VIRGILE ROSTAND

ANALYSIS OF DISCRETE FINITE ELEMENT SHALLOW-WATER MODELS

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de doctorat en mathématiques
pour l'obtention du grade de philosophiea doctor (Ph.D.)

FACULTÉ DES SCIENCES ET GÉNIE
UNIVERSITÉ LAVAL

NOVEMBRE 2007

RÉSUMÉ

Les équations de Saint-Venant sont un système aux dérivées partielles jouant un rôle central dans la modélisation des écoulements océaniques. La méthode des éléments finis est particulièrement adaptée pour résoudre les équations de Saint-Venant car elle offre une grande flexibilité sur les domaines irréguliers ainsi qu'une variété d'espaces pour l'approximation de la solution. Or, la qualité de la solution numérique dépend de l'interaction entre ces espaces. Pour certaines combinaisons ou paires d'élément finis la solution numérique peut présenter des oscillations articiellement introduites par la discrétisation. Cette thèse porte sur le comportement numérique des solutions aux équations de Saint-Venant obtenues par différentes paires d'éléments finis. Tout d'abord, une étude sur la dispersion des ondes d'inertie-gravité est présentée pour une sélection de neuf paires d'éléments finis. Un ensemble de trois propriétés est ensuite mis en évidence afin que la discrétisation respecte le comportement des équations analytiques. Une méthode basée sur le calcul des noyaux est utilisée pour caractériser les modes stationnaires correspondant aux écoulements géostrophiques. Finalement, les espaces vectoriels de Raviart-Thomas et Brezzi-Douglas-Marini sont analysés.

ABSTRACT

The shallow-water equations system plays a central role in numerical oceanic models. The finite element method is particularly well suited to solve the shallow-water equations as it works on irregular meshes with a variety of approximation spaces. However, the behavior of the numerical solution highly depends on the interaction between these approximation spaces. For specific finite element pairs the solution may exhibit spurious oscillations induced by the discretization scheme. In this thesis, we analyze these oscillations for a wide selection of finite element pairs. The numerical dispersion of inertia-gravity waves is quantified with dispersion analyses. A constructive linear algebra approach is developed to compute the kernels of the discretized operators. The results are used to characterize the smallest representable vortices on both structured and unstructured meshes. A special attention is given to the Raviart-Thomas and Brezzi-Douglas-Marini approximation spaces.

PREFACE

This thesis is divided into three main chapters and each one corresponds to an article published or submitted in a peer reviewed journal. The contribution of the candidate for each article is specified here:

Chapter 1: The candidate computed the analytical dispersion relations with the MAPLE software. He brought the idea of the canal simulation to test inertia-gravity waves dispersion. He produced all the figures except those issue from the numerical simulations. D. Y. Le Roux, the thesis supervisor is the primary author and B. Pouliot is a master degree student. This paper was published in SIAM Journal on Scientific Computing (SISC) in 2007.

Chapter 2: The candidate highlighted a set of three kernel relations necessary for the discretized system to share the same stationary properties as the continuous system. He developed a complete finite element code in MATLAB to test these relations. He produced almost all the figures and the text under the supervision of D. Y. Le Roux. Professor G. F. Carey is the CFDlab director at the University of Texas at Austin. The candidate is the primary author and this paper is submitted to SISC.

Chapter 3: The candidate computed the analytical dispersion relations with the MAPLE software and he implemented the numerical code in FORTRAN. He produced all the figures along with the text under the supervision of D. Y. Le Roux. The candidate is the primary author and this paper is submitted to International Journal for Numerical Methods in Fluids (IJNMF).

REMERCIEMENTS

J'aimerais remercier l'ensemble des gens qui m'ont aidé de près ou de loin à compléter cette thèse. En particulier et non exclusivement:

Mon directeur de thèse, Daniel Le Roux, pour ses conseils, son attention et son intégrité.

Les professeurs du département de mathématiques et statistique de l'Université Laval.

Le FQRNT pour son soutient financier.

Mes amis, ma famille et ma conjointe Karine.

UN BIEN GROS MERCI :-)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

0. INTRODUCTION

The World Ocean is a body of saline water covering approximately 70% of the Earth surface and constituting 96.5% of the hydrosphere. The ocean holds powerful currents carrying a huge amount of heat all over the globe [46]. These currents influence the global climate and a small change in their behavior can have dreadful consequences on the weather, sea level, and marine ecosystems. Experimental data about the oceanic currents is generally limited due to the vastness and deepness of the ocean and so a numerical model is needed to analyse and simulate the oceanic circulation.

The ocean circulation is driven by the water temperature and salinity, the gravity force, the surface winds, and the Coriolis force (trajectory deflection induced by the Earth rotation). The direction and amplitude of the flow vary with the depth. The vertical motion of the water is mainly induced by its density. Cold and salty water is heavier than fresh and warm water and so differences in the water temperature and salinity trigger mixing and upwelling. On the other hand, horizontal flows are mainly affected by the gravity and Coriolis forces. The water flow is said to be in geostrophic equilibrium when the circular motion balances with the surface elevation, e. g. in the Northern Atlantic the water flows in a clockwise circular motion that pushes the water toward the center making the surface elevation higher in the center of the ocean than on the coastline. Most of the flow in the ocean is in geostrophic equilibrium.

The average oceanic depth is approximately 3,790 meters but the oceanic reach is in thousand of kilometers. So the ocean is in proportion a relatively thin fluid laying on a rotating sphere. Consequently, a simplified 2D model of the flow can be obtained under reasonable approximations based on this specific shape of the fluid. The resulting model leads to the shallow-water equations (SWE).

0.1 Shallow-water equations

The SWE are usually derived from the 3D Navier-Stokes equations under general physical approximations. Two particularly important approximations are the Boussinesq and the hydrostatic approximations. The first one restricts the effects of the density to terms where the gravitational force is involved. The second one states that the pressure should vary linearly with the depth. The simplified equations are integrated in the vertical direction and the resulting system yields to the 2D SWE. The physical intuition along with the mathematical details behind this procedure can be found in several textbooks [37, 43, 21, 63, 30].

Although the SWE are a 2D system, they are playing a central role in several 3D oceanic circulation models (for example in QUODDY [40], FEOM [22], and SLIM [69]). These models are made up of layers corresponding to SWE systems with extra terms for the flow in the vertical direction. The computational cost of 3D models is generally high and therefore these models will benefit from an efficient numerical scheme to solve the SWE.

The solutions to the SWE equations can be separated into two distinct categories. The first one is constituted of inertia-gravity waves. These solutions are fast propagating waves and they are mainly driven by the gravity force, e. g. tsunamis are inertia-gravity waves. The second category contains vortex flows where the Coriolis force balances with the gravity force. These flows are in geostrophic equilibrium and correspond to Rossby waves. The behavior of the waves belonging to these two distinct groups largely differs. Therefore different techniques and approaches have to be set to study both inertia-gravity waves and Rossby waves.

0.2 Finite element method

The finite element method [49, 18] is an attractive numerical method to compute the solution to the SWE. Finite elements are generally triangular or quadrangular, but triangles are particularly well suited to approach irregular domains boundary like coastlines. The theoretical framework around the finite element method is well established and it provides error estimates that can be used to locally refine the mesh in order to increase the precision of the solution at a reasonable computational cost. The approximation spaces can be chosen in various ways and they offer a great flexibility to compute the solution.

Usually, the finite element spaces contain low order polynomials defined piecewise by element. The piecewise continuous polynomials of the first and second order are typical spaces and denoted by P_1 and P_2, respectively. The continuity restrictions between elements can also be weaken to obtain semi-continuous or discontinuous spaces like the P_1^{NC} and P_0 spaces. The discretization of the SWE with the finite element method involves choosing two finite element spaces; one for the velocity field and another one for the surface-elevation. The coupling between these two spaces constitutes a finite element pair. We use the notation $P_n - P_m$ to designate the pair whose velocity field and surface-elevation belong to the P_n and P_m spaces, respectively.

The numerical behavior of finite element pairs highly depends on the coupling between the velocity and surface-elevation spaces. The SWE are of hyperbolic type (opposed to elliptic) and so the famous inf-sup stability criterion from Ladyzhenskaya, Brezzi and Babuska (LBB) [14] does not apply for SWE. So different tools are required to analyze the finite element pairs used to solve these equations. In this thesis, we use a variety of methods to characterize and quantify the numerical behavior of twelve finite element pairs.

