High-Order Numerical Methods in Lake Modelling

by

Derek Steinmoeller

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2014

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The physical processes in lakes remain only partially understood despite successful data collection from a variety of sources spanning several decades. Although numerical models are already frequently employed to simulate the physics of lakes, especially in the context of water quality management, improved methods are necessary to better capture the wide array of dynamically important physical processes, spanning length scales from $\sim$ 10 km (basin-scale oscillations) – 1 m (short internal waves). In this thesis, high-order numerical methods are explored for specialized model equations of lakes, so that their use can be taken into consideration in the next generation of more sophisticated models that will better capture important small scale features than their present day counterparts.

The full three-dimensional incompressible density-stratified Navier–Stokes equations remain too computationally expensive to be solved for situations that involve both complicated geometries and require resolution of features at length-scales spanning four orders of magnitude. The main source of computational expense lay with the requirement of having to solve a three-dimensional Poisson equation for pressure at every time-step. Simplified model equations are thus the only way that numerical lake modelling can be carried out at present time, and progress can be made by seeking intelligent parameterizations as a means of capturing more physics within the framework of such simplified equation sets. In this thesis, we employ the long-accepted practice of sub-dividing the lake into vertical layers of different constant densities as an approximation to continuous vertical stratification. We build on this approach by including weakly non-hydrostatic dispersive correction terms in the model equations in order to parameterize the effects of small vertical accelerations that are often disregarded by operational models. Favouring the inclusion of weakly non-hydrostatic effects over the more popular hydrostatic approximation allows these models to capture the emergence of small-scale internal wave phenomena, such as internal solitary waves and undular bores, that are missed by purely hydrostatic models.

The Fourier and Chebyshev pseudospectral methods are employed for these weakly non-hydrostatic layered models in simple idealized lake geometries, e.g., doubly periodic domains, periodic channels, and annular domains, for a set of test problems relevant to lake dynamics since they offer excellent resolution characteristics at minimal memory costs. This feature makes them an excellent benchmark to compare other methods against. The Discontinuous Galerkin Finite Element Method (DG-FEM) is then explored as a mid- to high-order method that can be used in arbitrary lake geometries. The DG-FEM can be interpreted as a domain-decomposition extension of a polynomial pseudospectral method

and shares many of the same attractive features, such as fast convergence rates and the ability to resolve small-scale features with a relatively low number of grid points when compared to a low-order method. The DG-FEM is further complemented by certain desirable attributes it shares with the finite volume method, such as the freedom to specify upwind-biased numerical flux functions for advection-dominated flows, the flexibility to deal with complicated geometries, and the notion that each element (or cell) can be regarded as a control volume for conserved fluid quantities. Practical implementation details of the numerical methods used in this thesis are discussed, and the various modelling and methodology choices that have been made in the course of this work are justified as the difficulties that these choices address are revealed to the reader. Theoretical calculations are intermittently carried out throughout the thesis to help improve intuition in situations where numerical methods alone fall short of giving complete explanations of the physical processes under consideration.

The utility of the DG-FEM method beyond purely hyperbolic systems is also a recurring theme in this thesis. The DG-FEM method is applied to dispersive shallow water type systems as well as incompressible flow situations. Furthermore, it is employed for eigenvalue problems where orthogonal bases must be constructed from the eigenspaces of elliptic operators. The technique is applied to the problem calculating the free modes of oscillation in rotating basins with irregular geometries where the corresponding linear operator is not self-adjoint.

Acknowledgements

I would like to thank Marek Stastna for continued great discussions and above and beyond supervision on these projects since the beginning of my Ph.D. work. Your forward thinking and hands-on approaches to science were what made much of the innovations and key results of this thesis possible. I'd like to thank Kevin Lamb for his excellent supervision and help with direction and seeing the big picture. You have set an excellent example for myself and other young scientists to follow by being sharp and focused on research at all times, while maintaining high standards on scientific products and pedagogy.

I would like to thank the Ontario Graduate Scholarship (OGS), the Ontario Graduate Scholarship in Science & Technology (OGSST), the Queen Elizabeth II Scholarship in Science & Technology, and the Shared Hierarchical Research Network (SHARCNET) for funding my research.

I owe many thanks to the organizers of the 2012 Gene Golub SIAM Summer School (G2S3): Frank Giraldo, Randall LeVeque, Michael Bader, and Jorn Behrens for offering me the unique opportunity to discuss related interests with other students and scientists who carry out scientific computing research in the geosciences. The chance to sharpen my skills in numerical methods while restoring confidence in myself as a programmer was well worth the trip. I'm especially grateful to Frank Giraldo for sharing some of his wisdom on discontinuous and continuous Galerkin methods with me and for agreeing to be my external examiner. His suggestion to me in 2012 to "suck it up and include the sound waves" was one piece of advice that resonated particularly strongly with me, since much of my research in the months that followed considered the artificial compressibility equations as a potential avenue for the work in this thesis (however, that work did not make the final cut!).

I would also like to thank Christopher Subich, Peter Diamessis, Jorge Escobar-Vargas, and Sumedh Joshi for being strong influences in the exciting research area of high-order numerical methods. They certainly made me feel not alone when struggling with the various intricacies of the field. Without the guidance of Christopher, my numerical linear algebra abilities would be quite un-scalable and stone-aged.

Much thanks is owed to the past and present administrative staff of the applied math department for making the administrative parts of graduate student life run smoothly. In particular, I'd like to thank Helen Warren, Ann Puncher, Rina Salazar, Carol Seely-Morrison, Marie O'Brien-Stockie, Stephanie Martin, Laura Frazee, and Maureen Fraser.

Finally, I would like to thank the many fellow UW applied math grad students that made life as a grad student quite enjoyable throughout my Master's and Ph.D., be it by

spending late nights working on assignments together or sharing a pitcher at the grad house. I would like to thank: Michael Dunphy, Killian Miller, Alex Shum, Kris Rowe, Nancy Soontiens, Anton Baglaenko, Jared Penney, Tim Rees, Vladimir Gerasik, Scott Rostrup, Antonio Sanchez, Rob Huneault, Ryan Morris, Easwar Magesan, Wentao Liu, Cam Christou and anyone else I might have missed.

Dedication

This thesis is dedicated to my wife, Jen, for always believing me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter, the key questions this thesis will attempt to answer are introduced and discussed. First, we explain the motivation of this work by introducing the problems associated with understanding the fluid dynamics of lakes, then possible numerical modelling techniques are explained. We then proceed by summarizing how the remainder of this thesis is structured. The goals and scope of the thesis are outlined, and the basic modelling assumptions are stated.

1.1 Motivation

Understanding the physical processes in lakes is of fundamental importance in a vast array of applications, ranging from water quality management to bio-geochemical cycling. Numerical modelling is perhaps the best tool available for improving the current partial understanding of lake dynamics.

The majority of mid-latitude lakes during the ice free seasons are primarily forced by the wind and surface heating, which injects energy on large scales [132]. Due to nonlinear steepening of basin-scale waves, these motions are subsequently cascaded over a vast range of length scales ranging from $\sim$ 10 km (basin-scale waves)–1 cm (small-scale turbulence), eventually leading to dissipation of the mechanical energy at the Kolmogorov length scale [27, 132]. While some of the resulting motions are quite irregular, many are coherent, with prominent examples being internal waves, Langmuir circulations and surface waves.

Due to the wide array of dynamics across multiple length scales, the physical processes in lakes remain only partially understood, despite successful data collection from a variety

of sources spanning several decades [57, 74, 91, 96, 97, 99, 132]. Many *in situ* measurements in lakes are made with equipment such as thermistor chains, water-level meters, and Acoustic Doppler Current Profilers. However, due to obvious real-world restrictions (e.g., budget constraints), these measurement tools can only be deployed in a relatively small number of locations in a given lake, leaving the vast majority of the lake unmeasured. Numerical models, thus, play a key role in helping to provide context and physical interpretations to the relatively sparse collection of data that can be feasibly measured in a lake. At the same time, these models provide insight into how the lake moves as a whole.

Modelling the motion of lakes numerically is a very difficult proposition. The dynamics of a lake are three-dimensional and are affected by rotation, density stratification, free-surface (moving boundary) effects, as well as the irregularly shaped bottom and side boundaries. As a result, the full Navier-Stokes equations remain too complicated to numerically solve directly, and therefore approximations must be made, as explained below.

Most numerical models of natural water bodies in three dimensions do not treat the free surface as a true moving boundary and use other techniques to approximate it [44]. Additionally, these models typically neglect vertical accelerations by invoking the hydrostatic approximation [44, 61, 26] in order to reduce the computationally expensive three-dimensional Poisson problem for the pressure to a two-dimensional one. Although the hydrostatic approximation is widely used, since the depth-to-length aspect ratio of lakes is typically small, i.e., $(H/L) \ll 1$, its use precludes the modelling of short wave dispersion. That is, the tendency for short waves to travel slower than long waves is not modelled. It is for this reason that, while existing numerical models such as the MITgcm [44] and ELCOM [61] when run 'in hydrostatic mode' are useful in predicting the large-scale motions of lakes, they fail to represent features with length scales $H \sim L \sim 10$ m, e.g., internal solitary waves and undular bores, which are inherently *non-hydrostatic*. These features are often either unresolved or misrepresented (artificially damped) by these models. It is worth mentioning however, that MITgcm does have fully non-hydrostatic capabilities [44], and with sufficient resolution, it should be expected to capture non-hydrostatic phenomena without issue.

The importance of understanding non-hydrostatic internal wave phenomena in lakes has become increasingly apparent in the physical limnology literature. For instance, it is known that internal wave energy is typically converted to small-scale turbulence and mixing by interaction with the bottom, shear instabilities, and wave breaking. These processes were reviewed by Wüest and Lorke [132]. This implies that non-hydrostatic internal waves represent a fundamental link between basin-scale and small-scale turbulent motions. The very recent study by Pannard *et al.* [91] has suggested that mixing due to recurrent internal waves in a small-sized lake has strong ecological consequences for phytoplankton

blooms that in turn affects the entire food web of the lake. The significance of a modelling methodology that can capture both the large scale motions and the emergence of non-hydrostatic internal wave phenomena is thus clear.

1.2 Numerical modelling techniques

If we assume that a decision has been made with respect to which model equations to solve, the next issue to address is which numerical method to use. If we assume the boundary is rectangular or can be easily mapped to a rectangle, Fourier and/or Chebyshev pseudospectral constitute a good choice of method, due to the fact that they give the highest order of approximation possible, have excellent resolution characteristics, and have low amounts of inherent dissipation [14]. If the goal is to represent the complicated geometry of an actual lake in a numerical model, then pseudospectral methods are no longer of utility. Moreover, it rapidly becomes clear that for general geometries, an element-based approach is needed to represent a complicated boundary in a piece-wise fashion.

In the physical oceanography and limnology communities, the most common choice of method is the finite volume method (FVM) (e.g., used in the MITgcm [44], ELCOM [61], and DIECAST [31] models). This popularity is owed to the fact that the FVM is relatively computationally inexpensive, allowing for three-dimensional computations. In addition, FVM is typically robust to a wide range of physical parameters, remaining stable where other methods would become unstable. Furthermore, it is possible to derive FVM schemes that conserve higher-order moments of the flow variables such as energy and enstrophy [4]. The FVM is also considered simple to implement in comparison to other methods. The trade-off of using FVM lies with the fact that FVM is constrained to low orders of approximation in general geometries [75], and as a result it is inherently dissipative. Although dissipation is computationally beneficial since it increases numerical stability, it may remove, or at the very least significantly alter, important physical features of advection-dominated flows.

As an alternative to the FVM, one may consider the use of the classical finite element method (FEM) since it allows for high orders of approximation on general geometries by adding more degrees of freedom to an element. Although it may be ideal for steady problems, the main difficulty with the finite element discretization of an unsteady problem is that the time-derivative operator is multiplied by the global mass matrix. This implies that the large global mass matrix must be inverted. At moderate to high resolution this may be too expensive to be done directly, and as a result an explicit semi-discrete scheme cannot be recovered [60]. On the other hand, it is worth noting that if one uses high-order

nodal finite element methods (i.e., spectral elements) with inexact quadrature integration then the global mass matrix is diagonal, and the issue is avoided. Another difficulty with FEM, which arises when the method is applied to advection equations, lies with the fact that the basis functions are symmetric in space. With the FVM, advection can be handled appropriately by using an upwind approximation, i.e., taking information from where it comes from [75]. In contrast, due to the spatial symmetry of the basis functions, the FEM is not ideal for problems where information has a preferred direction of propagation, and the resulting solutions may be inaccurate or unstable [60]. It should be noted that modern treatments of FEM seek to overcome this short-coming for advective problems by considering stabilization techniques such as the SUPG (streamline upwind/Petrov–Galerkin) method [62] as well as the class of entropy-based viscosity methods [88].

In this thesis, we choose the Discontinuous Galerkin finite element method (DG-FEM) as a mid- to high-order alternative to FVM for lake modelling of general geometries in Chapter 4. The idea behind DG-FEM is to make the FVM high-order by adding more degrees of freedom to a cell. A purely local scheme is recovered by duplicating the nodes that are shared between elements, and continuity of the solution between elements is imposed only in a weak sense using a suitable numerical flux [60]. The DG-FEM discretizations of partial differential equations (PDEs) then resemble FEM discretizations with the added benefit that global mass and stiffness matrices are not present, and only their local versions are required. Furthermore, the freedom in the choice of numerical flux allows schemes that are suited to problems where the information has a preferred direction of propagation. High-order DG-FEM solutions to non-hydrostatic coastal engineering problems have been previously considered by Eskilsson and Sherwin [42], Karniadakis and Sherwin [68], and Engsig-Karup *et. al.* [37]. Hydrostatic DG-FEM models for both baroclinic ocean [11, 21] and coastal ocean [66] flows have been considered in the framework of the hydrostatic primitive equations. Extension of the DG-FEM methodology to non-hydrostatic lake simulations thus appears natural.

1.3 Thesis goals and key contributions

We have so far pinpointed our interest on developing and applying numerical models that can capture both basin-scale and small-scale non-hydrostatic internal wave phenomena, but it has not yet been made clear how this can be accomplished. In this thesis, we consider the use of weakly non-hydrostatic layered models as a means of recovering a simplified set of model equations that can capture both large- and small-scale internal wave phenomena. Such a methodology has been previously employed by Brandt *et al.* [15] in their study of

internal waves in the Strait of Messina, by de la Fuente *et al.* [28] in their study of Poincaré and Kelvin waves in a rotating circular basin, and by Tomasson and Melville [120] in their process study of nonlinear and dispersive effects in the geostrophic adjustment of internal waves in a channel. Multilayer modelling approaches for both surface and internal waves have been considered by Lynett and Liu [81, 80].

Weakly non-hydrostatic layered models are said to be of 'Boussinesq-type', and are best described as an extension of traditional shallow water equations to include dispersive corrections to the hydrostatic pressure gradient. The original concept of a Boussinesq-type model can be traced back to the approximation made by the model's name-sake, Boussinesq [12] in an attempt to model surface solitary waves. A large portion of this thesis is dedicated to solving one- and two-layer Boussinesq-type model equations in physical situations corresponding to parameter values appropriate for small- to mid-sized lakes (Length scales between 5–25 km). The novelty of this work comes from attempting to apply mid- to high-order (polynomial orders between 4 and 8) numerical methods wherever possible. Additionally, two relevant side projects have been conducted during the course of the research that lead to this thesis, as will be explained below. Following **Chapter 2**, that covers background material, the main contributions to knowledge of this thesis (by chapter) are:

1. **Chapter 3: Pseudospectral methods for layered models in simple geometries**

   - The solution of a weakly non-hydrostatic single-layer model in periodic and annular domains with the high-order Fourier and Chebyshev pseudospectral methods. This work has appeared in the following publications:
     - D.T. Steinmoeller, M. Stastna, and K.G. Lamb. Fourier pseudospectral methods for 2D Boussinesq-type equations. *Ocean Modelling*, **52–53**:76–89, 2012 [110].
     - D.T. Steinmoeller, M. Stastna, and K.G. Lamb. Pseudospectral methods for Boussinesq-type equations in an annular domain with applications to mid-sized Lakes. *J. Comp. Sci.,* **4**:3–11, 2013 [111].
   - The solution of a weakly non-hydrostatic two-layer model in a periodic channel domain and the subsequent modelling of the emergence of undular bores and solitary waves. This work has not yet appeared in a publication.

2. **Chapter 4: Mid- to High-Order Discontinuous Galerkin methods for one-layer dispersive models in complex geometries**

- The solution of a weakly non-hydrostatic single-layer model in simple and irregular geometries with the Discontinuous Galerkin finite element method (DG-FEM). This work has not yet been published.

3. **Chapter 5: On Discontinuous Galerkin Methods for Incompressible Flow**

   - A linear eigenvalue analysis of the DG-FEM discretization of the pressure projection method for incompressible flow is carried out.
   - Simulations of the density-stratified inviscid incompressible Euler equations are performed with the DG-FEM. This work represents a natural extension of the publication where the linear eigenvalue analysis first appeared:
     - D.T. Steinmoeller, M. Stastna, and K.G. Lamb: A short note on the discontinous Galerkin discretization of the pressure projection operator in incompressible flow. *J. Comp. Phys.,* **251C**:480–486, 2013 [112].
   - Solutions of a weakly non-hydrostatic two-layer rigid-lid model with the DG-FEM are explored in relatively simple geometries. This work has not yet been published.

4. **Chapter 6: Calculating free modes of oscillation in complex geometries**

   - A robust method for calculating the free modes of oscillation in closed basins with arbitrary shape is presented. The underlying spatial discretization method is the DG-FEM method introduced in Chapter 4. This work has not yet been published.

The thesis is then closed with a summary and main conclusions given in Chapter 7. In the following section, the modelling choices made throughout the course of research are explained.

1.4 Modelling choices

The key modelling choices made throughout this work are introduced and explained point by point below.

1. **The numerical methods used in simulations are mid- to high-order.**

In simple geometries, pseudospectral methods are used to obtain the highest order of accuracy theoretically possible. In irregular geometries, the order of the local approximating polynomials is typically between 4 and 8.

High-order methods offer many advantages over low-order methods such as excellent resolution characteristics, fast convergence rates, and consequently, memory efficiency. Furthermore, low-order finite difference and finite volume numerical methods for geophysical flows represent a thoroughly studied and widely used set of numerical methods. A Ph.D. thesis represents an excellent opportunity to explore new methodologies, and here an attempt is made to show the utility of high-order pseudospectral and Discontinuous Galerkin methods in practical situations.

2. **The model equations do not contain damping terms.**

   The underlying reason for this choice is to avoid the possibility of arriving at a numerical model that is over-damped and thereby misses important physics. It should be noted that if a numerical method poorly treats the advective terms in fluid equations, then the method might be unstable. The easiest solution to address such a problem is to include damping terms that stabilize the scheme, rather than address the underlying issues with the advection scheme.

   In a sense, we seek *minimally dissipative*, stable numerical solutions. In the context of pseudospectral methods, this is achieved by using the smallest amount of modal filtering that is possible without observing instabilities for the physical situations of interest. With the high-order DG-FEM, a weakly diffusive numerical flux function (for inter-element coupling) combined with a local modal filter is used.

   On a side note, the physics of bottom roughness and viscosity can be included in the numerical codes at a later point if desired, and all the tools to do so are presented in this thesis. One complexity with including horizontal viscosity terms lies with the imposition of lateral boundary conditions. That is, on the basin scale, what is the appropriate second set of boundary conditions (in addition to no normal flow through vertical walls) to impose on the velocity field given that the no-slip boundary layers that would be present due to a molecular kinematic viscosity of $\nu \sim 10^{-6}$ $\mathrm{m^2 s^{-1}}$ would be impossible to resolve?

   One answer would be to impose zero shear stress at solid walls on the grounds that the no-slip layer is not be resolved. On the other hand, turbulence may still cause boundary-layer effects to be important in basin-scale lake simulations and parameterizations are often sought out [44]. We leave this interesting problem to other studies.

3. **The dispersive terms used in weakly non-hydrostatic layered models do not come from rigorous derivations.**

   The use of mixed space/time derivatives to model horizontal dispersive effects due to weak vertical accelerations are included in an *ad hoc* manner and are chosen to, at best, give a first-order correction to cause the dispersion relation to better agree with the linear dispersion characteristics from fully non-hydrostatic flow. In the surface waves literature there are well-established equation sets in the 'Boussinesq' hierarchy that are used in dispersive wave modelling [80, 90, 93, 83]. For models of internal waves, there is not yet a simple set of dispersive equations that is generally agreed upon and used in the literature. Indeed, it appears to be a somewhat controversial area (see [81] and [6] for an example of the controversy).

   The use of *ad hoc* dispersion terms can be seen as a starting point from which strides can be made to solve equations that are more asymptotically correct. The choice to use such terms in this work is mainly motivated by numerical feasibility, and to avoid the complexities associated with non-conservative products of the flow variables, third-order spatial derivatives, and large coupled linear systems of equations. These difficulties will be discussed in the thesis as well.

   On the other hand, it should be noted that *ad hoc* parameterizations and terms based on empirical models are a ubiquitous part of modern numerical models. One common example is the quadratic bottom drag law $\tau = C_D \mathbf{u}\|\mathbf{u}\|/h$ [7] that is used repeatedly, but never derived from the Navier–Stokes equations. See Baglaenko *et al.* [5] for a detailed discussion of the quadratic bottom drag law. If it is acceptable to use *ad hoc* damping terms, then why not *ad hoc* dispersive terms as long as they are physically motivated?

4. **The solutions of interest are assumed to be *smooth*.**

   This assumption goes hand-in-hand with the use of high-order methods. It is quite well-known that if a function contains a discontinuity, then its projection onto a finite set of smooth approximating basis functions (e.g., Fourier modes or polynomials), will contain Gibbs oscillations and the convergence rate of the approximation will be slow.

   It will be demonstrated that the use of dispersive terms precludes the formation of shocks from smooth initial data to motivate this assumption. However, there are exceptions to this caveat that will be discussed as well, such as in the vicinity of wetting/drying fronts where dispersive effects are negligible, or when the flow speed exceeds the breaking limit. In such cases where shocks are present in the physically relevant weak solution, it is often argued that a piece-wise linear approximation combined with slope limiting is the best numerical approximation one can make [75].

Chapter 2

Model Equations and Basic theory

To carry out an extensive discussion of possibilities for model equations in lakes, we will begin with the most general three-dimensional equations, the incompressible Navier–Stokes equations, and discuss common simplifications and reductions. The concept of internal wave dispersion is introduced, followed by the introduction of layered shallow water models as the depth-averaged form of the inviscid equations of motion. We conclude by looking into retaining important wave-dispersion effects that are typically ignored by layered models.

2.1 Equations for continuously stratified flow under the Boussinesq approximation

A derivation of the full 3D incompressible Navier–Stokes equations is available in any elementary fluid mechanics textbook (e.g., [70]). We state the equations here as:

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla p + \rho \mathbf{g} + \mu \nabla^2 \mathbf{u} \,, \tag{2.1}$$

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{2.2}$$

$$\frac{D\rho}{Dt} = \kappa \nabla^2 \rho \,, \tag{2.3}$$

where $\mathbf{u} = (u(x,y,z,t), v(x,y,z,t), w(x,y,z,t))$, ρ is the fluid density, $\mathbf{g} = -g\mathbf{k}$ is the gravitational acceleration, where $g = 9.81$ ms^{-2}, and μ is the molecular dynamic viscosity, κ is the molecular diffusivity, and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \,, \tag{2.4}$$

is the material (or Eulerian) derivative that accounts for the rate of change following the flow. Conservation of momentum is given by equation (2.1), with the terms on the right-hand side corresponding to pressure gradient, buoyancy, and viscous terms (from left to right). Equation (2.2), the incompressibility constraint, corresponds to conservation of mass and comes from a variety of assumptions, including: fluid velocities and wave propagation speeds are small in comparison to the speed of sound (low Mach number); small temperature variations; and limited vertical excursions . Equation (2.3) corresponds to scalar transport, and it comes from the heat equation. It simply states that the tracer (density) is spread out by molecular diffusion (which is typically small) in a frame of reference moving with the flow.

The Boussinesq approximation is a key simplification that comes from using the fact that density changes are typically small in oceanic and limnic flows (typically $O(1\%)$ or less). The density can thus be decomposed as

$$\rho(x, y, z, t) = \rho_0 + \rho'(x, y, z, t) \, , \tag{2.5}$$

where ρ_0 is a constant background density and ρ' is a small perturbation density, i.e., $\rho'/\rho_0 \ll 1$. The pressure may be correspondingly decomposed as:

$$p(x, y, z, t) = \bar{p}(z) + p'(x, y, z, t) \, , \tag{2.6}$$

where $p'(x, y, z, t)$ is the perturbation pressure, and $\bar{p}(z)$ is a steady background pressure that is assumed to be in hydrostatic balance with the background density ρ_0, i.e.,

$$\nabla \bar{p}(z) - \rho_0 \mathbf{g} = 0 \, , \tag{2.7}$$

Substituting these decompositions into the momentum equation (2.1) and dividing the equation by ρ_0 gives

$$\left(1 + \frac{\rho'}{\rho_0}\right) \frac{D\mathbf{u}}{Dt} = -\frac{\nabla \bar{p}}{\rho_0} - \frac{\nabla p'}{\rho_0} + \mathbf{g} + \frac{\rho'}{\rho_0}\mathbf{g} + \nu \nabla^2 \mathbf{u} \, , \tag{2.8}$$

where $\nu = \mu/\rho_0$ is the kinematic viscosity. Using the hydrostatic relation (2.7) together with the approximation that $\left(1 + \frac{\rho'}{\rho_0}\right) \approx 1$ yields

$$\frac{D\mathbf{u}}{Dt} = -\frac{\nabla p'}{\rho_0} + \frac{\rho' \mathbf{g}}{\rho_0} + \nu \nabla^2 \mathbf{u} \, , \tag{2.9}$$

the stratified incompressible Navier–Stokes equations under the Boussinesq equation. The incompressible Euler equations are found by neglecting viscosity ($\nu = 0$) and thermal

diffusivity ($\kappa = 0$). It is sometimes desirable to work with the 'full' density and pressure, ρ and p instead of the perturbation fields. In which case, the hydrostatic relation (2.7) may be added back onto the right hand side (since it is zero), giving:

$$\frac{D\mathbf{u}}{Dt} = -\frac{\nabla p}{\rho_0} + \frac{\rho \mathbf{g}}{\rho_0} + \nu \nabla^2 \mathbf{u} \,, \tag{2.10}$$

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{2.11}$$

$$\frac{D\rho}{Dt} = \kappa \nabla^2 \rho \,, \tag{2.12}$$

where we have re-written the conservation of mass and scalar transport equations to show the full closed system of equations.

2.1.1 Dipsersion Relation of Internal Waves in a Continuously Stratified Fluid

Following Kundu and Cohen [70], the dispersion characteristics in a continuously stratified fluid can be derived by first linearizing the stratified incompressible Euler equations ('Euler' since we neglect viscosity and diffusion terms). First, we drop nonlinear terms in the momentum equations, (and we denote partial derivatives with subscripts)

$$\mathbf{u}_t = -\frac{\nabla p}{\rho_0} + \frac{\rho \mathbf{g}}{\rho_0} \,, \tag{2.13}$$

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{2.14}$$

$$\frac{D\rho}{Dt} = 0 \,, \tag{2.15}$$

and note that we have effectively linearized about $\mathbf{u} = 0$ (no background flow). It remains to suitably linearize the density equation (2.15), and this can be achieved by decomposing the pressure and density into a background state that is in hydrostatic balance plus a small perturbation:

$$p = \bar{p}(z) + p'(x, y, z, t) \,, \tag{2.16}$$

$$\rho = \bar{\rho}(z) + \rho'(x, y, z, t) \,. \tag{2.17}$$

Here, the background density $\bar{\rho}$ is allowed to vary with z and thus primed variables have a different interpretation than in equations (2.5)–(2.6). Substitution into equations (2.13)–

(2.15) and dropping terms that are quadratic in the perturbation variables yields

$$u_t = -\frac{1}{\rho_0} p'_x \,, \tag{2.18}$$

$$v_t = -\frac{1}{\rho_0} p'_y \,, \tag{2.19}$$

$$w_t = -\frac{1}{\rho_0} p'_z - \frac{\rho' g}{\rho_0} \,, \tag{2.20}$$

$$\rho'_t - \frac{N^2 \rho_0}{g} w = 0 \,, \tag{2.21}$$

$$u_x + v_y + w_z = 0 \,, \tag{2.22}$$

where

$$N^2 = -\frac{g}{\rho_0} \frac{d\bar{\rho}}{dz} \,, \tag{2.23}$$

is the definition of the Brunt–Väisälä (or buoyancy) frequency, $N(z)$.

We can look for linear travelling wave solutions to the linearized equations by assuming a "normal modes" solution of the form

$$\begin{pmatrix} u \\ v \\ w \\ \rho' \\ p' \end{pmatrix} = \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\rho} \\ \hat{p} \end{pmatrix} \mathrm{e}^{i(kx+ly+mz-\sigma t)} \,. \tag{2.24}$$

Substitution of the ansatz (2.24) into the linearized equations (2.18)–(2.22) yields the following algebraic system of equations

$$\begin{pmatrix} -i\sigma & 0 & 0 & 0 & \frac{-ik}{\rho_0} \\ 0 & -i\sigma & 0 & 0 & \frac{-il}{\rho_0} \\ 0 & 0 & -i\sigma & \frac{g}{\rho_0} & \frac{-im}{\rho_0} \\ 0 & 0 & -\frac{\rho_0 N^2}{g} & -i\sigma & 0 \\ ik & il & im & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\rho} \\ \hat{p} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \,. \tag{2.25}$$

Non-trivial solutions to this system exist if and only if the determinant of the matrix is zero. Taking the determinant, setting it equal to zero, and solving for σ gives

$$\sigma^3 = N^2 \frac{k^2 + l^2}{k^2 + l^2 + m^2} \, \sigma, \tag{2.26}$$

which, upon dividing by σ (we are interested in solutions with $\sigma \neq 0$) and taking the positive root, can be written as

$$\sigma = N \cos\theta \ , \tag{2.27}$$

where θ is the angle between the wavenumber vector $\mathbf{k} = (k, l, m)$ and the horizontal. The key result to notice is that the frequency of internal waves depends on the direction of the wavenumber vector and not its magnitude. As we shall see later in this chapter, when continuously stratified internal waves are modelled as waves at the interface of a density jump, the frequency depends purely on the magnitude of the wavenumber. A few other interesting points are: (1) the upper bound on frequency is given by N, and (2) the group velocity (velocity of energy transport) and phase velocity (velocity of wave crests) are perpendicular. Here, the phase velocity represents the velocity of individual wave crests and is given by [70]

$$\mathbf{c} \equiv \frac{\sigma}{K}\frac{\mathbf{K}}{K} = \frac{\sigma}{K^2}\left(\mathbf{i}_x k + \mathbf{i}_z m\right) \ , \tag{2.28}$$

and the group velocity gives the velocity of energy transport

$$\mathbf{c_g} \equiv \mathbf{i}_x \frac{\partial \sigma}{\partial k} + \mathbf{i}_y \frac{\partial \sigma}{\partial l} + \mathbf{i}_z \frac{\partial \sigma}{\partial m} = \frac{Nm}{K^3}\left(\mathbf{i}_x m - \mathbf{i}_z k\right) \ . \tag{2.29}$$

where we have assumed the motion is in the xz-plane so ($l = 0$), and $\mathbf{K} = (k, m)$ with $K = \|\mathbf{K}\|$. See [70] for a thorough discussion of the implications of the dispersion relation (2.26).

2.2 Hydrostatic primitive equations

Many 3D numerical lake and ocean models [61, 44] avoid the difficulty of having to repeatedly solve a 3D Poisson problem for pressure p, by making the hydrostatic approximation and neglecting accelerations in the vertical momentum equation, i.e.,

$$0 = -\frac{1}{\rho_0}\frac{\partial p}{\partial z} - \frac{\rho g}{\rho_0} \ , \tag{2.30}$$

and we have again neglected the molecular viscosity by taking $\nu = 0$. This equation can be integrated from a rigid-lid boundary at $z' = 0$ to an arbitrary depth $z' = z$ to give

$$\int_{z'=0}^{z'=z} \frac{\partial p}{\partial z'}\, dz' = \int_{z'=0}^{z'=z} -\rho(x,y,z,t) g\, dz' \,, \tag{2.31}$$

$$\Rightarrow p|_{z'=z} = p_{lid}(x,y,t) + \int_{z'=0}^{z'=z} -\rho(x,y,z,t) g\, dz' \,. \tag{2.32}$$

where $p_{lid}(x,y,t) = p(x,y,z=0,t)$. The z-integral of ρ can be computed numerically if ρ is known, and the only unknown remaining is the two-dimensional rigid-lid pressure p_{lid}. The computationally expensive 3D Poisson problem is then replaced by a 2D Poisson problem for p_{lid}. In a similar manner, the vertical velocity w can be computed numerically by integrating the continuity equation

$$\int_{z'=0}^{z'=z} w_z\, dz' = -\int_{z'=0}^{z'=z} u_x + v_y\, dz' \,. \tag{2.33}$$

$$w(x,y,z,t) = \int_{z'=0}^{z'=z} u_x + v_y\, dz' \,. \tag{2.34}$$

since $w = 0$ at the surface by virtue of the rigid-lid approximation.

An alternative to using the rigid-lid approximation is to retain a free-surface at $z = \eta(x,y,t)$. In which case, there is no Poisson problem to solve, but rather an evolution equation for η that results from applying the kinematic boundary condition

$$\frac{D(z-\eta)}{Dt} = 0 \quad \text{at} \quad z = \eta \,,$$

$$\Rightarrow w = \frac{D\eta}{Dt} \quad \text{at} \quad z = \eta \,, \tag{2.35}$$

must be evolved along with the other equations. The kinematic boundary condition simply states that the position of the free-surface is a constant in a frame of reference that moves with the flow. Said another way: "a fluid particle on the free surface stays on the free surface" [70] since fluid particle is a classical definition going back to Euler. For reasons that will become clear later in this chapter, this form of the hydrostatic primitive equations is often referred to as the "three-dimensional shallow water equations."

2.2.1 Internal wave dispersion relation under the hydrostatic approximation

The dispersion relation for internal waves in the hydrostatic primative equations can be determined by following the same procedure explained in section 2.1.1 with the w_t term dropped from the vertical momentum equation. In which case, we find

$$\sigma = \pm N \frac{\sqrt{k^2 + l^2}}{m} \ . \tag{2.36}$$

Thus, if we fix a vertical wavenumber (or 'mode number', in a bounded geometry) $m = m_0$, the horizontal phase and group velocities are:

$$\mathbf{c} = \mathbf{c}_g = \frac{N}{m_0}(\mathbf{i} + \mathbf{j}) \ , \tag{2.37}$$

and we find that there is no dependence on horizontal wavenumber, and these waves can be thought of as horizontally non-dispersive. Upon comparing the relation (2.36) to (2.26), it is clear that the hydrostatic approximation is valid when $\sqrt{k^2 + l^2} \ll m$ (i.e., $H \ll L$), or when $\sigma \ll N$.

2.3 Derivation of the 1-layer shallow water equation

The shallow water equations form the prototypical 2D layered model, and it is instructive to see their derivation as a first step before discussing dispersive and multi-layer models. The underlying assumption is that the horizontal velocity components u and v have been averaged over depth so that they do not depend on z, the fluid-layer is homogeneous, and the aspect ratio is small $(H/L) \ll 1$ so that the hydrostatic assumption (see Section 2.2) can be invoked. We also assume that the amplitude of the water waves is small in comparison to the water depth $(a/H) \ll 1$ so that breaking and overturning effects can be neglected and hence, the free surface displacement can be represented only as a function of the horizontal coordinates and time. The basic setup of the flow is illustrated in Figure 2.1.

We define the undisturbed depth as $H(x, y)$ and $\eta(x, y, t)$ is the deviation of the free surface from a flat state at $z = 0$, so that the total depth is given by $h(x, y, t) = H(x, y) + \eta(x, y, t)$. The undisturbed depth profile $H(x, y)$ is sometimes written as a constant depth less a perturbation bathymetry profile $H_0 - b(x, y)$. We mention this here since it is the convention adopted for two-layer models in section 2.8.

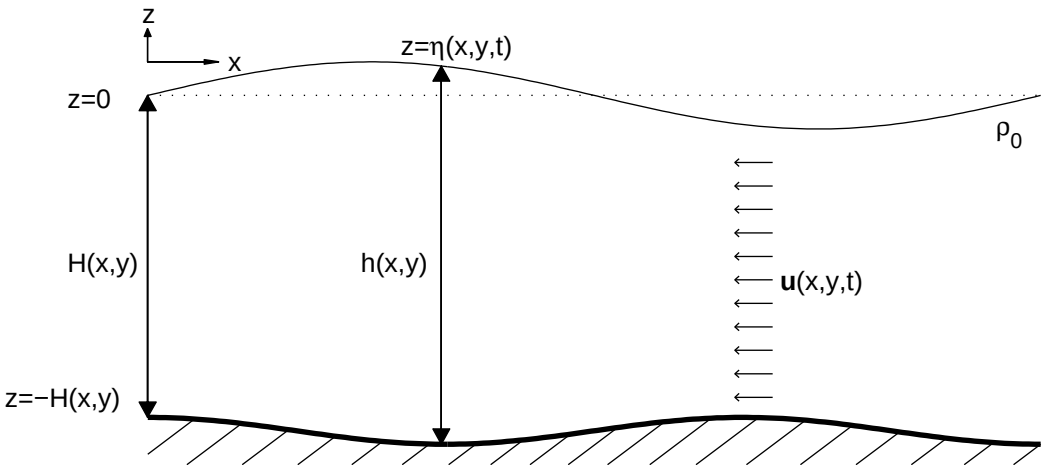

Figure 2.1: Schematic diagram for the 1-layer shallow water equations.

We begin by deriving the equations for conservation of mass. Given that in three-dimensions the flow is incompressible, the continuty equation is

$$u_x + u_y + w_z = 0 \ . \tag{2.38}$$

To find a depth-averaged statement of conservation of mass within each layer, we integrate equation (2.38) from the bottom at $z = -H(x, y)$ to the top at $z = \eta(x, y, t)$

$$\int_{z=-H(x,y)}^{z=\eta(x,y,t)} w_z \, dz = \int_{z=-H(x,y)}^{z=\eta(x,y,t)} -(u_x + v_y) \, dz \ . \tag{2.39}$$

Since the integrand on the right does not depend on z, we have

$$w(z = \eta(x, y, t)) - w(z = -H(x, y)) = h \nabla \cdot \mathbf{u} \ , \tag{2.40}$$

where we have used $h = H + \eta$ and we have introduced $\nabla \cdot \mathbf{u} = u_x + v_y$ as the horizontal divergence for notational convience.

As explained for the hydrostatic primitive equations, the kinematic boundary condition (2.35) can be invoked at the free surface to give

$$w = \frac{D\eta}{Dt} \ , \quad \text{at} \quad z = \eta \ . \tag{2.41}$$

A similar line of reasoning applies for the bottom boundary condition. However, the expression is simplified since H does not dependent on time

$$w = \mathbf{u} \cdot \nabla(-H) \; , \quad \text{at} \quad z = -H \; . \tag{2.42}$$

Substituting (2.41) and (2.42) into (2.40) gives

$$\frac{\partial \eta}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0 \; , \tag{2.43}$$

after simplifying with the product rule. Noting again that H is time-independent, this equation can also be written as

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0 \; , \tag{2.44}$$

so that η does not appear in the equation.

At this point, an expression for the hydrostatic pressure is required for use in the depth-averaged horizontal momentum equations. Integrating the hydrostatic pressure relation from the top $z' = \eta$ to an arbitrary depth z we have

$$\int_{z'=\eta}^{z'=z} \frac{\partial p}{\partial z'} \, dz' = \int_{z'=\eta}^{z'=z} -\rho_0 g \, dz' \; , \tag{2.45}$$

$$\Rightarrow p(z' = z) - p(z' = \eta) = -\rho_0 g z + \rho_0 g \eta \; . \tag{2.46}$$

Applying continuity of pressure at the free surface gives $p(z = \eta) = p_{atm} = constant$. Therefore,

$$p(x, y, z, t) = p_{atm} + \rho_0 g \eta(x, y, t) - \rho_0 g z \; , \tag{2.47}$$

and the horizontal pressure gradient is

$$\nabla p(x, y, t) = \rho_0 g \nabla \eta(x, y, t) \; , \quad \text{for} \quad -H \leq z \leq \eta \; . \tag{2.48}$$

Substituting the expression for pressure into the horizontal components of (2.10) with viscosity neglected gives the momentum equations. The shallow water equations are thus,

$$\frac{D\mathbf{u}}{Dt} = -g \nabla \eta \; , \tag{2.49}$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0 \; , \tag{2.50}$$

where $\mathbf{u} = (u(x,y,t), v(x,y,t))$. The momentum equation (2.49) can be put in the form of a conservation law by multiplying it by h and adding it to (2.50) multiplied by $\mathbf{u}$. Simplifying gives

$$\frac{\partial(h\mathbf{u})}{\partial t} + \nabla \cdot (h\mathbf{u} \otimes \mathbf{u}) = -gh\nabla\eta \; . \tag{2.51}$$

Here, $\otimes$ represents a Kronecker product. The free surface η can be eliminated by noting (from the diagram in Figure 2.1) that

$$h = H + \eta \; , \tag{2.52}$$

$$\Rightarrow \nabla\eta = \nabla h - \nabla H \tag{2.53}$$

giving

$$\frac{\partial(h\mathbf{u})}{\partial t} + \nabla \cdot (h\mathbf{u} \otimes \mathbf{u}) + \nabla\left(\frac{1}{2}gh^2\right) = gh\nabla H \; . \tag{2.54}$$

2.3.1 Linear Wave Characteristics

If we restrict the flow to one direction (x) and assume a flat bottom topography, the shallow water equations (2.49)–(2.50) become

$$u_t + uu_x = -g\eta_x \; , \tag{2.55}$$

$$\eta_t = -(H+\eta)u_x \; . \tag{2.56}$$

If we linearize these equations about a constant background flow U, i.e., $u(x,t) = U + u'(x,t)$ $\eta(x,t) = \eta'(x,t)$, we find

$$u'_t + Uu'_x = -g\eta'_x \; , \tag{2.57}$$

$$\eta'_t = -Hu'_x \; . \tag{2.58}$$

Next, if we assume each of η' and u' are proportional to $\mathrm{e}^{i(kx-\sigma t)}$ as above in Section 2.2.1, we find

$$\sigma = Uk \pm \sqrt{gH}k \; . \tag{2.59}$$

Hence, the phase and group speeds are

$$c = c_g = U \pm \sqrt{gH} \; . \tag{2.60}$$

Here, $\sqrt{gH}$ is called the long wave speed, and we see that under this linearization, the background current U simply Doppler shifts the waves. The relation $c = c_g$ holds only when the wave solutions for a given system of equations are *non-dispersive*. That is, since c and c_g do not depend on the wavenumber (hence, wavelength), all waves, regardless of their wavelength must travel at the same speed. This is also a common feature of all linear hyperbolic PDEs with constant coefficients.

2.3.2 Conservative and Quasi-linear Forms of the shallow water systems

Due to the wealth of literature on the numerical solution of PDE systems that are hyperbolic conservation laws [75, 121, 35], it is useful to note that the shallow water equations can be written in the general conservative form

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial(\mathbf{F}(\mathbf{q}))}{\partial x} + \frac{\partial(\mathbf{G}(\mathbf{q}))}{\partial y} = \mathbf{S}(q) \ , \tag{2.61}$$

where

$$\mathbf{q} = \begin{pmatrix} h \\ hu \\ hv \end{pmatrix} \ , \quad \mathbf{F}(\mathbf{q}) = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{pmatrix} \ , \tag{2.62}$$

$$\mathbf{G}(\mathbf{q}) = \begin{pmatrix} hv \\ huv \\ \frac{1}{2}gh^2 + hv^2 \end{pmatrix} \ , \quad \mathbf{S}(\mathbf{q}) = \begin{pmatrix} 0 \\ ghH_x \\ ghH_y \end{pmatrix} \ . \tag{2.63}$$

where $(\mathbf{F}, \mathbf{G})$ is the flux. Applying the chain rule leads to the quasi-linear form

$$\frac{\partial \mathbf{q}}{\partial t} + A(\mathbf{q})\mathbf{q}_x + B(\mathbf{q})\mathbf{q}_y = \mathbf{S}(\mathbf{q}) \ , \tag{2.64}$$

where

$$A(\mathbf{q}) = \begin{pmatrix} 0 & 1 & 0 \\ gh - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \ , \text{and } B(\mathbf{q}) = \begin{pmatrix} 0 & 0 & 1 \\ -uv & v & u \\ gh - v^2 & 0 & 2v \end{pmatrix} \ . \tag{2.65}$$

In principle, insight into the solutions of the full quasi-linear system can be obtained by considering the equations along a single direction and appropriately diagonalizing the system so that the the method of characteristics can be applied to a set of decoupled nonlinear wave equations. In the following sections, we apply this procedure in the simplified one dimensional linearized case where closed-form analytical solutions are tractable.

2.3.3 1D Hyperbolic Theory

Many numerical methods (including finite volume [75] and discontinuous Galerkin [20] methods) rely on the well-established hyperbolic theory of conservation laws to exchange information between computational cells since hyperbolic theory gives insights into how

information propagates as time evolves. The ideas are based upon the method of characteristics that, in 1D, looks for curves in the x-t plane along which the solution is constant, called *characteristic curves* or simply *characteristics*. Once the characteristic curves are known, the full solution can be determined by following the characteristics forward in time beginning from the initial data.

Although many PDE systems of practical interest are not fully hyperbolic (e.g., dispersive wave equations or equations for incompressible flow), their numerical treatments typically "split" terms up [42, 1] such that at least one step in the solution procedure resembles solving a hyperbolic system of equations, and the importance of understanding some hyperbolic theory for the purposes of numerical methods is thus clear. In this section, we give a brief introduction to the utility of hyperbolic theory for the 1D linear shallow water equations.

The 1D linear shallow water equations can be written as

$$\mathbf{q}_t + (\mathbf{F}(\mathbf{q}))_x = \mathbf{S}(\mathbf{q}) \ , \tag{2.66}$$

where $\mathbf{q} = (h, u)^\mathsf{T}$, and the flux vector is given by

$$\mathbf{F}(\mathbf{q}) = \begin{pmatrix} Hu \\ gh \end{pmatrix} \ . \tag{2.67}$$

$\mathbf{S}(\mathbf{q})$ corresponds to source terms, which in this context would be due to variable bathymetry. Since source terms do not affect the characteristics of the system, for simplicity we can assume a flat bottom, hence $\mathbf{S}(\mathbf{q}) = 0$. Applying the chain rule to (2.66) gives

$$\mathbf{q}_t + A\mathbf{q}_x = 0 \ . \tag{2.68}$$

where

$$A = \frac{\partial \mathbf{F}}{\partial \mathbf{q}} \ , \tag{2.69}$$

is the flux Jacobian. Here,

$$A = \begin{pmatrix} 0 & H \\ g & 0 \end{pmatrix} \ . \tag{2.70}$$

A first-order system is said to be hyperbolic if A is diagonalizable with real eigenvalues. The characteristic wave speeds are then given by the eigenvalues $\lambda_{1,2}$ of A. This fact can be seen by realizing that diagonalizing the PDE system would result in a set of decoupled linear advection equations each with its own linear wave speed. This result can be

further motivated by presuming that a wave speed λ is known and making the coordinate substitution $\xi = x - \lambda t$, in which case the system would read

$$\begin{pmatrix} -\lambda & H \\ g & -\lambda \end{pmatrix} \begin{pmatrix} h \\ u \end{pmatrix}_{\xi} = 0 \,, \tag{2.71}$$

and it becomes clear that interesting solutions require the determinant of the matrix in equation (2.71) to be zero. The characteristic equation $\lambda^2 - gH = 0$ is solved quite simply to give the characteristic wave speeds

$$\lambda_1 = -\sqrt{gH} \,, \quad \lambda_2 = \sqrt{gH} \,, \tag{2.72}$$

corresponding to leftward and rightward propagating waves, respectively. The eigenvectors $\mathbf{v}_{1,2}$ can also be determined from their definition $A\mathbf{v} = \lambda\mathbf{v}$, and since they are determined only up to a multiplicative constant we can choose their first entry be 1 without loss of generality. Hence,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -\sqrt{g/H} \end{pmatrix} \,, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ \sqrt{g/H} \end{pmatrix} \tag{2.73}$$

The eigenvectors dictate what the weighting of flow variables should be such that the solution profile is a single translating wave for all times. For example, if $h(x,0)$ is the initial depth profile, then setting the initial velocity to $u(x,0) = \left(\sqrt{g/H}\right) h(x,0)$[1] would result in a solution profile that translates to the right with the linear long-wave speed $\lambda_2 = \sqrt{gH}$. Physically, the eigenvectors dictate what the current induced by a single wave propagating either to the left or right would be.

Linear Riemann Problem

The Riemann problem is a well-studied applied mathematics problem that asks the question of how an initial jump between two constant states would evolve forward in time in a hyperbolic PDE system, and it is fundamental in the formulation of numerical methods where information must be exchanged between computational cells when strong continuity is not enforced at cell interfaces.

Without loss of generality, we may assume that the jump is centered at $x = 0$ and that the initial data is given by

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_l \,, & x \leq 0 \,, \\ \mathbf{q}_r \,, & x > 0 \,. \end{cases} \tag{2.74}$$

[1] One could also take $u(x,0) = \sqrt{g/H}\eta(x,0)$ so that there is no mean background flow.

As shown in Figure 2.2, the method of characteristics implies that for the linear shallow water system the solution should split into three regions for $t > 0$, corresponding to the initial left $\mathbf{q}_l$ and right states $\mathbf{q}_r$, and an intermediate state $\mathbf{q}^*$. We are charged with determining the value of $\mathbf{q}^*$ so that the full solution will be known for all times following the initialization.

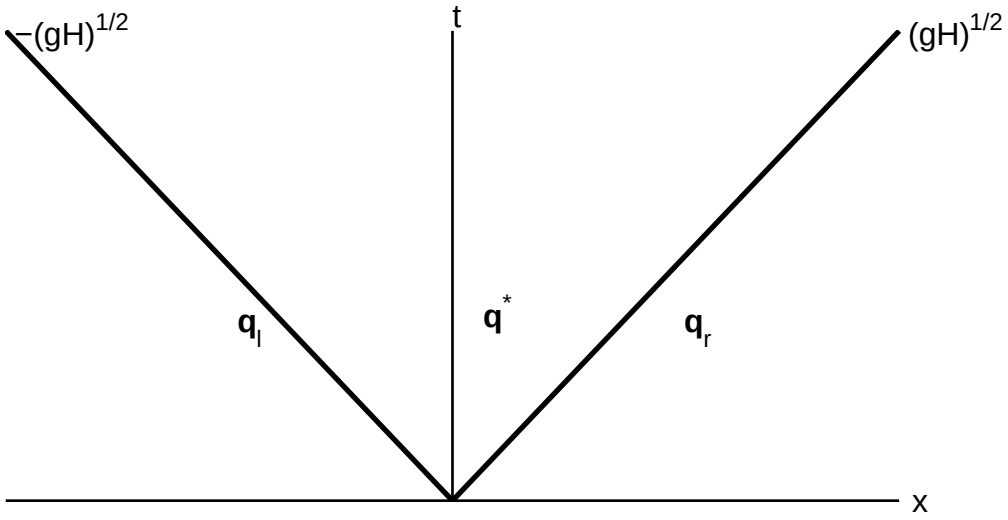

Figure 2.2: Illustration of the Riemann problem in the xt-plane. Immediately after the initialization, a third state ($\mathbf{q}^*$) appears in the solution that must be determined.

As explained in [75], the solution of the Riemann problem can be obtained by using the fact that the initial jump $\delta\mathbf{q} = \mathbf{q}_r - \mathbf{q}_l$ can be decomposed in terms of the eigenstructure found above in Section 2.3.3. That is, we wish to find constants α_1 and α_2 such that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \delta\mathbf{q} \; . \tag{2.75}$$

This is a matrix problem of the form

$$R\boldsymbol{\alpha} = \delta\mathbf{q} \; , \tag{2.76}$$

where $R = [\mathbf{v}_1 \; \mathbf{v}_2]$. Its solution is

$$\alpha_1 = -\frac{1}{2}\sqrt{\frac{g}{H}}\,(h_r - h_l) + \frac{1}{2}\,(u_r - u_l) \; , \tag{2.77}$$

$$\alpha_2 = \frac{1}{2}\sqrt{\frac{g}{H}}\,(h_r - h_l) + \frac{1}{2}\,(u_r - u_l) \; . \tag{2.78}$$

The terms on the left-hand side of eqn. (2.75) can be interpreted as individual waves $W_1 = \alpha_1 \mathbf{v}_1$ and $W_2 = \alpha_2 \mathbf{v}_2$, respectively. The value of $\mathbf{q}^*$ can be determined by noticing that $\mathbf{q}^* = \mathbf{q}_l + W_1$ (or $\mathbf{q}^* = \mathbf{q}_r + W_2$). We thus find

$$h^* = \frac{h_l + h_r}{2} + \frac{1}{2}\sqrt{\frac{H}{g}}\,(u_r - u_l) \ , \tag{2.79}$$

$$u^* = \frac{u_l + u_r}{2} + \frac{1}{2}\sqrt{\frac{g}{H}}\,(h_r - h_l) \ , \tag{2.80}$$

and the full solution is

$$\mathbf{q}(x,t) = \begin{cases} \mathbf{q}_l \ , & x < -\sqrt{gH}t \ , \\ \mathbf{q}^* \ , & -\sqrt{gH}t \le x \le \sqrt{gH}t \ , \\ \mathbf{q}_r \ , & x > \sqrt{gH}t \ . \end{cases} \tag{2.81}$$

Many numerical methods are concerned with the form of the flux vector in the star region $\mathbf{F}(\mathbf{q}^*)$ that is also known as the *numerical flux function*. Here, we find

$$\mathbf{F}(\mathbf{q}^*) = \begin{pmatrix} \frac{1}{2}(Hu_l + Hu_r) + \frac{1}{2}\sqrt{gH}\,(h_r - h_l) \\ \frac{1}{2}(gh_l + gh_r) + \frac{1}{2}\sqrt{gH}\,(u_r - u_l) \end{pmatrix} \ , \tag{2.82}$$

which may be re-written as

$$\mathbf{F}(\mathbf{q}^*) = \frac{\mathbf{F}(\mathbf{q}_l) + \mathbf{F}(\mathbf{q}_r)}{2} + \frac{\lambda_2}{2}\,(\mathbf{q}_r - \mathbf{q}_l) \ . \tag{2.83}$$

It turns out that (2.83) can be used in much more general situations and even forms the basis for the Lax-Friedrichs/Rusanov class of approximate Riemann solvers for nonlinear problems for which closed-form solutions become intractable [60, 121].

2.4 Effects of rotation

Since lake-scale physical processes in medium and large sized lakes at mid- and high-latitudes are affected by the Earth's rotation [115, 36], it is necessary to account for the Coriolis pseudo-force in our models. While Stewart and Dellar [113, 114] have derived shallow water type models that account for the complete Coriolis force, we make use of the very popular traditional approximation that neglects the locally horizontal part of the rotation vector. We further simplify matters by making the f-plane approximation that

neglects meridional variations in the full Coriolis frequency $f = 2\Omega \sin\theta$. The f-plane is known to be an accurate approximation to the full rotating equations of motion on the surface of a sphere provided that horizontal length scales and time scales are not too large [70]. From Kundu and Cohen [70], if we assume that departures from a central latitude θ_0 are small the shallow water equations on a sphere rotating with angular frequency Ω are well-approximated by

$$\frac{D\mathbf{u}}{Dt} + f\,(\mathbf{k}\times\mathbf{u}) = -g\nabla\eta\ , \tag{2.84}$$

$$\frac{\partial h}{\partial t} + \nabla\cdot(h\mathbf{u}) = 0\ , \tag{2.85}$$

where $f = f_0 = 2\Omega\sin\theta_0$ is the Coriolis parameter assumed to be a constant in the f-plane approximation.

2.4.1 Rotating Gravity Waves in a channel

It is important to understand the impact of the Coriolis (f-plane) effect on shallow water waves to build intuition on the behaviour of waves in lakes. An example of a classical, analytically tractable problem is that of free linear wave oscillations in a periodic or infinite channel.

In this section, we look for analytical solutions to the linearized rotating shallow water equations on a channel given by $[0, L_x]\times[-\infty,\infty]$. The governing equations are

$$u_t - fv = -g\eta_x\ , \tag{2.86}$$

$$v_t + fu = -g\eta_y\ , \tag{2.87}$$

$$\eta_t = -Hu_x - Hv_y\ . \tag{2.88}$$

We look for normal modes solutions of the form

$$\begin{pmatrix} u\\ v\\ \eta\end{pmatrix} = \mathrm{e}^{i(ly-\sigma t)}\begin{pmatrix}\hat{u}(x)\\ \hat{v}(x)\\ \hat{\eta}(x)\end{pmatrix}\ , \tag{2.89}$$

where $\sigma\in\mathbb{R}$, and $l = 2\pi m/L_y$, $m\in\mathbb{Z}$ since the channel is periodic and bounded in y. Under the ansatz (2.89), the system (2.86)-(2.88) reduces to the eigenvalue problem

$$\begin{pmatrix} 0 & f & -g\frac{d}{dx}\\ -f & 0 & -gil\\ -H\frac{d}{dx} & -Hil & 0\end{pmatrix}\begin{pmatrix}\hat{u}\\ \hat{v}\\ \hat{\eta}\end{pmatrix} = -i\sigma\begin{pmatrix}\hat{u}\\ \hat{v}\\ \hat{\eta}\end{pmatrix}\ . \tag{2.90}$$

The boundary conditions are given by no normal flow through the channel walls, i.e.,

$$\hat{u} = 0 \text{ at } x = 0, L_x \ . \tag{2.91}$$

Suitable conditions on $\hat{v}$ and $\hat{\eta}$ can be determined by substituting $\hat{u}(x = 0, L_x) = 0$ into the eigenvalue problem. After some light algebra, we recover the Robin type boundary conditions

$$fl\hat{\eta} - \sigma\hat{\eta}_x = 0 \text{ at } x = 0, L_x \ , \tag{2.92}$$

$$fl\hat{v} - \sigma\hat{v}_x = 0 \text{ at } x = 0, L_x \ . \tag{2.93}$$

Eliminating $\hat{u}$ and $\hat{v}$ from the eigenproblem (2.90) gives the single equation for $\hat{\eta}$

$$\hat{\eta}_{xx} + \left(\frac{\sigma^2 - f^2}{c_0^2} - l^2 \right) \hat{\eta} = 0 \ , \tag{2.94}$$

where $c_0 = \sqrt{gH}$. We have hence recovered the classical Sturm-Liouville differential eigenproblem, which may be solved by considering separate cases based on the sign of the quantity in parentheses in (2.94).

Case I: $\frac{\sigma^2 - f^2}{c_0^2} - l^2 \geq 0$

In this case, the eigenproblem can be written as

$$\hat{\eta}_{xx} + \lambda^2 \hat{\eta} = 0 \ , \tag{2.95}$$

where

$$\lambda = \sqrt{\frac{\sigma^2 - f^2}{c_0^2} - l^2} \ . \tag{2.96}$$

The general solution is

$$\hat{\eta} = A \sin(\lambda x) + B \cos(\lambda x) \ . \tag{2.97}$$

Imposing the boundary condition (2.92) at $x = 0$ yields

$$A = \frac{fl}{\sigma\lambda} B \ , \tag{2.98}$$

hence,

$$\hat{\eta} = B \left[\frac{fl}{\sigma\lambda} \sin(\lambda x) + \cos(\lambda x) \right] \ . \tag{2.99}$$

Imposing (2.92) at $x = L_x$ gives the equation

$$\left(f^2 l^2 + \sigma^2 \lambda^2 \right) \sin(\lambda L_x) = 0 \ , \tag{2.100}$$

for which there are a multitude of possible roots. First, assume the quantity multiplying $\sin(\lambda L_x)$ is zero, i.e.,

$$f^2 l^2 + \sigma^2 \left(\frac{\sigma^2 - f^2}{c_0^2} - l^2 \right) = 0 \,, \tag{2.101}$$

which may be rewritten as a quadratic equation in powers of σ^2,

$$\sigma^4 - (f^2 + c_0^2 l^2)\sigma^2 + f^2 l^2 c_0^2 = 0 \,. \tag{2.102}$$

Thus, the roots are given by

$$\sigma^2 = \frac{f^2 + c_0^2 l^2 \pm \sqrt{\left(f^2 - c_0^2\, l^2\right)^2}}{2} \,. \tag{2.103}$$

Therefore,

$$\sigma^2 = f^2 \,, \quad \sigma^2 = c_0^2\, l^2 \,. \tag{2.104}$$

However, these roots are unphysical for eigenfunctions of the form (2.97) since they break the assumption that λ is real and positive, so they must be discarded.

Returning to equation (2.100), the other possible roots are given by the zeros of the sine functions, that occur for

$$\lambda L_x = n\pi \,, \; n = 0, 1, \cdots \,. \tag{2.105}$$

Squaring this relation and substituting the expression for λ gives

$$\sigma^2 = f^2 + c_0^2 (k^2 + l^2) \,, \tag{2.106}$$

where

$$k = \frac{n\pi}{L_x} \,, \; n = 0, 1, \cdots \,, \tag{2.107}$$

represent the across-channel wavenumber. The dispersion relation (2.106) is the classical result for Poincaré (rotating gravity) waves. It is worth mentioning that the case where $\lambda = 0$ is recovered when $k = 0$, a wave with constant cross-channel structure.

Case II: $\frac{\sigma^2 - f^2}{c_0^2} - l^2 < 0$

In this case, the eigenproblem can be written as

$$\hat{\eta}_{xx} - \lambda^2 \hat{\eta} = 0 \,, \tag{2.108}$$

where

$$\lambda = \sqrt{l^2 - \frac{\sigma^2 - f^2}{c_0^2}} \,. \tag{2.109}$$

The general solution is

$$\hat{\eta} = C\sinh(\lambda x) + D\cosh(\lambda x)\ . \tag{2.110}$$

Imposing the boundary condition (2.92) at $x = 0$ yields

$$C = \frac{fl}{\sigma\lambda} D\ , \tag{2.111}$$

as before. Hence,

$$\hat{\eta} = D\left[\frac{fl}{\sigma\lambda}\sinh(\lambda x) + \cosh(\lambda x)\right]\ . \tag{2.112}$$

Imposing (2.92) at $x = L_x$ gives

$$\left(f^2 l^2 - \lambda^2\sigma^2\right)\sinh(\lambda L_x) = 0\ . \tag{2.113}$$

Since $\sinh(\lambda L_x)$ cannot be zero ($L_x, \lambda > 0$), it follows that the eigenvalues are given by the equation

$$f^2 l^2 + \sigma^2\left(\frac{\sigma^2 - f^2}{c_0^2} - l^2\right) = 0\ , \tag{2.114}$$

which was solved in Case I,

$$\sigma^2 = f^2\ , \quad \sigma^2 = c_0^2\, l^2\ . \tag{2.115}$$

In this case, these values of sigma do not yield a contradiction, and the following eigenfunctions are recovered in each case

$$\sigma = f \Rightarrow \lambda = \frac{f}{c_0} \Rightarrow \hat{\eta} = E\mathrm{e}^{-\lambda(L_x - x)}\ , \tag{2.116}$$

$$\sigma = -f \Rightarrow \lambda = \frac{f}{c_0} \Rightarrow \hat{\eta} = D\mathrm{e}^{-\lambda x}\ , \tag{2.117}$$

$$\sigma = c_0\, l \Rightarrow \lambda = \frac{f}{c_0} \Rightarrow \hat{\eta} = E\mathrm{e}^{-\lambda(L_x - x)}\ , \tag{2.118}$$

$$\sigma = -c_0\, l \Rightarrow \lambda = \frac{f}{c_0} \Rightarrow \hat{\eta} = D\mathrm{e}^{-\lambda x}\ . \tag{2.119}$$

Here, we introduced $E = D\mathrm{e}^{\lambda L_x}$ for ease of interpretation. The quantity c_0/f is the *external Rossby deformation radius*, and represents the decay length scale of these coastally trapped eigenmodes. Here, the $\sigma = \pm f$ solutions correspond to the so-called *inertial oscillation* and the $\sigma = \pm c_0 l$ solutions correspond to the *Kelvin wave*. Kelvin waves possess the property that their frequency can be less than the inertial frequency, and for closed basins this feature is typically used to identify which modes are Kelvin modes [115].

2.5 Dispersive shallow water model

Wave dispersion can be included in layered model equations so that weakly non-hydrostatic effects can be captured by the model. The underlying idea is that weak vertical accelerations are parameterized as horizontal dispersion. The governing equations used by de la Fuente *et. al.* [28] in their study of internal waves in a circular basin for a single fluid layer are

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0\,, \tag{2.120}$$

$$\frac{\partial (uh)}{\partial t} + \nabla \cdot ((uh)\mathbf{u}) = -gh\frac{\partial \eta}{\partial x} + fvh + \frac{H^2}{6}\frac{\partial}{\partial x}\left(\nabla \cdot \frac{\partial (\mathbf{u}h)}{\partial t}\right)\,, \tag{2.121}$$

$$\frac{\partial (vh)}{\partial t} + \nabla \cdot ((vh)\mathbf{u}) = -gh\frac{\partial \eta}{\partial y} - fuh + \frac{H^2}{6}\frac{\partial}{\partial y}\left(\nabla \cdot \frac{\partial (\mathbf{u}h)}{\partial t}\right)\,, \tag{2.122}$$

where $\mathbf{u} = (u(x,y,t), v(x,y,t))$ is the velocity field, $h(x,y,t) = H(x,y) + \eta(x,y,t)$ is the total depth with H representing the undisturbed depth, and η is the free surface displacement. The constants g and f are the acceleration due to gravity and the Coriolis frequency, respectively. The difference between the set of equations (2.120)–(2.122) and the traditional shallow water model is the inclusion of the dispersive terms $\frac{H^2}{6}\nabla(\nabla \cdot (\mathbf{u}h)_t)$ found in the momentum equations (2.121) & (2.122). The above system was first proposed by Brandt *et. al.* [15] in their study of internal waves in the Strait of Messina. This system is derived by a perturbation expansion in powers of the small dimensionless parameter $\mu = (H/L)$, and therefore is only physically relevant if $\mu \ll 1$ [28].

We have neglected bottom and surface stresses in equations (2.120)–(2.122) since their inclusion into numerical schemes is conceptually simple and contributes little to the discussion.

The linearized form of equations (2.120)–(2.122) in one space dimension with $f = 0$ take the form

$$\eta_t + (Hu)_x = 0\,, \tag{2.123}$$

$$u_t = -g\eta_x + \frac{H^2}{6}u_{xxt}\,, \tag{2.124}$$

Assuming a flat bottom ($H = constant.$) and that η and u are proportional to $\mathrm{e}^{i(kx-\sigma t)}$ leads to the dispersion relation

$$\sigma^2 = \frac{gHk^2}{1 + \frac{H^2}{6}k^2}\,, \tag{2.125}$$

with phase speed

$$c_p = \frac{\sigma}{k} = \frac{\sqrt{gH}}{\left(1 + \frac{H^2}{6}(k^2 + l^2)\right)^{\frac{1}{2}}} \ . \tag{2.126}$$

The expressions (2.125) and (2.126) should be compared to the exact dispersion relation for gravity waves in a single layer fluid from potential flow theory [70]:

$$\sigma = \pm\sqrt{gk\tanh(kH)} \ . \tag{2.127}$$

The dispersion relations (2.59), (2.127), and (2.127) are compared in Figure 2.3. It is

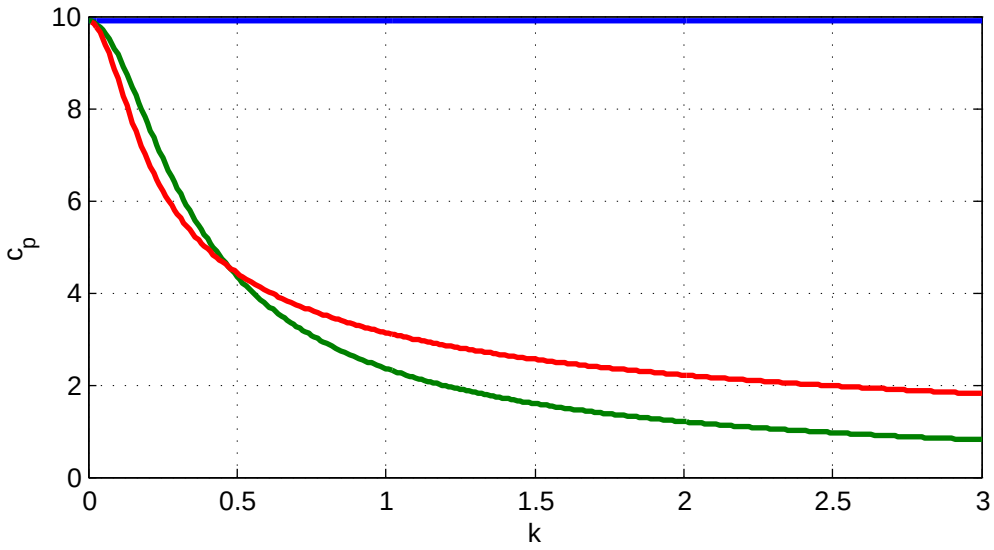

Figure 2.3: Comparison of phase speeds from the traditional shallow water mode (blue), the dispersive shallow water model (green), and the full dispersion relation from potential flow theory (red).

shown that all three dispersion relations agree in the long wave limit $k \to 0$, and that the dispersive shallow water model's dispersion relation agrees qualitatively with the behaviour of the full dispersion relation for gravity waves (2.127).

2.6 Energy considerations

Straightforward manipulation of the system (2.120)–(2.122) yields the following evolution equation for the total energy per unit area

$$\frac{\partial}{\partial t}\left(\frac{1}{2}h\|\mathbf{u}\|^2+\frac{1}{2}g\eta^2\right)=-\nabla\cdot\left(\left[\frac{1}{2}h\|\mathbf{u}\|^2+gh\eta\right]\mathbf{u}\right)+\frac{H^2}{6}\mathbf{u}\cdot\nabla\left(\nabla\cdot(\mathbf{u}h)_t\right)\ , \tag{2.128}$$

where $\|\mathbf{u}\|=\sqrt{u^2+v^2}$. It is instructive to consider the total energy balance for a given domain Ω. Integrating over the whole domain and applying the divergence theorem gives

$$\begin{aligned}\frac{d}{dt}\iint\limits_{\Omega}\frac{1}{2}h\|\mathbf{u}\|^2+\frac{1}{2}g\eta^2 dS= &-\oint\limits_{\partial\Omega}\left(\frac{1}{2}h\|\mathbf{u}\|^2+gh\eta\right)\mathbf{u}\cdot\hat{\mathbf{n}}\,d\ell\\ &+\iint\limits_{\Omega}\frac{H^2}{6}\mathbf{u}\cdot\nabla\left(\nabla\cdot(\mathbf{u}h)_t\right)dS\ .\end{aligned} \tag{2.129}$$

Imposing no normal flow or periodic boundary conditions causes the line integral contribution to the right-hand side to vanish. On the other hand, the integral that results from the dispersive terms has an integrand that cannot be written as a conserved quantity. As a result, the Boussinesq type system (2.120)–(2.122) will not conserve energy on periodic or impermeable domains as is the case with the traditional SWM. In fact, the total energy is oscillatory in time as is demonstrated in section 4.9. There are Boussinesq-type models that conserve energy [17], but they are generally not practical for numerical purposes because they tend to possess 3rd and 4th order spatial derivatives.

It is worth making special mention of the quantity

$$APE=\iint\limits_{\Omega}\frac{1}{2}g\eta^2\,dS\ , \tag{2.130}$$

that appears in the energy equation. It is the "available potential energy." It can also be defined as the time varying part of the potential energy since it can be shown that [86]

$$\frac{d(PE)}{dt}=\frac{d(APE)}{dt}\ . \tag{2.131}$$

Therefore, changes in APE are the only way that dynamical changes can occur to the kinetic energy,

$$KE=\iint\limits_{\Omega}\frac{1}{2}hu^2\,dS\ . \tag{2.132}$$

2.7 Spectral properties

Before discussing numerical solution procedures for the dispersively modified shallow water equations, it is useful to gain insight into the spectral properties of the dispersive shallow water model (2.120)–(2.122). To achieve this, we consider the equations (2.120)–(2.122) for a constant undisturbed depth $H = H_0$ linearized about the background state $\mathbf{u} = 0$ in one spatial dimension

$$\frac{\partial \eta}{\partial t} = -H_0 \frac{\partial u}{\partial x} , \tag{2.133}$$

$$\frac{\partial u}{\partial t} = -g\frac{\partial \eta}{\partial x} + \frac{H_0^2}{6}\frac{\partial^3 u}{\partial^2 x \partial t} . \tag{2.134}$$

If we assume a normal mode solution of the form

$$\begin{pmatrix} \eta \\ u \end{pmatrix} = \begin{pmatrix} \hat{\eta}(t) \\ \hat{u}(t) \end{pmatrix} \mathrm{e}^{ikx} , \tag{2.135}$$

where k is the wavenumber, we obtain

$$\frac{d}{dt}\begin{pmatrix} \hat{\eta}(t) \\ \hat{u}(t) \end{pmatrix} = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}\begin{pmatrix} \hat{\eta}(t) \\ \hat{u}(t) \end{pmatrix} , \tag{2.136}$$

where

$$\alpha = \frac{-igk}{1 + \frac{H_0^2}{6}k^2} , \beta = -ikH . \tag{2.137}$$

From (2.136), one readily obtains the eigenvalues

$$i\omega = \pm \frac{i\sqrt{gH_0}k}{\sqrt{1 + \frac{H_0^2}{6}k^2}} . \tag{2.138}$$

Inspecting the set of eigenvalues (2.138), it is clear that the maximum speed at which information propagates is the expected long wave speed $\sqrt{gH_0}$ as $k \to 0$. Furthermore, one should notice that for $k > 0$, the characteristic wave speed becomes monotonically slower with decreasing wavelength, reflecting the desired dispersive behaviour of short waves.

From a numerical point of view, the form of the eigenvalues (2.138) suggests that all the eigenvalues of the spatial discretization operator should lie on the imaginary interval

$$I = \left[\frac{-i\sqrt{gH_0}k_{max}}{\sqrt{1 + \frac{H_0^2}{6}k_{max}^2}}, \frac{i\sqrt{gH_0}k_{max}}{\sqrt{1 + \frac{H_0^2}{6}k_{max}^2}} \right] . \tag{2.139}$$

If we contrast this interval with the analogous interval for the traditional non-dispersive shallow water model with $i\omega = \pm i\sqrt{gH_0}k$

$$I = \left[-i\sqrt{gH_0}k_{max}, i\sqrt{gH_0}k_{max}\right] \ , \tag{2.140}$$

where k_{max} is the wavenumber of the shortest representable wave as determined by the grid spacing, we notice that the eigenvalues of the Boussinesq-type model are clustered more closely to the origin than the eigenvalues of the traditional shallow water model. To ensure the stability of a numerical ODE integrator, all the eigenvalues of the spatial discretization operator, when scaled by the timestep Δt, are required to lie within its stability region (see Trefethen [76]). Equations (2.139) & (2.140) imply that the Boussinesq-type model will be stable for a larger Δt than the traditional shallow water model.

2.8 Derivation of the two-layer shallow water equations

For the sake of physical interpretation, it is instructive to derive a two-layer model from first principles before proceeding with discussions of including weakly non-hydrostatic effects and numerical methods. Conceptually, the two-layer model is obtained by "stacking" two single-layer shallow water models of different constant density, as illustrated in Figure 2.4. We notice that all flow variables become duplicated and are labeled as subscript '1' if they describe the upper layer, and subscript '2' if they describe the lower layer.

We begin our derivation by considering the equations for conservation of mass. Following an analogous argument used for the one-layer shallow water equations derivation, we can derive statements of conservation of mass in each layer, i.e.,

$$\frac{\partial h_1}{\partial t} + \nabla \cdot (h\mathbf{u}_1) = 0 \ , \tag{2.141}$$

$$\frac{\partial h_2}{\partial t} + \nabla \cdot (h\mathbf{u}_2) = 0 \ . \tag{2.142}$$

where one integration of the three-dimensional incompressibility constraint has been carried out from the bottom at $z = -H + b(x, y)$ to the interface at $z = -H_1 + \zeta(x, y, t)$, and a second from the interface to the top at $z = \eta(x, y, t)$

Once again, the horizontal momentum equations are vertically averaged over each layer and accelerations due to the vertical velocity are neglected. As a final step, appropriate

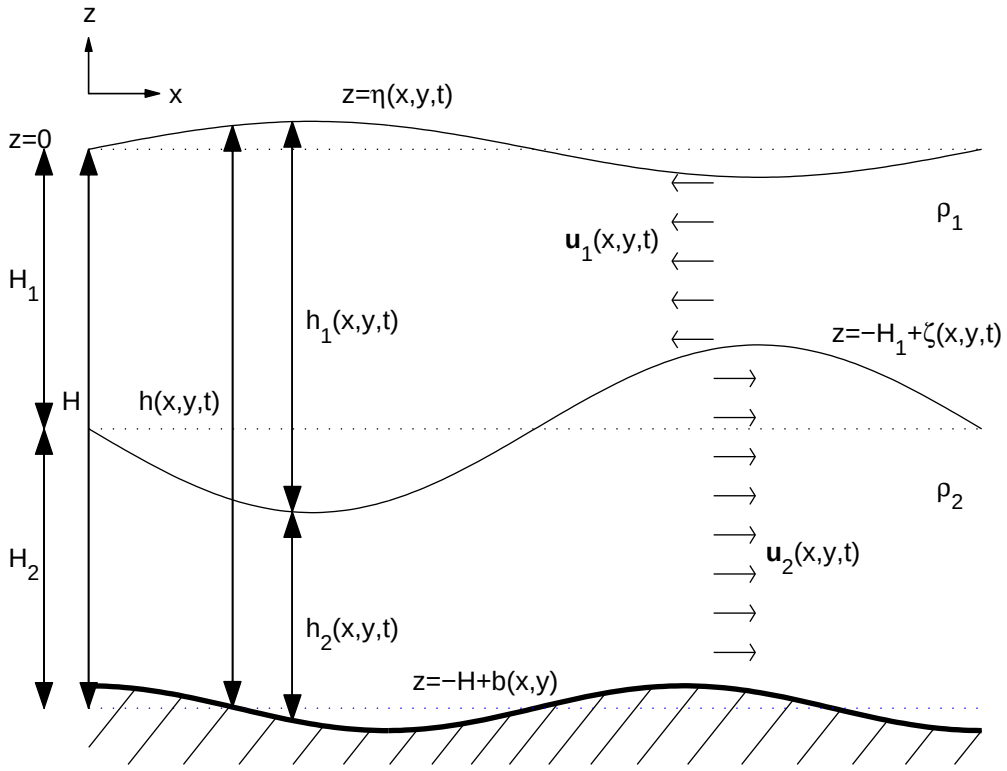

Figure 2.4: Schematic diagram for the two-layer shallow water equations.

expressions for the hydrostatic pressure in each layer must be determined. In the upper layer, $-H_1 + \zeta \leq z \leq \eta$, the pressure takes the same form as in the single-layer shallow water equations. Integrating the hydrostatic pressure relation from the top $z' = \eta$ to an arbitrary depth z in the upper layer, we have

$$p(x, y, z, t) = \rho_1 g\eta(x, y, t) - \rho_1 g z \ , \tag{2.143}$$

and the horizontal pressure gradient is

$$\nabla p(x, y, t) = \rho_1 g \nabla\eta(x, y, t) \ , \quad \text{for} \quad -H_1 + \zeta \leq z \leq \eta \ . \tag{2.144}$$

This is often re-written in terms of the layer thicknesses by noting (from the diagram) that

$$\begin{aligned} h &= h_1 + h_2 = H_1 + H_2 + \eta - b \ , & (2.145) \\ \Rightarrow \nabla\eta &= \nabla h + \nabla b = \nabla h_1 + \nabla h_2 + \nabla b \ . & (2.146) \end{aligned}$$

The pressure gradient in the upper layer then takes the form

$$\nabla p(x, y, t) = \rho_1 g \nabla h + \rho_1 g \nabla b = \rho_1 g \nabla h_1 + \rho_1 g \nabla h_2 + \rho_1 g \nabla b \ . \tag{2.147}$$

This rearrangement is useful for numerical methods that require the equation to be in conservative form, since it separates the terms that involve conserved quantities (layer thicknesses) from source terms (bathymetry).

