. Suppose that $\exists x_0 \in \Omega$ such that $x_0 \notin \mathcal{H}_\Delta(\tilde{\Omega})$. Let $\Gamma$ be a loop in $\tilde{\Omega}$ wrapping around $x_0$ at a distance greater than $\left(\|u_h^k\|_{L^\infty(\Omega)} \Delta t \right)$ from both $x_0$ and $\partial \tilde{\Omega}$. But, $\mathcal{H}_\Delta(\Gamma)$ is a loop in $\mathcal{H}_\Delta(\tilde{\Omega})$ which, by (5.21), still wraps around $x_0 \notin \mathcal{H}_\Delta(\tilde{\Omega})$. So we have a contradiction with our supposition since $\mathcal{H}_\Delta(\tilde{\Omega})$ is simply-connected (so it can’t have any holes, viz, it can’t exclude the point $x_0$) and thus $\Omega \subset \mathcal{H}_\Delta(\tilde{\Omega})$. Therefore, from periodicity, $\mathcal{H}_\Delta$ is onto. □
Now we can obtain the following generalization of a lemma due to Dawson et al [21].

**Lemma 5.1** Assume that \( u_h^k(x) \) has bounded first partial derivatives in space \( \forall k \), (assumption S1). Then, for \( \Delta t \) sufficiently small, an arbitrary function \( f \in L^2(\Omega) \) satisfies

\[
\frac{1}{2\Delta t} \left[ (\dot{f}, \dot{f}) - (f, f) \right] \leq K_1 \|f\|^2 + \epsilon \|f\|^2.
\]

where,

\[
K_1 = K_1 \left( \|\nabla \cdot u_h\|_{L^\infty(\Omega)} \right).
\]

**Proof:** Following closely the arguments of Lemma 3.1 in [21], let \( G_{\dot{x}}(x, t_k) = \dot{x} = x - u_h^k(x)\Delta t \). For each \( x \in \Omega \), and some fixed \( t \in J, t \geq \Delta t \), let

\[
y = G_{\dot{x}}(x, t) = \mathcal{H}_{\dot{x}}(x).
\]

From (5.16) and the boundedness assumptions on the first-order spatial partial derivatives of \( u_h \), observe that the inverse of the Jacobian is simply

\[
J_{\mathcal{H}_{\dot{x}}(x)}^{-1} = 1 + \nabla \cdot u_h^k(x) \Delta t + \mathcal{O}(\Delta t^2).
\]

Therefore, given that the differentiable homeomorphism \( \mathcal{H}_{\dot{x}}(x) \) maps the periodic \( \Omega \) into itself, consider the following change of variables:

\[
(\dot{f}, \dot{f}) = \int_{\Omega} f(y)f(y)dx = \int_{\Omega} f(y)f(y) \left| J_{\mathcal{H}_{\dot{x}}(x)}^{-1} \right| dy
\]

\[
= \int_{\Omega} f(y)f(y) \left[ 1 + \nabla \cdot u_h^k(x) \Delta t + \mathcal{O}(\Delta t^2) \right] dy.
\]

Now, subtracting \((f, f)\) from \((\dot{f}, \dot{f})\) yields

\[
\frac{1}{2\Delta t} \left[ (\dot{f}, \dot{f}) - (f, f) \right] = \frac{1}{2\Delta t} \left\{ \int_{\Omega} f(y)f(y) \left[ 1 + \nabla \cdot u_h^k(x) \Delta t + \mathcal{O}(\Delta t^2) \right] dy \right.
\]

\[
- \int_{\Omega} f(x)f(x)dx \right\}
\]
\[ W_1 = \frac{1}{2\Delta t} \left\{ \int_\Omega f(y)f(y) \left[ 1 + \nabla \cdot u_h^k(x) \Delta t + O(\Delta t^2) \right] \, dy \right. \\
\left. - \int_\Omega f(y)f(y) \, dy \right\} \\
= \frac{1}{2} \int_\Omega f(y)f(y) \left[ \nabla \cdot u_h^k(x) + O(\Delta t) \right] \, dy \\
= W_1 + W_2, \]

where the second equality comes from the fact that \( H_\Delta(x) \) is a differentiable homeomorphism onto itself.