0.3 Overview

Chapter 1: We present a quantitative analysis of numerically induced dispersion on inertia-gravity waves for a selection of nine finite element pairs. We consider a regular mesh made up of right biased triangles and we assemble the stencils for the discrete SWE. The stencils are then used to compute the dispersion relations and the discrete frequencies of inertia-gravity waves. The results are compared analytically and graphically.

Chapter 2: A linear algebra approach is developed to characterize the kernel of the discrete SWE. Three kernel relations are identified as necessary conditions for the discretized system to share the same properties as the continuous system. This matrix kernel scheme is implemented using MATLAB and applied to ascertain the presence, number, and structure of spurious modes arising in typical finite difference and finite element schemes. The kernel concept is then used to characterize the smallest representable vortices. Both uniform and unstructured mesh situations are considered and compared. Issues such as modal decoupling in vortex modes are also examined.

Chapter 3: An analysis of the discrete SWE using the Raviart-Thomas and Brezzi-Douglas-Marini finite elements is presented. For inertia-gravity waves, the discretization schemes are developed in details to quantify the level of numerical dispersion. Two meshes made up of equilateral and right biased triangles are considered. A linear algebra approach is used to ascertain the presence and the form of spurious modes arising from the discretization. The geostrophic balance is examined and the smallest representable vortices are characterised on both structured and unstructured meshes. Relevant numerical tests are presented.

1. ANALYSIS OF NUMERICALLY INDUCED OSCILLATIONS IN 2D FINITE ELEMENT SHALLOW-WATER MODELS PART I: INERTIA-GRAVITY WAVES

D. Y. Le Roux, V. Rostand, and B. Pouliot.
Published in SIAM Journal on Scientific Computing 29, 331 (2007).
Keywords: shallow-water equations, finite-element method, dispersion analysis, surface waves.

1.1 Résumé

L'approximation numérique des solutions aux équations de Saint-Veant est un problème délicat. Pour un bon nombre de schémas de discrétisation numérique, le couplage entre l'équation des moments et l'équation de continuité produit des oscillations non physiques dans la solution. Cet article présente une analyse de dispersion utilisée pour déterminer la présence et la forme de ces oscillations se produisant dans les schémas numériques basés sur la méthode de Galerkin et des éléments finis mixtes. Neuf paires d'éléments finis bien connues sont considérées dans le calcul des relations de dispersion. Ces relations sont ensuite comparées avec la solution exacte analytiquement et graphiquement. Pour certaines paires d'élément finis, il est montré que l'erreur numérique sur le calcul de la vitesse de phase ainsi que la vitesse de groupe pour les ondes rapides est significative et parfois même non physique. Les paires $P_1^{NC} - P_1$ et RT_0 sont identifiées comme étant des choix intéressants pourvu que le la résolution de la grille soit grande lorsque comparée au rayon de déformation de Rossby pour la paire RT_0. Finalement, les solutions de deux simulations numériques sont présentées et corroborent les résultats analytiques.

1.2 Abstract

The numerical approximation of shallow-water models is a delicate problem. For most of the discretization schemes, the coupling between the momentum and the continuity equations usually leads to anomalous dispersion in the representation of fast waves. A dispersion relation analysis is employed here to ascertain the presence and determine the form of spurious modes as well as the dispersive nature of the finite-element Galerkin mixed formulation of the two-dimensional linearized shallow-water equations. Nine popular finite-element pairs are considered using a variety of mixed interpolation schemes. For each pair the frequency or dispersion relation is obtained and analyzed, and the dispersion properties are compared analytically and graphically with the continuous case. It is shown that certain choices of mixed interpolation schemes may lead to significant phase and group velocity errors and spurious solutions in the calculation of fast waves. The $P_1^{NC} - P_1$ and RT_0 pairs are identified as a promising compromise, provided the grid resolution is high relative to the Rossby radius of deformation for the RT_0 element. The numerical solutions of two test problems to simulate fast waves are in good agreement with the analytical results.

1.3 Introduction

The shallow-water (SW) equations describe the behavior of a shallow homogeneous incompressible and inviscid fluid layer. They are derived from the depth-averaged Navier–Stokes equations under Boussinesq and hydrostatic pressure assumptions. The SW system is extensively used in environmental studies to model hydrodynamics in lakes, estuaries, coastal regions, and other applications.

Galerkin techniques have gradually evolved to become a popular method for this problem class including the finite-element (FE) [15, 22, 34, 45, 58, 39, 68], the spectral-element [35, 61], the finite-volume (FV) [4, 16, 17, 23, 33], and the discontinuous Galerkin [3, 7, 19, 24, 59] methods. Indeed, unstructured elements offer the enhanced flexibility of using grids of variable sizes, shapes, and orientation for representing the boundaries of complex domains and a natural treatment of boundary conditions. Calculations on graded meshes of unstructured grids are hence standard practice [5, 22, 23, 24, 54, 64, 67].

For most numerical methods, including the finite-difference (FD), FV, and FE schemes, the coupling between the momentum and continuity equations is a delicate problem.

One of the issues associated with mixed formulations is the possibility of spurious modes that may arise for certain choices of grids and bases. For example, the piecewise linear Galerkin discretization of the SW equations (for velocity and elevation variables) is usually plagued by spurious oscillations [35, 65]. The appearance of such oscillations is mainly due to an inappropriate placement of variables on the grid and/or a bad choice of approximation function spaces. Improvements have been achieved through the use of suitable FD grids [6, 47], a variety of mixed-order FE interpolation schemes [2, 35, 34, 58, 54, 70], and a wave equation formulation [36, 39].

This difficulty with mixed methods is not specific to the SW problem alone but is well known also in other contexts such as primitive variable viscous flows where it has been extensively investigated for the Galerkin FE method [14, 28]. However, the situation in the mixed SW problem is different and, as noted initially in [65] and in subsequent studies [34, 58, 67], an analysis of the dispersion relation for a given formulation will explicitly ascertain the presence and determine the form of spurious modes as well as the dissipative/dispersive nature of a given formulation.

The dispersion relation analysis assumes that the solution of the differential equations can be separated and formulated as periodic in space and time. It differs from the Fourier analysis in that the time variable remains continuous by using the harmonic form of the equations. Consequently, a Fourier expansion is performed only for the nodal unknowns that appear in the discrete equations. The analysis results in a dispersion relation where the magnitude of the temporal frequency is expressed in terms of wave numbers. Such a technique was applied early on to the SW model in [44].

Two-dimensional dispersion analysis of FE SW discretizations using the same linear basis functions for all variables was examined first. One grid configuration is considered in [65], while the influence of grid configuration on the dispersion properties is analyzed in [9, 10, 25] for the mixed primitive variable form and in [42] in the case of the wave equation formulation. Phase behavior of FV SW discretizations has been studied in [29, 60]. To our knowledge, the two-dimensional dispersion relation analysis of the SW equations using the $P_1^{NC} - P_1$ FE pair [52] was the first study analyzing the dispersion relation of a FE scheme using different approximation spaces for the velocity and surface-elevation fields on meshes made up of triangles. The aim of the present study is to extend the two-dimensional dispersion relation analysis of the FE SW equations to a variety of mixed-order interpolation schemes. Because all the aforementioned problems occur primarily in the context of linear formulations, solving linear equations is sufficient for our purpose.

The paper is developed as follows: The linear SW equations and the Galerkin FE dis-

cretization schemes are presented in section 1.4. The discrete operators are computed in section 1.5 and the dispersion relations are obtained. The latter are compared analytically and graphically with the continuous case in section 1.6. This dispersion analysis is followed by numerical tests in section 1.7. Some concluding remarks complete the study.

1.4 Discretization of the linear SW equations

1.4.1 Governing equations

Let Ω be the model domain with boundary Γ. The inviscid linear SW equations are expressed in Cartesian coordinates [37] as

$$\bar{\mathbf{u}}_t + f\,\mathbf{k}\times\bar{\mathbf{u}} + g\,\nabla\bar{\eta} = 0\,, \tag{1.1}$$

$$\bar{\eta}_t + H\,\nabla\cdot\bar{\mathbf{u}} = 0\,, \tag{1.2}$$

where $\bar{\mathbf{u}} = (\bar{u},\bar{v})$ is the velocity field, $\bar{\eta}$ is the surface elevation with respect to the reference level $z = 0$, g is the gravitational acceleration, $\mathbf{k}$ is a unit vector in the vertical direction, and the mean depth H and the Coriolis parameter f are assumed constant. Note that $\bar{\eta}$ would be the pressure in the Navier–Stokes equations. For a contained flow, (1.1) and (1.2) are solved subject to the no-normal flow boundary condition $\bar{\mathbf{u}}\cdot\mathbf{n} = 0$ on Γ, where $\mathbf{n}$ is the outward pointing normal at the boundary.