To determine the form of the horizontal pressure gradient in the lower layer $-H_1 - H_2 - b \leq z \leq -H_1 + \zeta$, we integrate the hydrostatic pressure relation from the interface $z' = -H_1 + \eta$ to an arbitrary depth z in the lower layer

$$\begin{aligned} \int_{z'=-H_1+\zeta}^{z'=z} \frac{\partial p}{\partial z'} \, dz' &= \int_{z'=-H_1+\zeta}^{z'=z} -\rho_2 g \, dz' \ , & (2.148) \\ \Rightarrow p(z' = z) - p(z' = -H_1 + \zeta) &= -\rho_2 g z + \rho_2 g \left(-H_1 + \zeta\right) \ . & (2.149) \end{aligned}$$

The value of p at the interface can be found by applying continuity of pressure between the two layers. Equation (2.143) gives us the dynamic boundary condition

$$p(x, y, -H_1 + \zeta, t) = \rho_1 g\eta(x, y, t) + \rho_1 g H_1 - \rho_1 g\zeta \ . \tag{2.150}$$

Substituting (2.150) into (2.149) gives

$$p(x, y, z, t) = \rho_1 g\eta(x, y, t) + (\rho_2 - \rho_1) g\zeta - (\rho_2 - \rho_1) g H_1 - \rho_2 g z \ , \tag{2.151}$$

with horizontal pressure gradient

$$\nabla p = \rho_1 g \nabla \eta + (\rho_2 - \rho_1) g \nabla \zeta \ , \quad \text{for} \quad -H_1 - H_2 - b \leq z \leq -H_1 + \zeta \ . \tag{2.152}$$

As with the upper layer, the pressure gradient in the lower layer can be written in terms of the layer thicknesses as

$$\nabla p = \rho_1 g \nabla h_1 + \rho_2 g \nabla h_2 + \rho_2 g \nabla b \ . \tag{2.153}$$

By substituting in the expressions for the horizontal pressure gradients in the upper (2.147) and lower (2.153) layers, the depth-averaged momentum equations

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla p \ , \tag{2.154}$$

take the form

$$\rho_1 \frac{D\mathbf{u}_1}{Dt} = -\left[\rho_1 g \nabla h_1 + \rho_1 g \nabla h_2 + \rho_1 g \nabla b\right] \ , \tag{2.155}$$

$$\rho_2 \frac{D\mathbf{u}_2}{Dt} = -\left[\rho_1 g \nabla h_1 + \rho_2 g \nabla h_2 + \rho_2 g \nabla b\right] \ , \tag{2.156}$$

in the upper and lower layers, respectively. That is,

$$\frac{D\mathbf{u}_1}{Dt} = -g \nabla h_1 - g \nabla h_2 - g \nabla b \ , \tag{2.157}$$

$$\frac{D\mathbf{u}_2}{Dt} = -r g \nabla h_1 - g \nabla h_2 - g \nabla b \ , \tag{2.158}$$

where $r = \rho_1 / \rho_2$. For a stable stratification, we require $r < 1$ and in limnic/oceanic flows we expect $\rho_1 / \rho_2 \sim 1$ (small density differences, typically $O(1\%)$ or less).

For the purposes of many numerical methods, the momentum equations are more usefully written in terms of the transport (or pseudo-momentum) in each layer, $h_1 \mathbf{u}_1$ and $h_2 \mathbf{u}_2$. Evolution equations for the layer-wise transport can be derived by multiplying each momentum equation by its respective layer thickness and adding its corresponding continuity equation multiplied by the layer's fluid velocity. Straightforward algebraic manipulations yield

$$\frac{\partial \left(h_1 \mathbf{u}_1\right)}{\partial t} + \nabla \cdot \left(h_1 \mathbf{u}_1 \otimes \mathbf{u}_1\right) + \nabla \left(\frac{1}{2} g h_1^2\right) = -g h_1 \nabla h_2 - g h_1 \nabla b \ , \tag{2.159}$$

$$\frac{\partial \left(h_2 \mathbf{u}_2\right)}{\partial t} + \nabla \cdot \left(h_2 \mathbf{u}_2 \otimes \mathbf{u}_2\right) + \nabla \left(\frac{1}{2} g h_2^2\right) = -r g h_2 \nabla h_1 - g h_2 \nabla b \ . \tag{2.160}$$

2.8.1 1D Hyperbolic Theory for the two-layer system

In 1D, the two-layer shallow water system (2.141),(2.142), (2.159), (2.160) takes the form

$$\frac{\partial h_1}{\partial t} + \frac{\partial (h_1 u_1)}{\partial x} = 0 \,, \tag{2.161}$$

$$\frac{\partial h_2}{\partial t} + \frac{\partial (h_2 u_2)}{\partial x} = 0 \,, \tag{2.162}$$

$$\frac{\partial (h_1 u_1)}{\partial t} + \frac{\partial \left(h_1 u_1^2 + \frac{1}{2} g h_1^2\right)}{\partial x} = -g h_1 h_{2x} - g h_1 b_x \,, \tag{2.163}$$

$$\frac{\partial (h_2 u_2)}{\partial t} + \frac{\partial \left(h_2 u_2^2 + \frac{1}{2} g h_2^2\right)}{\partial x} = -r g h_2 h_{1x} - g h_2 b_x \,. \tag{2.164}$$

We notice that equations (2.161)–(2.164) are not a system of conservation laws (see Section 2.3.3) due to the nonlinear coupling terms on the right-hand side of (2.163)–(2.164). Physically, this makes sense because we cannot expect momentum to be conserved within a single layer since momentum must be allowed to be exchanged between the two layers. To investigate the hyperbolicity of the system, we must examine the eigenstructure of the equations in quasilinear form $\mathbf{q}_t + A\mathbf{q}_x = \mathbf{S}(\mathbf{q})$ where $A = \frac{\partial \mathbf{F}}{\partial \mathbf{q}}$, $\mathbf{q} = (h_1, h_1 u_1, h_2, h_2 u_2)^T$, and

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ g h_1 - u_1^2 & 2u_1 & g h_1 & 0 \\ 0 & 0 & 0 & 1 \\ r g h_2 & 0 & g h_2 - u_2^2 & 2u_2 \end{pmatrix} . \tag{2.165}$$

The characteristic equation is given by

$$((\lambda - u_1)^2 - g h_1)((\lambda - u_2)^2 - g h_2) - r g^2 h_1 h_2 = 0 \tag{2.166}$$

As explained by Mandli [85], exact roots are complicated, and it is useful to make approximations. One set of roots is attainable via a velocity difference expansion, i.e., we assume $u_1 - u_2 = \epsilon$ is small, then

$$\lambda_{bt}^{\pm} \approx \frac{h_1 u_1 + h_2 u_2}{h_1 + h_2} \pm \sqrt{g(h_1 + h_2)} \,. \tag{2.167}$$

If $u_1 = u_2$, $\lambda_{bt}^{\pm}$ would reduce to the long linear wave speed from the single-layer shallow water equations and corresponds to the propagation speed of surface (or barotropic) waves in a two-layer fluid.

For the second set of roots (the baroclinic or internal mode), it is necessary to invoke the rigid-lid approximation that is explained in further detail in Section 2.10, yielding

$$\lambda_{bc}^{\pm} \approx \frac{h_1 u_2 + h_2 u_1}{h_1 + h_2} \pm \sqrt{g(1-r)\frac{h_1 h_2}{h_1 + h_2}\left[1 - \frac{(u_1 - u_2)^2}{g(1-r)(h_1 + h_2)}\right]} \,. \tag{2.168}$$

This expression allows us to define an approximate hyperbolicity condition:

$$\kappa = \frac{(u_1 - u_2)^2}{g(1-r)(h_1 + h_2)} \leq 1$$

Since $\kappa > 1$ yields complex eigenvalues (wave speeds), the governing equations are not hyperbolic in such situations. Physically, these complex eigenvalues correspond to Kelvin–Helmholtz or Miles–Howard Instability [85].

2.9 Boussinesq approximation in 2-layer flow

As mentioned above, the density difference between each layer is expected to be small. A common approximation is thus to replace the density factor on the left-hand sides of (2.155) and (2.156) with a constant reference value ρ_0, such that $\rho_1/\rho_0 \sim \rho_2/\rho_0 \sim 1$. If we write the pressure gradients in terms of the free-surface and interfacial displacements, equations (2.155) and (2.156) become

$$\frac{D\mathbf{u}_1}{Dt} = -g\nabla\eta \,, \tag{2.169}$$

$$\frac{D\mathbf{u}_2}{Dt} = -g\nabla\eta - g'\nabla\zeta \,, \tag{2.170}$$

where $g' = g(\rho_2 - \rho_1)/\rho_0 \approx g(\rho_2 - \rho_1)/\rho_2$[2] is the *reduced gravity*. In terms of the layer thicknesses, we have

$$\frac{D\mathbf{u}_1}{Dt} = -g\nabla h_1 - g\nabla h_2 - g\nabla b \,, \tag{2.171}$$

$$\frac{D\mathbf{u}_2}{Dt} = -g\nabla h_1 - (g+g')\nabla h_2 - (g+g')\nabla b \,, \tag{2.172}$$

[2]It is instructive to write ρ_2 in the denominator, so that one recovers $g' \to g$ in non-Boussinesq situations where $\rho_1 \ll \rho_2$, such as an air-water interface.

and the equations for the layer-wise transport are

$$\frac{\partial\left(h_1\mathbf{u}_1\right)}{\partial t}+\nabla\cdot\left(h_1\mathbf{u}_1\otimes\mathbf{u}_1\right)+\nabla\left(\frac{1}{2}gh_1^2\right) = - gh_1\nabla h_2 - gh_1\nabla b\ , \quad (2.173)$$

$$\frac{\partial\left(h_2\mathbf{u}_2\right)}{\partial t}+\nabla\cdot\left(h_2\mathbf{u}_2\otimes\mathbf{u}_2\right)+\nabla\left(\frac{1}{2}(g+g')h_2^2\right) = - gh_2\nabla h_1 - (g+g')h_2\nabla b\ . \quad (2.174)$$

2.9.1 Equation for the barotropic mode

In the parlance of two-layer flow, physical insights into the flow structure can be made by decomposing the velocity profile into its depth-averaged and shear components. These are referred to as the barotropic $\mathbf{u}_{bt}$ and baroclinic velocity $\mathbf{u}_{bc}$, respectively. Following the conventions of Gill [48], these velocities take the form:

$$\mathbf{u}_{bt} = \frac{h_1\mathbf{u}_1+h_2\mathbf{u}_2}{h}\ , \quad (2.175)$$

$$\mathbf{u}_{bc} = \mathbf{u}_1-\mathbf{u}_2\ . \quad (2.176)$$

The barotropic velocity $\mathbf{u}_{bt}$ is a measure of the depth-independent part of the flow (the part of the flow that behaves as a one-layer fluid), while $\mathbf{u}_{bc}$ gives a measure of the strength and direction of the vertical shear (the depth-dependent part of the flow) and is of utmost importance in the description of internal waves. Since under an appropriate linearization, any two-layer flow can be decomposed into linearly independent barotropic and baroclinic components, these two velocities are often called the barotropic and baroclinic *modes* [48].

An important consequence of using the Boussinesq approximation in two-layer flow is that the corresponding equation for the barotropic mode takes the form of a hyperbolic conservation law (plus source terms for variable bottom bathymetry). Adding equations (2.173)–(2.174) and recognizing that $h_1\mathbf{u}_1+h_2\mathbf{u}_2=h\mathbf{u}_{bt}$ gives

$$\frac{\partial(h\mathbf{u}_{bt})}{\partial t}+\nabla\cdot(h_1\mathbf{u}_1\otimes\mathbf{u}_1+h_2\mathbf{u}_2\otimes\mathbf{u}_2))+\nabla\left(\frac{1}{2}gh^2+\frac{1}{2}g'h_2^2\right)=-gh\nabla b-g'h_2\nabla b\ . \quad (2.177)$$

This equation is similar to the transport equation for the 1-layer shallow water equations. Its nonlinear advection terms can be further decomposed into barotropic and shear components (see [26]), but there is little to be gained by carrying out this exercise here. It is worth noting that a conservative form for the barotropic mode in the non-Boussinesq formulation (2.157)–(2.158) cannot be directly recovered for $h\mathbf{u}_{bt}$ since the parameter $r\neq 1$, and to

achieve such a conservative form requires the non-Boussinesq system to be put back into the form (2.155)–(2.156), and the conserved quantity in that case is $(\rho_1 h_1 \mathbf{u}_1 + \rho_2 h_2 \mathbf{u}_2)$.

The continuity equations in each layer may be added as well, yielding

$$\frac{\partial h}{\partial t} + \nabla \cdot (h \mathbf{u}_{bt}) = 0 \, . \tag{2.178}$$

The fact that equations (2.177) and (2.178) are so similar to the one-layer shallow water equations has consequences in the formulation of numerical methods. In particular, it leads to a technique called *mode splitting* [26] where the fast surface gravity waves are evolved separately from the internal waves, saving the computational effort of time-stepping all six equations with the small time-step associated the surface gravity wave speed $\sqrt{gH}$.

2.10 Rigid-lid approximation

In many applications, surface gravity waves may be too fast to be important in the dominant physics of the flow. Since this thesis is primarily concerned with developing simple models for internal waves on the basin scale, much focus is placed on the *rigid-lid approximation* where the free-surface is replaced by a rigid horizontal wall that bounds the flow at $z = 0$, as depicted in the schematic in Figure 2.5.

Such an abstraction of the free surface effectively 'filters out' the surface gravity waves, or equivalently, gives them an infinite propagation speed. The corresponding re-derivation of the two-layer shallow water equations requires special attention to be paid to the depth-integrated hydrostatic relation in upper layer

$$\int_{z'=0}^{z'=z} \frac{\partial p}{\partial z'} \, dz' = \int_{z'=0}^{z'=z} -\rho_1 g \, dz' \, , \tag{2.179}$$

$$\Rightarrow p(z' = z) = p_{lid} - \rho_1 g z \, , \tag{2.180}$$

where $p_{lid} = p(z' = 0)$ is the rigid-lid pressure. We conclude that $p_{lid} = p_{lid}(x, y, t)$, otherwise, there would be no mechanism for inducing horizontal accelerations in the upper-layer. The horizontal pressure gradient in the upper layer is then simply ∇p_{lid}.

If we follow an analogous line of reasoning that lead us to the expression for the pressure gradient in the lower layer in Section 2.8 and invoke the Boussinesq approximation as in

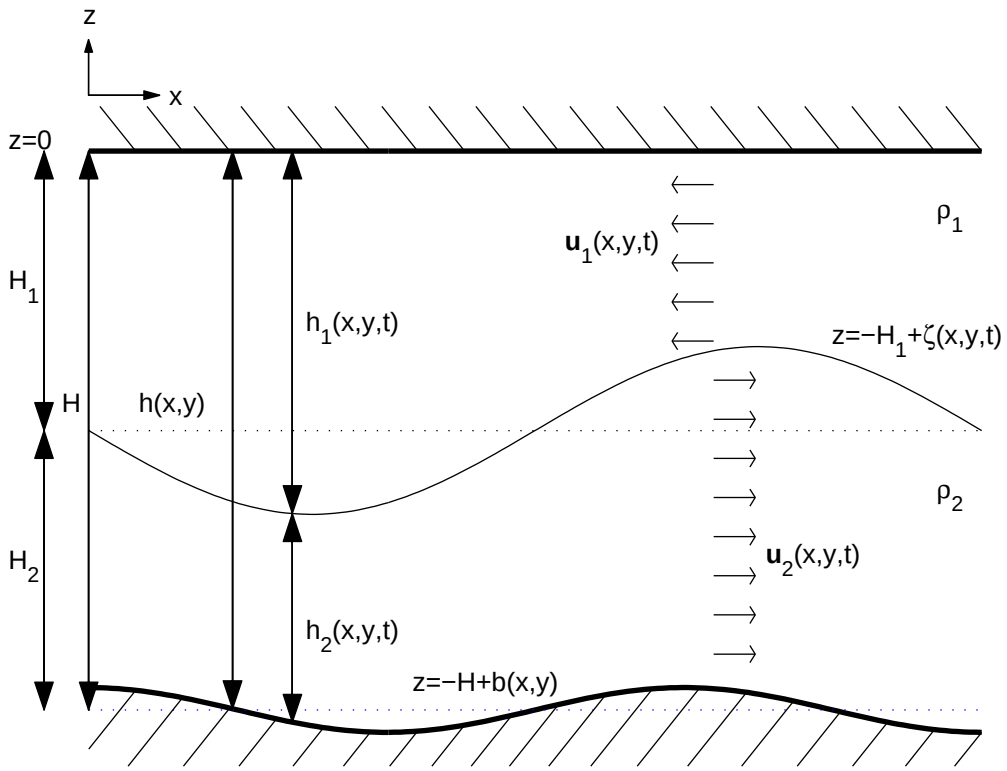

Figure 2.5: Schematic diagram for the two-layer shallow water equations under the rigid-lid approximation.

Section 2.9, we arrive at the momentum equations

$$\frac{D\mathbf{u}_1}{Dt} = -\nabla\tilde{p}_{lid} \,, \tag{2.181}$$

$$\frac{D\mathbf{u}_2}{Dt} = -\nabla\tilde{p}_{lid} - g'\nabla\zeta \,. \tag{2.182}$$

Here, $\tilde{p}_{lid} = p_{lid}/\rho_0$, and we drop the tilde hereafter for notational convenience. The equations for the layer-wise transport are again found by multiplying (2.181) and (2.182) by $h_1 = H_1 - \zeta$ and $h_2 = H_2 + \zeta - b$, and adding them to their corresponding continuity equation scaled by the corresponding velocity. Doing so gives

$$\frac{\partial\left(h_1\mathbf{u}_1\right)}{\partial t} + \nabla\cdot\left(h_1\mathbf{u}_1\otimes\mathbf{u}_1\right) = -h_1\nabla p_{lid} \,, \tag{2.183}$$

$$\frac{\partial\left(h_2\mathbf{u}_2\right)}{\partial t} + \nabla\cdot\left(h_2\mathbf{u}_2\otimes\mathbf{u}_2\right) + \nabla\left(\frac{1}{2}g'h_2^2\right) = -h_2\nabla p_{lid} - g'h_2\nabla b \,. \tag{2.184}$$

2.10.1 Rigid-lid barotropic mode

As in Section 2.9.1, equations (2.183) and (2.184) may be added to form an equation for the barotropic transport $h\mathbf{u}_{bt}$, where in the rigid-lid case $h(x,y) = H_1 + H_2 - b(x,y)$ is the total depth of the fluid, yielding

$$\frac{\partial(h\mathbf{u}_{bt})}{\partial t} + \nabla\cdot\left(h_1\mathbf{u}_1\otimes\mathbf{u}_1 + h_2\mathbf{u}_2\otimes\mathbf{u}_2\right)) + \nabla\left(\frac{1}{2}g'h_2^2\right) = -h\nabla p_{lid} - g'h_2\nabla b \,. \tag{2.185}$$

Similarly, the continuity equations in each layer may be added

$$h_t + \nabla\cdot\left(h\mathbf{u}_{bt}\right) = 0 \,. \tag{2.186}$$

However, the rigid-lid approximation implies that the total depth h is time-independent ($h = h(x,y)$ only), so we have

$$\nabla\cdot\left(h\mathbf{u}_{bt}\right) = \nabla\cdot\left(h_1\mathbf{u}_1 + h_2\mathbf{u}_2\right) = 0, \ , \tag{2.187}$$

i.e. the barotropic transport is *incompressible*. This result makes the barotropic mode under the rigid-lid approximation qualitatively different from the barotropic mode derived when the upper boundary was a free surface in Section (2.9.1). In that case, the barotropic mode satisfied a set of equations similar to the single-layer shallow water equations. Equations (2.185) and (2.187), on the other hand, are more similar to the horizontal incompressible Euler equations.

The kinematic constraint (2.187) assigns a more clear interpretation to the rigid-lid pressure: that it is present in the equations only as a means of enforcing incompressibility on the barotropic transport and has no other physical meaning.

2.10.2 Hyperbolic theory – Rigid-lid internal wave speed

Long [79] carried out one of the earliest works to derive the nonlinear characteristic speeds for a two-layer fluid bounded above by a rigid lid. He achieved this derivation by noting that in 1D, the incompressibility constraint

$$(h_1 u_1 + h_2 u_2)_x = 0 \,, \tag{2.188}$$

can be simply integrated to give

$$(H - h_2) u_1 + h_2 u_2 = K \,, \tag{2.189}$$

where K is a constant. In a fluid bounded by vertical walls, $K = 0$, and for an unbounded domain, K is expected to be at least time-independent. He proceeds by noting that when the momentum equations in each layer are written in terms of the velocities

$$u_{1t} + u_1 u_{1x} = -p_x \,, \tag{2.190}$$
$$u_{2t} + u_2 u_{2x} = -p_x - g' h_{2x} \,, \tag{2.191}$$

that pressure can be eliminated by subtracting the equations, yielding

$$u_{2t} + u_2 u_{2x} - u_{1t} - u_1 u_{1x} + g' h_{2x} = 0 \,. \tag{2.192}$$

u_1 can then be eliminated from this equation by rearranging (2.189) for u_1 and substituting into (2.192). Once this equation has been simplified, it can be combined with the lower layer's continuity equation to form a closed quasilinear system of equations for h_2 and u_2. By computing the eigenvalues of this system, Long [79] calculated the characteristic wave speeds to be

$$\lambda_{bc}^{\pm} \approx \frac{h_1 u_2 + h_2 u_1}{h_1 + h_2} \pm \sqrt{g' \frac{h_1 h_2}{h_1 + h_2} \left[1 - \frac{(u_1 - u_2)^2}{g'(h_1 + h_2)} \right]} \,. \tag{2.193}$$

This expression was used to approximate the baroclinic characteristic speeds in the full two-layer system in Section 2.8.1 assuming that $\sqrt{g' H_e} \ll \sqrt{gH}$, i.e., the linear baroclinic long-wave speed is much less than the linear barotropic long wave-speed. This can be seen by noting that if the time-scale of interest is set by the internal wave speed, then the barotropic continuity equation can be non-dimensionalized via

$$\mathbf{u}_{bt} = \left(\frac{a}{H} \right) c_{bt} \tilde{\mathbf{u}}_b t \,, \quad \eta = a\tilde{\eta} \,, \quad t = \frac{L}{c_{bc}} \tilde{t} \,, \quad h = H\tilde{h} \,, x = L\tilde{x} \tag{2.194}$$

to obtain

$$\epsilon\frac{\partial\tilde{\eta}}{\partial\tilde{t}} + \nabla\cdot\left(\tilde{h}\tilde{\mathbf{u}_{bt}}\right) = 0\ , \tag{2.195}$$

where $\epsilon = \sqrt{g'H_e}/\sqrt{gH}$. Thus, the rigid-lid equations would represent the $O(1)$ contribution to a singular perturbation expansion in ϵ to the full two-layer equations. For a typical lake of depth $H = 25.6$ m and temperature difference $\Delta T = 19°$C we have $H_1 = 14.5$ m, $H_2 = 11.1$ m and $g' = 0.0025g = 0.024525$ ms^{-2}. In which case, $c_{bt} = 15.8473$ ms^{-1} and $c_{bc} = 0.3927$ ms^{-1} so $\epsilon = 0.0248$ and the use of the rigid-lid approximation appears to be well motivated in this case.

2.11 Weakly non-hydrostatic 2-layer rigid-lid model

A suitable two-layer model with weakly non-hydrostatic effects that approximates the free surface with a rigid lid is given by

$$\frac{\partial h_i}{\partial t} + \nabla\cdot(h_i\mathbf{u}_i) = 0\ , \quad i = 1,2\ , \tag{2.196}$$

$$\frac{\partial(h_1u_1)}{\partial t} + \nabla\cdot((h_1u_1)\mathbf{u}_1) = -h_1\frac{\partial p}{\partial x} + fh_1v_1\ , \tag{2.197}$$

$$\frac{\partial(h_1v_1)}{\partial t} + \nabla\cdot((h_1v_1)\mathbf{u}_1) = -h_1\frac{\partial p}{\partial y} - fh_1u_1\ , \tag{2.198}$$

$$\begin{aligned}\frac{\partial(h_2u_2)}{\partial t} + \nabla\cdot((h_2u_2)\mathbf{u}_2) &= -g'h_2\frac{\partial\zeta}{\partial x} - h_2\frac{\partial p}{\partial x}\\ &+ fh_2v_2 + \gamma\nabla\cdot(h_2\mathbf{u}_2)_{xt}\ ,\end{aligned} \tag{2.199}$$

$$\begin{aligned}\frac{\partial(h_2v_2)}{\partial t} + \nabla\cdot((h_2v_2)\mathbf{u}_2) &= -g'h_2\frac{\partial\zeta}{\partial y} - h_2\frac{\partial p}{\partial y}\\ &- fh_2u_2 + \gamma\nabla\cdot(h_2\mathbf{u}_2)_{yt}\ ,\end{aligned} \tag{2.200}$$

corresponding to conservation of mass in the upper and lower layers (2.196), conservation of momentum in the upper layer (2.197)-(2.198), and conservation of momentum in the lower layer (2.199)-(2.200). Here, p is the pressure at the rigid lid, ζ is the interfacial displacement, and $h_1 = H_1 - \zeta$, $h_2 = H_2 + \zeta$ so that $h_1 + h_2 = H_1 + H_2 = H(x,y)$ gives the total depth that is constant in time. The reduced gravity is $g' = (\rho_2 - \rho_1)g/\rho_2$ and, f is the Coriolis parameter. The parameter $\gamma = H_1H_2/3 + H_2^2/3$ is chosen for reasons that will be made clear in the section below.

A kinematic constraint corresponding to the global conservation of mass can be obtained by adding the equations for conservation of mass in each layer (2.196) and recognizing that

$(h_1 + h_2)_t = H_t = 0$, hence,

$$\nabla \cdot (h_1 \mathbf{u}_1 + h_2 \mathbf{u}_2) = 0 \ . \tag{2.201}$$

This result is reminiscent of the conservation of mass constraint in the full incompressible Navier Stokes equations, $\nabla \cdot \mathbf{u} = 0$.

Non-hydrostatic effects have only been included in the lower layer for simplicity of the numerical algorithm, since including non-hydrostatic effects would require the numerical algorithm to solve three elliptic problems per time-step instead of two. Additionally, it will be shown in the section below on dispersion characteristics that non-hydrostatic effects need only be included in one of the layers in order to achieve the first non-hydrostatic correction to the hydrostatic dispersion relation.

2.11.1 Dispersion Characteristics

The dispersion relation for internal waves at the interface between two immiscible fluid layers (bounded by straight, rigid walls) of undisturbed depths H_1 and H_2 is given by

$$\sigma^2 = \frac{(\rho_2 - \rho_1) g k \tanh(kH_1) \tanh(kH_2)}{\rho_2 \tanh(kH_1) + \rho_1 \tanh(kH_2)} \ , \tag{2.202}$$

which is derived from the linearized Bernoulli equations in a 2-layer fluid (see appendix A). Under the Boussinesq approximation, $\rho_2/\rho_1 \approx 1$, hence

$$\sigma^2 \approx \frac{g' k \tanh(kH_1) \tanh(kH_2)}{\tanh(kH_1) + \tanh(kH_2)} \ , \tag{2.203}$$

where $g' = (\rho_2 - \rho_1) g / \rho_2$ is the reduced gravity, as in the above section. Since no approximations have been made with respect to aspect ratio (or the ratio of depth to wavelength), this relation is fully non-hydrostatic. Its Taylor series expanded about the hydrostatic or long-wave limit $k \ll 1$ is given by

$$\sigma^2 \sim g' \frac{H_1 H_2}{H_1 + H_2} k^2 - \frac{1}{3} g' \frac{H_1^2 H_2^2}{H_1 + H_2} k^4 + O(k^6) \ , \text{ as } k \to 0 \ . \tag{2.204}$$

We now wish to show that the dispersion relation of the model (2.196)-(2.200) agrees with (2.204) to O(k^4).

Reducing the model to one dimension, assuming a flat bottom, and linearizing gives

$$
\begin{aligned}
u_{1t} &= -p_x \ , && (2.205)\\
u_{2t} &= -p_x - g'\eta_x + \gamma u_{2xxt} \ , && (2.206)\\
\eta_t &= -\frac{1}{2}\left(H_2 u_{2x} - H_1 u_{1x}\right) \ , && (2.207)\\
H_1 u_{1x} + H_2 u_{2x} &= 0 \ . && (2.208)
\end{aligned}
$$

Here, the evolution equations for h_1 and h_2 have been added to yield a single prognostic equation for η, although under the rigid lid constraint (2.208), two additional forms for the η equation are possible and equivalent.

To find the dispersion relation of the linearized system, we assume a normal modes form for the solution

$$
\begin{pmatrix} u_1 \\ u_2 \\ \eta \\ p \end{pmatrix} = \begin{pmatrix} \hat{u_1} \\ \hat{u_2} \\ \hat{\eta} \\ \hat{p} \end{pmatrix} \mathrm{e}^{i(kx-\sigma t)} \ . \tag{2.209}
$$

Under this ansatz, the system (2.205)-(2.208) in matrix form becomes

$$
\begin{pmatrix} \sigma & 0 & 0 & -k \\ 0 & \sigma + \gamma\sigma k^2 & -gk & -k \\ \frac{kH_1}{2} & -\frac{kH_2}{2} & \sigma & 0 \\ H_1 & H_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{u_1} \\ \hat{u_2} \\ \hat{\eta} \\ \hat{p} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \ . \tag{2.210}
$$

Non-trivial solutions are obtained by setting the determinant of the matrix equal to zero. Doing so gives

$$
\begin{aligned}
\sigma^2 &= g' \frac{H_1 H_2}{H_1 + H_2 + H_1\left(\frac{1}{3}H_1 H_2 + \frac{1}{3}H_2^2\right)k^2} k^2 \ , \\
&\sim g' \frac{H_1 H_2}{H_1 + H_2} k^2 - \frac{1}{3} g' \frac{H_1^2 H_2^2}{H_1 + H_2} k^4 + O(k^6) \ , \ \text{as } k \to 0 \ , && (2.211)
\end{aligned}
$$

where we have used $\gamma = H_1 H_2/3 + H_2^2/3$. This expansion agrees with the fully non-hydrostatic dispersion relation (2.204) to $O(k^4)$.

It is worth asking what might be gained by including dispersive effects in both the upper- and lower-layer momentum equations. To investigate this question, let's consider the situation in which the term $\alpha_1 u_{1xxt}$ is present in equation (2.205) and the term $\alpha_2 u_{2xxt}$

is present in equation (2.206), and α_1, α_2 have yet to be determined. After substituting the normal modes ansatz, we recover the matrix form

$$\begin{pmatrix} \sigma+\alpha_1\sigma k^2 & 0 & 0 & -k \\ 0 & \sigma+\alpha_2\sigma k^2 & -gk & -k \\ \frac{kH_1}{2} & -\frac{kH_2}{2} & \sigma & 0 \\ H_1 & H_2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{u_1} \\ \hat{u_2} \\ \hat{\eta} \\ \hat{p} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} . \tag{2.212}$$

Taking the determinant of this matrix and setting it equal to zero gives the dispersion relation

$$\begin{aligned} \sigma^2 &= g' \frac{H_1 H_2}{H_1 + H_2 + \alpha_1 H_2 k^2 + \alpha_2 H_1 k^2} k^2 , \\ &\sim g' \frac{H_1 H_2}{H_1 + H_2} k^2 - g' \frac{H_1 H_2 (\alpha_1 H_2 + \alpha_2 H_1)}{(H_1 + H_2)^2} k^4 + O(k^6) , \text{ as } k \to 0 . \end{aligned} \tag{2.213}$$

To match with the $O(k^4)$ term from the non-hydrostatic dispersion relation (2.204), we need

$$\frac{1}{3} H_1 H_2 (H_1 + H_2) = H_2 \alpha_1 + H_1 \alpha_2 , \tag{2.214}$$

and we find that we are left with a free parameter. A natural way to close the system for (α_1, α_2) would be to match the dispersion relation to $O(k^6)$. However, since we are satisfied with agreement up to $O(k^4)$ we can set $\alpha_1 = 0$, and find $\alpha_2 = \gamma = H_1 H_2/3 + H_2^2/3$ and recover a simpler set of model equations. Other choices of note are $\alpha_1 = \alpha_2 = H_1 H_2/3$ (same coefficient in each layer) and $\alpha_2 = 0$, $\alpha_1 = H_1 H_2/3 + H_1^2/3$ (dispersive effects in upper layer only). It is worth noting that the true reason behind the occurrence of the free parameter is that if a free surface were included, the second parameter would be required to capture the $O(k^4)$ dispersion characteristics in the surface gravity waves. The fact that we have 'filtered out' the free surface waves has provided us with the freedom of placing a computationally expensive mixed-derivative term in just one of the layers.

2.11.2 Wave speeds and Stability Properties in the presence of background shear flow

If we consider 1D hydrostatic 2-layer rigid lid model and linearize about the background state $u_1 = U_1$ and $u_2 = U_2$, where U_1 and U_2 are constants, i.e.,

$$\begin{aligned} u_1 &= U_1 + u_1' \,, && (2.215)\\ u_2 &= U_2 + u_2' \,, && (2.216)\\ \eta &= \eta' \,. && (2.217)\\ p &= p' \,. && (2.218) \end{aligned}$$

In the case where $U_1 \neq U_2$ we have a simple piece-wise constant model of shear flow in a 2-layer fluid. The linearized equations become

$$\begin{aligned} u_{1t}' + U_1 u_{1x}' + p'_x &= 0 \,, && (2.219)\\ u_{2t}' + U_2 u_{2x}' + p'_x + g'\eta'_x &= 0 \,, && (2.220)\\ \eta'_t + \frac{1}{2}\left(U_1\eta'_x + U_2\eta'_x + H_2 u_{2x}' - H_1 u_{1x}'\right) &= 0 \,, && (2.221)\\ H_1 u_{1x}' + H_2 u_{2x}' + (U_2 - U_1)\eta'_x &= 0 \,, && (2.222) \end{aligned}$$

Substituting the normal modes ansatz (2.209) into these equations yields the system

$$\begin{pmatrix} \sigma - kU_1 & 0 & 0 & -k \\ 0 & \sigma - kU_2 & -gk & -k \\ \frac{kH_1}{2} & -\frac{kH_2}{2} & \sigma - \frac{k}{2}(U_1 + U_2) & 0 \\ H_1 & H_2 & (U_2 - U_1) & 0 \end{pmatrix} \begin{pmatrix} \hat{u_1} \\ \hat{u_2} \\ \hat{\eta} \\ \hat{p} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \,, \qquad (2.223)$$

Once again, the existence of non-trivial algebraic solutions implies that the determinant of the matrix is zero. Thus we recover the dispersion relation

$$\sigma = \left(\frac{H_1U_2 + H_2U_1}{H}\right) k \pm k\sqrt{g'H_e - \frac{H_e}{H}(U_1 - U_2)^2} \,, \qquad (2.224)$$

where $H = H_1 + H_2$ and $H_e = H_1H_2/H$ are the total and equivalent depths respectively, and we have assumed $k \geq 0$.

Notice that for $|U_1 - U_2|$ sufficiently large, the quantity under the square root becomes negative yielding an imaginary value for the frequency, σ. This situation corresponds to the physical Kelvin-Helmholtz instability. However, in the 2-layer shallow water framework,

there are no mechanisms to arrest the exponential growth of such instabilities. Therefore, having such an instability occur during a numerical simulation would be catastrophic. For the purposes of practical simulation of 2-layer flow, the modeller must restrict their attention to the case of sufficiently weak shear. If physical instabilities are desired in simulations, a more "full" system of equations such as the stratified incompressible Navier-Stokes equations under the Boussinesq approximation must be solved. It is also worth noting that in this simple model, the criteria for instability does not depend on the wavenumber, k.

From the dispersion relation (2.224), it is straight-forward to recover the phase speed of the internal waves as

$$c = \frac{\sigma}{k} = \left(\frac{H_1 U_2 + H_2 U_1}{H} \right) \pm \sqrt{g' H_e - \frac{H_e}{H} (U_1 - U_2)^2} \, . \tag{2.225}$$

It is interesting to note that the presence of background shear flow does not imply a simple Doppler shift in the speed of the waves as one would find in the special case where $U_1 = U_2 = U$ and the term itself reduces to U and the second term under the square root vanishes. Since here the waves are non-dispersive, we also have

$$c_g = \frac{d\sigma}{dk} = c \, , \tag{2.226}$$

for the group speed, which can easily be seen since σ depends linearly on k.

If physical dispersion is included in the model equations, we expect the first term under the square root to be reduced for shorter waves. Thus, for short waves, less shear would be required to drive Kelvin–Helmholtz instabilities than for long waves which agrees well with physical intuition.

Also of note is the fact that (2.225) provides a simple upper-bound on the speed of internal waves in the presence of background flow

$$c \leq \left(\frac{H_1 U_2 + H_2 U_1}{H} \right) + \sqrt{g' H_e} \, , \tag{2.227}$$

using the fact that the second quantity under the square root is negative definite.

2.11.3 Energy considerations

Following the procedure for a 1-layer fluid, energy equations may be derived for the available potential energy density, and the kinetic energies in the upper and lower layers. They read

$$\frac{\partial}{\partial t}\left(\frac{1}{2}g'\eta^2\right) + g'\eta\nabla\cdot(h_2\mathbf{u}_2) = 0\,, \tag{2.228}$$

$$\frac{\partial}{\partial t}\left(\frac{1}{2}h_1\|\mathbf{u}_1\|^2\right) + \nabla\cdot\left(\frac{1}{2}h_1\|\mathbf{u}_1\|^2\mathbf{u}_1\right) = -h_1\mathbf{u}_1\cdot\nabla p\,, \tag{2.229}$$

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{1}{2}h_2\|\mathbf{u}_2\|^2\right) + \nabla\cdot\left(\frac{1}{2}h_2\|\mathbf{u}_2\|^2\mathbf{u}_2\right) &= -h_2\mathbf{u}_2\cdot\nabla p \\ &- g'h_2\mathbf{u}_2\cdot\nabla\eta\,.\end{aligned} \tag{2.230}$$

The total energy density can be found by adding these three equations

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{1}{2}h_1\|\mathbf{u}_1\|^2 + \frac{1}{2}h_2\|\mathbf{u}_2\|^2 + \frac{1}{2}g'\eta^2\right) &= -\nabla\cdot\left(\frac{1}{2}h_1\|\mathbf{u}_1\|^2\mathbf{u}_1\right) \\ &- \nabla\cdot\left(\left[\frac{1}{2}h_2\|\mathbf{u}_2\|^2 + g'h_2\eta\right]\mathbf{u}_2\right) \\ &- (h_1\mathbf{u}_1 + h_2\mathbf{u}_2)\cdot\nabla p \\ &+ \gamma\nabla\left(\nabla\cdot(\mathbf{u_2})_t\right)\cdot h_2\mathbf{u}_2\,.\end{aligned} \tag{2.231}$$

Upon integrating over the whole domain, the first two terms on the right-hand side vanish after applying the divergence theorem along with the no normal flow boundary conditions on $\mathbf{u}_1$ and $\mathbf{u}_2$. The pressure term vanishes after integrating by parts, then invoking the rigid-lid constraint on the resulting domain integral and no normal flow on the line integral. We see that just as in the single-layer case, the dispersive terms cause energy to only be conserved in the long-wave limit where dispersive effects may be neglected.

2.12 Other model choices in the literature

It is worth noting that the choice of dispersive shallow water models discussed above are not unique, and may very well not accurately capture dispersive properties in situations involving variable depth, H.

In the surface waves literature, there are a host of equations referred to as the Boussinesq equations. Peregrine's equations [94] are a 2D dispersive shallow water model for variable

depth derived using Boussinesq's idea [12] of expanding the velocity potential in a Taylor series about the bottom, resulting in a dispersive evolution equation for the depth-averaged velocity. Nwogu's "extended" Boussinesq system [90] attempts to improve the Boussinesq equations by evaluating the velocity at an arbitrary depth-level. This arbitrary depth-level provides a free paramater that is then optimized for accuracy. The "high-order" Boussinesq equations of Madsen *et. al.* [83, 84] add additional dispersive terms with coefficients chosen in such a way as to increase the agreement of the dispersion with linear wave theory. In a review in the thesis of Walkley [126], it is demonstrated on what range of wavelengths each system's dispersion relation agrees well with the dispersion relation for inertia-gravity waves from the full Navier-Stokes equations.

To familiarize the reader with some of these alternative models, we present Peregrine's 2D Boussinesq equations [94] where the depth-averaged velocity $\mathbf{u}$ satisfies

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\,\mathbf{u} = -g\nabla\eta + \frac{H}{2}\nabla\left(\nabla\cdot\left(H\frac{\partial \mathbf{u}}{\partial t}\right)\right) - \frac{H^2}{6}\nabla\left(\nabla\cdot\left(\frac{\partial \mathbf{u}}{\partial t}\right)\right), \tag{2.232}$$

and the fluid column height h satisfies (2.120). In the case of a flat bottom, the dispersive terms can be grouped together as $\frac{H^2}{3}\nabla\nabla\cdot\mathbf{u}_t$. In an analysis by Walkley [126], it is shown that when the dispersion is compared to linear theory, the phase velocity error is over 5% for depths over 3/10 of the wavelength, while the group velocity error is about 10% for depths over 2/10 of the wavelength. The shoaling gradient is also shown to quickly diverge from that predicted from the linear theory at depths over 15% of the wavelength.

A more accurate set of equations are Nwogu's extended Boussinesq system [90] where the velocity is evaluated at an arbitrary depth $z = \theta h$:

$$\frac{\partial h}{\partial t} + \nabla\cdot(h\mathbf{u}) + \nabla\cdot\left(B_1H^3\nabla(\nabla\cdot\mathbf{u}) + B_2H^2\nabla(\nabla\cdot(H\mathbf{u}))\right) = 0, \tag{2.233}$$

$$\frac{\partial \mathbf{u}}{\partial t} + g\nabla\eta + (\mathbf{u}\cdot\nabla)\mathbf{u} + A_1H^2\nabla\left(\nabla\cdot\frac{\partial \mathbf{u}}{\partial t}\right) + A_2H\nabla\left(\nabla\cdot\left(H\frac{\partial \mathbf{u}}{\partial t}\right)\right) = 0, \tag{2.234}$$

where $A_1 = \theta^2/2$, $A_2 = \theta$, $B_1 = \theta^2/2 - 1/6$, and $B_2 = \theta + 1/2$. Given an optimal choice of θ determined by Nwogu [90], Walkley [126] demonstrates that the phase velocity error is less than 1% for depths up to 1/2 of the wavelength, and the group velocity error is at most 7% for depths up to 35% of the wavelength, while the shoaling gradient is accurate for depths up to 3/10 of the wavelength.

Although Nwogu's equations are considerably more accurate than Peregrine's, the additional dispersive terms present in the continuity equation (2.233) add considerable complexity to a numerical solution method, since the dispersive terms represent third-order

spatial derivatives that impose the rather undesireable restriction that the time-step Δt scales like the cube of the grid spacing [60]. Numerical solutions using low order finite-element discretizations where the semi-discrete schemes are fully implicit are presented by Walkley [126]. Low-order finite difference discretizations for Nwogu's equations are considered by Wei and Kirby [129]. At present, solving Nwogu's equations with high order spatial discretization methods seems far too computationally expensive to be practical.

The equations represented by (2.120)–(2.122) are taken as a starting point in modelling dispersion in a single-layer shallow water framework as they are the simplest dispersive shallow water equations available in the literature. At present, the goal of this work is not to give the most accurate description of dispersive waves, but rather to have a simple representation of short wave dispersion that also provides a mechanism for preventing unphysical shocks from forming.

A benefit of using equations (2.120)–(2.122) is that they allow for the momentum equations to be decoupled using the scalar approach of Eskilsson and Sherwin [42] who used this approach to speed up the numerical solution of Peregrine's [94] Boussinesq equations with a high-order modal DG-FEM method. This approach will be applied for pseudospectral methods in Chapter 3 and used again with the nodal DG-FEM method in Chapter 4.

Since this thesis is concerned with developing a simple model for internal waves, at some point it is necessary to study multi-layer dispersive shallow water models as a more accurate description of internal waves. A single layer reduced gravity model approximates the free surface as being infinitely far from the interface, and is thus a crude model of interfacial waves.

While it appears that there are sets of equations that are generally agreed upon for modelling dispersive surface waves, the same cannot be said for dispersive models of internal waves. For example, there is the model proposed by Lynett and Liu [81] that in non-dimensional form reads

$$\eta_t + \nabla \cdot \mathbf{M} = 0 \,, \tag{2.235}$$

$$\begin{aligned}
\mathbf{M}_t &+ \left(\frac{1}{h_1}+\frac{1}{h_2}\right)^{-1} \nabla\eta + \epsilon\left(\frac{1}{h_1}-\frac{1}{h_2}\right)\left[\mathbf{M}\cdot\nabla\mathbf{M} - (\mathbf{M}\eta)_t\right] \\
&- \epsilon\frac{1}{h_1^3}\left(\frac{1}{h_1}+\frac{1}{h_2}\right)^{-1}\left[(\mathbf{M}\cdot\nabla h_2)\mathbf{M}\right] \\
&+ \epsilon^2\left(\frac{1}{h_1^2}-\frac{1}{h_2 h_1}+\frac{1}{h_2^2}\right)\left[(\mathbf{M}\eta^2)_t - (\mathbf{M}\cdot\nabla)(\eta\mathbf{M}) - \eta(\mathbf{M}\cdot\nabla)\mathbf{M}\right] + \mu^2\frac{h_1 h_2}{3}\nabla\left(\nabla\cdot\mathbf{M}_t\right) \\
&- \mu^2\left(\frac{1}{h_1}+\frac{1}{h_2}\right)^{-1}\left[(\nabla\cdot\mathbf{M}_t)\nabla h_2 - \frac{2}{h_2}\left(\mathbf{M}_t\cdot\nabla h_2\right)\nabla h_2 + \nabla(\mathbf{M}_t\cdot\nabla h_2)\right] = 0 \,,
\end{aligned} \tag{2.236}$$

and

$$\mathbf{M} = (h_1 - \epsilon\eta)\mathbf{U}_1 = -(h_2 + \epsilon\eta)\mathbf{U}_2 \ . \tag{2.238}$$

Here $\mathbf{U}_1$ and $\mathbf{U}_2$ are the depth-averaged velocities in the upper and lower layer respectively. h_1 and h_2 are the non-dimensional undisturbed layer thicknesses that have been scaled by the equivalent depth $h_0 = H_1H_2/(H_1 + H_2)$. The nonlinearity parameter is $\epsilon = a_0/h_0$, where a_0 is a characteristic wave amplitude and $\mu = h_0/L$ is the aspect ratio.

While Lynett and Liu [81] have successfully obtained numerical solutions to equations (2.235)–(2.237) in two spatial dimensions using a predictor-corrector type finite difference formulation, the equations are somewhat controversial in the literature since Bai and Song [6] have suggested there are several misprints in Lynett and Liu's model equations that need to be corrected. Other two-layer dispersive and nonlinear models have been derived by Sebsarma *et al.* [29] and Ye [134]. In future work, the solution procedures for two-layer equations presented in Chapters 3 and 5 may be extended to one of these models. Further research must be conducted before deciding on a set of two-layer equations that offers a middle ground between accuracy and practicality.

2.13 Chapter summary and discussion

In this chapter the full Navier–Stokes equations were introduced alongside the Boussinesq approximation that is popular for flows with small density changes. The concept of a layered model as a common simplification was introduced, and the method of normal modes was used to analyze the dispersive behaviour of the waves associated with these models of varying complexity. The idea of using a dispersive shallow water model as a simplified model for internal waves was proposed, and energy diagnostics were carried out on these dispersive layered models. Hyperbolic equation theory was applied to both the single layer and two layer non-dispersive shallow water models, and it was demonstrated that the two layer model breaks down (loses hyperbolicity) in regions where the magnitude of vertical shear exceeds the speed of long internal gravity waves. Additionally, the fact that the two layer system does not comprise a set of conservation laws (as the single layer system does) turns out to be a limitation in the formulation of numerical methods as will be discussed in the chapters that follow. As a final remark, more exotic and complicated two-dimensional models compared to those studied in this thesis were introduced. These more accurate models represent a possible avenue for future work.

Chapter 3

Pseudospectral Methods for Layered Models in Simple Geometries

In this chapter, we consider the numerical solution of the single-layer weakly non-hydrostatic (Boussinesq-type) shallow water system (2.120)–(2.122) and the two-layer extension (2.196)–(2.200) in idealized lake geometries. The geometry choices considered here are: a doubly periodic domain, an annular (circular) geometry as an idealized closed basin, and a periodic channel.

Given that the domains are simple in geometric structure, and that we are interested in resolving small-scale internal wave phenomena, the Fourier and Chebyshev pseudospectral spatial discretization methods are a clear choice due to the fact that they give the highest order of accuracy possible for a fixed number of grid points, have excellent resolution characteristics, and have small amounts of inherent dissipation (see, for example, [14]).

In the following sections, we described temporal and spatial discretization methods separately, as is appropriate in the context of the *method of lines* [122]. Practical implementation details, such as numerical linear algebra techniques (e.g., iterative solvers and preconditioning), and spectral filtering are explained as well. Validation of the numerical methodology is carried out by grid convergence studies and, where possible, comparisons to analytical solutions.

The resulting wave dynamical models are applied in a variety of simulations relevant to internal wave dynamics in small- to mid-sized lakes that are strongly influenced by stratification during the ice-free seasons, and rotation (especially for the mid-sized lakes). The results presented here give some insights into what conditions are necessary for the

emergence of small-scale internal wave phenomena via nonlinear steepening of long (or basin-scale) waves.

The present work suggests that scalable Fast Fourier Transform (FFT) based methods for water wave equations can be extended to physical cases involving non-constant bathymetry and can also be an effective tool for solving elliptic problems with non-constant coefficients provided they are used alongside an appropriate iterative linear solver with pre-conditioning. Given the highly accurate nature of pseudospectral methods, the results presented here may be seen as a benchmark for lower-order spatial discretization techniques such as DG-FEM and FVM, and allow for rational hypotheses to be formulated for subsequent testing against field data.

3.1 Time-stepping techniques

For the moment, we will assume that we have spatially discretized the system (2.120)–(2.122) using a method of lines approach as discussed by [122]. That is, the flow variables of interest (h, u, v) have been discretized on N grid-points and are now represented by the $N \times 1$ vectors $(\mathbf{h}, \mathbf{u}, \mathbf{v}) = ([h_1, \cdots, h_N]^\mathsf{T}, [u_1, \cdots, u_N]^\mathsf{T}, [v_1, \cdots, v_N]^\mathsf{T})$, where we adopt the notation that bold-faced variables refer to the discretized approximate solution fields of the system (2.120)-(2.122). We further assume that the continuous spatial derivative operators $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial y^2}$, $\frac{\partial^2}{\partial xy}$ have been replaced by the $N \times N$ matrices D_x, D_y, D_{xx}, D_{yy}, D_{xy} or that the required matrix-vector products are attainable by other means, such as the pseudospectral technique [95].

We employ the second-order accurate Leapfrog scheme for the temporal discretization of the model equations that is commonly used in atmospheric/oceanic general circulation models [131, 2]. Although it is only second-order accurate, Leapfrog offers benefits in the form of requiring less memory than the corresponding linear multi-step methods (i.e., Adams–Bashforth) and fewer computations than a multi-stage Runge-Kutta method. Although the Leapfrog method is typically known to require the Robert–Asselin filter for stability in finite difference schemes [35], we have found that the Leapfrog computational mode is not a source of difficulty for the high-order methods discussed here. We believe this result is due to the spatial filtering techniques we employ that are discussed in Section 3.2.3.

To keep the discussion as general as possible, we do not specify which spatial discretization scheme we are using since the following time-stepping schemes may be applied to a number of spatial discretization methods including finite difference methods, the Fourier pseudospectral method, the Chebyshev spectral collocation method, and DG-FEM [122].

Upon applying the method of lines to the system (2.120)-(2.122), we recover the semi-discrete system of equations

$$\frac{d\mathbf{h}}{dt} = -D_x(\mathbf{uh}) - D_y(\mathbf{vh}) \,, \tag{3.1}$$

$$\begin{aligned} \frac{d(\mathbf{uh})}{dt} - \frac{\mathbf{H}^2}{6}\frac{d}{dt}\left(D_{xx}(\mathbf{uh}) + D_{xy}(\mathbf{vh})\right) &= -D_x(\mathbf{uuh}) - D_y(\mathbf{uvh}) \\ &\quad -g\mathbf{h}D_x\boldsymbol{\eta} + f\mathbf{vh} \,, \end{aligned} \tag{3.2}$$

$$\begin{aligned} \frac{d(\mathbf{vh})}{dt} - \frac{\mathbf{H}^2}{6}\frac{d}{dt}\left(D_{xy}(\mathbf{uh}) + D_{yy}(\mathbf{vh})\right) &= -D_x(\mathbf{vuh}) - D_y(\mathbf{vvh}) \\ &\quad -g\mathbf{h}D_y\boldsymbol{\eta} - f\mathbf{uh} \,, \end{aligned} \tag{3.3}$$

where we have regrouped terms for later convenience. For notational brevity, we adopt the convention that vector products of the form $\mathbf{ab}$ refer to the Schur product, i.e.,

$$\mathbf{ab} = [a_1 b_1, \cdots, a_N b_N]^{\mathsf{T}} \,.$$

The question that remains is "how to choose the time-discretization to allow for a stable and efficient scheme?". The most obvious choice is to apply the same numerical ODE integrator to all instances of $\frac{d}{dt}$ in equations (3.1)-(3.2). If we discretize the flow variables $(\mathbf{h}, \mathbf{uh}, \mathbf{vh})$ at the time levels

$$t_n = n\Delta t, \quad k = 0, 1, \cdots \,, \tag{3.4}$$

where Δt represents the time-step, and we adopt the notation that superscript n denotes the n^{th} time-step. Applying the Leapfrog formula to equations (3.1)-(3.3) results in the scheme

$$\mathbf{h}^{n+1} = \mathbf{h}^{n-1} + 2\Delta t(-D_x(\mathbf{uh})^n - D_y(\mathbf{vh})^n) \,, \tag{3.5}$$

$$\begin{pmatrix} I - \frac{\mathcal{H}^2}{6}D_{xx} & -\frac{\mathcal{H}^2}{6}D_{xy} \\ -\frac{\mathcal{H}^2}{6}D_{xy} & I - \frac{\mathcal{H}^2}{6}D_{yy} \end{pmatrix} \begin{pmatrix} (\mathbf{uh})^{n+1} \\ (\mathbf{vh})^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{RHS}_1^{n,n-1} \\ \mathbf{RHS}_2^{n,n-1} \end{pmatrix} \,, \tag{3.6}$$

where

$$\begin{aligned} \mathbf{RHS}_1^{n,n-1} &= (\mathbf{uh})^{n-1} - \frac{\mathbf{H}^2}{6}D_{xx}(\mathbf{uh})^{n-1} - \frac{\mathbf{H}^2}{6}D_{xy}(\mathbf{vh})^{n-1} \\ &\quad +2\Delta t(-D_x(\mathbf{uuh})^n - D_y(\mathbf{uvh})^n - g\mathbf{h}^n D_x\boldsymbol{\eta}^n + f(\mathbf{vh})^n) \,, \end{aligned} \tag{3.7}$$

$$\begin{aligned} \mathbf{RHS}_2^{n,n-1} &= (\mathbf{vh})^{n-1} - \frac{\mathbf{H}^2}{6}D_{xy}(\mathbf{uh})^{n-1} - \frac{\mathbf{H}^2}{6}D_{yy}(\mathbf{vh})^{n-1} \\ &\quad +2\Delta t(-D_x(\mathbf{vuh})^n - D_y(\mathbf{vvh})^n - g\mathbf{h}^n D_y\boldsymbol{\eta}^n - f(\mathbf{uh})^n) \,, \end{aligned} \tag{3.8}$$

$\mathcal{H}_{ij} = \text{diag}H_i$ is the $N \times N$ matrix with the entries of $\mathbf{H} = [H_1, \cdots, H_N]^T$ along its diagonal, and I is the $N \times N$ identity matrix. Due to the coupled nature of the semi-discrete momentum equations (3.2)-(3.3), a block matrix of size $2N \times 2N$ appears in the scheme despite our choice of an explicit numerical ODE integrator. An approach for reducing the dimension of the required linear system by a factor of 2 is discussed below.

3.1.1 The Scalar Approach

Although there is nothing wrong with the scheme represented by (3.5)-(3.6), it is desirable to find an alternative scheme that involves solving a smaller linear system of equations, if possible. Such a scheme can be obtained by adding an auxiliary elliptic equation to the Boussinesq system. The resulting linear system is $N \times N$. This was demonstrated by Eskilsson and Sherwin [42] where the DG-FEM method was used to solve the equations of Peregrine [93] that are similar to the system (2.120)-(2.122).

The approach begins by introducing the scalar variable

$$z = \nabla \cdot (\mathbf{u}h)_t \,, \tag{3.9}$$

which represents the time rate of change of momentum divergence. If we then take the divergence of the vector form of the momentum equations (2.121)-(2.122), we arrive at the elliptic equation

$$\nabla \cdot \left(\frac{H^2}{6} \nabla z \right) - z = -\nabla \cdot \mathbf{a} \,, \tag{3.10}$$

that is referred to as a *wave continuity* equation in [42]. The vector $\mathbf{a} = (a_1, a_2)^T$ is given by the flux terms in equation (2.121)-(2.122), i.e.,

$$\mathbf{a} = \begin{pmatrix} -\nabla \cdot ((uh)\mathbf{u}) - gh\eta_x + fvh \\ -\nabla \cdot ((vh)\mathbf{u}) - gh\eta_y - fuh \end{pmatrix} . \tag{3.11}$$

Applying the method of lines to the augmented system represented by equations (2.120)-

(2.122) and (3.10) gives the semi-discrete equations

$$\frac{d\mathbf{h}}{dt} = -D_x(\mathbf{uh}) - D_y(\mathbf{vh}) \,, \tag{3.12}$$

$$\frac{d(\mathbf{uh})}{dt} = -D_x(\mathbf{uuh}) - D_y(\mathbf{uvh}) - g\mathbf{h}D_x\boldsymbol{\eta} + f\mathbf{vh} + \frac{\mathbf{H}^2}{6}D_x\mathbf{z} \,, \tag{3.13}$$

$$\frac{d(\mathbf{vh})}{dt} = -D_y(\mathbf{vuh}) - D_y(\mathbf{vvh}) - g\mathbf{h}D_y\boldsymbol{\eta} - f\mathbf{uh} + \frac{\mathbf{H}^2}{6}D_y\mathbf{z} \,, \tag{3.14}$$

$$\begin{aligned} \frac{\mathbf{H}^2}{6}(D_{xx}\mathbf{z} &+ D_{yy}\mathbf{z}) - \mathbf{z} \\ &+ \frac{1}{6}\left(D_x(\mathbf{H}^2)D_x\mathbf{z} + D_y(\mathbf{H}^2)D_y\mathbf{z}\right) = -(D_x\mathbf{a_1} + D_y\mathbf{a_2}), \end{aligned} \tag{3.15}$$

where we have first applied the product rule to equation (3.10) in arriving at (3.15). The left-hand side of equation (3.15) may be factored to resemble a linear system of equations of the form

$$\mathcal{A}\mathbf{z} = \mathbf{b} \,, \tag{3.16}$$

with

$$\mathcal{A} = \frac{\mathcal{H}^2}{6}(D_{xx} + D_{yy}) - I + \frac{1}{6}\left(D_x(\mathcal{H}^2)D_x + D_y(\mathcal{H}^2)D_y\right) \,, \tag{3.17}$$

$$\mathbf{b} = -(D_x\mathbf{a_1} + D_y\mathbf{a_2}) \,, \tag{3.18}$$

where $\mathcal{H}_{ij} = \mathrm{diag}H_i$. We can then obtain an appropriate numerical scheme by applying the Leapfrog formula to equations (3.12)-(3.14) and using time-splitting so that the equation for $\mathbf{z}$ may be inverted using the most recent information available. The resulting scheme at each time-step is

$$\mathbf{h}^{n+1} = \mathbf{h}^{n-1} + 2\Delta t(-D_x(\mathbf{uh})^n - D_y(\mathbf{vh})^n) \,, \tag{3.19}$$

$$(\mathbf{uh})^\dagger = (\mathbf{uh})^{n-1} + 2\Delta t{\mathbf{a_1}}^n \,, \tag{3.20}$$

$$(\mathbf{vh})^\dagger = (\mathbf{vh})^{n-1} + 2\Delta t{\mathbf{a_2}}^n \,, \tag{3.21}$$

$$\mathbf{z}^\dagger = \mathcal{A}^{-1}\mathbf{b}^\dagger \,, \tag{3.22}$$

$$(\mathbf{uh})^{n+1} = (\mathbf{uh})^\dagger + 2\Delta t\frac{\mathbf{H}^2}{6}D_x\mathbf{z}^\dagger \,, \tag{3.23}$$

$$(\mathbf{vh})^{n+1} = (\mathbf{vh})^\dagger + 2\Delta t\frac{\mathbf{H}^2}{6}D_y\mathbf{z}^\dagger \,, \tag{3.24}$$

where $\mathbf{b}^\dagger$ is the vector $\mathbf{b}$ evaluated using $(\mathbf{uh})^\dagger$, $(\mathbf{vh})^\dagger$, and $\mathbf{h}^{n+1}$. An alternative method that requires fewer computations at the cost of slightly worse accuracy is to compute $\mathbf{z}^\dagger$ first

using only information from the n^{th} time-step, and then to compute $(\mathbf{h}^{n+1}, (\mathbf{uh})^{n+1}, (\mathbf{vh})^{n+1})$ without time-splitting. Our numerical experiments revealed negligible differences between the two methods.

The most expensive part of the algorithm is step (3.22), in which we solve the linear system $\mathcal{A}\mathbf{z} = \mathbf{b}$. For pseudospectral methods, the matrix $\mathcal{A}$ is dense, and due to memory restrictions, direct methods such as LU-factorizations become impractical at high resolutions [14]. To overcome this issue, it is necessary to consider iterative methods such as the generalized minimum residual method (GMRES) and pre-conditioning to reduce the required number of iterations. In Section 3.2.2, we illustrate how to construct a suitable pre-conditioner using a finite difference approximation.

The schemes presented above are not self-starting. Therefore, they must be started by taking either a single time-step with the first order accurate Forward Euler method or a higher order Runge-Kutta method.

3.2 Fourier method for doubly-periodic domains

We now present the Fourier spatial discretization method applied to the scheme represented by equations (3.19)-(3.24). We begin by discretizing the periodic rectangular domain $\Omega = [0, L_x] \times [0, L_y]$ by constructing a tensor-product grid from the one-dimensional equidistant grids

$$x_i = i\Delta x, \quad i = 0, \cdots, N_x - 1 \,, \tag{3.25}$$

$$y_j = j\Delta y, \quad j = 0, \cdots, N_y - 1 \,, \tag{3.26}$$

where $\Delta x = L_x/N_x$ and $\Delta y = L_y/N_y$ represent the grid spacing in the x and y directions, respectively. The resulting two-dimensional grid then has $N = N_x N_y$ total grid points. It is also useful to define the discrete wavenumber vectors $\mathbf{k}$ and $\mathbf{l}$ as

$$k_i = \frac{2\pi}{L_x} i, \quad i = -\frac{N_x}{2} + 1, \cdots, \frac{N_x}{2} \,, \tag{3.27}$$

$$l_j = \frac{2\pi}{L_y} j, \quad j = -\frac{N_y}{2} + 1, \cdots, \frac{N_y}{2} \,, \tag{3.28}$$

Given these definitions, it is useful to take note that $\pi/\Delta x$ and $\pi/\Delta y$ are the shortest wavelengths resolvable in the x- and y-directions, respectively.

Rather than using differentiation matrices to compute the approximate derivatives in the schemes presented above, we employ the "pseudospectral technique" as described by

Peyret [95]. That is, differentiation is performed in spectral space (the space of the Fourier coefficients) with the fast discrete Fourier transform (FFT) while products are performed in physical space. Doing so allows one to avoid the expense of directly computing convolution sums in the space of the Fourier coefficients, as the nonlinear terms would require. Pseudospectral differentiation is also faster than explicitly calculating matrix-vector products that require $O(N^2)$ floating-point operations (FLOPS) since the FFT requires $O(N \log N)$ FLOPS, and Schur products requires $O(N)$ FLOPS [95].

For the purposes of pseudospectral differentiation, it is useful to consider the flow fields as $N_y \times N_x$ matrices instead of $N_x N_y \times 1$ vectors. For a given discretized field $\boldsymbol{\phi}$ which may represent a flow variable or a product of flow variables, we approximate its discrete derivatives as

$$\boldsymbol{\phi}_x = \mathcal{F}_x^{-1}\left(\mathfrak{i}\mathcal{K}\mathcal{F}_x(\boldsymbol{\phi})\right) , \tag{3.29}$$

$$\boldsymbol{\phi}_y = \mathcal{F}_y^{-1}\left(\mathfrak{i}\mathcal{L}\mathcal{F}_y(\boldsymbol{\phi})\right) , \tag{3.30}$$

$$\boldsymbol{\phi}_{xx} = \mathcal{F}_x^{-1}\left(-\mathcal{K}^2\mathcal{F}_x(\boldsymbol{\phi})\right) , \tag{3.31}$$

$$\boldsymbol{\phi}_{yy} = \mathcal{F}_y^{-1}\left(-\mathcal{L}^2\mathcal{F}_y(\boldsymbol{\phi})\right) , \tag{3.32}$$

$$\boldsymbol{\phi}_{xy} = \mathcal{F}_y^{-1}\left(\mathfrak{i}\mathcal{L}\mathcal{F}_y(\boldsymbol{\phi}_x)\right) , \tag{3.33}$$

where $\mathcal{F}_x$ and $\mathcal{F}_y$ represent the discrete Fourier transforms with respect to x and y, respectively, $\mathfrak{i} = \sqrt{-1}$, and superscript -1 refers to the inverse transform. The wavenumber matrices $\mathcal{K}$ and $\mathcal{L}$ are of size $N_y \times N_x$ with entries $\mathcal{K}_{ij} = k_j$, $\mathcal{L}_{ij} = l_i$. All of the products in (3.29)-(3.33) are Schur products.

The underlying assumptions used in this spatial discretization are that the solution fields are smooth and periodic in space in both directions, and hence that they are well represented by a sinusoidal basis. Given these assumptions, the Fourier pseudospectral spatial discretization method guarantees an exponential convergence rate [14]. If one or both of these assumptions are broken, Gibbs oscillations are introduced into the solution and the convergence rate is reduced to polynomial order.

For a one dimensional function, $f(x)$ being discretely approximated using the Fourier method by $\mathbf{f}$, we also have the continuous approximation which we denote $f_h(x)$

$$f(x) \approx f_h(x) = \sum_{k=0}^{N_x-1} \hat{f}_k \mathrm{e}^{\frac{2\pi \mathfrak{i} k x}{L_x}} , \tag{3.34}$$

where the $\hat{f}_k$'s are the Fourier coefficients and can be obtained by the FFT as a faster method than direct quadrature integration. The continuous approximation $f_h(x)$ is useful

if the solution needs to be interpolated to values of x that are not contained on the computational grid. This idea may be extended to the two-dimensional case by considering a double-series expansion and the two-dimensional FFT.

3.2.1 Solving the Linear System

In order to solve the linear system (3.16), one may be tempted to explicitly build the large matrix $\mathcal{A}$ using two-dimensional spectral differentiation matrices. However, this is typically not a good idea due to memory restrictions. Two-dimensional spectral differentiation matrices can be built from Kronecker products between the 1D differentiation matrix and the appropriate identity matrix, with a memory requirement that is $O(N_x N_y (N_x + N_y))$. If mixed spatial derivatives are required, the situation can be the worst case, namely an $O(N_x^2 N_y^2)$ memory requirement, which is certainly not reasonable. It is clear that iterative methods for solving the system (3.16) are required in the case of a pseudospectral spatial discretization.

In doubly-periodic cases with a flat bottom, the mean depth H is a constant and the linear system (3.16) may be solved efficiently as explained by Peyret [95] using the pseudospectral technique by first taking its Fourier transform, yielding

$$\widehat{\mathcal{A}} \mathcal{F}_{xy}(\mathbf{z}) = \mathcal{F}_{xy}(\mathbf{b}) \ , \tag{3.35}$$

where

$$\widehat{\mathcal{A}} = \left(-\frac{H^2}{6} \left(\mathcal{K}^2 + \mathcal{L}^2 \right) - \mathbf{1} \right) \ , \tag{3.36}$$

$\mathbf{1}$ is the $N_y \times N_x$ matrix of all ones, and $\mathcal{F}_{xy}$ is the double discrete Fourier transform. To solve the system, we take the Schur product of both sides with $\widehat{\mathcal{A}}^{-1}$, defined as

$$\widehat{\mathcal{A}}_{ij}^{-1} = \frac{1}{\widehat{\mathcal{A}}_{ij}} \ , \tag{3.37}$$

the multiplicative inverse of $\widehat{\mathcal{A}}$. Hence,

$$\mathbf{z} = \mathcal{F}_{xy}^{-1} \left(\widehat{\mathcal{A}}^{-1} \mathcal{F}_{xy}(\mathbf{b}) \right) \ . \tag{3.38}$$

This situation is ideal, since we are able to effectively solve a large, dense linear system with $O(N_x N_y (N_x + N_y))$ entries in $O(N_x N_y \log(N_x N_y))$ FLOPS. We have successfully exploited the fact that since the equation is linear with constant coefficients, the Fourier modes are

not coupled and the problem in Fourier space reduces to a single algebraic equation that can be inverted for each Fourier coefficient. In cases where the bottom is not flat, the technique represented by (3.35)-(3.36) is not available since point-wise products become convolutions in Fourier space and the modes become coupled, so a different method must be sought.

Iterative linear system solutions appear to be our only course of action in the case of variable depth. Since the Krylov subspace methods do not explicitly require the entries of the matrix $\mathcal{A}$ [52], they are a clear choice. Furthermore, given that the matrices being solved are not guaranteed to be symmetric nor skew-symmetric [122], a good choice of iterative linear solver is the generalized minimum residual method (GMRES) [123].

The main difficulty with using such iterative solvers, is that the linear systems to be solved can be quite poorly conditioned, driving the number of iterations up, and indeed to the same order as the problem's dimension. This issue typically gets worse at higher resolutions [14]. To overcome this, it is useful to pre-condition the linear system to obtain convergence at a relatively small number of iterations as discussed below.

3.2.2 Finite Differences Pre-Conditioner

Since the linear system to be solved is the result of a high-order PDE spatial discretization, a popular and effective choice of pre-conditioner is a low-order spatial discretization of the PDE [123]. The underlying assumption is that the low-order approximation to the operator is a reasonable estimate of the high-order approximation to the operator.

A finite differences discretization is a natural choice since it allows one to fix the order of approximation independently of the number of grid points and the grid-spacing used [76]. The resulting spatial-discretization operators are typically very sparse and banded, and as a result can be solved or factored quite easily using sparse matrix manipulation software libraries.

To construct a pre-conditioner for solving the linear system (3.16), we employ the second-order centered differences formulas given by Leveque [76] to construct the $N \times N$ differentiation matrices $D_x^{(2)}$, $D_y^{(2)}$, $D_{xx}^{(2)}$, $D_{yy}^{(2)}$, where superscript "(2)" refers to the order of approximation used. A second order approximation to the matrix $\mathcal{A}$, denoted $\mathcal{A}^{(2)}$, can then be constructed using the formula (3.17). The resulting matrix is pentadiagonal, requiring $O(5N)$ memory since its construction relies on the 5-point finite difference stencil for the Laplacian [63].

Since $\mathcal{A}^{(2)}$ is an approximation of $\mathcal{A}$, we can imagine left-multiplying (3.16) by $(\mathcal{A}^{(2)})^{-1}$

$$(\mathcal{A}^{(2)})^{-1}\mathcal{A}\mathbf{z} = (\mathcal{A}^{(2)})^{-1}\mathbf{b}\ , \tag{3.39}$$

to obtain a more well-conditioned linear system since $(\mathcal{A}^{(2)})^{-1}\mathcal{A} \approx I$. Of course, this is merely illustrative since $\mathcal{A}$ is not explicitly built and computing the explicit inverse of $\mathcal{A}^{(2)}$ is impractical. Instead, the fact that we are using the GMRES method requires that linear systems of the form $\mathcal{A}^{(2)}\tilde{\mathbf{z}} = \tilde{\mathbf{b}}$ be solved at each iteration. In order to ensure linear systems of this form can be solved effectively, it is useful to compute the LU-factorization of $\mathcal{A}^{(2)}$ in the pre-processing stage and to simply reuse its factors at each GMRES iteration.

It has been found that using the factors returned by the sparse-LU factorization routine provided in the UMFPACK library yields very fast convergence of solutions to $\mathcal{A}^{(2)}\tilde{\mathbf{z}} = \tilde{\mathbf{b}}$. In addition to the lower- L and upper-triangular U factors, partial pivoting is also performed with a permutation matrix P and column-reordering matrix Q so that

$$P\mathcal{A}^{(2)}Q = LU\ . \tag{3.40}$$

The main cost of using this technique is in storing the factors L and U, which in the worst case, can be the same cost as storing a full $N \times N$ matrix. At high resolutions, storing the factors may become unfeasible, and incomplete LU-factorizations may be used instead with a drop-tolerance tuned to give a balance between memory usage and iteration count. At even higher resolutions, such a balance may not exist, and more memory efficient techniques such as geometric multigrid [124] or multi-level domain decomposition algorithms [108] should be considered. The geometric multigrid pre-conditioner has recently been used successfully in the context of pseudospectral methods for the three-dimensional incompressible Navier–Stokes equations by Subich [117, 118].

3.2.3 Filter Stabilization of Aliasing-driven Instabilities

The equations do not possess any viscosity terms and thus lack any physical energy dissipation mechanism. As a result, the quadratic nonlinearity terms can cause energy to accumulate at the small scales in an unphysical manner. Additionally, aliasing errors that occur due to the "pointwise product" treatment of the nonlinear terms can drive weak numerical instabilities that can destroy the numerical solutions [60]. An example of an aliasing-driven instability (or blow-up) occurs when a short travelling wave is misrepresented on a coarse grid as a longer growing (hence, unstable) wave.

The process of aliasing can be illustrated as follows. Consider two polynomials, $f(x)$ and $g(x)$, of degree N. By uniqueness, these polynomial can be completely recovered from their

discretizations $\mathbf{f}$ and $\mathbf{g}$ on $N+1$ grid points using the Lagrange interpolating polynomials. Now consider the product $f(x)g(x)$. Analytically, this product is a polynomial of degree $2N$. However, in the methods presented above, this product is approximated by the Schur product $\mathbf{fg}$ represented on the original $N+1$ grid points. Since $N+1$ points is not enough to uniquely specify a polynomial of degree $2N$, an aliasing error ensues.

There are methods for completely removing aliasing, thereby preventing aliasing-driven instabilities. One method is to simply evaluate the convolution sum in spectral space, rather than compute the product in physical space. A popular method for dealiasing quadratic nonlinearities, referred to as Orszag's 3/2's rule [95], is to extend (zero-pad) the spectrum by 3/2's its original size, perform the product in physical space, and then to truncate the extended spectrum back to its original size. Although these methods do prevent weak instabilities due to aliasing, they can be computationally expensive, and they do not, in and of themselves, preclude situations in which small-scale noise persists in the numerical solution, eventually destroying the physical relevance of the solution.

In light of these issues, filtering is taken as a procedure to dissipate energy as it accumulates at the small scales thereby preventing aliasing errors from driving weak instabilities. This can be achieved by applying a low-pass wavenumber filter of the form

$$\sigma(k) = \begin{cases} 1, & 0 \leq k < k_{crit} \\ \exp\left(-\alpha\left(\frac{k-k_{crit}}{k_{max}-k_{crit}}\right)^s\right), & k_{crit} \leq k \leq k_{max} \end{cases} \tag{3.41}$$

in each direction in spectral space to the solution fields after each time-step. A similar filter is used by Hesthaven and Warburton [60] in the nodal DG-FEM framework. Typical parameters used in the simulations presented in this section are $k_{crit} = 0.65k_{max}$, $s = 4$, $\alpha = 18.4$. The parameters α, s, and k_{crit} are tunable and, in general, their values must be determined through experimentation. As a rule of thumb, we can only refer the reader to the Ockamian proverb:

> "Filter as little as possible
> ... but as much as is needed." [60]

One might be tempted to reject the approach of spectral filtering in favour of a low order finite difference method that does not need filtering since the approach represents an *ad hoc* solution to the weak instabilities problem. However, one should recall that, by Taylor series, a first-order upwind finite difference aproximation to $\frac{df}{dx}$ has truncation error of $\frac{d^2f}{dx^2}\frac{\Delta x}{2} + O(\Delta x^2)$. Thus, using such a low order method inherently adds an undetermined amount of unphysical second order diffusion to the numerical solution (which prevents aliasing-driven

instabilities). Therefore, the filtering answer to the weak instabilities problem should be embraced, not rejected, since it allows one to directly control the amount of numerical dissipation used, rather than have it be unknown.

3.2.4 Pre-conditioner performance test

To confirm the performance of our pre-conditioned GMRES algorithm for solving equation (3.15), we have compared it to the GMRES method without pre-conditioning. The problem we consider for inversion here corresponds to the first time-step of the full simulation presented below in Section 3.2.10. Hence, the variable coefficients on the left-hand side as well as the function on the right-hand side of (3.15) are both non-trivial. The convergence criterion was taken to require the magnitude of the relative residual to be below 10^{-9}, and this value was also used for all simulations with a non-flat bottom, i.e., whenever GMRES was used. The problem was solved at a variety of grid resolutions, ranging from 16×16 to 1024×1024.

Iteration counts for both the 'GMRES with pre-conditioning' (GMRESP) and 'GMRES without pre-conditioning' (GMRESNP) methods along with the ratio of their run-times are listed at all resolutions considered in Table 3.1. The corresponding tests were carried out in Matlab, and it was found that the values for the run-times fluctuated somewhat due to the performance of Matlab's built-in parallelization's dependence on processor load. Hence, the run-time values used in Table 3.1 were averaged over ten runs to average out these fluctuations.

Table 3.1 shows the pre-conditioner's ability to keep the iteration count relatively low in comparison to the case where pre-conditioning is not used. At low resolutions (128×128 and below), the reduction in iteration count does not overcome the computational cost of using a pre-conditioner since the run-time ratio is less than one. However, at higher resolutions the savings are considerable, and the high iteration count at high resolutions makes the 'GMRESNP' method impractical for use in simulations due to the unreasonable amount of computational time required. For example, the GMRESNP method at 1024×1024 resolution typically took about 280 s to converge. Having large iteration counts with GMRES is known to be problematic for the additional reason that the amount of work for the N^{th} GMRES iteration scales like $O(N^2)$ [123].

	GMRESP	GMRESNP	
Resolution	Iteration Count	Iteration Count	Run-Time Ratio
16×16	5	5	0.71
32×32	6	7	0.77
64×64	9	10	0.78
128×128	9	15	0.94
256×256	12	32	1.75
512×512	16	70	4.27
1024×1024	19	147	16.2

Table 3.1: Iteration count vs. grid resolution for the 'GMRES with pre-conditioning' (GMRESP) and 'GMRES without pre-conditioning' (GMRESNP) methods. At each resolution, the run-time ratio is given by the time taken for the GMRESNP method to converge divided by time taken for the GMRESP method to converge.

3.2.5 Convergence test of the Fourier method

To verify that the Fourier spatial discretization method is giving the desired exponential convergence rate, we have performed a convergence study for the Helmholtz problem (3.10) in two dimensions, where we have chosen the form of the exact solution and variable coefficient *a priori*, and the right-hand side function was calculated analytically from the known functions. The problem we consider here is

$$\nabla \cdot (\alpha \nabla z) - z = f \ , \tag{3.42}$$

on the periodic square $\Omega = [0, 2\pi] \times [0, 2\pi]$, where

$$\alpha(x, y) = 2 + \sin(x)\cos(y) \ , \tag{3.43}$$

$$\begin{aligned} f(x, y) = & -\mathrm{sech}(10(y-\pi))^3[9\cos(y)\cos(x)\sin(9x)\cosh(10(y-\pi))^2 \\ & -10\sin(y)\sin(x)\cos(9x)\sinh(10(y-\pi))\cosh(10(y-\pi)) \\ & -19\cos(9x)\cos(y)\sin(x)\cosh(10(y-\pi))^2 \\ & -37\cos(9x)\cosh(10(y-\pi))^2 \\ & +200\cos(9x)\cos(y)\sin(x) + 400\cos(9x)] \ , \end{aligned} \tag{3.44}$$

so that the exact solution is given by

$$z(x, y) = \cos(9x)\,\mathrm{sech}(10(y-\pi)) \ . \tag{3.45}$$

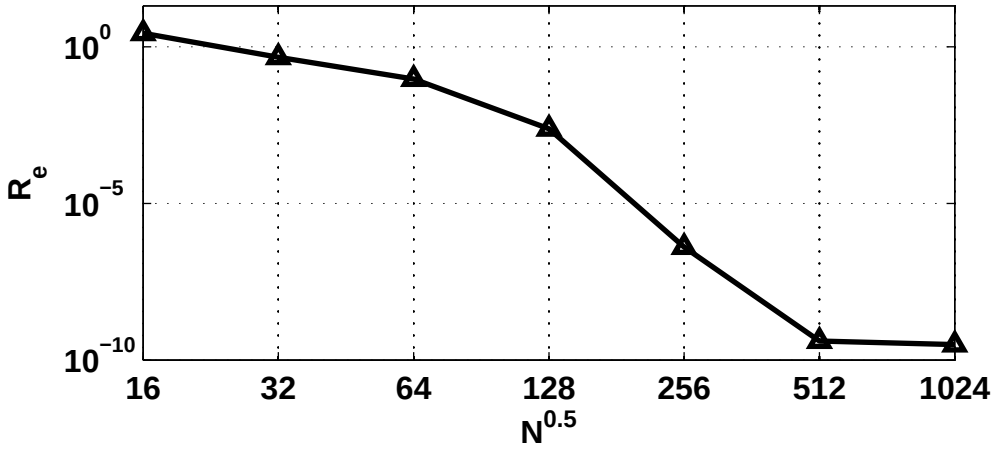

Figure 3.1: Relative error (R_e) between exact and numerical solutions to the 2D Helmholtz problem (3.42) vs. $N^{0.5}$, the square-root of the total number of grid points.