In [21], term \( W_1 \) (with \( \nabla \cdot u(x) \) instead of \( \nabla \cdot u_h^k(x) \)) is bounded by first adding and subtracting \( \nabla \cdot u(y) \) to get two terms \( W_{1a} \) and \( W_{1b} \). The second term, \( W_{1b} \) was straightforward to bound using the assumption that \( \nabla \cdot u \) is bounded in \( L^\infty(\Omega) \). The first term \( W_{1a} \) was bounded using the Mean-Value Theorem on \( \nabla \cdot u \), assuming that \( \nabla(\nabla \cdot u) \) exists and is bounded in \( L^\infty(\Omega) \). Here, we weaken these assumptions by not splitting \( W_1 \) into two terms and instead writing

\[ W_1 = \int_\Omega f(y)f(y)(\nabla \cdot u_h^k(x)) \, dy = \int_\Omega f(y)f(y) \left( \nabla \cdot u_h^k(H^{-1}(y)) \right) \, dy \leq ||f||^2 \left| | |\nabla \cdot u_h^k|||_{L^\infty(\Omega)} \right| \]

\[ = K_1 \left( ||\nabla \cdot u_h^k|||_{L^\infty(\Omega)} \right) ||f||^2. \]

Now, note that for \( \Delta t \) sufficiently small,

\[ W_2 = \frac{O(\Delta t)}{2} \int_\Omega f(y)f(y) \, dy \leq \epsilon ||f||^2. \]

Therefore,

\[ \frac{1}{2\Delta t} \left[ (\dot{f}, \dot{f}) - (f, f) \right] \leq K_1 \left( ||\nabla \cdot u_h^k|||_{L^\infty(\Omega)} \right) ||f||^2 + \epsilon ||f||^2. \]

Remark: It is important to note that this lemma is a generalized form of Lemma 3.1 (here, with the coefficient for the time derivative equal to unity) in [21] in the sense that we don't assume the true velocity \( u \) is given.
We will also need to develop another technique based on the definition of the characteristic map, as done in Russell et al in [50, 24, 21]. First, let us introduce the following ideas.

Let $\alpha(x), \beta(x) \in \Omega$. For a general function $f(x)$ defined over $\Omega$ an application of Taylor’s theorem with integral remainder gives

$$f(\alpha(x)) - f(\beta(x)) = \int_{\beta(x)}^{\alpha(x)} \frac{\partial f}{\partial z}(z)\,dz$$

where $z$ is the unit vector in the direction $\beta(x) - \alpha(x)$. Letting $\bar{z} \in [0, 1]$ parametrize the segment from $\beta(x)[\bar{z} = 0]$ to $\alpha(x)[\bar{z} = 1]$, then

$$f(\alpha(x)) - f(\beta(x)) = \left[ \int_0^1 \frac{\partial f}{\partial z} \left((1 - \bar{z})\beta(x) + \bar{z}\alpha(x)\right) d\bar{z} \right] |\alpha(x) - \beta(x)|$$

$$\equiv I_f(\alpha(x), \beta(x)) |\alpha(x) - \beta(x)|, \quad (5.22)$$

where

$$I_f(\alpha(x), \beta(x)) = \int_0^1 \frac{\partial f}{\partial z} \left((1 - \bar{z})\beta(x) + \bar{z}\alpha(x)\right) d\bar{z} = \int_0^1 \frac{\partial f}{\partial z} \left(F_{\alpha, \beta}(x)\right) d\bar{z}.$$ 

with $F_{\alpha, \beta}(x) = (1 - \bar{z})\beta(x) + \bar{z}\alpha(x)$.

**Lemma 5.2** ($\alpha = \dot{x}, \beta = \ddot{x}$, See [24]) Let integer $p \in [1, \infty]$, then

$$\|I_f(\dot{x}, \ddot{x})\|_{C^p(\Omega)} \leq K \|\nabla f\|_{C^p(\Omega)}$$

**Proof.** Russell et al [24] establish that $F_{\dot{x}, \ddot{x}}(x)$ is a differentiable homeomorphism (except they use an extrapolated approximate velocity instead of the approximate velocity proper.) In fact, following their arguments outlined in the proof of Theorem 5.2, we see, from assumption R5 that

$$\left| J_{F_{\dot{x}, \ddot{x}}(x)} - 1 \right| \leq K \Delta t, \quad \forall x \in \Omega. \quad (5.23)$$
Moreover, for \( x, x' \in \Omega \),
\[
F_{\hat{x}, \hat{x}}(x) - F_{\hat{x}, \hat{x}}(x') = x - x' - \left[ (1 - \bar{z}) \left( u_h^k(x) - u_h^k(x') \right) + \bar{z} \left( u^k(x) - u^k(x') \right) \right] \Delta t \\
= (x - x') \left[ 1 - (1 - \bar{z}) \nabla u_h^k(\bar{x}) \Delta t - \bar{z} \nabla u^k(\bar{x}) \Delta t \right] \tag{5.24}
\]
where \( \hat{x}, \hat{x} \in (x, x') \).