For this analysis we seek periodic solutions of (1.1)–(1.2) of the form

$$\bar{\mathbf{u}}(x,y,t) = \mathbf{u}(x,y)e^{i\omega t}, \quad \bar{\eta}(x,y,t) = \eta(x,y)e^{i\omega t}, \tag{1.3}$$

where $\mathbf{u} = (u,v)$, η are amplitudes, and ω is the angular frequency, and we obtain

$$i\omega\,\mathbf{u} + f\,\mathbf{k}\times\mathbf{u} + g\,\nabla\eta = 0, \tag{1.4}$$

$$i\omega\,\eta + H\,\nabla\cdot\mathbf{u} = 0. \tag{1.5}$$

1.4.2 Spatial discretization

The weak formulation

We assume $\mathbf{u}$ and η belong to the spaces $\mathbf{V}$ and Q, respectively, with Q and $\mathbf{V}$ being either the square-integrable space $L^2(\Omega)$ or the Sobolev space $H^1(\Omega)$, i.e., the space of functions in $L^2(\Omega)$ whose first derivatives belong to $L^2(\Omega)$. The weak formulation of (1.4) and (1.5) requires the test functions $\boldsymbol{\varphi}$ (whose x- or y-component is formally denoted by φ) and ψ to be sufficiently regular and to belong, respectively, to the same function space as $\mathbf{u}$ and η, such that

$$i\omega \int_\Omega \mathbf{u}\cdot\boldsymbol{\varphi}\, d\Omega + \int_\Omega f(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\varphi}\, d\Omega + g\int_\Omega \nabla\eta\cdot\boldsymbol{\varphi}\, d\Omega = 0, \tag{1.6}$$

$$i\omega \int_\Omega \eta\,\psi\, d\Omega + H\int_\Omega \nabla\cdot\mathbf{u}\,\psi\, d\Omega = 0, \tag{1.7}$$

where $d\Omega$ is the area element.

Depending on the regularity of functions $\mathbf{u}$ and η, the terms containing derivatives in (1.6) and (1.7) may be integrated by parts using Green's theorem. In this way, (1.6) is rewritten if needed as

$$i\omega \int_\Omega \mathbf{u}\cdot\boldsymbol{\varphi}\, d\Omega + \int_\Omega f(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\varphi}\, d\Omega - g\int_\Omega \eta\,\nabla\cdot\boldsymbol{\varphi}\, d\Omega = 0, \tag{1.8}$$

and for (1.7) we obtain

$$i\omega \int_\Omega \eta\,\psi\, d\Omega - H\int_\Omega \mathbf{u}\cdot\nabla\psi\, d\Omega = 0, \tag{1.9}$$

by letting $\mathbf{u}\cdot\mathbf{n} = 0$ on Γ for all $\mathbf{u}$ belonging to $\mathbf{V}$. Hence, we also have $\boldsymbol{\varphi}\cdot\mathbf{n} = 0$ on Γ.

Galerkin FE discretization

The Galerkin method approximates the solution of (1.6) and (1.7), and eventually (1.8) and (1.9) if integration by parts needs to be performed, in finite-dimensional subspaces. Consider a FE triangulation $\mathcal{T}_h$, of the polygonal domain Ω, where h is a representative meshlength parameter that measures resolution. For triangle $K \in \mathcal{T}_h$, let $P_n(K)$ denote the space of polynomials of degree n on K.

The discrete solutions $\mathbf{u}_h$ and η_h sought belong to finite-dimensional spaces $\mathbf{V}_h$ and Q_h, respectively, whose restrictions on K belong to $P_k(K) \times P_k(K)$ for $\mathbf{u}_h$ and to $P_l(K)$ for η_h. The components of $\mathbf{u}_h$ and η_h are represented over a triangle K_i by interpolating functions $\varphi(x, y)$ of degree k and $\psi(x, y)$ of degree l (in the x- and y-components), respectively, with $\boldsymbol{\varphi}(x, y)$ belonging to $\mathbf{V}_h$ and $\psi(x, y)$ belonging to Q_h. We thus have

$$\mathbf{u}_h = \sum_{j \in S^{\mathbf{u}}_{K_i}} \mathbf{u}_j \, \boldsymbol{\varphi}_j, \quad \eta_h = \sum_{j \in S^{\eta}_{K_i}} = \eta_j \, \psi_j, \tag{1.10}$$

where j represents a node of K_i and $S^{\mathbf{u}}_{K_i}$ and $S^{\eta}_{K_i}$ denote the set of nodes of K_i.

The expansions of $\mathbf{u}_h$ and η_h over the whole domain Ω are then obtained by summing (1.10) over all triangles K_i of the triangulation $\mathcal{T}_h$. Introducing the FE basis leads to a FE statement as in (1.6) and (1.7), and eventually in (1.8) and (1.9), but with $\mathbf{u}, \eta$ replaced by the FE trial functions $\mathbf{u}_h, \eta_h$ and $\boldsymbol{\varphi}, \psi$ replaced by the corresponding FE test functions. We then decompose the integrals in (1.6) and (1.7) into triangle contributions.

For the purposes of the following analysis we consider a uniform mesh made up of biased right triangles as in the figures of section 1.5, and h is thus taken as a constant in the x- and y-directions.

The FE pairs

We now introduce the schemes that are used for the spatial discretization. Nine candidate FE pairs for representing velocity and surface elevation are described and evaluated in the remainder of this paper. Conventional FE terminology is adopted to describe the FE pairs. The nomenclature $P_m - P_n$ means that velocity components and surface elevation are represented, respectively, as piecewise-defined polynomials of degree m and n. Enhancements of this basic terminology are introduced as needed.

Common to the first six FE velocity/surface-elevation pairs (as shown in Figures 1.7(d) and 1.8) is a piecewise-linear continuous representation of surface elevation, and they differ from one another in their representation of velocity. The $P_1 - P_1$ pair has velocity colocated at triangle vertices, and the corresponding basis functions are piecewise-linear. The MINI element [8] also has continuous piecewise-linear basis functions at the vertices, but bubble functions are added at the barycenters for velocity in order to stabilize the pair. The $P_1^{NC} - P_1$ pair [34, 52] has velocity nodes at triangle edge midpoints, and linear basis functions are used to approximate the two velocity components on the element's two-triangle support. Since this particular representa-

tion of velocity is continuous only across triangle boundaries at midpoint nodes, and discontinuous everywhere else around a triangle boundary, this element is termed non-conforming (NC) in the FE literature. The P_1 iso P_2-P_1 element pair [12] has piecewise linear basis functions for velocity on a refined triangulation obtained by dividing each triangle into four subtriangles using the midpoints of triangle sides. There are thus six velocity nodes over each unrefined triangle, the same as for a quadratic approximation of velocity, termed P_2. The designation P_1 in P_1 iso P_2 denotes linear velocity elements on subtriangles, whereas iso P_2 indicates that the nodal placement is that associated with quadratic elements on unrefined triangles. The P_2-P_1 pair [32], also known in the literature as the Taylor–Hood element, has quadratic velocity basis functions. Finally, the P_0-P_1 pair, shown in Figure 1.7(d), has a piecewise-constant representation of velocity.

Common to the last three pairs (shown in Figures 1.7(a)–(c)) is a discontinuous piecewise-constant representation of surface elevation. The RT0 element, also called low-order Raviart–Thomas element [48], is based on flux conservation on element edges and has normal velocity components at triangle midedge points. Finally, the $P_1^{NC}-P_0$ and P_2-P_0 pairs have, respectively, P_1^{NC} and P_2 representations of velocity.