The problem was first solved on a coarse 16×16 grid. The resolution was then doubled in each direction and the problem was solved repeatedly until convergence to numerical precision was reached. In each case, the number of grid points was kept the same in each direction, i.e., $N_x = N_y = N^{0.5}$, where N is the total number of grid points. In Figure 3.1, we plot the relative L^2 error, R_e, vs. $N^{0.5}$, where

$$R_e = \left[\frac{\iint\limits_{\Omega} (z_{Num} - z)^2 \, dS}{\iint\limits_{\Omega} z^2 dS} \right]^{\frac{1}{2}} , \tag{3.46}$$

and z_{Num} is the solution computed numerically via the pre-conditioned GMRES method. The integral in the numerator of equation (3.46) was computed numerically using the Fourier expansion coefficients (obtained with FFT), and the integral in the denominator was calculated exactly.

Figure 3.1 reveals that the Fourier method converges very fast to the exact solution, as expected, and that convergence to within 10^{-9} has been reached at a grid resolution of 512×512.

3.2.6 Comparison of numerical code to approximate analytical solutions with WKB analysis

In the next step towards validating our numerical methodology for the Fourier spatial discretization method, we compared numerical solutions obtained from our numerical code to approximate analytical solutions obtained using the WKB (Wentzel-Kramers-Brillouin) approximation for situations involving variable depth in 1D. The approximation is valid in situations where the depth H varies more slowly in space than the free surface η. Hence, we assume that H depends only on a slow coordinate defined by.

$$\chi = \epsilon x \ , \tag{3.47}$$

where ϵ is a small parameter.

Starting from the linearized 1D governing equations

$$\eta_t = (-Hu)_x \ , \tag{3.48}$$

$$u_t = -g\eta_x + \frac{H^2}{6} u_{xxt} \ , \tag{3.49}$$

and making the substitution $\chi = \epsilon x$, the chain rule gives

$$\eta_t = -\epsilon (H(\chi)u)_\chi \ , \tag{3.50}$$

$$u_t = -\epsilon g \eta_\chi + \epsilon^2 \frac{H(\chi)^2}{6} u_{\chi\chi t} \ , \tag{3.51}$$

After differentiating (3.50) with respect to t, substituting (3.51), and retaining only terms of order ϵ^2 and lower, we have the variable-speed 1D wave equation in terms of η

$$\eta_{tt} - \epsilon^2 \left(gH\eta_\chi \right)_\chi = 0 \ , \tag{3.52}$$

where $H = H(\chi)$ only. It is worth noting that this approximate equation does not contain any dispersive terms such as those included in the full system (2.120)–(2.122), so the approximation is only expected to be accurate for waves that are sufficiently long with respect to the water depth.

The solution, $\eta(\chi, t)$, may then be separated into the product of a sinusoidal time-dependent component and an unknown spatial structure, $\psi(\chi)$, as

$$\eta = \psi(\chi) \mathrm{e}^{-i\sigma t} \ , \tag{3.53}$$

where we are considering waves of a single frequency, σ. The spatial structure of the free surface is then assumed to have the form of the WKB ansatz

$$\psi(\chi) = \mathrm{e}^{i\left(\frac{S_0}{\epsilon}(\chi)+S_1(\chi)+\epsilon S_2(\chi)+\cdots\right)} , \tag{3.54}$$

such that

$$\frac{S_0}{\epsilon} \gg S_1 \gg \epsilon S_2 \gg \cdots , \tag{3.55}$$

$$\epsilon S_2 \ll 1 , \text{ as } \epsilon \to 0 . \tag{3.56}$$

Substituting the ansatz (3.54) into the wave equation (3.52) and solving the resulting problems at orders 1 and ϵ yields the WKB solution

$$S_0(\chi) = \pm \int_0^\chi \frac{\sigma}{\sqrt{gH(\zeta)}} d\zeta , \tag{3.57}$$

$$S_1(\chi) = \frac{i}{2} \ln |H S_0'| = \frac{i}{2} \ln \left| \sigma \sqrt{\frac{H}{g}} \right| , \tag{3.58}$$

where prime ($'$) denotes differentiation with respect to χ. Thus, we have

$$\eta(x,y,t) \sim A(\chi) \mathrm{e}^{i\left(\frac{S_0}{\epsilon}(\epsilon x)-\sigma t\right)} , \text{ as } \epsilon \to 0 , \tag{3.59}$$

where $A(\chi) = a_0 H^{-\frac{1}{4}}$ and a_0 is an arbitrary constant. Since the problem is linear, it is a straight-forward task to show that

$$u(x,y,t) = \sqrt{\frac{g}{H}} \eta(x,y,t) . \tag{3.60}$$

To compare our numerical code with the WKB solution (3.59)–(3.60), we initialized the numerical solver with the real part of the WKB solution (with $S_0 > 0$) taken at $t = 0$, stepped the solution forward in time for five wave periods, and compared the numerical solution to the approximate analytical solution at the final time. We chose the slowly varying depth profile

$$H(\epsilon x) = H_0 - \Delta H \sin(\epsilon x) , \tag{3.61}$$

where $\epsilon = 2\pi/L_x$ is the wavenumber of the longest sinusoidal wave that fits in the domain. Here, $L_x = 3000$ m and $H_0 = 15$ m. We have varied the parameter ΔH from 0–2.5 m, expecting the two solutions to agree best in the limit that $\Delta H \to 0$ (a flat bottom). We set $a_0 = 10^{-4} H_{min}$ in all cases to ensure that nonlinear effects in the numerical solution were

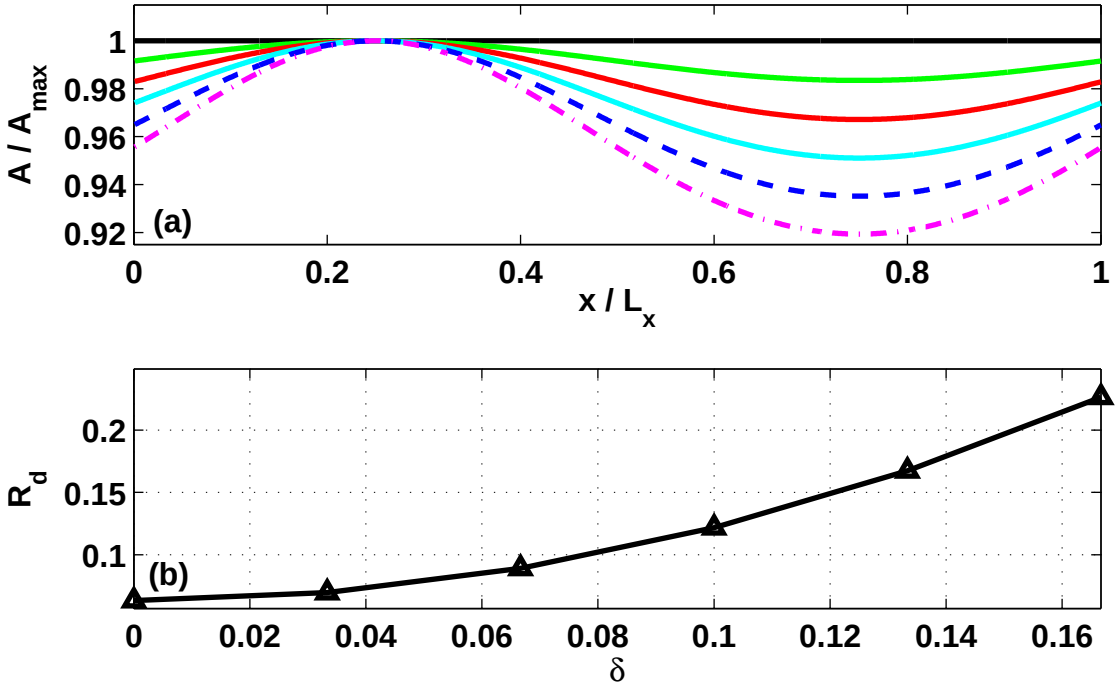

Figure 3.2: Panel **(a)**: Envelopes of the WKB solution, scaled by their maximum value, for the values of $\delta = \Delta H/H_0 = 0$ (solid, black), 1/30 (solid, green), 1/15 (solid, red), 1/10 (solid ,cyan), 2/15 (blue, dashed), 1/6 (magenta, dash-dotted). Panel **(b)**: Relative difference (R_d) between the numerical solution and the WKB solution after five wave periods vs. δ.

negligible. The numerical grid was taken to have 1024 points (grid halving experiments suggest that the simulations are numerically converged upon reaching 256 points), and the time-step was taken to be

$$\Delta t = \frac{1}{20}\frac{\Delta x}{\sqrt{gH_0}}\ , \tag{3.62}$$

where Δx is the uniform grid spacing. The time-step was taken to be smaller than what is typically required for numerical stability. This was done in order to minimize the amount of error introduced during the numerical time-integration process.

The function $S_0(\epsilon x)$ was calculated numerically using quadrature rules for integration, since a closed-form analytical expression is not available for our choice of $H(\epsilon x)$. We chose the value

$$\sigma = \sqrt{gH_0}\left(\frac{10\pi}{L_x}\right)\ , \tag{3.63}$$

for the frequency of the waves. If the bottom is flat, this choice represents the frequency of a sinusoidal wave whose wavelength is a factor of five shorter than the longest wavelength that fits in the domain.

After time-stepping was completed, the relative L^2 difference

$$R_d = \left[\frac{\int_0^{L_x}\left(\eta_{Num}-\eta_{WKB}\right)^2 dx}{\int_0^{L_x}(\eta_{WKB})^2 dx}\right]^{\frac{1}{2}}\ , \tag{3.64}$$

was calculated where η_{Num} and η_{WKB} represent the numerical and WKB η fields, respectively. The integrals were evaluated using the Fourier expansion coefficients of each integrand (obtained with FFT).

Close agreement between the two solutions in the limit that $\Delta H \to 0$ is illustrated in Figure 3.2 where we have introduced the non-dimensional parameter $\delta = \Delta H/H_0$. Panel **(a)** shows the shape of the spatially dependent wave amplitude function, $H^{-\frac{1}{4}}$, for several choices of δ, and panel **(b)** shows the decline in the relative difference between the analytical and numerical solution as $\delta \to 0$. The agreement was found to improve somewhat by increasing the domain length while keeping the depth fixed, but the difference was less than an order of magnitude.

3.2.7 Comparison to a quasi-analytical solutions of dispersive short wave packets

In this section, we consider the ability of the algorithm to numerically capture dispersive behaviour, using a simulation where a packet of short waves of two distinct wavelengths is

released from rest. The domain was taken to be periodic and $L_x = 4000$ m in length, the depth was fixed at $H = 5$ m, and the acceleration due to gravity was taken to be $g = 9.81$ m s^{-2}. The initial condition is

$$\eta(x,0) = \eta_0 \cos(0.15x)\cos(0.05x)\mathrm{e}^{-5\left(\frac{x-0.5L_x}{400}\right)^2} , \tag{3.65}$$

$$u(x,0) = 0 , \tag{3.66}$$

where $\eta_0 = 0.1$ m. The amplitude of the wave packet was chosen to be small enough so that linear wave theory would be a good predictor of the group velocities. Due to dispersion, we expect the longer waves to overtake and lead the shorter waves, after sufficient time has passed, since the linear group velocity of the longer waves is $c_g \approx 9.31$ m s^{-1} while the group velocity of the shorter waves is $c_g \approx 6.14$ m s^{-1}. We use this run to validate our numerical method in the regime where nonlinear effects are negligible and the bottom is flat. Under these simplifying assumptions, the governing equations may be linearized and solved exactly in Fourier space.

The analytical solution procedure begins by considering the linearized one-dimensional form of (2.120)-(2.121)

$$u_t = -g\eta_x + \gamma u_{xxt} , \tag{3.67}$$

$$\eta_t = -Hu_x . \tag{3.68}$$

where $\gamma = H^2/6$ is a constant. Taking the Fourier transform of this system with respect to x gives

$$\hat{u}_t = -gik\hat{\eta} - \gamma k^2 \hat{u}_t , \tag{3.69}$$

$$\hat{\eta}_t = -Hik\hat{u} , \tag{3.70}$$

where hats denote the Fourier transform with respect to x, and k is the wavenumber. Combining these equations to eliminate $\hat{\eta}$ yields

$$\hat{u}_{tt} = \sigma^2(k)\hat{u} , \tag{3.71}$$

where

$$\sigma^2(k) = \frac{gHk^2}{1+\gamma k^2} , \tag{3.72}$$

is the dispersion relation. The solution to (3.71) is

$$\hat{u}(k,t) = A(k)\cos(\sigma t) + B(k)\sin(\sigma t) , \tag{3.73}$$

where $A(k)$ and $B(k)$ are unknown functions of k. Since the wave packet is released from rest, we have as an initial condition $u(x,0) = 0$, i.e., $\hat{u}(k,0) = 0$, which implies that $A(k) \equiv 0$. An expression for $\hat{\eta}$ is obtained by substituting the solution for $\hat{u}(k,t)$ into (3.69). After simplifying, we have

$$\hat{\eta}(k,t) = \frac{ikH}{\sigma} B(k) \cos(\sigma t) \ . \tag{3.74}$$

Applying the initial condition $\hat{\eta}(k,0) = \hat{\eta}_0(k)$ (the Fourier transform of (4.70), we find

$$B(k) = \frac{\sigma}{ikH} \hat{\eta}_0(k) \ . \tag{3.75}$$

In summary, the solution in Fourier space reads

$$\hat{\eta}(k,t) = \hat{\eta}_0(k) \cos(\sigma t) \ , \tag{3.76}$$

$$\hat{u}(k,t) = -\frac{i\sigma}{Hk} \hat{\eta}_0(k) \sin(\sigma t) \ , \tag{3.77}$$

where $\sigma(k)$ is given by the dispersion relation

$$\sigma = \pm \frac{\sqrt{gH}k}{\sqrt{1 + \frac{H^2}{6}k^2}} \ . \tag{3.78}$$

Since we have only obtained the solution in Fourier space, we use the FFT to obtain the solution in physical space for comparison to the numerical solution. Since this last step is numerical, we have deemed this solution to be *quasi-analytical*.

In Figure 4.3, the numerical solution (Fourier in space and Leapfrog in time) is compared to the quasi-analytical solution. Good agreement is shown between the numerical and quasi-analytical solution at the final time $t = 100$ s in panels **(b)**-**(c)**. The zoomed-in view of the wavetrain in panel **(c)** illustrates that there is some phase error between the two solutions that is driven purely by temporal discretization errors, since the Leapfrog method is only a second-order accurate finite difference formula in time.

The time series of the domain-integrated total energy, defined by

$$E(t) = \int_0^{L_x} \frac{1}{2} h u^2 + \frac{1}{2} g \eta^2 \, dx \ , \tag{3.79}$$

is plotted in panel **(c)**. It has been shown in Section 2.6 that the governing equations (2.120)-(2.122) do not conserve energy (as is true of most Boussinesq-type systems), and

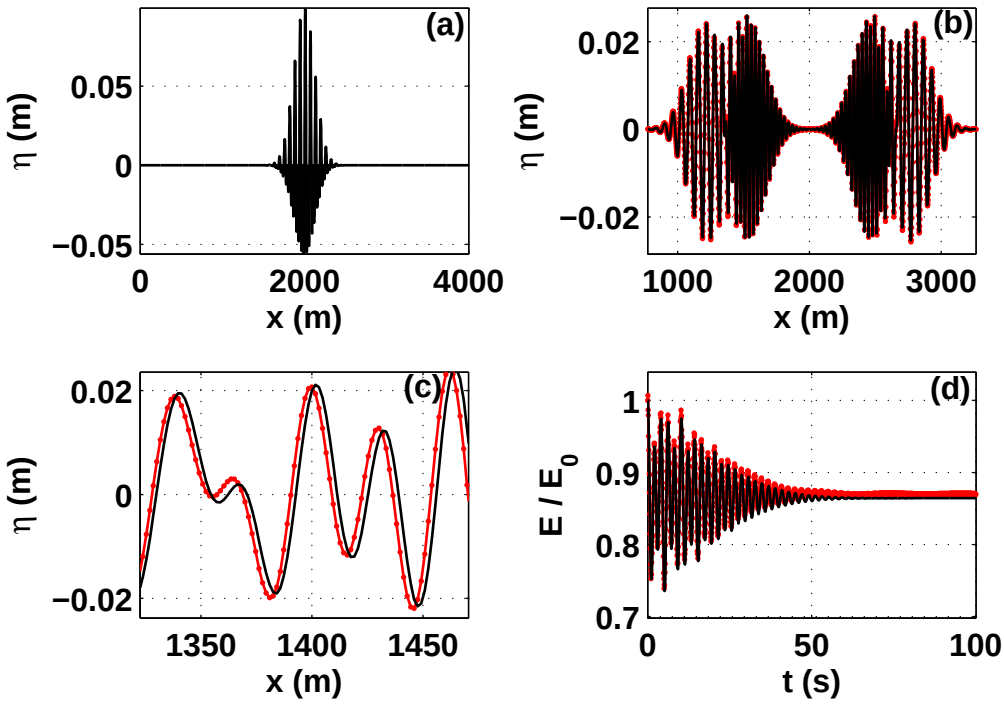

Figure 3.3: Results for the *dispersive short wave packets* run. Panel **(a)**: Initial free surface displacement $\eta(x, 0)$. Panel **(b)**: numerical solution (black, solid) and quasi-analytical solution (red, dots) at $t = 100$. Panel **(c)** is a zoomed-in view of the dispersive wavetrain from panel **(b)**, and panel **(d)** shows the total (domain-integrated) energy scaled by its initial value E_0 vs. time for the numerical solution (black, solid) and quasi-analytical solution (red, dots).

thus even exact solutions would not satisfy the physical property $E(t)$ =constant. Indeed, the plot in Figure 4.3 **(c)** confirms that $E(t)$ is oscillatory. This oscillatory behaviour is a well-known consequence of using Boussinesq-type systems, and Boussinesq-type systems that conserve energy exactly have been proposed [17]. However, these energy-conserving systems tend to be undesirable for numerical integration due to the presence of third-order spatial derivatives.

3.2.8 Grid-convergence study using a simulation of 1D wave-topography interaction

We next focus our attention on a 1D simulation of nonlinear and dispersive waves repeatedly propagating over a ridge with the Fourier method. Since analytical solutions are not available to confirm the validity of the results, we rely on grid-doubling experiments to illustrate the method's convergence in the well-resolved limit.

We begin by considering a periodic domain of length $L_x = 2$ km. The depth profile is given by

$$H(x,y) = H_1 - \Delta H \mathrm{e}^{-5(\frac{x-0.5L_x}{100})^4} \ , \tag{3.80}$$

with $H_1 = 10$ m and $\Delta H = 2$ m, reflecting a pre-dominantly flat bottom with a 2 m tall ridge in the center of the domain. The simulation was initialized using the initial conditions

$$\eta(x,0) = \eta_0 \mathrm{e}^{-(\frac{x-0.25L_x}{100})^2} \ , \tag{3.81}$$

$$u(x,0) = \sqrt{\frac{g}{H_1}}\eta(x,0) \ , \tag{3.82}$$

with $\eta_0 = 1$ m, representing a single wave of elevation, initialized to propagate in the positive x-direction with the off-ridge long wave speed $\sqrt{gH_1}$. A schematic diagram of the initialization is shown in Figure 3.4**(a)**. The governing equations were stepped forward until a final time of $t = T_{final} = 605$ s was reached. The final time was chosen such that a linear wave would traverse the length of the domain three times. The time-step was taken to be

$$\Delta t = \frac{1}{2}\frac{\Delta x}{\sqrt{gH_1}} \ , \tag{3.83}$$

and numerical instabilities were prevented by employing the spatial filtering methodology discussed in Section 3.2.3.

In this simulation, nonlinearity plays a key role in the evolution of the flow. The initial wave immediately begins to steepen, and the steepening is further enhanced due to shoaling

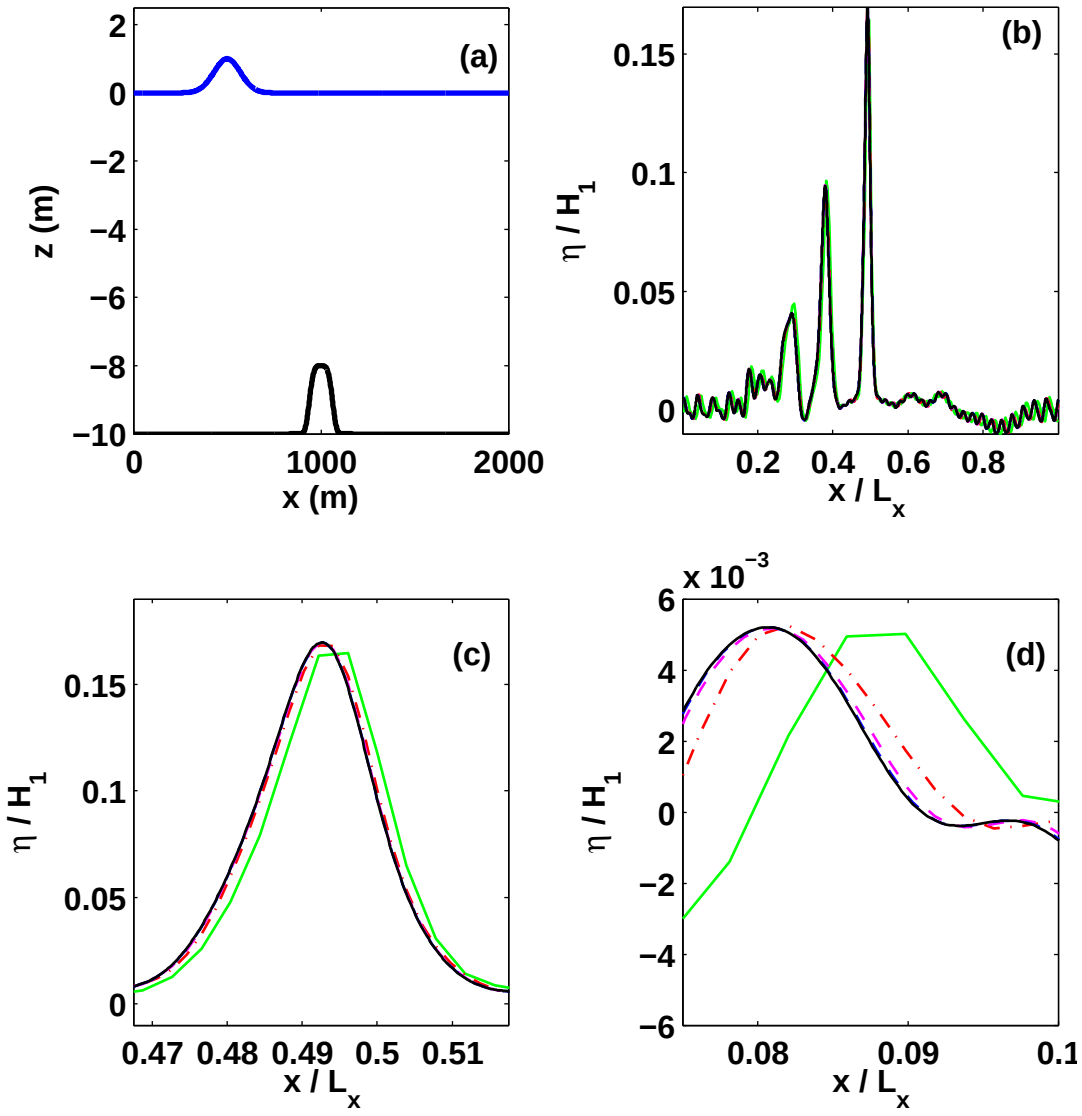

Figure 3.4: Results for the *1D wave-topography interaction* run. Panel **(a)**: Plot of the initialization, showing the topography $z = -H(x)$ and the initial free surface displacement $z = \eta(x, 0)$. Panels **(b)**–**(d)**: η at $t = T_{final}$ at resolutions $N_x = 256$ (solid, green), $N_x = 512$ (dash-dotted, red), $N_x = 1024$ (dashed, magenta), $N_x = 2048$ (dashed, blue), $N_x = 4096$ (solid, black). Panel **(c)** is zoomed-in on the leading solitary wave, and panel **(d)** is zoomed-in on a section of the dispersive tail. In panels **(b)**–**(d)**, the variable η has been made dimensionless by dividing by the off-ridge water depth, $H_1 = 10$ m.

as the wave propagates over the ridge (not shown). Dispersion then acts to balance the nonlinearity and prevent the formation of shocks. The final result is a collection of three solitary waves propagating in the positive x-direction followed by a dispersive wavetrain. These solitary waves are similar in shape to the sech^2 solitons predicted by Korteweg-de Vries (KdV) theory [130]. It can be shown that such solitons are approximate solutions to the governing equations (2.120)–(2.122) under the assumption of a flat-bottom, as was done for a similar system by [129].

Details of the η field at $t = T_{final}$ at several resolutions are depicted in Figure 3.4**(b)**–**(d)**. Inspecting the various plots suggests that grid convergence has been reached when $N_x = 2048$ grid points are used, since doubling the resolution once more to $N_x = 4096$ only yields minute differences in the fine-scale features of the η field (see Figure 3.4**(d)**).

3.2.9 A 2D simulation of wave generation by flow over topography

In our next test-case, we present a two-dimensional simulation of forced surface waves interacting with bottom topography to illustrate the numerical model's applicability to real-world problems in water wave dynamics. It is quite well known that when the inflow speed approaches the long wave speed, upstream propagating nonlinear waves are generated. This process is referred to as resonant generation [55]. Non-dispersive shallow water dynamics for flow over axisymmetric obstacles has been discussed by [43] using finite volume methods.

The physical parameters were set to: $g = 9.81$ m s^{-2}, $f = 0$ (no rotation), and $L_x = L_y = 2$ km, reflecting a (periodic) square domain. The grid was taken to have 2048 points in the x-direction and 256 points in the y-direction. Modal filtering in each direction was carried out using the parameters discussed in Section 3.2.3. The depth profile was taken to be

$$H(x, y) = H_1 - \Delta H \mathrm{e}^{-5(\frac{x-0.5L_x}{100})^4 - 5(\frac{y-0.5L_y}{200})^4}, \tag{3.84}$$

with $H_1 = 20$ m and $\Delta H = 2$ m. This is essentially a two-dimensional version of the depth-profile used in Section 3.2.8, i.e., a predominantly flat profile with a square-shaped ridge in the center of the domain. The simulation was initialized from quiescent conditions and forced by adding the body forcing term hF_x to the right-hand side of equation (2.121), where

$$F_x = \begin{cases} \frac{\sqrt{gH_1}}{\beta}, & 0 \leq t < 10 \text{ s} \\ 0, & t \geq 10 \text{ s} \end{cases} \tag{3.85}$$

and $\beta = 50/3$ s is a time-scale. The forcing is constant in space and piece-wise constant in time. Its effect is to induce a flow over the topography in the positive x-direction,

that is constant upstream of the topography. The value of β was chosen so that the final upstream velocity is equal to three-fifths of the off-ridge long wave speed, and hence the flow is formally sub-critical[1]. Since the addition of body forcing simply represents a source term in the governing equations, it was added to the time-stepping procedure using a straight-forward explicit evaluation.

Snapshots of the developing η-field are shown in Figure 3.5. In addition to a trapped wave of depression generated over the ridge, an upstream-propagating wavefront of elevation (with a slightly depressed tail, or possibly a second wave) can be seen emanating from the ridge, and travelling westward. This wavefront can be seen losing amplitude as time progresses. This is due to to radial spreading, or in other words geometric decay. The extent of this decrease in amplitude (energy density) due to geometric decay can be illustrated qualitatively by comparing this 2D simulation to an analogous 1D simulation where variations in the y-direction are neglected. This comparison is carried out in Figure 3.6, where it can be seen that in the 1D case (panel **(b)**), the upstream-propagating wave front better retains its amplitude than in the 2D case (panel **(a)**) since there is no radiation in the y-direction.

3.2.10 A 2D simulation of wave propagation over a shoal

In our final simulation in a doubly periodic domain, we follow the evolution of a wave front with an initially one-dimensional shape as it propagates over a shoal that is partitioned by a deep region in the center of the domain. The physical parameters (g, f, L_x, L_y) were chosen to be the same as in the previous simulation (see Section 3.2.9), as were the filtering parameters. The grid was taken to have 1024 points in each direction. The depth-profile is

$$H = H_1 - \Delta H \left[\operatorname{sech}\left(\frac{y - 0.5L_y}{500} \right) - \mathrm{e}^{-\left(\frac{r_0}{200}\right)^2} \right] . \tag{3.86}$$

Here, $r_0 = \sqrt{(x - 0.5L_x)^2 + 0.25(y - 0.5L_y)^2}$, $H_1 = 20$ m, and $\Delta H = 10$ m. The depth-profile is shown in Figure 3.7. The initial conditions were set to

$$\eta(x, y, 0) = \eta_0 \mathrm{e}^{-\frac{(y-0.1L_y)^2}{3200}} , \tag{3.87}$$

$$v(x, y, 0) = \sqrt{\frac{g}{H_1}} \eta(x, y, 0) , \tag{3.88}$$

$$u(x, y, 0) = 0 . \tag{3.89}$$

[1] A flow is said to be subcritical whenever the Froude number, $Fr = U/\sqrt{gH}$, is less than 1 and super-critical when $Fr > 1$.[70]

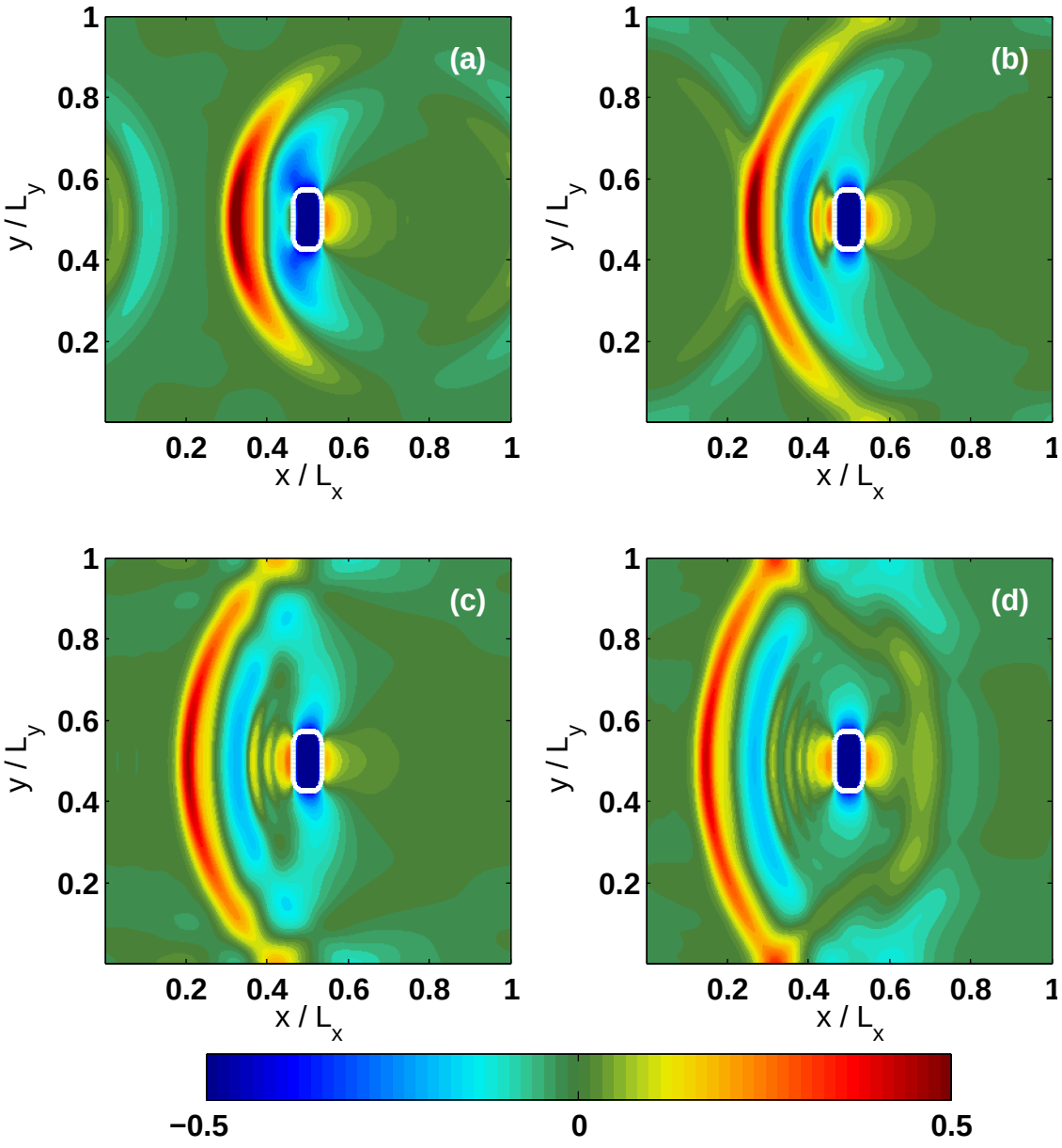

Figure 3.5: Fixed time snapshots of the free surface displacement at **(a)** $t = 60$ s, **(b)** $t = 80$ s, **(c)** $t = 100$ s, and **(d)** $t = 120$ s in the *2D wave generation by flow over topography* run. The solid-white line is the depth contour $H = 19.5$ m, indicating the location of the ridge.

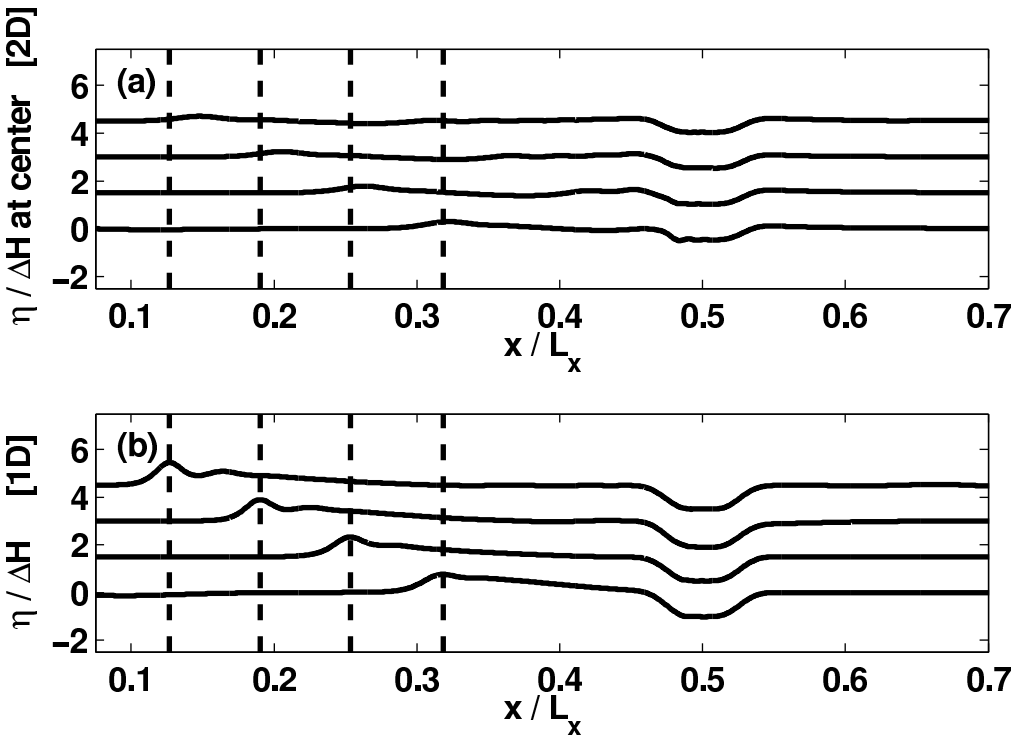

Figure 3.6: Panel **(a)**: 1D slices of the snapshots presented in Figure 3.5 through the line $y = 1$ km. Panel **(b)**: Snapshots of the η field for an analogous 1D simulation, where variations in y have been neglected. In each panel, a single curve corresponds to a time in Figure 3.5, with the lowest curve giving a slice through the snapshot taken at $t = 60$ s and the uppermost curve giving a slice through the snapshot taken at $t = 120$ s. Each curve has been shifted upwards by $3/40\,(t - 60)$ units. Dashed vertical lines represent the location of the maximum height of the upstream-propagating wavefront at each snapshot from the 1D simulation. The variable η has been made dimensionless by dividing by the ridge height, $\Delta H = 2$ m.

where $\eta_0 = 0.25\,(H_1 - \Delta H)$, reflecting a one-dimensional (symmetric in x) wave propagating in the positive y-direction at the linear long wave speed.

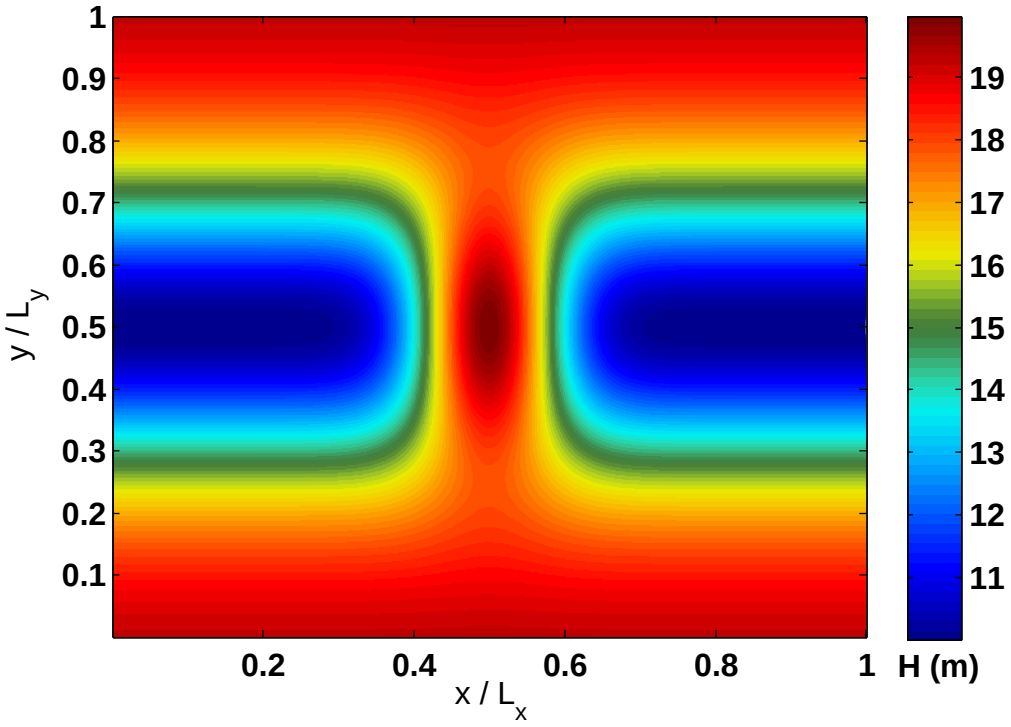

Figure 3.7: The depth-profile used in the *wave propagation over topography* run, corresponding to equation (3.86).

Snapshots of the evolving η field are shown in Figure 3.8. By $t = 70$ s (panel **(b)**), the symmetry in the x-direction has been broken due to refraction as the portion of the wave front propagating over the deep region in the center of the domain (near $x = 0.5L_x$) has a faster effective wave speed than the portion of the wave front that is propagating over the shoal. As the initial wave shoals, solitary waves emerge. The solitary waves are most evident in panels **(b)**,**(d)**, and **(f)** when the main wave front is situated on top of the shoal. At the later times (panels **(c)**-**(f)**), an interesting interference pattern of relatively weak waves follows the main wave front due to wave scattering and solitary wave fissioning induced by the topography.

While the decrease in relative amplitude of the portion of the main wave front over the deep region to that over the shallow region is intuitive, it should be noted that the absolute

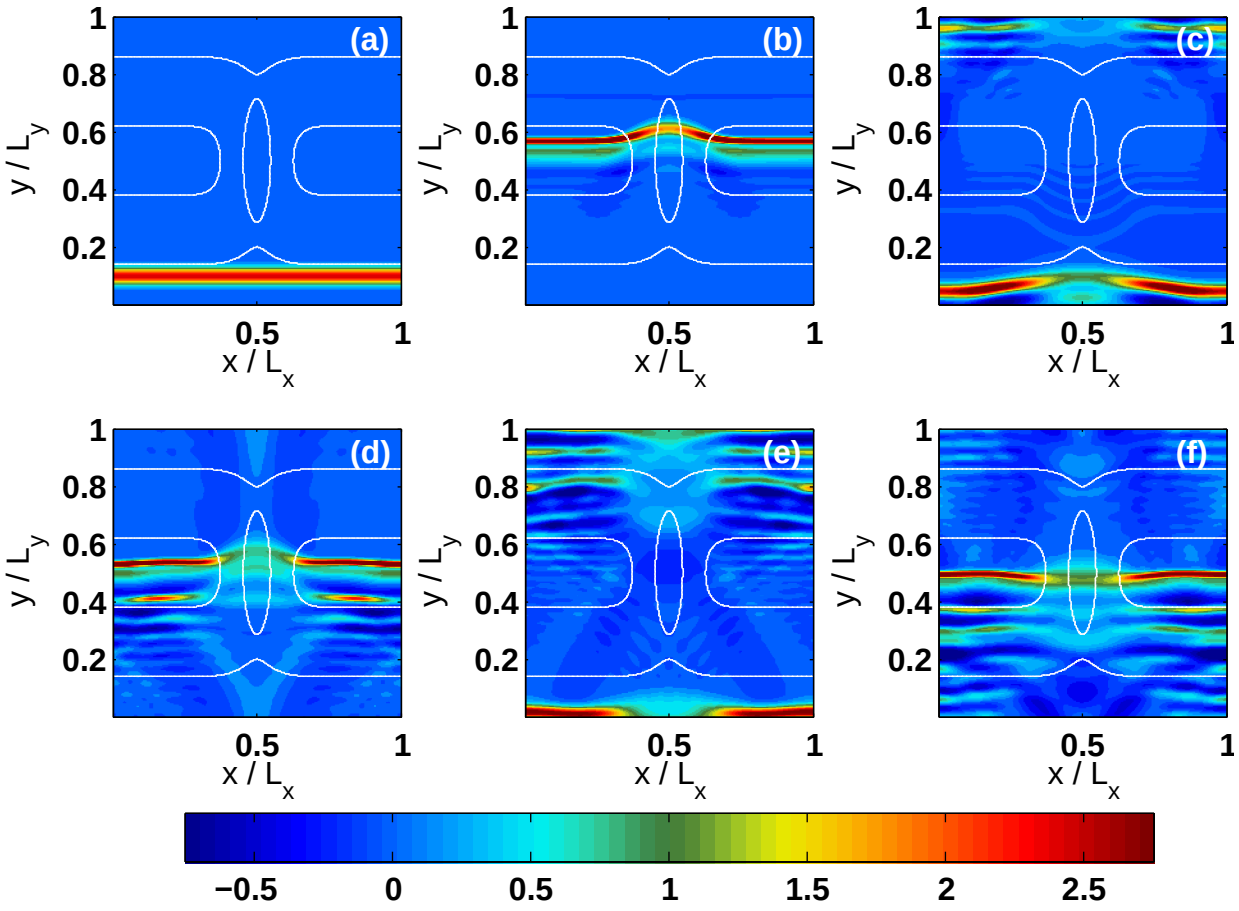

Figure 3.8: Fixed time snapshots of the free surface displacement at **(a)** $t = 0$, **(b)** $t = 70$ s, **(c)** $t = 140$ s, **(d)** $t = 210$ s, **(e)** $t = 280$ s, **(f)** $t = 350$ s in the *wave propagation over topography* run.

wave amplitude decreases over the deep region as well. This effect occurs due to the bending of wavefronts towards the lines of constant depth, which leads to a divergence of energy over the deep region. Thus, there is a corresponding focusing of energy near $0.3L_x < x < 0.4L_x$ and $0.6L_x < x < 0.7L_x$ that is more clearly seen in plots of u, the x-component of velocity, shown at times $t = 70$ s, $t = 210$ s, and $t = 350$ s in Figure 3.9. In field situations, this energy focusing could have implications for wave-boundary-layer interactions. It is also interesting to note that both the solitary wave widening and fissioning is qualitatively consistent with KdV theory [130].

In the following section, we explain how the Chebyshev method can be used in domains that are more general than the simple periodic ones considered above. In particular, we consider the influence of solid wall boundaries. Simulations with $f \neq 0$ are considered in the next section as well.

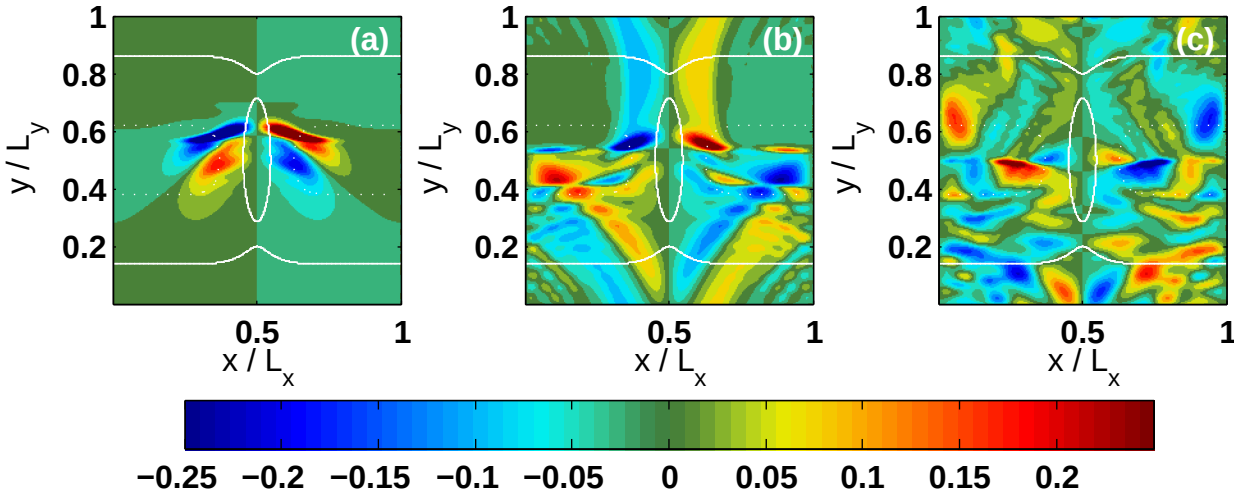

Figure 3.9: Fixed time snapshots of the u field (x-component of velocity) at **(a)** $t = 70$ s, **(b)** $t = 210$ s **(c)** $t = 350$ s in the *wave propagation over topography* run, corresponding to Figures 3.8**(b)**,**(d)**,**(f)**, respectively. The dotted and solid white lines correspond to the $H = 12$ m and $H = 18$ m depth contours, respectively, illustrating the shape of the topography.

3.3 Chebyshev method for domains with solid boundaries

As mentioned above, the Fourier method restricts the solution fields to be periodic in space. This may be a hindrance in problems where flow interactions with solid vertical boundaries, such as coastlines, are important.

One approach to use a pseudospectral method to represent wall boundaries is to use the Fourier method discussed above, but to expand the velocity in a sine-series about wall boundaries where the no normal flow condition (e.g. $u = 0$ at $x = 0$) is required. For consistency, this expansion choice then requires both the free-surface and the velocity tangent to the wall (e.g. v) to be expanded in cosine series about the wall. To see this, one could assume that u is an odd function of x in equation (2.121) and ask what the parity of v and η should be to ensure that all terms are odd. A main drawback to this approach is that if $f \neq 0$, the parities of η and v can not be chosen to ensure parity homogeneity. As a result, the solution would contain Gibbs oscillations and the rate of convergence would be reduced to algebraic order. Thus, the Fourier method is not useful in situations where the Coriolis effect is non-negligible.

It becomes evident that a more flexible basis, i.e., one that allows arbitrary boundary conditions, is necessary. Polynomials appear to be a clear choice for achieving this. The choice of which set of polynomials is driven by two factors: 1) the requirement that the polynomials are well-behaved in the well-resolved limit ($N_x \to \infty$); and 2) the desire to be able to perform "fast" operations as with the Fourier method. Given these requirements, the clear choice is the Chebyshev polynomials as is illustrated below.

Suppose we have a function $f(x)$ defined on $x \in [-1, 1]$. Its truncated Chebyshev series expansion is given by

$$f(x) \approx f_h(x) = \sum_{k=0}^{N_x-1} \hat{f}_k T_k(x) \ , \tag{3.90}$$

where $T_k(x)$ is the k^{th}-order Chebyshev polynomial of the first kind defined by $T_k(x) = \cos(k \cos^{-1}(x))$ [95]. If we make the change of variables $\theta = \cos^{-1}(x)$ the series becomes

$$f(x) \approx f_h(x) = \sum_{k=0}^{N_x-1} \hat{f}_k \cos(k\theta) \ . \tag{3.91}$$

Comparing the expansion (3.91) to (3.34) with $L_x = 2\pi$ it is clear that we have recovered the truncated cosine series on the interval $\theta \in [0, \pi]$. The relationship between the physical

values of $f(x)$ and the modes $\hat{f}_k$ can once again be established by the FFT. Pseudospectral differentiation can be performed as described in section 3.2 upon performing this change of variables.

Since the Fourier method works well for the equidistant nodes $\theta_i = \pi i/N_x$, $i = 0, \cdots, N_x - 1$, it is reasonable to assume that the Chebyshev basis should behave well and have good convergence properties at the nodes

$$x_i = \cos\left(\frac{\pi i}{N_x}\right), \quad i = 0, \cdots N_x - 1 \,, \tag{3.92}$$

where we have undone the above change of variable on the equidistant grid for $\theta \in [0, \pi]$. As it turns out, this set of points is the optimal choice of nodes for interpolation of the Chebyshev polynomials [122] and are often referred to as the Chebyshev points.

For simple rectangular domains, modes in the x and y direction are not coupled. As a result, it is easy to combine the Fourier and Chebyshev methods, using each basis along each dimension. This pratice may be physically relevant if vertical walls are only needed in one direction, as may be the case in a rectangular channel.

A periodic rectangular channel also serves as the boundary-following coordinates for circular geometries. The combination of the Fourier and Chebyshev methods thus allow for highly-accurate numerical modeling of closed (albeit specialized) lake basins. While circular basins may seem to be an esoteric case, they form a well studied class of problems in physical limnology dating back over a century ([119],[115]). High order numerical methods for such basins allow the robustness of classical solutions to be explored without the uncertainty associated with the inherent dissipation in many low order methods. This, in turn, allows for a rational set of hypotheses to be formulated for subsequent testing against field data.

In the following sub-sections, we discuss how the methodology presented above for doubly-periodic domains needs to be only slightly modified to solve the equations (2.120)–(2.122) on an annular (ring-shaped) domain. Here, we focus on ring-shaped domains since the polar coordinates mapping is singular and, hence, formulating an appropriate numerical treatment can become difficult when the domain contains the centre of the circle [122]. However, grid-stretching techniques can be used to overcome some of these issues as discussed in Boyd and Yu [13]. In the following sections, we first re-write the equations in polar coordinates and explain how the methods in section 3.2 need to be modified to deal with the solid-wall boundaries present at the inner and outer basin radii. Finally, the numerical model's potential for predicting and simulating wave motions in mid-sized lakes is illustrated with three simulation cases: 1) wave diffraction around an island and near-shore focusing; 2) the formation, propagation and destruction of wave trains and

solitary-like waves in rotating basins; and 3) the influence of bottom bathymetry on wave formation and propagation from a rotating seiche.

3.3.1 One-layer equations in polar form

In this section, we rewrite the one-layer dispersive shallow water system (2.120)–(2.122) in non-conservative form and in standard polar coordinates $(r,\theta) \in [r_{min}, r_{max}] \times [0, 2\pi]$. The resulting augmented system is

$$\frac{\partial h}{\partial t} + \frac{1}{r}\left(\frac{\partial (r h u_r)}{\partial r} + \frac{\partial (h u_\theta)}{\partial \theta}\right) = 0 \,, \tag{3.93}$$

$$\frac{\partial u_r}{\partial t} = a_r + \gamma \frac{\partial z}{\partial r} \,, \tag{3.94}$$

$$\frac{\partial u_\theta}{\partial t} = a_\theta + \frac{\gamma}{r}\frac{\partial z}{\partial \theta} \,, \tag{3.95}$$

$$\nabla\gamma \cdot \nabla z + \gamma \nabla^2 z = \nabla \cdot \mathbf{a} \,, \tag{3.96}$$

with

$$\nabla = \hat{\mathbf{r}}\frac{\partial}{\partial r} + \hat{\boldsymbol{\theta}}\frac{1}{r}\frac{\partial}{\partial \theta} \,, \tag{3.97}$$

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \,, \tag{3.98}$$

$$\nabla \cdot \mathbf{a} = \frac{1}{r}\left(\frac{\partial (r a_r)}{\partial r} + \frac{\partial a_\theta}{\partial \theta}\right) \,, \tag{3.99}$$

$$\mathbf{a} = a_r \hat{\mathbf{r}} + a_\theta \hat{\boldsymbol{\theta}}, \tag{3.100}$$

where the components of $\mathbf{a}$ are given by

$$a_r = -u_r \frac{\partial u_r}{\partial r} - \frac{u_\theta}{r}\frac{\partial u_r}{\partial \theta} + \frac{u_\theta^2}{r} - g'\frac{\partial \eta}{\partial r} + f u_\theta \,, \tag{3.101}$$

$$a_\theta = -\frac{g'}{r}\frac{\partial \eta}{\partial \theta} - f u_r - u_r \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r}\frac{\partial u_\theta}{\partial \theta} - \frac{u_r u_\theta}{r} \,, \tag{3.102}$$

where $\gamma = H^2/6$ is, in general, allowed to vary in space, and u_r and u_θ are the radial and annular velocities, respectively. In moving from the system (2.120)-(2.122) to (3.94)-(3.96), we have invoked the *mild slope* approximation $\nabla h \approx 0$ in the non-hydrostatic term, so that

$$\gamma \nabla \left(\nabla \cdot (\mathbf{u} h)_t\right) \approx \gamma h \nabla \left(\nabla \cdot \mathbf{u}_t\right) \,. \tag{3.103}$$

Such approximations are common in Boussinesq-type models ([28],[15]) where high accuracy modelling of short-wave/topography interactions is not of primary concern.

3.3.2 Fourier-Chebyshev Method for Annular Domains

As mentioned above, the system is discretized in space using a Fourier pseudospectral method in the annular direction and a Chebyshev pseudospectral method in the radial direction. In our simulations, we always take $r_{min} > 0$ to avoid the singularity and undesired clustering of grid points near $r = 0$ that are associated with the standard polar coordinates mapping. It is worth mentioning, however, that it is possible to reduce the amount of undesired clustering of Chebyshev points near the origin by using a mapped Chebyshev grid that is only heavily clustered near the outer boundary, as explored by Boyd and Yu [13]. As in Section 3.2, time-stepping is carried out with the second-order accurate Leapfrog method.

The boundary conditions imposed are reflective wall conditions at the inner and outer radii of the basin:

$$u_r = 0 \quad \text{at} \quad r = r_{min} \text{ and } r_{max} \ , \tag{3.104}$$

$$\frac{\partial \eta}{\partial r} = 0 \quad \text{at} \quad r = r_{min} \text{ and } r_{max} \ . \tag{3.105}$$

Since $\mathbf{u}$ and η are calculated with explicit time-stepping, these boundary conditions are enforced by over-writing the boundary values of $\mathbf{u}$ and η at the completion of each time-step. The enforcement of (3.105) relies on re-arranging the Chebyshev differentiation matrix [122] to solve for the unknown values of $\eta(r = r_{max}, \theta)$ and $\eta(r = r_{min}, \theta)$ such that (3.105) is satisfied.

A suitable time-dependent inhomogenous Neuman boundary condition for the elliptic variable z is obtained by taking the dot-product between the vector form of the momentum equations (3.94)-(3.95) and the unit outward boundary normal $\hat{\mathbf{n}}$ and solving for the normal derivative of z, i.e.,

$$\frac{\partial z}{\partial n} = -\frac{\mathbf{a} \cdot \hat{\mathbf{n}}}{\gamma} \ \text{ at } r = r_{min}, r_{max} \ , \tag{3.106}$$

where we have used the fact that $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ at $r = r_{min}$,r_{max}.

As with the Fourier-Fourier method discussed in 3.2 the spatial discretization method relies on the pseudospectral technique [95] for speed and memory efficiency. The primary difference here being that a fast cosine transform (implemented with FFT) must be used in the radial direction, however, the periodic coordinate is treated with the usual FFT. As mentioned previously, the pseudospectral technique relies on the point-wise calculation of nonlinear products on the physical grid which can in turn lead to aliasing errors that may be tamed by modal filtering. In this case, the equations in polar form (3.94)–(3.95)

possess additional nonlinearities over the cartesian equation (2.120)–(2.122) since the polar coordinates mapping is itself nonlinear. For this reason, we have found that stronger filtering is required to obtain stable solutions than what was previously needed.

As discussed previously, the elliptic problem (3.96) is solved iteratively with the generalized minimum residual method (GMRES) using the LU-factorization of the second-order central finite differences approximation [63] of the elliptic operator as a pre-conditioner in order to reduce the iteration count to a reasonable number. The only difference here is that the linear operator must be adjusted along the lake boundaries to enforce the boundary condition (3.106).

3.3.3 Model Parameters

In the results presented in this section, the model resolution was given by $N_r = 256$ radial (Chebyshev) points and $N_\theta = 1024$ annular (equispaced, Fourier) points. The spatial filtering parameters for the filter (3.41) were typically taken to be: $k_{crit} = 0.15 k_{max}$, $s = 4$, $\alpha = \ln 10^{-12}$, where k_{max} is the Nyquist wavenumber.

The GMRES convergence criterion was chosen to require the relative residual to be smaller than 10^{-8}. The number of GMRES iterations required per time-step varied by simulation, but was typically between 9 and 16 and fluctuated throughout run-time. A sudden growth in iteration count during time-stepping tended to indicate that the simulation was approaching a numerical instability, and thus that the model should be re-run with a different set of filtering parameters to ensure stability.

The time-step Δt was determined through experimentation, but the CFL condition was always employed as a first guess or "rule of thumb":

$$\Delta t \leq \frac{\min_{(r,\theta)}\{\Delta r, r\Delta\theta\}}{U_{max} + \sqrt{g'H}} \,, \tag{3.107}$$

where $U_{max} = \max_{(r,\theta)} \sqrt{(u_{r,0})^2 + (u_{\theta,0})^2}$ is the maximum speed of the initial conditions.

The numerical methodology was validated against approximate analytical solutions in one dimension and compared to numerical solutions in two dimensions obtained with the DG-FEM method at various orders of accuracy (see section 4.11). The pseudospectral method presented here was found to have better resolution and energy-conserving characteristics than the DG-FEM method in all cases. In particular, many fine scale flow features found in the results shown below are not present in the low order DG-FEM simulations, likely because they are 'smeared out' by numerical diffusion.

3.3.4 2D Simulation of Wave diffraction around an island and near-shore focusing (hereafter, WDAINF)

In the first simulation, the inner and outer basin radii were taken to be $r_{min} = 100$ m and $r_{max} = 1272$ m, respectively. The mean fluid depth was set to $H = 10$ m (constant), the reduced gravity was taken as $g' = 0.196\text{ms}^{-2}$, and rotation was turned off ($f = 0$). The initial conditions were taken to be

$$\eta_0 = \mathrm{e}^{-0.5(y+600)^2}\,, \tag{3.108}$$

$$u_{\theta,0} = 0\,, \tag{3.109}$$

$$u_{r,0} = 0\,, \tag{3.110}$$

representing a still fluid with a rectangular Gaussian interface perturbation stretching across the west-east length of the basin centered about the line $y = -600$ m. Here, subscript '0' denotes the respective flow variable at time $t = 0$.

Since the fluid is initially at rest, the initial condition splits into two component waves, one travelling northward and one travelling southward. At $t = 6.6$ min, the southward propagating wave begins to strengthen as it propagates longshore and approaches the southern end of the basin. Meanwhile, the northward propagating wave is diffracting around the circular island centered at $(x, y) = (0, 0)$. At $t = 13.3$ min, the initially northward propagating wave is beginning to focus as it approaches the northern end of the basin, while the initially southward propagating wave has reflected off the southern coastline and has begun propagating northward, now focused as a predominantly single egg-shaped elevation wave. The process continues for all other times shown, with the focused waves spreading and undergoing wave-wave interactions with waves propagating in the opposite direction.

3.3.5 2D simulation of the formation, propagation and destruction of wave trains and solitary-like waves, and the effect of rotation (hereafter, FPDWT)

In our next simulation, the physical parameters were chosen such that the idealized circular basin would be similar to the physical situation of Lake Kinneret, Israel by inferring approximate values from the data presented in [104]. To that end, the physical parameters were chosen as follows: $r_{max} = 8345$ m, $r_{min} = 1$ km, $g' = 0.024525$ m s^{-2}, $H = 12.8$ m, $f = 7.8828 \times 10^{-5}$ s^{-1}. Under these parameters, the Rossby deformation radius is

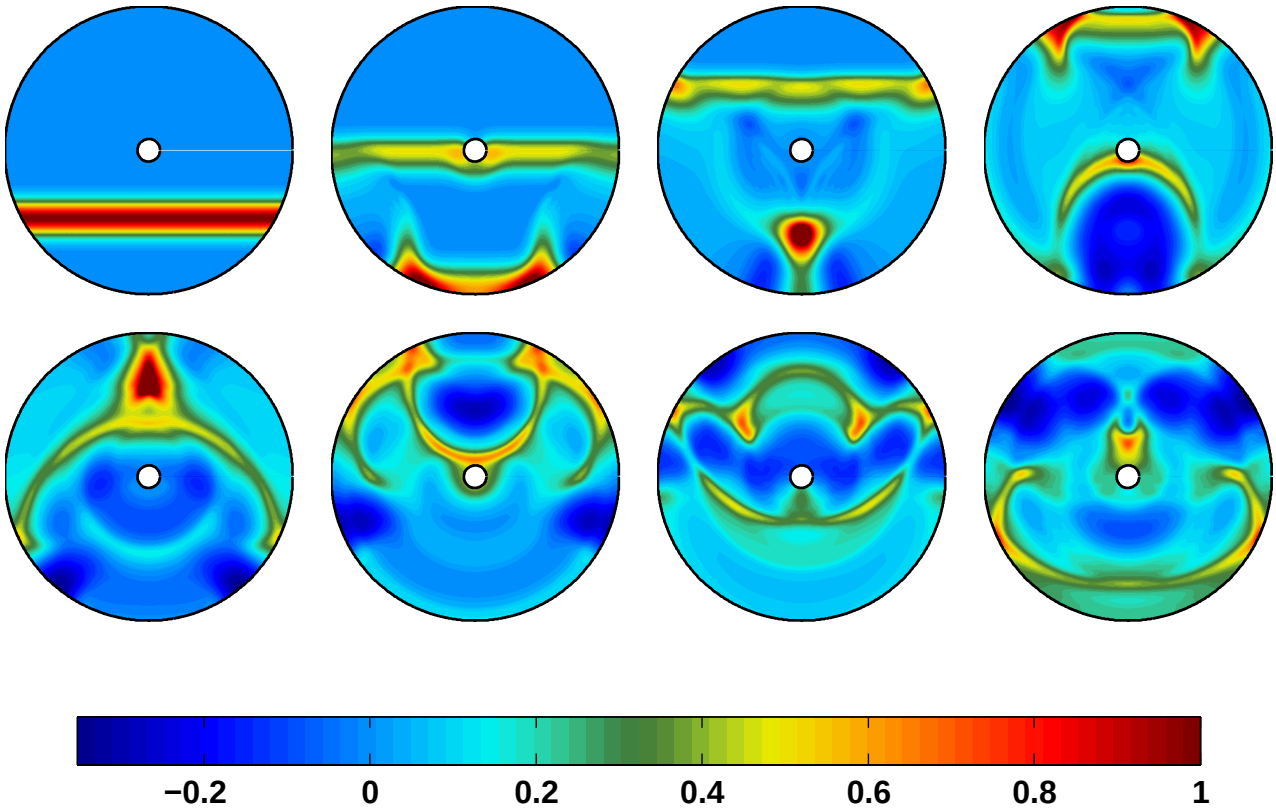

Figure 3.10: Selected snapshots of the η field in the *WDAINF* test-case at times $t = 0, 6.6, 13.3, 19.7$ min (top) and $t = 26.2, 32.8, 39.4, 45.9$ min (bottom). The contour interval is given by $[0,\ \eta_0^{max}]$. The inner and outer basin radii are given by $r_{min} = 100$ m and $r_{max} = 1272$ m, respectively. Rotation was turned off in this test-case ($f = 0$).

$L_D = \sqrt{g'H}/f = 7107$ m, which is quite similar to the distance between the inner and outer basin radii.

The initial conditions are taken to be

$$\eta_0 = \eta_0^{max}\, \mathrm{e}^{-10^{-7}(r-r_{max})^2 - 50\left(\theta - \frac{\pi}{2}\right)^2}\,, \tag{3.111}$$

$$u_{\theta,0} = \sqrt{\frac{g'}{H}}\eta_0\,, \tag{3.112}$$

$$u_{r,0} = 0\,, \tag{3.113}$$

representing a coastally localized interfacial perturbation propagating in the annular direction at the long internal wave speed $\sqrt{g'H}$.

In Figures 3.11 and 3.12 the maximum amplitude of the initial disturbance was taken to be $0.01H$ and $0.25H$, respectively to illustrate the effect of nonlinearity on a coastally-propagating wave in a lake similar to Lake Kinneret.

In Figure 3.11, the effects of nonlinearity are negligible, and as a result no steepening occurs. By $t = 14$ h, the initial shape of the disturbance has spread out considerably and a region of interfacial depression follows the primary elevation wave. By $t = 28$ h, the region behind the depression has become the highest point of interfacial elevation, and the region of largest amplitude. This process of the wavefront losing energy to the interior of the basin by long Poincaré waves continues throughout the evolution. At subsequent times, the resulting wave field is best described as a nearly basin-scale wavefront followed by an undular tail, with the wavelength decreasing towards the rear of the tail.

In Figure 3.12, the effects of nonlinearity play a key role in the evolution of the wave field. By $t = 7$ h, the initial disturbance has steepened up to a near shock. However, the formation of the shock is prevented by the dispersive terms in equations (2.121)-(2.122). When nonlinear steepening and dispersion are in balance, solitary waves become possible [130] and a collection of high-frequency solitary-like waves can be observed in the wave field; however, this is difficult to discern in Figure 3.12. When the flow is animated, it appears that two-dimensional interactions between the solitary-like waves occur, and that the waves lose energy by interacting with the interior of the basin. As in the linear case (Figure 3.11), the elevation at the front of the wave eventually decays sufficiently so that the elevated region behind the interfacial depression becomes the region of largest amplitude in the basin. After sufficient time has passed, solitary-like waves can be found in this secondary region of elevation due to nonlinear steepening, and the process continues with a further loss of energy downstream.

A more detailed look at the early time nonlinear evolution for the case when $\eta_0^{max} = 0.25H$ is shown in Figure 3.13. Figure 3.13**(a)** reveals the key role that dispersion plays in

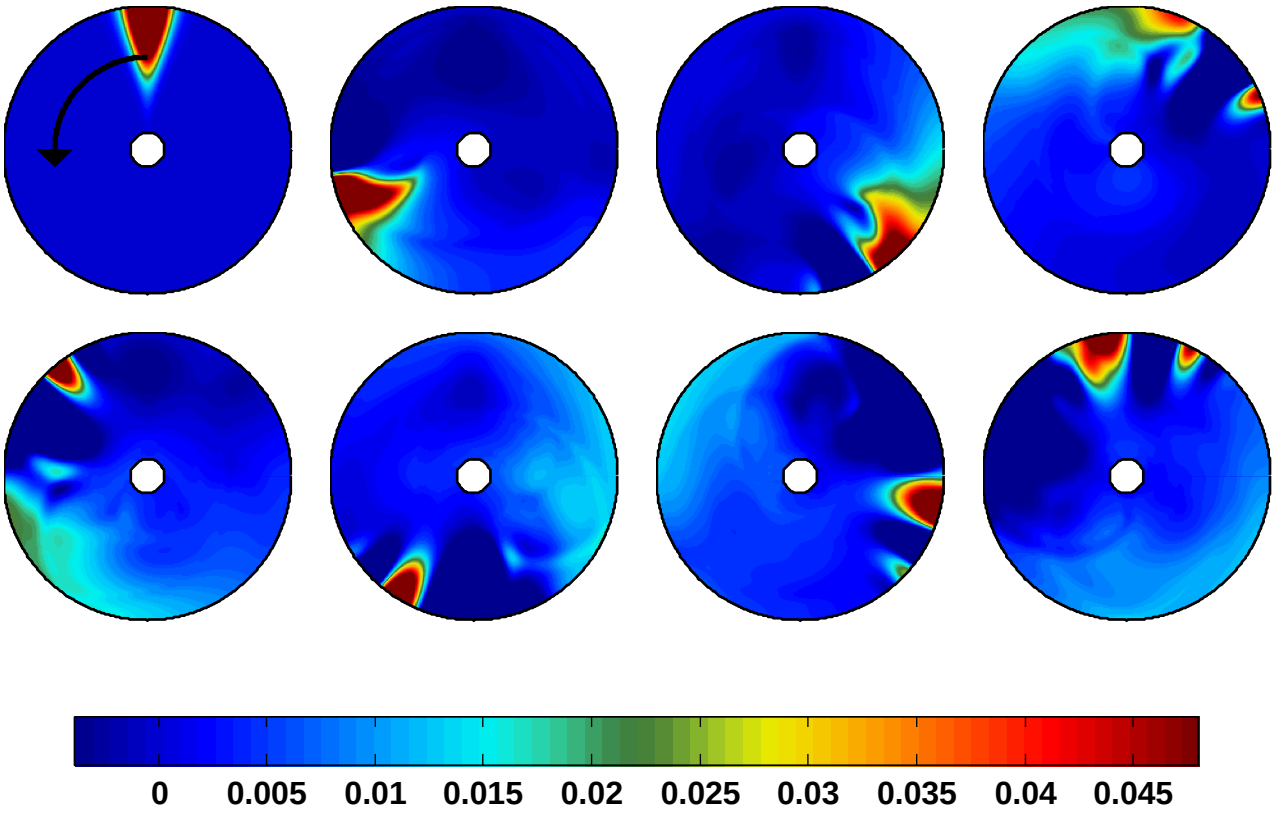

Figure 3.11: Evolution of an initial interfacial perturbation propagating in the counter-clockwise longshore direction in the *FPDWT* simulation with $f = 7.8828 \times 10^{-5}$ s^{-1}, $r_{min} = 1$ km, $r_{max} = 8345$ m. Snapshots were taken at $t = 0, 7, 14, 21$ h (top) and $t = 28, 35, 42, 49$ h (bottom). The maximum amplitude of the perturbation was taken to be $\eta_0^{max} = 0.01H$.

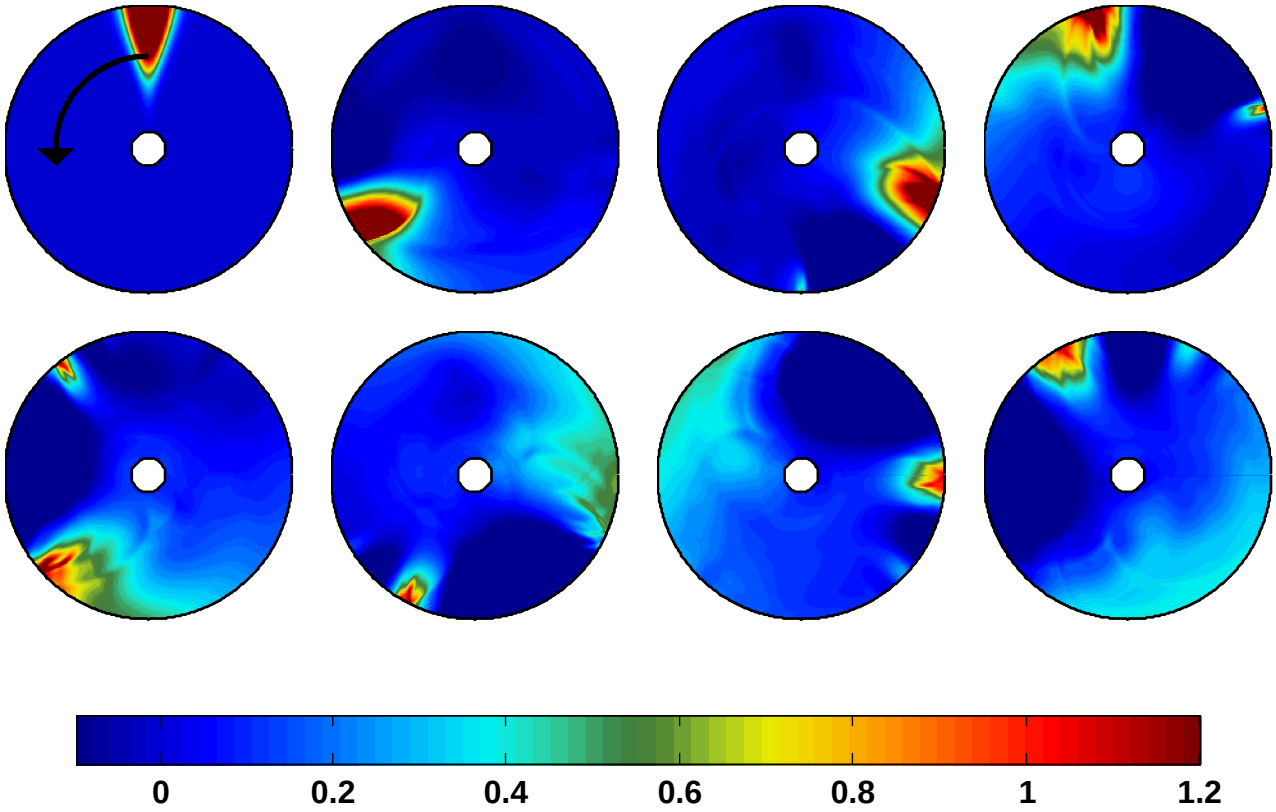

Figure 3.12: Like Figure 3.11, but with $\eta_0^{max} = 0.25H$.

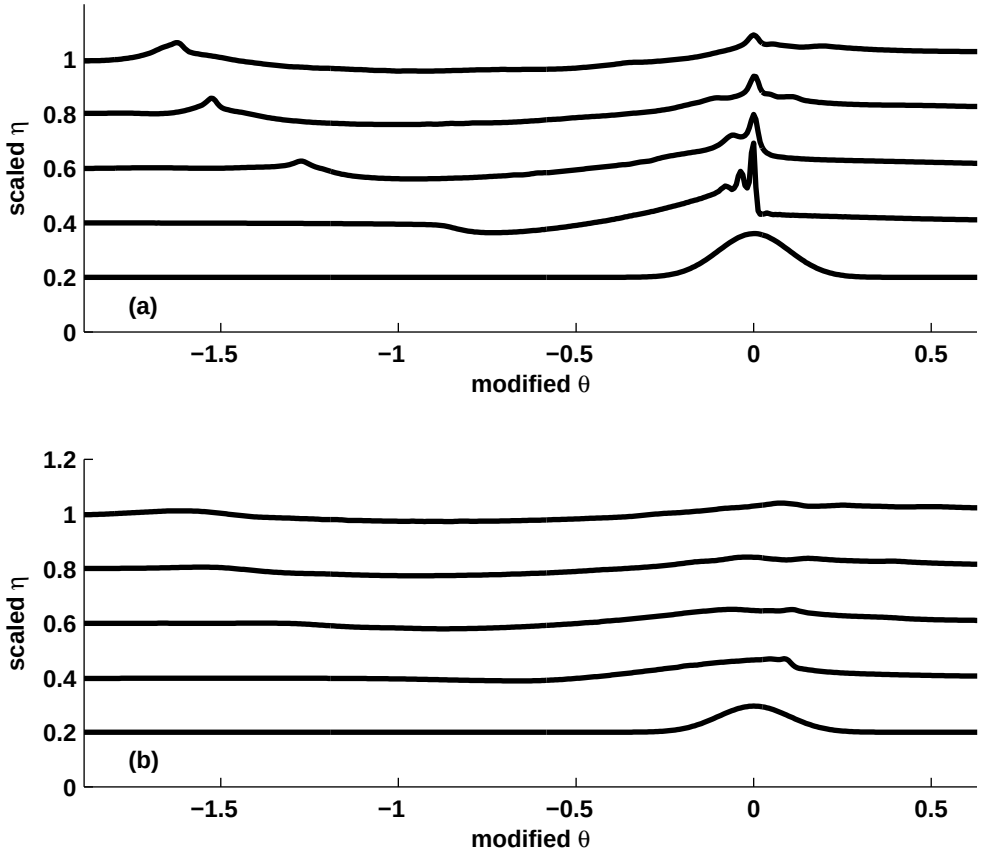

Figure 3.13: 1D slices of the snapshots presented in Figure 3.12 through the circles **(a)** $r = r_{max} = 8345$ m and **(b)** $r = 2268$ m. In each panel, a single curve corresponds to a time in Figure 3.12, with the lowest curve giving a slice through the initial condition ($t = 0$ h) and the uppermost curve giving a slice through the snapshot taken at $t = 28$ h. The slices have been shifted in such a way that the main wave-front is centered about $\theta = 0$, and η has been made dimensionless by dividing by the undisturbed layer thickness H.

the early evolution of the wave field: disallowing a shock in favour of a train of dispersive shortwaves and a coherent solitary-like wave. At later times, the primary wave-front is seen losing energy to its tail wave (of elevation). This effect was observed by Stastna *et al.* [109] and Helfrich [59] in simulations of the evolution of solitary waves affected by rotation. A comparison of panels **(a)** and **(b)** reveals decay in amplitude of the wave as we move inwards from the coast ($r = r_{max}$). Also visible is the fact that the portion of the wave in the interior of the basin leads the portion at the edge, due to the boundary curvature's influence on the portion of the wave nearest to the coast.

We now turn our attention to the effects of f-plane rotation on the evolution of the wave-field for an initial interfacial perturbation of fixed amplitude. In Figure 3.14, f is almost doubled to 1.4544×10^{-4} s^{-1}, reflecting the Coriolis frequency at the North Pole, the highest value of f possible on Earth. In this case, the Rossby deformation radius is $L_D = 3852$ m, or about half-way between the inner and outer basin radii.

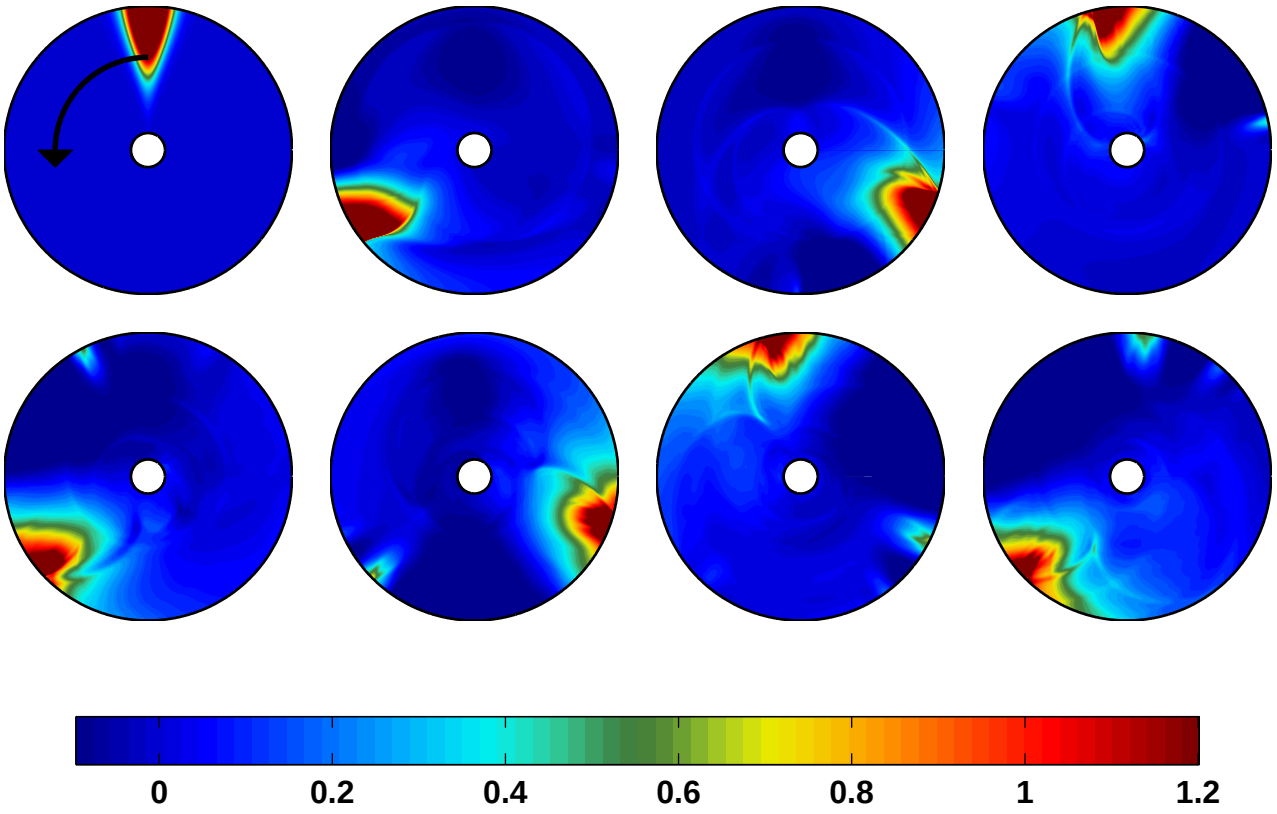

Figure 3.14: Like Figure 3.11, but with $\eta_0^{max} = 0.25H$ and $f = 1.4544 \times 10^{-4}$ s^{-1}.