Then, given \( x, x' \in \Omega \), \( \Delta t \) sufficiently small,
\[
\left| F_{\hat{x}, \hat{x}}(x) - F_{\hat{x}, \hat{x}}(x') \right| = |x - x'| \left| 1 - (1 - \bar{z}) \nabla u_h^k(\bar{x}) \Delta t - \bar{z} \nabla u^k(\bar{x}) \Delta t \right| \\
= |x - x'| \left( 1 - (1 - \bar{z}) \nabla u_h^k(\bar{x}) \Delta t - \bar{z} \nabla u^k(\bar{x}) \Delta t \right) \\
\geq |x - x'| \left( 1 - (1 - \bar{z}) \left\| \nabla u_h^k \right\|_{L^\infty(\Omega)} \Delta t \right) \\
- \bar{z} \left\| \nabla u^k \right\|_{L^\infty(\Omega)} \Delta t \right). \tag{5.25}
\]

Moreover, we have
\[
\left| F_{\hat{x}, \hat{x}}(x) - x \right| = \left| \bar{z} u^k(x) - (1 - \bar{z}) u_h^k(x) \right| \Delta t = O(\Delta t). \tag{5.26}
\]

Therefore,

Case \( \rho \in [1, \infty) \):
\[
\int_{\Omega} |I_f(x, \hat{x})|^\rho \, dx = \int_{\Omega} \left\| \int_0^1 \frac{\partial f}{\partial z} \left( F_{\hat{x}, \hat{x}}(x) \right) \, d\bar{z} \right\|^\rho \, dx \\
\leq \int_0^1 \int_{\Omega} \left\| \frac{\partial f}{\partial z} \left( F_{\hat{x}, \hat{x}}(x) \right) \right\|^\rho \, dx \, d\bar{z}.
\]

Letting \( y = F_{\hat{x}, \hat{x}}(x) \) and changing variables above, yields
\[
\int_{\Omega} |I_f(x, \hat{x})|^\rho \, dx \leq \int_0^1 \int_{\Omega} \left| \frac{\partial f}{\partial z} (y) \right|^\rho \left| J_{F_{\hat{x}, \hat{x}}(x)}^{-1} \right| \, dy \, d\bar{z} \\
\leq K \int_0^1 \int_{\Omega} \left| \frac{\partial f}{\partial z} (y) \right|^\rho \, dy \, d\bar{z}.
\]

Thus
\[
\int_{\Omega} |I_f(x, \hat{x})|^\rho \, dx \leq K \left\| \nabla f \right\|^\rho_{L^\rho(\Omega)}.
\]
Case $p = \infty$ is straightforward.

Thus, we obtain the following result.

**Lemma 5.3** Let $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. Let $g(x) \in L^p(\Omega)$, $I_f(x) \in L^q(\Omega)$, and $(\alpha(x) - \beta(x)) \in L^r(\Omega)$, with $(\alpha(x), \beta(x))$ as in Lemma 5.2. Then,

$$
\int_{\Omega} g(x) \left( f(\alpha(x)) - f(\beta(x)) \right) dx = \int_{\Omega} g(x) I_f(\alpha(x), \beta(x)) |\alpha(x) - \beta(x)| dx \\
\leq \|g\|_{L^p(\Omega)} \|I_f\|_{L^q(\Omega)} \|\alpha - \beta\|_{L^r(\Omega)} .
$$

### 5.2.3. Bounding the CE Error Equation

Choose test function $v = \psi_H^k$. Then, with $a = \psi_H^k$ and $b = \psi_H^{k-1}$, we have used the inequality $(a - b, a) \geq \frac{1}{2}(a^2 - b^2)$, and then added and subtracted $\|\psi_H^{k-1}\|^2$ to this inequality.

We can write (5.14) as follows,

$$
\frac{1}{2\Delta t} \left( \|\psi_H^k\|^2 - \|\psi_H^{k-1}\|^2 \right) \\
\leq \frac{1}{2\Delta t} \left( \|\hat{\psi}_H^{k-1}\|^2 - \|\psi_H^{k-1}\|^2 \right) \\
+ \left( \left( \frac{\theta_H^k - \theta_H^{k-1}}{\Delta t} \right), \psi_H^k \right) - \left( \left( \frac{\tilde{H}^{k-1} - \tilde{H}^{k-1}}{\Delta t} \right), \psi_H^k \right) \\
+ \left( (\theta_H^{k-1} - \psi_H^{k-1}) \nabla \cdot u_h^{k-1}, \psi_H^k \right) + \left( H^{k-1} \nabla \cdot (\theta_{u_{k-1}} - \psi_{u_{k-1}}), \psi_H^k \right) \\
- \left( \zeta^k, \psi_H^k \right) - \left( \omega^k, \psi_H^k \right) \\
= T_1 + \cdots + T_7. \tag{5.27}
$$