1.5 Computation of the dispersion relations

1.5.1 Matrix computations

For all FE pairs having a piecewise-constant representation of surface-elevation an integration by parts of $\nabla\eta$ in the left-hand side of (1.6) is performed, and hence (1.8) is employed. For the $P_1^{NC}-P_1$ and P_0-P_1 pairs, the term $\nabla\cdot\mathbf{u}$ appearing in (1.7) is integrated by parts to avoid computing u and v derivatives, and (1.9) is used instead.

After substitution of $\mathbf{u}_h$ and η_h from (1.10) into (1.6) and (1.7), the discrete formulation is obtained. We then need to compute

$$M_{i,j}=\int \boldsymbol{\varphi_i}\cdot\boldsymbol{\varphi_j}\,d\Omega,\quad G_{i,j}=\int \nabla\psi_i\cdot\boldsymbol{\varphi_j}\,d\Omega,\quad D_{i,j}=\int \nabla\cdot\boldsymbol{\varphi_i}\,\psi_j\,d\Omega, \tag{1.11}$$

where $M_{i,j}, G_{i,j}$, and $D_{i,j}$ represent the mass, gradient, and divergence elementary FE matrices, respectively. When (1.8) is employed instead of (1.6) the elementary gradient matrix is then $-D_{j,i}$, and if (1.9) is used, the elementary divergence matrix becomes $-G_{j,i}$. Assembling the elementary matrices leads to the stencils of Figures 1.1, 1.2, 1.3, 1.4, and 1.5 for the pairs examined here. The symbols • and black arrows indicate

the location of velocity and normal velocity nodes, respectively. The arrow points in the direction of the chosen normal. When these symbols are represented in grey this means that velocity and normal velocities have zero values at the corresponding nodes. The symbol ○ indicates nodes for surface elevation and dotted circles mean that the elevation is zero at those nodes.

For the subsequent dispersion analysis we consider two possible types of barycenters, corresponding to lower left and upper right triangles, and three types of faces: horizontal, vertical, and diagonal. Note that for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs in Figure 1.2, the x- (resp., y-) component of the surface elevation gradient is zero on horizontal (resp., vertical) faces. Further, in Figure 1.4 for the $P_2 - P_1$ pair, the surface-elevation gradient is zero at vertex nodes.

1.5.2 The continuous case

The free modes of (1.4)–(1.5) are examined by perturbing about the basic state $u = v = \eta = 0$. Because the governing equations are linear, the solution may be examined by considering the behavior of one Fourier mode. We then seek solutions of (1.4)–(1.5) of the form $(u, v, \eta) = (\tilde{u}, \tilde{v}, \tilde{\eta})\, e^{i(kx+ly)}$, where k and l are the wave numbers in the x- and y-directions, respectively. Substitution into (1.4)–(1.5) leads to a square matrix system for the amplitudes $\tilde{u}, \tilde{v}, \tilde{\eta}$. For a nontrivial solution to exist, the determinant of the matrix must equal zero, and this constraint leads to a relationship between the two wave numbers k and l and the frequency ω. The so-called dispersion relation is then obtained for the frequency

$$\omega\,(\,\omega^2 - f^2 - g\,H\,(\,k^2 + l^2\,)\,) = 0\,. \tag{1.12}$$

The first solution $\omega = 0$ is the geostrophic mode, and it would correspond to the slow Rossby mode on a β-plane, while the other two solutions

$$\omega_{AN} = \pm\sqrt{f^2 + g\,H\,(\,k^2 + l^2\,)} \tag{1.13}$$

correspond to the free-surface gravitational modes with rotational correction. Since ω is purely real, all modes are neutrally stable and neither amplify nor decay.

1.5.3 The discrete case

For each FE pair, the discrete momentum and continuity equations are obtained from the stencils of Figures 1.1, 1.2, 1.3, 1.4, and 1.5. As previously mentioned, the following

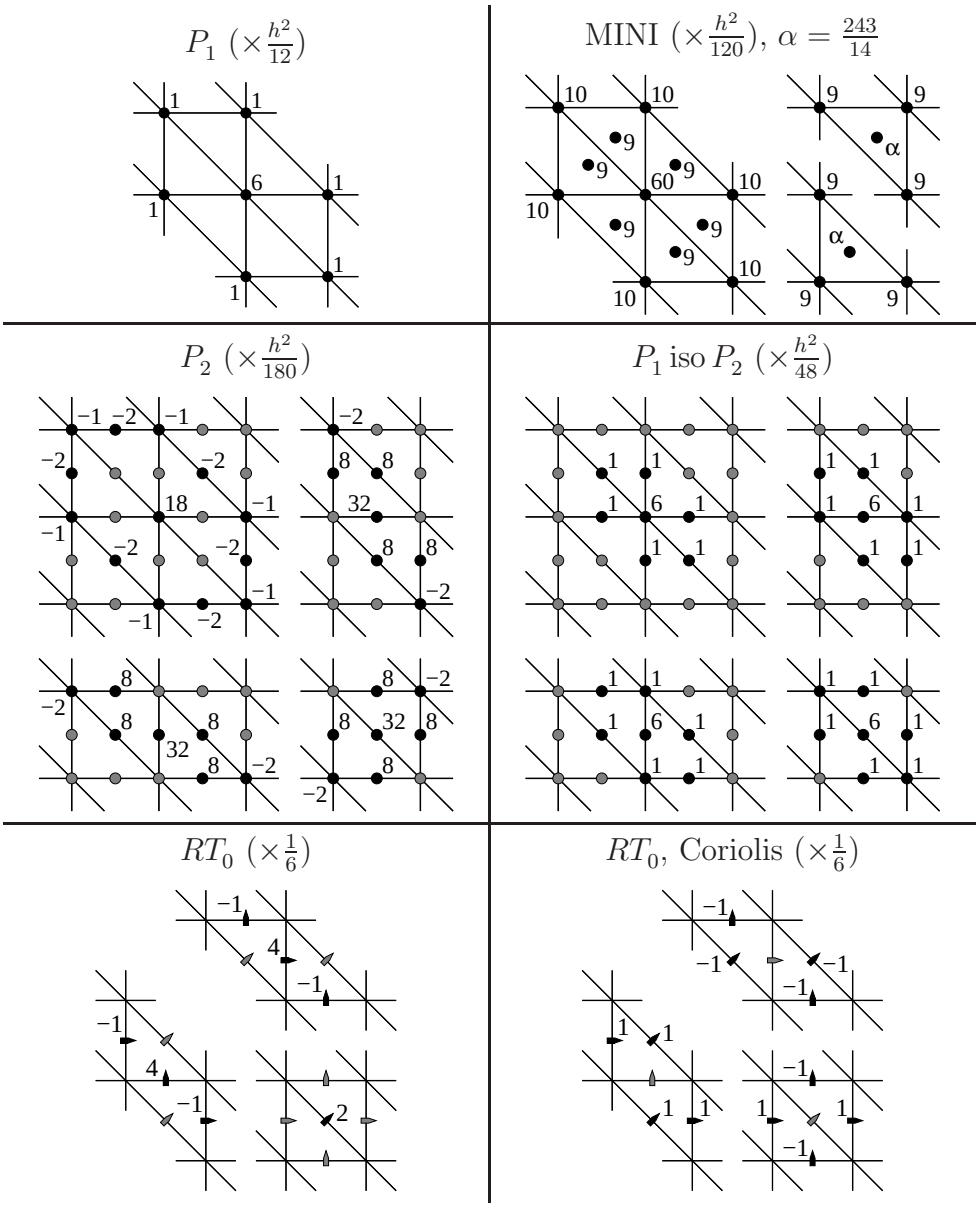


Figure 1.1: Mass stencils for several FE discretizations. The symbols • and black arrows indicate the location of velocity and normal velocity nodes, respectively. The arrow points in the direction of the chosen normal. Grey symbols mean that velocity and normal velocities have zero values at the corresponding nodes.