Due to the linear analysis performed by Stocker and Imberger [115], we expect increasing f to result in increasing the extent to which the lowest frequency gravity mode will be trapped along the coast, thereby weakening the interactions between our coastally propagating disturbance and the interior of the basin. Indeed, this effect is apparent in our results and we make two main observations: firstly, it is found that the length of time required for the primary wavefront to lose energy to its tail is increased substantially when compared to Figure 3.12; and secondly, the outward spreading of the initial disturbance to near basin scales takes considerably longer. This result may also be interpreted in terms of the Rossby deformation radius, L_D: a smaller value of L_D lowers the effect of boundary curvature, leading to a scenario more akin to a Kelvin wave propagating along a straight coastline, hence there are fewer radiating Poincaré waves. These effects can be visualized more clearly by directly comparing slices of the wave field at fixed times between the two simulations corresponding to $f = 1.4544 \times 10^{-4}$ s^{-1} and $f = 7.8828 \times 10^{-5}$ s^{-1} with $\eta_0^{max} = 0.25H$. This comparison is done in Figure 3.15. It can be seen that the simulation with the higher rotation rate maintains a coherent leading wave for twice as long as the lower rotation rate case. Furthermore, the steepening of the secondary wave is decreased in the high f case. Finally, in the high f case significantly more short wave activity upstream of the leading disturbance is evident (especially for the final three times shown).

3.3.6 2D simulation of the wave formation and propagation from an internal rotating seiche and the influence of bottom bathymetry. (hereafter, IKS)

In the final simulation employing annular domains, we consider the evolution of a basin wide linear tilt in the density interface with the fluid initially at rest. The corresponding initial conditions for a west-east tilt are:

$$\eta_0 = \frac{ax}{r_{max}}, \tag{3.114}$$

$$u_{\theta,0} = 0\,, \tag{3.115}$$

$$u_{r,0} = 0\,, \tag{3.116}$$

where a is the maximum amplitude of the tilt. The physical parameters $(f, g', H, r_{max}, r_{min})$ are given by the same values used in *FPDWT* simulation, intended to mimic the real-world situation of Lake Kinneret under summer stratification conditions.

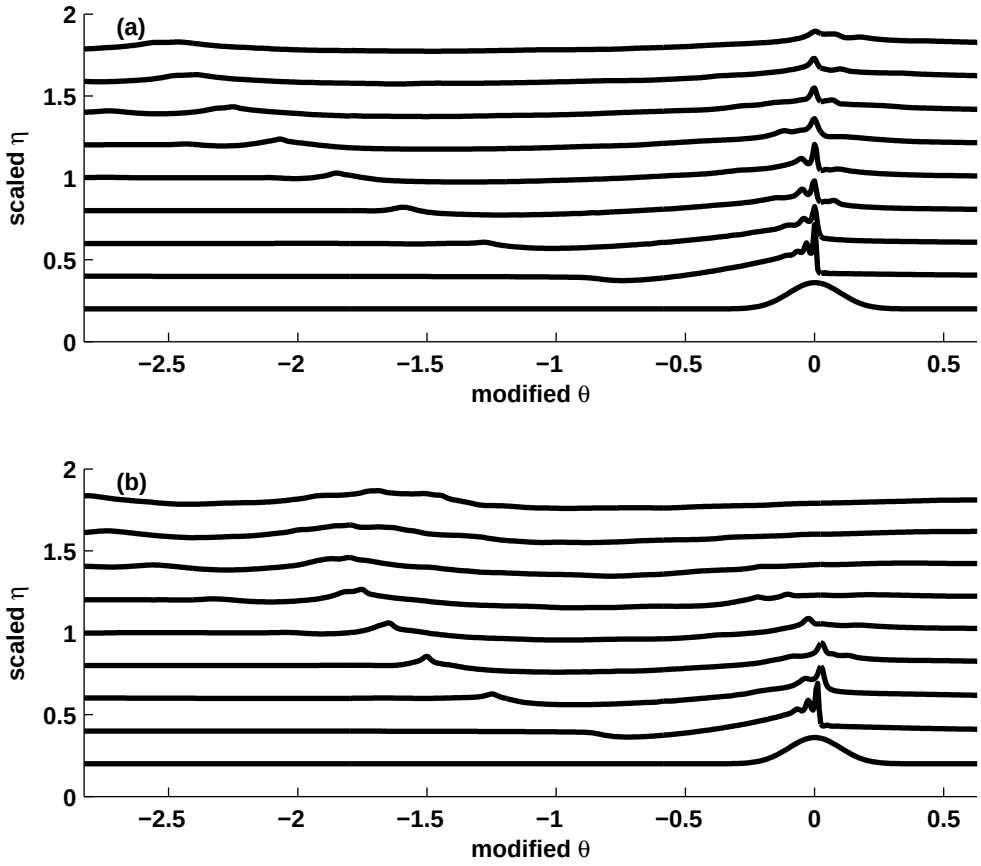

Figure 3.15: 1D slices of the snapshots presented in Figures **(a)** 3.14, and **(b)** 3.12 at $r = r_{max} = 8345$ m. In each panel, a single curve corresponds to a time in Figure 3.12, with the lowest curve giving a slice through the initial condition ($t = 0$ h) and the uppermost curve giving a slice through the final snapshot. The slices in panel **(a)** have been shifted in such a way that the main wave-front is centered about $\theta = 0$, and the same shift has been applied in panel **(b)**. Here, η has been made dimensionless by dividing by the undisturbed layer thickness H.

In the case where $f = 0$ (no rotation), one would expect a single basin-scale periodic standing wave (or seiche) to result [48]. However, in the present case where $f = 7.8828 \times 10^{-5}$ s^{-1} > 0 (Northern Hemisphere), we expect the ensuing basin-scale dynamics to be driven by the lowest frequency rotating gravity mode from the linear theory (see also Chapter 6). The ensuing cyclonic motion has also been referred to as a "Kelvin-seiche" [74]. However, the term "Kelvin wave" is usually reserved for gravity modes with frequencies lower than f.

Snapshots of this rotating seiche motion at fixed times are shown in Figures 3.16 and 3.17 corresponding to initial tilt amplitudes of $a = 0.1H$ and $a = 0.25H$, respectively. A comparison of the two figures reveals the effect of nonlinearity due to an increase in initial tilt-amplitude. In Figure 3.17, the effects of nonlinear steepening are visible as the clustering of η-contours at later times and the emergence of long "filament waves" stretching across the basin (particularly at $t = 10.67$ h and $t = 18.67$ h). As in the *FPDWT* simulation, the formation of shocks does not occur due to the dispersive nature of the model equations (dispersive effects in an actual lake would also preclude shock formation).

Some information about the effect of nonlinearity, and the details of the wave field can be obtained from the line plots shown in Figure 3.18. It can be seen that strongly non-sinusoidal waveform shapes occur in the larger amplitude case (panels **(a)** and **(c)**), with some short wave generation at later times. In contrast, the smaller amplitude case is nearly sinusoidal, for the times shown.

The final simulation shown here in Figure 3.19 **(a)** is intended to illustrate the numerical model's robustness in dealing with situations where the bottom of the basin is not flat. Here, the depth profile was taken to be

$$H = \overline{H}\left(1 - \frac{1}{4}\left(\frac{r}{r_{max}}\right)^2\right) , \tag{3.117}$$

where $\overline{H} = 12.8$ m. This profile reflects a lake with a bowl-shaped bottom topography and is a reasonable idealization for the large scale bathymetry of typical, real-world, mid-sized lakes. The effects of bottom topography are best observed by comparing the run that includes topography (Figure 3.19 **(a)**) to the same run but with a flat bottom (Figure 3.19 **(b)**). We compare these two runs at later times when considerable differences in the wave fields have had time to develop. Since the two depth profiles are similar near

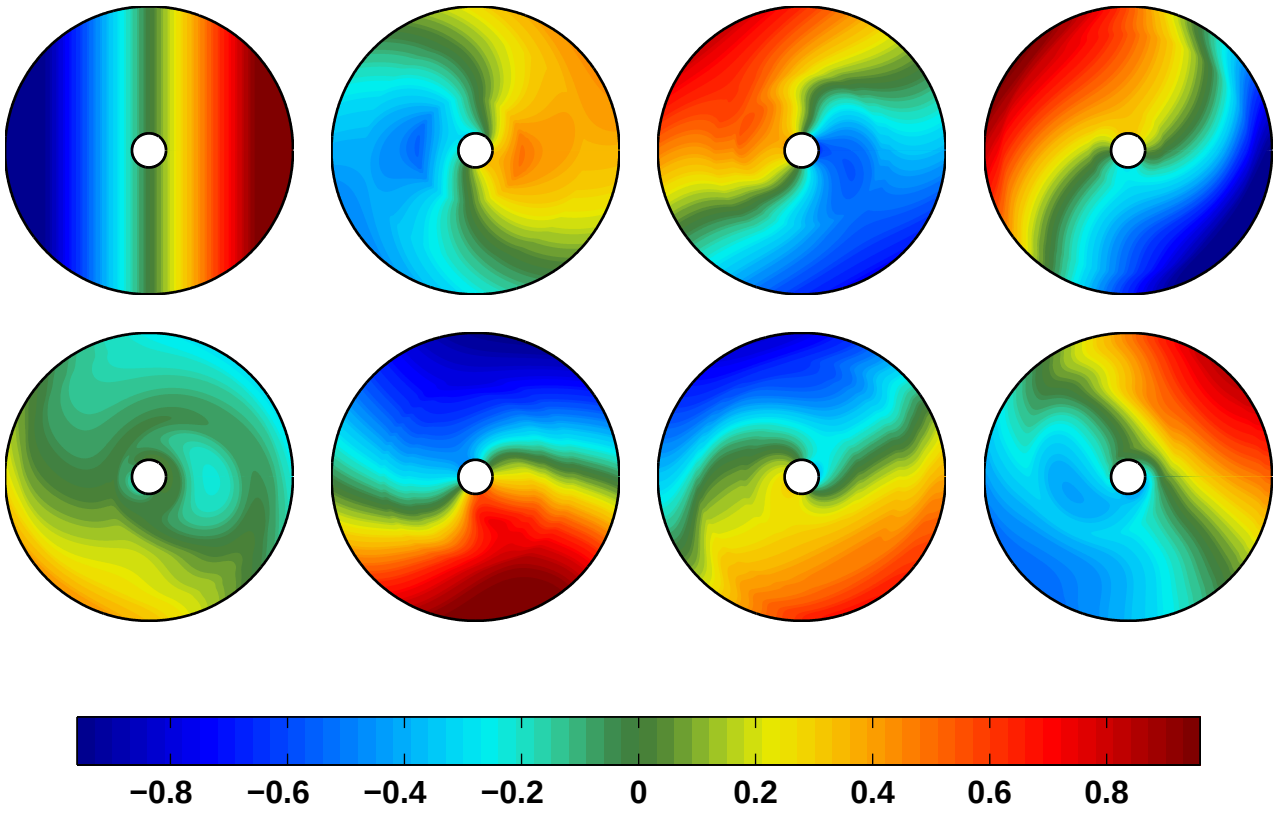

Figure 3.16: Selected snapshots of the η field in the *IKS* test-case with $a = 0.1H$ at times $t = 0, 2.67, 5.33, 8.00$ h (top) and $t = 10.67, 13.33, 16.00, 18.67$ h (bottom).

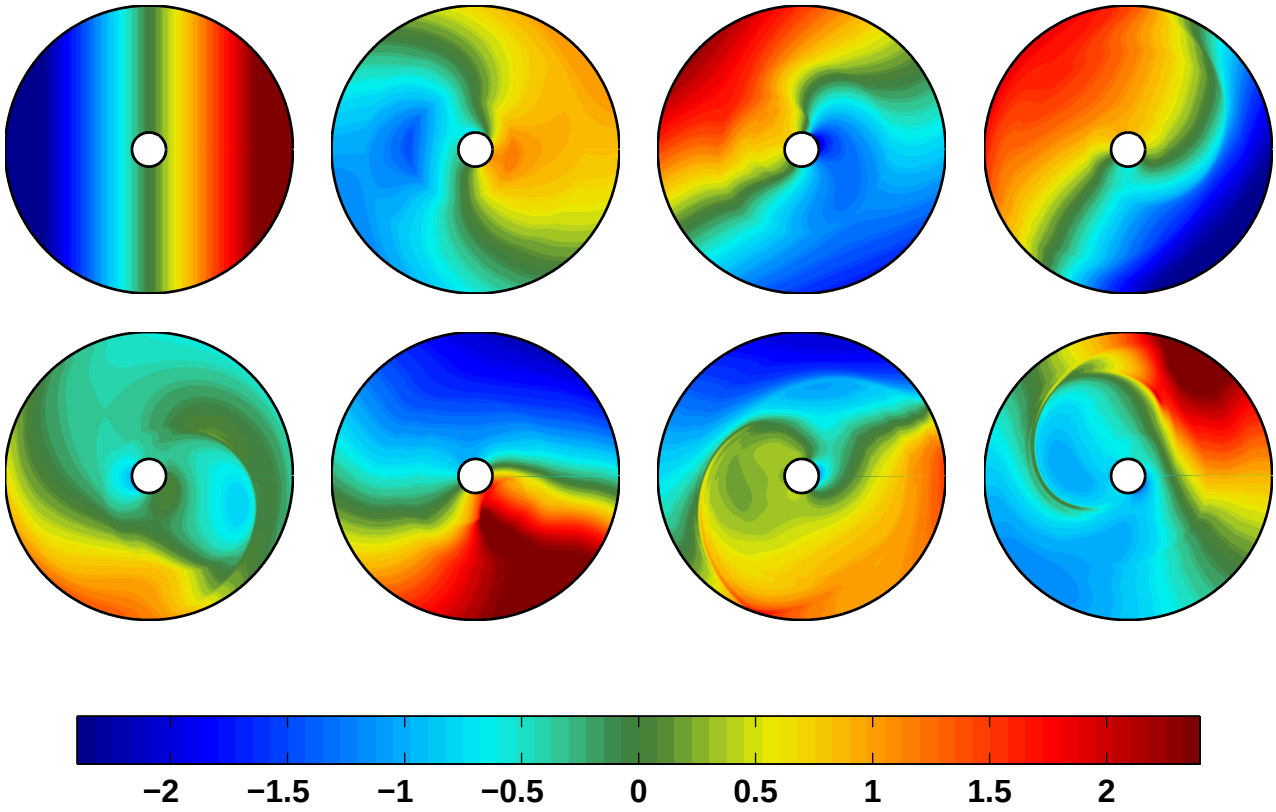

Figure 3.17: Like Figure 3.16, but with $a = 0.25H$.

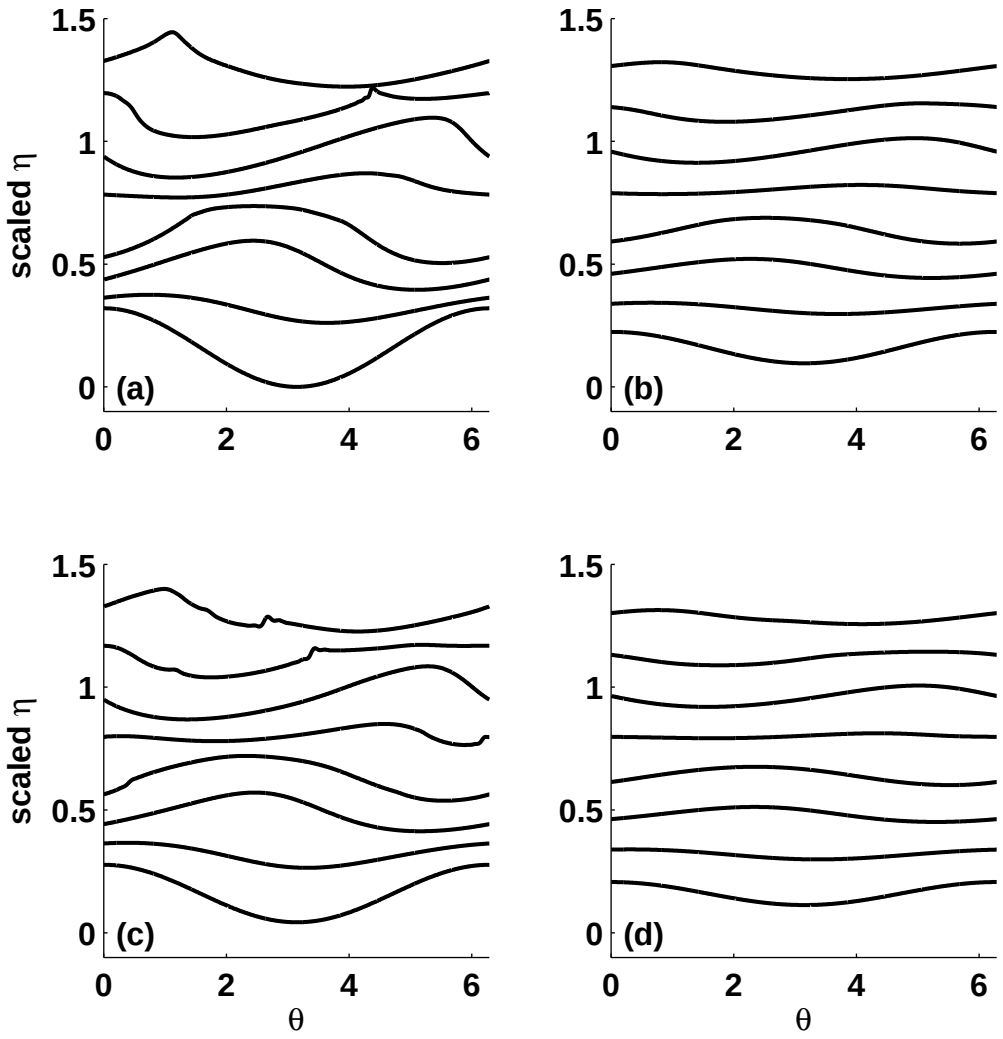

Figure 3.18: Scaled η at $r = r_{max} = 8435$ m (panels **(a)** and **(b)**) and $r = 6167$ m (panels **(c)** and **(d)**) versus θ corresponding to Figure 3.17 (panels **(a)** and **(c)**) and Figure 3.16 (panels **(b)** and **(d)**). In each panel, the bottom-most profile corresponds to the initial conditions ($t = 0$) while the top-most profile corresponds to the bottom rightmost panel in Figures 3.17 and 3.16. The corresponding times for the curves (from bottom to top) are thus $t = 0, 2.67, 5.33, 8.00, 10.67, 13.33, 16.00, 18.67$ h. The interfacial displacement η has been made dimensionless by dividing by the undisturbed layer thickness H.

the interior of the basin and most different at the edge of the basin, a primary difference between panels **(a)** and **(b)** in Figure 3.19 is in the position of the main wave front. With the parabolic bottom topography, the long internal wave speed $\sqrt{g'H(x,y)}$ is lower near the edge of the basin than in the interior (essentially a WKB approximation due to the gradual change of depth), therefore we expect the near-edge wave front to travel more slowly than in the flat bottom case. Furthermore, a close comparison of the nonlinear "filament" waves (Figure 3.19 **(c)** and **(d)**) that radiate towards the interior of the basin from the primary wave front reveals that these waves undergo wave refraction as they cross depth isolines. This effect is also expected, since these waves are essentially propagating from a slow medium to a fast one (from the edge of the lake to the interior).

3.4 Pseudospectral methods for two-layer models

The reduced-gravity models for internal waves discussed in the above sections may not be adequate models of internal waves in general since they do not account for baroclinic effects, such as the exchange of momentum between fluid layers [48]. The question as to whether the methods of this chapter can be extended to two-layer weakly non-hydrostatic equation sets is thus an important one to answer. If a free surface is to be retained in the physics, one possibility is the two-layer system solved by de la Fuente *et al.* [28] using finite volume methods

$$\frac{\partial h_i}{\partial t} + \nabla \cdot (h\mathbf{u}_i) = 0 \,, \quad i = 1,2 \,, \tag{3.118}$$

$$\begin{aligned} \frac{\partial (h_1\mathbf{u}_1)}{\partial t} + \nabla \cdot (h_1\mathbf{u}_1 \otimes \mathbf{u}_1) &= -gh_1\nabla\eta + fh_1\mathbf{u}_1 \times \mathbf{k} \\ &+ \alpha_1 \nabla \left(\nabla \cdot (h_1\mathbf{u}_1)_t\right) \,, \end{aligned} \tag{3.119}$$

$$\begin{aligned} \frac{\partial (h_2\mathbf{u}_2)}{\partial t} + \nabla \cdot (h_2\mathbf{u}_2 \otimes \mathbf{u}_2) &= -gh_2\nabla\eta - g'h_2\nabla\zeta + fh_2\mathbf{u}_2 \times \mathbf{k} \\ &+ \alpha_2 \nabla \left(\nabla \cdot (h_2\mathbf{u}_2)_t\right) \,, \end{aligned} \tag{3.120}$$

where $\alpha_1 = H_1^2/6$ and $\alpha_2 = H_2^2/3 + H_1H_2/2$. An inspection of these equations suggests that the methodologies discussed for the one-layer system generalize quite well for this two-layer system. The fact that the non-hydrostatic terms do not introduce additional coupling between the two layers implies that an auxiliary elliptic problem for the dispersive terms may be derived for each layer, resulting in two Helmholtz-type problems to be solved at each time-step. Indeed, this approach has been implemented by the author for doubly-periodic domains (not shown).

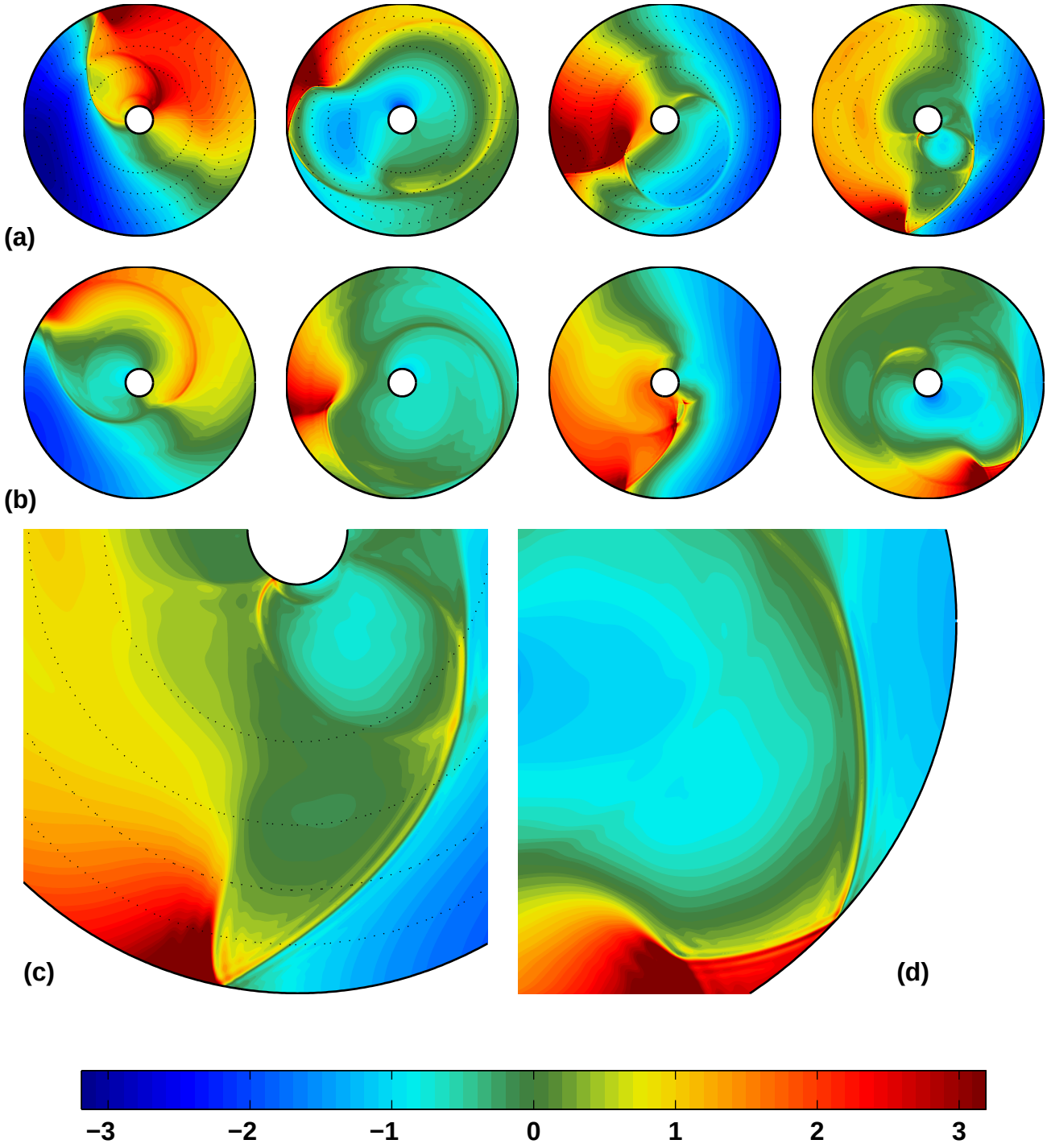

Figure 3.19: **(a)** Selected snapshots of the η field in the *IKS* test-case with parabolic bathymetry and $\eta_0^{max} = 0.25 H_{max}$ at times $t = 22.84, 25.53, 28.22, 30.91$ h. Four (dotted) depth contours are super-imposed on the plot, to indicate the shape of the bathymetry. **(b)** The same snapshots as **(a)** but without topography, i.e., the same run as Figure 3.17 but at later times. Panels **(c)** and **(d)** show a magnified comparison between the with-topography and no-topography runs at $t = 30.91$ h.

However, one problem with retaining a free-surface is that the time-step imposed by the surface gravity waves will be much smaller than the time-step imposed by the long internal gravity wave speed as a result of the CFL condition [35], i.e.,

$$\Delta t_{bt} < \frac{\Delta x}{\sqrt{gH}} \ll \Delta t_{bc} < \frac{\Delta x}{\sqrt{g' H_e}} \,, \tag{3.121}$$

where $H_e = H_1 H_2/(H_1 + H_2)$, and it is possible that this time-step is simply too restrictive to simulate the relevant dynamics of the internal wave field using a reasonable amount of computational effort. One possibility is to implicitly time-step the terms responsible for the propagation of the fast gravity waves to allow for a scheme that is stable for larger time-steps but would not accurately propagate the surface waves. Another technique that is very popular in three-dimensional ocean modelling [44], since it does not overly degrade the surface wave field, is called "mode splitting" where the equations are depth-averaged and a set of equations similar to the one-layer shallow water equations are time-stepped forward with a series of small "barotropic steps." Once the free-surface field is obtained at the later time, it can be substituted into the other equations to time-step the three-dimensional fields with the larger "baroclinic step." Though Cushman-Roisin and Beckers [26] have noted that mode splitting can also be applied to the multi-layer shallow water equations, we have not explored this idea in this thesis since it is not clear if an appreciable amount of computational effort can be saved for two-layer systems that also contain dispersive terms. The significance of retaining the free surface in a layered model can be further called into question by noting that since the free surface is represented as a function of the horizontal coordinates $\eta = \eta(x, y, t)$, the complex physics of wind waves (e.g., overturning and surface tension effects) will be misrepresented.

In the present work, we have opted to employ the approach of "filtering out" the fast gravity waves by invoking the rigid-lid approximation discussed in Section 2.10. Although this technique introduces the rigid-lid pressure that can only be calculated by solving an elliptic problem to enforce the incompressibility constraint on the barotropic mode, we find this approach to lead to a simple and efficient two-layer numerical model of internal waves that can be solved by employing the well-studied set of numerical techniques for incompressible flow as described in the following sections.

3.4.1 Splitting/Projection Time-stepping Method for the 2-layer rigid lid model

In this section we discuss how an augmented form of the widely-used splitting/projection method of Karniadakis *et al.* [67] can be applied to the inviscid weakly non-hydrostatic

two-layer rigid lid equations in the context of pseudospectral methods. The equations are given by (2.196)-(2.200), and we restate them here as

$$\frac{\partial h_i}{\partial t} + \nabla \cdot (h_i \mathbf{u}_i) = 0 \ , \quad i = 1, 2 \ , \tag{3.122}$$

$$\frac{\partial (h_1 u_1)}{\partial t} + \nabla \cdot ((h_1 u_1)\mathbf{u}_1) = -h_1 \frac{\partial p}{\partial x} + f h_1 v_1 \ , \tag{3.123}$$

$$\frac{\partial (h_1 v_1)}{\partial t} + \nabla \cdot ((h_1 v_1)\mathbf{u}_1) = -h_1 \frac{\partial p}{\partial y} - f h_1 u_1 \ , \tag{3.124}$$

$$\begin{aligned} \frac{\partial (h_2 u_2)}{\partial t} + \nabla \cdot ((h_2 u_2)\mathbf{u}_2) &= -g' h_2 \frac{\partial \zeta}{\partial x} - h_2 \frac{\partial p}{\partial x} \\ &+ f h_2 v_2 + \gamma \nabla \cdot (h_2 \mathbf{u}_2)_{xt} \ , \end{aligned} \tag{3.125}$$

$$\begin{aligned} \frac{\partial (h_2 v_2)}{\partial t} + \nabla \cdot ((h_2 v_2)\mathbf{u}_2) &= -g' h_2 \frac{\partial \zeta}{\partial y} - h_2 \frac{\partial p}{\partial y} \\ &- f h_2 u_2 + \gamma \nabla \cdot (h_2 \mathbf{u}_2)_{yt} \ , \end{aligned} \tag{3.126}$$

Once again, we will assume the equations have been discretized in time and we are charged with the task of evolving the physical fields from time level $t = t_n$ to $t = t_{n+1}$. For notational convenience, we assume the spatial coordinates are continuous and focus on the equations in semi-discrete form with the understanding that the spatial discretization approach discussed in section 3.2 applies here as well.

The first step in the splitting algorithm is the advective step where the nonlinear advection and source terms are evolved using a linear multi-step method

$$\mathbf{q}^\dagger = \mathbf{q}^n + \Delta t \sum_{j=0}^{N-1} b_j \mathbf{R}(\mathbf{q}^{n-j}) \ , \tag{3.127}$$

Here, $\mathbf{q} = (h_1, h_1 u_1, h_1 v_1, h_2, h_2 u_2, h_2 v_2)^\mathsf{T}$ and

$$\mathbf{R}(\mathbf{q}) = \begin{pmatrix} -\nabla \cdot (h_1 \mathbf{u}_1) \\ -\nabla \cdot (h_1 u_1 \mathbf{u}_1) + f h_1 v_1 \\ -\nabla \cdot (h_1 v_1 \mathbf{u}_1) - f h_1 u_1 \\ -\nabla \cdot (h_2 \mathbf{u}_2) \\ -\nabla \cdot (h_2 u_2 \mathbf{u}_2) + f h_2 v_2 - g h_2 \zeta_x \\ -\nabla \cdot (h_2 v_2 \mathbf{u}_2) - f h_2 u_2 - g h_2 \zeta_y \end{pmatrix} \ , \tag{3.128}$$

In the present work, we use the third-order Adams-Bashforth (AB3) method where $N = 3$ in equation (3.127) and the b_j's are given in Table 3.2. Since AB3 is not a self-starting

N	b_0	b_1	b_2
1	1		
2	3/2	-1/2	
3	23/12	-4/3	5/12

Table 3.2: Coefficients of the first three linear multi-step methods, obtained from Leveque [76].

scheme, time stepping is initialized using a Forward Euler step followed by an AB2 step, corresponding to $N = 1$ and $N = 2$ in Table 3.2, respectively.

Since the layer thicknesses h_1 and h_2 are effectively active tracers, we have completed their calculation at this stage. That is $h_{1,2}^{n+1} = h_{1,2}^{\dagger}$. In fact, some computational work can be saved in the advective step by realizing that once one of h_1 or h_2 is known, the other one is automatically known via the rigid-lid constraint. For example, if we compute h_1^{n+1} then

$$h_2^{n+1} = h - h_1^{n+1} \ , \tag{3.129}$$

since the total layer thickness $h = h(x, y)$ is time-independent (see Section 2.10). Hence, only one of the continuity equations needs to be time-stepped using AB3. The interfacial displacement ζ^{n+1} can be recovered by rearranging one of the layer thickness definitions, i.e., either of

$$h_1^{n+1} = H_1 - \zeta^{n+1} \ , \text{ or} \tag{3.130}$$

$$h_2^{n+1} = H_2 + \zeta^{n+1} - b \tag{3.131}$$

$$= \widetilde{H}_2 + \zeta^{n+1} \ , \tag{3.132}$$

where H_1 and H_2 are constants, $b = b(x, y)$ is the displacement of the bottom topography (see Fig. 2.5), and $\widetilde{H}_2(x, y) = H_2 - b(x, y)$ is an alternative definition of the undisturbed thickness of the lower layer that accounts for the bathymetry.

The time-discretized momentum equations in each layer are thus

$$(h_1\mathbf{u}_1)^{n+1} = (h_1\mathbf{u}_1)^{\dagger} - \Delta t\, h_1^{n+1}\nabla p^{n+1} \ , \tag{3.133}$$

$$(h_2\mathbf{u}_2)^{n+1} = (h_2\mathbf{u}_2)^{\dagger} - \Delta t\, h_2^{n+1}\nabla p^{n+1} + \Delta t\gamma\nabla z^{\dagger} \ , \tag{3.134}$$

where we have substituted in the intermediate transport terms $(h_1\mathbf{u}_1)^{\dagger}$ and $(h_2\mathbf{u}_2)^{\dagger}$. Here, $z = \nabla\cdot(h_2\mathbf{u}_2)_t$ is the auxiliary non-hydrostatic pressure variable and p^{n+1} is the rigid-lid pressure that will enforce the incompressibilitiy constraint on the barotropic transport at the new time-step $\nabla\cdot((h_1\mathbf{u}_1)^{n+1} + (h_2\mathbf{u}_2)^{n+1}) = \nabla\cdot(h\mathbf{u}_{bt})^{n+1} = 0$.

At this point, it must be determined which order each of the pressure terms should be evolved in. To answer this question, we first note that it is essential for the incompressibility constraint $\nabla \cdot (h\mathbf{u}_{bt})^{n+1} = 0$ to be satisfied to numerical precision at the end of each time-step for the scheme to be consistent. Realizing that the non-hydrostatic term $\gamma \nabla z$ is proportional to a gradient, it is quite trivial to see that the weakly non-hydrostatic pressure term is compressible in general and evolving it after the rigid-lid pressure will not lead to a consistent scheme. Therefore, we evolve the weakly non-hydrostatic pressure term before enforcing incompressibility.

To evolve the non-hydrostatic pressure, we follow the same procedure presented in Section 3.2 and derive a wave continuity equation by taking the divergence of the continuous form of the lower layer momentum equations with the rigid-lid pressure ignored. Doing so gives

$$\nabla \cdot (\gamma \nabla z) - z = -\nabla \cdot \mathbf{a} \,, \tag{3.135}$$

where

$$\mathbf{a} = \begin{pmatrix} -\nabla \cdot (h_2 u_2 \mathbf{u}_2) + f h_2 v_2 - g h_2 \zeta_x \\ -\nabla \cdot (h_2 v_2 \mathbf{u}_2) - f h_2 u_2 - g h_2 \zeta_y \end{pmatrix} \,, \tag{3.136}$$

i.e., the last two entries of the vector $\mathbf{R}(\mathbf{q})$. In our scheme, the vector $\mathbf{a}$ is evaluated using $\mathbf{q}^\dagger$. Equation (3.135) can then be inverted to compute $z^\dagger$ using the same methods described in Sections 3.2 and 3.3. The intermediate transport in the lower layer is then updated via

$$(h_2 \mathbf{u}_2)^\star = (h_2 \mathbf{u}_2)^\dagger + \Delta t \gamma \nabla z^\dagger \,, \tag{3.137}$$

and for notational convenience we let $(h_1 \mathbf{u}_1)^\star = (h_1 \mathbf{u}_1)^\dagger$. The starred transports can be thought of as "predicted" transports that must be "corrected" by projecting the predicted barotropic transport onto the space of divergence-free vector fields. This procedure is directly analogous to the commonly used splitting/projection method for the incompressible Euler and Navier–Stokes equations [67, 1].

The projection method can be derived by considering the pressure step

$$(h_1 \mathbf{u}_1)^{n+1} = (h_1 \mathbf{u}_1)^* - \Delta t h_1^{n+1} \nabla p^* \,, \tag{3.138}$$

$$(h_2 \mathbf{u}_2)^{n+1} = (h_2 \mathbf{u}_2)^* - \Delta t h_2^{n+1} \nabla p^* \,. \tag{3.139}$$

Adding these two equations and recalling that $h\mathbf{u}_{bt} = h_1 \mathbf{u}_1 + h_2 \mathbf{u}_2$, we have

$$(h\mathbf{u}_{bt})^{n+1} = (h\mathbf{u}_{bt})^* - \Delta t h \nabla p^* \,, \tag{3.140}$$

and we have once again used the fact that $h_1 + h_2 = h$, is constant in time. Taking the divergence of this equation for the barotropic transport gives

$$\frac{1}{\Delta t} \left[\nabla \cdot (h\mathbf{u}_{bt})^{n+1} - \nabla \cdot (h\mathbf{u}_{bt})^* \right] = -\nabla \cdot \left(h \nabla p^{n+1} \right) \,. \tag{3.141}$$

To project onto the space of solutions that conserve mass, we employ the rigid-lid constraint (2.201) at the new time-step, giving

$$\frac{1}{\Delta t}\nabla\cdot(h\mathbf{u}_{bt})^{*} = \nabla\cdot\left(h\nabla p^{n+1}\right)\ . \tag{3.142}$$

This elliptic equation is solved for p^{n+1}, and used in (3.138)-(3.139) to recover the transport in each layer at the new time-step $((h_1\mathbf{u}_1)^{n+1}, (h_2\mathbf{u}_2)^{n+1})$.

For domains with solid wall boundaries, a suitable boundary condition on p^{n+1} can be derived by taking the dot product of (3.140) with the unit outward normal $\hat{\mathbf{n}}$ and enforcing the no normal flow condition $(h\mathbf{u}_{bt})^{n+1}\cdot\hat{\mathbf{n}} = 0$ on the boundary $\partial\Omega$, yielding

$$\frac{\partial p^{n+1}}{\partial n} = \frac{1}{\Delta t}\mathbf{u}^{*}_{bt}\cdot\hat{\mathbf{n}}\ , \quad \text{on} \quad \partial\Omega\ . \tag{3.143}$$

We have hence described all of the steps that are necessary to evolve the two-layer weakly non-hydrostatic model equations forwards in time. As it turns out, the solution of the Poisson–Neumann problem (3.142)–(3.143) is somewhat non-trivial in comparison to the Helmholtz type problem (3.135), and special attention is paid to it below.

3.4.2 Solvability and Compatibility of the Poisson–Neumann Problem

It is a well-known issue that when the domain consists of only solid-wall or periodic boundaries (or a combination thereof), the solution to the Poisson-Neumann problem (3.142)–(3.143) is only determined up to an additive constant, i.e., there is no unique solution. The non-uniqueness trait stems from the fact that in incompressible flow there is no thermodynamic definition for pressure, as is the case for compressible flow [70]. Hence, only the *gradient* of pressure is a physically meaningful quantity.

The question then arises as to how to modify the Poisson–Neumann problem to guarantee a unique solution for p since the issue of incompatibility between the Poisson equation and its boundary condition may arise due to such a modification [124]. To see this, consider the Poisson problem re-written as

$$\nabla\cdot(h\nabla p) = f\ , \quad \text{on} \quad \Omega\ , \tag{3.144}$$

$$h\nabla p\cdot\hat{\mathbf{n}} = g\ , \quad \text{on} \quad \partial\Omega\ . \tag{3.145}$$

Integrating each of these equations gives

$$\iint_{\Omega} \nabla \cdot (h\nabla p)\, dS = \iint_{\Omega} f\, dS\,, \tag{3.146}$$

$$\int_{\partial\Omega} h\nabla p \cdot \hat{\mathbf{n}}\, ds = \int_{\partial\Omega} g\, ds\,. \tag{3.147}$$

Now, due to Gauss' divergence theorem we must have

$$\iint_{\Omega} \nabla \cdot (h\nabla p)\, dS = \int_{\partial\Omega} h\nabla p \cdot \hat{\mathbf{n}}\, ds\,. \tag{3.148}$$

It follows that f and g must satisfy

$$\iint_{\Omega} f\, dS = \int_{\partial\Omega} g\, ds\,, \tag{3.149}$$

a constraint called the *compatibility condition*. A simple technique to guarantee a unique solution for p would be to impose a zero-Dirichlet condition on p at a single grid point on the boundary. However, as a result p would not satisfy the discrete Poisson equation at this point, and as a consequence, incompressibility would not be satisfied at the point [117].

In this work, we have employed the remedy suggested in [124], and used successfully in the context of pseudospectral methods by Subich [117, 118], of adding a small unknown quantity, σ to (3.144) and (3.145), i.e.,

$$\nabla \cdot (h\nabla p) = f - \sigma\,, \quad \text{on} \quad \Omega\,, \tag{3.150}$$

$$h\nabla p \cdot \hat{\mathbf{n}} = g - \sigma\,, \quad \text{on} \quad \partial\Omega\,, \tag{3.151}$$

to impose the additional constraint of zero mean pressure

$$\iint_{\Omega} p\, dS = 0\,, \tag{3.152}$$

chosen to guarantee a unique solution. The result of using the augmented Poisson–Neumann equation (3.150)–(3.152) is that the effect of any incompatibility is shared by all of the grid points in the domain (and not just one single point). Thus, as long as the

magnitude of σ is small (i.e., below discretization error), we expect there to be no ill effects on the numerical solution.

Implementing this technique numerically amounts to "bordering" the linear system of equations for p. That is, if we start with (3.144)–(3.145) discretized as, for example, the matrix-vector equation

$$L\mathbf{p} = \mathbf{r} \,, \tag{3.153}$$

where L is and $N \times N$ matrix, and $\mathbf{p}$ and $\mathbf{r}$ are $N \times 1$ vectors, then we would instead use the bordered system

$$\begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1N} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ L_{N1} & L_{N2} & \cdots & L_{NN} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_N \\ \sigma \end{pmatrix} = \begin{pmatrix} r_1 \\ \vdots \\ r_N \\ 0 \end{pmatrix} \,. \tag{3.154}$$

For simplicity, we have used the sum of the grid points of p instead of a discretized version of the integral constraint (3.152) that would replace the bottom row with a set of quadrature weights associated with each grid point. This alternate constraint is a sensible choice since there is nothing special or physical about the zero mean pressure constraint, and it was merely chosen as a constraint that incorporates all of the grid-points of p, as does

$$\sum_{i=1}^{N} p_i = 0 \,. \tag{3.155}$$

Augmenting the Poisson–Neumann operator in this manner can also be achieved using iterative methods, and this approach has been successfully carried out using GMRES in our MATLAB implementation of the solver described in Section 3.4.1. To lower the required number of iterations, a pre-conditioner has been constructed using the bordered matrix form of the second-order finite differences approximation to the Poisson–Neumann operator.

While the method of "bordering" the linear system offers a simple strategy for addressing the solvability issue, it is not the only technique available. Recent work by Escobar-Vargas [38] gives a detailed review of other strategies in the context of the spectral multidomain penalty method. That work advocates slightly modifying the right-hand side such that the solution vector will be orthogonal to the left null singular vector of the Poisson-Neumann operator, so as to guarantee both a solvable and strongly consistent linear system.

3.4.3 Simulations of Finite–Amplitude Wave Propagation in a two-layer channel

In this section, we apply the numerical solution technique for the rigid-lid system described above to study the evolution of an interfacial disturbance of finite amplitude propagating along a wall in a two-layer fluid in a periodic channel. These simulations are not unlike those carried out for a 1-layer fluid in an annular domain in Section 3.3.5. Here, we employ the Fourier method in the (periodic) y-direction, and use a Chebyshev method in the x-direction to impose no normal flow at the walls. Rotation was turned off ($f = 0$), and the reduced gravity was taken to be $g' = 0.02$ ms^{-2}. The domain was taken to be 3000 m $\times$ 3000 m in extent. The spectral filtering parameters were set to $k_{crit} = 0.4k_{max}$, $s = 4$ in all cases.

In our first simulation, we take the undisturbed layer thicknesses as $H_1 = 2.5$ m in the upper layer and $H_2 = 7.5$ m in the lower layer. The initial conditions are given by:

$$\zeta^0 = -a_0 \mathrm{sech}((y - \frac{L_y}{2})/250)^2 \sin(\frac{\pi y}{L_y}) \mathrm{e}^{-\frac{x}{800}} , \tag{3.156}$$

$$v_1^0 = (c_0/H_1)\zeta^0 , \tag{3.157}$$

$$v_2^0 = -(c_0/H_2)\zeta^0 , \tag{3.158}$$

$$u_1^0 = 0 , \tag{3.159}$$

$$u_2^0 = 0 , \tag{3.160}$$

where the maximum amplitude is given by $a_0 = 2$ m, and $c_0 = \sqrt{g'H_e}$ is the long linear internal wave speed. This initial profile reflects a depressed interfacial perturbation initialized to propagate in the negative y direction (according to linear theory). Fixed time snapshots of the interface displacement field ζ are shown in Figure 3.20, beginning with the initial condition in panel **(a)**.

Inspecting Figure 3.20, we notice that as the disturbance steepens, a bend appears in the main wave-front by $t = 1.09$ h (panel **(b)**. This bending effect is explained by noting that since the wave is of greater amplitude nearest the wall, it follows that its effective propagation speed will be fastest there and slower moving away from the wall. By $t = 3.28$ h (panel **(d)**, the disturbance appears to have clearly separated into a set of distinct solitary wave of depression ordered by their respective amplitudes from largest to smallest. As the train of solitary waves continues to propagate in the negative y direction, it is apparent that energy is being lost to cross-channel wave radiation that eventually forms a secondary disturbance along the opposite channel wall at $x = 3000$ m (see panel

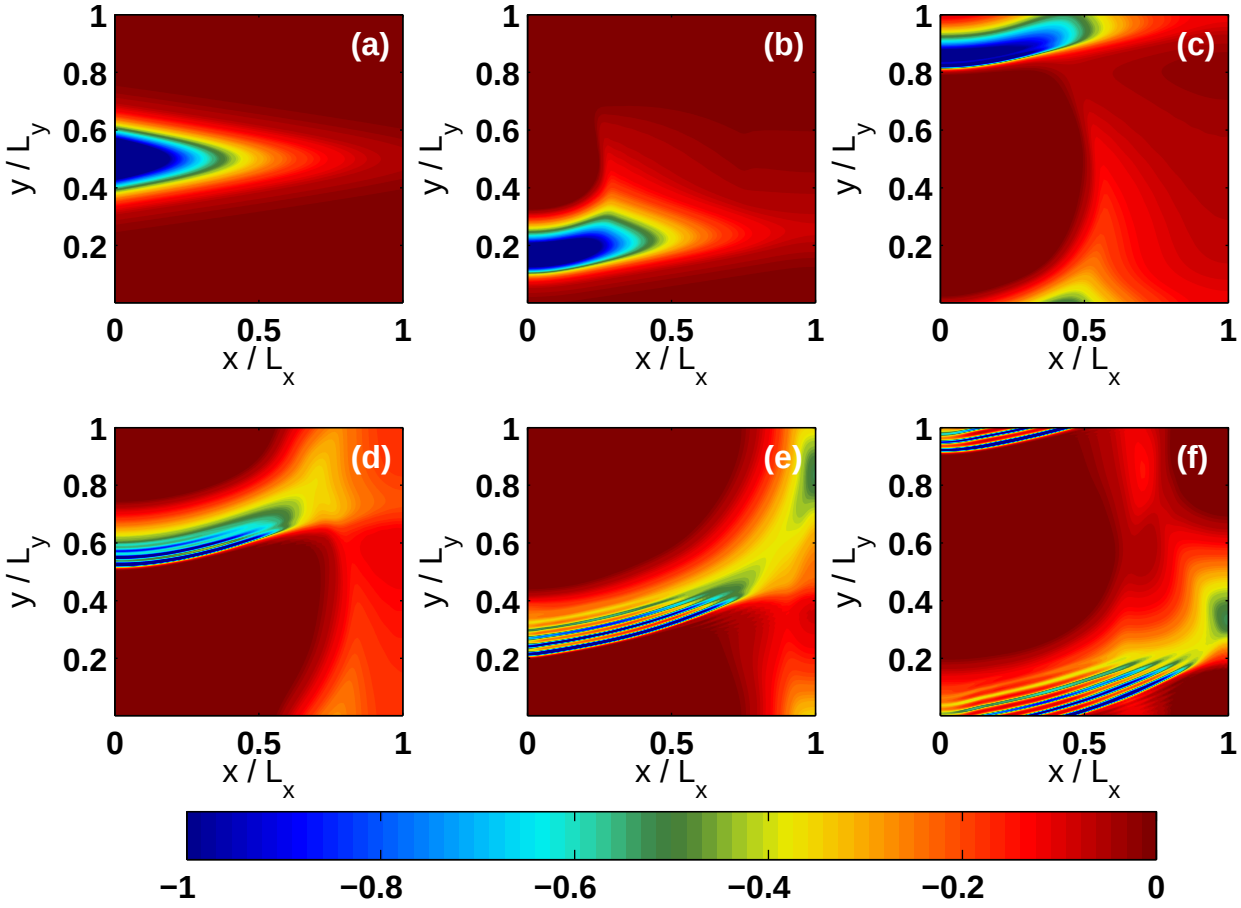

Figure 3.20: Fixed time snapshots of the interfacial displacement at **(a)** $t = 0$, **(b)** $t = 1.09$ h, **(c)** $t = 2.19$ h, **(d)** $t = 3.28$ h, **(e)** $t = 4.37$ h, and **(f)** $t = 5.47$ h in the *2-layer wave propagation in a channel* run. Here, $H_1 = 2.5$ m and $H_2 = 7.5$ m.

(**e**)). The solitary wavetrain and the secondary disturbance appear to continue to interact via wave radiation as the simulation unfolds.

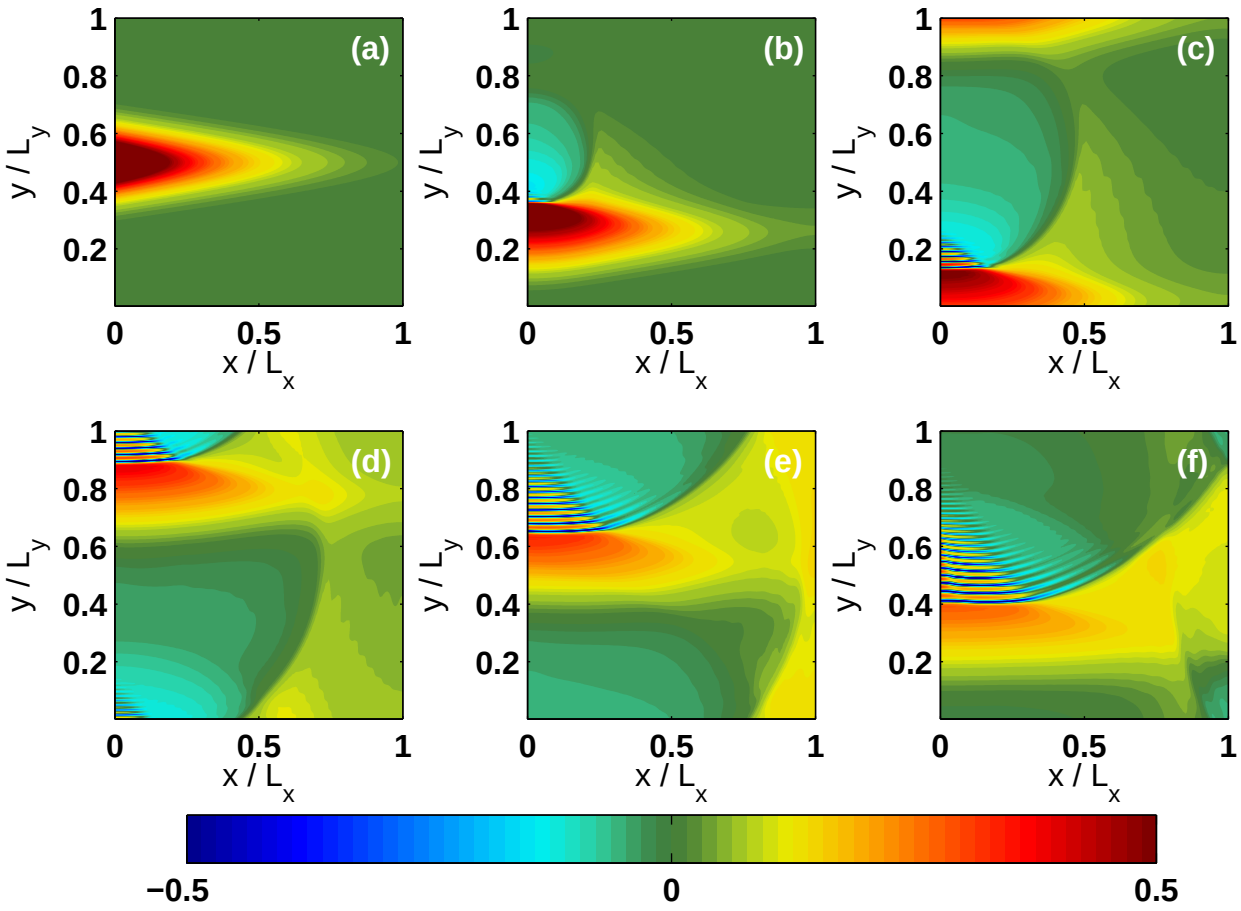

Figure 3.21: Like Figure 3.20, but the sign of ζ^0 was taken to be positive (polarity reversed from Fig. 3.20, and its amplitude was reduced by a factor of 2, i.e., $a_0 = 1$ m.

In Figure 3.21, we consider a very similar situation to that of Figure 3.20, but we have reversed the sign of the interfacial disturbance (such that it is a wave of elevation) and reduced its amplitude by a factor of 2. Examining the panels in Figure 3.21, we notice an evolution that is strikingly different from that of Figure 3.20. By panel (**b**)

it is apparent that the disturbance is steepening towards its *rear* and that the front of the wave is losing energy to radial spreading. Tracking the evolution further in panels **(c)**–**(f)** it is apparent that a collection of solitary waves is not emerging, but rather a completely different dispersive and nonlinear phenomenon is taking place. The elevated rear of the wave has steepened to the point of forming a smoothed shock (or hydraulic jump) that abruptly shoots below the $\zeta = 0$ level and is followed by a series of oscillations (or undulations) of successively smaller amplitude and wavelength.

This small-scale internal wave phenomenon is known as an *undular bore* and has been observed in relatively small-sized lakes such as Loch Ness [36]. A theoretical description of the undular bore as a modulated periodic wavetrain has been established using KdV theory by Gurevich and Pitaevskii [56] and Whitham [130]. The theory also provides insight into the long-time evolution of the undular bore. In particular, it is known that the wave envelope of the bore spreads out in space, and as $t \to \infty$ each wave crest will become a solitary wave imposed on a new background state [56]. This long-time behaviour has not been explored in the simulations herein, but represents an interesting avenue for future research.

A general heuristic (from KdV theory) for predicting which type of nonlinear waves will emerge in a two-layer fluid under the Boussinesq approximation is that if the lower (upper) layer is thicker, then solitary waves of depression (elevation) may emerge [130]. Since the lower layer is thicker in Figures 3.20 and 3.21, we find that an initially depressed perturbation has lead to solitary waves while the elevated perturbation has instead yielded an undular bore. These theoretical predictions have been further tested in Figures 3.22 and 3.23 where the simulations are directly analogous to those in Figs. 3.20 and 3.21, respectively, with the exception that the undisturbed layer thicknesses, H_1 and H_2, have been swapped so that the upper layer is now the thicker layer. As expected, the depressed initial condition now leads to the formation of an undular bore while the elevated disturbance breaks down into solitary waves.

3.5 Chapter summary and conclusions

In this chapter, we have discussed solving dispersive shallow water models of the Boussinesq-type using Fourier and Chebyshev pseudospectral methods. In line with previous studies [42, 68], we discussed two approaches for the time-discretization method in Section 3.1, the so-called "coupled" and "scalar" approaches. Although both methods are stable, the scalar approach reduces the dimension of the resulting linear systems to be solved by a factor of two, and transforms the problem of time-stepping mixed space-time derivatives to a

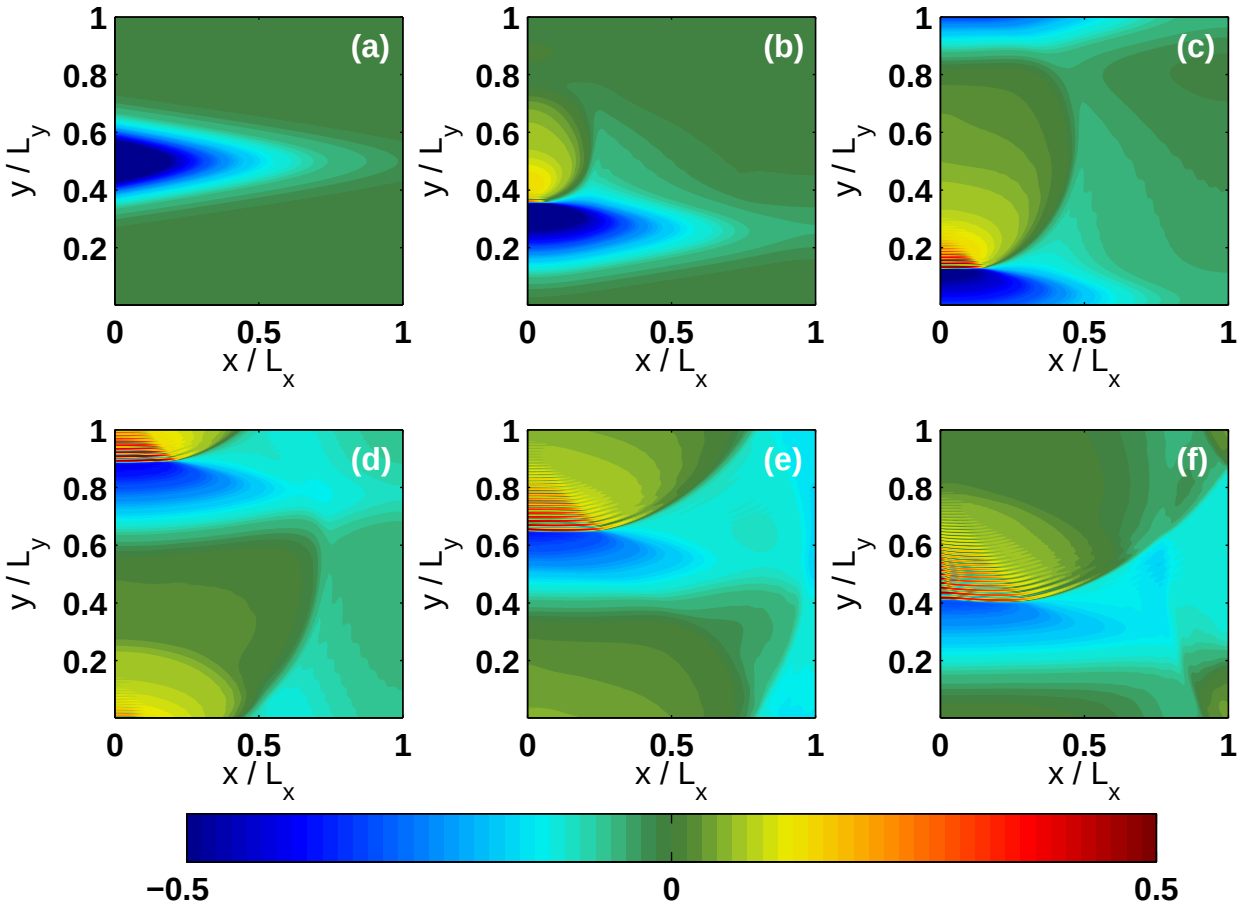

Figure 3.22: Like Fig. 3.20, but the undisturbed layer thicknesses have been switched, i.e., $H_1 = 7.5$ m and $H_2 = 2.5$ m, and $a_0 = 1$ m.

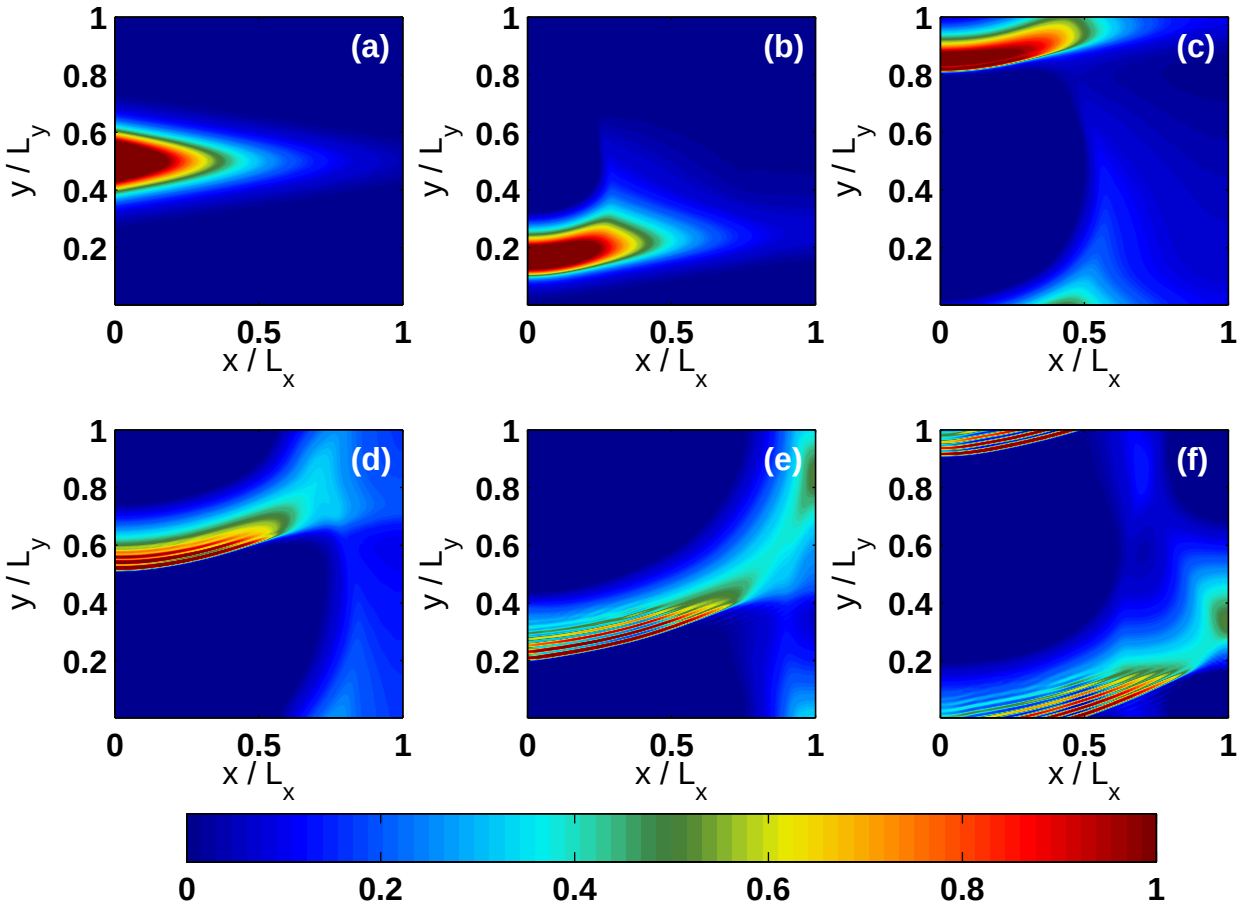

Figure 3.23: Like Fig. 3.21, but the undisturbed layer thicknesses have been switched, i.e., $H_1 = 7.5$ m and $H_2 = 2.5$ m, and $a_0 = 2$ m.

familiar pressure-type elliptic problem. In Section 3.2, we employed the Fourier method to solve the single-layer equations of motion on periodic domains with variable depth. Practical details of implementation were discussed in Sections 3.2.1–3.2.3, including details of obtaining efficient solutions to the aforementioned linear systems with numerical linear algebra techniques and pre-conditioning, or discrete Fourier transforms where appropriate. Other practical considerations, such as filter stabilization of aliasing/nonlinearity-driven numerical instabilities were outlined as well. In light of these methods presented, it is clear that FFT-based methods can be extended to problems involving variable bathymetry and can also be a highly-accurate means of solving elliptic problems with variable coefficients if used in conjunction with iterative linear system solvers and pre-conditioning.

Our numerical methodology was validated in one dimension against approximate analytical solutions for the cases of dispersive short waves over a flat-bottom in Section 3.2.7 and long waves over a slowly varying bottom in Section 3.2.6. Very fast convergence rates of the Fourier spatial discretization method was demonstrated in two-dimensions by comparing numerical solutions to the exact solution of an elliptic problem in Section 3.2.5.

Grid convergence of the Fourier method was illustrated in Section 3.2.8 for the test-case of a long wave steepening and propagating over topography leading to the emergence of solitary waves. This test case was important because it showed that in the well-resolved limit the numerical model is accurate in situations where both dispersion and nonlinearity are prevalent in the dynamics.

These results indicate that the Fourier method is an excellent choice of benchmark for lower-order methods (DG-FEM, FVM) that can be used in much more general geometries than the Fourier method. Furthermore, the high accuracy of the Fourier method allows classical water-wave solutions to be explored without the uncertainty associated with the numerical dissipation inherent in low-order methods.

A two-dimensional dynamical simulation of waves driven by flow over topography was carried out in Section 3.2.9 to illustrate how the proposed numerical model may be used in practical physical limnology problems. A set of rich wave dynamics, including topographically-trapped waves, upstream propagating waves, and waves radiating in the cross-stream direction, was observed. Our results agreed qualitatively with past analytical and numerical results of resonant wave generation by flow over topography [55, 43].

A second two-dimensional simulation corresponding to a long wave propagating over a shoal was carried out in Section 3.2.10. Interesting linear and nonlinear phenomena such as wave scattering, steepening, and the emergence of fissioning solitary waves (only in sufficiently shallow regions) rapidly broke the symmetry of the initial conditions resulting in a rather complicated final wave field with a variety of fine scale features.

A pseudospectral method was then presented for solving the one-layer weakly non-hydrostatic rotating shallow water system in annular domains in Section 3.3. The methodology employs a Fourier pseudospectral spatial discretization in the azimuthal direction and a Chebyshev pseudospectral spatial discretization in the radial direction after the model equations have been transformed to polar coordinates (r, θ). It was explained how only minimal modifications needed to be made to the method presented in 3.2 to deal with the annular geometry.

The resulting model of annular lakes was shown to be useful in performing simulations of internal wave dynamics in mid-sized lakes during the ice-free seasons when lakes are expected to possess a thermal stratification and (depending on their size) be affected by the Earth's rotation. The robustness of the numerical model was illustrated in Sections 3.3.4–3.3.6 by considering different wave dynamical situations corresponding to different sets of initial conditions while varying the physical parameters of the equations over a wide range of physically-relevant choices. These parameter choices included: 1) the strength of nonlinearity, i.e., extent of wave steepening, 2) the strength of the Coriolis force, i.e., the angular frequency of the rotating frame of reference, and 3) the bottom topography of the circular lake. The main focus of our simulations was on the transfer of energy from large-(or basin-scale) waves to small-scale features such as dispersive wavetrains and high-frequency solitary-like waves.

The methods applied for the single-layer equations were extended to the case of a weakly non-hydrostatic two-layer fluid under the rigid-lid approximation in Section 3.4. In particular, we explained how the well-known splitting/projection time-stepping method of Karniadakis *et al.* [67] for incompressible flow can be applied to evolve the rigid-lid equations forward in time in Section 3.4.1. Two-dimensional simulations of a finite amplitude interfacial perturbation propagating along the wall of a periodic channel were then carried out in Section 3.4.3. In addition to solitary waves, the undular bore was also observed as another nonlinear internal wave phenomenon that can be modelled by the two-layer weakly non-hydrostatic rigid-lid equations.

The results of the simulations shown in this chapter appeared to agree well with intuition, well-known observations of wave dynamics in lakes and experiments [28, 132, 74, 36], as well as past numerical simulations [15, 109, 59, 36] and mathematical theory [115, 130, 56] of internal waves. This agreement with previous results in the literature suggests that the models presented here may well prove useful in practical physical limnology applications. In particular, the lack of inherent numerical dissipation allows for the construction of a set of rational hypotheses regarding the behaviour of waves in mid-sized lakes that can be subsequently tested against field data.

Chapter 4

Mid- to High-Order Discontinuous Galerkin Methods for One-Layer Dispersive Models in Complex Geometries

Since a lake's coastal boundary generally specifies a physical domain with complex/irregular boundaries, the methods presented in section 3.2 and 3.3 are not sufficient for modelling real-world lakes. To represent more general geometries, we now turn to the discontinuous Galerkin finite element method (DG-FEM) as a high-order alternative to the low-order finite volume and finite element methods that are commonly used for irregular geometries. The results presented in this chapter typically use local polynomial orders between $N = 4$ and $N = 8$, the methods are thus high-order in contrast to traditional finite element methods that typically use piece-wise linear or quadratic basis functions. See for instance Walkley [126], who solved a Boussinesq-type system with a low-order finite element method (FEM).

It is worth stressing that the high-order DG-FEM is not the same as the spectral element method (SEM) [68] that represents the high-order extension of the traditional FEM. Both FEM and SEM are continuous Galkerin formulation which requires C^0 continuity at element interfaces. Although DG-FEM and SEM both use a high-order orthogonal polynomial basis, the DG-FEM only imposes continuity in a weak sense through the specification of a numerical flux function at element edges in order to allow for stable advection schemes [20, 18, 60]. The requirement of C^0 continuity in the SEM means that the method is not

ideal for advection problems since an upwind-type scheme cannot be formulated to account for the preferred direction of propagation of information [60]. This shortcoming can lead to situations where Gibbs oscillations are trapped at element interfaces, as has been illustrated for the spectral element ocean model by Levin *et al.* [77]. However, stabilization and upwinding techniques do exist for continuous Galerkin formulations, as has been discussed by Nazarov and Popov [88] and Hughes [62], for example.

The specification of an upwind-biased numerical flux is usually furnished through the well-established theory on approximate Riemann solvers [121] that are commonly used in the formulation of finite volume methods [75] in order to propagate information between finite volume cells. It is for this reason that DG-FEM with piece-wise constant basis functions (order $N = 0$) is identical to the low-order finite volume method [60].

In the following sections, we follow the techniques and developments for nodal DG-FEM as presented by Hesthaven and Warburton [60], and we build upon their techniques where necessary. In the following sections, we explain the basic nodal DG-FEM formulation as the spatial discretization method for both hyperbolic and elliptic systems and the corresponding reduction to local operators in the context of the one-layer dispersive shallow water model (2.120)-(2.122). Following this, comparisons with the pseudospectral methods of Chapter 3 are carried out in 1D and 2D as a means of validating the numerical schemes and illustrating the resolution characteristics of the DG-FEM at varying polynomial orders. The necessity of curvilinear elements for general situations is illustrated due to singular/spurious flow features that can emerge due to a piece-wise linear representation of the boundary. It is then explained how the nodal DG-FEM method should be augmented with high-order cubature and quadrature integration rules to deal with the non-constant mapping Jacobians introduced by curvilinear elements. Highlight simulations are provided alongside the developments, and we conclude with a simulation in a real-world lake: Pinehurst Lake, Alberta.

4.1 Discontinuous Galerkin finite element method for the one-layer equations

Since the DG-FEM method is primarily suited to solving hyperbolic conservation laws, it is useful to consider the augmented system (2.120)-(2.122) in the conservation form

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial y} = \mathbf{B} + \mathbf{C} + \mathbf{N} \,, \tag{4.1}$$

with

$$\mathbf{Q} = \begin{pmatrix} h \\ hu \\ hv \end{pmatrix}, \ \mathbf{F} = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{pmatrix}, \ \mathbf{G} = \begin{pmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{pmatrix}. \tag{4.2}$$

The terms

$$\mathbf{B} = gh \begin{pmatrix} 0 \\ \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial y} \end{pmatrix}, \ \mathbf{C} = f \begin{pmatrix} 0 \\ vh \\ -uh \end{pmatrix}, \ \mathbf{N} = \frac{H^2}{6} \begin{pmatrix} 0 \\ \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \end{pmatrix}, \tag{4.3}$$

are the bed slopes, Coriolis terms, and the dispersive terms, respectively. As explained in 3.1.1, The variable $z = \nabla \cdot (\mathbf{u}h)_t$ is governed by the Helmholtz problem (3.10) restated here as

$$\nabla \cdot (\gamma \nabla z) - z = -\nabla \cdot \mathbf{a}\,, \tag{4.4}$$

with $\gamma = H^2/6$ and $\mathbf{a}$ is as defined previously by equation (3.11). In the case of the traditional non-rotating hydrostatic shallow water model with a flat bottom, the right-hand side of (4.1) vanishes.

To apply the DG-FEM method in two-dimensions, we first assume that the domain Ω can be triangulated using K elements (or sub-domains) and that the triangulation is assumed to be geometrically conforming. That is, the boundary $\partial\Omega$ is represented by a piece-wise linear approximation with each line segment belonging to a side of a triangle. The boundary can be approximated in a piece-wise curvilinear sense by using triangles with curved edges as explored later in Section 4.14, but this approach is typically much more expensive and does not lend itself easily to the nodal approach [60] owing mainly to the fact that the mapping Jacobian to the reference triangle is non-constant. Additionally, we assume that the nodes along a triangle edge that are shared between two elements are duplicated, so as to ensure that a purely local scheme can be recovered. This is a fundamental difference between DG-FEM and FEM, which uses shared nodes along a shared edge.

In each element $\mathbf{D}^k$, we form the approximate local solution $(h_h^k, (hu)_h^k, (hv)_h^k, z_h^k)$ with nodal representations

$$h_h^k(\mathbf{x}, t) = \sum_{i=1}^{N_p} h_h^k(\mathbf{x}_i^k, t)\ell_i^k(\mathbf{x})\,, \tag{4.5}$$

and similarly for the other fields. Here, $\ell_i^k(x)$ represents the i^{th} order two-dimensional Lagrange interpolating polynomial, $\mathbf{x} = (x, y)$, and N_p is the number of points within an element. It is assumed that this number is the same for all elements in the domain,

although this is not required. The $\mathbf{x}_i^k$'s refer to the local grid points on element $\mathbf{D}^k$ with a distribution that we leave unspecified for the time being.

The nodal DG-FEM weak integral form statement is obtained by substituting the approximate local solutions into each of equations (4.1), multiplying by a member of the space of local test functions $V_h^k = \{\ell_j^k\}_{j=1}^{N_p}$, and integrating the flux terms by parts. If we neglect the $\mathbf{B}$ and $\mathbf{N}$ terms and focus only on (4.1) for the moment, this gives

$$\int_{\mathbf{D}^k} \frac{\partial \mathbf{Q}_h^k}{\partial t}\ell_j^k - \mathbf{F}_h^k\frac{\partial \ell_j^k}{\partial x} - \mathbf{G}_h^k\frac{\partial \ell_j^k}{\partial y} - \mathbf{C}_h^k\ell_j^k d\mathbf{x} = -\int_{\partial \mathbf{D}^k} \ell_j^k\left(\mathbf{F}^*, \mathbf{G}^*\right)\cdot\hat{\mathbf{n}} d\mathbf{x} \tag{4.6}$$

where $\hat{\mathbf{n}}$ is the unit outward-pointing normal. Due to the fact that we do not require the solution to be continuous between elements, the value of $(\mathbf{F}, \mathbf{G})$ in the surface integral term on the right-hand side is not unique. Therefore, we have introduced $(\mathbf{F}^*, \mathbf{G}^*)$ as the numerical flux vector that represents some linear combination of information interior to the element $(\mathbf{F}^-, \mathbf{G}^-)$ and exterior information $(\mathbf{F}^+, \mathbf{G}^+)$. Since we have not explicitly imposed continuity at element interfaces, the numerical flux is our means for imposing continuity in a weak sense. Without it, the elements would completely decouple and a meaningful global solution would not be recovered. The numerical flux is typically chosen in a way that "mimics the flow of information in the underlying PDE" to ensure a stable and accurate scheme [60]. The choice of numerical flux mainly considered in this thesis is the local Lax–Friedrichs (L-F) flux

$$(\hat{n}_x\mathbf{F}_h + \hat{n}_y\mathbf{G}_h)^* = \hat{n}_x\{\!\{\mathbf{F}_h\}\!\} + \hat{n}_y\{\!\{\mathbf{G}_h\}\!\} + \frac{\lambda}{2}[\![\mathbf{Q}_h]\!]\,, \tag{4.7}$$

where

$$\{\!\{\mathbf{u}\}\!\} = \frac{\mathbf{u}^- + \mathbf{u}^+}{2}, \qquad [\![\mathbf{u}]\!] = \hat{\mathbf{n}}^-\cdot\mathbf{u}^- + \hat{\mathbf{n}}^+\cdot\mathbf{u}^+\,, \tag{4.8}$$

are the average and jump in $\mathbf{u}$ across the interface, respectively. The numerical flux choice (4.7) represents the local Lax-Friedrichs flux where λ is an approximation to the maximum linearized wave speed

$$\lambda = \max_{\mathbf{s}\in[\mathbf{Q}_h^-, \mathbf{Q}_h^+]}\left(\|\mathbf{u}(\mathbf{s})\| + \sqrt{gh(\mathbf{s})}\right)\,. \tag{4.9}$$

A sketch of the justification for using the Lax–Friedrichs flux was given in Section 2.3.3.

Returning to our weak statement (4.6), we can obtain the strong DG form by integrating the flux terms by parts again, yielding

$$\int_{\mathbf{D}^k}\left(\frac{\partial \mathbf{Q}_h^k}{\partial t} + \frac{\partial \mathbf{F}_h^k}{\partial x} + \frac{\partial \mathbf{G}_h^k}{\partial y} - \mathbf{C}_h^k\right)\ell_j^k d\mathbf{x} = -\int_{\partial \mathbf{D}^k}\ell_j^k(\mathbf{F}_h^k - \mathbf{F}^*, \mathbf{G}_h^k - \mathbf{G}^*)\cdot\hat{\mathbf{n}} d\mathbf{x}\,, \tag{4.10}$$

the weak and strong form are analytically equivalent, but for computational and conceptual reasons we mainly use the strong form in our numerical methods [60].

The "Galerkin approach" outlined above is motivated by attempting to make the residual, i.e., the function left over when the approximate solution is substituted into the PDE system, vanish in a general sense by imposing that the residual should be orthogonal to a space of test functions [60]. Methods that do not take the space of test functions to be the same as the set of basis functions are possible, but are referred to as the more general class of *weighted residual* methods [95]. For example, another possible choice of test functions may be the space of shifted dirac delta functions centered about a discrete set of points. The resulting method is known as the collocation method (e.g., see section 3.3) that requires the residual to identically vanish at the collocation points [95].

In order to reduce the statement (4.10) to a form useful for numerical computations, it is important to rewrite it in terms of matrices wherever possible. As an illustration, consider the first component of (4.10) with nodal expansions of the form (6.48) explicitly substituted in. We can then write this first component as

$$\mathcal{M}^k \frac{d\mathbf{h}^k}{dt} = -\mathcal{S}_x^k (\mathbf{hu})^k - \mathcal{S}_y^k (\mathbf{hv})^k + \int_{\partial \mathbf{D}^k} \ell_j^k \left((hu)_h^k - (hu)^*, (hv)_h^k - (hv)^* \right) \cdot \hat{\mathbf{n}} d\mathbf{x} \tag{4.11}$$

where

$$\mathbf{h}^k = \left[h_h^k(\mathbf{x}_1) \cdots h_h^k(\mathbf{x}_{N_p}) \right]^\mathsf{T} , \tag{4.12}$$

$$(\mathbf{hu})^k = \left[(hu)_h^k(\mathbf{x}_1) \cdots (hu)_h^k(\mathbf{x}_{N_p}) \right]^\mathsf{T} , \tag{4.13}$$

$$(\mathbf{hv})^k = \left[(hv)_h^k(\mathbf{x}_1) \cdots (hv)_h^k(\mathbf{x}_{N_p}) \right]^\mathsf{T} , \tag{4.14}$$

and we have left the surface integral contribution alone for now. Here the local mass matrix is given by

$$\mathcal{M}_{ij}^k = \int_{\mathbf{D}^k} \ell_i^k(\mathbf{x}) \ell_j^k(\mathbf{x}) d\mathbf{x} = J^k \int_{\mathbf{I}} \ell_i(\mathbf{r}) \ell_j(\mathbf{r}) d\mathbf{r} = J^k \mathcal{M} , \tag{4.15}$$

where $J^k = x_r^k y_s^k - x_s^k y_r^k$ is the (constant) Jacobian of the linear mapping from the element $\mathbf{D}^k$ to the reference element $\mathbf{I} = \{\mathbf{r} = (r,s) | (r,s) \geq -1; r+s \leq 0\}$, and we have also introduced the mass matrix on the reference triangle, $\mathcal{M}$.