Thus, from Lemma 5.1, with $f = \psi_H^{k-1}$, we immediately have that

$$
T_1 \leq K_1 \|\psi_H^{k-1}\|^2 + \epsilon \|\psi_H^{k-1}\|^2 .
$$
The bounds on $T_4, T_5$ are straightforward, using Cauchy-Schwartz and the arithmetic-geometric mean inequality. Let $K_2 = K_2 \left( \| H^{k-1} \|_{L^\infty(\Omega)} \right)$, then
\[
T_4 \leq K_1 \left( \| \theta_H^{k-1} \| \right)^2 + K_1 \left( \| \psi_H^{k-1} \| \right)^2 + K_1 \left( \| \psi_H^{k} \| \right)^2,
\]
\[
T_5 \leq \epsilon \left( \| \nabla \psi_u^{k-1} \| \right)^2 + K_2 \left( \| \nabla \cdot \theta_u^{k-1} \| \right)^2 + K_2 \left( \| \psi_H^{k} \| \right)^2.
\]

Let $K_3 = K_3 \left( \| u^k \|_{L^\infty(\Omega)} \right)$. From the definition of the $L^2$ projection and continuity of $\theta_H^{k-1}$, we get
\[
T_2 = \frac{1}{\Delta t} \int_{\Omega} \left( \theta_H^{k} - \theta_H^{k-1} \right) \psi_H^k \, dx = \frac{1}{\Delta t} \int_{\Omega} \left( \theta_H^{k} - \theta_H^{k-1} \right) \psi_H^k \, dx
\leq \frac{1}{\Delta t} \left( \| \nabla \theta_H^{k-1} \| \right) \left( \| x - \hat{x} \|_{L^\infty(\Omega)} \right) \left( \| \psi_H^k \| \right) = \left( \| \nabla \theta_H^{k-1} \| \right) \left( \| u^k \|_{L^\infty(\Omega)} \right) \left( \| \psi_H^{k} \| \right)
\leq K_3 \left( \| \nabla \theta_H^{k-1} \| \right)^2 + K_3 \left( \| \psi_H^{k} \| \right)^2.
\]

Note, that the term $\| \nabla \theta_H^{k-1} \|$ accounts for the suboptimality of the error estimate we will derive. Russell et al in [50] and [21], handle a similar term by bounding $\left( \frac{\theta_H^{k} - \theta_H^{k-1}}{\Delta t} \right)$ in the $H^{-1}$ norm since then $\| \theta_H^{k} - \theta_H^{k-1} \|_{H^{-1}(\Omega)} \approx \| \theta_H^{k-1} \|$. However, the test function must then be measured in the $H^1$ norm. Since we won’t have a term on the LHS of the error equations in which to hide this latter term, we do not find it useful to apply a duality argument.

The bound on $T_3$ can be determined using a parametrization argument made in [24, 50]. In Lemma 5.3, choose $g = \psi_H^k$, $f = \Delta t^{-1} \tilde{H}^{k-1}$, $\alpha(x) = \tilde{x}$, and $\beta(x) = \hat{x}$, to get
\[
T_3 = \int_{\Omega} \left( \frac{\tilde{H}^{k-1} - \tilde{H}^{k-1}}{\Delta t} \right) \psi_H^k \, dx \leq \left( \| \psi_H^k \| \right) \left( \| \mathcal{I}_{\tilde{H}^{k-1}} \|_{L^\infty(\Omega)} \right) \left( \| u^k - u_h^k \| \right).
\]
where
\[
\mathcal{I}_{\tilde{H}^{k-1}}(\tilde{x}, \hat{x}) = \int_0^1 \frac{\partial \tilde{H}^{k-1}}{\partial \tilde{z}} ((1 - \tilde{z}) \hat{x} + \tilde{z} \tilde{x}) \, d\tilde{z},
\] and with
\[
\| \mathcal{I}_{\tilde{H}^{k-1}} \|_{L^\infty(\Omega)} \leq K \left( \| \nabla \tilde{H}^{k-1} \|_{L^\infty(\Omega)} \right).
\]
Finally, recalling Lemma 3.2. using the AGMI, and letting $K_4 = K_4 \left( \| \nabla \hat{H}^{-1} \|_{L^{\infty}(\Omega)} \right)$, we get

$$T_3 \leq K \| \psi_H^k \| \| \nabla \hat{H}^{-1} \|_{L^{\infty}(\Omega)} \| u_h^k - u^k \| \leq K_4 \| \psi_H^k \|^2 + K_4 \| \psi_u^k \|^2 + K_4 \| \theta_u^k \|^2.$$