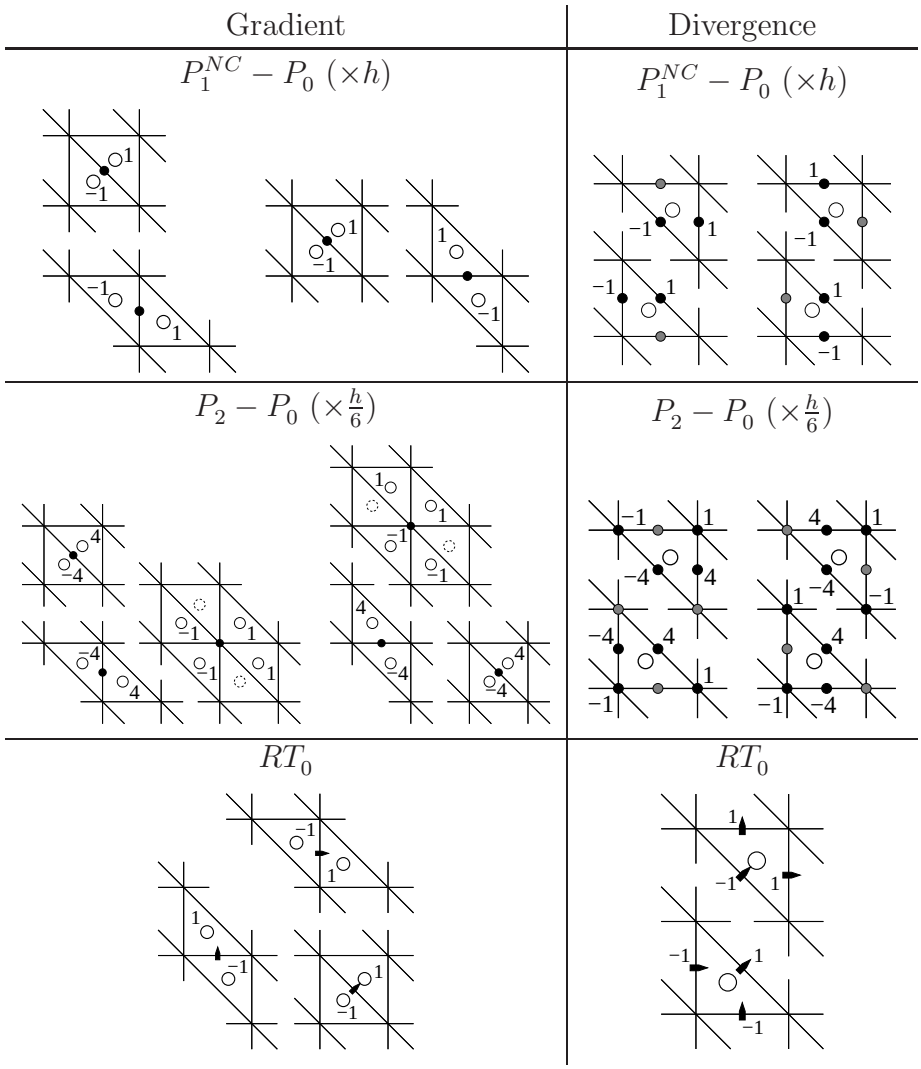


Figure 1.2: As for Figure 1.1 but for the gradient and divergence stencils of the $P_1^{NC}-P_0$, P_2-P_0, and RT0 pairs. For the $P_1^{NC}-P_0$ and P_2-P_0 pairs, the x- and y-components are on the left and right sides, respectively. The symbol $\bigcirc$ indicates nodes for surface elevation and dotted circles mean that the elevation is zero at those nodes.

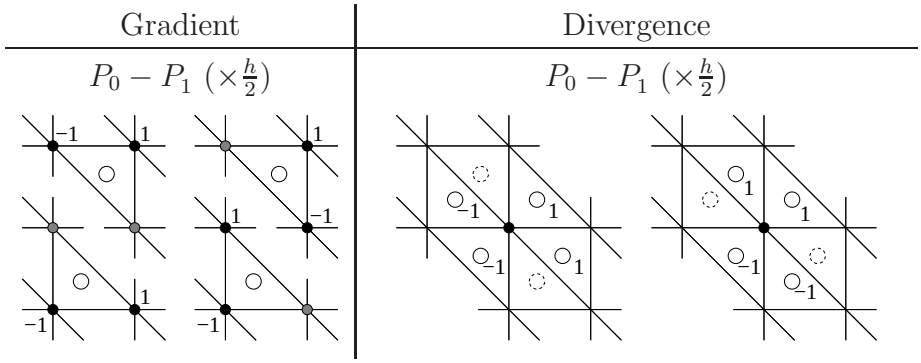


Figure 1.3: As for Figure 1.2 but for the $P_0 - P_1$ pair.

analysis will consider a uniform mesh made up of biased right triangles as in the figures of section 1.5, and h is thus taken as a constant in the x- and y-directions. Because nodal unknowns may be located on different types of nodes, i.e., vertices, faces, and barycenters, selected discrete equations for each type of node are retained. For example, the P_2-P_1 pair leads to considering four discrete momentum equations, one at a typical vertex node and three on the three possible types of faces, i.e., horizontal, vertical, and diagonal (written as H, V, and D, respectively, in the following), and only one discrete continuity equation at a typical vertex node (written as S in the following). For the $P_0 - P_1$ pair, two discrete momentum equations are retained at the two possible types of barycenters (corresponding to lower left and upper right triangles, written as C_1 and C_2, respectively, in the following). More details are given in [10, 25] for the $P_1 - P_1$ pair, and in [52] for the $P_1^{NC} - P_1$ one, to compute their dispersion relations. For all pairs, the typical nodes belonging to the same set (vertices, faces, midedge points, or barycenters) are hence distributed on a regular grid of size h.

As for the continuum case, the dispersion relation for the discrete scheme is found through a Fourier expansion. The discrete solutions corresponding to $(u_j, v_j, \eta_j) = (\tilde{u}, \tilde{v}, \tilde{\eta})\, e^{i(kx_j+ly_j)}$ are sought at node j ($j = 1, 2, 3\ldots$), where (u_j, v_j, η_j) are the nodal unknowns that appear in the selected discrete equations and $(\tilde{u}, \tilde{v}, \tilde{\eta})$ are amplitudes. The (x_j, y_j) coordinates are expressed in terms of a distance to a reference node. Substitution into the discrete equations leads to a square matrix system for the Fourier amplitudes. The dispersion relation is then obtained by setting the determinant of the matrix system to zero.

We let I_n be the $n \times n$ identity matrix, M an $m \times n$ matrix with m and n two positive integers, and $M^* = \overline{M}^T$ the conjugate transpose of M, and we define

$$a = \frac{1}{3}(\,3 + \cos kh + \cos lh + \cos(k-l)h\,), \quad A = \begin{pmatrix} i\omega & -f \\ f & i\omega \end{pmatrix},$$