The local stiffness matrix $\mathcal{S}_x^k$ is given as follows:

$$\mathcal{S}_{x,ij}^k = \int_{\mathbf{D}^k} \ell_i^k(\mathbf{x}) \frac{\partial \ell_j^k}{\partial x} d\mathbf{x} = J^k \int_{\mathbf{I}} \ell_i(\mathbf{r}) \left(\frac{\partial \ell}{\partial r} r_x^k + \frac{\partial \ell}{\partial s} s_x^k \right) d\mathbf{r} \,, \tag{4.16}$$

$$= J^k \int_{\mathbf{I}} \ell_i(\mathbf{r}) \left(\frac{\partial \ell}{\partial r} \frac{y_s^k}{J^k} - \frac{\partial \ell}{\partial s} \frac{y_r^k}{J^k} \right) d\mathbf{r} \,, \tag{4.17}$$

$$= y_s^k \mathcal{S}_r - y_r^k \mathcal{S}_s \tag{4.18}$$

where we have used the fact that the Jacobian matrices have the inverse property, i.e.,

$$\frac{\partial \mathbf{x}}{\partial \mathbf{r}} \frac{\partial \mathbf{r}}{\partial \mathbf{x}} = \begin{bmatrix} x_r & x_s \\ y_r & y_s \end{bmatrix} \begin{bmatrix} r_x & r_y \\ s_x & s_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \,, \tag{4.19}$$

hence,

$$r_x = \frac{y_s}{J}, \; r_y = -\frac{x_s}{J}, \; s_x = -\frac{y_r}{J}, \; s_y = \frac{x_r}{J} \,, \tag{4.20}$$

in going from (4.16) to (4.17). Similarly, for $\mathcal{S}_y^k$, we have

$$\mathcal{S}_{y,ij}^k = \int_{\mathbf{D}^k} \ell_i^k(\mathbf{x}) \frac{\partial \ell_j^k}{\partial x} d\mathbf{x} = -x_s^k \mathcal{S}_r + x_r^k \mathcal{S}_s \,. \tag{4.21}$$

The stiffness matrices defined on the standard triangle $\mathbf{I}$ are given by

$$\mathcal{S}_{r,ij} = \int_{\mathbf{I}} \ell_i(\mathbf{r}) \frac{\partial \ell_j}{\partial r} d\mathbf{r} \,, \quad \mathcal{S}_{s,ij} = \int_{\mathbf{I}} \ell_i(\mathbf{r}) \frac{\partial \ell_j}{\partial s} d\mathbf{r} \,. \tag{4.22}$$

We have hence written all local mass and stiffness matrices in terms of inner products over the standard triangle $\mathbf{I}$. For the moment, however, it is unclear how to evaluate these inner products since the explicit form of the two-dimensional Lagrange polynomials on a triangle are not known [60]. The developments by Hesthaven and Warburton [60] ensure that the evaluation of these inner products can be performed implicitly by considering an appropriate modal expansion that can be evaluated in a general way for arbitrary orders of approximation.

4.1.1 Evaluating the Inner Products: Modal approach vs. Nodal approach

We follow Hesthaven and Warburton's discussion by introducing a modal expansion for each solution field as an alternative to the nodal representations (e.g. (6.48)). For example, for

an arbitrary field $u(\mathbf{r})$ defined on $\mathbf{I}$, we have

$$u(\mathbf{r}) \approx u_h(\mathbf{r}) = \sum_{n=1}^{N_p} \hat{u}_n \psi_n(\mathbf{r}) = \sum_{i=1}^{N_p} u(\mathbf{r}_i) \ell_i(\mathbf{r}) \ , \tag{4.23}$$

where $\{\psi(\mathbf{r})\}_{i=1}^{N_p}$ is a two-dimensional basis. The relationship between the modes $\hat{u}_n$ and the nodes $u(\mathbf{r}_i)$ can be established by an L^2-projection onto a particular member of the basis ψ_m , i.e.,

$$\int_{\mathbf{I}} u(\mathbf{r}) \psi_m(\mathbf{r}) d\mathbf{r} = \sum_{n=1}^{N_p} \hat{u}_n \int_{\mathbf{I}} \psi_n(\mathbf{r}) \psi_m(\mathbf{r}) d\mathbf{r} \ , \tag{4.24}$$

or, in matrix-vector notation,

$$\mathbf{v} = \mathcal{H} \hat{\mathbf{u}} \ , \tag{4.25}$$

where

$$\hat{\mathbf{u}} = [\hat{u}_1, \cdots, \hat{u}_{N_p}], \ \ \mathcal{H}_{ij} = \int_{\mathbf{I}} \psi_i \psi_j d\mathbf{r}, \ \ \mathbf{v}_i = \int_{\mathbf{I}} u \psi_i d\mathbf{r} \ . \tag{4.26}$$

In order to ensure that $\mathcal{H}$ is well-conditioned (i.e., the basis functions are well-behaved) for an arbitrary-sized basis, it is instructive to choose the basis $\{\psi(\mathbf{r})\}_{i=1}^{N_p}$ to be orthonormal. It is then clear that $\mathcal{H}$ will reduce to the identity matrix. An appropriate basis can be found be applying the Gram-Schmidt process to the monomial basis $r^i s^j$ where $0 \leq i + j \leq N$, yielding [60]

$$\psi_m(\mathbf{r}) = \sqrt{2} P_i(a) P_j^{(2i+1,0)}(b) (1-b)^i \ , \tag{4.27}$$

where

$$a = 2 \frac{1+r}{1-s} - 1, \ \ b = s \ , \tag{4.28}$$

and $P_n^{(\alpha,\beta)}$ is the n^{th}-order Jacobi polynomial and $P_n = P_n^{(0,0)}$ is the n^{th}-order Legendre polynomial. In one space dimension, the relationship between the order of the highest-degree basis polynomial and the number of points on the element is given by $N_p = N + 1$. On the triangle, however, the relationship is given by the $(N+1)^{st}$ triangular number

$$N_p = \sum_{i=1}^{N+1} i = \left(\begin{array}{c} N+2 \\ 2 \end{array} \right) \ , \tag{4.29}$$

that can be derived by counting the number of basis polynomials of degree at most N. The connection to the binomial coefficient can be established by counting the number of distinct handshakes between $N + 2$ people in two different ways.

The only remaining question is how to evaluate the inner products on the left hand-side of the projection (4.24). If the numerical method uses a purely modal approach, one may consider using a cubature (2D quadrature) formula at the nodes, i.e.

$$\hat{u}_n \approx \sum_{i=1}^{N_p} u(\mathbf{r}_i)\psi_n(\mathbf{r}_i)w_i \ , \tag{4.30}$$

where the r_i's must be taken to be cubature points, and the w_i's are the associated cubature weights. Although the modal DG method is not considered in detail in this thesis, we later employ cubature and quadrature integration rules as a means of removing aliasing errors when curvilinear elements are employed in Section 4.14.

In the nodal approach [60], we assume the modal expansion interpolates u_h at the nodes $\mathbf{r}_i$, i.e.,

$$u_h(\mathbf{r}_i) = \sum_{n=1}^{N_p} \hat{u}_n\psi_n(\mathbf{r}_i) \ . \tag{4.31}$$

It follows that the relationship between the nodes and the modes can be established via the generalized Vandermonde matrix $\mathcal{V}$, that is

$$\mathcal{V}\hat{\mathbf{u}} = \mathbf{u} \ , \tag{4.32}$$

where $\mathcal{V}_{ij} = \psi_j(\mathbf{r}_i)$, $\hat{\mathbf{u}}_i = \hat{u}_i$, and $\mathbf{u}_i = u_h(\mathbf{r}_i)$. Combining (4.32) with the uniqueness statement (4.23), one can obtain the following useful formula for the Lagrange polynomials in terms of the basis polynomials

$$\ell_i(\mathbf{r}) = \sum_{n=1}^{N_p} \left(\mathcal{V}^{\mathsf{T}}\right)^{-1}_{in} \psi_n(\mathbf{r}) \ . \tag{4.33}$$

4.2 Polynomial interpolation nodes in 2D

It remains to ask what the optimal choice of points $\mathbf{r}_i$ should be to ensure that the system (4.32) is well-conditioned for arbitrary orders of approximation N. In one-dimension, one can show [60] that the optimal choice of points for the orthogonal Legendre polynomials is the Legendre–Gauss–Lobotto (LGL) or Gaussian quadrature points that bear a similar resemblence to the Chebyshev points (3.92). The LGL nodes can be computed quickly by numerically calculating the eigenvalues of a tridiagonal system [122, 60] and they ensure

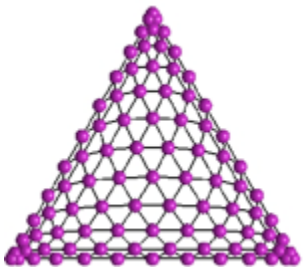

Figure 4.1: Warburton's "near-optimal" polynomial interpolation nodes on an equilateral triangle for order $N = 12$. Image from www.caam.rice.edu/∼timwar.

good conditioning by requiring the determinant of $\mathcal{V}$ to grow unbounded monotonically for large N.

On the triangle, the optimal choice of points is much more difficult to derive analytically. Until recently, this represented an open problem with the calculation of nodal sets on the triangle often leading to expensive implicit computations. The method of Warburton [128] considers a simple and inexpensive explicit approach for selecting the points, and we briefly explain this technique here. The resulting points are referred to as "near-optimal" since optimality has not been theoretically proven. However, the desired growth of the Vandermonde matrix determinant has been established [128]. The resulting set of points are not associated with quadrature formulas like the 1D LGL quadrature nodes, but are closely related to the one-dimensional LGL nodes. An illustration of these points on an equilateral triangle for polynomial order $N = 12$ is shown in Figure 4.1.

The underlying idea behind choosing the interpolation points on the triangle is to use the fact that along a triangle's edge, we expect the nodes to be distributed by arc-length in the same manner as the 1D LGL nodes on the interval $[-1, 1]$. Using this fact, a 1D deformation field called the *warp factor* can be computed that describes how the 1D equidistant nodes should be displaced to arrive at the LGL nodes along the edge. The final step relies on a technique known as Gordon–Hall blending. This technique blends the 1D deformation to the interior points by choosing an appropriate weighting function (called the *blend factor*) in the edge normal direction. The blend factor is based on barycentric coordinates to obey symmetry properties of the equilateral triangle [53, 60]. The resultant 2D displacement field is then applied to the equidistant points on the triangle, and the process is repeated for all edges of the triangle.

Though the method of Gordon and Hall was originally intended for constructing curvilinear coordinate systems, and we use it for this very purpose later on in section 4.14, it

has been shown to be useful as well for the purpose of choosing polynomial interpolation nodes in 2D and 3D geometries, and in particular on the triangle [128, 60].

4.3 Local operators for the nodal approach

We can now apply all of the developments of the above section to the evaluation of the inner products posed by equations (4.15) and (4.22). Substituting the formula (4.33) into the expression for the standard local mass matrix, we recover

$$\mathcal{M} = \left(\mathcal{V}\mathcal{V}^{\mathsf{T}}\right)^{-1} . \tag{4.34}$$

Before considering the stiffness matrices, it is useful to define the differentiation matrices

$$\mathcal{D}_{r,ij} = \left.\frac{\partial \ell_j}{\partial r}\right|_{\mathbf{r}_i}, \quad \mathcal{D}_{s,ij} = \left.\frac{\partial \ell_j}{\partial s}\right|_{\mathbf{r}_i} \tag{4.35}$$

whose entries may be furnished directly by appropriate differentiation of the formula (4.33). It can be shown [60] that the local stiffness matrices can be recovered by

$$\mathcal{M}\mathcal{D}_r = \mathcal{S}_r, \quad \mathcal{M}\mathcal{D}_s = \mathcal{S}_s . \tag{4.36}$$

In other words,

$$\mathcal{D}_r = \mathcal{M}^{-1}\mathcal{S}_r, \quad \mathcal{D}_s = \mathcal{M}^{-1}\mathcal{S}_s . \tag{4.37}$$

This is useful because it implies that an explicit semi-discrete scheme can be obtained by multiplying (4.11) by $(\mathcal{M}^k)^{-1} = \frac{1}{J^k}\mathcal{M}^{-1}$. As a consequence of the fact that the local mass matrix only varies by a constant factor on each element, it follows that this operation is computationally cheap since $\mathcal{M}$ is an $N_p \times N_p$ matrix. For example, with order $N = 8$ basis functions, the local mass matrix is a 45×45 full matrix. This is another key difference between DG-FEM and the classical FEM, where explicit semi-discrete schemes often cannot be recovered since the time-derivative operator is multiplied by the global mass matrix, that may be large and expensive to invert explicitly. It is worth mentioning that if quadrilateral elements are combined with inexact quadrature rules to evaluate the integrals, the global mass matrix is diagonalized in the traditional FEM [51], and the issue is avoided.

4.3.1 Surface Integral Contributions

To close our numerical scheme, it remains to discuss the surface integral term in equation (4.11)

$$\int_{\partial \mathbf{D}^k} \ell_j^k(\mathbf{x}) \mathbf{g}_h \cdot \hat{\mathbf{n}} d\mathbf{x} \tag{4.38}$$

where $\mathbf{g_h} = \left((hu)_h^k - (hu)^*, (hv)_h^k - (hv)^*\right)$ represents the jump in flux across an interface. Since the normal $\hat{\mathbf{n}}$ is constant along each edge, it is useful to break this expression up into three integrals

$$\int_{\partial \mathbf{D}^k} \ell_j^k(\mathbf{x}) \mathbf{g}_h \cdot \hat{\mathbf{n}} d\mathbf{x} = \sum_{e=1}^{3} \hat{\mathbf{n}}_e \cdot \int_{\text{edge}_e} \ell_j^k(\mathbf{x}) \mathbf{g}_h d\mathbf{x} \ . \tag{4.39}$$

If we substitute the nodal expression $\mathbf{g}_h = \sum_{i=1}^{N+1} \ell_i^k(\mathbf{x}) \mathbf{g}_i$ the right hand side reduces to

$$\sum_{e=1}^{3} \sum_{i=1}^{N+1} \hat{\mathbf{n}} \cdot \mathbf{g}_i \mathcal{M}_{ij}^{k,e} \ , \tag{4.40}$$

where we have introduced the $(N+1) \times (N+1)$ edge mass matrix

$$\mathcal{M}_{ij}^{k,e} = \int_{edge_e} \ell_j^k(\mathbf{x}) \ell_i^k(\mathbf{x}) \ d\mathbf{x} = J^{k,e,1} \mathcal{M}^1 \ , \tag{4.41}$$

where $J^{k,e,1}$ is the Jacobian of the mapping from the edge to the standard interval $[-1, 1]$. Using the 1D developments in [60], the standard 1D mass matrix is related to the Vandermonde matrix for 1D polynomial interpolation by $\mathcal{M}^1 = \left(\mathcal{V}^1 (\mathcal{V}^1)^T\right)^{-1}$.

4.4 Boundary conditions

The freedom in the numerical flux choice gives us a convenient way to impose boundary conditions through appropriately choosing imaginary "ghost" states, i.e. the '+' traces along boundary edges. For a purely reflective wall with no flow going through it, we impose

$$h^+ = h^- \ , \tag{4.42}$$
$$hu^+ = hu^- - 2(n_x hu + n_y hv) n_x \ , \tag{4.43}$$
$$hv^+ = hv^- - 2(n_x hu + n_y hv) n_y \ , \tag{4.44}$$

The second and third conditions are equivalent to imposing no normal $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ along the wall. The first condition is equivalent to imposing $\nabla h \cdot \hat{\mathbf{n}} = 0$ which states that the interface should be parallel to the bathymetry at the wall.

4.5 Bathymetry and non-hydrostatic terms

So far, we have not discussed the treatment of the bathymetry and nonhydrostatic terms contained in the vectors $\mathbf{B}$ and $\mathbf{N}$, respectively. We have avoided these terms so far because they cannot be addressed by the standard nodal DG-FEM treatment.

As an example of the issues that arise, let us consider the second entry of $\mathbf{B}_h$. If we remove the subscript-h notation for clarity, multiply by ℓ_j^k, and integrate over the element, the following integrals appear in the strong DG statement

$$\int_{\mathbf{D}^k} gh^k \frac{\partial H^k}{\partial x} \ell_j(x) d\mathbf{x} - \int_{\partial \mathbf{D}^k} gh \left(H^k - H^* \right) d\mathbf{x} \, , \tag{4.45}$$

the surface integral term does not pose a problem, and in the case where H is continuous across element interfaces, it vanishes. The first term, on the other hand does pose a problem because we cannot write it in terms of the local stiffness matrix $\mathcal{S}_x^k$. To see this, let us substitute a nodal expansion in for H, yielding (after index relabelling)

$$\begin{aligned} \int_{\mathbf{D}^k} gh^k \frac{\partial H^k}{\partial x} \ell_j(\mathbf{x}) d\mathbf{x} &= \sum_{j=1}^{N_p} gH^k(x_j) \int_{\mathbf{D}^k} h^k \ell_i(\mathbf{x}) \frac{\partial \ell_j}{\partial x} d\mathbf{x} \, , \\ &= \sum_{j=1}^{N_p} gH^k(x_j) \mathcal{S}_{ij}^{k,h} \, , \qquad (4.46) \\ &= g\mathcal{S}^{k,h} \mathbf{H}^k \, , \qquad (4.47) \end{aligned}$$

where we have taken the integral on the right to be the modified local stiffness matrix, which depends on h. Since h is a function of both space and time, this approach is computationally expensive since the local stiffness matrix is now different on every element and must be updated after each time-step. Although this approach is computationally expensive in general, it should not be completely disregarded since it becomes necessary in situations where curvilinear elements are used since the mapping Jacobian is no longer constant. See Sections 4.14–4.14.2.

To pursue a less expensive approach, let us introduce the auxiliary variable

$$\kappa(\mathbf{x}) = \frac{\partial H}{\partial x} \, . \tag{4.48}$$

Following previous discussion, we can approximate κ with the DG-FEM method by

$$\mathcal{M}^k \boldsymbol{\kappa} = \mathcal{S}_x \mathbf{H}^k - \int_{\partial \mathbf{D}^k} \left(H^k - H^* \right) n_x \, d\mathbf{x} \, , \tag{4.49}$$

or,

$$\boldsymbol{\kappa} = \mathcal{D}_x \mathbf{H}^k - \left(\mathcal{M}^k\right)^{-1} \int_{\partial \mathbf{D}^k} \left(H^k - H^*\right) n_x \, d\mathbf{x} \, . \tag{4.50}$$

If we now return to the bathymetry terms, we are charged with computing the integral

$$\int_{\mathbf{D}^k} g h^k(\mathbf{x}) \kappa^k(\mathbf{x}) \ell_j(\mathbf{x}) d\mathbf{x} \, . \tag{4.51}$$

We could proceed as before and simply substitute in the nodal expansion for κ^k, we would then be left with a modified mass matrix $\mathcal{M}^{k,h}$, and we will not have gained much. On the other hand, if we approximate the nodal expansion product $h^k \kappa^k$ in the following manner

$$h^k(\mathbf{x}) \kappa^k(\mathbf{x}) \approx \sum_{i=1}^{N_p} h^k(\mathbf{x}_i) \kappa^k(\mathbf{x}_i) \ell_i^k(\mathbf{x}) \, , \tag{4.52}$$

i.e., we approximate the function product with a point-wise (or Schur) product, we then recover the scheme

$$\begin{aligned} \int_{\mathbf{D}^k} g h^k \frac{\partial H^k}{\partial x} \ell_j(\mathbf{x}) d\mathbf{x} &\approx \sum_{j=1}^{N_p} g \kappa^k(\mathbf{x}_j) h^k(\mathbf{x}_j) \int_{\mathbf{D}^k} \ell_i(\mathbf{x}) \ell_j(\mathbf{x}) d\mathbf{x} \, , \\ &= g \mathcal{M}^k \left(\boldsymbol{\kappa} \mathbf{h}\right)^k \, , \end{aligned} \tag{4.53}$$

which is less computationally expensive than the former scheme since the local mass matrix only varies by a constant value between elements.

The price we pay when using this approximation is that we have essentially committed a couple of "variational crimes" [60]. Aliasing errors result from two distinct sources: 1) the fact that a product of two functions cannot be completely recovered by a point-wise product between the nodal values; and 2) the fact that the interpolant of a derivative is not the same thing as the derivative of an interpolant. As discussed above for the pseudospectral methods in section 3.2.3, modal filtering can be used to prevent these aliasing errors from driving weak instabilities. The form of the filter used is precisely the same as in section 3.2.3 with the exception that it is applied to the 2D basis on each element.

The inexpensive nodal approach presented here is used in time-stepping both the bathymetry terms $gh\nabla H$ and the non-hydrostatic terms $\gamma \nabla z$. Both of these terms may be regarded as source terms in the DG-FEM formulation assuming that h and z are known. The gradient of z may be either evaluated using the central flux $z^* = \{\!\{z\}\!\}$ or the purely internal choice $z^* = z^-$. We explain how z is calculated by solving the elliptic problem (4.4) within the DG-FEM framework below.

4.6 DG-FEM for elliptic problems

At first, it may not be clear how the DG-FEM can be applied to second-order elliptic equations since such equations are not hyperbolic and thus do not have well-posed Riemann problems that may be considered to impose weak continuity across element interfaces. However, it is explained in [60] that elliptic equations may be recast as a first-order system of equations and appropriate numerical fluxes can be obtained from penalty methods.

To re-write (4.4) as a first-order system, we introduce the auxiliary variable

$$\mathbf{q} = (q_x, q_y) = \sqrt{\gamma}\nabla z \,, \tag{4.54}$$

yielding the system

$$\nabla \cdot (\sqrt{\gamma}\mathbf{q}) - z = -\nabla \cdot \mathbf{a} \,, \tag{4.55}$$

$$q_x = \sqrt{\gamma}\frac{\partial z}{\partial x} \,, \tag{4.56}$$

$$q_y = \sqrt{\gamma}\frac{\partial z}{\partial y} \,. \tag{4.57}$$

Inspecting the system (6.45)-(6.47) it may be unclear how, given an input right-hand side $-\nabla \cdot \mathbf{a}$, one can recover z. This is generally achieved by considering the inverse situation, i.e., if z is known, then $\mathbf{q}$ can be computed by solving equations (6.46)-(6.47), and $-\nabla \cdot \mathbf{a}$ can be recovered using (6.45). This set of operations can be considered a non-singular linear transformation, and hence there must exist an inverse transformation.

The strong DG formulation of (6.46)–(6.47) together with the weak formulation of (6.45) is given by

$$\mathcal{M}^k \mathbf{q_x}^k = \sqrt{\boldsymbol{\gamma}}^k \mathcal{S}_x \mathbf{z}^k - \sqrt{\boldsymbol{\gamma}}^k \int_{\partial \mathbf{D}^k} \ell_j^k \left(z^k - z^*\right) n_x \, d\mathbf{x} \,, \tag{4.58}$$

$$\mathcal{M}^k \mathbf{q_y}^k = \sqrt{\boldsymbol{\gamma}}^k \mathcal{S}_y \mathbf{z}^k - \sqrt{\boldsymbol{\gamma}}^k \int_{\partial \mathbf{D}^k} \ell_j^k \left(z^k - z^*\right) n_y \, d\mathbf{x} \,, \tag{4.59}$$

$$\begin{aligned} -\left(\mathcal{S}_x^k\right)^T \left(\sqrt{\boldsymbol{\gamma}}\mathbf{q_x}\right)^k - \left(\mathcal{S}_y^k\right)^T \left(\sqrt{\boldsymbol{\gamma}}\mathbf{q_y}\right)^k + \int_{\partial \mathbf{D}^k} \ell_j^k (\sqrt{\gamma}\mathbf{q})^* \cdot \hat{\mathbf{n}} \, d\mathbf{x} - \mathcal{M}^k \mathbf{z}^k = \\ \left(\mathcal{S}_x^k\right)^T \mathbf{a_x}^k + \left(\mathcal{S}_y^k\right)^\mathsf{T} \mathbf{a_y}^k - \int_{\partial \mathbf{D}^k} \ell_j^k \mathbf{a}^* \cdot \hat{\mathbf{n}} \, d\mathbf{x} \,, \end{aligned} \tag{4.60}$$

Here, we choose the central flux for the right-hand side, i.e. $\mathbf{a}^* = \{\!\{\mathbf{a}\}\!\}$ together with the interior penalty (IP) flux for the elliptic operator, i.e. $z^* = \{\!\{z\}\!\}$, $(\sqrt{\gamma}\mathbf{q})^* = \{\!\{\sqrt{\gamma}\nabla z\}\!\} -$

$\tau[\![z]\!]$, $\tau > 0$. The point of the penalty term is to penalize large jumps at the element interfaces. If $\tau = 0$, a numerical calculation of the eigenfunctions of the Laplacian would reveal a spurious $\lambda = 0$ mode with all elements completely de-coupled, and the system would be singular [60]. The use of the penalty term pushes the spurious eigenmode out of the operator's null space to guarantee invertibility. In general, a sufficiently large penalty parameter will suppress any other spurious modes to the high-λ part of the eigenspectrum as well. This property represents an advantage over continuous Galerkin discretizations of elliptic operators that often possess spurious *convergent* modes whose corresponding eigenvalues can lie within the physical range of the eigenspectrum [24].

Other flux possibilities for DG-discretized Laplacian and Helmholtz operators include the penalized central flux $z^* = \{\!\{z\}\!\}$, $(\sqrt{\gamma}\mathbf{q})^* = \{\!\{\sqrt{\gamma}\mathbf{q}\}\!\} - \tau[\![z]\!]$ and the local discontinuous Galerkin (LDG) flux $z^* = \{\!\{z\}\!\} + \hat{\mathbf{n}}[\![z]\!]$, $(\sqrt{\gamma}\mathbf{q})^* = \{\!\{\sqrt{\gamma}\mathbf{q}\}\!\} - [\![\sqrt{\gamma}\mathbf{q}]\!] \cdot \hat{\mathbf{n}} - \tau[\![z]\!]$. The central flux should not be used in general since it only has optimal convergence characteristics for even polynomial orders N, and its matrix-form is not as sparse as the other choices (i.e., large stencil). The LDG flux, on the other hand, has optimal convergence rates at all orders and offers the most sparse representation, but it is known to be the most poorly conditioned operator. The IP flux offers a balance between these two, giving optimal convergence at all orders, a middle-ground in terms of sparsity, and similar condition numbers to the central-flux operator [60]. Furthermore, with some algebraic manipulations, the auxiliary variable $\mathbf{q}$ can be eliminated locally, and this allows the operator to be efficiently set-up directly as a symmetric sparse matrix. Such an elimination gives the following statement in terms of local operators

$$
\begin{aligned}
- \quad & \left(\left(\mathcal{D}_x^k\right)^T \mathcal{M}^k \Gamma^k \left(\mathcal{M}^k\right)^{-1} \Gamma^k \mathcal{M}^k \mathcal{D}_x^k + \left(\mathcal{D}_y^k\right)^T \mathcal{M}^k \Gamma^k \left(\mathcal{M}^k\right)^{-1} \Gamma^k \mathcal{M}^k \mathcal{D}_y^k + \mathcal{M}^k \right) \mathbf{z}^k \\
+ \quad & \sum_{e=1}^{3} \left(\mathcal{D}_n^{k,e}\right)^T \mathcal{M}^{k,e} \Gamma^{k,e} \left(\mathcal{M}^{k,e}\right)^{-1} \Gamma^{k,e} \mathcal{M}^{k,e} \left(\frac{\mathbf{z}^- - \mathbf{z}^+}{2} \right) \\
+ \quad & \sum_{e=1}^{3} \mathcal{M}^{k,e} \left[\Gamma^{k,e} \left(\mathcal{D}_n^{k,e} \left(\frac{\mathbf{z}^- - \mathbf{z}^+}{2} \right) + \tau(\mathbf{z}^- - \mathbf{z}^+) \right) \right] = \text{RHS} \, ,
\end{aligned}
\tag{4.61}
$$

where Γ^k is the diagonal matrix with the entries of $\sqrt{\gamma}^k$ written along its diagonal and $\mathcal{D}_n^{k,e} = \mathcal{D}_x^k n_x^{k,e} + \mathcal{D}_y^k n_y^{k,e}$ is the discretized normal derivative along edge e of element k.

The discontinuous Galerkin IP discretization method has become known as the 'symmetric interior penalty discontinuous Galerkin' (SIP-DG) method in the literature, and has been applied to the pressure Poisson equation and viscous operator of the incompressible Navier–Stokes equations [46, 105].

4.7 Briefly on the complexity of the DG-FEM

Since we have replaced all of the operators in the DG-FEM formulation with matrices, we are now in a position to analyze the amount of work a DG-FEM scheme requires at each time-step.

Given that the surface-integral contributions involve only three $(N+1) \times (N+1)$ matrices on each element, it is clear that the application of the differentiation matrices $(N_p \times N_p)$ to each element is the most expensive operation. Thus, a single time-step of one scalar (purely hyperbolic) equation discretized with DG-FEM would require $O(KN_p^2)$ FLOPs where K is the number of elements and N_p is the number of points per element. That is, we should expect the amount of work to scale linearly with the number of elements, and quadratically with the number of points per element.

Since the modeller's input does not directly control the number of points per element, it is useful to consider the amount of work in terms of the order of the basis polynomials. Using the identity (4.29), it is clear that the amount of work is

$$O\left(K\left(\begin{array}{c} N+2 \\ 2 \end{array}\right)^2\right) \text{ FLOPs.}$$

If we use the fact that $\left(\begin{array}{c} N+2 \\ 2 \end{array}\right) = O(N^2)$, we find the amount of work is $O(KN^4)$ FLOPs. Thus, the amount of work scales quartically with the order of the basis polynomials (in two-dimensions).

4.8 Time-stepping method

The time-stepping technique applied to the DG-FEM discretized version of the one-layer model closely follows the 'scalar approach' used for the pseudospectral discretization in Chapter 3 where splitting is applied such that advective and source terms are time-stepped first, followed by the dispersive terms.

Neglecting the dispersive terms for the time-being since they are not a part of the first splitting step, the method of lines [122] can be applied by noticing that once the DG-FEM integral form (4.10) has been written purely in terms of matrix operators as discussed in the above sections, we recover the system of ordinary differential equations

$$\frac{d\mathbf{Q}}{dt} = \mathcal{R}(\mathbf{Q}) \,, \tag{4.62}$$

where $\mathbf{Q} = (h, hu, hv)^\mathsf{T}$ is the vector of unknowns and $\mathcal{R}$ is the DG spatial discretization operator for the advection, Coriolis, and bathymetry source terms. It is assumed that we have left-multiplied the inverse of the local mass matrix, i.e., $\left(\mathcal{M}^k\right)^{-1}$, in arriving at $\mathcal{R}$ so that there is no matrix operator acting on the time-derivative on the left-hand side of (4.62). We have once again followed Eskilsson and Sherwin [42] and time-discretized (4.62) beginning at time-level $t_n = n\Delta t$ using the three-stage third-order strong stability preserving Runge-Kutta (SSP-RK) method [107], yielding

$$\mathbf{V}^{(1)} = \mathbf{Q}^n + \Delta t \mathcal{R}(\mathbf{Q}^n) \,, \tag{4.63}$$

$$\mathbf{V}^{(2)} = \frac{1}{4}\left(3\mathbf{Q}^n + \mathbf{V}^{(1)} + \Delta t \mathcal{R}(\mathbf{V}^{(1)})\right) \,, \tag{4.64}$$

$$\mathbf{Q}^\dagger = \frac{1}{3}\left(\mathbf{Q}^n + 2\mathbf{V}^{(2)} + 2\Delta t \mathcal{R}(\mathbf{V}^{(2)})\right) \,. \tag{4.65}$$

Modal filtering is applied to the spatial discretization operator $\mathcal{R}$ after each stage to help tame aliasing and nonlinearity-driven instabilities as explained in Section 3.2.3. The choice of SSP-RK time-stepper here is not a unique one, and we have mainly used it here since it offers third-order accuracy and allows for a simple adaptive time-stepping scheme. That is, Δt can be adjusted after each time-step without changing the coefficients of the scheme. The SSP-RK methods have gained favour in the DG-FEM literature [60, 20] since it guarantees no oscillations are introduced as a result of time-stepping for problems involving discontinuities and shocks. As mentioned previously, however, shocks and discontinuous features are not of concern for the equations under consideration here.

As discussed in Section 3.1.1, the next step in the 'scalar approach' is to solve the wave continuity equation. Its continuous form is given by

$$\nabla \cdot \left(\gamma \nabla z^\dagger\right) - z^\dagger = -\nabla \cdot \mathbf{a}^\dagger \,, \tag{4.66}$$

with $\gamma = H^2/6$ and $\mathbf{a}$ is as defined previously by equation (3.11). The spatially discretized vector $\mathbf{a}^\dagger$ can be computed quite simply by evaluating $\mathcal{R}(\mathbf{Q}^\dagger)$ for only the $\mathbf{hu}$ and $\mathbf{hv}$ equations. The auxiliary variable $\mathbf{z}^\dagger$ is then computed by inverting the matrix representation of the SIP-DG formulation of the Helmholtz operator (4.61). The momentum equations are then updated via

$$(h\mathbf{u})^{n+1} = (h\mathbf{u})^\dagger + \gamma \Delta t \nabla z^{n+1} \,, \tag{4.67}$$

where the DG-FEM discretization of the source terms involving nonlinear products with gradients of known quantities is discussed in Section 4.5. Hence, the vector of unknowns at time t_{n+1} is updated via $\mathbf{Q}^{n+1} = (h^\dagger, hu^{n+1}, hv^{n+1})^\mathsf{T}$.

For the simulations considered in this thesis, we have found that the SIP-DG sparse matrix can be factored using the sparse LU-decomposition that was explained in Section 3.2.2 so that the factors may be re-used at each time-step for fast inversions, and we have side-stepped the issue of using an iterative solver such as GMRES that was required for the pseudospectral methods of Chapter 3. At high resolutions, a linear iterative method will certainly be necessary due to memory restrictions prohibiting the storage of the LU factors. While preliminary experiments with iterative solvers and pre-conditioning for the DG-FEM have been carried out by the author, it is an issue worthy of further research.

4.9 DG-FEM vs. Fourier method in 1D

As highlighted previously in Chapter 3, the Fourier/Chebyshev pseudospectral methods are good methods to benchmark against due to their high-order accuracy and small amounts of inherent dissipation. Since the majority of future work will be in developing DG-FEM methods for Boussinesq-type systems, we have decided to compare results from the DG-FEM method to the Fourier method in 1D with a flat bottom to better understand the performance of the DG-FEM method at varying orders of accuracy. In 1D, the system (2.120)-(2.122) reduces to

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0 \,, \tag{4.68}$$

$$\frac{\partial (hu)}{\partial x} + \frac{\partial \left(hu^2\right)}{\partial x} = -\frac{\partial}{\partial x}\left(\frac{1}{2} g h^2\right) + \frac{H^2}{6}\frac{\partial (hu)}{\partial x^2 t} \,, \tag{4.69}$$

where H is a constant.

To perform this comparison, we have decided to run a simulation where a packet of short waves of two distinct wavelengths is released from rest. The domain was taken to be periodic and $L_x = 4000$ m in length, the depth was fixed at $H = 5$ m, and the acceleration due to gravity was taken to be $g = 9.81$ m s^{-2}. The initial condition is

$$\eta(x,0) = \eta_0 \cos(0.15x)\cos(0.05x) \mathrm{e}^{-5\left(\frac{x-0.5L_x}{400}\right)^2} \,, \tag{4.70}$$

$$u(x,0) = 0 \,, \tag{4.71}$$

where $\eta_0 = 0.1$ m. The amplitude of the wave packet was chosen to be small enough so that linear wave theory would be a good predictor of the group velocities. This was confirmed by solving the linearized equations exactly in Fourier space and comparing with the numerical solution (not shown). Due to dispersion, we expect the longer waves to

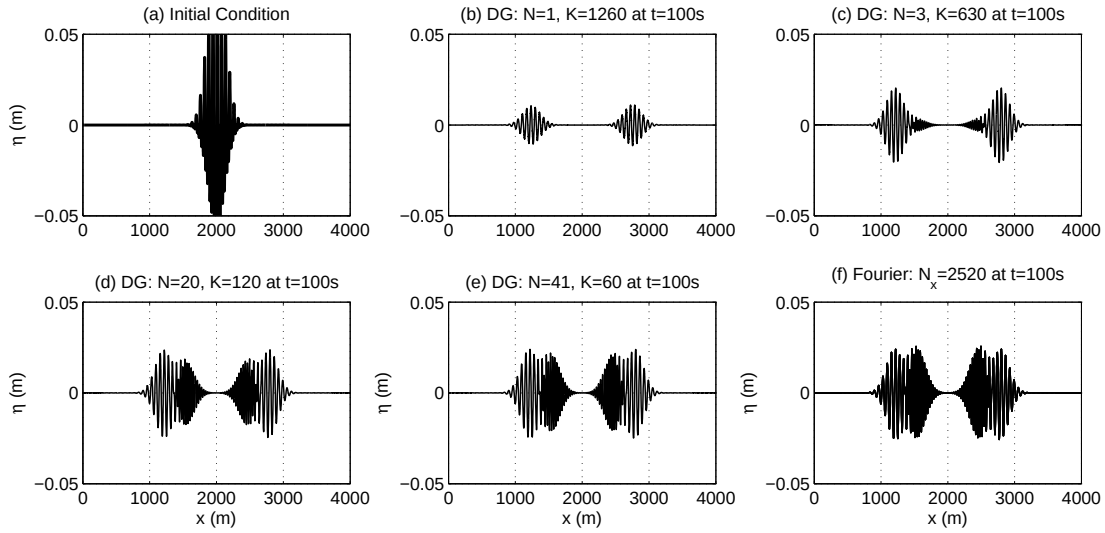

Figure 4.2: Fixed time snapshots of the free-surface displacement at various orders of approximation for the 1D dispersive short-waves run. Panels (b)-(f) are all at time $t = 100$s. (a) η at $t = 0$. (b) DG-FEM $N = 1$ result. (c) DG-FEM $N = 3$ result. (d) DG-FEM $N = 20$ result. (e) DG-FEM $N = 41$ result. (f) Fourier Method with $N_x = 2520$ grid points result

overtake and lead the shorter waves, after sufficient time has passed, since the linear group velocity of the longer waves is $c_g \approx 9.31$ m s^{-1} while the group velocity of the shorter waves is $c_g \approx 6.14$ m s^{-1}. This run was also used to validate the numerical methods in the regime where nonlinear effects are negligible and the bottom is flat.

In Figure 4.2, the results of the runs are displayed at various orders of accuracy. The values of K (total number of elements) and N (order of the basis functions) were chosen such that the total number of points used in the DG-FEM method would be fixed at $N_{dof} = K(N + 1) = 2520$. Modal filtering was not used in any of the runs, since the choice of small-amplitude waves over a flat bottom remove most, if not all, of the sources of nonlinearity and aliasing errors. A striking observation is that for the low-order runs, the shorter waves are dissipated to a very large degree, and in the $N = 1$ case, in effect entirely.

The plots reveal the difference in numerical dissipation between the DG-FEM method at different orders when compared to the Fourier method. Even at a very high order of

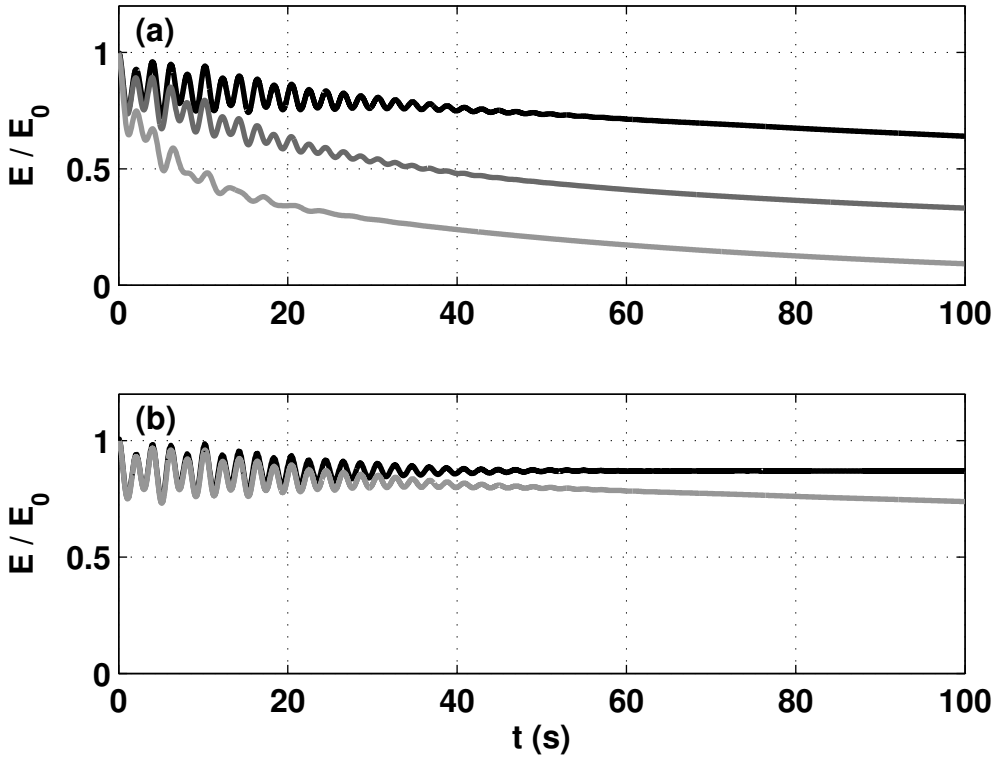

Figure 4.3: Domain-integrated total energy time series for the *1D dispersive short-waves* run with (a) the DG-FEM method at orders $N = 1$ (light grey), $N = 3$ (dark grey), and $N = 20$ (black), and (b) the DG-FEM method at order $N = 41$ (grey) and the Fourier method with $N_x = 2520$ points (black). The domain-integrated total energy E has been scaled by E_0, its value at $t = 0$. The number of grid points (degrees of freedom) is fixed at $N_{dof} = 2520$ in all cases.

$N = 41$, the DG-FEM method cannot match the energy-conserving qualities of the Fourier method. This fact is likely owed to the numerical dissipation introduced by using the numerical flux function (4.7) that for stability, is chosen to contribute a non-positive value to the global energy balance at each time-step [60]. Regardless of this fact, for a fixed number of degrees of freedom ($N_{dof} = 2520$), one still expects the DG-FEM method's result to converge to the Fourier method's result in the high-order limit $(N, K) = (2519, 1)$ where the number of interior elemental interfaces is zero.

4.10 Alternative advective numerical fluxes

Given the results of the above section, it is reasonable to ask if better DG-FEM results become possible with a different choice of numerical flux function for the advective terms by considering a more accurate approximate Riemann solver. To answer this question, we have implemented the HLLC flux (Harten–Lax–Van Leer Contact wave flux) as was also done for a Boussinesq-type system in [42]. A derivation of the HLLC flux is available in Toro [121] and involves making assumptions about the wave pattern arising from the nonlinear Riemann problem for the shallow water equations in quasi-linear form.

The HLLC solver approximates wave speeds along the interior $(-)$ edge and exterior $(+)$ edge and in the intermediate $(*)$-region by,

$$S^{\pm} = \left|u_{\perp}^{\pm}\right| \pm \sqrt{g h^{\pm}} s^{\pm} , \tag{4.72}$$

$$S^{*} = \frac{S^{-} h^{+} \left(u_{\perp}^{+} - S^{+}\right) - S^{+} h^{-} \left(u_{\perp}^{-} - S^{-}\right)}{h^{+} \left(u_{\perp}^{+} - S^{+}\right) - h^{-} \left(u_{\perp}^{-} - S^{-}\right)} , \tag{4.73}$$

where $u_{\perp} = \mathbf{u} \cdot \hat{\mathbf{n}}$ and $u_{\parallel} = \mathbf{u} \cdot (-n_y, n_x)$ are the velocities in the edge-normal and edge-tangent direction, and

$$s^{\pm} = \begin{cases} \sqrt{\left((h^{*})^2 + h^{*} h^{\pm}\right) / 2 (h^{\pm})^2}, & h^{*} > h^{\pm} \\ 1, & h^{*} \leq h^{\pm} \end{cases} . \tag{4.74}$$

The water column height in the star region is approximated with

$$h^{*} = \frac{1}{g} \left(\frac{1}{2} \left(\sqrt{g h^{-}} + \sqrt{g h^{+}} \right) + \frac{1}{4} \left(u_{\perp}^{-} - u_{\perp}^{+} \right) \right)^2 . \tag{4.75}$$

Along the edge-normal $\hat{\mathbf{n}}$ direction, the HLLC flux is given by

$$\mathbf{F}^*(\mathbf{Q}^-, \mathbf{Q}^+) = \begin{cases} \mathbf{F}(\mathbf{Q}^-), & S^- \geq 0 \\ \mathbf{F}(\mathbf{Q}^-) + S^- \left(\mathbf{Q}^{*,-} - \mathbf{Q}^-\right), & S^- \leq 0 \leq S^* \\ \mathbf{F}(\mathbf{Q}^+) + S^+ \left(\mathbf{Q}^{*,+} - \mathbf{Q}^+\right), & S^* \leq 0 \leq S^+ \\ \mathbf{F}(\mathbf{Q}^+), & S^+ \leq 0 \, . \end{cases} \tag{4.76}$$

The intermediate state vectors $\mathbf{Q}^{*,\pm}$ are given by

$$\mathbf{Q}^{*,\pm} = h^\pm \left(\frac{S^\pm - u_\perp^\pm}{S^\pm - S^*} \right) \begin{bmatrix} 1 \\ S^* \\ u_\parallel^\pm \end{bmatrix} \, , \tag{4.77}$$

In Figure 4.4, we compare the HLLC and Lax–Friedrichs (LF) fluxes at varying polynomial orders for the same simulation shown in Section 4.9 by once again plotting the total energy time series. A notable difference between the LF and HLLC schemes are only found in the low-order $N = 1$ case in panel **(a)** where the HLLC flux yields significantly less numerical dissipation than the LF flux. As the order is increased in panels **(b)**–**(c)**, the energy profiles of the two schemes converge.

In retrospect, the result of Figure 4.4, should not be surprising since the choice of numerical flux only alters the DG-FEM method along the edges of elements. As N increases, the number of interior points increases as well and the contribution of the interfacial treatment to the overall scheme is marginalized. It is for this reason that it becomes important to use modal filtering for high-order nonlinear simulations (see sections below), since we cannot expect the numerical flux on its own to stabilize the numerical method, especially in the presence of nonlinear polynomial aliasing errors. The dissipative contribution of numerical flux functions remains important, however, since it is known that polynomial filtering becomes least effective at the edges of elements where the polynomials are most poorly behaved [14].

Hesthaven and Warburton [60] have also suggested that the choice of numerical flux only becomes important in situations involving shocks. In such situations, a limiter must be used to preserve monotonicity near the shock. Hence, the scheme is reduced to a low-order finite volume method in the vicinity of shocks, and the choice of Riemann solver once again becomes important. The fact that a limiters will always locally degrade the order of the solution is problematic, and recent works have proposed the use of localized artificial viscosity instead [135]. Since the governing equations in this thesis contain dispersive terms that preclude the formation of purely discontinuous shocks, the issue of slope limiting is side-stepped, and we can expect to obtain stable and accurate solutions with high polynomial orders N and the LF flux, provided sufficient modal filtering is employed.

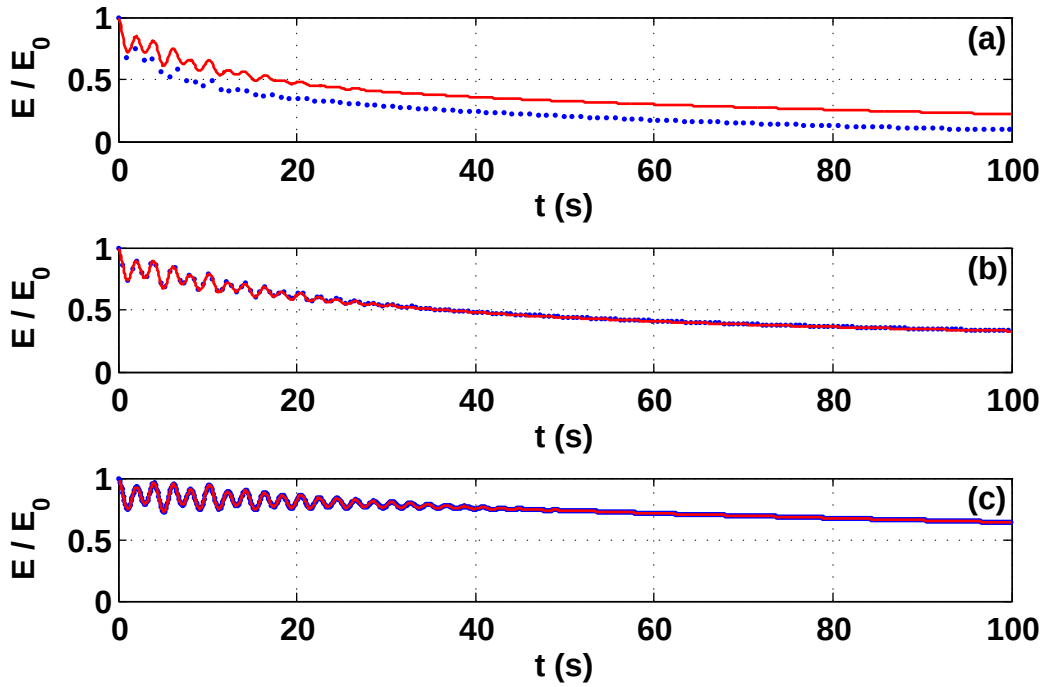

Figure 4.4: Total energy time series for the *1D dispersive short-waves* with the DG-FEM method using the Lax–Friedrichs flux (blue, dots) and the HLLC flux (red, solid). Polynomial orders are given by **(a)** $N = 1$, **(b)** $N = 3$, **(c)** $N = 20$.

4.11 DG-FEM vs. Fourier–Chebyshev method in 2D

In this section, we validate our DG-FEM solver for the one-layer weakly non-hydrostatic model equations (2.120)–(2.122) against the Fourier–Chebyshev method presented in Section 3.3 for the particular test-case in Section 3.3.6 where an initial tilt of amplitude $a = 0.25H$ is released from rest (see Figure 3.17, where the depth is $H = 12.8$ m). The Coriolis frequency is given by $f = 7.8828 \times 10^{-5}$ s^{-1} and the reduced gravity is $g' = 0.024525$ m s^{-2}.

Unlike the 1D comparison from Section 4.9, it should be noted that the comparison between the Fourier method and the DG methods at various orders here is not a "fair" one, since the number of degrees of freedom has not been held fixed in all cases. Here, the point is to illustrate that spectral-like resolution characteristics can become possible with increasing polynomial orders on a fixed finite element mesh triangulation. The mesh used here consists of $K = 1330$ triangular elements, and it was chosen to directly correspond with the annular-shaped lake of Section 3.3.6. The mesh, shown in Fig. 4.6 **(b)** was generated with the `mesh2d` MATLAB algorithm that uses an adaptive Delaunay-based triangulation algorithm implemented using quadtrees. Modal filtering was applied using an exponential cut-off filter analogous to the one used for the pseudospectral methods (3.41). Here, the cut-off polynomial order was set to $N_c = 3$ and the filter order was set to $s = 4$, and the same filtering parameters were used in both the $N = 4$ and $N = 8$ cases.

The energy characteristics of the various methods from Figure 4.5 are compared in Figure 4.6 **(a)** by plotting the scaled total energy against time. We see that the initial energy rapidly dissipated in the low-order $N = 1$ case. The $N = 4$ and $N = 8$ cases exhibit nearly identical energy profiles, however a more detailed view would reveal that the $N = 8$ line is slightly above the $N = 4$ line. As expected, the Fourier–Chebyshev method outperforms all of the DG methods with the least energy lost. Once again, we illustrate here the utility of pseudospectral methods as a benchmark numerical method. These results also validate the DG-FEM method since the details of the nonlinear wave fronts in Figure 4.5 are reasonably represented for orders $N \geq 4$, and the amount of numerical dissipation approaches that of the pseudospectral method for increasing N (see Figure 4.6 **(a)**.

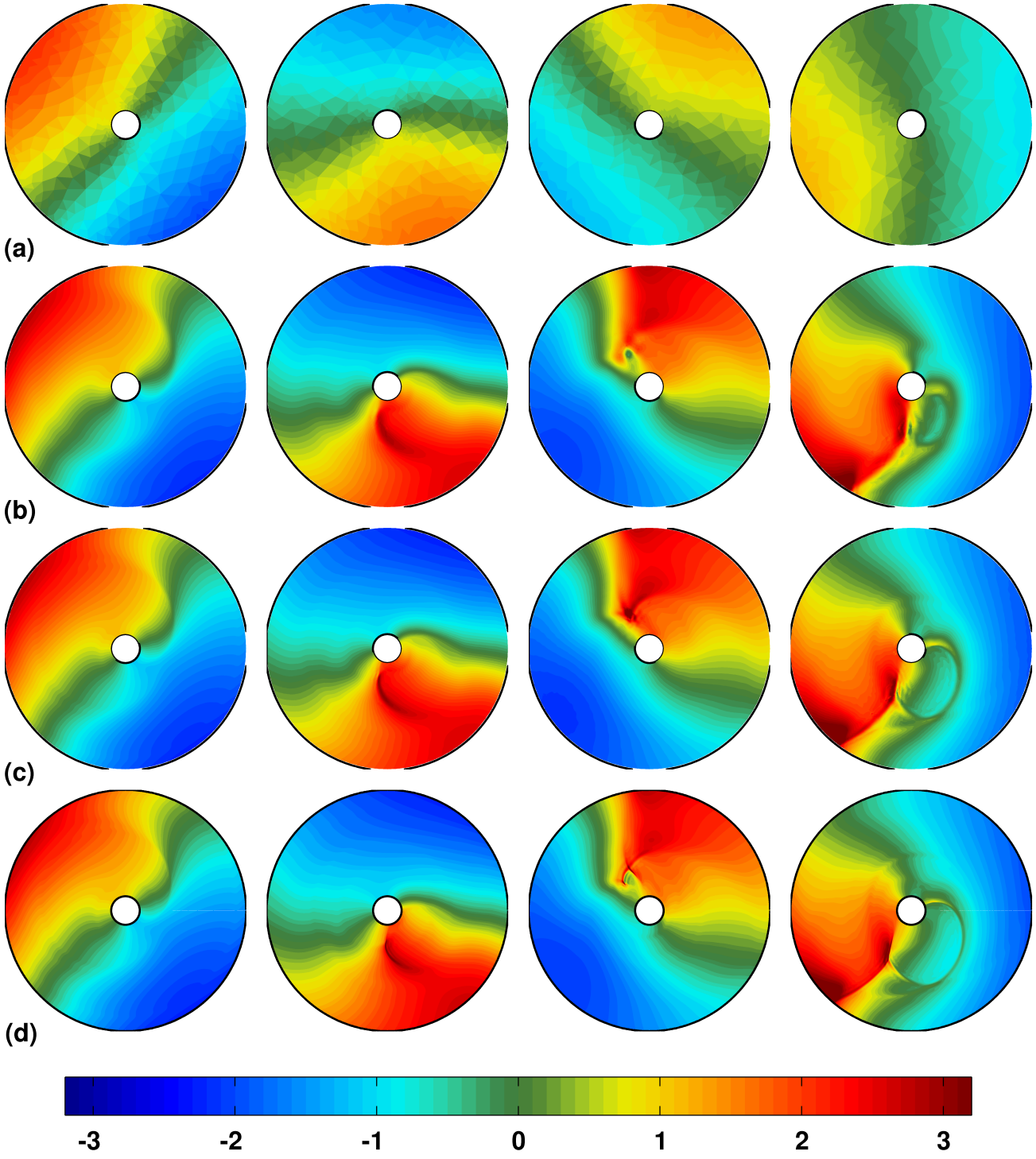

Figure 4.5: Comparison between the DG-FEM method at orders **(a)** $N = 1$, **(b)** $N = 4$, and **(c)** $N = 8$ to the Fourier–Chebyshev pseudospectral method (row **(d)**) for the IKS simulation presented in Section 3.3.6 with $N_r \times N_\theta = 256 \times 1024$ points. The number of elements in the DG simulations was $K = 1330$ in all cases. Row **(d)** corresponds to the same simulation presented in Figure 3.17, but at different times. In all rows, snapshots of the η field are given at times (from left-to-right): $t = 7$ h, $t = 14$ h, $t = 20$ h, $t = 27$ h.

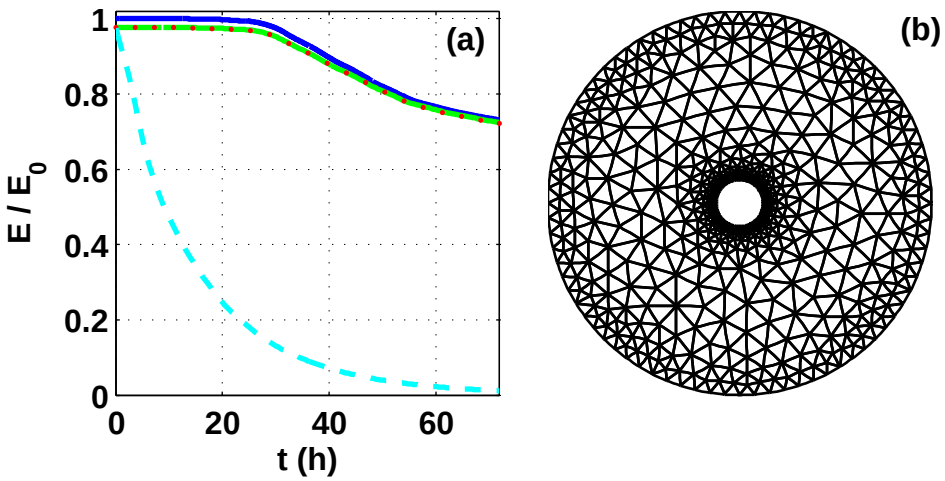

Figure 4.6: Panel **(a)**: Scaled domain-integrated total energy (E/E_0) time-series for the simulations in Fig. 4.5. The lines correspond to the Fourier–Chebyshev method with 256×1024 points (blue, solid), DG-FEM with $N = 8$ (green, solid), DG-FEM with $N = 4$ (red, dots), DG-FEM with $N = 1$ (cyan, dashed). Panel **(b)**: Finite element mesh with $K = 1330$ elements used in the DG-FEM simulations.

4.12 Internal rotating seiche in a lake with a wavy coastline

Since we are now comfortable that the one-layer DG-FEM solver has been validated, we can now enjoy the benefits that the DG method offers in terms of geometric flexibility to explore lakes with irregular shapes. In this section, we consider a simple perturbation of the annular geometry studied in sections 4.11 and 3.3 by taking the outer basin radius to depend on θ, i.e.,

$$r_{outer} = 8345 + 1000\sin(\theta)\ , \tag{4.78}$$

representing a dimpled circle. Once again, the mesh was generated using `mesh2d` and

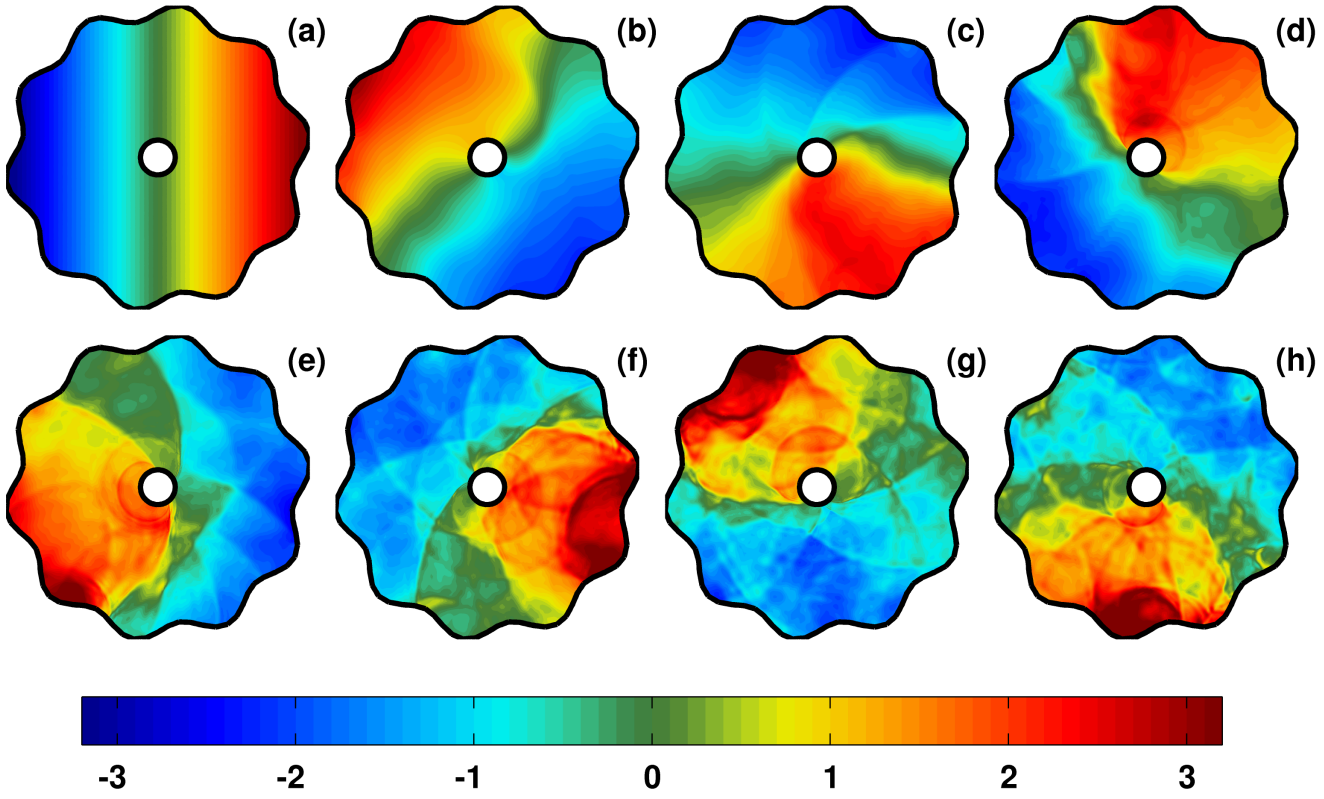

Figure 4.7: Evolution of an initial linear interfacial tilt in a lake with a wavy coastline at times **(a)** $t = 0$ h, **(b)**, $t = 6.8$ h, **(c)** $t = 14.0$ h, **(d)** $t = 20.9$ h, **(e)** $t = 28.1$ h, **(f)** $t = 34.9$ h, **(g)** $t = 42.1$ h, **(h)** $t = 49.0$ h.

has $K = 1394$ triangular elements (see Fig. 4.8 **(a)**). The run was carried out with polynomial order $N = 8$ and the filtering parameters were the same as given above in Section 4.11. Snapshots of the interfacial displacement are shown in Figure 4.7 at roughly 7 h intervals. Panels **(b)-(e)** should be compared and contrasted with the snapshots shown of the annular/circular geometry in Figure 4.5 since the times shown are quite similar. As expected, the results shown here suggest that a more irregular coastline gives rise to a more active wave field. In particular, the "dimpled" nature of the coast allows for more situations involving near-shore focusing and wave-steepening, giving rise to more small-scale nonlinear wave features than is possible with the simpler circular case. Since there are more opportunities for nonlinear short waves to emerge, it is not surprising that they become a more ubiquitous feature in this lake than a somewhat more localized feature in the circular lake.

Field scientists are often charged with the very difficult task of taking a time-series corresponding to one or more single-point measurements and explaining what physical processes are at work in the lake. Numerical simulations offer us the exciting opportunity to consider the inverse problem: Given a model lake, where physical processes can be either prescribed or easily discerned by the modeller, what are the corresponding (noiseless) single-point measurements? This is a powerful idea for testing the way we interpret field data, and we explore this idea in Figure 4.8 **(b)**–**(d)** where we plot the time-series and the corresponding power spectra from four synthetic data probes that measure the interfacial height.

Without the full simulation results from Figure 4.7, the interpretation of the synthetic measurements in Figure 4.8 is complicated. Despite the nonlinear nature of the simulation, the three most prominent peaks in the power spectrum correspond well to the lowest frequency free modes of oscillation that can be calculated using linear theory calculations of Rao and Schwab [100] (modified for the DG-FEM). The frequencies of the first three (vertical mode 1) rotating internal gravity modes (not pictured here) are: 0.055 cph, 0.089 cph, and 0.10 cph. The peak most accurately predicted by linear theory is the frequency corresponding to the gravest mode, 0.055 cph. These results illustrate that the linear regime remains important to study since low-frequency Kelvin and Poincaré waves will always be the mechanisms that drive the large-scale motions of a lake, even following the emergence of a large set of nonlinear waves.

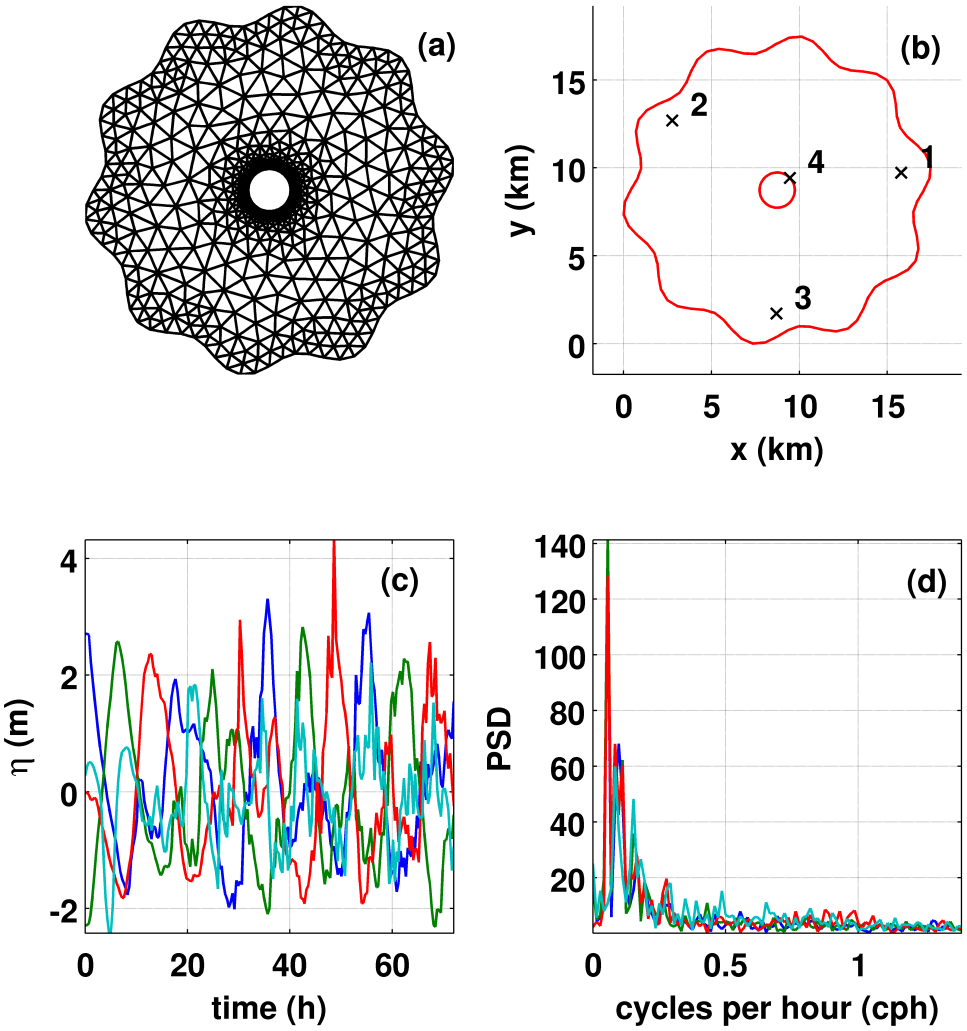

Figure 4.8: Panel **(a)**: Finite element mesh with $K = 1394$ elements. Panel **(b)** shows the theoretical locations of four synthetic probes numbered from 1–4. Panels **(c)** and **(d)** show the time-series and power spectral density of interfacial height, respectively, recorded by meter 1 (blue), 2 (green), 3 (red), 4 (cyan).

4.13 Spurious eddies in inviscid DG-FEM solutions

Exploring other geometries with the DG-FEM code, it was found that under certain conditions spurious eddies, corresponding to an unphysical production of vorticity, appear to form in the domain near obstacles resembling re-entrant corners (i.e., boundary corners that protrude inwards into the domain). This effect is illustrated in Figure 4.9 where our annular basin has been perturbed to include a peninsula. The DG-FEM solver with polynomial order $N = 4$ was initialized with the initial conditions used in Sections 4.11 and 4.12. A numerical instability occurred shortly after $t = 27$ h, preventing further time-stepping, though the reason for the instability was evident earlier due to the sharp gradients visible near the corner in Figure 4.9.

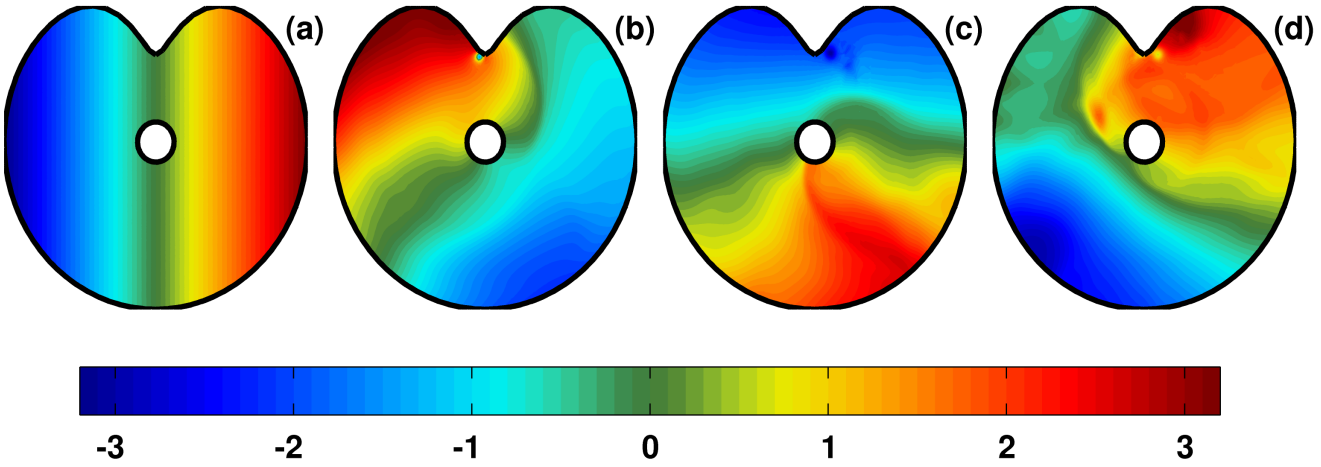

Figure 4.9: Snapshots of the η-field in the order $N = 4$ DG-FEM simulation of a rotating seiche on a perturbed circular domain with a re-entrant peninsula at **(a)** $t = 0$ h, **(b)** $t = 6.8$ h, **(c)** $t = 14.0$ h, **(d)** $t = 20.9$ h. Note the apparent separation eddies visible near the peninsula in panels **(b)**–**(d)**.

The eddies bear a striking resemblence to boundary-layer separation eddies that would occur due to flow past an obstacle in viscous flow [70]. However, since our model equations do not contain any viscous terms, the formation of a viscous boundary-layer is not possible and hence boundary-layer separation should not be possible. These spurious eddies are thus artifacts, and appear to coincide with the presence of a sharp re-entrant corner. Even in the cases where the actual boundary is smooth, re-entrant corners at the element-scale may result as a consequence of the piece-wise linear representation of the boundary assumed

in mesh generation. Although these artifacts are spurious in the sense that inviscid flow around an obstacle should not separate, from a theoretical stand-point they should be expected. Below, we explain why this is the case and propose methods for remedying the situation

Consider the simpler situation of inviscid, incompressible and irrotational (i.e., potential) flow. That is, we assume

$$\nabla \cdot \mathbf{u} = 0 \ , \quad \nabla \times \mathbf{u} = 0 \ . \tag{4.79}$$

In general, these constraints will certainly not be true during simulations and we presently make this assumption simply to gain insight from theory. Under these constraints, the velocity field has both a velocity potential and a streamfunction, i.e.,

$$\mathbf{u} = \nabla\phi = \nabla^{\perp}\psi \ , \tag{4.80}$$

where $\nabla^{\perp} = \mathbf{k} \times \nabla = \left(-\frac{\partial}{\partial y}, \frac{\partial}{\partial x}\right)$. Suitably differentiating and adding the various components (or by using vector calculus identities), it is straightforward to show that ϕ and ψ both satisfy Laplace's equation

$$\nabla^2\phi = 0 \ , \quad \nabla^2\psi = 0 \ . \tag{4.81}$$

The no-normal flow boundary conditions imply that

$$\frac{\partial\phi}{\partial n} = 0 \ , \quad \text{on} \quad \partial\Omega \ , \tag{4.82}$$

$$\psi = \text{constant} \ , \quad \text{on} \quad \partial\Omega \ . \tag{4.83}$$

It is useful to define a complex velocity potential

$$w(z) = \phi + i\psi \ , \tag{4.84}$$

where $z = x + iy$, from which the velocity components (u, v) can be recovered by noticing that

$$\frac{dw}{dz} = \frac{\partial(\phi + i\psi)}{\partial x} = u - iv \ , \tag{4.85}$$

where we have used the fact that the derivative of any complex-valued function is independent of the direction taken in the xy-plane.

Progress was made in the area of potential flow theory not by solving Laplace's equation, but rather by using the result from complex analysis that any analytic function (a complex-valued function whose derivative exists) has real and imaginary parts that satisfy Laplace's

equation [70]. Thus, any analytic function that is known represents a potential flow solution for some particular situation. What type of obstacle the fluid is flowing around depends on properties of the function, $w(z)$, itself.

It is known that the complex potential for flow around a wall angle $\alpha = \pi/n$ is given by [70]:

$$w = Az^n \ , \ \left(n \geq \frac{1}{2}\right) \ . \tag{4.86}$$

In polar coordinates, $z = re^{i\theta}$,

$$w = A\left(re^{i\theta}\right)^n = Ar^n(\cos(n\theta) + i\sin(n\theta) \ , \tag{4.87}$$

hence

$$\phi = Ar^n \cos(n\theta) \ , \quad \psi = Ar^n \sin(n\theta) \ . \tag{4.88}$$

Noticing that $\psi = 0$ for $\theta = 0, \pi/n$ and using the fact that any streamline along which $\psi = constant$ can represent a boundary, we recognize that this particular complex potential corresponds to flow around a sharp corner with angle $\theta = \pi/n$.

At this point, we can recover the velocity using equation (4.85)

$$u = \Re\left(nAz^{n-1}\right) \ = \ \Re\left(\frac{A\pi}{\alpha} z^{\frac{\pi-\alpha}{\alpha}}\right) \ , \tag{4.89}$$

$$v = \Im\left(nAz^{n-1}\right) \ = \ \Im\left(\frac{A\pi}{\alpha} z^{\frac{\pi-\alpha}{\alpha}}\right) \ . \tag{4.90}$$

Paying close attention to the expression on the right-hand side, it is clear that the exponent to which z is raised is negative whenever the wall-angle exceeds π, and thus the velocity is not defined when $z = 0$ (the corner).

Even in the case where $0 < \alpha < \pi$, the point $z = 0$ is a singular point since the velocity's derivatives do not exist. We thus reach the main point of this discussion, that is also highlighted by Kundu [70]: the velocity at a wall-corner is infinite if the wall-angle is greater than 180° and is zero for wall-angles less than 180°. Thus, near re-entrant corners the numerical solution should be *expected* to be poorly behaved since the exact potential flow solution is also. Although the velocity derivatives do not exist at corners less than 180°, this does not appear to be an issue for the numerical solution.

The "spurious eddies" encountered in simulations begin as very steep free-surface depressions that diffuse away from the boundary, to understand this phenomenon, consider Bernoulli's equation

$$p + \frac{1}{2}\rho_0 \|\mathbf{u}\|^2 = const. \ , \tag{4.91}$$

where p is the pressure. Taking the gradient and re-arranging gives

$$-\nabla p = \frac{1}{2}\rho_0 \nabla \left(\|\mathbf{u}\|^2 \right) \ . \tag{4.92}$$

Since the flow is infinite at the corner corresponding to $z = 0$, the upstream fluid must

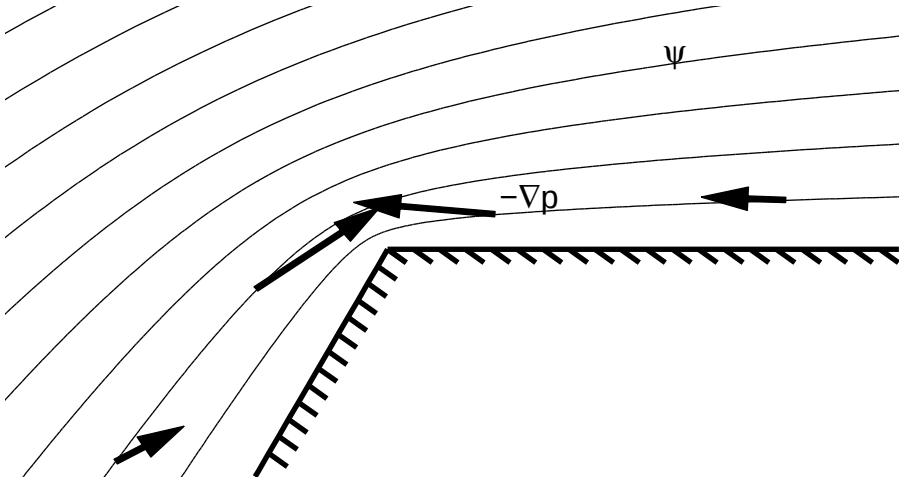

Figure 4.10: Cartoon diagram of potential flow around a wall corner of angle $\alpha > \pi$. Contours depict streamlines (lines where $\psi = const.$) and arrows illustrate the relative strength and direction of $-\nabla p$.

accelerate as it approaches the corner and decelerate as it moves away from it. This fact, together with equation (4.92), implies that $-\nabla p$ points in the direction with the flow upstream of the corner (i.e., a favourable pressure gradient) and *against* the flow downstream of the corner (i.e., an adverse pressure gradient). Hence, a region of very low pressure must exist in the vicinity of the corner. It is thus clear then why the singularity would manifest itself as a large (localized) free-surface depression since in the shallow water framework $p = \rho_0 g \eta$. In general, we found that the free-surface depression is advected downstream of the corner immediately after appearing, since the numerical method cannot compensate for the singularity.

In real flow around a corner, the region of adverse pressure gradient would cause the flow to separate from the corner resulting in the formation of eddies due to effects in the viscous boundary layer [70]. In the DG-FEM simulations discussed, the observed eddies

are a result of the local modal filtering that attempts to stabilize the pressure singularity by diffusing it away from the boundary, taking over the role of viscosity in realistic flows. This effect of the filter was discovered by turning off the filter and observing singular growth at the corner that leads to numerical blow-up with no eddy introduced. It was also found that spurious eddy generation is more prominent in simulations where nonlinear effects are non-negligible. The fact that a standard filter coupled with the presence of re-entrant corners will typically lead to spurious eddies is a dangerous feature of the numerical model, since a modeller may be led to believe that these eddies are physical, when in fact they are the result of the filter's action on a part of the solution that is singular. For instance, in [136], spurious eddies due to a limiter are presented as physical for the situation of supersonic compressible flow past an equilateral triangle. Despite the effort of filtering, it has been found that this singular behaviour can still lead to numerical blow-up. Thus, some effort must be taken to remedy this problem, as discussed below.

4.14 Curvilinear elements

In addition to solution singularities, it is also known that the convergence rates of a high-order method may be limited to sub-optimal rates as a result of an inaccurate representation of the boundary. This fact was demonstrated in [60] who demonstrated poor convergence rates for the solution to Maxwell's equations on a circular domain with a piece-wise linear representation of the boundary. Dupont [34] has also suggested that rounding singular corners is necessary to suppress poor polynomial behaviour resulting from the high-order DG-FEM in his inter-model comparison of the oceanic shallow water equations. It is thus apparent that a high-order method begs for a smooth and accurate representation of the boundary, and hence, deformed or curvilinear elements along the boundary will be necessary to achieve accurate solutions on general lake geometries with the high-order discontinuous Galerkin method.