In bounding $T_6$, let $K_5 = K_5 \left( \| \alpha^4 \|_{L^{\infty}(\Omega)} \right)$, and recall from [24] (obtained by a parametrization and change-of-variable argument) that

$$\| \zeta^k \|^2 \leq K_5 \int_{t_{k-1}}^{t_k} \left| \frac{\partial^2 H}{\partial t^2} \right|^2 dt$$

to get

$$T_6 \leq K_5 \int_{t_{k-1}}^{t_k} \left| \frac{\partial^2 H}{\partial t^2} \right|^2 dt + K_5 \| \psi_H^k \|^2.$$

To bound $T_7$, recall that we can write

$$\| \delta_1^k \|^2 \leq \Delta t \int_{t_{k-1}}^{t_k} \left| \frac{\partial H(\nabla \cdot u)}{\partial t} \right|^2 dt.$$

So that,

$$T_7 \leq K_4 \Delta t \int_{t_{k-1}}^{t_k} \left| \frac{\partial H(\nabla \cdot u)}{\partial t} \right|^2 dt + K_4 \| \psi_H^k \|^2.$$

Finally, we obtain that the RHS of (5.27) is bounded above by

$$(4K_1 + 2K_2 + K_3 + K_4 + K + \epsilon) \| \psi_H^k \|^2 + \epsilon \left( \| \nabla \psi_u^{k-1} \|^2 + K_4 \| \psi_u^k \|^2 + K_4 \| \theta_u^{k-1} \|^2 + K_2 \| \nabla \theta_u^k \|^2 \right.$$

$$+ K_1 \| \theta_H^{k-1} \|^2 + K_3 \| \nabla \theta_H^{k-1} \|^2 + K_5 \int_{t_{k-1}}^{t_k} \left| \frac{\partial^2 H}{\partial t^2} \right|^2 dt$$

$$+ K \Delta t \int_{t_{k-1}}^{t_k} \left| \frac{\partial H(\nabla \cdot u)}{\partial t} \right|^2 dt.$$  \hspace{1cm} (5.28)
5.2.4. Bounding the NCME Error Equation

Choose test function $w = \psi_u^k$. Then (5.15) can be written as follows

$$\frac{1}{2\Delta t} \left( \|\psi_u^k\|^2 \right) - \|\psi_u^{k-1}\|^2) + \mu \|\nabla \psi_u^k\|^2 + \|\tau_h \psi_u^k\|^2$$

$$\leq \frac{1}{2\Delta t} \left( \|\dot{\psi}_u^{k-1}\|^2 - \|\psi_u^{k-1}\|^2 \right)$$

$$+ \left( \left( \frac{\theta_u^k - \theta_u^{k-1}}{\Delta t} , \psi_u^k \right) - \left( \left( \frac{\dot{\theta}^{k-1}}{\Delta t} , \psi_u^k \right) 

- (g(\theta_H^k - \psi_H^k), \nabla \cdot \psi_u^k) + \mu (\nabla \theta_u^k, \nabla \psi_u^k) + (\tau_h^k \psi_u^k - \tau_h^{k-1} \psi_u^k, \psi_u^k)$$

$$+ (f \cdot k \times \theta_u^k, \psi_u^k) - (\sigma^k, \psi_u^k) \right)$$

$$= S_1 + \cdots + S_8. \quad (5.29)$$

Most of the terms on the RHS of (5.29) are bounded similarly to those on the RHS of (5.27). To get the bound on $S_6$, we recall the bound obtained for a similar term in [17]. Let $K_8 = K_8 \left( \|\tilde{u}^k\|_{L^\infty(\Omega)} \right)$, $K_9 = K_9(K_{\ast \ast}), K_{10} = K_{10} \left( \|u_h^k\|_{L^\infty(\Omega)} \right)$. Then,

$$S_6 \leq K_8 K_9 K_{10} \left[ \|\psi_H^k\|^2 + \|\theta^k_H\|^2 + \|\psi_u^k\|^2 \right]$$

$$+ K_8 K_9 K_{10} \left[ \|\tilde{u}^k\|_{L^\infty(\Omega)} \right]^2 + \|\psi_u^k\|^2 \right)$$

$$+ K_8 K_9 \left[ \|\theta_u^k\|^2 + \|\psi_u^k\|^2 \right]$$

Now, let $K_6 = K_6(g), K_7 = K_7 \left( \|\tilde{u}^k\|_{L^\infty(\Omega)} \right), K_{11} = K_{11}(f_c), K_{12} = K_{12}(\mu).$ After some effort, we obtain that the RHS of (5.29) is bounded above by