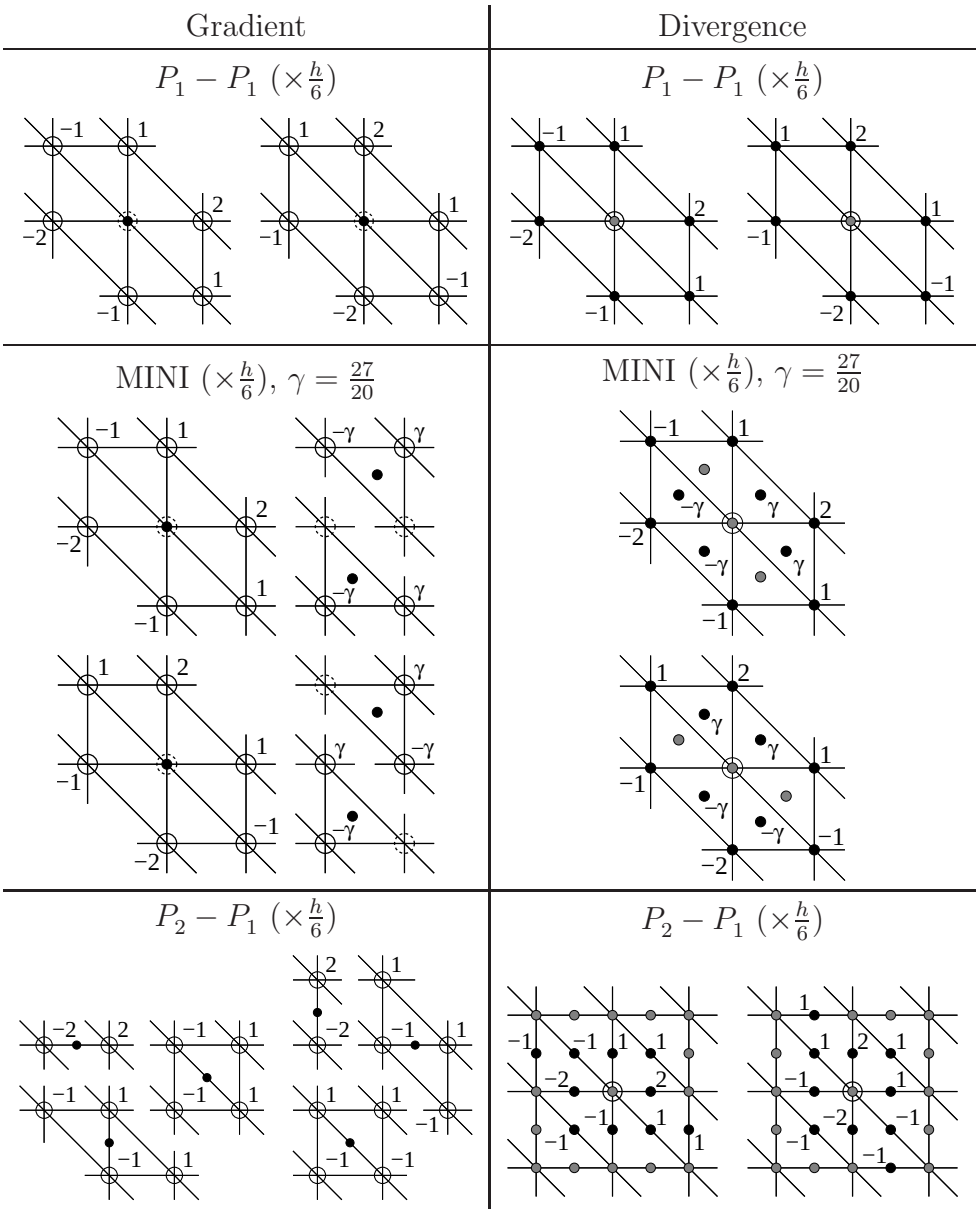


Figure 1.4: As for Figure 1.2 but for the $P_1 - P_1$, MINI and $P_2 - P_1$ pairs. For the $P_1 - P_1$ and $P_2 - P_1$ pairs, the x- and y-components are on the left and right sides, respectively, and they are at the top and bottom, respectively, for MINI.

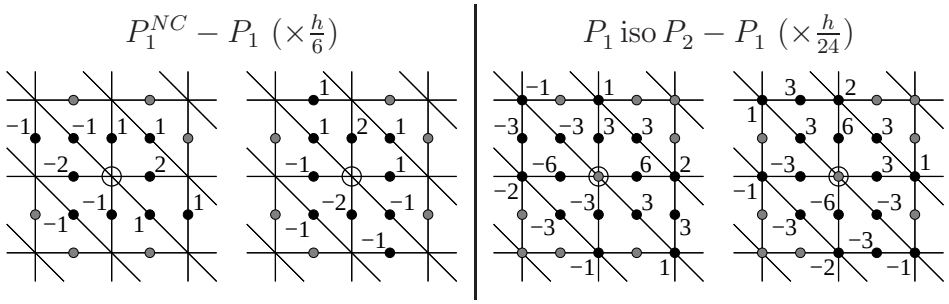


Figure 1.5: Divergence stencils for the $P_1^{NC} - P_1$ and P_1 iso $P_2 - P_1$ pairs. For each pair, the x- and y-components are on the left and right sides, respectively.

with $a \geq \frac{1}{2}$ for all k and l.

We now examine the dispersion relations corresponding to the nine pairs that are considered in this study.

1.5.4 The $P_1^{NC} - P_0$ pair

The selected discrete equations are written in the matrix form

$$\begin{pmatrix} A & 0 & 0 & \\ 0 & A & 0 & -g\,B^* \\ 0 & 0 & A & \\ & HB & & \frac{3}{2}i\omega I_2 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_H \\ \tilde{\mathbf{u}}_V \\ \tilde{\mathbf{u}}_D \\ \tilde{\boldsymbol{\eta}} \end{pmatrix} = 0, \tag{1.14}$$

where $\tilde{\boldsymbol{\eta}} = (\tilde{\eta}_{C_1}, \tilde{\eta}_{C_2})$, and

$$B = \frac{3}{h} \begin{pmatrix} 0 & -b_1 & -\overline{b}_3 & 0 & b_2 & b_2 \\ 0 & \overline{b}_1 & b_3 & 0 & -\overline{b}_2 & -\overline{b}_2 \end{pmatrix},$$

with

$$b_1 = e^{i(k-2l)\frac{h}{6}}, \quad b_2 = e^{i(k+l)\frac{h}{6}}, \quad b_3 = e^{i(2k-l)\frac{h}{6}}.$$

Vanishing the above 8×8 determinant leads to

$$\omega_{1,2} = 0, \;\; \omega_{3,4} = \pm f, \;\; \omega_{5,6} = \pm\sqrt{f^2 + 6\frac{gH}{h^2}\,(4-\alpha)}, \;\; \omega_{7,8} = \pm\sqrt{f^2 + 6\frac{gH}{h^2}\,(4+\alpha)},$$

with $\alpha = \sqrt{2(3a + \cos kh + \cos lh)}$. For infinitesimal mesh spacing we obtain

$$\omega_{5,6} = \pm\sqrt{f^2 + gH\left(\frac{9}{4}(k^2+l^2) - \frac{3}{2}kl\right)} + O(h^2), \tag{1.15}$$

$$\omega_{7,8} = \pm 4\sqrt{3}\frac{\sqrt{gH}}{h} + O(h). \tag{1.16}$$

Note that neither $\omega_{5,6}$ nor $\omega_{7,8}$ coincide with the continuous solution obtained from (1.12) in the limit as mesh spacing $h \to 0$. The roots $\omega_{7,8}$ correspond to a spurious mode from the $P_1^{NC} - P_0$ discretization scheme. Such a mode of type $O(\frac{1}{h})$ has already been observed in [53], where the one-dimensional SW equations are discretized using the discontinuous Galerkin method. Finally, the roots $\omega_{3,4}$ represent spurious inertial oscillations. It has been shown in [1] for the FD CD-grid and in [52] for the $P_1^{NC}-P_1$ pair that such spurious modes take the form of propagating inertial oscillations and have no particular spatial characteristics. They are a consequence of using more velocity nodes than surface-elevation nodes.

1.5.5 The $P_2 - P_0$ pair

The selected discrete equations lead to

$$\begin{pmatrix} a_1A & a_3A & a_4A & a_2A & \\ a_3A & a_8A & a_7A & a_6A & -gB^* \\ a_4A & a_7A & a_8A & a_5A & \\ a_2A & a_6A & a_5A & a_8A & \\ & HB & & & \frac{i\omega}{2}I_2 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_S \\ \tilde{\mathbf{u}}_H \\ \tilde{\mathbf{u}}_V \\ \tilde{\mathbf{u}}_D \\ \tilde{\boldsymbol{\eta}} \end{pmatrix} = 0, \tag{1.17}$$

where $\tilde{\boldsymbol{\eta}} = (\tilde{\eta}_{C_1}, \tilde{\eta}_{C_2})$,

$$B = \frac{1}{6h}\begin{pmatrix} -\overline{b}_1 + \overline{b}_2 & -\overline{b}_1 + \overline{b}_5 & 0 & -4\overline{b}_6 & -4\overline{b}_4 & 0 & 4\overline{b}_3 & 4\overline{b}_3 \\ b_1 - b_2 & b_1 - b_5 & 0 & 4b_6 & 4b_4 & 0 & -4b_3 & -4b_3 \end{pmatrix},$$

with $a_1 = \frac{1}{30}(4-a)$, $a_8 = \frac{8}{45}$, and

$$\begin{aligned} a_2 &= -\tfrac{1}{45}\cos\tfrac{(k+l)h}{2}, & a_5 &= \tfrac{4}{45}\cos\tfrac{kh}{2}, & b_1 &= e^{i(k+l)\frac{h}{3}}, & b_4 &= e^{i(2k-l)\frac{h}{6}}, \\ a_3 &= -\tfrac{1}{45}\cos\tfrac{(k-2l)h}{2}, & a_6 &= \tfrac{4}{45}\cos\tfrac{lh}{2}, & b_2 &= e^{-i(2k-l)\frac{h}{3}}, & b_5 &= e^{i(k-2l)\frac{h}{3}}, \\ a_4 &= -\tfrac{1}{45}\cos\tfrac{(2k-l)h}{2}, & a_7 &= \tfrac{4}{45}\cos\tfrac{(k-l)h}{2}, & b_3 &= e^{-i(k+l)\frac{h}{6}}, & b_6 &= e^{-i(k-2l)\frac{h}{6}}. \end{aligned}$$