4.14.1 Constructing coordinates systems for curvilinear elements

We have adopted the approach in [60] that avoids some of the difficulties and cumbersome work associated with explicitly constructing two-dimensional mapping functions for high-order curvilinear elements, e.g., explicitly calculating high-order "shape-functions". The technique discussed here generalizes well to elements with an arbitrary number of nodes and thus allows for the robust construction of high-order curvilinear elements. The method

discussed here represents an extension of the technique used in [60] for circular boundaries, since we consider arbitrary domain boundaries represented by cubic splines.

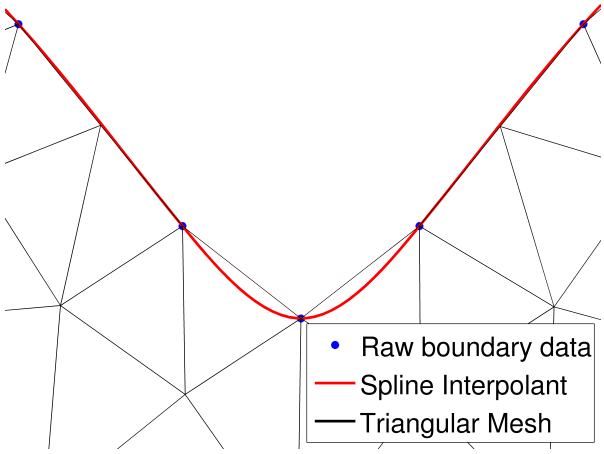

Figure 4.11: Illustration of straight-sided element mesh along with a smooth representation of the boundary, the spline interpolant, that will be used to produce deformed elements.

Assume we have generated a straight-sided finite element mesh that approximates the boundary in a piece-wise linear manner, and also assume we have retained a smooth representation of the boundary in a parameterized curve $\mathcal{C}$: $\mathbf{x}_b(t) = (x_b(t), y_b(t))$ (see Fig. 4.11), that we assume to be parameterized by arc-length $0 \leq t \leq S$. In practice, we have found taking $\mathcal{C}$ to be a parametric cubic-spline interpolant of the boundary to be a simple and effective choice. The algorithm for a particular element that is to be curved is as follows:

1. Adjust the straight-sided finite element mesh by moving the vertices (i.e., end-points only) of the straight-sided element's boundary edge e so that they lie exactly at points on $\mathcal{C}$, say $\mathbf{x}_b(t_1)$ and $\mathbf{x}_b(t_2)$

2. Distribute the 1D LGL nodes along the curved edge by arc-length using the parameterization $\mathbf{x}_b(t)$ for $t_1 \leq t \leq t_2$ to obtain new local coordinates along the curved edge, denoted $\mathbf{x}_{curved}(r,s)|_e$, where (r,s) are the coordinates of the reference triangle (see Fig. 4.12).

3. Calculate the deformation (displacement field) in moving from the edge nodes from the straight edge to the curve $\mathcal{C}$, i.e., $\mathbf{w}(r,s) = \mathbf{x}_{curved}(r,s)|_e - \mathbf{x}_{straight}(r,s)|_e$, also called the *warp* factor.

4. "Blend" the edge deformation to the interior nodes using Gordon–Hall blending (see below) to obtain new local coordinates for the whole element: $\mathbf{x}_{curved}(r,s) = \mathbf{x}_{straight}(r,s) + b(r,s)\mathbf{w}(r,s)$, where $b(r,s)$ is a *blending* function.

5. Compute local metric factors, i.e., x_r,y_r,x_s,y_s, and Jacobian $J = x_r y_s - x_s y_r$, numerically using the differentiation matrices on the reference element $\mathcal{D}_r$, $\mathcal{D}_s$.

The one point that requires further attention is how to choose a blending function $b(r,s)$ to appropriately "blend" the edge deformation on the element boundary to the interior of the element. To motivate our discussion, consider the simplistic one-dimensional case where two function values f_0 and f_1 are known at points x_0 and x_1 and we wish find a function $f(x)$ to interpolate to points inside the interval $[x_0, x_1]$. If, for additional simplicity, we assume $f_0 = 0$, we realize that the only way to interpolate to interior points with the information that we have is by the linear Lagrange interpolant $\ell_1(x) = (x - x_0)/(x_1 - x_0)$, i.e.,

$$f(x) = \left(\frac{x - x_0}{x_1 - x_0} \right) f_1 \ . \tag{4.93}$$

In a sense, we have found the appropriate blending function to be $\ell_1(x)$ since this function satisfies the desired properties: $\ell_1(x_1) = 1$, $\ell_1(x_0) = 0$.

Now consider the two-dimensional case where, for example, our edge deformation $\mathbf{w}(r,s)$ is known along the triangle edge corresponding to the line $r = -1$ for $-1 \leq s \leq 1$ on the reference element (Fig. 4.12). Clearly, we require the blending function to satisfy $b(r = -1, s) = 1$ since this is the only region where information is known. It also seems sensible that the effect of the edge-deformation would decay to zero at the opposite triangle edge that lies on the line $s = -r$ for $-1 \leq r \leq 1$, leading us to define the blending function as

$$b(r,s) = \left(\frac{s + r}{s - 1} \right) \ , \tag{4.94}$$

that satisfies $b(r = -1, s) = 1$ and $b(r, s = -r) = 0$, as required. The one issue that remains is the apparent singularity at the point $(-1, 1)$. This point corresponds to a location where $\mathbf{w} = 0$ since it is a vertex of the finite element mesh that does not need to be deformed. Thus, we can simply apply the blending at nodal points not corresponding to the singular point in step 4 above.

While the "blending" procedure discussed is a straight-forward extension of linear Lagrange interpolation to two-dimensions, one subtle difference between Lagrange interpolation is that the two-dimensional blending function is chosen to be zero or one along entire *line segments*, and not at points in space. It is for this reason that the technique has been

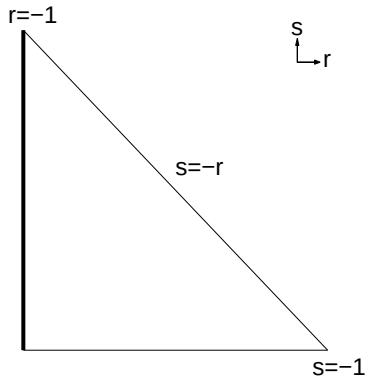

Figure 4.12: Diagram of the reference triangle and illustration of (r, s) coordinates.

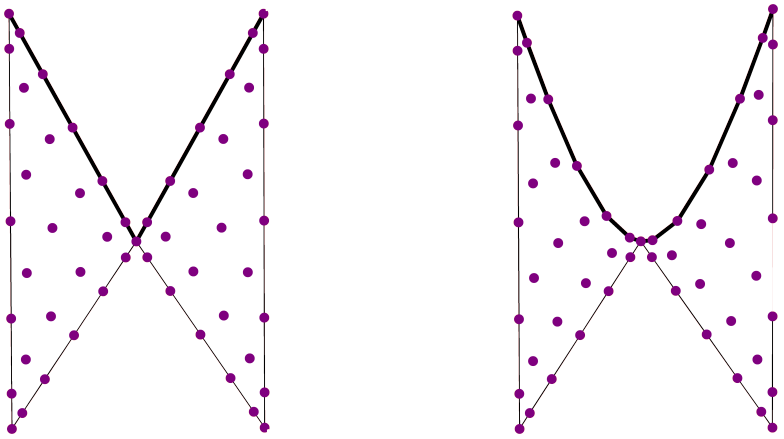

Figure 4.13: **Left**: A pair of elements before being deformed. **Right**: The same elements after being deformed to match the cubic-spline representation of the boundary with interior nodes re-distributed via Gordon–Hall blending.

referred to as "transfinite interpolation" [53], since in general the data is being sampled over a continuum and not just at a finite set of points.

4.14.2 Cubature and Quadrature Integration

The nodal approach described in the above sections relies heavily on the assumption that the Jacobian of the mapping from a particular element to the standard element is a constant, and hence may be brought outside of the integrals in the nodal DG-FEM formulation. Once we have introduced curvilinear elements, the Jacobian on these elements is no longer constant and we must thus pay a computational price. Firstly, a separate mass matrix must be stored for each curvilinear element, thereby driving up computational storage costs. Secondly, the Jacobians of the mappings used here are rational functions of the standard element's coordinates, and their product with solution fields will in general lead to aliasing errors that can drive numerical instabilities.

Nonlinearities involving rational functions cannot be de-aliased completely (as, for example, a quadratic nonlinearity could) since their polynomial representation would consist of a Taylor series with infinitely many terms. Nevertheless, a great deal of aliasing error can be removed by evaluating the integrals with cubature rules that are of higher order than the approximating polynomials. Here, 'cubature' refers to the higher-dimensional analogy to 1D quadrature rules. For polynomials of order N, we follow [60] and evaluate these integrals with cubature rules for the reference triangle of order $3(N+1)$. A general inner-product of two functions f and g are evaluated using a cubature rule by

$$\int_{\mathbf{D}^k} f g \, d\mathbf{x} \approx \sum_{i=1}^{N_c} f(\mathbf{r}_i^c) g(\mathbf{r}_i^c) J_i^k w_i^c \ , \tag{4.95}$$

where J_i^k is the Jacobian of the mapping from the standard element D^k, w_i^c are the cubature weights associated with cubature nodes $\{\mathbf{r_i^c}\}_{i=1}^{N_c}$. The cubature nodes and weights are provided by the symmetric rules in [127] and implemented in MATLAB in [60].

The use of cubature integration makes the evaluation of the local mass and stiffness matrices more computationally expensive, since additional interpolation operations must be carried out to interpolate integrands to the cubature nodes. In particular, we define the $N_c \times N_p$ interpolation matrix $V_{ij}^c = \ell_j(\mathbf{r}_i^c)$ to interpolate functions defined at the polynomial interpolation nodes to the cubature nodes. The $N_p \times N_p$ mass matrix can then be found

using cubature integration as follows

$$\mathcal{M}^k_{lm} = \int_{\mathbf{D}^k} \ell^k_l(\mathbf{x})\ell^k_m(\mathbf{x})\, d\mathbf{x} \tag{4.96}$$

$$\approx \sum_{i=1}^{N_c} \ell_l(\mathbf{r}^c_i)\ell_m(\mathbf{r}^c_i) w^c_i J^{k,c}_i \ , \tag{4.97}$$

Hence,

$$\mathcal{M}^k = (\mathcal{V}^c)^{\mathsf{T}}\, \mathcal{W}^k \mathcal{V}^c \ . \tag{4.98}$$

where $\mathcal{W}^k$ is the $N_c \times N_c$ diagonal matrix with entries $\mathcal{W}^k_{ii} = w^c_i J^{k,c}_i$. For the local stiffness matrix,

$$\mathcal{S}^k_{x,nm} = \int_{\mathbf{D}^k} \ell^k_n(\mathbf{x}) \frac{\partial \ell^k_m}{\partial x}(\mathbf{x})\, d\mathbf{x} \tag{4.99}$$

we must invoke the chain rule to express the operators in terms of the $N_p \times N_p$ differentiation matrices on the reference triangle, $\mathcal{D}_r$ and $\mathcal{D}_s$, yielding

$$\mathcal{S}^k_x = (\mathcal{V}^c)^{\mathsf{T}}\, \mathcal{W}^k \left(\mathrm{diag}(r^k_x(\mathbf{r}^c_i))\, \mathcal{V}^c \mathcal{D}_r + \mathrm{diag}(s^k_x(\mathbf{r}^c_i))\, \mathcal{V}^c \mathcal{D}_s \right) \ . \tag{4.100}$$

An identical argument gives

$$\mathcal{S}^k_y = (\mathcal{V}^c)^{\mathsf{T}}\, \mathcal{W}^k \left(\mathrm{diag}(r^k_y(\mathbf{r}^c_i))\, \mathcal{V}^c \mathcal{D}_r + \mathrm{diag}(s^k_y(\mathbf{r}^c_i))\, \mathcal{V}^c \mathcal{D}_s \right) \ . \tag{4.101}$$

In addition to volume (two-dimensional) integrals, surface integral (element-coupling) terms must also be computed using Gaussian quadrature, with analogous two-dimensional interpolation operators used to evaluate the integrand at the appropriate quadrature points along an edge. We again follow [60] and use order $N_G = 2(N+1)$ Gaussian quadrature along the edges.

4.15 Internal rotating seiche simulation using curvilinear elements

We now consider the same simulation shown in Section 4.13 where a circular basin has been perturbed to include a peninsula. The difference here is that we employ the developments

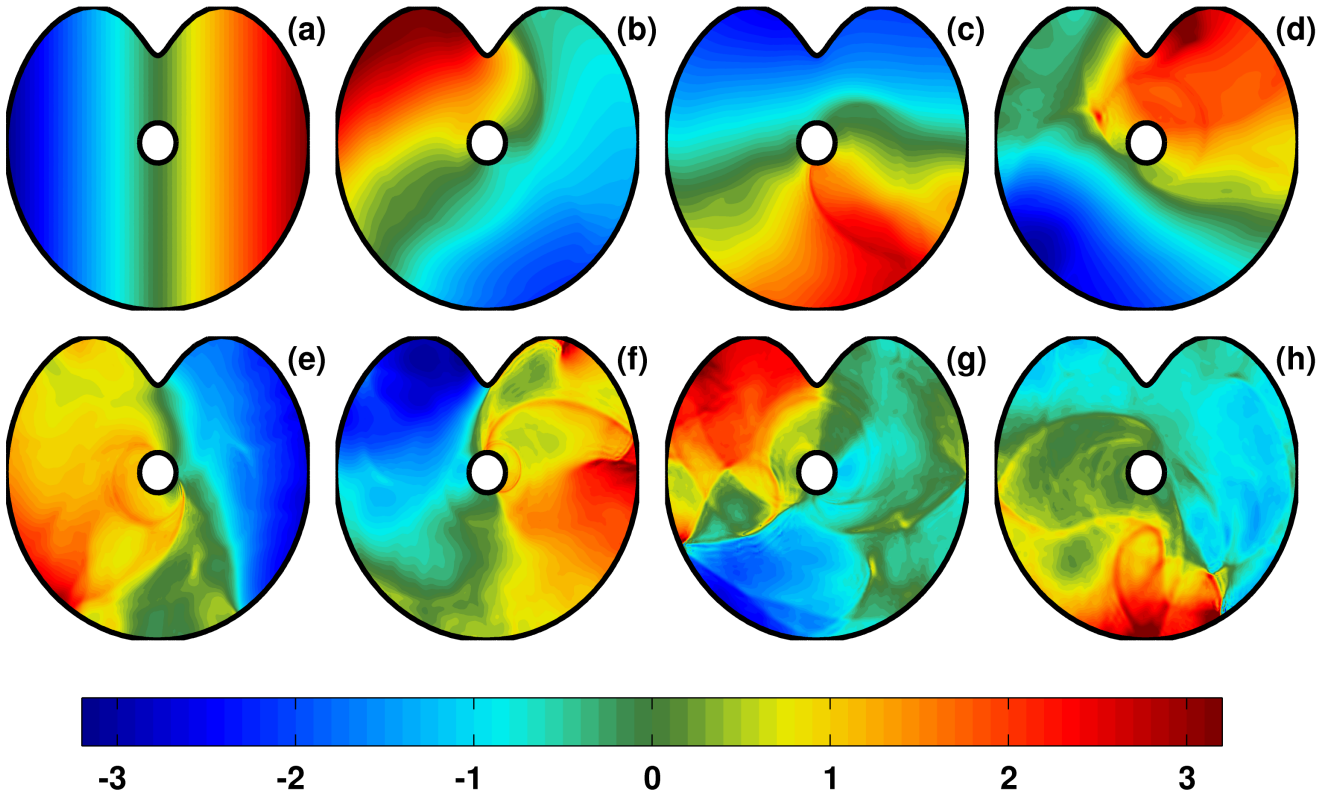

Figure 4.14: Panels **(a)**–**(d)**: Like Figure 4.9 but with curvilinear elements along the boundary. The other panels correspond to the later times **(e)** $t = 28.1$ h, **(f)** $t = 34.9$ h, **(g)** $t = 42.1$ h, **(h)** $t = 49.0$ h.

on curvilinear elements described in the above sections along with polynomial order $N = 8$. All boundary elements have been deformed such that their boundary edges conform to a cubic spline interpolant of the boundary.

The results of the simulation are shown in Figure 4.14. In addition to finding that the simulation is apparently long-term stable, unlike in the straight-sided case, we also note that the spurious eddies associated with the sharp re-entrant corner have been suppressed since the peninsula is now represented in a geometrically smooth way.

It is important to note that although the spurious eddies have been suppressed, the region of the flow at the tip of the peninsula still represents a geometric feature where a strong adverse pressure gradient must appear in order to decelerate flow around the obstacle. Indeed, an adverse η-gradient appears between $t = 2.5$ h and $t = 5.4$ h (not shown). It is possible that such high gradients in the physical fields can still lead to instabilities and this turns out to be the case when we explore two-layer flows in Section 5.6.

As discussed in Section 4.13, in real-world flows it is certainly reasonable to expect the flow around the peninsula to separate and generate eddies due to viscous boundary-layer effects, but since we have not included a physical model for such processes we are left in the somewhat precarious situation in which we demand the flow to remain 'attached' to the peninsula in all cases.

4.16 Internal rotating seiche simulation in a real-world lake

In this section, we provide proof-of-concept that the high-order DG-FEM methodology of this chapter can be applied to real-world lake geometries involving irregular coastlines. Bathymetry data at a resolution of 50 m for the mid-sized Pinehurst Lake, Alberta has been obtained from the Alberta Geological Survey website `http://www.ags.gov.ab.ca/`. Pinehurst Lake was chosen since its data set was freely available and its size is such that both rotation and stratification effects are expected to be important in the summer months. The raw data consists of a cartesian grid with 216×245 data points containing both land and water measurements. A plot of the 50 m bathymetry data is shown in Figure 4.15 where land values have been set to zero.

A parametric representation of the coastline was obtained using the data returned by MATLAB's `contour` function used to obtain the zero-depth contour and is shown in Figure 4.15**(b)**. It was found that finite element meshes generated from the raw data

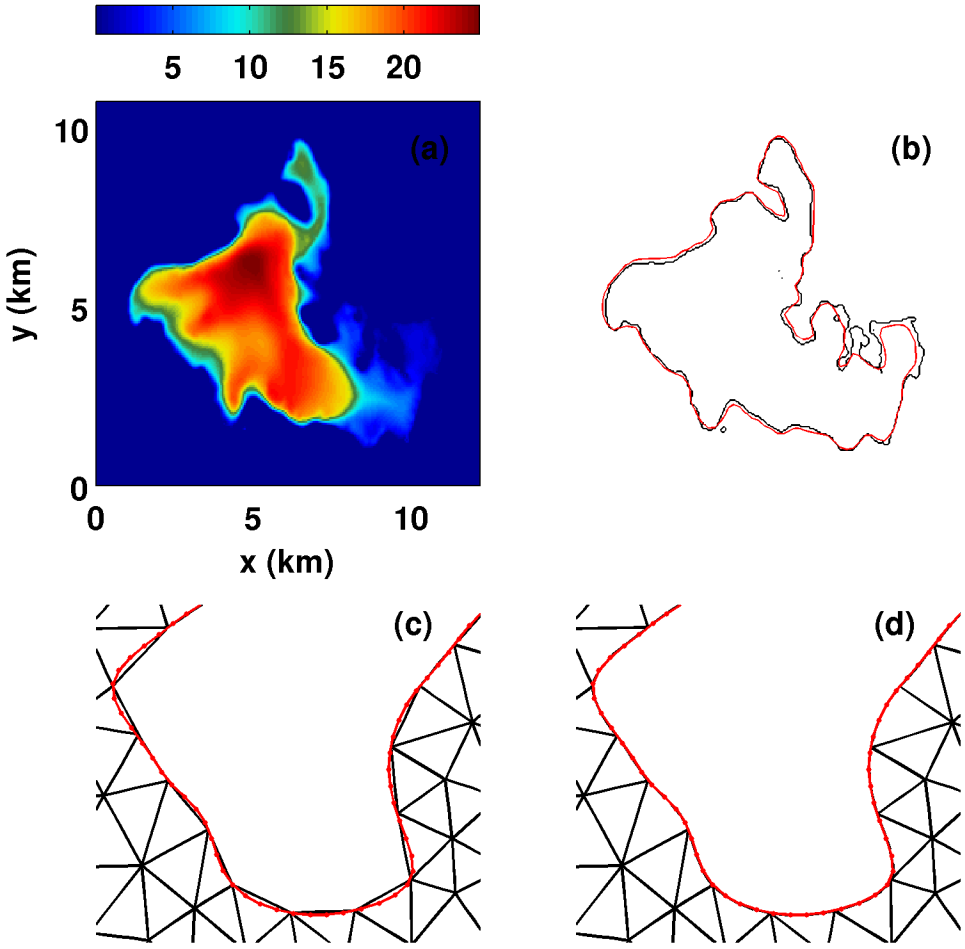

Figure 4.15: Panel **(a)**: Depth (in m) of Pinehurst Lake, AB from raw 50 m bathymetry data, and panel **(b)**: corresponding $H = 0$ contour (black) with smoothed coastline super-imposed (red). The lower panels show a zoomed-in section of the **(c)** straight-sided and **(d)** curved ($N = 6$) finite element mesh with $K = 1807$ elements near $(x, y) = (7 \text{ km}, 5 \text{ km})$ with cubic spline interpolant super-imposed (red).

contained $O(10,000)$ elements and possessed poor mesh quality (i.e., large aspect ratio triangles and large element size gradients) since the raw 0-depth contour is far from smooth. A smoothed piece of coastline is shown in Figure 4.15**(c)** with corresponding $N = 6$ curved finite element mesh in panel **(d)**. The smoothed coastline was found by convolving the two-dimensional bathymetry data with the 2D cardinal B-spline 16 times and sub-sampling the result to a 200 m resolution data set. A piece-wise cubic spline interpolant of the coastline was then constructed so that boundary elements could be deformed using the techniques explained in Section 4.14. The straight-sided finite element mesh, that is later deformed by our DG-FEM solver, was constructed using the open-source `gmsh` software [47] that was found to give better quality meshes than `mesh2d` in this case. Finally, the depth-profile $H(x, y)$ was linearly interpolated from the Cartesian data to our unstructured DG-FEM mesh for use during simulations. The depth-profile $H(x, y)$ was capped at a minimum depth of 6 m to avoid dry states that would drive the DG-FEM solver unstable.

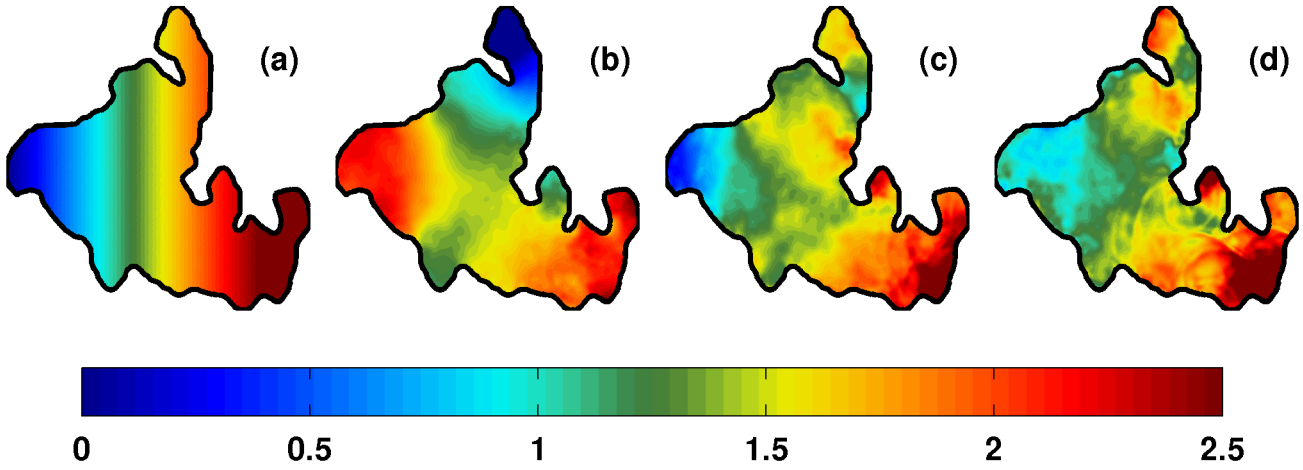

Figure 4.16: Evolution of an interfacial tilt in Pinehurst Lake, AB using the $N = 6$ DG-FEM with curvilinear boundary elements at times **(a)** $t = 0$ h, **(b)** $t = 19.4$ h, **(c)** $t = 39.3$ h, **(d)** $t = 62.7$ h.

As in previous simulations, the reduced gravity is $g' = (\Delta\rho/\rho_0)g = 0.024525$ ms^{-2}, where $(\Delta\rho/\rho_0) = 0.0025$. The Coriolis parameter was taken to be $f = 1.1863 \times 10^{-4}$ s^{-1}, corresponding to the 54.65° latitude of Pinehurst Lake. Results of an $N = 6$ DG-FEM simulation from an initial east-west interfacial tilt taken to increase linearly from $\eta = 0$ to $\eta = 2.5$ m are shown in Figure 4.16 that illustrates the evolving density interface at fixed-time snapshots with the initial condition plotted in panel **(a)**. Since the relative amplitude of the initial condition compared to the depth is, on average, not as large as in

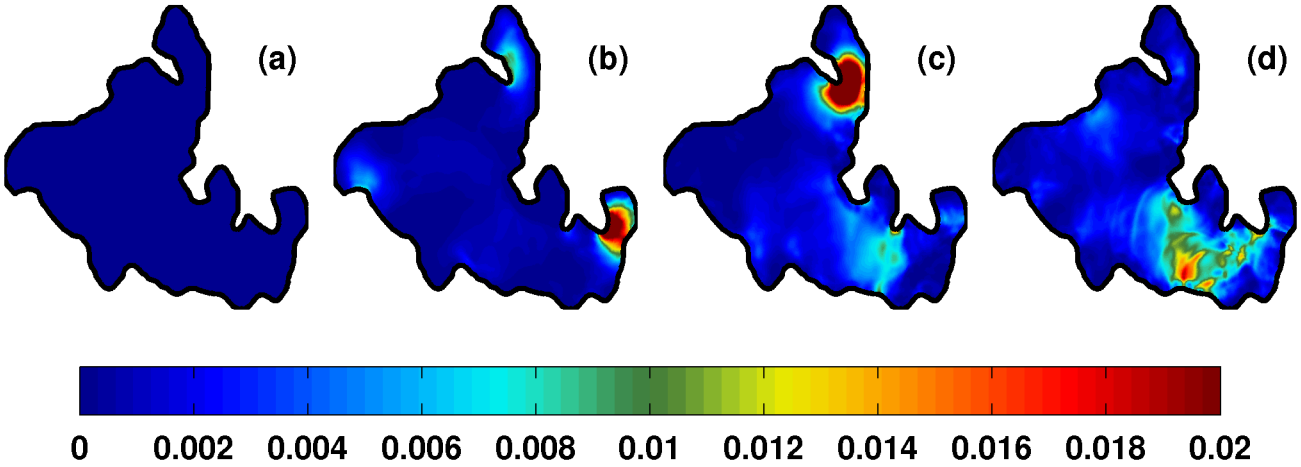

Figure 4.17: Like Fig. 4.16, except the kinetic energy density, $\frac{1}{2}h(u^2+v^2)$ is plotted.

previous simulations in this chapter, nonlinear effects are expected to be weaker. In spite of this fact, panels **(c)** and **(d)** show that nonlinear (high-amplitude) waves emerge in the shallows in the southeastern part of the basin after sufficient time has passed. As a result, small scale waves have profilerated throughout the entire basin by $t = 62.7$ h.

Figure 4.16 should be compared closely to Figure 4.17 where the kinetic energy density is plotted at the same times. At the earlier times (panels **(b)** and **(c)**), the most energetic features correspond to attached flow around peninsulas or other coastal obstacles. It is apparent that geometric focusing intensifies such features when they occur in narrow, confined parts of the basin. Panel **(d)** illustrates the kinetic energy fingerprint of small scale internal wave activity localized in the shallow eastern end of the lake at later times.

4.17 Chapter summary and conclusions

In this chapter, we have discussed solving dispersive shallow water models of the Boussinesq-type using the discontinuous Galerkin finite element method (DG-FEM) for general geometries. The DG-FEM represents a high-order alternative to finite volume or finite element methods that allows for both high-order polynomial approximations with upwind biased advection schemes [60].

The basics of the DG-FEM integral formulation for the dispersive shallow water system and its reduction to local matrix operators was discussed in Sections 4.1–4.6, and we

commented on the scaling of computational work with the DG-FEM in section 4.7. The time-stepping method used with the DG-FEM, that closely mirrors the scalar approach from Chapter 3, was presented in Section 4.8.

Comparisons between the pseudospectral methods of Chapter 3 and the DG-FEM methods developed here were carried out in 1D and 2D in Sections 4.9 and 4.11, respectively. From these comparisons, we conjecture that the DG-FEM can reach comparable resolution and energy-conserving characteristics to the Fourier method for sufficiently high polynomial order N. In Section 4.10, the impact of numerical flux choice on the DG-FEM schemes was explored by comparing the simple local Lax–Friedrichs flux to the HLLC flux. It was found that the benefits of a more sophisticated numerical flux (i.e., Riemann solver) are significant mainly in the low-order ($N = 1$) limit.

DG-FEM simulations in lakes with more general coastlines were considered beginning in Section 4.12 where a circular lake perturbed to have a wavy coastline was considered, and questions relevant to physical limnology applications were explored. It was then demonstrated that the DG-FEM is poorly behaved in the neighborhood of sharp re-entrant corners in Section 4.13, since sharp gradients and spurious eddies appear. An explanation of this phenomenon in terms of potential flow theory was offered. The remedy of rounding the corners using curvilinear elements along the boundary was proposed and the implementation was explained in Section 4.14. The conclusion we draw from that discussion is that general coastlines need a more computationally expensive treatment than simple circular geometries since the integrals in the DG-FEM formulations must be evaluated with cubature and quadrature rules of higher order than the approximating basis polynomials.

Finally, applications using the curvilinear element methodology were carried out. In Section 4.15, it was illustrated that the spurious eddies reported in 4.14 did not manifest when the re-entrant corner was represented in a smooth manner. The same methodology was then applied to the real-world situation of Pinehurst Lake, Alberta. The resulting high-resolution numerical solution was able to pinpoint a hot-spot of small-scale wave activity in the shallow eastern end of the basin. From this, we conclude that the DG-FEM solution of a weakly non-hydrostatic layered model may be a useful tool in helping to identify regions in lakes where internal wave-induced mixing is most dominant. The ability to identify such hot-spots has strong ecological consequences, as noted by Pannard *et al.* [91].

Chapter 5

On Discontinuous Galerkin Methods for Incompressible Flow

In this chapter, the possibility of using the discontinuous Galerkin finite element method (DG-FEM) for the solution to incompressible flow problems is explored. The idea of the pressure projection is revisited and an eigenvalue analysis of its DG-FEM discretization is carried out. The eigenvalue analysis demonstrates that the pressure projection does not properly enforce the incompressibility constraint on the velocity field. We attempt to correct this shortcoming by using a local post-processing projection to enforce incompressibility to numerical precision. However, our simulations suggest that this technique is not sufficient to guarantee stability and that there are other discretization issues at work. As a case study, simulations of an internal solitary wave as a solution to the inviscid incompressible Euler equations under the Boussinesq approximation are carried out. The work on the stratified Euler equations is motivated by the desire to solve the two-layer weakly non-hydrostatic shallow water equations under the rigid-lid approximation. Examples of successful discretization of these equations is presented along with a description of some of the pitfalls faced.

5.1 Motivation

Our main motivation comes from the desire to solve a two-layer dispersive shallow water system under the Boussinesq and rigid-lid approximation in general geometries. Equations

(2.196)–(2.200) can be written as

$$\frac{\partial h_i}{\partial t} + \nabla \cdot (h_i \mathbf{u}_1) = 0 , \quad i = 1, 2 \tag{5.1}$$

$$\frac{\partial (h_1 \mathbf{u}_1)}{\partial t} + \nabla \cdot (h_1 \mathbf{u}_1 \otimes \mathbf{u}_1) = -h_1 \nabla p , \tag{5.2}$$

$$\frac{\partial (h_2 \mathbf{u}_2)}{\partial t} + \nabla \cdot (h_2 \mathbf{u}_2 \otimes \mathbf{u}_2) = -h_2 \nabla p - g' h_2 \nabla \zeta + \gamma \nabla \left(\nabla \cdot (h_2 \mathbf{u}_2)_t \right) , \tag{5.3}$$

where $h_1(x, y, t) = H_1 - \zeta(x, y, t)$, $h_2(x, y, t) = H_2(x, y) + \zeta(x, y, t)$, and $h_1 + h_2 = h(x, y)$ is a constant with respect to time, and we have neglected rotation for the moment. As briefly discussed in Chapter 3, adding the two continuity equations (5.1) yields the incompressibility constraint on the barotropic transport

$$\nabla \cdot (h \mathbf{u}_{bt}) = 0 , \tag{5.4}$$

where $\mathbf{u}_{bt} = (h_1 \mathbf{u}_1 + h_2 \mathbf{u}_2)/h$. In incompressible flow problems, the constraint (5.4) is often imposed with a splitting/projection method [1, 67], where the rigid-lid pressure p is solved for from a Poisson equation such that eqn. (5.4) is satisfied at time level $t_{n+1} = t_n + \Delta t$. The full time-discrete algorithm is explained for the two-layer rigid-lid equations in Section 3.4.1.

The pressure projection method can be applied to any system that corresponds to incompressible flow. It is most often studied in the context of the incompressible Navier–Stokes equations or their inviscid counterpart, the incompressible Euler equations. For a large portion of this chapter, we consider the pressure projection (PP) method in the context of stratified incompressible flow under the Boussinesq approximation. We focus on difficulties encountered when using the standard PP for the INS equations in velocity-pressure formulation for long time integrations of non-hydrostatic stratified flow simulations under the Boussinesq approximation (BA). These difficulties are especially prevalent in under- to marginally-resolved inviscid or low-viscosity situations. The first contribution of this chapter is a spectral analysis of the DG PP operator that is furnished by the numerical calculation of the eigenvalues and eigenfunctions on a unit square domain with solid-wall boundary conditions. We consider the use of a local post-processing projection to exactly enforce incompressibility, in hopes that it can enhance the stability properties of the scheme. Following this, a series of numerical experiments are carried out and discussed for the (inviscid) incompressible Euler equations under the Boussinesq approximation. We conclude with a simulation of the two-layer weakly non-hydrostatic rigid-lid equations, and the implications of the results presented here are discussed.

5.2 Problem background

Recent interest in the possibility of using the discontinuous Galerkin (DG) method for numerical solutions to the unsteady incompressible Navier–Stokes (INS) equations has been sparked by its success in compressible flow simulations [9, 49, 60] and its many attractive features such as geometric flexibility, stencil locality/compactness, upwind-biased fluxes for advection-dominated flows, and high-order accuracy [60, 18, 20]. The desire for high-order DG solutions is further motivated by the Ladyzhenskaya-Babuska-Brezzi (LBB) or inf-sup stability problem that plagues many element-based formulations due to spurious pressure modes, resulting in stability for only certain spatially mixed-order velocity-pressure formulations [24, 58, 64, 65].

However, since the INS equations do not comprise a hyperbolic system it is unclear if the DG method can be used to recover consistent, stable, and accurate numerical solutions in general, since DG methods only weakly impose continuity across elemental interfaces, typically by solving Riemann problems to specify an appropriate numerical flux function [60]. Indeed, for the spectral element ocean model (SEOM), Levin *et al.* [77] chose a DG formulation for scalar transport equations but not for the momentum equations, citing the ill-posedness of the Riemann problem as a central difficulty. On the other hand, some studies have successfully obtained DG solutions, using the PP method and equal-order approximations, to some standard test cases for the INS equations, such as the Taylor-Green Vortex [60, 46, 105], laminar Kovasznay flow [60], and flow past square [46, 105] and circular [60] cylinders. These test cases, however, typically focus on short time integrations or highly viscous/damped situations, and the long-term stability properties of the methods remain unclear.

More work needs to be done to explore weakly damped and inviscid cases with DG, since some promise has already been shown using the closely-related spectral multi-domain penalty method (SMPM) in the vertical coordinate for simulations of stratified turbulence in incompressible flow [30]. The SMPM differs from the DG method only in the sense that a collocation formulation is used instead of a Galerkin formulation, but the other basic concepts of the schemes (e.g., locally high-order polynomial methods, continuity between subdomains only weakly enforced) are the same [38]. The DG and SMPM methods are compared directly for the shallow water equations by Escobar-Vargas *et al.* [39].

5.3 Methods

The non-dimensional stratified INS equations under the BA are [70]

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u} \otimes \mathbf{u}) = -\nabla p + \frac{1}{Re}\nabla^2 \mathbf{u} - \frac{1}{Fr^2}\rho \mathbf{k} \,, \tag{5.5}$$

$$\nabla \cdot \mathbf{u} = 0 \,, \tag{5.6}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \,, \tag{5.7}$$

corresponding to conservation of momentum (5.5) and mass (5.6), and scalar transport (5.7). Here, $\mathbf{u}$ is the dimensionless velocity field with components $(u(\mathbf{x},t), w(\mathbf{x},t))$ (and $\mathbf{x} = (x,z)$) in two dimensions, $\mathbf{k}$ is the unit vector pointing along the positive z-axis, and $p = p(\mathbf{x},t)$ and $\rho = \rho(\mathbf{x},t)$ are the dimensionless pressure and density, respectively. $Re = UL/\nu$ is the Reynolds number as a function of kinematic viscosity ν and typical velocity and length scales U and L, respectively. The Froude number is defined by $Fr = U/\sqrt{gL}$, where g is the gravitational acceleration, and we have chosen the advective time-scale $T = L/U$ in our non-dimensionalization. The inviscid incompressible Euler (IE) equations are recovered in the limit $Re \to \infty$, and the case of non-buoyancy-driven flow is recovered by taking $Fr \to \infty$. The essence of the Boussinesq approximation (BA) is that in the momentum equations, the density is taken as a constant in all but the buoyancy term.

The PP time-stepping algorithm has been thoroughly explained in the literature (see e.g., [1]). A very popular approach is the high-order stiffly-stable splitting algorithm due to Karniadakis *et al.* [67]. The stiffly-stable time-stepping scheme takes the following general form

$$\frac{\left(\gamma_0 \mathbf{u}^{n+1} - \sum_{q=0}^{N_q-1} \alpha_q \mathbf{u}^{n-q}\right)}{\Delta t} = \sum_{q=0}^{N_q-1} \beta_q \mathbf{N}(\mathbf{u}^{n-q}, \rho^{n-q}) - \nabla p^{n+1} + \frac{1}{Re}\nabla^2 \mathbf{u}^{n+1} \,, \tag{5.8}$$

where

$$\mathbf{N}(\mathbf{u}, \rho) = -\nabla \cdot (\mathbf{u} \otimes \mathbf{u}) - \frac{1}{Fr^2}\rho \mathbf{k} \,. \tag{5.9}$$

are the nonlinear advection and buoyancy source terms. The coefficients γ_0, α_q and β_q for orders $N_q = 1$ to 3 are listed in [67]. The basic idea behind the scheme is to discretize the implicit pressure and viscous terms with a standard stiffly-stable backward differentiation formula (see, e.g., [76]) and apply a linear multi-step method to the advection and source terms.

The first step in the splitting scheme is to form a predicted velocity by evolving the explicit terms using information from the previous time-levels.

$$\hat{\mathbf{u}} = \frac{1}{\gamma_0}\left[\sum_{q=0}^{N_q-1}\alpha_q \mathbf{u}^{n-q} + \Delta t \sum_{q=0}^{N_q-1}\beta_q \mathbf{N}(\mathbf{u}^{n-q}, \rho^{n-q})\right] . \tag{5.10}$$

The next step is to project the predicted velocity $\hat{\mathbf{u}}$ onto the space of divergence-free velocities [1]. This projection can be achieved by considering the pressure split-step

$$\hat{\hat{\mathbf{u}}} = \hat{\mathbf{u}} - \frac{\Delta t}{\gamma_0}\nabla p^{n+1} . \tag{5.11}$$

Taking the divergence of (5.11) results in the following Poisson equation for the pressure at time $t_{n+1} = t_n + \Delta t$:

$$\nabla\cdot\hat{\hat{\mathbf{u}}} - \nabla\cdot\hat{\mathbf{u}} = -\nabla^2\tilde{p}^{n+1} . \tag{5.12}$$

where we've introduced the re-scaled pressure $\tilde{p} = (\Delta t/\gamma_0)p$ for convenience. Here $\hat{\hat{\mathbf{u}}}$ is an intermediate velocity that is formed after the pressure terms have been evolved, but before the viscous step. The constraint (5.6) is enforced by removing the $\nabla\cdot\hat{\hat{\mathbf{u}}}$ term in (5.12):

$$\nabla\cdot\hat{\mathbf{u}} = \nabla^2\tilde{p}^{n+1} . \tag{5.13}$$

thereby projecting the approximate solution onto the space of approximately non-divergent velocity fields. The Poisson equation (5.13) must be solved along with boundary conditions. On solid walls, an appropriate condition can be found by enforcing no normal flow on the corrected velocity field in (5.11). Taking the dot product of (5.11) with the unit outward normal $\hat{\mathbf{n}}$ along the boundary $\partial\Omega$ and taking $\hat{\hat{\mathbf{u}}}\cdot\hat{\mathbf{n}} = 0$ gives

$$\frac{\partial\tilde{p}^{n+1}}{\partial n} = \hat{\mathbf{u}}\cdot\hat{\mathbf{n}} , \quad \text{on} \quad \partial\Omega . \tag{5.14}$$

Karniadakis *et al.* [67] have argued this boundary condition should also account for the viscous term $(1/Re)\nabla^2\mathbf{u}^{n+1}$ in (5.8) to not compromise on the high-order accuracy of the splitting scheme. However, since $\mathbf{u}^{n+1}$ is not known, it must be extrapolated from previous time-levels. The suggested boundary condition is thus

$$\frac{\partial\tilde{p}^{n+1}}{\partial n} = \hat{\mathbf{u}}\cdot\hat{\mathbf{n}} - \frac{\Delta t}{\gamma_0 Re}\sum_{q=0}^{N_q}\beta_q\left[\nabla\times\left(\nabla\times\mathbf{u}^{n-q}\right)\right]\cdot\hat{\mathbf{n}} , \tag{5.15}$$

where the identity $\nabla^2\mathbf{u}^{n+1} = \nabla(\nabla\cdot\mathbf{u}^{n+1}) - \nabla\times(\nabla\times\mathbf{u}^{n+1}) = -\nabla\times(\nabla\times\mathbf{u}^{n+1})$ has been applied with incompressibility enforced so that any contributions of non-zero divergence

from previous time-levels is removed. It is worth mentioning that the simpler condition (5.14) is recovered in the inviscid limit, $Re \to \infty$.

Once $\tilde{p}^{n+1}$ has been computed, the corrected (divergence-free) velocity field is obtained from eqn. (5.11), i.e.,

$$\hat{\hat{\mathbf{u}}} = \hat{\mathbf{u}} - \nabla \tilde{p}^{n+1} \, . \tag{5.16}$$

The final step is to evolve the viscous terms that have been implicitly discretized in time. From equation (5.8), we have

$$\left[1 - \left(\frac{\Delta t}{\gamma_0 Re} \right) \nabla^2 \right] \mathbf{u}^{n+1} = \hat{\hat{\mathbf{u}}} \, , \tag{5.17}$$

usually subject to the no-slip boundary condition $\mathbf{u} = 0$ on solid boundaries. In the DG-FEM framework, the Helmholtz operator on the right-hand side is typically discretized using the SIP-DG discretization method [60, 46, 105] that was introduced in Section 4.6.

In the context of the DG-FEM, the difficulty highlighted in this work stems from the discretization of the left-hand side of (5.12), and we pre-suppose that the right-hand side is discretized via the nodal symmetric interior penalty DG (SIP-DG) method applied to the Laplacian. The weak DG formulation of the left-hand side can be found by considering the local solution to (5.12) on a particular element (or sub-domain of Ω) D^k (where $k = 1, \cdots, K$), multiplying by a member of the space of local test-functions $\{\ell_j^k\}_{j=1}^{N_p}$ and integrating by parts to yield

$$\left(\int_{\partial D^k} \left(\ell_j^k \hat{\hat{\mathbf{u}}} \right)^* \cdot \hat{\mathbf{n}} \, d\mathbf{x} - \int_{D^k} \hat{\hat{\mathbf{u}}}^k \cdot \nabla \ell_j^k \, d\mathbf{x} \right) - \left(\int_{\partial D^k} (\ell_j \hat{\mathbf{u}})^* \cdot \hat{\mathbf{n}} \, d\mathbf{x} - \int_{D^k} \hat{\mathbf{u}}^k \cdot \nabla \ell_j^k \, d\mathbf{x} \right) \, , \tag{5.18}$$

where superscript $*$ denotes an appropriate numerical flux function chosen to impose weak continuity across element interfaces in a way that is consistent with the underlying dynamics of the INS equations.

We notice that the PP method results in simply removing the first two terms in eqn. (5.18). Thus, the divergence-free constraint is only being imposed upon the corrected velocity $\hat{\hat{\mathbf{u}}}$ in a weak and local sense, and the overall impact on the resulting DG-FEM scheme remains unclear. There has been some suggestion that the standard PP scheme does not result in sufficient coupling between the velocity and pressure [8, 10] fields at element interfaces. Here, we attempt to understand the effects of using the PP with DG via a numerical eigenvalue analysis as explained below.

5.3.1 Numerical Method for the Eigenvalue Analysis

Beginning with the nodal DG implementation of the INS solver presented in [60], we have computed the eigenvalues of the PP operator $\mathbb{P} : \hat{\mathbf{u}} \mapsto \mathbf{u}^{n+1}$ that carries out the following linear operations in a single MATLAB function:

1. Given the input $\hat{\mathbf{u}}$, solve for $\tilde{p}^{n+1}$ using the SIP-DG discretization of the Poisson problem (5.13).

2. Calculate the discretized form of $\nabla\tilde{p}^{n+1}$, and use eqn. (5.16) to obtain $\hat{\hat{\mathbf{u}}}$.

3. If Re is finite, advance the viscous term using SIP-DG discretization of the viscosity operator (5.17) to recover the output, $\mathbf{u}^{n+1}$.
   If Re is infinite, set $\mathbf{u}^{n+1} = \hat{\hat{\mathbf{u}}}$.

The corresponding function-handle is then passed into MATLAB's `eig` eigenvalue solver that calculates eigenvalues using the implicitly-restarted Arnoldi method implemented in the `ARPACK` Fortran77 library [73]. It is worth noting that since the pressure variable $\tilde{p}$ is not a prognostic flow variable and its purpose is to simply enforce incompressibility on $\mathbf{u}$ (see [1]), it is treated as an auxiliary field by the $\mathbb{P}$ operator.

The domain under consideration is the closed unit square $\Omega = [0, 1]^2$ subject to no normal flow (no slip) boundary conditions for the inviscid (viscous) case. Since the $\mathbb{P}$ operator only takes $\hat{\mathbf{u}}$ as input, we impose the simple inviscid pressure boundary condition (5.14) in solving the Poisson problem (5.12) in all cases. For a domain consisting entirely of solid walls, only Neumann boundary conditions are imposed on pressure and there is no unique solution to (5.12) [67]. To address this issue, we have adopted the approach of [124], where a small additive scalar unknown is added to (5.12) in order to impose the additional constraint of zero mean pressure. This approach has been introduced earlier in Section 3.4.2. Alternatives to this approach, including null singular vector removal, are possible as well. See [38] for an overview in the context of the SMPM method.

Equation (5.12) itself is solved by computing the LU-factorization of the discrete Laplacian during pre-processing for re-use during each `eig` iteration. Throughout this chapter, we have chosen to solve the linear systems directly. We leave the issue of using an iterative linear solver, and the associated complexities (e.g., pre-conditioning), to a future work.

5.4 Results

5.4.1 DG Simulations using the PP method

We have carried out long-time integrations of the homogeneous ($\rho = constant$, $Fr \to \infty$) form of (5.5)-(5.6) using the nodal DG implementation in [60] and found spikes in the solution that form at element interfaces and eventually lead to numerical instability. Modal filtering [60] alleviates the issue somewhat, but does not appear to prevent instabilities in general. Plotting the divergence $\nabla \cdot \mathbf{u}$ suggests the issue may be related to spurious compressibility artifacts near element interfaces that are not zero to numerical precision.

For stratified flow simulations under the BA, the situation is worse since the presence of an active density tracer, ρ, in the vertical momentum equation implies that any numerically-driven perturbation to ρ will cause spurious vertical motion. In under-resolved cases, we found that spurious compressions caused regions of high density to artificially emerge over regions of low density at certain element interfaces, resulting in unphysical grid-scale Rayleigh-Taylor (RT) instabilities [70] that destroyed the numerical solution. In marginally- to well-resolved simulations, the unphysical RT instabilities appear to be suppressed. However, the long-term stability properties are again uncertain due to non-zero divergence.

5.4.2 Spectral analysis of the DG PP operator and proposed remedies to the problem

The unit square domain Ω was partitioned into 8 uniform triangular elements, and the eigenvalues and eigenfunctions (velocities), $(\lambda_i, \mathbf{u}_{\phi_i})_{i=1}^{N_e}$ of the PP operator $\mathbb{P}$ were computed for polynomial order $N = 8$ corresponding to $N_p = 45$ nodal points on each triangle [60], for a total of 360 grid points yielding $N_e = 720$ eigenmodes. The eigenspectrum was computed for a variety of Re as well as for the inviscid case. Each velocity field was scaled such that $\int_\Omega \|\mathbf{u}_{\phi_i}\|^2 \, d\mathbf{x} = 1$.

To assess the incompressibility properties of the eigenmodes, in Fig. 5.1 we plot the quantity

$$D_i = \left(\int_\Omega \left[\nabla \cdot \mathbf{u}_{\phi_i} \right]^2 \, d\mathbf{x} \right)^{\frac{1}{2}} \left[\max_{1 \le j \le N_e} \left(\int_\Omega \left[\nabla \cdot \mathbf{u}_{\phi_j} \right]^2 \, d\mathbf{x} \right)^{\frac{1}{2}} \right]^{-1}, \tag{5.19}$$

the scaled L^2-norm of divergence in each velocity eigenfunction $\mathbf{u}_{\phi_i}$ against its corresponding eigenvalue, λ_i. The integrals in eqn. (5.19) were evaluated using the orthogonality of the basis functions in the local modal expansion.

Perhaps the most revealing result in Fig. 5.1 is the inviscid case shown in panel **(c)**. Since the inviscid $\mathbb{P}$ operator includes only the projection and not viscosity, we find that there are only two possible eigenvalues, $\lambda = 1$ and $\lambda = 0$, each with its own collection of corresponding eigenfunctions. To understand the structure of the degeneracy of the eigenspectrum of the inviscid projection operator, it is useful to recall the Helmholtz decomposition that guarantees any vector field may be decomposed into the sum of curl-free and divergence-free components

$$\mathbf{u} = -\nabla\varphi + \nabla \times (\psi\mathbf{j}) \ , \tag{5.20}$$

where the unit vector $\mathbf{j} = (0, 1, 0)$ appears since we assume $\mathbf{u}$ lies in the xz-plane. The $\mathbb{P}$ operator should map vector fields of the general form (5.20) to the divergence-free part $\nabla \times (\psi\mathbf{j})$. Two special cases that arise from this fact are: (1) If $\mathbf{u}_{\phi_i}$ is an eigenfunction of $\mathbb{P}$ with no curl-free part, then $\mathbb{P}$ should not change it, hence $\lambda = 1$, and (2) if $\mathbf{u}_{\phi_i}$ is an eigenfunction of $\mathbb{P}$ with no divergence-free part, then $\mathbb{P}$ should map it to $\mathbf{0}$, i.e., it belongs in the null space ($\lambda = 0$) of $\mathbb{P}$. Therefore, the inviscid $\lambda = 0$ eigenmodes should be interpreted as a basis of curl-free velocities while the inviscid $\lambda = 1$ eigenmodes should be interpreted as a basis of divergence-free velocities.

The key thing to notice here is that the discretized eigenfunctions forming a basis for incompressible velocity fields are themselves not divergence-free since they have $D_i \neq 0$. In Fig. 5.2, we show contour plots of the scaled absolute value of the divergence of two such eigenfunctions to illustrate this undesired behaviour that is worst at element interfaces. In light of these results, it is not clear if we can represent an incompressible velocity field using such a basis and expect it to be genuinely divergence-free. The finite Re cases in Fig. 5.1 show that viscosity introduces eigenmodes with $0 < \lambda < 1$. These eigenvalues between 0 and 1 should be expected since in the simplified case of periodic boundary conditions, the viscosity operator would effectively multiply the sinusoid e^{ikx} by $[1 - \Delta t k^2/(\gamma_0 Re)]$. Although viscosity does not yield eigenmodes with smaller values of D_i, sufficiently small Re will ensure that the eigenspectrum is structured such that smaller eigenvalues ($\lambda_i < 1$) are assigned to eigenfunctions with larger values of D_i. This can be seen as beneficial since the most poorly behaved eigenmodes are marginalized in the eigen decomposition of $\mathbb{P}$. Of particular note is the absence of $O(1)$ values of D_i for $\lambda_i \approx 1$ in the $Re = 1$ and $Re = 40$ cases. Despite these results, it remains unclear how much viscosity is required to attain long term stability in general because viscosity does not correct the problem of

spurious divergence. Instead, it provides a rather general mechanism for damping small-scale numerical artifacts. The question of how strong the viscosity operator should be to attain stability has been recently examined by Ferrer [45].

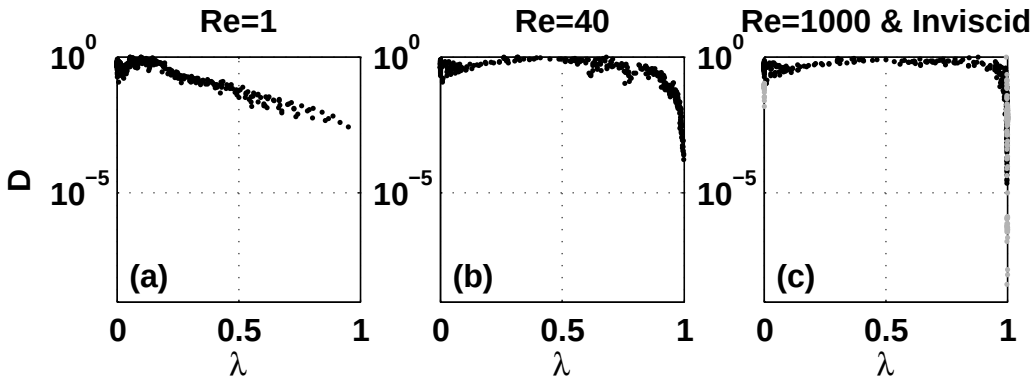

Figure 5.1: D_i, eigenfunction's scaled maximum L^2-norm of divergence vs. corresponding eigenvalue λ_i at selected Re and in the inviscid case, $Re \to \infty$ (panel **(c)**, grey dots). For finite Re, we set $\Delta t/\gamma_0 = 10^{-3}$.

Approaches adopted in the literature to circumvent the problem highlighted above avoid the PP altogether, and ensure that the weak form of the divergence-free constraint (5.6) explicitly appears in the scheme along with a suitable numerical flux function [8, 10, 19]. Cockburn *et al.* [19] use a pressure stabilization term in their numerical flux choice for the weak DG form of (5.6) along with a post-processing procedure to obtain exactly non-divergent approximate velocity fields to solve the steady INS equations. The method of Bassi *et al.* [8, 10] recovers a well-posed Riemann problem in the imposition of (5.6) by considering a numerical flux function from the artificial compressibility equations that are the same as (5.5)-(5.7) except (5.6) is replaced with

$$\epsilon^2 \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{u} \;=\; 0 \,, \tag{5.21}$$

where $\epsilon = U/c$ is a Mach number and c is an artificial sound speed. However, the method appears somewhat costly since all terms are discretized implicitly in time and exact Riemann problems must be solved numerically by nonlinear Newton iterations at each time-step. Other possibilities lie with the recently discussed class of hybridizable DG (HDG) methods [89, 102], that impose strong continuity only in the normal component of numerical fluxes. Finally, for strictly two-dimensional flow, a streamfunction-vorticity formulation could be adopted. This idea has been explored in a DG context by Liu and Shu [78].

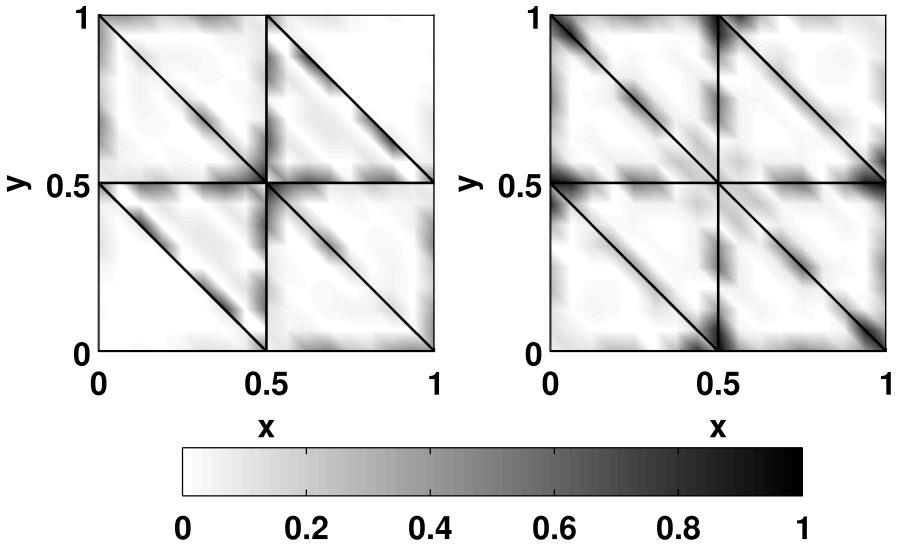

Figure 5.2: Absolute value of divergence $|\nabla \cdot \mathbf{u}_{\phi_i}|$ (re-scaled to have maximum value of 1) for two selected eigenfunctions of the inviscid $\mathbb{P}$ operator (see Fig. 5.1 **(c)**) with corresponding eigenvalue $\lambda_i = 1$ (to within 5 decimal places).

In the present work, we have followed up on the theoretical developments in [19] wherein an inf-sup condition was derived for the DG-discretized INS equations, and it was proven that stable equal-order DG schemes for the steady INS equations require (i) a pressure stabilization term in the DG discretization of (5.6), and (ii) a local post-processing operation to recover an exactly non-divergent velocity field from the weakly non-divergent velocity field. Here, we explored these ideas in the unsteady inviscid case. In the SIP-DG framework discussed above, condition (i) is already satisfied since the SIP-DG discretization of the Laplacian uses a stabilization (or penalty) term in the numerical flux function of an auxiliary vector variable $\mathbf{q} = \nabla p$ to penalize large jumps in p. Here, the SIP-DG numerical flux functions are given by $p^* = \{\!\{ p \}\!\}$ and $\mathbf{q}^* = \{\!\{\nabla p\}\!\} - \tau [\![p]\!]$, where $\tau > 0$ is the penalty parameter [60], and the operators $\{\!\{\cdot\}\!\}$ and $[\![\cdot]\!]$ denote the average and jump across an interface, respectively. Therefore, it appears that the missing ingredient in the scheme discussed above is the local post-processing projection, and we have sought to rectify this issue.

A locally non-divergent velocity basis on the reference element can be constructed by an

appropriate differentiation of the modal basis functions ψ_i that, as in [60], are orthogonal polynomials of order N. In two-dimensions, the velocity basis is taken to be

$$\mathbf{u}_{\psi_i} = \nabla \times (\psi_i \mathbf{j}) \ , \quad i = 1, \cdots, N_m \ , \tag{5.22}$$

where N_m is the number of local modal basis functions (we omit the $\psi_i = constant$ mode), and $\mathbf{j} = (0, 1, 0)$. The unit vector $\mathbf{j}$ appears here since the resulting velocity basis should lie in the xz-plane, i.e., the same plane as the velocity field. By construction, we have $\nabla \cdot \mathbf{u}_{\psi_i} = 0$ for $i = 1, \cdots, N_m$, and we have verified that the local discrete differential operators satisfy this condition to numerical precision. Although the basis (5.22) is a set of vectors and not scalars as is common with DG and CG methods, a Galerkin-type projection can be furnished on the reference element by considering an arbitrary divergence-free velocity field $\mathbf{v}$ expanded in terms of the basis (5.22):

$$\mathbf{v} = \sum_{j=1}^{N_m} c_j \mathbf{u}_{\psi_j} \ , \quad j = 1, \cdots, N_m \ . \tag{5.23}$$

The c_j's can be computed by taking the dot product of (5.23) with a member of the divergence-free velocity basis $\mathbf{u}_{\psi_i}$, integrating over the reference element Γ, and inverting the resulting linear system of equations

$$\sum_{i=1}^{N_m} M_{ij} c_j = l_i \ , \tag{5.24}$$

where

$$M_{ij} = \int_\Gamma \mathbf{u}_{\psi_i} \cdot \mathbf{u}_{\psi_j} \, d\mathbf{x} \ , \quad l_i = \int_\Gamma \mathbf{u}_{\psi_i} \cdot \mathbf{v} \, d\mathbf{x} \ . \tag{5.25}$$

Therefore, any arbitrary velocity $\mathbf{u}$ defined on the reference element can be projected onto the space of exactly non-divergent velocities by solving (5.24), with $\mathbf{v}$ replaced by $\mathbf{u}$ in (5.25), and summing the right-hand side of (5.23) to recover $\mathbf{v}$, an exactly divergence-free approximation to the non-divergent part of $\mathbf{u}$. As in [19], the operation of mapping $\mathbf{u}$ to $\mathbf{v}$ is completely local and can be carried out in an element-by-element fashion once the local velocity on each element has been transformed to the standard element's coordinate system. Should this projection be extended to three dimensions, it is worth noting that special care should be taken to ensure that the solenoidal basis spans all of three-space. The basis of divergence-free vectors would need to be made as much as three times larger than the corresponding three-dimensional scalar basis.

In Fig. 5.3, we show the strong impact that applying the post-processing projection has on the eigenspectrum of the $\mathbb{P}$ operator. All eigenmodes except those corresponding to the null space of $\mathbb{P}$ satisfy $\nabla \cdot \mathbf{u}_\phi = 0$ to numerical precision.

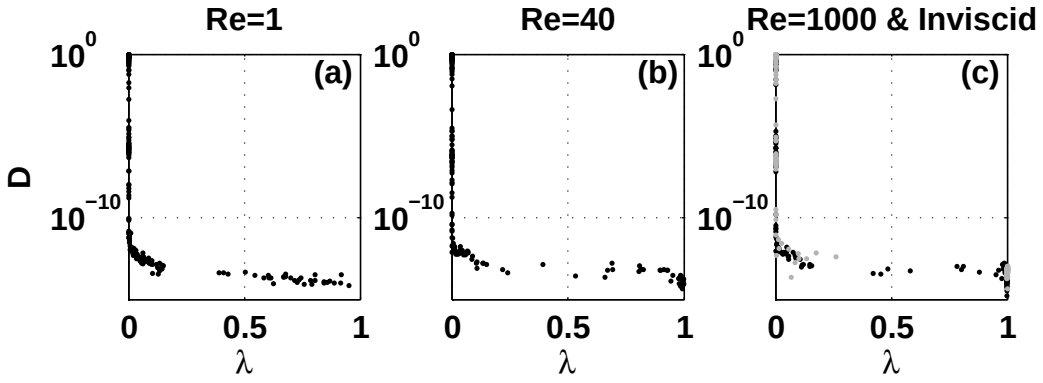

Figure 5.3: Like Fig. 5.1, but the post-processing operator has been applied at the end of the usual $\mathbb{P}$ operation.

5.5 Inviscid stratified incompressible Euler simulations in two space dimensions

Since we are primarily interested in the ability to carry out inviscid simulations, we have modified the INS solver presented in [60] to solve the incompressible Euler (IE) equations under the Boussinesq Approximation (BA). Since the full high-order splitting algorithm (5.8) is intended for use in viscous simulations (finite Re), it cannot be directly applied to the inviscid case. An appropriately modified scheme removes the backward differentiation formula (BDF) coefficients, i.e., $\alpha_q = 0$ for all indices q) $\gamma_0 = 1$, and the linear multistep coefficients β_q can be taken to be given by the third-order Adams-Bashforth method (See Table 3.2). The Lax-Friedrichs advective flux is employed in the DG discretization of the advective terms, as in [60]. An exponential cut-off filter function is applied to the local modal coefficients of the full solution fields after the advective step to prevent aliasing and nonlinearity errors from driving weak instabilities (see Section 3.2.3).

We have successfully validated the method for the stratified IE BA equations against the Fourier spectral method benchmark laboratory-scale internal solitary wave (ISW) solutions to the Dubreil-Jacotin-Long (DJL) equation found in [33]. Although the solution

profile simply translates to the right with constant speed $c = 0.1042$ m s^{-1}, the dynamics are driven by nonlinear and non-hydrostatic effects. The geometry is given by a simple rectangle with dimensions 5 m $\times$0.15 m taken to be periodic in x and bounded by rigid horizontal walls at $z = 0$ and $z = -0.15$ m; nevertheless, this test-case is challenging since the horizontal:vertical aspect ratio is $\sim$33:1, and high amounts of vertical resolution are required to properly resolve the thin pycnocline.

We have constructed a structured grid of $1,089$ rectangular elements and carried out DG simulations using polynomial orders $N = 4$, $N = 6$, and $N = 8$, both with the local projection (WLP) discussed in Sec. 5.4.2 and without (NLP). Here, we have used rectangular elements since we were able to reach the requisite resolution with a smaller number of elements than would be necessary with a structured or unstructured triangular mesh. The basis functions were thus taken to be the two-dimensional tensor product of the one-dimensional Legendre polynomials [68], and the rectangular nodal points were taken to be given by the tensor-product of two one-dimensional Legendre-Gauss-Lobotto (LGL) grids. For the cases considered here, the added computational time associated with including the local projection was negligible. The local modal filtering parameters were taken to be $(s, N_c) = (4, 4)$ in the $N = 4$ case and $(s, N_c) = (8, 6)$ for $N = 8$. Here, s is the order of the exponential filter function and N_c is the cut-off order, below which no filtering takes place. In particular, these parameters imply that for the $N = 4$ simulations, 9 of 25 modes are unaffected by the filter, while 18 of 49 and 30 of 81 modes are unaffected in the $N = 6$ and $N = 8$ cases, respectively. The initial solution fields (ρ, u, w) were interpolated from the equispaced Fourier grid to our DG grids via a band-limited spectral interpolation code[1]. All runs were time-stepped toward a target final time of $t = 600$ s, when the wave has travelled a distance of 62.52 m (or ≈ 12.5 domain lengths).

The results shown in Fig. 5.4**(a)-(b)** demonstrate that the method accurately captures the propagation of the ISW. The long-term stability properties of the NLP methods are assessed in panel **(c)** where we plot the L^2-norm of the divergence vs. time in each case. For the WLP methods, the divergence was $O(10^{-12})$ or less for all times, and the divergence could not be used as a stability indicator. The $N = 4$ NLP case becomes unstable by $t = 160$ s, and the $N = 4$ WLP case becomes unstable by $t = 240$ s. This result indicates that the local projection has improved the stability properties of the $N = 4$ scheme somewhat, but it is not sufficient for long-term stability in this case due to grid-scale noise introduced as a result of under-resolution. Contrary to the $N = 4$ result, the $N = 6$ WLP case becomes unstable by $t = 368$ s while the $N = 6$ NLP case does not and remains stable until $t = 600$ s. However, the L^2-norm of the divergence appears to grow for later times and it is unclear if this will lead to an instability. Both the WLP and NLP $N = 8$ cases appear

[1] The code `bandLimFourierInterp` is freely available at: https://github.com/dsteinmo/bandLimFourierInterp/.

to be long-term stable, since the L^2-norm of divergence in the NLP case demonstrates a damped limit-cycle type behaviour in Fig. 5.4, and the WLP method did not undergo a numerical instability or appear overly polluted by grid-scale noise.

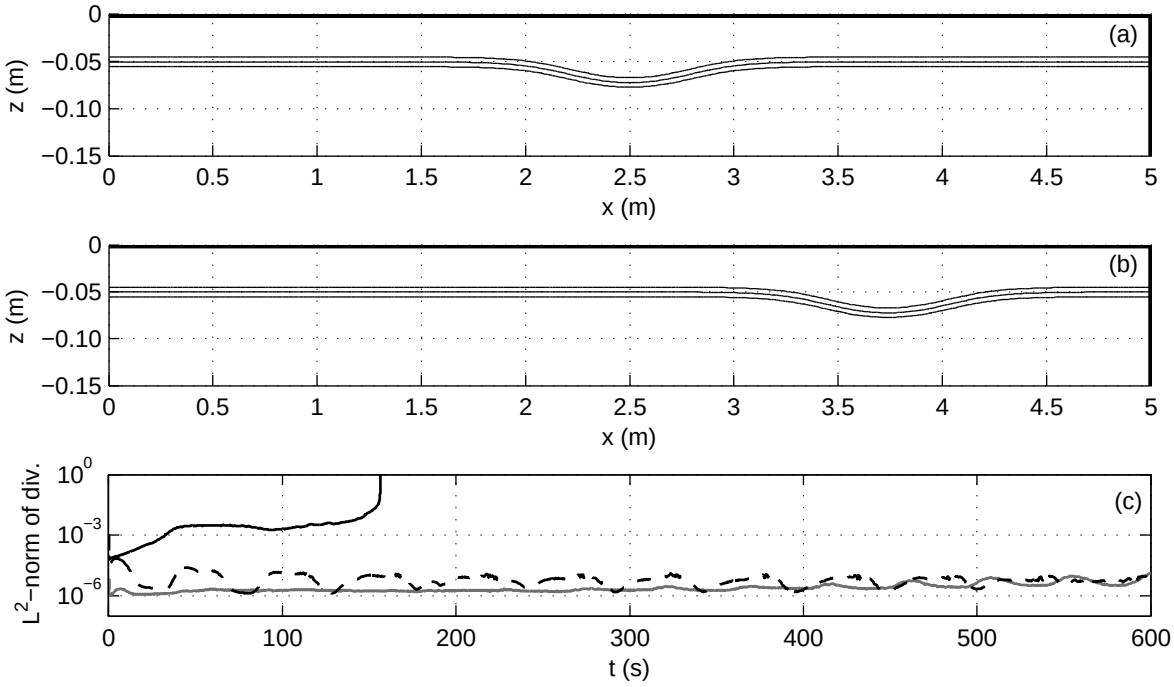

Figure 5.4: The position of the thin pycnocline is indicated by three contour lines of non-dimensional density ρ for the DJL ISW test case at times **(a)** $t = 0$ s and **(b)** $t = 540$ s using the $N = 8$ WLP method. Panel **(c)** shows the L^2-norm of divergence vs. time for the NLP methods: $N = 4$ (black, solid), $N = 6$ (dark, grey), $N = 8$ (grey, solid), and $N = 8$ NLP (black, dashed).

In Figure 5.5, we plot the vertical velocity w at $t = 112$ s for both the NLP and WLP methods for polynomial order $N = 4$. We focus on the w field since it was typically found to be the first field to be visibly affected by grid-scale noise. In this case, the local divergence-free projection suppresses some of the noise found in the case where the local projection has not been used. As mentioned above, this has allowed for a 50% increase in the length of the time-integration. However, the local post-processing projection does not appear to guarantee long-term stability in this case. A sharp contrast to this result is shown for the order $N = 6$ in Figure 5.6 where it is shown that the result from the WLP method is corroded by grid-scale noise. This is a puzzling result, since it appears that the

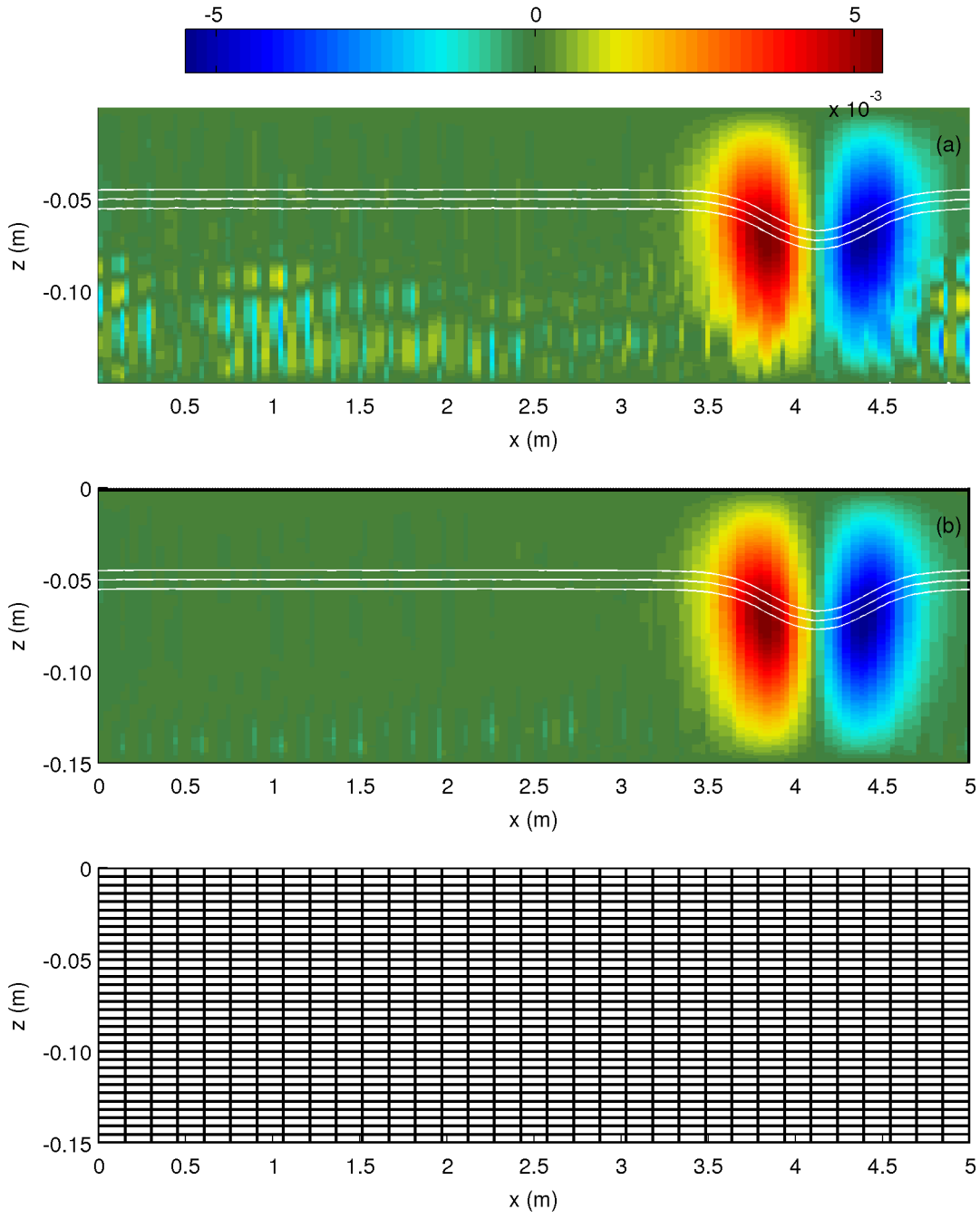

Figure 5.5: Contour plots of the vertical velocity (w) at $t = 112$ s for polynomial order $N = 4$ for **(a)** NLP method and **(b)** WLP method. Three density contours of non-dimensional density ρ have been super-imposed to indicated the location of the pycnocline. The bottom panel shows the finite element mesh used in all cases with $K = 1089$ elements.

local projection has introduced noise in the $N = 6$ case, while it has suppressed it in the $N = 4$ case.

One possible explanation for the above result is that the local post-processing projection introduces error which disregards the weak continuity properties imposed at elemental interfaces by the SIP-DG penalty method. The effect of this error on the corresponding pressure and velocity fields is unclear, and it depends on the spectral properties of the SIP-DG discretization of the Laplacian. It is possible that poorly-behaved spurious pressure modes are excited for $N = 6$, but not for $N = 4$, for reasons that are neither obvious nor intuitive.

	WLP	NLP
N	t_{blowup}	t_{blowup}
4	240 s	160 s
6	368 s	-
8	-	-

Table 5.1: Summary of stability results from the various runs for the DJL ISW test case, where t_{blowup} indicates the time of numerical instability (if any). Empty entries indicate the method was stable until the target time of $t = 600$ s.

In Figure 5.7, we plot the w field from the $N = 8$ WLP run at the same time from Figure 5.4 **(b)** where ρ was plotted. Panel **(a)** reveals a rather clean result, but a more selective choice of contour interval illustrates that grid-scale noise is indeed present in the solution. As discussed above, it was found that the local projection did not destabilize the $N = 8$ case as it did in the $N = 6$ case, however, an analogous plot to Figure 5.7 for the NLP case shown in Figure 5.8 reveals that the WLP method introduces slightly more noise near the bottom boundary.

5.6 Simulations of the two-layer rigid-lid equations

In addition to simulations of the stratified incompressible Euler equations, we have also carried out simulations of the two-layer weakly non-hydrostatic rigid-lid model (5.1)-(5.3) in the same geometries as in Chapter 4 using triangular meshes. In general, it was found that the rigid-lid equations were more sensitive to under-resolution and poor mesh quality than the single-layer weakly non-hydrostatic model. This behaviour is perhaps not surprising given the results discussed in the above section for the IE equations. In particular, results

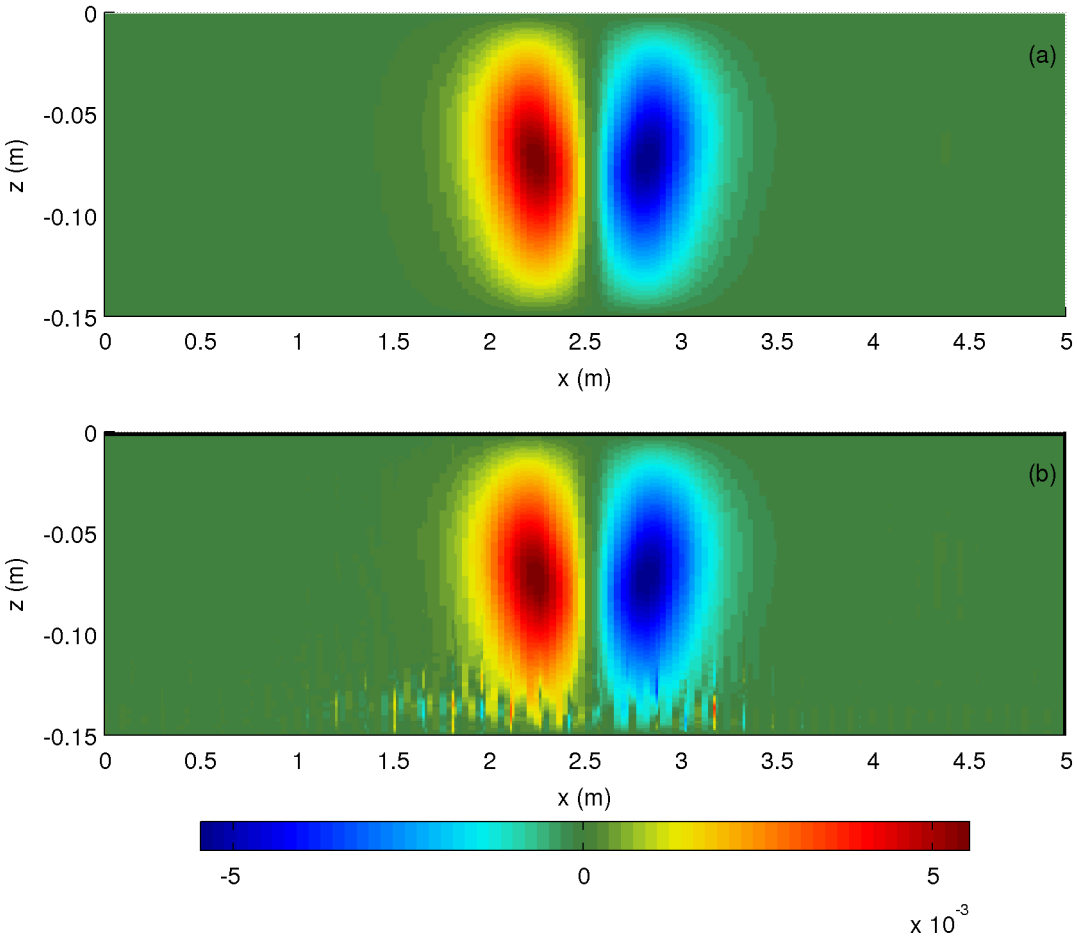

Figure 5.6: Contours plots of the vertical velocity (w) at $t = 337$ s for polynomial order $N = 6$ for **(a)** NLP method and **(b)** WLP method.