$$(K_1 + K_3 + 2K_7 + K_8 K_9 (K_3 + K_{10} + K) + K_{11} + K + \epsilon) \|\psi_u^k\|^2$$

$$+ \epsilon \|\nabla \psi_u^k\|^2 + (K_6 + K_8 K_9 (K_3 + K_{10})) \|\psi_H^k\|^2$$

$$+ (K_6 + K_8 K_9 (K_3 + K_{10})) \|\theta_H^k\|^2 + (K_3 + K_12) \|\nabla \theta_u^k\|^2$$

$$+ (K_7 + K_8 K_9 + K_{11}) \|\theta_u^k\|^2 + K_8 \Delta t \int_{t_{k-1}}^{t_k} \left\| \frac{\partial^2 u}{\partial t^2} \right\|^2 dt. \quad (5.30)$$
5.2.5. Bounding the Sum of the Error Equations

Finally, add together (5.27) and (5.29), using bounds (5.28) and (5.30). Multiplying
the resulting inequality by $\Delta t$ and sum over $k, k = 1, \ldots, N$ to obtain

$$
\frac{1}{2} \left| \psi^N_H \right|^2 + \frac{1}{2} \left| \Psi^N_u \right|^2 + \sum_{k=1}^{N} \left| \sqrt{\tau_k^*} \psi^k_u \right|^2 \Delta t + \mu \sum_{k=1}^{N} \left| \nabla \psi^k_u \right|^2 \Delta t
\leq \frac{1}{2} \left| \psi^0_H \right|^2 + \frac{1}{2} \left| \Psi^0_u \right|^2 + \epsilon \sum_{k=1}^{N} \left| \nabla \psi^k_u \right|^2 \Delta t + K_{13} \sum_{k=1}^{N} \left| \psi^k_H \right|^2 \Delta t
+ K_{14} \sum_{k=1}^{N} \left| \psi^k_u \right|^2 \Delta t + K_{15} \sum_{k=1}^{N} \left| \theta^k_H \right|^2_{H^1(\Omega)} \Delta t + K_{16} \sum_{k=1}^{N} \left| \theta^k_u \right|^2_{H^1(\Omega)} \Delta t
+ K_{17} \left| \frac{\partial u}{\partial t} \cdot \nabla h^0_b \right|^2_{L^2((0,T); L^2(\Omega))} + K_{18} \left| \frac{\partial H(\nabla u)}{\partial t} \right|^2_{L^2((0,T); L^2(\Omega))}
+ K_{19} \Delta t^2 \left| \frac{\partial^2 u}{\partial t^2} \right|^2_{L^2((0,T); L^2(\Omega))} + K_{20} \Delta t^2 \left| \frac{\partial^2 u}{\partial t^2} \right|^2_{L^2((0,T); L^2(\Omega))}.
$$

(5.31)

where

$$
K_{13} = K_{13} (K_1, \ldots, K_4, K_6, \ldots, K_{10}, K, \epsilon);
$$

$$
K_{14} = K_{14} (K_1, K_3, K_4, K_7, \ldots, K_{11}, K, \epsilon);
$$

$$
K_{15} = K_{15} (K_1, K_3, K_6, K_8, \ldots, K_{10});
$$

$$
K_{16} = K_{16} (K_2, \ldots, K_4, K_7, \ldots, K_9, K_{11}, K_{12}).
$$

Hide $\epsilon \sum_{k=1}^{N} \left| \nabla \psi^k_u \right|^2 \Delta t$ on the LHS of (5.31) and use the fact that $\psi^0_H = 0, \psi^0_u = 0$, to get

$$
\frac{1}{2} \left| \psi^N_H \right|^2 + \frac{1}{2} \left| \Psi^N_u \right|^2 + \left| \sqrt{\tau_k^*} \psi^k_u \right|_{L^2(J^0_{\psi}; L^2(\Omega))}^2 + \frac{\mu}{2} \left| \nabla \psi^k_u \right|_{L^2(J^0_{\psi}; L^2(\Omega))}^2
\leq K_{13} \sum_{k=1}^{N} \left| \psi^k_H \right|^2 \Delta t + K_{14} \sum_{k=1}^{N} \left| \psi^k_u \right|^2 \Delta t + K_{15} \left| \theta^k_H \right|_{L^2(J^0_{\theta}; H^1(\Omega))}^2
+ K_{16} \left| \theta^k_u \right|_{L^2(J^0_{\theta}; H^1(\Omega))}^2 + K \Delta t^2 \left| \frac{\partial u}{\partial t} \cdot \nabla h^b \right|_{L^2((0,T); L^2(\Omega))}^2.
$$
\[ + K_6 \Delta t^2 \left\| \frac{\partial H(\nabla \cdot u)}{\partial t} \right\|_{L^2((0,T);L^2(\Omega))}^2 + K_5 \Delta t^2 \left\| \frac{\partial^2 H}{\partial t^2} \right\|_{L^2((0,T);L^2(\Omega))}^2 \]