Vanishing the above 10×10 determinant leads to a polynomial in ω of the form

$$(\omega^2 - f^2)^2 \left(c_6\,\omega^6 + c_4\,\omega^4 + c_2\,\omega^2 + c_0 \right) = 0,$$

where c_0, c_2, c_4, and c_6 are functions of k and l. We then obtain

$$\omega_{1,2} = \pm O(h), \quad \omega_{3,4,5,6} = \pm f \text{ (double root)}, \quad \omega_{7,8} = \pm O(1), \quad \omega_{9,10} = \pm O\left(\frac{1}{h}\right).$$

We have found that $c_0 \neq 0$, and hence $\omega = 0$ is no longer a solution to the dispersion relation, and the possibility of damped or growing slow modes occurs. However, we note the appearance of modes $\omega_{1,2} = \pm O(h)$ which are expected to be close to zero in the limit as mesh spacing $h \to 0$. As for the $P_1^{NC} - P_0$ pair, the roots $\omega_{3,4,5,6}$ represent spurious inertial oscillations. They are again due to an imbalance in the number of degrees of freedom between velocity and surface-elevation nodes. We have also found, as for the $P_1^{NC} - P_0$ pair, that the frequencies corresponding to inertia-gravity modes, here $\omega_{7,8}$, do not coincide with the continuous solution obtained from (1.12) in the limit as mesh spacing $h \to 0$. Finally, the frequencies $\omega_{9,10}$ again correspond to a spurious mode from the $P_2 - P_0$ discretization scheme, as for the $P_1^{NC} - P_0$ pair.

1.5.6 The RT_0 pair

The discrete equations lead to the following system:

$$\begin{pmatrix} A_{\text{RT0}} & -g\,B^* \\ HB & i\omega\frac{h^2}{2} I_2 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{J}} \\ \tilde{\boldsymbol{\eta}} \end{pmatrix} = 0, \tag{1.18}$$

where $\tilde{\mathbf{J}} = (\tilde{J}_H, \tilde{J}_D, \tilde{J}_V)$ and J_p is the velocity flux through the edge containing node p (here H, D, or V), with $J_p = \tilde{J}_p\, e^{i(kx_p + ly_p)}$. We also have $\tilde{\boldsymbol{\eta}} = (\tilde{\eta}_{C_1}, \tilde{\eta}_{C_2})$,

$$A_{\text{RT0}} = \frac{1}{3} \begin{pmatrix} 2i\omega & a_2 f & -a_3(i\omega - f) \\ -a_2 f & i\omega & a_1 f \\ -a_3(i\omega + f) & -a_1 f & 2i\omega \end{pmatrix}, \quad B = \begin{pmatrix} -b_1 & b_2 & -b_3 \\ \overline{b}_1 & -\overline{b}_2 & \overline{b}_3 \end{pmatrix},$$

and

$$\begin{aligned} a_1 &= \cos\tfrac{kh}{2}, & a_2 &= \cos\tfrac{lh}{2}, & a_3 &= \cos\tfrac{(k-l)h}{2}, \\ b_1 &= e^{i(k-2l)\frac{h}{6}}, & b_2 &= e^{i(k+l)\frac{h}{6}}, & b_3 &= e^{-i(2k-l)\frac{h}{6}}. \end{aligned}$$

For the 5×5 determinant to vanish we obtain $\omega = 0$ and four additional roots corresponding to gravity modes. For infinitesimal mesh spacing we have

$$\omega_1 = 0, \quad \omega_{2,3} = \omega_{AN} + O(h^2), \quad \omega_{4,5} = \pm 6\frac{\sqrt{gH}}{h} + O(h).$$

Note that only $\omega_{2,3}$ coincide with the continuous solution obtained from (1.12) in the limit as mesh spacing $h \to 0$, while $\omega_{4,5}$ presumably correspond to spurious modes from the RT_0 discretization scheme.

1.5.7 The $P_0 - P_1$ pair

The selected discrete equations are written in the matrix form

$$\begin{pmatrix} A & 0 & \\ 0 & A & -gB^* \\ & HB & ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_{C_1} \\ \tilde{\mathbf{u}}_{C_2} \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.19}$$

where $B = \frac{1}{h} \begin{pmatrix} b_1 & b_2 & -\overline{b}_1 & -\overline{b}_2 \end{pmatrix}$, with

$$b_1 = e^{i(k+l)\frac{h}{3}}(1 - e^{-ikh}), \quad b_2 = e^{i(k+l)\frac{h}{3}}(1 - e^{-ilh}).$$

Vanishing the above 5×5 determinant leads to

$$\omega_1 = 0, \quad \omega_{2,3} = \pm f, \quad \omega_{4,5} = \pm\sqrt{f^2 + 4\frac{gH}{h^2}\left(\frac{2 - \cos kh - \cos lh}{a}\right)}.$$

For infinitesimal mesh spacing the frequencies $\omega_{4,5}$ coincide with the continuous solution obtained from (1.12) in the limit as mesh spacing $h \to 0$, and we have $\omega_{4,5} = \omega_{AN} + O(h^2)$. Again, the roots $\omega_{2,3}$ represent spurious inertial oscillations.

1.5.8 The $P_1 - P_1$ pair

The selected discrete equations lead to

$$\begin{pmatrix} aA & -g\,B^* \\ H\,B & ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_S \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.20}$$

where $B = \frac{2i}{3h} \begin{pmatrix} b_1 & b_2 \end{pmatrix}$, with

$$b_1 = 2\sin kh + \sin lh + \sin(k-l)h, \quad b_2 = \sin kh + 2\sin lh - \sin(k-l)h.$$

For a nontrivial solution to exist, the 3×3 determinant of the coefficient matrix above must vanish. This condition implies

$$\omega_1 = 0, \quad \omega_{2,3} = \pm\sqrt{f^2 + \frac{4}{9}\frac{gH}{h^2}\frac{(b_1^2 + b_2^2)}{a^2}}.$$

For infinitesimal mesh spacing we obtain $\omega_{2,3} = \omega_{AN} + O(h^4)$. In the case $f = 0$, corresponding to the propagation of pure gravity waves, we may have $\omega_{2,3} = 0$ if $b_1 = b_2 = 0$. This occurs, in particular, when $kh = lh = \pi$ (wavelength $2h$) or when $kh = -lh = \frac{2\pi}{3}$ (wavelength $3h$). The waves of length $2h$ and $3h$ hence lead to $\omega_{2,3} = 0$ and do not propagate, but they are trapped within the model grid. They are numerical artifacts introduced by the spatial discretization scheme and describe oscillations of wavelength $2h$ and $3h$. Such waves are identified as spurious surface-elevation modes corresponding to physical eigenmodes which have their phase speed reduced to zero by the numerical scheme and appear as stationary oscillations [65, 25]. They can cause an accumulation of energy in the smallest-resolvable scale, leading to noisy solutions.

1.5.9 The MINI pair

If a cubic bubble is employed, the discrete equations lead to the following system:

$$\begin{pmatrix} aA & \overline{a}_1 A & a_1 A & \\ a_1 A & a_2 A & 0 & -g\,B^* \\ \overline{a}_1 A & 0 & a_2 A & \\ & HB & & ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_S \\ \tilde{\mathbf{u}}_{C_1} \\ \tilde{\mathbf{u}}_{C_2} \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.21}$$

where $B = \frac{1}{h} \begin{pmatrix} b_1 & b_2 & b_3 & b_4 & -\overline{b}_3 & -\overline{b}_4 \end{pmatrix}$, with

$$\begin{aligned} a_1 &= \tfrac{3}{20} e^{i(k+l)\frac{h}{6}} \left(2\cos\tfrac{(k-l)h}{2} + e^{\frac{-i}{2}(k+l)h} \right), & a_2 &= \tfrac{81}{280}, \\ b_1 &= \tfrac{2i}{3}(2\sin kh + \sin lh + \sin(k-l)h), & b_3 &= \tfrac{9}{20} e^{i(k+l)\frac{h}{3}}(1 - e^{-ikh}), \\ b_2 &= \tfrac{2i}{3}(2\sin lh + \sin kh + \sin(l-k)h), & b_4 &= \tfrac{9}{20} e^{i(k+l)\frac{h}{3}}(1 - e^{-ilh}). \end{aligned}$$

For the 7×7 determinant to vanish we obtain

$$\omega_1 = 0, \quad \omega_{2,3,4,5} = \pm f \text{ (double root)}, \quad \omega_{6,7} = O(1).$$

For infinitesimal mesh spacing we have $\omega_{6,7} = \omega_{AN} + O(h^2)$, and the roots $\omega_{2,3,4,5}$ represent spurious inertial oscillations.