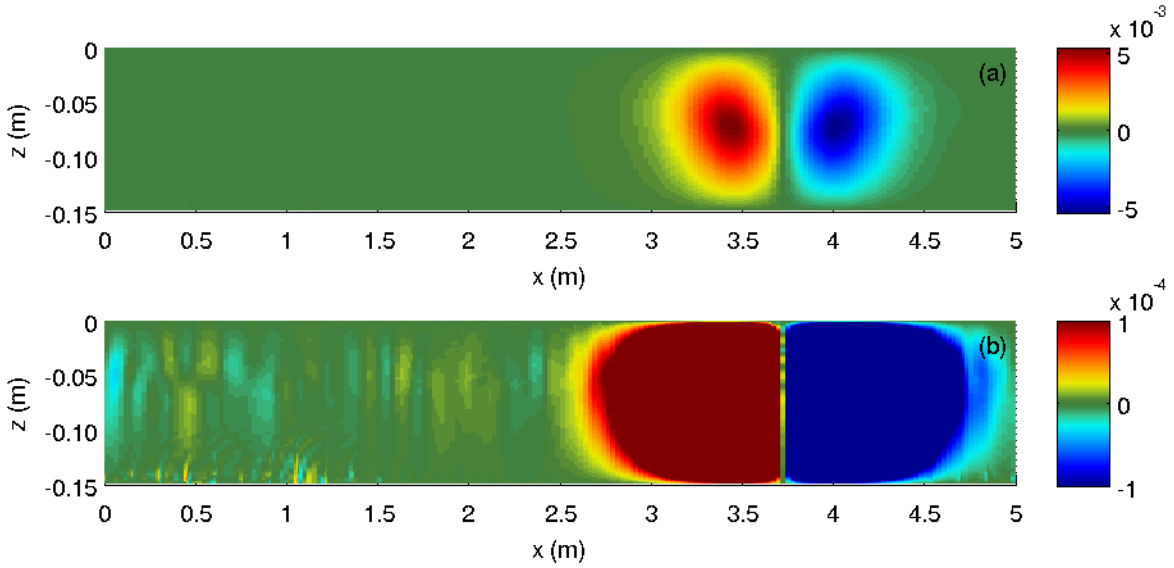

Figure 5.7: **(a)** Contour plots of the vertical velocity (w) at $t = 540$ s for the $N = 8$ WLP method, and **(b)** the same as **(a)** but the field has been saturated to a maximum absolute value of 10^{-4}.

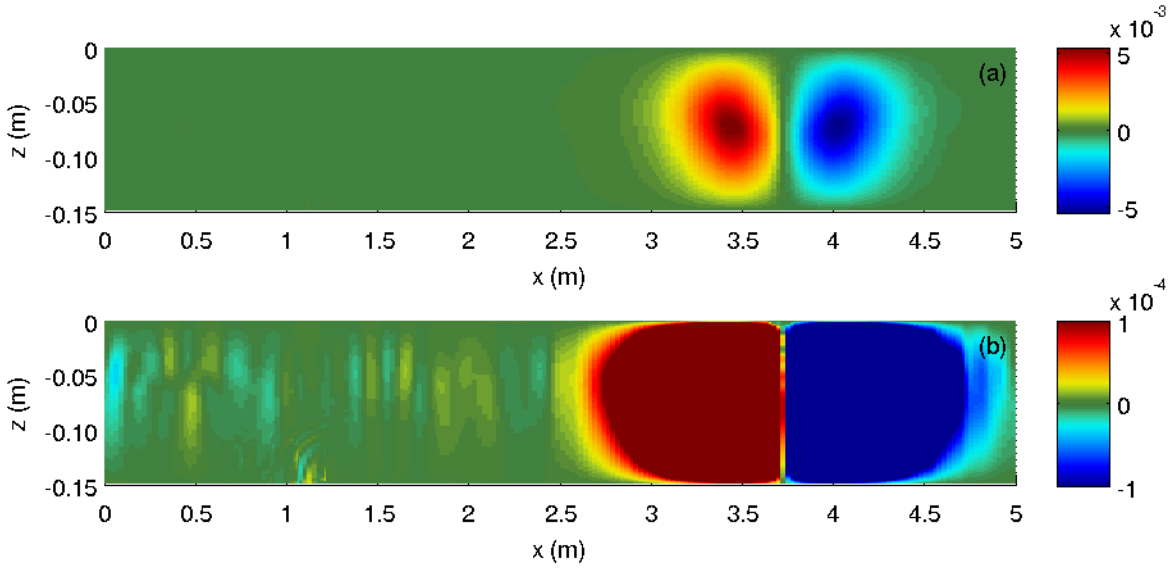

Figure 5.8: Like Figure 5.7, except with the $N = 8$ NLP method.

could not be obtained on the idealized mesh representing Pinehurst Lake that was discussed in Section 4.16 since grid-scale noise was introduced in the vicinity of re-entrant obstacles where element Jacobians had very large gradients. The inclusion of the variable bottom bathymetry of Pinehurst Lake drove additional noise that destroyed the solutions for early times. We believe the inclusion of source terms associated with variable depth is directly analogous to the inclusion of the source term for an active density tracer, which raised the demand on resolution for the incompressible Euler equations as discussed in Section 5.4.1 when the effects of density stratification were included.

As a lake model, two-layer model equations are inherently limited in a variety of key physical situations. For instance, it is quite common for the thermocline to intersect the bottom boundary in the later summer when the surface mixed layer deepens during nighttime convection (see for example [22] for measurements in Lake Simcoe, Canada). The area over which this occurs will be enlarged during seiching, leading to the emergence of turbulent boluses and other fluid motions too complex to be included in two layer models (again see the measurements in [22] for an example). As discussed in Chapter 2, two-layer equations exhibit instabilities in the presence of sufficiently strong vertical shear. Furthermore, the lower layer intersecting the bottom ($h_2 = 0$) is problematic since a positivity-preserving wetting and drying treatment would be necessary to prevent instabilities associated with a negative layer thickness. Such a wetting and drying treatment has been successfully employed by, e.g., Bunya *et al.* [16], for the $N = 1$ DG-FEM solution to the single-layer shallow water equations by retaining an infinitesimal amount of fluid in dry regions. However, even with such a treatment, the correct bottom boundary-layer physics would be misrepresented by such a layered model, and the method would be reduced to first order in areas where such a treatment was implemented. Taken together the two issues of interlayer shear and dry regions become intertwined, since the infinitesimal layer thickness approach to wetting and drying implies that while the transport $h\mathbf{u}$ is well behaved in dry regions, the velocity $\mathbf{u}$ is not guaranteed to be. Hence, in a two-layer setting, such a treatment would undoubtedly lead to regions of high shear that would drive numerical instabilities. In the present work, we accept these physical limitations associated with drying and/or high vertical shear, and focus our attention on wave propagation in parameter regimes where shear is sufficiently low and the layers are sufficiently thick.

The numerical method was implemented using the same techniques as for the IE equations discussion in Section 5.5. That is, time-stepping was carried out with AB3, modal filtering was applied to the full solution fields after the advective-step, and the local Lax–Friedrichs flux was used in the discretization of the advection terms. The pressure projection method was used to enforce incompressibility on the barotropic transport, and once the pressure had been determined its gradients were time-stepped within each layer. The

methodology for cuvilinear elements discussed in Section 4.14 was employed to suppress (at least in part) singular behaviour at re-entrant corners.

In Figure 5.9 results are shown for the evolution of an initial interfacial tilt in a two-layer fluid. The undisturbed layer thicknesses in the upper and lower layers taken to be $H_1 = 12.8$ m and $H_2 = 7.2$ m, respectively. The domain under consideration is the same as that from Section 4.12 except the 'island' in the center of the domain has been removed. It was found that retaining the island led to instabilities driven by under-resolution. The triangular mesh used here has $K = 1360$ elements. The east-west interfacial tilt was taken to increase linearly from $\zeta = -3$ m to $\zeta = 3$ m across the width of the basin. As in Sections 4.11 and 4.12, the Coriolis frequency is given by $f = 7.8828 \times 10^{-5}$ s^{-1} and the reduced gravity is $g' = 0.024525$ m s^{-2}. The simulation was carried out with order $N = 8$ and the filtering parameters were taken to be $(s, N_c) = (8, 4)$ such that 28 of 45 modes were unaffected by the filter.

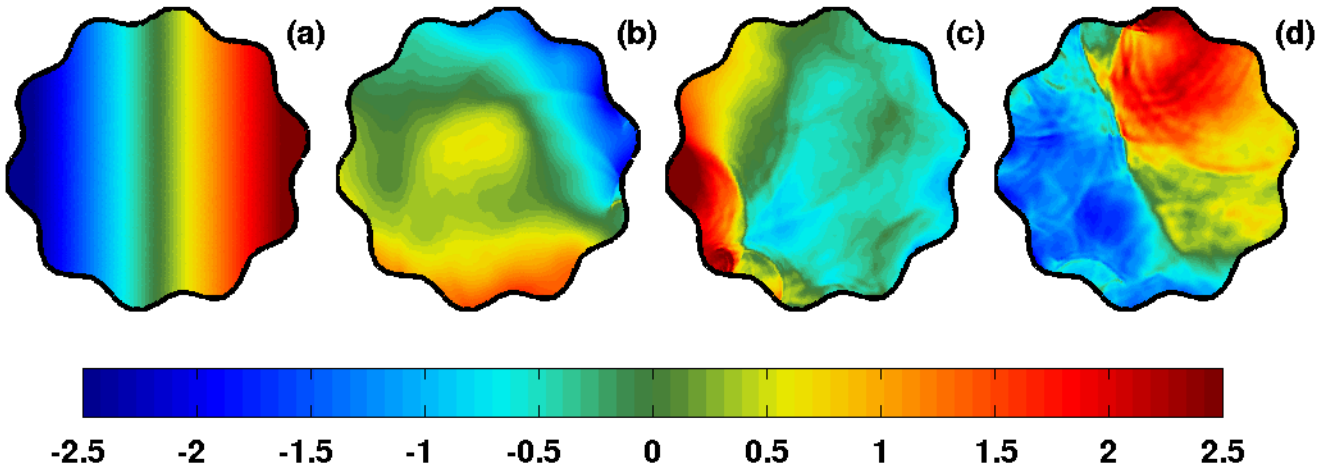

Figure 5.9: Evolution of an initial linear interfacial tilt in a lake with a wavy coastline using a two-layer rigid-lid model at times **(a)** $t = 0$ h, **(b)**, $t = 23.7$ h, **(c)** $t = 48.3$ h, **(d)** $t = 72.0$ h.

As discussed in Section 4.12, the undulating coastline enhances the near-shore focusing and nonlinear steepening of wave fronts. The emergence of some localized small-scale nonlinear waves is visible in Figure 5.9 **(b)** at time $t = 23.7$ h, and the wave field becomes highly dominated by small scale nonlinear and non-hydrostatic internal waves for the later times shown.

In Figure 5.10, we plot the x-component of the barotropic velocity $u_{bt} = (h_1 u_1 + h_2 u_2)/h$

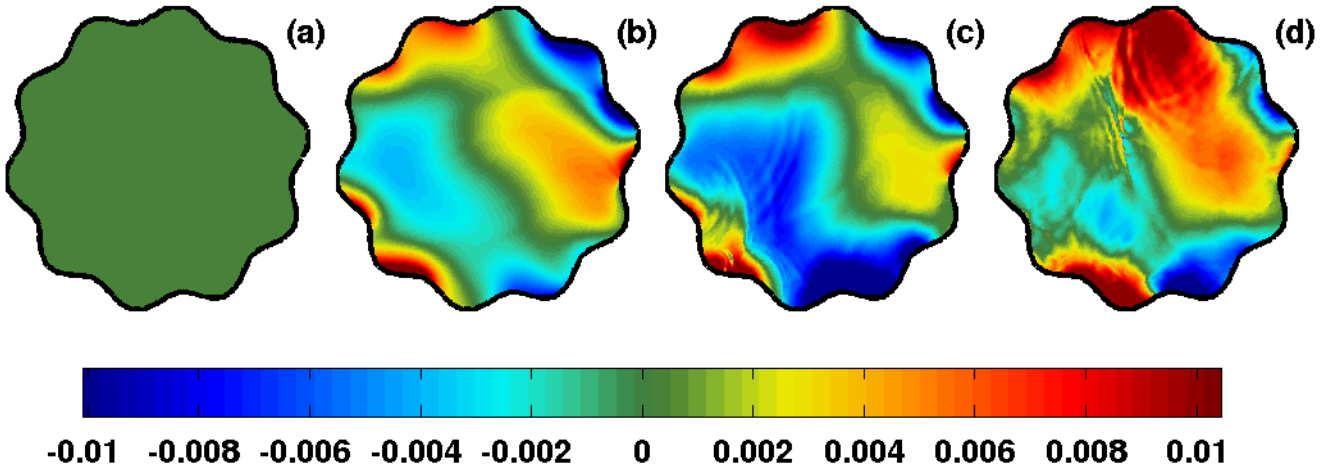

Figure 5.10: Like Figure 5.9, but the x-component of the barotropic velocity, u_{bt} is plotted.

at the same times as in Figure 5.9. Since the barotropic transport $h\mathbf{u}_{bt}$ must be divergence-free, we expect that any grid-scale noise due to the pressure projection method should appear in the components of the barotropic velocity. For the times shown, u_{bt} does not appear to be adversely affected by grid-scale noise since all of the small-scale features correspond well to small-scale waves in the interfacial displacement field.

In the final simulation of the chapter, we follow the evolution of a linear interfacial tilt on a circular lake with a ridge located at the southern end of the basin. Here, the tilt was taken to decrease from $\zeta = 6$ m to $\zeta = -6$ m. The depth profile is given by

$$H(x,y) = H_0 - \Delta H \operatorname{sech}^2\left(\frac{x}{0.25 r_0}\right) \operatorname{sech}^2\left(\frac{y + r_0}{0.65 r_0}\right) , \tag{5.26}$$

where $r_0 = 8345$ m is the radius of the basin, $H_0 = 25.6$ m, and $\Delta H = 4$ m is the height of the ridge. The upper layer thickness is taken as $H_1 = 14.5$ m so that all the depth variations would be felt in the lower layer $H_2(x,y) = H(x,y) - H_1$. The Coriolis parameter and the reduced gravity were taken to be the same as in the previous case. Once again the $N = 8$ DG-FEM was used with the $(s, N_c) = (8, 4)$ filter. The finite element mesh in this case has $K = 1400$ elements.

In Figure 5.11, we focus on a dynamical event that occurs as a result of the ridge. At $t = 24.0$ h (panel **(b)**), we see that the basin-scale seiche has steepened as in previous simulations. Here, we see that the seiche-induced flow over the ridge has caused a sharp interfacial depression to form just upstream of the peak of the ridge. Shortly afterwards

at $t = 28.2$ h (panel **(c)**), the main elevated wavefront pushes the low downstream and the two fronts interact. By $t = 32.4$ h, we see that the initially trapped low has been completely displaced as the main wave front has propagated over the ridge. At later times, an undular bore forms at the tail of the main wave front. This phenomenon is displayed in Figure 5.12, where the bore is shown propagating over the ridge. At $t = 63.6$ h, (panel **(a)**), we see that the main wave front is passing over the ridge and the undular bore lies behind the ridge. At $t = 67.8$ h (panel **(b)**), the bore has propagated over the ridge peak, and shoaling has caused its energy to focus into the leading crest. Panel **(c)** shows the undular bore at $t = 70.8$ h after it is completely downstream of the ridge.

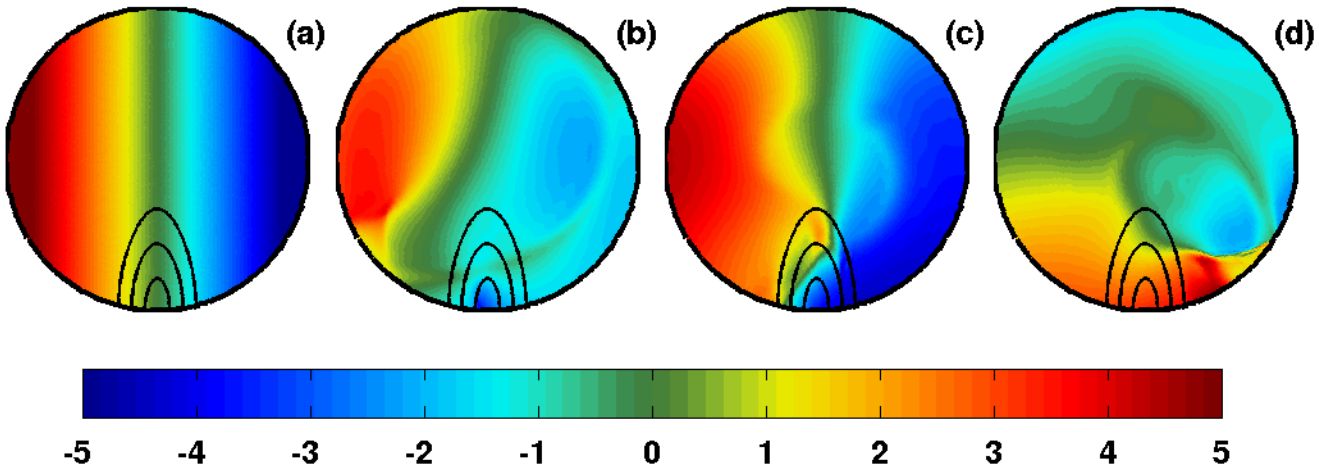

Figure 5.11: Evolution of an interfacial tilt in a two-layer circular basin with a ridge at times **(a)** $t = 0$, **(b)** $t = 24.0$ h, **(c)** $t = 28.2$ h, and **(d)** $t = 32.4$ h. The position of the ridge is indicated by three black depth contours at $H = 22$ m, 23 m, and 24 m.

5.7 Discussion and conclusions

Through a numerical eigenvalue analysis, we have shown that the eigenfunctions of the DG PP operator are themselves not divergence-free. As a result, our numerical simulations have shown that instabilities can occur due to the presence of spurious compressibility artifacts. The prominence of these instabilities appears to be worse in low-order or poorly resolved simulations. It was thought that enforcing incompressibility exactly after each time-step by a local post-processing projection could help improve the stability properties of the method. While some promise was shown for the $N = 4$ method, the local projection did

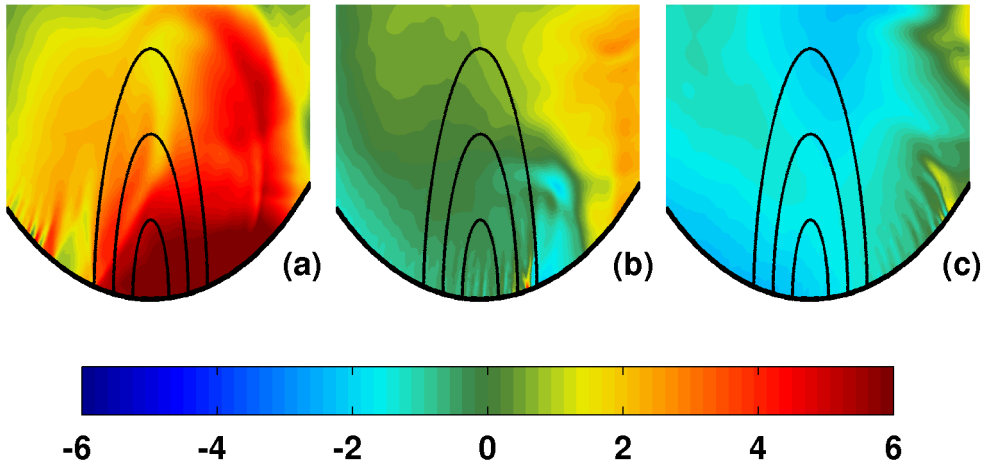

Figure 5.12: The same case shown in Figure 5.11, but at the later times **(a)** $t = 63.6$ h, **(b)** $t = 67.8$ h, and **(c)** $t = 70.8$ h. All panels have been zoomed-in on the ridge.

not necessarily improve the stability of the higher-order, $N = 6$ and 8, schemes. Through a series of simulations and discussions, we have reached the following conjectures about equal-order DG-FEM simulations of inviscid incompressible flow:

1. Solutions are very adversely affected by under-resolution and/or poor mesh quality.
2. The inclusion of source terms further increases the demand on resolution since numerically-driven perturbations can give rise to spurious unbalanced motions.
3. Long time integrations can be achieved through a combination of using higher-order polynomials together with appropriate local modal filtering.

The source of the difficulties illustrated in this chapter are almost certainly due to the LBB/inf-sup stability condition that usually implies that equal-order element methods for incompressible flow are unstable [24]. However, more work needs to be done to confirm that this is indeed the case. The signature of the LBB problem is grid-scale noise that is introduced due to spurious modes in the discretization of the Laplacian. While the SIP-DG formulation guarantees that unphysical modes are suppressed to the small-scale portion of the eigenspectrum of the Laplacian, it does not remove them. Hence, the LBB problem remains an active one. The spurious compressibility artifacts that we have focused on in detail are likely a consequence of this short-coming of the pressure discretization. Removing

these artifacts does not guarantee that the spurious pressure modes will not be excited and pollute the solution, in general.

In a recent thesis, Ferrer [45] applied an equal order DG-FEM method to solve the INS equations with the standard PP method in a variety of engineering configurations. In that work, it was suggested that using polynomial orders of $N \geq 5$ would improve the LBB properties of the scheme. Additionally, it was reported that sufficiently strong viscosity (either by small Re or large Δt) would, not surprisingly, lead to enhanced stability. Since our simulations in Section 5.5 do not contain a viscosity operator, the reported stability for the $N = 6$ and $N = 8$ methods is owed entirely to local polynomial filtering applied to the full fields that takes over the stabilizing role of viscosity in these cases.

The fact that local filtering can have such a postitive effect on equal-order DG-FEM inviscid incompressible flow simulations raises a series of interesting follow-up questions for ongoing work such as:

1. Could results be improved by an adaptive filtering strategy that minimizes the amount of numerical dissipation?

2. Could an element-overlapping filtering strategy be devised to improve the behaviour of the schemes at interfaces?

3. Applying the same filtering methodologies above, what is the maximum allowable Reynolds number Re that would allow for a stable viscous simulation with minimal noise?

Regarding point 2, it is worth noting that Diamessis *et al.* [30] have used a technique referred to as 'adaptive interfacial averaging' to suppress solution spikes at element interfaces in simulations of stratified incompressible turbulence using the spectral multi-domain penalty method. It would be interesting to see if the technique would be beneficial in the DG-FEM context considered here.

It may be possible to further remedy the stability issues highlighted here by considering a methodology outside of the standard PP framework. Remedies discussed in other works include considering the artificial compressibility equations to formulate a suitable numerical flux function [8, 10], thereby ensuring stronger coupling between pressure and velocity at element interfaces. It may also be possible to pursue an HDG spatial discretization method [89, 102] that enforces continuity only in the normal component of numerical fluxes, while allowing for jumps in the tangential direction at interfaces. Unequal order approximations for velocity and pressure may be a possibility as well. However, Shahbazi *et al.* [105] showed

poorer stability characteristics than their equal-order method when taking the order of their pressure approximation to be one less than the order of the velocity approximation. Thus, the utility of mixed-order DG formulations for incompressible flow problems does not appear to be well-understood in the literature.

In light of the results presented in this chapter the utility of the DG-FEM for more complicated incompressible flow situations (e.g., three-dimensional ocean and lake models) has become questionable. In particular, the difficulties encountered due to under-resolution are particularly troublesome since real-world flow situations would involve such features in the form of coastal boundary layers, bottom boundary layers, wetting and drying regions (particularly where pycnoclines intersect the boundary), small-scale turbulence, *etc.* On the other hand, since the results in Chapter 4 have shown that the DG-FEM method can quite comfortably handle hyperbolic (or nearly hyperbolic) model equations, three dimensional modelling may be within reach by considering a set of equations that include compressibility effects. However, since the inclusion of compressibility results in the presence of sound waves with short time-scales, an implicit-explicit (IMEX) type scheme should be employed so that computational time is not wasted due to the very restrictive time-step imposed by the sound waves. Such schemes have been explored in atmospheric models with the DG-FEM in the non-hydrostatic unified model of the atmosphere (NUMA) by Restelli and Giraldo [101]. Incompressible flow simulations can thus be explored by considering the artificial compressibility equations with a suitably chosen artificial sound speed with the pressure waves discretized implicitly. The construction of sub grid scale parametrizations that accurately model physics while maintaining the hyperbolic mathematical structure remain an open problem.

Chapter 6

Calculating Free Modes of Oscillation in Complex Geometries

In this chapter, the calculation of free modes of oscillation in the linear rotating shallow water framework is considered. The DG-FEM method used in Chapters 4 and 5 is employed so that its geometric flexibility can be exploited to explore closed basins of arbitrary shape. In the following sections, the numerical methodology is introduced and explained. The method is then validated against analytical solutions and applied to basins with both simple and irregular geometries.

6.1 History and motivation

The calculation of free modes of oscillation in closed rotating basins represents a thoroughly-studied problem with previous works dating back well over a century [71, 119]. The importance of free modes of oscillation in understanding basin-scale dynamics in large lakes is thus well-known, and a method that can calculate these modes on numerical grids that accurately reflect the actual geometry of real-world lakes is very desirable since these numerically calculated modes allow hypotheses regarding the basin-scale dynamics of lakes to be formulated and tested against field data.

Analytical formulae for the free modes of oscillations in infinite non-rotating and rotating channels were among the first successfully calculated. It was also attempted to use so-called one-dimensional channel theory to estimate the longitudinal oscillations of Lake Erie [96] and the other great lakes [103]. Analytical expressions were also calculated for

idealized rotating circular [71, 25, 115] and elliptical basins [3]. Calculations of the internal free modes of oscillation have also been explored in idealized basins by means of a normal modes decomposition in [25, 3].

Numerical calculations were then performed by [98, 100] using finite difference methods. The method of Rao and Schwab [100] uses an analytical reduction by using a series decomposition in terms of divergence-free and irrotational basis functions. The main difficulty with the method of Rao and Schwab lay with ensuring the eigenspectrum of basis functions is orthogonal. While their method is able to compute modes in general (arbitrarily-shaped) basins, the method can only represent the coastline in a piece-wise constant manner. Hamblin [57] computed the free modes of oscillation by re-casting the equations as a variational optimization problem that was solved using the finite element method, allowing for more flexibility in accurately representing coastlines in a piece-wise linear and curvilinear manner. Due to the limited availability of computing power in the 1970–1980 time-frame, it is perhaps not surprising that these calculations were done with very coarse grids by today's standards, and the problem is worthy of being revisited with modern computers and discretization methods.

In this work, we employ the analytical reduction methodology of [100] in closed basins represented by triangular grids using the symmetric interior penalty discontinuous Galerkin finite element method (hereafter, SIP-DG). The key ideas of the methodology thus remains the same as in [100], and only the numerical details differ. The method represents an improvement over the method of [100] since it allows the coastline to be represented in a piece-wise linear manner. Furthermore, the SIP-DG method allows more accurate solutions to be obtained quite simply by increasing the order of the local basis polynomials used on each triangle without affecting the size of the computational stencil. The methodology of [100] also requires special care in constructing the numerical grid, since a staggered grid is required to ensure the spatial discretization operators are symmetric to guarantee the orthogonality of the basis of eigenfunctions. With the SIP-DG method, the spatial discretization operator is, by construction, guaranteed to be symmetric regardless of where the grid points are collocated. This additional flexibility requires little work to be done by the individual performing the calculation of the modes aside from generating the triangular mesh.

An alternative to the SIP-DG method that one may consider to obtain solutions on unstructured triangular meshes is the Continuous Galerkin (CG) finite element method (FEM) or its high-order extension, the spectral element method (SEM) [68]. However, it is known that the CG-based methods can result in spurious (numerical) modes that can pollute the physical part of the eigenspectrum [60]. In some cases, such spurious modes can be convergent [60] as one refines the computational grid and may lead to the

rather undesirable result where they are mistaken for physical modes. Additionally, FEM discretizations may also contain so-called "pesky modes" that converge to being members of the operator null space as the grid is refined [24]), leading to difficulties for iterative solvers. The SIP-DG method circumvents this issue by suppressing any spurious modes to unphysically high frequencies, ensuring that they will not be present in any reasonable truncation of the eigenspectrum. This suppression of spurious modes is furnished by the use of a penalty parameter, that guarantees that the spurious modes do not pollute the physical part of the spectrum provided that it is taken to be sufficiently large.

In the following, we present our numerical methodology for calculating the free modes of oscillation in rotating basins of arbitrary geometry followed by validation cases, a case study of an idealized basin, and applications to real-world lakes.

6.2 Methods

The linearized shallow water equations on the rotating f-plane can be written as

$$\frac{\partial \mathbf{M}}{\partial t} + f\mathbf{M}^{\perp} = -gH\nabla\eta \tag{6.1}$$

$$\frac{\partial \eta}{\partial t} + \nabla \cdot \mathbf{M} = 0 \tag{6.2}$$

where $H(\mathbf{x})$ is allowed to vary in space in terms of the Cartesian position vector $\mathbf{x} = (x, y)$, $\eta(\mathbf{x}, t)$ is the free surface elevation above the undisturbed state, $\mathbf{M} = H\mathbf{u}$ is the volume transport vector, $\mathbf{u} = (u(\mathbf{x}, t), v(\mathbf{x}, t))$ is the fluid velocity field, and superscript $\perp$ denotes rotation by 90° counter-clockwise, so that $\mathbf{M}^{\perp} = H(-v, u)$. The parameters g and f represent the acceleration due to gravity and the Coriolis parameter, respectively. The fluid depth is non-dimensionalized via $H(\mathbf{x}) = \overline{H}h(\mathbf{x})$, where $\overline{H}$ is the mean basin depth and $h(\mathbf{x})$ is dimensionless.

As explained in [100], combining equations (6.1)–(6.2) to find an equation for η yields an operator that is not self-adjoint. Further difficulty results from the fact that the no normal flow boundary condition in terms of η contains real and imagninary components and is coupled to the eigenvalue (the frequency). To circumvent these difficulties, a Helmholtz decomposition is applied to the volume transport vector in order to express it in terms of a velocity potential and a streamfunction, i.e.,

$$\mathbf{M} = -h\nabla\phi + \nabla^{\perp}\psi \ , \tag{6.3}$$

where $\nabla^\perp = \left(-\frac{\partial}{\partial y}, \frac{\partial}{\partial x}\right)$, and ϕ and ψ satisfy the no normal flow boundary condition

$$\frac{\partial \phi}{\partial n} = \psi = 0 \quad \text{on} \quad \partial\Omega \, . \tag{6.4}$$

where $\frac{\partial \phi}{\partial n} = \nabla\phi \cdot \hat{\mathbf{n}}$ represents the derivative taken in the direction normal to $\partial\Omega$, the boundary of a closed basin (domain), Ω. The no-normal flow condition on ψ actually requires ψ to be constant along the boundary. However, since only the derivatives of ψ are of physical relevance, the constant is set to zero without loss of generality.

The decomposition (6.3) is unique. That is, assuming $\mathbf{M}$ is known, ϕ is the unique solution of

$$\begin{aligned} \nabla \cdot (h\nabla\phi) + \nabla \cdot \mathbf{M} &= 0 \, , \quad \text{on} \quad \Omega \, , \\ \frac{\partial \phi}{\partial n} &= 0 \, , \quad \text{on} \quad \partial\Omega \, , \end{aligned} \tag{6.5}$$

as can be seen by taking the divergence of (6.3). Similarly, ψ is the unique solution of

$$\begin{aligned} \nabla \times \left(\frac{1}{h}\nabla^\perp \psi\right) - \nabla \times \left(\frac{1}{h}\mathbf{M}\right) &= 0 \, , \quad \text{on} \quad \Omega \, , \\ \psi &= 0 \, , \quad \text{on} \quad \partial\Omega \, . \end{aligned} \tag{6.6}$$

The next step is to form two sets of basis functions based on two eigenvalue problems. We define the set $\{\phi_\alpha, \lambda_\alpha\}_{\alpha=1}^{N_\phi}$ given by the eigenproblem

$$\nabla \cdot (h\nabla\phi_\alpha) + \lambda_\alpha \phi_\alpha = 0 \, , \quad \text{on} \quad \Omega \, , \tag{6.7}$$

$$\nabla\phi_\alpha \cdot \hat{\mathbf{n}} = 0 \, , \quad \text{on} \quad \partial\Omega \, , \tag{6.8}$$

and the set $\{\psi_\alpha, \mu_\alpha\}_{\alpha=1}^{N_\psi}$ given by

$$\nabla \cdot \left(\frac{1}{h}\nabla\psi_\alpha\right) + \mu_\alpha \psi_\alpha = 0 \, , \quad \text{on} \quad \Omega \, , \tag{6.9}$$

$$\psi_\alpha = 0 \, , \quad \text{on} \quad \partial\Omega \, . \tag{6.10}$$

Here, N_ϕ and N_ψ are taken as finite integers to truncate the bases for computational reasons. It can be shown using dimensional analysis that a suitable normalization for the basis functions is

$$\iint_\Omega h\nabla\phi_k \cdot \nabla\phi_l dA = \lambda_l \iint_\Omega \phi_k \phi_l dA = Ac^2 \overline{H}^2 \delta_{k,l} \, , \tag{6.11}$$

$$\iint_\Omega \frac{1}{h}\nabla^\perp\psi_k \cdot \nabla^\perp\psi_l dA = \mu_l \iint_\Omega \psi_k \psi_l dA = Ac^2 \overline{H}^2 \delta_{k,l} \, . \tag{6.12}$$

Next, ϕ and ψ are expanded in terms of the basis functions via

$$\phi = \sum_\alpha P_\alpha(t)\phi_\alpha \,, \tag{6.13}$$

$$\psi = \sum_\alpha Q_\alpha(t)\psi_\alpha \,. \tag{6.14}$$

Now, since

$$\eta_t = -\nabla\cdot\mathbf{M} = \nabla\cdot(h\nabla\phi) = -\sum_\alpha P_\alpha\lambda_\alpha\phi_\alpha \,, \tag{6.15}$$

it is clear that η should be expanded in terms of the ϕ_α's. Rao and Schwab [100] set

$$\eta = \sum_\alpha R_\alpha(t)\eta_\alpha = \sum_\alpha R_\alpha(t)\frac{\lambda_\alpha^{\frac{1}{2}}}{c}\phi_\alpha \,, \tag{6.16}$$

such that the R_α's are dimensionless, just as the P_α's and Q_α's are, by construction.

For the momentum equation (6.1), we can use

$$\mathbf{M} = -h\nabla\phi + \nabla^\perp\psi \,, \tag{6.17}$$

$$= -h\sum_\alpha P_\alpha(t)\nabla\phi_\alpha + \sum Q_\alpha(t)\nabla^\perp\psi_\alpha \,, \tag{6.18}$$

and

$$\mathbf{M}^\perp = (-h\nabla\phi + \nabla^\perp\psi)^\perp = -h\nabla^\perp\phi - \nabla\psi \,, \tag{6.19}$$

to find

$$\begin{aligned} -\sum_\alpha \frac{dP_\alpha}{dt}h\nabla\phi_\alpha + \sum_\alpha \frac{dQ_\alpha}{dt}\nabla^\perp\psi_\alpha &- fh\sum_\alpha P_\alpha\nabla^\perp\phi_\alpha - f\sum_\alpha Q_\alpha\nabla\psi_\alpha \\ &= -g\overline{H}h\sum_\alpha R_\alpha\frac{\lambda_\alpha^{\frac{1}{2}}}{c}\nabla\phi_\alpha \,. \end{aligned} \tag{6.20}$$

Equation (6.20) can be simplified by taking the dot product with $\nabla\phi_\beta$, using the normalization (6.11), and the orthogonality relation

$$\iint_\Omega \nabla^\perp\psi_\alpha\cdot\nabla\phi_\beta\, dA = 0 \,, \tag{6.21}$$

that can be established using integration by parts. This simplification yields

$$\begin{aligned}\frac{dP_\beta}{dt} &+ \sum_\alpha P_\alpha \left(\frac{1}{Ac^2\overline{H}^2} \iint\limits_\Omega fh\nabla\phi_\beta \cdot \nabla\phi_\alpha^\perp \, dA \right) \\ &+ \sum_\alpha Q_\alpha \left(\frac{1}{Ac^2\overline{H}^2} \iint\limits_\Omega f\nabla\psi_\alpha \cdot \nabla\phi_\beta \, dA \right) = c\lambda_\beta^{\frac{1}{2}} R_\beta \ , \end{aligned} \tag{6.22}$$

which may be written in terms of matrices as

$$\frac{dP_\beta}{dt} - \sum_\alpha A_{\beta\alpha} P_\alpha - \sum_\alpha B_{\beta\alpha} Q\alpha - \nu_\beta R_\beta = 0 \ , (\text{no sum over } \beta) \ , \tag{6.23}$$

where

$$A_{\beta\alpha} = -\frac{1}{Ac^2\overline{H}^2} \iint\limits_\Omega fh\nabla\phi_\beta \cdot \nabla^\perp\phi_\alpha \, dA \ , \tag{6.24}$$

$$B_{\beta\alpha} = -\frac{1}{Ac^2\overline{H}^2} \iint\limits_\Omega f\nabla\psi_\alpha \cdot \nabla\phi_\beta \, dA \ , \tag{6.25}$$

$$\nu_\beta = c\lambda_\beta^{\frac{1}{2}} \ . \tag{6.26}$$

Similarly, $\frac{dP_\alpha}{dt}$ can be eliminated from (6.20) by taking its dot product with $\frac{1}{h}\nabla^\perp\psi_\beta$ along with the normalization (6.12) and the orthogonality relation

$$\iint\limits_\Omega \nabla^\perp\psi_\beta \cdot \nabla\phi_\alpha \, dA = 0 \ . \tag{6.27}$$

Doing so gives

$$Ac^2\overline{H}^2 \frac{dQ_\beta}{dt} - \sum_\alpha P_\alpha \iint\limits_\Omega \nabla^\perp\phi_\alpha \cdot \nabla^\perp\psi_\beta \, dA - \sum_\alpha Q_\alpha \iint\limits_\Omega f\nabla\psi_\alpha \cdot \frac{1}{h}\nabla^\perp\psi_\beta \, dA = 0 \ , \tag{6.28}$$

or in matrix form

$$\frac{dQ_\beta}{dt} - \sum_\alpha C_{\beta\alpha} P_\alpha - \sum_\alpha D_{\beta\alpha} Q_\alpha = 0 \ . \tag{6.29}$$

where

$$\begin{aligned} C_{\beta\alpha} &= \frac{1}{Ac^2\overline{H}^2}\iint\limits_{\Omega} f\nabla^{\perp}\phi_{\alpha}\cdot\nabla^{\perp}\psi_{\beta}\,dA\;, \\ &= \frac{1}{Ac^2\overline{H}^2}\iint\limits_{\Omega} f\nabla\phi_{\alpha}\cdot\nabla\psi_{\beta}\,dA\;, \\ &= -B_{\alpha\beta}\;, \end{aligned} \tag{6.30}$$

and

$$D_{\beta\alpha} = \frac{1}{Ac^2\overline{H}^2}\iint\limits_{\Omega} f\frac{1}{h}\nabla\psi_{\alpha}\cdot\nabla^{\perp}\psi_{\beta}\;. \tag{6.31}$$

The fact that C and B share this symmetry property will have important consequences momentarily, along with the fact that A and D are anti-symmetric, i.e.,

$$A_{\beta\alpha} = -A_{\alpha\beta}\;, \tag{6.32}$$

$$D_{\beta\alpha} = -D_{\alpha\beta}\;, \tag{6.33}$$

which can be demonstrated by first noticing that $\nabla\phi_{\beta}\cdot\nabla^{\perp}\phi_{\alpha} = \nabla\phi_{\alpha}\cdot\left(-\nabla^{\perp}\phi_{\beta}\right)$. The system of ordinary differential equations for $(P_{\alpha}, Q_{\alpha}, R_{\alpha})$ must be closed by substituting the expansion (6.16) into equation (6.2), resulting in

$$\sum_{\alpha}\frac{dR_{\alpha}}{dt}\frac{\lambda_{\alpha}^{\frac{1}{2}}}{c}\phi_{\alpha} + \sum_{\alpha}P_{\alpha}\nabla\cdot(-h\nabla\phi_{\alpha}) = 0\;. \tag{6.34}$$

This equation can be simplified by multiplying by ϕ_{β}, integrating over the domain, and exploiting orthogonality, via

$$\sum_{\alpha}\frac{dR_{\alpha}}{dt}\frac{\lambda_{\beta}^{\frac{1}{2}}}{c}\frac{Ac^2\overline{H}^2}{\lambda_{\beta}}\delta_{\alpha\beta} - \sum_{\alpha}P_{\alpha}\iint\limits_{\Omega}\phi_{\beta}\nabla\cdot(h\nabla\phi_{\alpha})\;dA = 0\;. \tag{6.35}$$

After simplifying, we find

$$\frac{dR_{\beta}}{dt} + \nu_{\beta}P_{\beta} = 0\;, \quad \text{(no sum)}\;. \tag{6.36}$$

Equations (6.23), (6.29), and (6.36) may be combined into the standard form of a system of ODEs

$$\frac{d}{dt}\mathbf{V} + E\mathbf{V} = 0\;, \tag{6.37}$$

where

$$\mathbf{V} = \begin{pmatrix} P \\ Q \\ R \end{pmatrix} , \quad \text{and} \quad E = \begin{pmatrix} -A & -B & -<\nu> \\ -C & -D & 0 \\ <\nu> & 0 & 0 \end{pmatrix} , \tag{6.38}$$

and

$$<\nu>= \begin{pmatrix} \nu_1 & 0 & \dots & 0 \\ 0 & \nu_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \nu_{N_\phi} \end{pmatrix} = \text{diag}\,(\nu_i) \ . \tag{6.39}$$

The eigenvalue problem for the horizontal modes of the system can be recovered by looking for time-periodic solutions, i.e., by assuming

$$\mathbf{V}(t) = \mathrm{e}^{i\sigma t}\hat{\mathbf{V}} \ , \tag{6.40}$$

where

$$\hat{\mathbf{V}} = \begin{pmatrix} \hat{P} \\ \hat{Q} \\ \hat{R} \end{pmatrix} \tag{6.41}$$

is constant in time. The substitution of this ansatz yields the eigenvalue problem

$$iE\hat{\mathbf{V}} = \sigma\hat{\mathbf{V}} \ . \tag{6.42}$$

Inspecting the structure of E and recalling that the matrices A and D are symmetric, $B = -C^{\mathsf{T}}$, it follows that E is anti-symmetric. Therefore, iE is Hermitian and the eigenvalues σ are real. Furthermore, since iE is purely imaginary and σ is purely real, it follows that all eigenvectors must have real and imaginary components in order to satisfy the eigenproblem (6.42). Physically, this property corresponds to the various basis functions being out of phase with one another.

6.2.1 Obtaining basis functions from the Symmetric Interior Penalty Discontinuous Galerkin (SIP-DG) method

The above section deals with the theory behind the analytical reduction methodology used to compute the modes. It assumes that two orthogonal sets of basis functions resulting from homogenous eigenvalue problems are known *a priori*. In basins of general shape, the basis functions must be constructed numerically, and we achieve this through the symmetric interior penalty disconuous Galerkin (SIP-DG) as we explain below.

Consider, without loss of generality, the eigenvalue problem for the unknown eigenpair (μ, ϕ),

$$\nabla \cdot (h \nabla \phi) = \mu \phi \; . \tag{6.43}$$

As explained by [60], in order to discretize an elliptic operator with the DG method it is necessary to re-write the problem as a first-order system of equations by introducing the auxiliary variable

$$\mathbf{q} = (q_x, q_y) = \sqrt{h} \nabla \phi \; . \tag{6.44}$$

After having done so, the first-order system reads

$$\nabla \cdot \left(\sqrt{h} \mathbf{q} \right) = \mu \phi \; , \tag{6.45}$$

$$q_x = \sqrt{h} \frac{\partial \phi}{\partial x} \; , \tag{6.46}$$

$$q_y = \sqrt{h} \frac{\partial \phi}{\partial y} \; . \tag{6.47}$$

It is then assumed the domain Ω is partitioned into K triangular sub-domains (elements). We apply the nodal DG methodology, and assume that each of the unknown approximate solution fields can be represented locally on a particular triangle $\mathbf{D}^k$ in terms of two-dimensional Lagrange interpolating polynomials, i.e.,

$$\phi^k(\mathbf{x}) = \sum_{i=1}^{N_p} \phi^k(\mathbf{x}_i^k) \ell_i^k(\mathbf{x}) \; , \tag{6.48}$$

and similar definitions exist for q_x^k and q_y^k. Here $\{\mathbf{x}_i^k\}_{i=1}^{N_p}$ is the set of N_p nodes used to represent the solution on a triangle $\mathbf{D}^k$ with N_p dependent on the order of approximation. The weak form of equation (6.45) is obtained by substituting the approximate local solution functions, multiplying by a member of the space of local test functions $V^k = \{\ell_j^k\}_{j=1}^{N_p}$ and integrating by parts, yielding

$$\int_{\partial \mathbf{D}^k} \ell_j^k (\sqrt{h}\mathbf{q})^* \cdot \hat{\mathbf{n}}^k \, d\mathbf{x} - \int_{\mathbf{D}^k} \left(\sqrt{h}^k \mathbf{q}^k \right) \cdot \nabla \ell_j^k \, d\mathbf{x} = \mu \int_{\mathbf{D}^k} \ell_j^k \phi^k \, d\mathbf{x} \; , \tag{6.49}$$

where $(\sqrt{h}\mathbf{q})^*$ is a numerical flux function used to weakly impose continuity across element interfaces and $\hat{\mathbf{n}} = (n_x, n_y)$ is the unit outward normal to an element edge. For computational convenience, we calculate the strong form of equations (6.46)–(6.47) where

integration by parts is performed twice instead of once. This gives

$$\int_{\mathbf{D}^k} \ell_j^k q_x^k d\mathbf{x} = \int_{\mathbf{D}^k} \ell_j^k \sqrt{h}^k \frac{\partial \phi^k}{\partial x}\, d\mathbf{x} - \int_{\partial \mathbf{D}^k} \ell_j^k \left((\sqrt{h}\phi)^k - (\sqrt{h}\phi)^* \right) n_x^k \, d\mathbf{x} \,, \tag{6.50}$$

$$\int_{\mathbf{D}^k} \ell_j^k q_y^k d\mathbf{x} = \int_{\mathbf{D}^k} \ell_j^k \sqrt{h}^k \frac{\partial \phi^k}{\partial y}\, d\mathbf{x} - \int_{\partial \mathbf{D}^k} \ell_j^k \left((\sqrt{h}\phi)^k - (\sqrt{h}\phi)^* \right) n_y^k \, d\mathbf{x} \,. \tag{6.51}$$

In the symmetric interior penalty (SIP-DG) method, the numerical fluxes are taken to be

$$(\sqrt{h}\mathbf{q})^* = \{\!\{\sqrt{h}\nabla\phi\}\!\} - \tau [\![\sqrt{h}\phi]\!] \,, \tag{6.52}$$

$$(\sqrt{h}\phi)^* = \{\!\{\sqrt{h}\phi\}\!\} \,. \tag{6.53}$$

Here,

$$\{\!\{\mathbf{u}\}\!\} = \frac{\mathbf{u}^- + \mathbf{u}^+}{2}, \qquad [\![\mathbf{u}]\!] = \hat{\mathbf{n}}^- \cdot \mathbf{u}^- + \hat{\mathbf{n}}^+ \cdot \mathbf{u}^+ \,, \tag{6.54}$$

represent the average and jump in some quantity, $\mathbf{u}$, across the interface, with superscripts $-$ and $+$ referring to information interior to and exterior to the element, respectively.

Following the discussion in Section 4.6, the integral formulations are written in terms of local matrix operators and the auxiliary variable $\mathbf{q}$ is elminated locally so that a global stiffness matrix operator can be set up directly in sparse matrix format. Similarly, a global mass matrix operator may be constructed trivially since it is block diagonal with each block given by $\mathcal{M}^k$. In terms of global matrix operators, we see that the discretized Poisson eigenvalue problem takes the form of a generalized matrix eigenvalue problem,

$$K\boldsymbol{\phi} = \mu M \boldsymbol{\phi} \,, \tag{6.55}$$

where K is the global stiffness matrix and M is the global mass matrix. Both are guaranteed to be symmetric, and $\boldsymbol{\phi} = \bigoplus_{k=1}^{K} \boldsymbol{\phi}^k$ is the global solution vector.

6.2.2 Constructing the Modes

Computing the eigenvalues of the eigenvalue problem (6.42) yields the frequencies σ and eigenvectors for $2N_\phi$ gravitational modes and N_ψ rotational modes. Rotational modes exist purely due to changes in potential vorticity in the background state that, under

quasi-geostrophic theory [70], can only occur due to changes in the lake's depth H or variations in the Coriolis parameter f. In this work, we ignore variations in f by invoking the f-plane approximation, and we only find rotational modes when bathymetric variations are included.

The computed frequencies (eigenvalues) come in pairs with one the negative of the other, and their corresponding eigenvectors are complex conjugates of one another. This is to be expected since one can show that if $(\sigma, \hat{\mathbf{V}})$ is an eigenpair of (6.42), then $(-\sigma, \overline{\hat{\mathbf{V}}})$ is also an eigenpair, where $\overline{}$ denotes complex conjugation. Physically, these two eigenpairs correspond to the same mode since

$$\Re\left(\hat{\eta}(x,y)\mathrm{e}^{i\sigma t}\right) = \Re\left(\overline{\hat{\eta}(x,y)\mathrm{e}^{i\sigma t}}\right) = \Re\left(\overline{\hat{\eta}(x,y)}\mathrm{e}^{i(-\sigma)t}\right) \ . \tag{6.56}$$

Hence, of the $2N_\phi + N_\psi$ computed modes, only half of them are physically distinct.

The spatial structure of each mode $(\hat{\eta}, \hat{\phi}, \hat{\psi})$ can be computed by summing the series (6.13)-(6.14) and (6.16) using the coefficients for $(\hat{P}, \hat{Q}, \hat{R})$ that are given by the entries of $\hat{\mathbf{V}}$. Since the computed modes contain both real and imaginary parts, it is useful to write the free surface in terms of an amplitude $A(x,y)$ and a phase angle $\theta(x,y)$ as follows:

$$\begin{aligned} \eta(x,y,t) &= \Re\left(\hat{\eta}(x,y)\mathrm{e}^{i\sigma t}\right) \\ &= A(x,y)\cos(\sigma t - \theta(x,y)) \ , \end{aligned} \tag{6.57}$$

where

$$\hat{\eta}(x,y) = \hat{\eta}_r(x,y) + i\hat{\eta}_i(x,y) \ , \tag{6.58}$$

$$A(x,y) = \sqrt{\hat{\eta}_r^2 + \hat{\eta}_i^2} \ , \tag{6.59}$$

$$\theta(x,y) = \tan^{-1}\left(-\frac{\hat{\eta}_i}{\hat{\eta}_r}\right) \ . \tag{6.60}$$

In the results below, we numerically compute $\theta(x,y)$ using the four-quadrant inverse tangent function to obtain angles between -180° and 180°. By construction, high water propagates in the direction of increasing θ, and hence θ gives insight into the direction of rotation and propagation properties of the modes.

6.3 Validation – internal Kelvin and Poincaré waves in a circular basin

As a means of validating our numerical method, we have reproduced the analytical calculations performed by [25] for a two-layer flat-bottomed circular basin of radius $r_0 = 67.5$ km representing a model of a large mid-latitude lake. The acceleration due to gravity and Coriolis parameter were taken to be $g = 9.81$ ms^{-2}, $f = 10^{-4}$ s^{-1}, respectively. The upper-layer and lower-layer thicknesses were taken to be $H_1 = 15$ m and $H_2 = 60$ m, and the density jump between the upper and lower layers was taken to be $\Delta\rho = 1.74$ kg m^{-3} with a reference density of $\rho_0 = 1000$ kg m^{-3}. The calculations follow the normal modes decomposition in the vertical direction [25], so that the barotropic (surface) horizontal free modes of oscillation are computed using the long wave speed for surfaces wave $c_{bt} = \sqrt{gH}$ where $H = H_1 + H_2$ is the total depth, and the baroclinic (internal, vertical mode 1) horizontal modes are computed using the long internal wave speed $c_{bc} = \sqrt{gH_e}$ where $H_e = (\Delta\rho/\rho_0)\, H_1 H_2/H$ is the equivalent depth. To facilitate a direct comparison, we have adopted Csanady's convention [25] of absorbing all of the stratification details into H_e, rather than defining a reduced gravity, g' as is usually done throughout this thesis. The importance of the density stratification relative to the Earth's rotation is captured by the non-dimensional quantity

$$S = \frac{c}{f r_0}\,, \tag{6.61}$$

known as the Burger number, that may be interpreted as the Rossby deformation radius, c/f, that has been non-dimensionalized by the basin length scale. The assumptions that underlie the normal modes decomposition are that finite amplitude (nonlinear) effects are negligible and that the bottom is flat. Although both barotropic and baroclinic modes are accessible using this technique, in this test case we present results for a subset of the baroclinic modes.

Since there are not any depth variations in this case, there will not be any rotational modes and all modes will be gravitational. In other words, it can be shown that the coefficients matrix $D = 0$. However, we do expect two distinct types of gravitational modes – a discrete finite number of Kelvin modes that are sub-inertial ($|\sigma/f| < 1$) and a countably infinite set of Poincaré modes whose frequencies are above the inertial frequency ($|\sigma/f| > 1$).

Kelvin modes may be further characterized by: (1) they decay exponentially away from the coastline (i.e., coastally trapped), and (2) the direction of propagation is counter-clockwise in the northern hemisphere (so-called right-bounded), and clock-wise in the

southern hemisphere (left-bounded). It was first demonstrated by Lamb [71] that for circular basins, Kelvin modes are present whenever $S < 1/\sqrt{2}$. For the set of physical parameters described above, $S = 0.067 < 1/\sqrt{2}$, and we find that there are fourteen Kelvin modes.

	Current Method	Analytical	
s	σ/f	σ/f	Relative Error
1	0.069514	0.069418	0.0013844
2	0.13904	0.13883	0.0015106
3	0.20853	0.20824	0.0013735
4	0.27807	0.27764	0.0015312
5	0.34752	0.34703	0.0014061
6	0.41709	0.41641	0.0016186
7	0.48648	0.48577	0.0014589
8	0.55707	0.55512	0.0035239
9	0.6254	0.62444	0.0015399
10	0.69476	0.69375	0.001461
11	0.7643	0.76303	0.001665
12	0.83357	0.83229	0.0015304
13	0.9032	0.90153	0.0018462
14	0.9724	0.97075	0.0017009

Table 6.1: Analytical and numerical values for the non-dimensionalized frequencies (σ/f) of Kelvin modes with azimuthal mode number s in a stratified rotating basin with Burger number $S = 0.067$.

To verify the numerical method, we have compared analytically calculated Kelvin mode frequencies to those obtained numerically. The analytical calculations are obtained using a standard separation of variables technique, the details of which may be found in [25]. The resulting equation for the eigenvalues (i.e., frequencies) is in terms of Bessel functions, and thus the roots of these equations must be calculated numerically. This procedure was carried out using GNU Scientific Library (GSL) routines for rooting finding and evaluating the Bessel functions. Numerical calculations were carried out using a variety of grid parameters and expansion sizes for the streamfunction and velocity potential basis functions. In general, it is found that refining the mesh (i.e., adding more elements, K), for fixed polynomial order (N), and increasing the size of the bases all improved the agreement of the frequencies between the two methods. In the results below, the parameters were taken

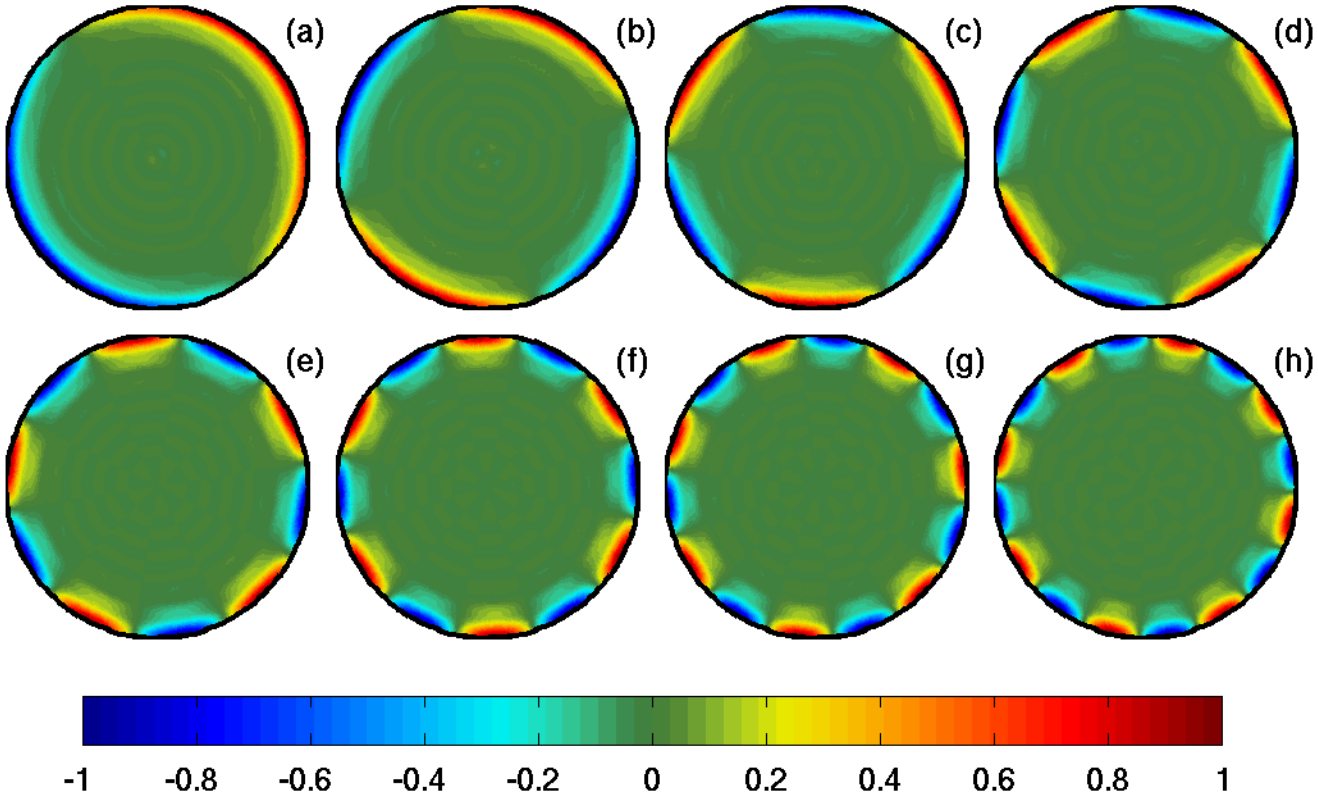

Figure 6.1: Numerically computed real part of the vertical mode-1 displacement, $\hat{\eta}_r$, for the first six Kelvin modes of a model large circular mid-latitude lake. The computed scaled frequencies (σ/f) for the modes shown in panels **(a)**–**(h)** are given in the first 8 rows in Table 6.1, respectively.

to be $N = 4$, $K = 710$, and $N_\phi = N_\psi = 200$. Table 6.1 lists the scaled Kelvin mode frequencies computed using both the analytical and numerical methods, along with the relative error. It is shown that the relative error in the computed Kelvin mode frequencies is $O(0.4\%)$ or less.

In Figure 6.1, the real parts of the spatial structure $(\hat{\eta}_r)$ of the first 8 numerically computed Kelvin modes are shown. In this case, we chose to plot the function $\hat{\eta}_r$, which may be interpreted as a snapshot of the interfacial displacement field taken at an arbitrary time, instead of plotting the amplitude and phase angles. We make this choice since the amplitudes $A(x, y)$ all carry the same structure for the Kelvin modes, and the phase angles $\theta(x, y)$ are highly irregular in regions away from the coastline where the amplitudes are

negligible.

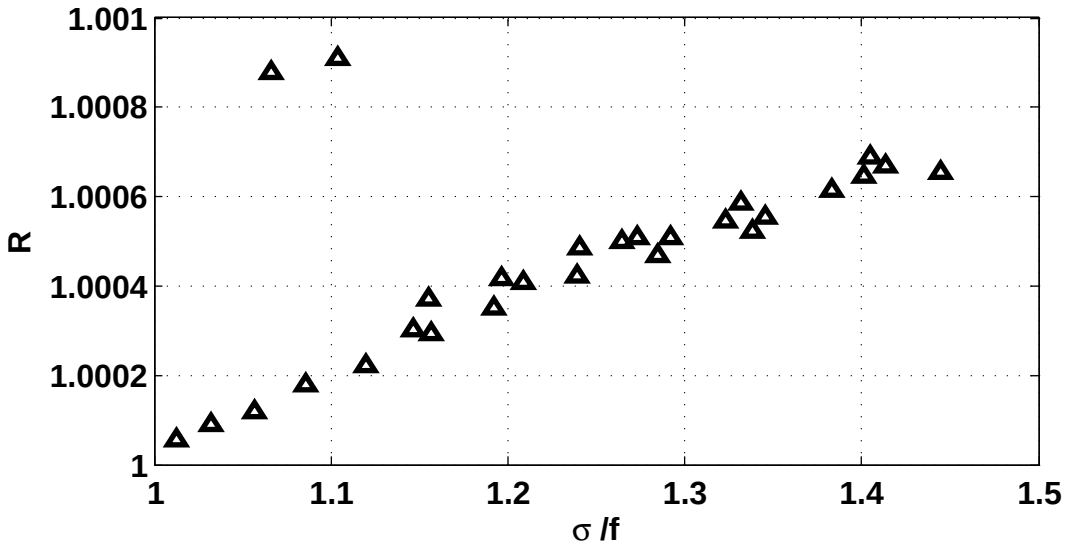

Figure 6.2: R, ratio of computed frequency to analytical frequency (i.e., $R = \sigma_{numerical}/\sigma_{analytical}$) vs. $\sigma_{analytical}/f$

In addition to Kelvin modes, we have also calculated a number of Poincaré modes and compared their frequencies to those computed analytically. The results are shown in Figure 6.2 where R, the ratio of computed frequency to analytical frequency, is plotted against the scaled analytically calculated frequency $\sigma_{analytical}/f$ for the first 28 Poincaré modes. Good agreement is found between the analytical and numerical methods since R does not exceed 1.001 for all modes considered here.

6.4 Validation – normal modes in a basin with parabolic bottom

The method remains to be validated in cases that involve variable depth where rotational modes exist, along with gravitational modes. To carry out such a validation, we refer to the analytical result of Lamb [71] for a circular basin with a parabolic bottom. The

depth-profile is defined by

$$H(\mathbf{x}) = H_0 \left[1 - \left(\frac{r}{r_0} \right)^2 \right] , \tag{6.62}$$

where H_0 is the depth of the centre of the basin, r_0 is the radius of the lake and $r = \sqrt{x^2 + y^2}$. The fact that $H \to 0$ is problematic for the method described in Section 6.2 since the homogeneous eigenvalue problems (6.7) and (6.9) are singular at $r = r_0$ and the numerical eigenvalue solver used to construct the expansion bases does not converge. To side-step this difficulty, we perform our calculations on the perturbed depth profile $\tilde{H}(\mathbf{x}) = H(\mathbf{x}) + 0.005H_0$, under the assumption that such a small adjustment to the depth would not drastically alter the solutions. It was found that reducing the size of the perturbation polluted the spatial structure of the eigenmodes, and the frequencies converge to unreliable values. Examining the $\beta = 40$ case for the clockwise rotating mode with linear structure, it was found that decreasing the size of the perturbation caused the scheme to converge to an incorrect frequency of $\sigma/f = -1.0471$, while the correct frequency is $\sigma/f = -1.048$.

Lamb [71] lists frequencies for three values of the non-dimensional parameter $\beta = (fr_0/c_0)^2$, i.e., the square of the inverse Burger number with $c_0 = \sqrt{gH_0}$, and two specialized mode structures with azimuthal mode number $s = 1$. The radial structure of these modes is given by a terminating hypergeometric series. The first specialized mode has a linear radial structure, hence η is represented by a plane, and the second mode considered has a cubic structure in r. The plane mode appears twice in the set of modes, once as a clockwise rotating ($\sigma < 0$) gravity mode and once as a counter-clockwise rotating ($\sigma > 0$) rotational mode. The cubic mode, on the other hand, appears three times: once as a counter-clockwise rotational mode and also as a pair of counter-rotating gravity modes of different absolute frequency ($|\sigma|/f > 1$). The frequencies calculated by our method are compared to those given by Lamb in Table 6.2. Fair agreement of $O(1\%)$ or less is found in all cases except for the counter-clockwise cubic mode with $\beta = 40$ where the error was found to be $O(6\%)$. It is believed that the overall agreement is poorer in this case than for the Kelvin modes case (Section 6.3) since it was necessary to use a perturbed form of the depth profile used in the analytical solution, and the overall sensitivity of each mode to the perturbation for different β values is unclear.

s	Radial Structure	β	Current Method	Lamb (1932)	Relative Error
1	linear	2	-1.6229	-1.618	0.0030
			0.62366	0.618	0.0092
		6	-1.2658	-1.264	0.0014
			0.26657	0.264	0.0097
		40	-1.0476	-1.048	0.00038
			0.04837	0.048	0.0077
1	cubic	2	-2.9166	-2.889	0.0096
			0.62366	0.618	0.0092
			2.7912	2.764	0.0098
		6	-1.887	-1.874	0.0069
			0.10027	0.100	0.0027
			1.7858	1.774	0.0067
		40	-1.1801	-1.183	0.0025
			0.03757	0.040	0.0608
			1.1411	1.143	0.0017

Table 6.2: Comparison between scaled frequencies σ/f calculated using the present method and the analytical results of [71] for the free modes of oscillation in a circular lake with parabolic bottom. The parameter s is the azimuthal mode number, and β is the square of the inverse Burger number, i.e., $(fr_0/c_0)^2$.

6.5 Comparison with simulations: Kelvin waves on a model large circular mid-latitude lake

The method discussed above can also be thought of as a validation tool for numerical simulations of shallow water models where closed-form analytical solutions are not available. It can also be used to help interpret simulation results by determining which free modes of oscillation have been excited in the solution. In this section, we initialize the one-layer DG-FEM model from Chapter 4 with the super-position of the real part of the first three Kelvin modes from Csanady's large circular mid-latitude lake, shown in Figure 6.1**(a)**–**(c)**. The initial velocity $\mathbf{u}$ of the vertical mode-1 Kelvin waves were recovered from the stream-function and potential using the Helmholtz decomposition (6.3). The initialization thus takes the form

$$\begin{pmatrix} \eta_0 \\ \mathbf{u}_0 \end{pmatrix} = \sum_{i=1}^{3} \begin{pmatrix} \eta^{(i)} \\ \mathbf{u}^{(i)} \end{pmatrix} , \tag{6.63}$$

where superscript (i) represents the i^{th} gravitational mode. Here, each mode $(\eta^{(i)}, \mathbf{u}^{(i)})^{\mathsf{T}}$ has been scaled such that $\max(\eta^{(i)}) = 10^{-5} H_e$ so that nonlinear effects would be negligible. The order $N = 4$ DG-FEM method was used along with the same $K = 710$ element triangular mesh used to compute the modes. The solution was integrated out to 4 periods of the lowest frequency Kelvin wave.

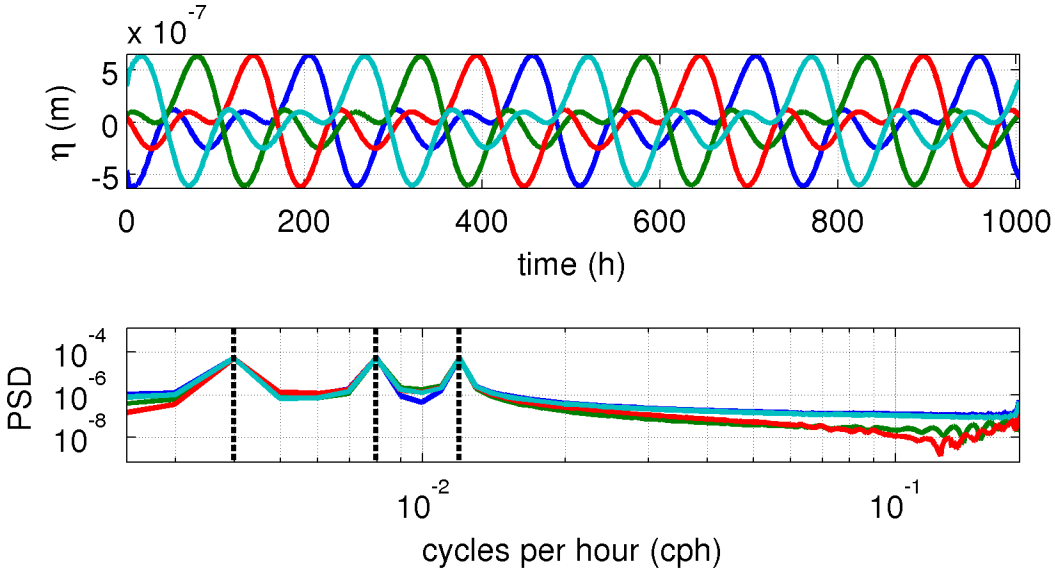

Figure 6.3: Top: Time series taken of η from the 'Kelvin waves on a model large circular mid-latitude lake' simulation taken at the points $(x, y) = (0, r_0)$ (blue), $(0, -r_0)$ (green), $(r_0, 0)$ (red), $(-r_0, 0)$ (cyan). Bottom: Corresponding power spectral density for each of the curves in the top panel. Dashed curves represent theoretical frequencies of the first three Kelvin modes with calculated periods of $T = 251$ h, $T = 125.5$ h, and $T = 83.7$ h.

In Figure 6.3, we plot the time series and corresponding power spectral density of the interfacial displacement η at the four points corresponding to the ends of the basin in the four compass directions. We find that the spectral analysis recovers the theoretically predicted frequencies of the three Kelvin waves quite well, since the three peaks in the spectrum coincide directly with the calculated periods of $T = 251$ h, $T = 125.5$ h, and $T = 83.7$ h.

6.6 Normal modes in a lake with a wavy coastline

In this section, we revisit the idealized basin considered in Section 4.12 that represents a rose-petal-like perturbation on the classical circular geometry. Here, we choose the same physical parameters used in Chapter 4. The Coriolis frequency is given by $f = 7.8828 \times 10^{-5}$ s^{-1}, the depth is $H = 12.8$ m, and the reduced gravity is $g' = 0.024525$ m s^{-2}. Given that the typical length scale is $r_0 = 8345$ m, the Burger number in this case is $S = 0.852$.

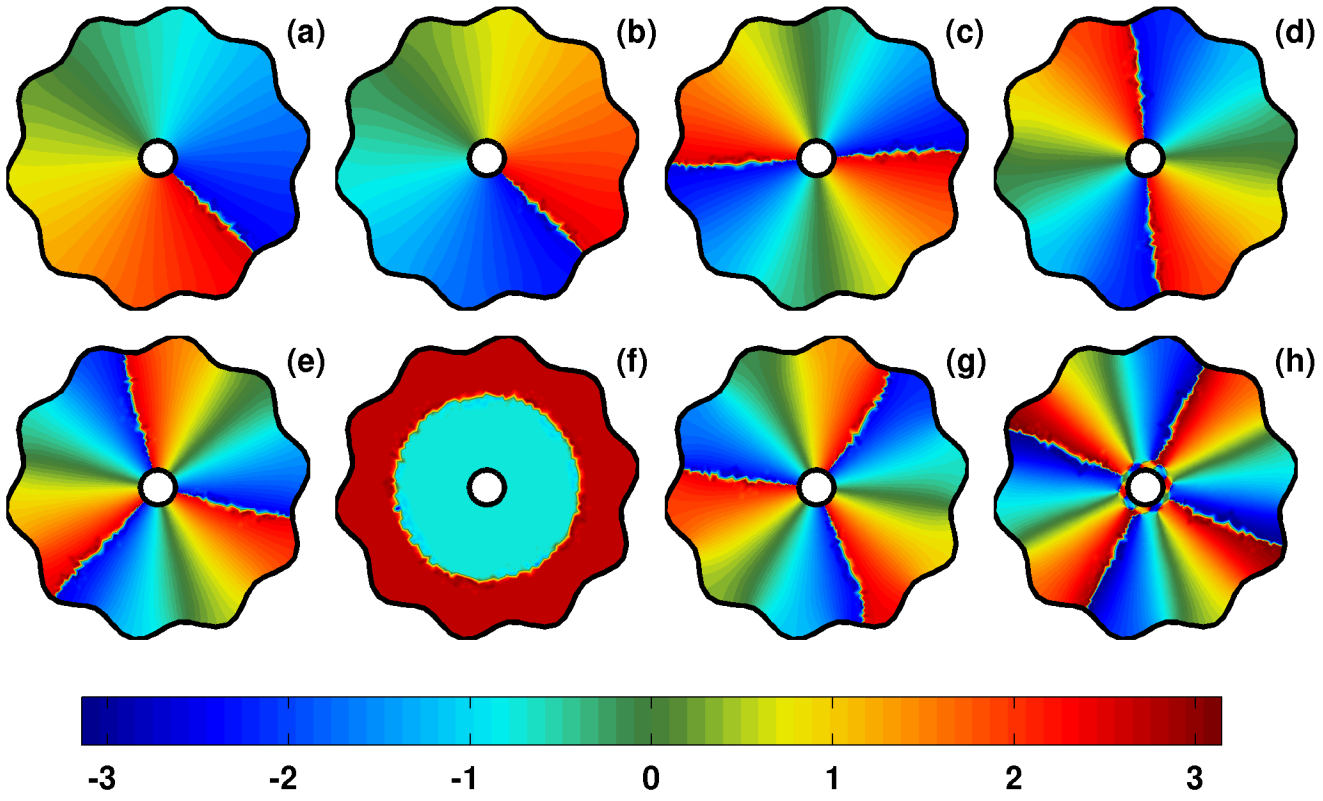

Figure 6.4: Phase angles $\theta(x, y)$ of 8 gravitational modes of the model lake with a wavy coastline. The corresponding periods of the modes are given by **(a)** $T = 18.21$ h, **(b)** $T = 11.26$ h, **(c)** $T = 9.89$ h, **(d)** $T = 7.37$ h, **(e)** $T = 6.97$ h, **(f)** $T = 6.45$ h, **(g)** $T = 5.71$ h, **(h)** $T = 5.49$ h.

The free modes of oscillation were computed using the $K = 1330$ element mesh used previously for this geometry with polynomial order $N = 4$ and $N_\psi = N_\phi = 200$ basis

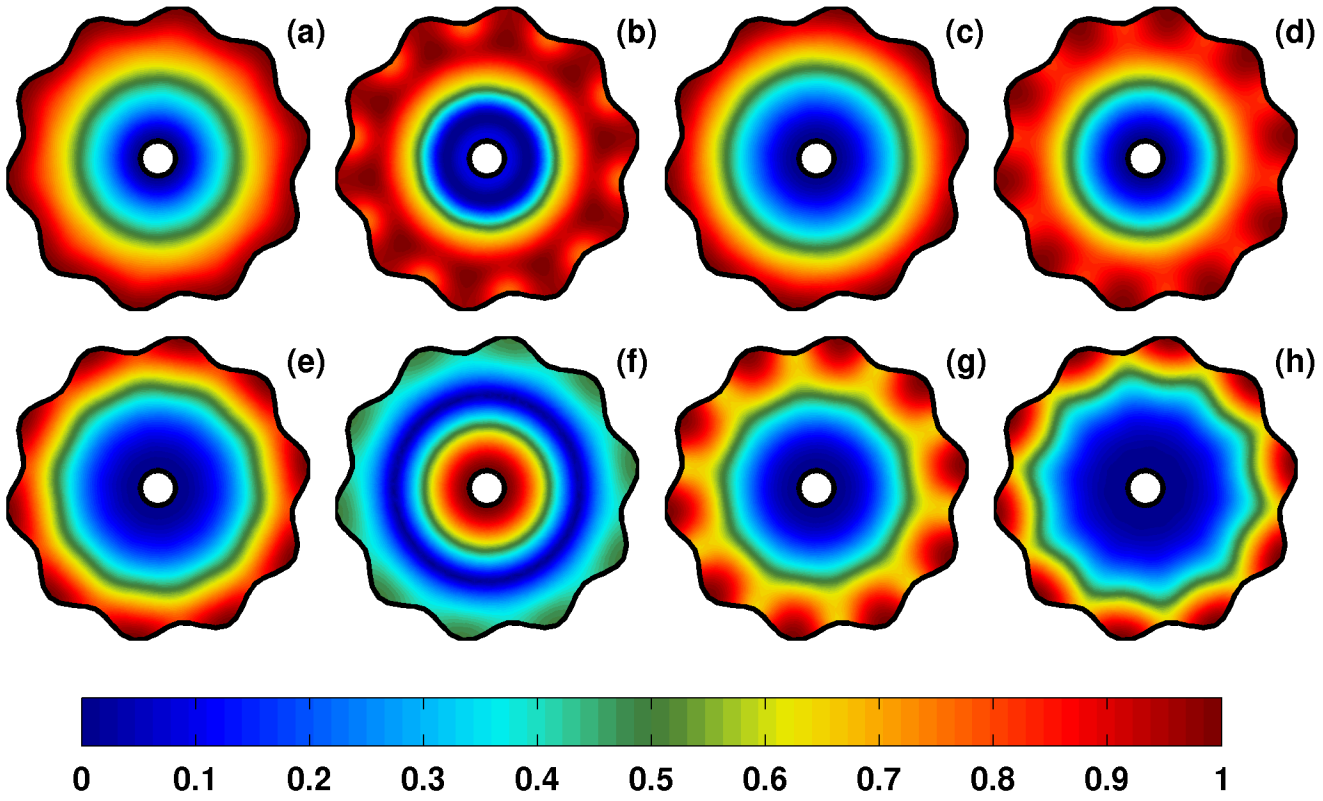

Figure 6.5: Like Figure 6.4, but the amplitude $A(x, y)$ is plotted. $A(x, y)$ is scaled to have a maximum value of 1.

functions. The phase angles $\theta(x, y)$ and amplitudes $A(x, y)$ of the first 8 free modes of oscillation are presented in Figures 6.4 and 6.5, respectively. The plotted modes have periods between 18.21 h and 5.49 h, all of which are below the inertial period since the lowest frequency mode has $\sigma/f = 1.22 > 1$. Hence, these modes can be thought of as Poincaré waves [70, 115].

Recalling that high water propagates in the direction of increasing θ, we find that the first four free modes (Fig. 6.4 **(a)**–**(d)**) appear in counter-rotating pairs of different frequency with the lowest frequency mode rotating counter-clockwise. Interestingly, a radially-symmetric mode (panel **(f)**) appears between the azimuthal mode-3 pair (panels **(e)** and **(g)**). The azimuthal mode-4 mode (panel **(h)**) shows a small cell near the centre island that is out of phase with the dominant outer motion, though both propagate counter-clockwise. Inspecting the amplitudes in Figure 6.5, we find that nodal points

($A = 0$) typically occur at or near the centre of the basin for these large-scale modes, except in the case of the radially symmetric mode (panel **(f)**). The different concentrations of amplitude along the outer coast are also worth noting, with some modes showing stronger intensification in the small dimples or 'bays.'

6.6.1 Effects of stronger rotation

To illustrate the impact of stronger planetary rotation (or equivalently, weaker stratification), we have re-computed the free modes of oscillation on the same basin considered in the above section but f has been increased to $f = 1.4554 \times 10^{-4}$ s^{-1} (North Pole rotation) and $f = 10^{-3}$ s^{-1} (an artificially high value) such that the corresponding Burger numbers are $S = 0.462$ and $S = 0.067$. In Figure 6.6, we show a comparison between the amplitudes of the lowest frequency gravity modes for $S = 0.852$ (same as Figure 6.5 **(a)**) and these two new values for the Burger number. In Figure 6.6 **(b)**, we see that increasing the Burger number to $S = 0.0462$ has decreased the scaled frequency from $\sigma/f = 1.22$ to $\sigma/f = 0.581$, but only slightly altered the amplitude distribution such that the $O(1)$ values of $A(x, y)$ are shifted closer to the outer coastline. Decreasing S to 0.067 in panel **(c)** causes the mode to be highly trapped along the outer coastline and its scaled frequency decreases to $\sigma/f = 0.0652$.

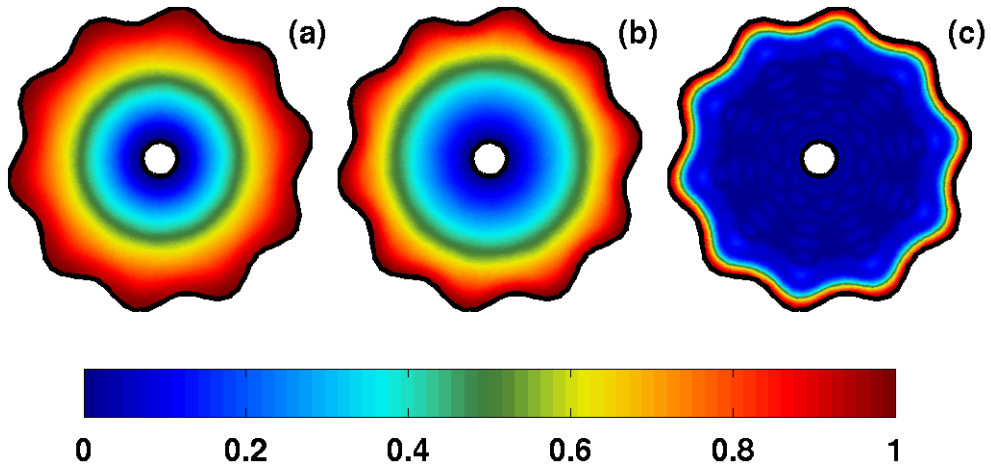

Figure 6.6: Amplitudes $A(x, y)$ of the lowest frequency modes on the lake with a wavy coastline and varying Burger numbers S. Each panel corresponds to **(a)** $S = 0.852$, $\sigma/f = 1.22$, **(b)** $S = 0.462$, $\sigma/f = 0.581$, **(c)** $S = 0.067$, $\sigma/f = 0.0652$.