\[ + K_3 \Delta t^2 \left\| \frac{\partial^2 u}{\partial t^2} \right\|_{L^2((0,T);L^2(\Omega))}^2. \]

Finally, apply the generalized discrete Gronwall's Lemma to obtain

\[ \left\| \psi_H^N \right\|^2 + \left\| \psi_u^N \right\|^2 + \left\| \sqrt{\tau_n} \psi_u \right\|_{L^2(J_N;C^2(\Omega))}^2 + \left\| \nabla \psi_u \right\|_{L^2(J_N;C^2(\Omega))}^2 \]

\[ \leq \hat{K} \left[ K_{15} \left\| \theta_H^k \right\|_{L^2(J_N;H^1(\Omega))}^2 + K_{16} \left\| \theta_u^k \right\|_{L^2(J_N;H^1(\Omega))}^2 \right. \]

\[ + K_6 \Delta t^2 \left( \left\| \frac{\partial^2 H}{\partial t^2} \right\|_{L^2((0,T);L^2(\Omega))}^2 + \left\| \frac{\partial^2 u}{\partial t^2} \right\|_{L^2((0,T);L^2(\Omega))}^2 \right) \]

\[ + K \Delta t^2 \left( \left\| \frac{\partial u}{\partial t} \cdot \nabla h_b \right\|_{L^2((0,T);L^2(\Omega))}^2 + \left\| \frac{\partial H(\nabla \cdot u)}{\partial t} \right\|_{L^2((0,T);L^2(\Omega))}^2 \right), \tag{5.33} \]

where \( K_{17} = 2 \max\{K_{13}, K_{14}\} \), \( \hat{K} = \exp \left( \sum_{k=1}^{N} \left( \frac{K_{17}}{1 - \Delta t K_{17}} \right) \Delta t \right) \), and \( \Delta t \) is sufficiently small.

Therefore,

\[ \left\| \psi_H^N \right\| + \left\| \psi_u^N \right\| + \left\| \sqrt{\tau_n} \psi_u \right\|_{L^2(J_N;C^2(\Omega))} + \left\| \nabla \psi_u \right\|_{L^2(J_N;C^2(\Omega))} \leq \hat{K} \left( h^{\ell-1} + \Delta t \right). \]

To complete the proof, we use the same argument as in [17], namely, when \( \Delta t = o(h), s_1 \geq 3, \ell > 2 \), we obtain that

\[ \left| \phi^N \right| \leq \left| \psi_H^N \right| + \left| \bar{\phi}^N \right|, \quad \phi = \{H_h, u_h\} \]

\[ \leq K h^{-1} \left( h^{\ell-1} + \Delta t \right) + K^* \]

\[ < K^{**}, \]

and

\[ H_h^N = \psi_H^N + \bar{H}^N \]

\[ \geq -K h^{-1} (h^{\ell-1} + \Delta t) + K_*, \]

\[ > K^{**}. \]
Finally, balancing $\Delta t$ and $h$, we find that if $\Delta t = o(h^2)$, $s_1 \geq 4, \ell > 3$, then
\[
|\nabla \cdot u_h^N| \leq |\nabla u_h^N| \\
\leq |\nabla \psi_u^N| + |\nabla \tilde{u}^N| \\
\leq K \Delta t^{-1/2} h^{-1} \left( \sum_{k=0}^{N} ||\nabla \psi_u||^2 \Delta t \right)^{1/2} + K^* \\
\leq K \Delta t^{-1/2} h^{-1} (h^{l-1} + \Delta t) + K^* \\
< K^{**},
\]

Thus, we have proved the following:
Theorem 5.3 Let $0 \leq s_0 < \ell \leq s_1$. Let $\Omega$ be periodic. Let $(H^k, u^k)$ be the solution to (5.9)-(5.10) at time $t = t^k$. Let $(H^k_h, u^k_h)$ be the Characteristic-Galerkin approximations to $(H^k, u^k)$. If $H^k \in H^1(\Omega) \cap H^\ell(\Omega) \cap W_1^1(\Omega)$, $u^k \in H^1(\Omega) \cap H^\ell(\Omega) \cap W_1^1(\Omega)$, $H^k_h \in S_h(\Omega)$, $u^k_h \in S_h(\Omega)$ for each $k$; with reasonable assumptions on surface and body forces, and with $\Delta t$ sufficiently small, in particular $\Delta t = o(h)$, then $\exists$ a constant $\bar{K} = \bar{K}(T, s_1, K_*, K^*, K^{**})$ such that

$$
\|H^k_h - H^N_h\| + \|u^k_h - u^N_h\| \\
+ \left\|\sqrt{\tau_h}(u - u_h)\right\|_{L^2(\Omega)} + \|\nabla u - \nabla u_h\|_{L^2(\Omega)} \\
\leq \bar{K} \left(h^{\ell-1} + \Delta t\right).
$$

If $h, \Delta t$ are sufficiently small, $s_1 \geq 3$ and $\ell > 2$, then we can remove the boundedness assumptions on $u_h$ and on $H_h$, but not on $\nabla \cdot u_h$. Finally, for $h, \Delta t$ sufficiently small, $s_1 \geq 4$ and $\ell > 3$, then

$$
\bar{K} = \bar{K}(T, s_1, K_*, K^*).
$$

Remark: To improve the time-step size, it may be necessary to couple (5.9)-(5.10). In doing so, the truncation term $\delta_1$ reduce to zero (and don’t appear in the development of the estimate), removing from the estimate the contribution of the constants bounding these norms and thus weakening the assumptions of the solutions.
Chapter 6

Extensions and Future Directions

In this section, we ask various questions that arise in the work presented in this proposal.

Wave Formulation:

1. Can we numerically confirm the \textit{a priori} estimates developed here? Lynch and Gray [44] propose a standard framework to test convergence rates of numerical methods in which analytical solutions are known. However, this framework is based on the linearized form of the shallow water equations suggesting that we could only test the \textit{a priori} error estimate of the linearized model.

2. Can we easily extend the error estimates to include physical boundary conditions?

3. Can we analyze higher-order time-stepping schemes in the discrete-time case?

4. Immediately evident from the temporal discretization in (4.20)-(4.23) as well as from that given in the hydrodynamic simulator is the need to (a) make the GWCE more implicit, and (b) strongly couple the equations through the advective terms. This tight coupling would mean, of course, that we would have to iteratively solve, say by GMRES or some other Krylov-Newton iterative solver, a large nonlinear system of equations at every time-step (since now we have to include the coupled contributions from the GWCE).

Although it may be possible to improve the time-step size, \textit{ADCIRC} will still be unable to handle advective-dominated flow.
5. Can we apply a discontinuous-Galerkin method to the shallow water equations? Recently, the discontinuous-Galerkin method has been applied to Hyperbolic Conservation Laws with some success, Bey and Oden [10, 11] give a brief review.

The discontinuous-Galerkin method can be viewed as an application of the Galerkin method on each element so that, compared to the standard Galerkin, continuity across element boundaries is not enforced and therefore the elements are decoupled. Instead, solutions on neighboring elements are weakly coupled through flux across element boundaries. An obvious advantage of this approach would be that mass conservation could be enforced element by element.

There are various interesting issues to contend with that can be inferred from the derivation of this error estimate.

**Characteristic Formulation**

1. The most obvious first step to take is to try to numerically verify the *a priori* error estimate.

2. Can the periodicity assumption on $\Omega$ be removed? That is, can we handle realistic boundary conditions in the *a priori* error estimate and, if so, how will these affect the properties of the characteristic mappings, that is, will $f^k(\hat{x})$ always be defined?

Arbogast and Wheeler [6] proposed a characteristic-mixed method in which a mixed finite element method (with lowest-order Raviart-Thomas spaces) is used for the spatial discretization. They are able to implement more general boundary conditions with some care. Moreover, using a mixed finite element method as
the spatial discretization preserves local mass conservation whereas a Galerkin method does not.

3. Can we determine if the Characteristic-Galerkin approximation to the primitive formulation of the shallow water equations is still prone to spurious oscillations or are these removed?

4. Since, Zienkiewicz and Ortiz [64, 65] claim that their characteristic-based finite-element approach matches tidal data taken from the Severn Estuary. can we understand their approach on a theoretical basis? How does the Chorin projection help their method achieve large time-steps?

5. Does it serve any purpose to extend the method of characteristics to the wave formulation of the shallow water equations? Besides improving time-step size and handling of the advective terms, it is a more difficult matter to handle the second-directional derivative since the second directional derivative now looks like a plane.
Bibliography