The value $S = 0.067$ corresponds to the same value of S as Csanady's model large circular mid-latitude lake in Section 6.5, and the mode pictured in panel **(c)** is analogous to an azimuthal mode-1 Kelvin wave on a circular basin. The results shown here illustrate the transition of the lowest frequency free mode of oscillation on a nearly circular basin from the super-inertial $|\sigma/f| > 1$ to the sub-inertial $|\sigma/f| < 1$ frequency regimes. In other words, the mode transitions from a Poincaré wave to a Kelvin wave. While the results shown here indicate that the lowest frequency mode always possesses azimuthal mode-1 shape and rotates counter-clockwise, the effect of increased rotation (or decreased stratification) tends to increase the extent to which the mode is trapped along the outer coastline, i.e., to behave more like a Kelvin wave propagating along a straight and infinite coast.

6.6.2 Effects of variable bathymetry

Up until this point, we have focused purely on lakes with a flat bottom, that is $H = const.$ We now turn our attention to the effects of bottom topography on the free modes of oscillation in the same idealized geometry as above, but we define the depth profile as

$$H = H_0(3 - \tanh(x/\lambda))/4 \,, \tag{6.64}$$

with $H_0 = 12.8$ m and $\lambda = 2$ km, such that the depth decreases by a factor of two (from $H = 12.8$ m to $H = 6.4$ m) across a thin transition region centred about $x = 0$. The linear modes calculation was carried out with $N_\phi = N_\psi = 300$ basis functions and the mesh and polynomial order were taken to be the same as in previous sections: $(N, K) = (4, 1394)$. Figures 6.7 and 6.8 plot the amplitudes and phases of the first 6 gravitational modes, respectively. Here, the Burger number is $S = \sqrt{g'\overline{H}}/fr_0 = 0.7376$, where $r_0 = 8345$ m, $\overline{H} = 9.6$ m, $g' = 0.0025g = 0.025$ ms^{-2} and $f = 7.8828 \times 10^{-5}$ s^{-1}.

Figures 6.7 and 6.8 should be compared and contrasted with the flat bottom case shown previously in Figures 6.4 and 6.5. A main result is that the amplitude of each mode tends to be stronger at the eastern end of the basin where the depth is lower. Furthermore, nodal points appear to be shifted to the right, whereas many of the modes in the flat bottom case were centred about $(x, y) = (0, 0)$. Comparing the phase angle plots, the breaking of symmetry due to changes in depth is a clearly evident aspect of the results (see for example, panel **(e)**), and such an effect is indicative of the fact that wave refraction plays an important role in dictating the spatial structure of the modes. More rapid phase changes in shallow water than in deep water are visible as well.

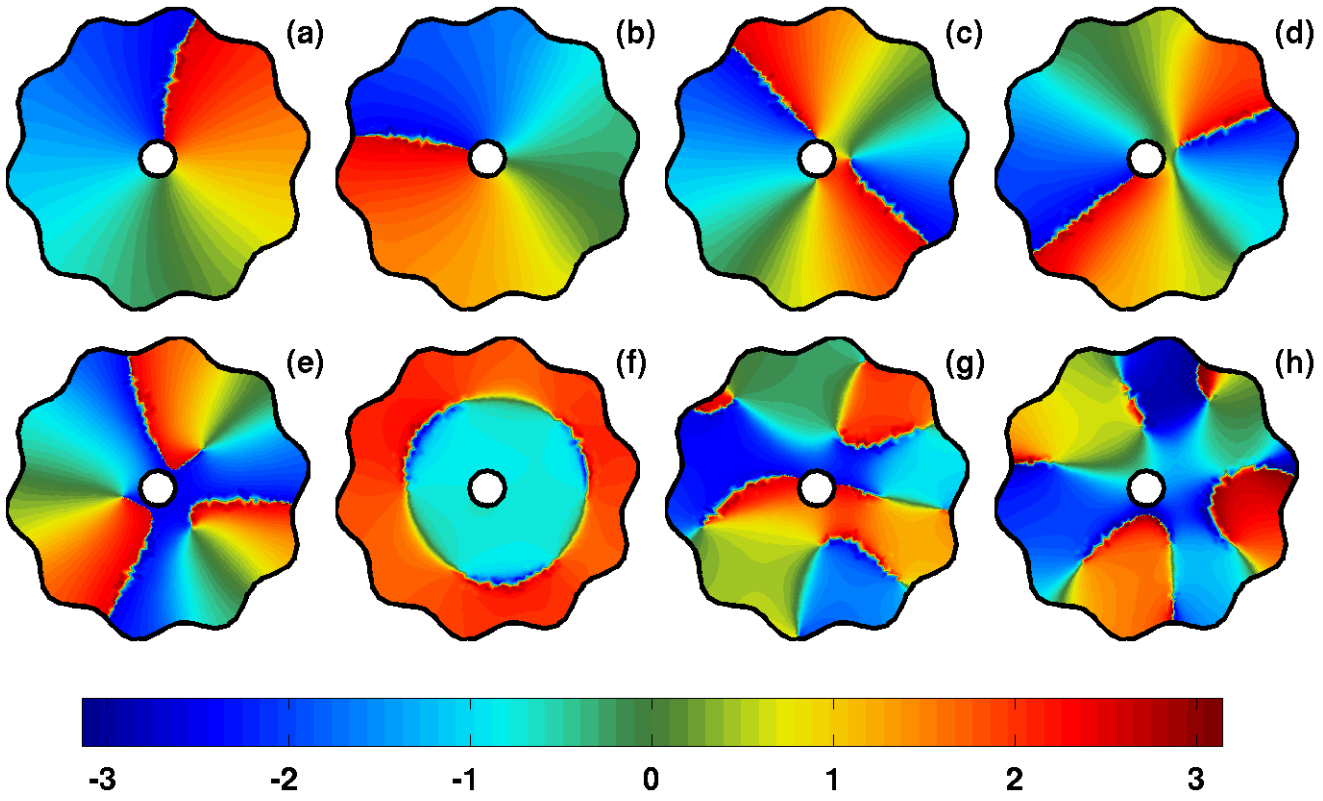

Figure 6.7: Phase angles $\theta(x, y)$ of 8 gravitational modes of the model lake with a wavy coastline and variable bathymetry. The corresponding periods of the modes are given by **(a)** $T = 22.41$ h, **(b)** $T = 12.79$ h, **(c)** $T = 12.09$ h, **(d)** $T = 8.58$ h, **(e)** $T = 8.50$ h, **(f)** $T = 7.78$ h, **(g)** $T = 6.83$ h, **(h)** $T = 6.62$ h.

Rotational modes

In addition to rotating gravity modes, 150 distinct rotational modes were calculated using the method. This is the first occurrence of rotational modes presented thus far, and in the absence of a variation in the Coriolis parameter, the presence of such modes is solely due to a non-flat bottom. Here, we choose the mode with the largest spatial structure for examination. Since it has the largest structure, it is presumed to be the most important rotational linear mode in the physics of the model lake considered in this section because it would be forced by large scale wind patterns.

In Figure 6.9, snapshots of $\Re(\eta)/\max_{x,y}(|\hat{\eta}|)$ are shown at fixed times spanning one full

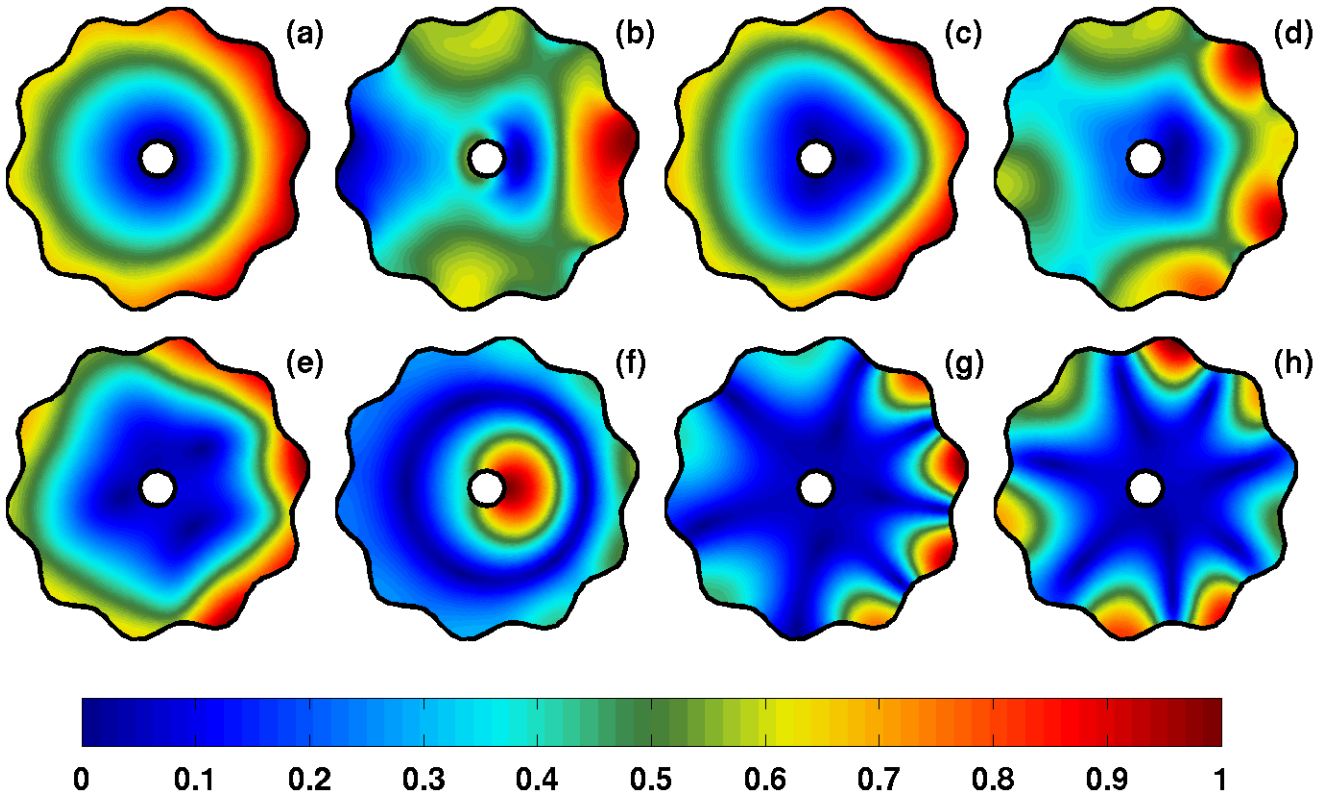

Figure 6.8: Like Figure 6.7, but the amplitude $A(x, y)$ is plotted. $A(x, y)$ is scaled to have a maximum value of 1.

wave period of $T = 195.6$ h. The mode behaves as a topographically trapped wave that propagates along the region centred about $x = 0$, where the gradient of the depth profile is highest. The existence of such a wave is analogous to 'Rossby waves' on the beta plane [92, 70]; westward propagating vorticity waves that exist due to a northward gradient in planetary vorticity. The mode depicted in Figure 6.9 is a northward propagating wave that exists due to an eastward (positive x) depth gradient that corresponds to a westward gradient in potential vorticity. Due to this analogy, these waves have been referred to as 'topographic Rossby waves' [116]. The traditional (non-dispersive) shallow water model guarantees that potential vorticity defined by $q = (\zeta + f)/H$ is conserved following the flow [92], i.e.,

$$\frac{Dq}{Dt} = 0 \,, \tag{6.65}$$

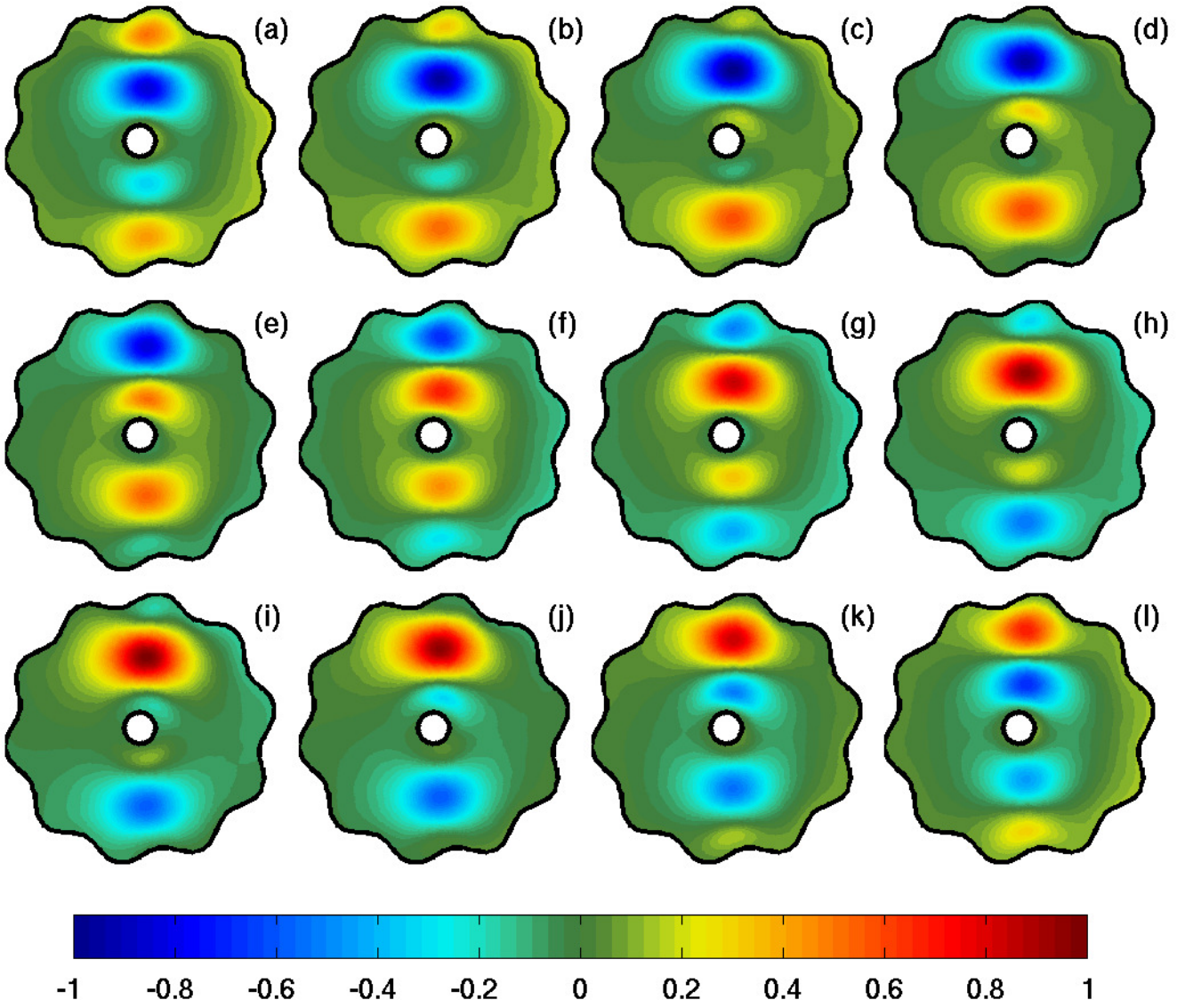

Figure 6.9: Snapshots of one full wave period of $\Re(\eta)/\max_{x,y}(|\hat{\eta}|)$ of the largest scale rotational mode computed in the perturbed circular basin with variable bathymetry. The times of each panel are given by **(a)** $t = 0$, **(b)** $t = 0.083T$, **(c)** $t = 0.167T$, **(d)** $t = 0.25T$, **(e)** $t = 0.333T$, **(f)** $t = 0.417T$, **(g)** $t = 0.5T$, **(h)** $t = 0.583T$, **(i)** $t = 0.667T$,**(j)** $t = 0.75T$, **(k)** $t = 0.833T$,**(l)** $t = 0.917T$. Here, $T = 195.6$ h.

where $\zeta = v_x - u_y$ is the relative vorticity.

6.7 Comparison with simulations: Poincaré waves on a lake with a wavy coastline

In this section, we carry out a simulation where we initialize the single-layer model from Chapter 4 on the lake with a wavy coastline discussed in Section 4.12 with the lowest frequency gravity mode, and we scale its maximum amplitude to $a = 10^{-5}H$ so that nonlinear effects are negligible. It is assumed that since the mode is large-scale, that dispersive effects should be negligible as well and the model should behave as the linear shallow water equations. The model was initialized with the real part of the lowest frequency mode with period $T = 18.21$ h. The phase angle and amplitude of this mode is depicted in Figures 6.4 **(a)** and 6.5 **(a)**, respectively. The simulation was time-stepped to $t = 182$ h, or ten periods of the lowest frequency mode.

Here, in Figure 6.10 **(a)**, we plot the time series of the interfacial height η as measured from four points throughout the basin. The corresponding power spectral density is shown in panel **(b)**, where the most prominent peak in all curves agrees quite well with the frequency predicted from the linear theory calculation, and there do not appear to be any other well-defined peaks in the spectrum.

In Figure 6.11, the effects of nonlinearity are explored for the same geometric configuration by increasing the maximum amplitude of the initial η to $a = 0.1H$. The initial velocity has been re-scaled appropriately as well. We have time-stepped the equations for four periods of the lowest frequency mode, or twice as long as in Figure 6.10. We once again find that the most prominent peak in all power spectral curves corresponds to the lowest frequency predicted by the linear theory calculation. For the time series taken near the outer coastline, we find that nonlinear effects have excited a number of other peaks in the spectrum. We do not find these same peaks for the cyan curve that is taken near the inner island due to the fact that nonlinear steepening is localized along the outer coast, and the dynamics in the centre of the basin remain essentially linear. It is worth noticing that for the timescale of the simulation carried out here, the peaks excited due to nonlinearities do not appear to directly coincide with the frequencies of the modes predicted by linear theory. Snapshots of the evolving η field are shown in Figure 6.12 to illustrate the effects of local nonlinear steepening on the basin-scale wave. If the dynamics were purely linear, all panels would be identical to the initial condition (panel **(a)**) since the times in panels **(b)**–**(d)** are given at integer multiples of the mode's period. Instead, there is a clear

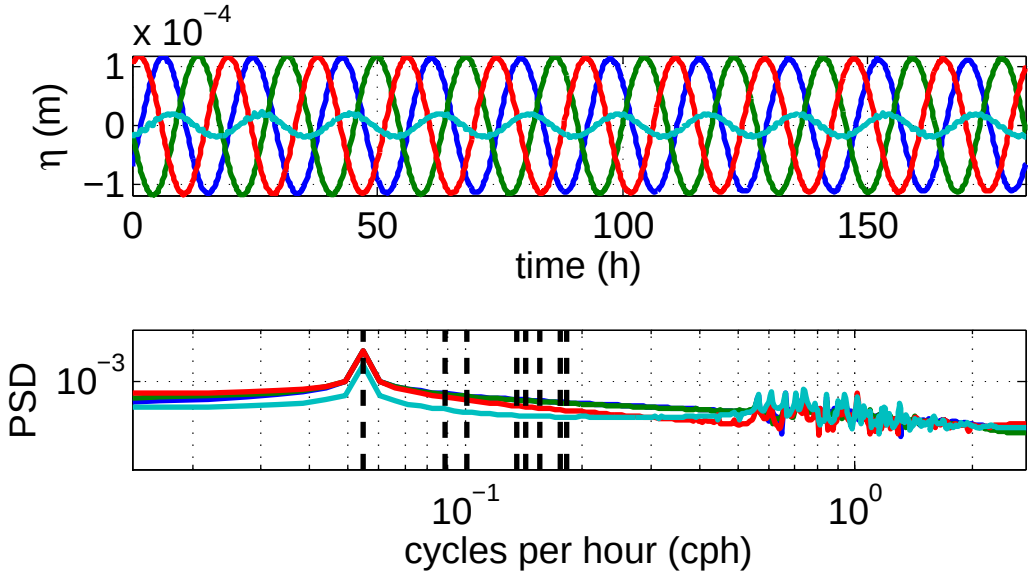

Figure 6.10: Panel **(a)**: Time series of η for a simulation initialized with the lowest-frequency Poincaré mode at the points shown in Figure 4.8 **(b)**. The blue, green, and red curves correspond to points near the outer coastline, while the cyan curve corresponds to a point near the centre island. The corresponding power spectral density is shown in panel **(b)** where the dashed vertical line represents the predicted frequency of 0.0549 cph (or $T = 18.21$ h).

steepening of the wave visible in panels **(b)** through **(d)**. This steepening preferentially occurs along the leading edge of the wave front, becoming more pronounced with each wave period, and it is clear by contrasting with Figure 6.5 that it does not project onto any of the first eight Poincaré modes in a clean manner.

In light of the above discussion, it is worth exploring whether one of these clockwise harmonics would remain coherent if a simulation was initialized with one. Such a case is considered in Figure 6.13 where we initialize with the Poincaré mode with the 4th lowest frequency, a mode that exhibits clockwise rotation (see Figure 6.4 **(d)**) with period $T = 7.37$ h. The simulation was time-stepped to $t = 29.48$ h, or four modal periods. In direct analogy with the results presented in Figure 6.10, we find a single peak in the spectrum that corresponds precisely to the frequency of the mode used to initialize the simulation.

In Figure 6.14, we consider the question of which gravity modes would be excited by an initial condition corresponding to a linear east-west tilt in the interface with no initial

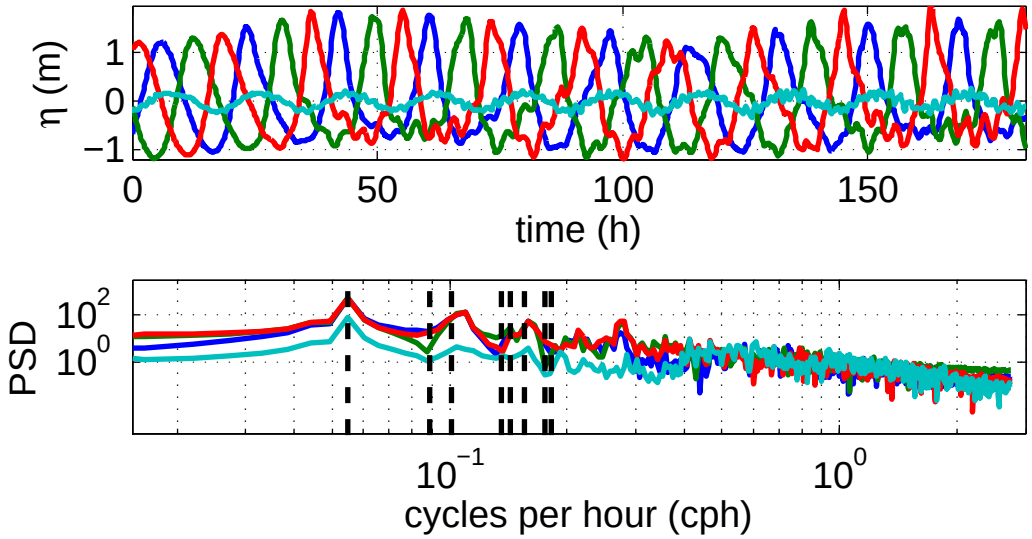

Figure 6.11: Like Figure 6.10, but the amplitude has been increased to $a = 10^{-1}H$ and the integration time has been doubled. The dashed lines correspond to the theoretically predicted frequencies of the first 8 modes.

velocity. Nonlinear effects were once again suppressed by taking the maximum amplitude to be $a = 10^{-5}H$. In this case, the time series plots suggest that the solution is comprised of the superposition of many modes since we do not observe simple sinusoidal wave forms as we did when we our initialization consisted of one or more of the modes themselves. The power spectra suggest that the two lowest frequency modes, both of azimuthal mode 1 shape, have been excited. Here, the second lowest mode rotates clockwise. Contrary to the other time series, the time series taken near the centre island (cyan curve) shows a stronger peak at the second mode than the first mode. This observation agrees with the fact that the amplitude $A(x, y)$ is lower at the centre island for the first mode than the second mode (cf. panels **(a)** and **(b)** of Figure 6.5). It appears that the third gravitational mode (counter-clockwise, azimuthal mode 2) may have been excited as well, however it is not as clear as for the first two modes.

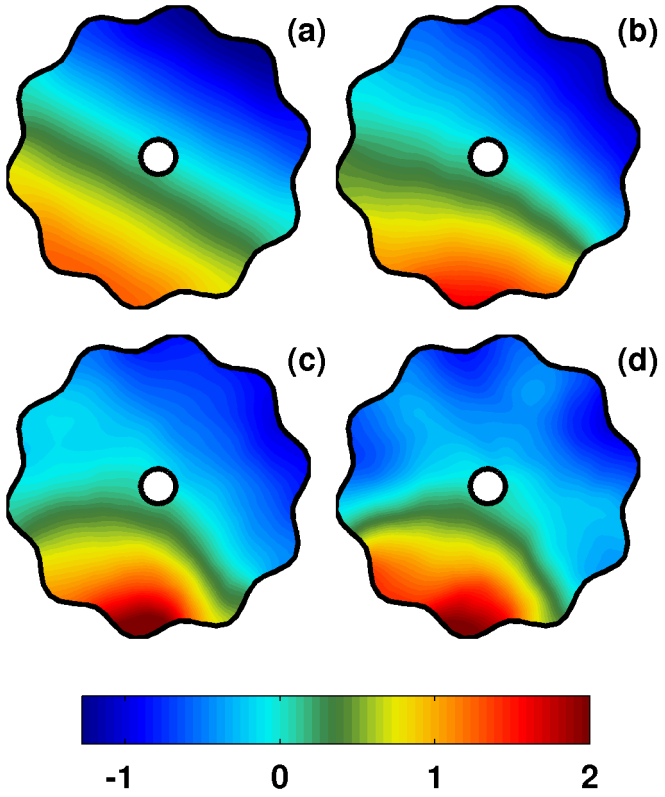

Figure 6.12: Snapshots of the η field from the simulation initialized with the lowest gravity mode and maximum amplitude $a = 0.1H$ at **(a)** $t = 0$, **(b)** $t = T$, **(c)** $t = 2T$, **(d)** $t = 3T$, where $T = 18.2$ h is the period of the mode as predicted from the linear theory calculation.

6.8 Free modes of oscillation in Pinehurst Lake, AB

We next return to the situation considered in Section 4.16 of Pinehurst Lake, Alberta. The mesh and bathymetry data used here are the same as those in Figure 4.15, and the physical parameters were taken to be the same as previously: $g' = (\Delta\rho/\rho_0)g = 0.024525$ m s^{-2}, where $(\Delta\rho/\rho_0) = 0.0025$, and $f = 1.1863 \times 10^{-4}$ s^{-1}. We have opted to use the same parameters as the simulation reported on in Chapter 4, since the details from the calculated modes may help to shed further light on the results of the simulation considered there. A rough estimate for the Burger number is $S \approx \sqrt{g'\overline{H}}/(f\sqrt{A}) \approx 0.74$, where $A = 39.0$ km^2 is the area of the basin and $\overline{H} = 12.4$ m is the mean depth, both were computed numerically.

The phase angles and amplitudes of the first 8 modes are shown in Figures 6.15 and 6.16,

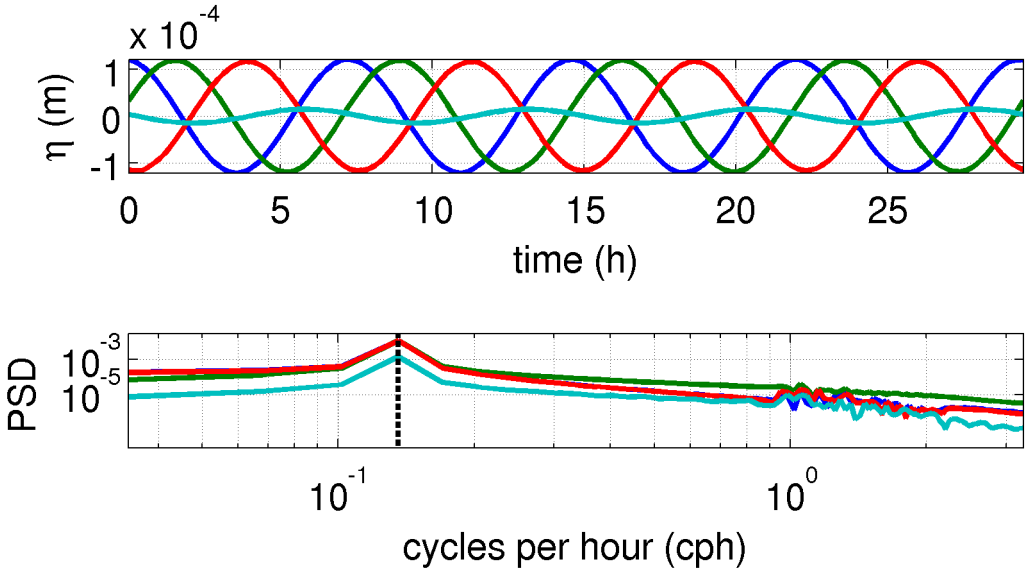

Figure 6.13: Like Figure 6.10, but the initial condition was taken to be the 4^{th} lowest frequency Poincaré mode with maximum amplitude $a = 10^{-5}H$. The dashed line corresponds to the theoretically predicted period of $T = 7.37$ h.

respectively. Unlike in the geometrically simpler case of the perturbed circular basin considered in Section 6.6, we do not simply find pairs of counter-rotating modes of different frequency and slightly different spatial structure since the basin's geometric complexity breaks the majority of the modes into smaller cells that exhibit their own local rotation/propagation properties. The exception is the lowest frequency mode shown in panel **(a)** with period $T = 12.84$ h. This mode exhibits basin-wide counter-clockwise rotation about a single nodal point near the centre of the basin.

Inspecting the amplitudes of the various basin modes in Figure 6.16, we notice that the regions of highest amplitude are typically found in confined, narrow bays, especially in the shallows in the eastern part of the basin. This result provides insight on the simulation presented in Section 4.16 where it was found that the seiche-induced flow past coastal obstacles was intensified by geometric focusing in narrow and confined parts of the basin. Such energy focusing coincides well with the excitation of the various low frequency modes as predicted by the linear theory calculation.

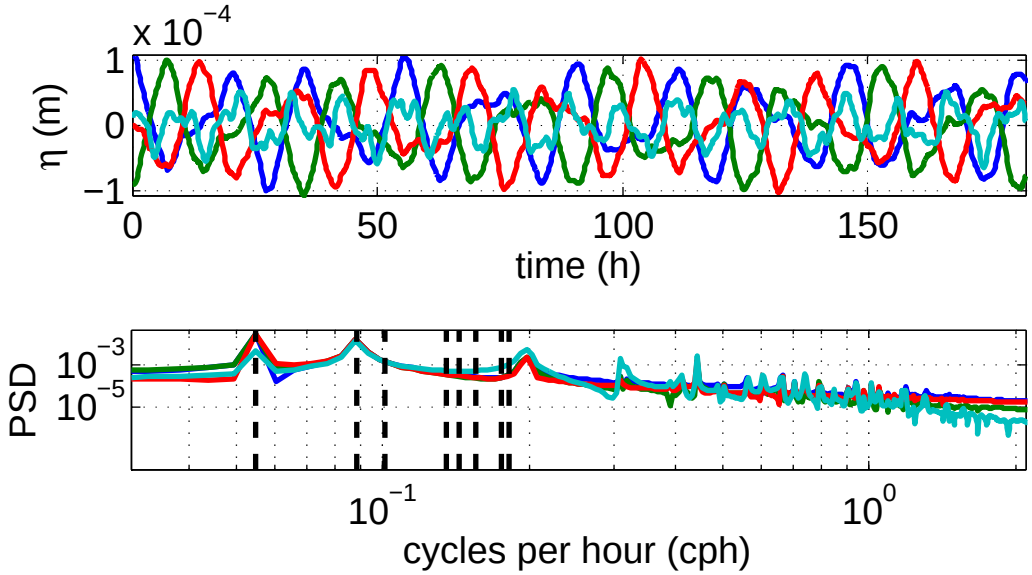

Figure 6.14: Like Figure 6.11, but the intial condition was taken to be an east-west linear tilt with maximum amplitude $a = 10^{-5}H$.

6.9 Discussion and conclusions

In this chapter, the calculation of the free modes of oscillation in arbitrarily shaped, closed basins with rotation was considered, and the modes themselves were studied. As explained in previous works, [99, 100], the standard Laplace equation eigenvalue problem obtained in the absence of rotation, no longer holds when rotation is included, and hence the inclusion of f-plane rotation complicates the calculation considerably. As a result, both Helmholtz and Galerkin decomposition techniques are required to recover a self-adjoint matrix eigenvalue problem for the unknown frequencies and Galerkin projection coefficients of the free modes of oscillation. In this work, the continuous differential operators have been discretized using the symmetric interior penalty discontinuous Galerkin (SIP-DG) method that was introduced previously in Chapter 4. The SIP-DG method allows for the construction of basis functions from discretized generalized eigenvalue problems that are specified in terms of symmetric matrices.

Once the linear free modes of oscillation are calculated for a given physical situation, two key questions arise. Firstly, are these modes generated in field situations when driven by wind forcing? Secondly, do the modes persist in the basin-scale physics once generated? In this chapter, we have primarily focused on the second question and have found that if

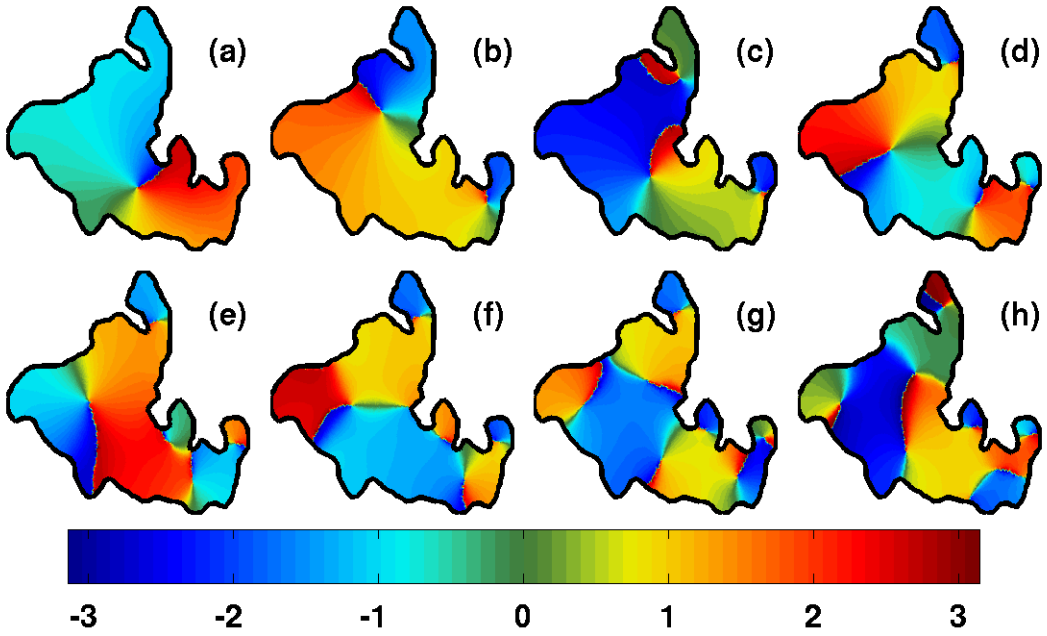

Figure 6.15: Phase angles $\theta(x, y)$ of the lowest 8 gravitational modes of Pinehurst Lake. The corresponding periods of the modes are given by **(a)** $T = 12.84$ h, **(b)** $T = 9.57$ h, **(c)** $T = 7.66$ h, **(d)** $T = 5.67$ h, **(e)** $T = 5.00$ h, **(f)** $T = 4.45$ h, **(g)** $T = 3.67$ h, **(h)** $T = 3.24$ h.

a simulation is initialized with a mode (or linear super-position of modes) then the mode remains intact throughout the simulation as long as nonlinear (finite amplitude) effects are negligible. In simulations where nonlinear effects were prevalent, low frequency modes were found to steepen and lead to the excitation of higher-frequency modes. It was also found that there were nonlinearly excited peaks in the power spectra that did not directly match any of the frequencies from linear theory. Question (1) could be the subject of future work and could be explored either by including a wind-forcing parameterization in the model simulations or by including forcing terms in the decomposition problem.

The method has been used to investigate the effects of rotation and variable bathymetry in a simple model lake. It was found that stronger rotation tends to introduce modes that are trapped along the coast. This observation is consistent with the theoretical predictions for the channel geometry (see 2.4.1). Variable bottom bathymetry, on the other hand, modifies gravity modes by enhancing their amplitude in more shallow regions. Variable bathymetry also causes a potential vorticity gradient in the background state that intro-

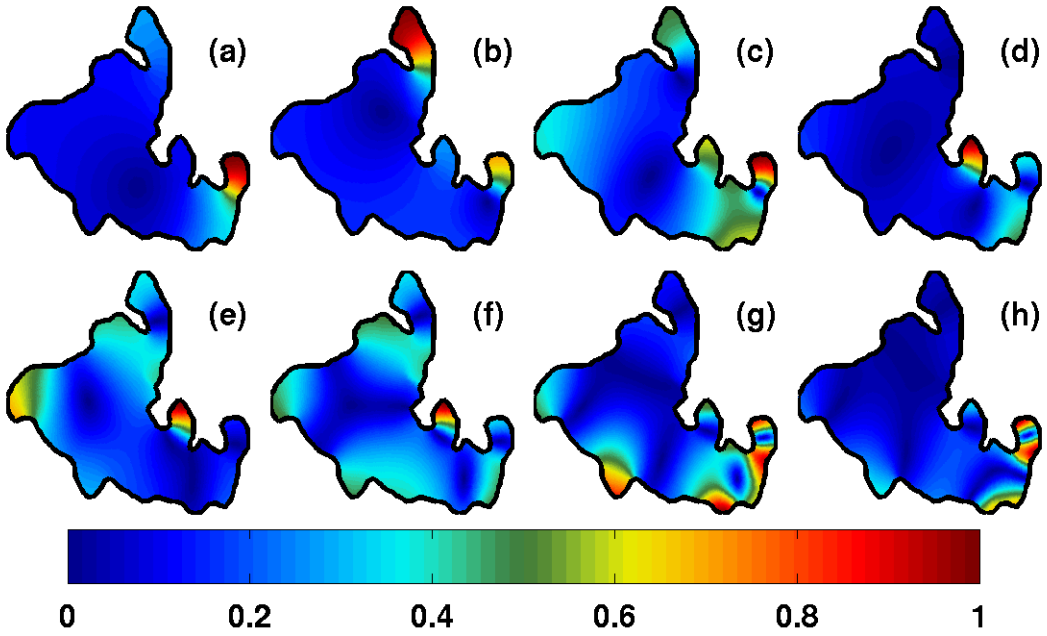

Figure 6.16: Like Figure 6.15, but the amplitudes $A(x, y)$ are plotted. $A(x, y)$ is scaled to have a maximum value of 1.

duces a corresponding family of rotational waves into the spectrum. These rotational modes are analogous to planetary-scale Rossby waves that exist due to the meridional gradient of the Coriolis parameter. The application of the method to Lake Pinehurst that features both irregular geometry and variable bottom bathymetry yielded modes that in general consisted of small (localized) cells that exhibit their own local rotation and propagation properties. The structure of the modes gave further insight into how wave phenomena would be amplified both in confined narrow bays of lakes and also in shallow regions, as discussed previously.

Chapter 7

Summary and Conclusions

In this thesis, problems concerned with numerical modelling of internal waves in lakes have been explored. In particular, major results have been collected in the areas of: 1) pseudospectral methods for weakly non-hydrostatic layered models; 2) DG-FEM methods for a one-layer weakly non-hydrostatic model; 3) DG-FEM methods for incompressible flow; and 4) the calculation of free modes of oscillation in arbitrary enclosed basins. The major results are now summarized, and possible future work is proposed.

7.1 Pseudospectral methods for weakly non-hydrostatic layered models

7.1.1 Contributions

1. The solution to dispersive shallow water models of the Boussinesq-type using Fourier and Chebyshev pseudospectral methods was considered. The "coupled" and "scalar" approaches were introduced as a means of time-stepping the dispersive terms. The scalar approach reduces the dimension of the resulting linear systems to be solved by a factor of two, and transforms the problem of time-stepping mixed space-time derivatives to a familiar pressure-type elliptic problem. The Fourier method was used to solve the single-layer equations of motion on periodic domains with variable depth. Practical details of implementation were discussed including details of obtaining efficient solutions to the aforementioned linear systems with numerical linear algebra techniques and pre-conditioning, or discrete Fourier transforms where appropriate.

2. A pseudospectral method was presented for solving the one-layer weakly non-hydrostatic rotating shallow water system in annular domains. The methodology employs a Fourier pseudospectral spatial discretization in the azimuthal direction and a Chebyshev pseudospectral spatial discretization in the radial direction, after the model equations have been transformed to polar coordinates (r, θ). The resulting numerical model allowed for simulations in the case of a specialized closed basin and simulations of physical interest to lake-scale phenomena were carried out.

3. The methods applied to the single-layer equations were extended to the case of a weakly non-hydrostatic two-layer fluid under the rigid-lid approximation. The well-known splitting/projection time-stepping method of Karniadakis *et al.* [67] for incompressible flow was applied to evolve the rigid-lid equations forward in time. Two-dimensional simulations of a finite amplitude interfacial perturbation propagating along the wall of a periodic channel were then carried out in Section 3.4.3. In addition to solitary waves, the undular bore was also observed as related nonlinear internal wave phenomenon that can be modelled by the two-layer weakly non-hydrostatic rigid-lid equations.

7.1.2 Future work

A source of ongoing work lies with porting the MATLAB code written for the one- and two-layer pseudospectral solvers to scalable parallel C++ codes. This can be achieved using the parallel SPINS (Spectral Parallel Incompressible Navier–Stokes) code [117, 118] as a framework. The SPINS solver uses a geometric multigrid solver [124] to invert the finite difference pre-conditioner thereby allowing for enhanced scalability over the method presented in Section 3.2.2 where the pre-conditioning operator was factorized in pre-processing and the factors were re-used at each GMRES iteration. Once the codes have been written and validated, higher resolution simulations will become possible via parallel runs on CPU clusters. The main hurdle in writing the codes is that the iterative solver in SPINS will need to be modified to solve elliptic problems with variable coefficients to deal with cases involving variable depth.

7.2 DG-FEM methods for one-layer weakly non-hydrostatic shallow water models

7.2.1 Contributions

1. Comparisons between the pseudospectral methods of Chapter 3 and the Discontinuous Galerkin Finite Element Method (DG-FEM) method were carried out in 1D and 2D in Sections 4.9 and 4.11, respectively. From these comparisons, we conjecture that the DG-FEM can reach comparable resolution and energy-conserving characteristics to the Fourier method for sufficiently high polynomial order N. However, it is more practical to employ DG-FEM for medium order ($N = 4$ to $N = 8$) two-dimensional simulations.

2. DG-FEM simulations in lakes with more general coastlines were considered. It was demonstrated that the DG-FEM is poorly behaved in the neighbourhood of sharp re-entrant corners in Section 4.13, since sharp gradients and spurious eddies appear. The remedy of rounding the corners using curvilinear elements along the boundary was proposed and the implementation was explained in Section 4.14. It was demonstrated that general coastlines need a more computationally expensive treatment than simple coastlines that do not contain re-entrant obstacles since the integrals in the DG-FEM formulations must be evaluated with cubature and quadrature rules of higher order than the approximating basis polynomials.

3. Applications using the curvilinear element methodology were carried out, and it was illustrated that the spurious eddies do not appear when re-entrant corners are represented in a smooth manner. The same methodology was then applied to the real-world situation of Pinehurst Lake, Alberta. The resulting high-resolution numerical solution was able to pinpoint a hot-spot of short length-scale wave activity in the shallow, eastern end of the basin. From this, we concluded that the DG-FEM solution of a weakly non-hydrostatic layered model may be a useful tool in helping to identify regions in lakes where internal wave-induced mixing is most dominant.

7.2.2 Future work

The DG-FEM codes developed in Chapter 4 presume that the elliptic operators can be factorized and stored in memory during pre-processing for re-use at each time-step. To

achieve a scalable code for higher resolutions, a parallelizable iterative solution procedure with appropriate pre-conditioning will be required. Software for addressing this need may be available through the open source PETSc (paralllel extensible toolkit for scientific computing) framework. A collection of Discontinuous Galerkin solvers for hyperbolic and elliptic problems is available in the HEDGE (Hybrid 'n' Easy Discontinuous Galerkin Environment) environment that uses Python as its front-end solver interface and has its core numerical/arithmetic functions implemented in C++ for speed. HEDGE also supports MPI and can be run on both CPU and GPU clusters. Thus, a future version of the solver in chapter 4 may be implemented in Python using HEDGE with elliptic solvers implemented using the PETSc libraries. It is also worth noting that Ferrer [45] has successfully carried out elliptic solves using the PARDISO (parallel direct/iterative solver) from the Intel MKL libraries in his work with the incompressible Navier–Stokes equations, and thus the PARDISO solver may be considered as an alternative to the open-source PETSc.

A second source of future work is the implementation of a wetting/drying treatment. A good candidate treatment is given by the positivity-preserving high-order limiter of Xing *et al.* [133], since many other treatments only function for linear ($N = 1$) elements.

7.3 DG-FEM methods for incompressible flow

7.3.1 Contributions

1. A numerical eigenvalue analysis was carried out to show that the eigenfunctions of the Discontinuous Galerkin pressure projection (PP) operator are themselves not divergence-free. As a result, our numerical simulations have shown that instabilities can occur due to the presence of spurious compressibility artifacts. The prominence of these instabilities appears to be worse in low-order or poorly resolved simulations. It was thought that enforcing incompressibility exactly after each time-step by a local post-processing projection could help improve the stability properties of the method. However, it was shown that under-resolution remains the main driver of instabilities regardless of whether spurious compressibility artifacts are present. Through a series of simulations and discussions, we have reached a set of conjectures about equal-order DG-FEM simulations of inviscid incompressible flow:

   (a) Solutions are very adversely affected by under-resolution and/or poor mesh quality.

(b) The inclusion of source terms further increases the demand on resolution since numerically-driven perturbations can give rise to spurious unbalanced motions.

(c) Long time integrations can be achieved through a combination of using higher-order polynomials ($N = 6$ to $N = 8$) together with appropriate local modal filtering.

2. The source of the difficulties illustrated in this chapter are almost certainly due to the LBB/inf-sup stability condition that usually implies that equal-order element methods for incompressible flow are unstable [24]. The signature of the LBB problem is grid-scale noise that is introduced due to spurious modes in the discretization of the Laplacian. While the SIP-DG formulation guarantees that unphysical modes are suppressed to the small-scale portion of the eigenspectrum of the Laplacian, it does not remove them. Hence, the LBB problem remains an active one. The spurious compressibility artifacts that we have focused on in detail are likely a consequence of this short-coming of the pressure discretization. Removing these artifacts does not guarantee that the spurious pressure modes will not be excited and pollute the solution, in general.

7.3.2 Future work

A natural extension of the incompressible Euler simulations presented in Chapter 5 would be to include viscous terms and carry out simulations using the incompressible Navier–Stokes equations. As discussed above in Section 7.2.2, it will eventually become necessary to implement the numerical solvers in a parallel computing framework where the key difficulty lies with parallelizing the elliptic solves. It would also be interesting to determine what type of wetting/drying treatment could be utilized in the two-layer setup where the model equations become ill-posed in regions of high vertical shear, or in cases where the bottom layer thickness vanishes.

7.4 Calculation of the free modes of oscillation in arbitrary enclosed basins

7.4.1 Contributions

1. The calculation of the free modes of oscillation in arbitrarily shaped, closed basins with rotation was considered, and the modes themselves were studied. Both Helmholtz

and Galerkin decomposition techniques are required to recover a self-adjoint matrix eigenvalue problem for the unknown frequencies and Galerkin projection coefficients of the free modes of oscillation. In Chapter 6, the continuous differential operators have been discretized using the symmetric interior penalty discontinuous Galerkin (SIP-DG) method that was introduced previously in Chapter 4. The SIP-DG method allows for the construction of basis functions from discretized generalized eigenvalue problems that are specified in terms of symmetric matrices.

2. We have found that if a simulation is initialized with a mode (or linear super-position of modes) then the mode remains intact throughout the simulation as long as nonlinear (finite amplitude) effects are negligible. In simulations where nonlinear effects were prevalent, low frequency modes were found to steepen and lead to the excitation of higher-frequency modes. It was also found that there were nonlinearly excited peaks in the power spectra that did not directly match any of the frequencies from linear theory.

3. The method has been used to investigate the effects of rotation and variable bathymetry in a simple model lake. It was found that stronger rotation tends to introduce modes that are trapped along the coast. This observation is consistent with the theoretical predictions for the channel geometry (see 2.4.1 or [96, 103]). Variable bottom bathymetry, on the other hand, modifies gravity modes by enhancing their amplitude in shallow regions thereby creating a general mechanism for wave refraction. Variable bathymetry also causes a potential vorticity gradient in the background state that introduces a corresponding family of rotational waves into the spectrum. These rotational modes are analogous to planetary-scale Rossby waves that exist due to the meridional gradient of the Coriolis parameter.

4. The application of the method to Pinehurst Lake, which features both irregular geometry and variable bottom bathymetry, yielded modes that in general consisted of small (localized) cells that exhibit their own local rotation and propagation properties. The structure of the modes gave further insight into how wave phenomena would be amplified both in confined narrow bays of lakes and also in shallow regions, as discussed previously in Chapter 4. Resonances of narrow bay modes due to interactions with basin-scale Kelvin modes have been investigated using a three-dimensional model by Kawamura *et al.* [69].

7.4.2 Future work

It remains to be determined whether the calculated modes are generated in field situations when driven by wind forcing. This issue can be addressed using simulations with a wind-forcing parameterization or by including forcing terms in the decomposition problem. Further work could also carry out a modal decomposition in the vertical direction. Such a dual decomposition has been explored by Shimizu [106].

The existing software could be properly organized into a Matlab graphical interface where the implemented method could act as a "black box" that simply requires the user to provide the finite element mesh and physical parameters in order to return a full set of linear modes.

One key physical situation that has been overlooked is the case where the pycnocline intersects the bottom boundary. In such cases, modes could be computed by re-meshing based on a new geometry where the boundary is defined by the contour where such an intersection takes place. Other possible directions include calculation of modes of a continuously stratified fluid where, once again, the linear operator is not self-adjoint.

Appendix A

Fully non-hydrostatic dispersion relation for waves at an interface of a two-layer fluid

In this section we derive the fully non-hydrostatic dispersion relation for internal waves in a two-layer flow where the background flow and stratification are constant in each layer and the full vertical structure of the flow is included so that the solution is valid in both shallow and deep water regimes. All rotational effects can be assumed to be at the interface (i.e., the interface is also a vortex sheet), and the solution is thus made analytically tractable since the flow is inviscid, irrotional, and incompressible in each layer. This problem appears as an exercise to the reader in [32].

We begin by assuming we have a two-layer fluid with horizontal walls at $z = -H_2$ and $z = -H_1$ and take the position of the undisturbed interface at $z = 0$. Assume a background flow of the form

$$U(z) = \begin{cases} U_1 \,, & 0 < z < H_1 \,, \\ U_2 \,, & -H_2 < z < 0 \,, \end{cases} \tag{A.1}$$

$$\rho(z) = \begin{cases} \rho_1 \,, & 0 < z < H_1 \,, \\ \rho_2 \,, & -H_2 < z < 0 \,, \end{cases} \tag{A.2}$$

If we presume the flow is inviscid and irrotational within each layer, we can define the velocity potentials ϕ_1 and ϕ_2 such that $(u_1, w_1) = (\phi_{1_x}, \phi_{1_z}) = \nabla\phi_1$ and $(u_2, w_2) = \nabla\phi_2$.

Assuming incompressibility within each layer implies that the velocity potentials are governed by Laplace's equation, i.e.,

$$\nabla^2\phi_1 = 0\,, \quad \text{for} \quad 0 < z < H_1\,, \tag{A.3}$$
$$\nabla^2\phi_2 = 0\,, \quad \text{for} \quad -H_2 < z < 0\,, \tag{A.4}$$

subject to the no normal flow boundary conditions at the walls

$$\phi_{1_z} = 0\,, \quad \text{at} \quad z = H_1\,, \tag{A.5}$$
$$\phi_{2_z} = 0\,, \quad \text{at} \quad z = -H_2\,. \tag{A.6}$$

Assuming finite amplitude effects are negligible, we also have the linearized kinematic boundary conditions,

$$\eta_t + U_1\eta_x - \phi_{1_z} = 0\,, \quad \text{at} \quad z = 0\,, \tag{A.7}$$
$$\eta_t + U_2\eta_x - \phi_{2_z} = 0\,, \quad \text{at} \quad z = 0\,, \tag{A.8}$$

above and below the interface, respectively.

The temporal variation of ϕ is governed by the unsteady Bernoulli equation within each layer

$$\rho_1\left(\phi_{1_t} + U_1\phi_{1_x} + g\eta\right) + p = 0\,, \quad \text{at} \quad z = 0\,, \tag{A.9}$$
$$\rho_2\left(\phi_{2_t} + U_2\phi_{2_x} + g\eta\right) + p = 0\,, \quad \text{at} \quad z = 0\,. \tag{A.10}$$

Since pressure must be continuous across the interface, the above two equations imply that we must have

$$\rho_2\left(\phi_{2_t} + U_2\phi_{2_x}\right) - \rho_1\left(\phi_{1_t} + U_2\phi_{1_x}\right) + (\rho_2 - \rho_1)g\eta = 0\,, \quad \text{at} \quad z = 0\,. \tag{A.11}$$

We look for separable solution of the form

$$\phi_1 = \hat{\phi}_1(z)\mathrm{e}^{i(kx-\sigma t)}\,, \tag{A.12}$$
$$\phi_2 = \hat{\phi}_2(z)\mathrm{e}^{i(kx-\sigma t)}\,, \tag{A.13}$$
$$\eta = C\mathrm{e}^{i(kx-\sigma t)}\,, \tag{A.14}$$

where C is a constant. The vertical structure of the flow can then be solved for since the Laplace equations (A.3) and (A.4) become

$$\hat{\phi}_1'' - k^2\hat{\phi}_1 = 0\,, \tag{A.15}$$
$$\hat{\phi}_2'' - k^2\hat{\phi}_2 = 0\,. \tag{A.16}$$

Solving these equations such that no normal boundary conditions (A.5) and (A.6) gives

$$\hat{\phi}_1 = A\cosh(k(z-H_1)) \ , \quad \hat{\phi}_2 = B\cosh(k(z+H_2)) \ . \tag{A.17}$$

Substituting the separable form for η and the expressions for $\hat{\phi}_1$ and $\hat{\phi}_2$ into the kinematic boundary conditions (A.7) and (A.8) gives

$$-i\sigma C + ikU_1 C - kA\sinh(kH_1) = 0 \ , -i\sigma C + ikU_2 C + kB\sinh(kH_2) = 0 \ . \tag{A.18}$$

Similarly, the dynamic condition (A.11) becomes

$$-i\rho_2(\sigma - U_2 k)B + i\rho_1(\sigma - U_1 k)A + (\rho_2 - \rho_1)gC = 0 \ . \tag{A.19}$$

Equations (A.18)–(A.19) can be written in matrix form as

$$\begin{pmatrix} -k\sinh(kH_1) & 0 & -i(\sigma - U_1 k) \\ 0 & k\sinh(kH_2) & -i(\sigma - U_2 k) \\ i\rho_1(\sigma - U_1 k) & -i\rho_2(\sigma - U_2 k) & (\rho_2 - \rho_1)g \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = 0 \ . \tag{A.20}$$

After taking the determinant equal to zero, solving for σ and simplifying, we have the dispersion relation

$$\begin{aligned} \sigma &= \left(\frac{\rho_1 U_1 \tanh(kH_2) + \rho_2 U_2 \tanh(kH_1)}{\rho_1 \tanh(kH_2) + \rho_2 \tanh(kH_1)} \right) k \\ &\pm \sqrt{\frac{(\rho_2 - \rho_1)gk\tanh(kH_1)\tanh(kH_2)}{\rho_1 \tanh(kH_2) + \rho_2 \tanh(kH_1)} - \frac{\rho_1\rho_2(U_1 - U_2)^2 k^2 \tanh(kH_1)\tanh(kH_2)}{(\rho_1 \tanh(kH_2) + \rho_2 \tanh(kH_1))^2}} \ . \end{aligned} \tag{A.21}$$

In the case of no background flow $U_1 = U_2 = 0$ we find

$$\sigma = \pm\sqrt{\frac{(\rho_2 - \rho_1)gk\tanh(kH_1)\tanh(kH_2)}{\rho_1 \tanh(kH_2) + \rho_2 \tanh(kH_1)}} \ . \tag{A.22}$$

Bibliography

[1] A.S. Almgren, J.B. Bell, and W.Y. Crutchfield. Approximate projection methods: Part I. Inviscid analysis. *SIAM J. Sci. Comput.*, 22(4):1139–1159, 2000.

[2] J. Amezcua, E. Kalnay, and P.D. Williams. The Effects of the RAW Filter on the Climatology and Forecast Skill of the SPEEDY Model. *Mon. Weather Rev.*, 139:608–619, 2011.

[3] J.P. Antenucci and J. Imberger. Energetics of long internal gravity waves in large lakes. *Limnol. Oceanogr.*, 46:1760–1773, 2001.

[4] A. Arakawa and V.R. Lamb. Computational design of the basic dynamical processes of the UCLA general circulation model. *Meth. Comput. Phys.*, 17:173–265, 1977.

[5] A. Baglaenko, M. Stastna, D. Steinmoeller, and F.J. Poulin. The effects of the spatial distribution of bottom topography and bottom drag on seiche-induced wave-train formation. *Water. Qual. Res. J. Can.*, 2012, in press.

[6] Y. Bai and J. Song. Modification on Lynett and Liu's model for internal solitary wave propagation over variable bathymetry. *Chin. J. Oceanol. and Limnol.*, 24:338–344, 2006.

[7] N.J. Balmforth and S. Mandre. Dynamics of roll waves. *J. Fluid Mech.*, 514:1–33, 2004.

[8] F. Bassi, A. Crivellini, D. A. Di Pietro, and S. Rebay. An artificial compressibility flux for the discontinuous Galerkin solution of the incompressible Navier–Stokes equations. *J. Comp. Phys*, 218(2):794–815, 2006.

[9] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comp. Phys.*, 131:267–279, 1997.

[10] Francesco Bassi, Andrea Crivellini, Daniele A. Di Pietro, and Stefano Rebay. An implicit high-order discontinuous Galerkin method for steady and unsteady incompressible flows. *Comput. Fluids*, 36(10):1529–1546, 2007.

[11] S. Blaise, R. Comblen, V. Legat, J.-F. Remacle, E. Deleersnijder, and J. Lambrechts. A discontinuous finite element baroclinic marine model on unstructured prismatic meshes. part I: space discretization. *Ocean Dynamics*, 60:1371–1393, 2010.

[12] J. Boussinesq. Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *Journal de Mathématique Pures et Apliquée, Deuxième Série*, **17**:55–108, 1872.

[13] J. P. Boyd and F. Yu. Comparing seven spectral methods for interpolation and for solving the Poisson equation in a disk: Zernike polynomials, Logan–Shepp ridge polynomials, Chebyshev–Fourier series, cylindrical Robert functions, Bessel–Fourier expansions, square-to-disk conformal mapping and radial basis functions. *J. Comput. Phys.*, 230(4):1408–1438, 2011.

[14] J.P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover Publications, 2nd edition, 2001.

[15] P. Brandt, A. Rubino, W. Alpers, and J. Backhaus. Internal waves in the Strait of Messina studied by a numerical model and synthetic aperture radar umages from *ERS 1/2* Satellites. *J. Phys. Oceangr.*, **27**:648–663, 1997.

[16] S. Bunya, E.J. Kubatko, J.J. Westerink, and C. Dawson. A wetting and drying treatment for the Runge-Kutta discontinuous Galerkin solution to the shallow water equations. *Comput. Methods Appl. Mech. Engrg.*, **198**:1548–1562, 2009.

[17] C.I. Christov. An energy-consistent dispersive shallow-water model. *Wave Motion*, **1018**:1–14, 2000.

[18] B Cockburn, SC Hou, and CW Shu. The Runge–Kutta local projection discontinuous Galerkin finite-element method for conservation-laws. 4. The multidimensional case. *Math. Comp*, 54(190):545–581, 1990.

[19] B. Cockburn, G. Kanschat, and D. Schötzau. An equal-order DG method for the incompressible Navier–Stokes equations. *J. Sci. Comput.*, 40:188–210, 2009.

[20] B Cockburn and CW Shu. TVB Runge–Kutta local projection discontinuous Galerkin finite-element method for conservation-laws. 2. General framework. *Math. Comp.*, 52(186):411–435, 1989.

[21] R. Comblen, S. Blaise, V. Legat, J.-F. Remacle, E. Deleersnijder, and J. Lambrechts. A discontinuous finite element baroclinic marine model on unstructured prismatic meshes. Part II: Implicit/explicit time discretization. *Ocean Dynamics*, 60:1395–1414, 2010.

[22] R. Cossu and M.G. Wells. The interaction of large amplitude internal seiches with a shallow sloping lakebed: Observations of benthic turbulence in lake simcoe, ontario, canada. *PLoS ONE*, 2013.

[23] C. J. Cotter, D. D. Holm, and J. R. Percival. The square root depth wave equations. *Proc. R. Soc. A*, 466:3621–3633, 2010.

[24] Colin J. Cotter, David A. Ham, Christopher C. Pain, and Sebastian Reich. LBB stability of a mixed Galerkin finite element pair for fluid flow simulations. *J. Comp. Phys.*, 228(2):336–348, 2009.

[25] G.T. Csanady. Large–Scale Motion in the Great Lakes. *J. Geophys. Res.*, 72(16):4151–4162, 1967.

[26] B. Cushman-Roisin and J.-M. Beckers. *Introduction to Geophysical Fluid Dynamics*. Academic Press, 2nd edition, 2011.

[27] P.A. Davidson. *Turbulence: An Introduction for Scientists and Engineers*. Oxford University Press, 1st edition, 2005.

[28] A. de la Fuente, K. Shimizu, J. Imberger, and Y. Niño. The evolution of internal waves in a rotating, stratified, circular basin and the influence of weakly nonlinear and nonhydrostatic accelerations. *Limnol. Oceanogr.*, 53(6):2738–2748, 2008.

[29] S. Debsarma, K.P. Das, and J.T. Kirby. Fully nonlinear higher-order model equations for long internal waves in a two-fluid system. *J. Fluid Mech.*, 654:281–303, 2010.

[30] P.J. Diamessis, J.A. Domaradzki, and J.S. Hesthaven. A spectral multidomain penalty method model for the simulation of high Reynolds number localized incompressible stratified turbulence. *J. Comp. Phys.*, 202:298–322, 2005.

[31] D.E. Dietrich. Application of a modified 'A' grid ocean model having reduced numerical dispersion to the Gulf of Mexico circulation. *Dyn. Atmos. Ocean*, 27:201–217, 1997.

[32] P.G. Drazin and W.H. Reid. *Hydrodynamic Stability*. Cambridge University Press, 2nd edition, 2004.

[33] M. Dunphy, C. Subich, and M. Stastna. Spectral methods for internal waves: indistinguishable density profiles and double-humped solitary waves. *Nonlin. Proc. Geoph.*, 18(3):351–358, 2011.

[34] Frédéric Dupont. *Comparison of numerical methods for modelling ocean circulation in basins with irregular coasts*. PhD thesis, McGill University, 2001.

[35] D.R. Durran. *Numerical Methods for Fluid Dynamics*. Springer, 2nd edition, 2010.

[36] K. Hutter (editor). *Nonlinear Internal Waves in Lakes (Advances in Geophysical and Environmental Mechanics and Mathematics)*. Springer, 2012.

[37] A. Engsig-Karup, J. Hesthaven, H. Bingham, and P. Madsen. Nodal DG-FEM solution of high-order Boussinesq-type equations. *J. Eng. Math.*, **56**:351–370, 2006.

[38] J.A. Escobar-Vargas. *A spectral multidomain penalty method solver for environmental flow processes*. PhD thesis, Cornell University, Ithica, NY, USA, 2012.

[39] J.A. Escobar-Vargas, P.J. Diamessis, and F.X. Giraldo. High-order discontinuous element-based schemes for the inviscid shallow water equations: Spectral multidomain penalty and discontinuous Galerkin methods. *Appl. Math. and Comput.*, 218:4825–4848, 2012.

[40] J.A. Escobar-Vargas, P.J. Diamessis, and T. Sakai. A spectral quadrilateral multidomain penalty method model for high-Reynolds number incompressible stratified flows. *J. Comp. Phys. (submitted)*, 2012.

[41] J.A. Escobar-Vargas, P.J. Diamessis, and C.F. Van Loan. The numerical solution of the pressure Poisson equation for the incompressible Navier–Stokes equations using a quadrilateral spectral multidomain penalty method. *J. Comp. Phys. (submitted)*, 2011.

[42] C. Eskilsson and S. Sherwin. Spectral/*hp* discontinuous Galerkin methods for modelling 2D Boussinesq equations. *J. Sci. Comp.*, **22**:269–288, 2005.

[43] J.G. Esler, O.J. Rump, and E.R. Johnson. Non-dispersive and weakly dispersive single-layer flow over an axisymmetric obstacle: the equivalent aerofoil formulation. *J. Fluid Mech.*, 574:209–237, 2007.

[44] A. Adcroft et. al. *MITgcm User Manual*. Massachusetts Institue of Technology, 2009.

[45] E. Ferrer. *A high order Discontinuous Galerkin - Fourier incompressible 3D Navier-Stokes solver with rotating sliding meshes for simulating cross-flow turbines*. PhD thesis, University of Oxford, UK, 2012.

[46] E. Ferrer and R.H.J. Willden. A high order Discontinuous Galerkin Finite Element solver for the incompressible Navier–Stokes equations. *Comput. Fluids*, 46:224–230, 2011.

[47] C. Geuzaine and J.-F. Remacle. A three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *Int. J. Numer. Methods Eng.*, 79:1309–1331, 2009.

[48] A. Gill. *Atmosphere-Ocean Dynamics*. Academic Press, 1st edition, 1982.

[49] F. X. Giraldo and M. Restelli. A study of spectral element and discontinuous Galerkin methods for the Navier-Stokes equations in nonhydrostatic mesoscale atmospheric modeling: Equation sets and test cases. *J. Comp. Phys.*, 227(8):3849–3877, 2008.

[50] F. X. Giraldo, M. Restelli, and M. Laeuter. Semi-implicit formulations of the Navier–Stokes equations: Application to nonhydrostatic atmospheric modeling. *SIAM J. Sci. Comp.*, 32(6):3394–3425, 2010.

[51] F.X. Giraldo. The lagrange-galerkin spectral element method on unstructured quadrilateral grids. *J. Comp. Phys.*, 147:114–146, 1998.

[52] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.

[53] W.N. Gordon and C.A. Hall. Construction of curvilinear coordinate systems and application to mesh generation. *Int. J. Numer. Meth. Eng.*, 7:461–477, 1973.

[54] W.J.F. Govaerts. *Numerical methods for bifurcations of dynamical equilibria*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

[55] R.H.J. Grimshaw and N. Smyth. Resonant flow of a stratified fluid over topography. *J. Fluid Mech.*, 169:429–464, 1986.

[56] A.V. Gurevich and L.P. Pitaevskii. Nonstationary structure of a collisionless shock wave. *Sov. Phys. JETP*, 38:291–297, 1974.

[57] P.F. Hamblin. On the free surface oscillations of Lake Ontario. *Limnol. Oceanogr.*, 27(6):1039–1049, 1982.

[58] E. Hanert, R.A. Walters, D.Y. Le Roux, and J. Pietrzak. A tale of two elements: P1ncP1 and RT0. *Ocean Modelling*, 28:24–33, 2009.

[59] K.R. Helfrich. Decay and return of internal solitary waves with rotation. *Phys. Fluids*, 19:026601, 2007.

[60] J. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods*. Springer, 2008.

[61] B. Hodges and C. Dallimore. *Estuary, Lake and Coastal Ocean Model: ELCOM*. Centre for Water Research, University of Western Australia, 2.2 edition, 2005.

[62] T.J. Hughes. Recent progress in the development and understanding of SUPG methods with special reference to the compressible euler and navier-stokes equations. *Int. J. Numer. Method. Fluids*, 7:1261–1275, 1987.

[63] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*. Cambridge University Press, 1996.

[64] M Iskandarani, DB Haidvogel, and JP Boyd. A staggered spectral element model with application to the oceanic shallow-water equations. *Int. J. Numer. Meth. Fluids*, 20(5):393–414, 1995.

[65] M Iskandarani, DB Haidvogel, and JC Levin. A three-dimensional spectral element model for the solution of the hydrostatic primitive equations. *J. Comp. Phys.*, 186(2):397–425, 2003.

[66] T. Kärnä, V. Legat, and E. Deleersnijder. A baroclinic discontinuous Galerkin finite element model for coastal flows. *Ocean Modelling*, 61:1–20, 2013.

[67] G.E. Karniadakis, M. Israeli, and S.A. Orszag. High-order splitting methods for the incompressible Navier–Stokes equations. *J. Comp. Phys.*, 97:414–443, 1991.

[68] G.E. Karniadakis and S.J. Sherwin. *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford University Press, USA, 2nd edition, 2005.

[69] Y. Kawamura, Y. Kitade, and M. Matsuyama. Scattering of semidiurnal internal Kelvin wave at step bottom topography. *J. Oceanogr.*, 61:59–68, 2005.

[70] PK Kundu and IM Cohen. *Fluid Mechanics*. Elsevier Academic Press, 4th edition, 2008.

[71] H. Lamb. *Hydrodynamics*. Dover Publications, 6th edition, 1932.

[72] K.G. Lamb and V.T. Nguyen. Calculating energy flux in internal solitary waves with an application to reflectance. *J. Phys. Oceanogr.*, 39:559–580, 2009.

[73] R. B. Lehoucq, D. C. Sorensen, and C. Yang. ARPACK users' guide: Solution of large scale eigenvalue problems with implicitly restarted Arnoldi methods, Oct 1997.

[74] U. Lemmin, C.H. Mortimer, and E. Bauerle. Internal seiche dynamics in Lake Geneva. *Limnol. Oceanogr.*, 501(1):207–216, 2005.

[75] R.J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.

[76] R.J. Leveque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*. Society for Industrial and Applied Mathematics, 2007.

[77] J. C. Levin, M. Iskandarani, and D. B. Haidvogel. To continue or discontinue: Comparisons of continuous and discontinuous Galerkin formulations in a spectral element ocean model. *Ocean Modelling*, 15(1-2):56–70, 2006.

[78] J.-G. Liu and C.-W. Shu. A high-order discontinuous Galerkin method for 2D incompressible flows. *J. Comp. Phys.*, 160:577–596, 2000.

[79] R.R. Long. Long waves in a two-fluid system. *J. Meteorol.*, 13:70–74, 1955.

[80] P. Lynett and P. L.-F. Liu. A two-layer approach to wave modelling. *Proc. R. Soc. Lond. A*, 460:2637–2669, 2004.

[81] P.J. Lynett and P.L.-F. Liu. A two-dimensional, depth-integrated model for internal wave propagation over variable bathymetry. *Wave Motion*, 36:221–240, 2002.

[82] P.A. Madsen, H.B. Bingham, and H. Liu. A new Boussinesq method for fully nonlinear waves from shallow to deep water. *J. Fluid Mech.*, **462**:1–30, 2002.

[83] P.A. Madsen, R. Murray, and O.R. Sørensen. A new form of the Boussinesq equations with improved linear dispersion characteristics. *Coastal Engineering*, **15**:371–388, 1991.

[84] P.A. Madsen and O.R. Sørensen. A new form of the Boussinesq equations with improved linear dispersion characteristics. Part 2. A slowly varying bathymetry. *Coastal Engineering*, **18**:183–204, 1992.

[85] K.T. Mandli. *Finite Volume Methods for the Multilayer Shallow Water Equations with Applications to Storm Surges*. PhD thesis, University of Washington, Seattle, Washington, USA, 2011.

[86] J.C. McWilliams. *Fundamentals of Geophysical Fluid Dynamics*. Cambridge University Press, 2011.

[87] C. Mirabito, C. Dawson, E.J. Kubatko, J.J. Westerink, and S. Bunya. Implementation of a discontinuous Galerkin morphological model on two-dimensional unstructured meshes. *Comput. Methods Appl. Mech. Engrg.*, 200:189–207, 2011.

[88] M. Nazarov and B. Popov. A posteriori error estimation for the compressible euler equations using entropy viscosity. *Comp. & Fluids*, 2014 (submitted).

[89] N.C. Nguyen and J. Peraire. Hybridizable discontinuous Galerkin method for partial differential equations in continuum mechanics. *J. Comp. Phys.*, 231:5955–5988, 2012.

[90] O. Nwogu. Alternative form of Boussinesq equations for nearshore wave propagation. *J. Waterw. Port Coast. Ocean Eng.*, **119**(6):618–638, 1993.

[91] A. Pannard, B.E. Beisner, D.F. Bird, J. Braun, D. Planas, and M. Bormans. Recurrent internal waves in a small lake: Potential ecological consequences for metalimnetic phytoplankton populations. *L&O:F&E*, 1:91–109, 2011.

[92] J. Pedlosky. *Ocean Circulation Theory*. Springer, 2nd edition, 2004.

[93] D.H. Peregrine. Long waves on a beach. *J. Fluid. Mech.*, **27**(4):815–827, 1967.

[94] D.H. Peregrine. *Waves on Beaches and Resulting Sediment Transport*, pages 95–122. Academic Press, 1972.

[95] R. Peyret. *Spectral Methods for Incompressible Viscous Flow*. Springer-Verlag New York, Inc., 2002.

[96] G.W. Platzman and D.B. Rao. Spectra of Lake Erie water levels. *J. Geophys. Res.*, 69(12):2525–2535, 1964.

[97] M. Preusse, M. Stastna, H. Freistühler, and F. Peeters. Intrinsic breaking of internal solitary waves in a deep lake. *PLoS One*, 7(7), 2012.

[98] D.B. Rao. *Free gravitational oscillations in rotating rectangular basins*. PhD thesis, University of Chicago, Dept. Geophys. Sci., Chicago, USA, 1965.

[99] D.B. Rao, C.H. Mortimer, and D.J. Schwab. Surface normal modes of Lake-Michigan – calculations compared with spectra of observed water level fluctuations. *J. Phys. Oceanogr.*, 6(4):575–588, 1976.

[100] D.B. Rao and D.J. Schwab. 2 dimensional normal modes in arbitrary enclosed basins on a rotating Earth - application to lakes Ontario and Superior. *Philos. Trans. Roy. Soc. A: Math. Phys. Eng. Sci.*, 281(1299):63–96, 1976.

[101] M. Restelli and F.X. Giraldo. A conservative semi-implicit discontinuous Galerkin method for the Navier–Stokes equations in nonhydrostatic mesoscale atmospheric modeling. *SIAM J. Sci. Comp.*, 31:2231–2257, 2009.

[102] S. Rhebergen and B. Cockburn. A space–time hybridizable discontinuous Galerkin method for incompressible flows on deforming domains. *J. Comp. Phys.*, 231:4185–4204, 2012.

[103] D.C. Rockwell. Theoretical free oscillations of the Great Lakes. *Proc. 9th Conf. Great Lakes Res., University of Michigan, Great Lakes Res. Div., Publ.*, 15:352–368, 1966.

[104] A. Saggio and J. Imberger. Internal wave weather in a stratified lake. *Limnol. Oceanogr.*, 43:1780–1795, 2001.

[105] K. Shahbazi, P.F. Fischer, and C.R. Ethier. A high-order discontinuous Galerkin method for the unsteady incompressible Navier–Stokes equations. *J. Comp. Phys*, 222:391–407, 2007.

[106] K. Shimizu. A theory of vertical modes in multilayer stratified fluids. *J. Phys. Oceanogr.*, 41:1694–1707, 2011.

[107] C.W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comp. Phys.*, 77:439–471, 1988.

[108] B. Smith, P. Bjorstad, and W. Gropp. *Domain Decomposition: Parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press, 2004.

[109] M. Stastna, F.J. Poulin, K.L. Rowe, and C. Subich. On fully nonlinear vertically trapped wave packets. *Phys. Fluids*, 21:106604, 2009.

[110] D.T. Steinmoeller, M. Stastna, and K.G. Lamb. Fourier pseudospectral methods for 2D Boussinesq-type equations. *Ocean Modelling*, 52–53:76–89, 2012.

[111] D.T. Steinmoeller, M. Stastna, and K.G. Lamb. Pseudospectral methods for Boussinesq-type equations in an annular domain with applications to mid-sized lakes. *J. Comp. Sci.*, 4:3–11, 2013.

[112] D.T. Steinmoeller, M. Stastna, and K.G. Lamb. A short note on the discontinous Galerkin discretization of the pressure projection operator in incompressible flow. *J. Comp. Phys*, 251C:480–486, 2013.

[113] A. L. Stewart and P. J. Dellar. Multilayer shallow water equations with complete Coriolis force. Part I. Derivation on a non-traditional beta-plane. *J. Fluid Mech.*, 651:387–413, 2010.

[114] A. L. Stewart and P. J. Dellar. Multilayer shallow water equations with complete coriolis force. Part 2. Linear plane waves. *J. Fluid Mech.*, 690:16–50, 2012.

[115] R. Stocker and J. Imberger. Energy partitioning and horizontal dispersion in a stratified rotating lake. *J. Phys. Oceanogr.*, 33:512–529, 2003.

[116] T. Stocker and K. Hutter. One-dimensional models for topographic Rossby waves in elongated basins on the f-plane. *J. Fluid Mech.*, 170:435–459, 1986.

[117] Christopher Subich. *Simulation of the Navier–Stokes Equations in Three Dimensions with a Spectral Collocation Method*. PhD thesis, University of Waterloo, 2011.

[118] C.J. Subich, K.G. Lamb, and M. Stastna. Simulation of the Navier–Stokes equations in three dimensions with a spectral collocation method. *Int. J. Numer. Meth. Fluids*, 73:103–129, 2013.

[119] W. Thomson. On gravitational oscillations of rotating water. *P. Roy. Soc. Edinb.*, **82**:92–100, 1872.

[120] G. Tomasson and W. Melville. Geostrophic adjustment in a channel: nonlinear and dispersive effects. *J. Fluid Mech.*, 241:23–48, 1992.

[121] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, 2nd edition, 1999.

[122] L.N. Trefethen. *Spectral Methods in MATLAB*. Society for Industrial and Applied Mathematics, 2000.

[123] L.N. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1997.

[124] U. Trottenberg, C.W. Oosterlee, and A. Schuller. *Multigrid*. Academic Press, 1st edition, 2000.

[125] S. Vitousek and O.B. Fringer. Physical vs. numerical dispersion in nonhydrostatic ocean modeling. *Ocean Modelling*, 40:72–86, 2011.

[126] M.A. Walkley. *A numerical method for extended Boussinesq shallow-water wave equations*. PhD thesis, University of Leeds, UK, 1999.

[127] S. Wandzura and H. Xiao. Symmetric quadrature rules on a triangle. *Comput. Math. Applica.*, 45:1829–1840, 2003.

[128] T. Warburton. An explicit construction of interpolation nodes on the simplex. *J. Eng. Math*, **56**:247–262, 2006.

[129] G. Wei and J.T. Kirby. Time-dependent numerical code for extended Boussinesq equations. *J. Waterw. Port Coast. Ocean Eng.*, **121**:251–261, 1995.

[130] G.B. Whitham. *Linear and Nonlinear Waves*. Wiley-Interscience, 1999.

[131] P.D. Williams. The RAW Filter: An Improvement ot the Robert–Asselin Filter in Semi-Implicit Integrations. *Mon. Weather Rev.*, 139:1996–2007, 2011.

[132] A. Wüest and A. Lorke. Small-scale hydrodynamics in lakes. *Annu. Rev. Fluid Mech.*, 35:373–412, 2003.

[133] Y. Xing, X. Zhang, and C.-W. Shu. Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations. *Adv. Water. Res.*, 33:1476–1493, 2010.

[134] F. Ye. *Derivation of a two-layer non-hydrostatic shallow water model.* PhD thesis, University of Hawaii, USA, 1995.

[135] M. Yu, F.X. Giraldo, M. Peng, and Z.J. Wang. Localized artificial viscosity stabilization of discontinuous galerkin methods for nonhydrostatic atmospheric modeling. *J. Comp. Phys.*, 2014 (submitted).

[136] X. Zhang, Y. Xia, and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high order Discontinuous Galerkin schemes for conservation laws on triangular meshes. *J. Sci. Comput.*, 50:29–62, 2012.