The MINI discretization offers the particularity that equations associated with bubble nodes on an element depend only on the values of the solution on that element. This permits us to eliminate the bubble unknowns at the element level through a procedure called static condensation [27]. The linear system (1.21) reduces to a 3×3 one after static condensation, and for the corresponding 3×3 determinant to vanish we exactly obtain

the solutions ω_1 and $\omega_{6,7}$. Because the factor $(\omega^2 - f^2)^{-2}$ arises during the procedure, the solutions $\omega_{2,3,4,5}$ also need to be considered. Finally, we have also performed the calculations and numerical experiments with a linear bubble, instead of a cubic one, and insignificant differences have been observed.

1.5.10 The $P_1^{NC} - P_1$ pair

The selected discrete equations are written in the matrix form

$$\begin{pmatrix} A & 0 & 0 & \\ 0 & A & 0 & -g\,B^* \\ 0 & 0 & A & \\ & HB & & \frac{3}{2}ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_H \\ \tilde{\mathbf{u}}_V \\ \tilde{\mathbf{u}}_D \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.22}$$

where $B = \frac{1}{h}\begin{pmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{pmatrix}$, with

$$b_1 = 2i\sin\tfrac{kh}{2}, \quad b_2 = 2i\sin\tfrac{lh}{2}\cos\tfrac{(k-l)h}{2}, \quad b_3 = 2i\sin\tfrac{kh}{2}\cos\tfrac{(k-l)h}{2},$$
$$b_4 = 2i\sin\tfrac{lh}{2}, \quad b_5 = 2i\sin\tfrac{kh}{2}\cos\tfrac{lh}{2}, \qquad b_6 = 2i\sin\tfrac{lh}{2}\cos\tfrac{kh}{2}.$$

Vanishing the above 7×7 determinant leads to

$$\omega_1 = 0, \quad \omega_{2,3,4,5} = \pm f \text{ (double root)}, \quad \omega_{6,7} = \pm\sqrt{f^2 + 4\frac{gH}{h^2}\alpha},$$

with

$$\alpha = \sin^2\frac{kh}{2} + \sin^2\frac{lh}{2} + \frac{2}{3a}\left[\sin^4\frac{kh}{2} + \sin^4\frac{lh}{2}\right].$$

The above frequencies coincide with those found in [52], where temporally discretized SW equations are considered, in the limit as time step $\Delta t \to 0$. For infinitesimal mesh spacing we obtain $\omega_{6,7} = \omega_{AN} + O(h^4)$.

1.5.11 The P_1 iso $P_2 - P_1$ pair

The selected discrete equations lead to

$$\begin{pmatrix} 3A & a_1A & a_2A & a_3A & \\ a_1A & 3A & a_3A & a_2A & -gB^* \\ a_2A & a_3A & 3A & a_1A & \\ a_3A & a_2A & a_1A & 3A & \\ & & HB & & 12ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_S \\ \tilde{\mathbf{u}}_H \\ \tilde{\mathbf{u}}_V \\ \tilde{\mathbf{u}}_D \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.23}$$

where $B = \frac{1}{h}\begin{pmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & b_8 \end{pmatrix}$, with

$$\begin{aligned}
&a_1 = \cos\tfrac{kh}{2}, \quad a_2 = \cos\tfrac{lh}{2}, \quad a_3 = \cos\tfrac{(k-l)h}{2}, && b_4 = 6i(\sin\tfrac{kh}{2} - \sin\tfrac{(k-2l)h}{2}),\\
&b_1 = 2i(2\sin kh + \sin lh + \sin(k-l)h), && b_5 = 6i(\sin\tfrac{lh}{2} + \sin\tfrac{(2k-l)h}{2}),\\
&b_2 = 2i(2\sin lh + \sin kh - \sin(k-l)h), && b_7 = 6i(\sin\tfrac{(k-l)h}{2} + \sin\tfrac{(k+l)h}{2}),\\
&b_3 = 12i\sin\tfrac{kh}{2}, \quad b_6 = 12i\sin\tfrac{lh}{2}, && b_8 = 6i(\sin\tfrac{(k+l)h}{2} - \sin\tfrac{(k-l)h}{2}).
\end{aligned}$$

For the 9×9 determinant to vanish we obtain

$$\omega_1 = 0, \quad \omega_{2,3,4,5,6,7} = \pm f \text{ (triple root)}, \quad \omega_{8,9} = O(1).$$

For infinitesimal mesh spacing we have $\omega_{8,9} = \omega_{AN} + O(h^2)$.

1.5.12 The $P_2 - P_1$ pair

The discrete equations lead to the following system:

$$\begin{pmatrix} a_1A & a_4A & a_5A & a_3A & \\ a_4A & a_2A & a_8A & a_7A & -gB^* \\ a_5A & a_8A & a_2A & a_6A & \\ a_3A & a_7A & a_6A & a_2A & \\ & HB & & & \frac{1}{2}ia\omega \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}}_S \\ \tilde{\mathbf{u}}_H \\ \tilde{\mathbf{u}}_V \\ \tilde{\mathbf{u}}_D \\ \tilde{\eta}_S \end{pmatrix} = 0, \tag{1.24}$$

where $B = \frac{1}{h}\begin{pmatrix} 0 & 0 & b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{pmatrix}$, with $a_1 = \frac{1}{30}(4-a)$, $a_2 = \frac{8}{45}$, and

$$\begin{aligned}
&a_3 = \tfrac{-1}{45}\cos\tfrac{(k+l)h}{2}, && a_4 = \tfrac{-1}{45}\cos\tfrac{(k-2l)h}{2}, && a_5 = \tfrac{-1}{45}\cos\tfrac{(2k-l)h}{2},\\
&a_6 = \tfrac{4}{45}\cos\tfrac{kh}{2}, && a_7 = \tfrac{4}{45}\cos\tfrac{lh}{2}, && a_8 = \tfrac{4}{45}\cos\tfrac{(k-l)h}{2},\\
&b_1 = \tfrac{2i}{3}\sin\tfrac{kh}{2}, && b_2 = \tfrac{2i}{3}\sin\tfrac{lh}{2}\cos\tfrac{(k-l)h}{2}, && b_3 = \tfrac{2i}{3}\sin\tfrac{kh}{2}\cos\tfrac{(k-l)h}{2},\\
&b_4 = \tfrac{2i}{3}\sin\tfrac{lh}{2}, && b_5 = \tfrac{2i}{3}\sin\tfrac{kh}{2}\cos\tfrac{lh}{2}, && b_6 = \tfrac{2i}{3}\sin\tfrac{lh}{2}\cos\tfrac{kh}{2}.
\end{aligned}$$

Vanishing the above 9×9 determinant leads to

$$\omega_1 = 0, \quad \omega_{2,3,4,5,6,7} = \pm f \text{ (triple root)}, \quad \omega_{8,9} = O(1).$$

For infinitesimal mesh spacing we obtain $\omega_{8,9} = \omega_{AN} + O(h^2)$.

Table 1.1: Number of frequencies of type $\omega = 0$, $O(1)$, $O(\frac{1}{h})$, $O(h)$, $\pm f$, solutions of the dispersion relation of degree n for the nine FE schemes examined in section 1.5.

FE pair	n	$\omega = 0$	$O(1)$	$O(1) \to \omega_{AN}$ when $h \to 0$	$O(\frac{1}{h})$	$O(h)$	f	$-f$
$P_1^{NC} - P_0$	8	2	2	no	2	0	1	1
$P_2 - P_0$	10	0	2	no	2	2	2	2
RT_0	5	1	2	yes	2	0	0	0
$P_0 - P_1$	5	1	2	yes	0	0	1	1
$P_1 - P_1$	3	1	2	yes	0	0	0	0
MINI	7	1	2	yes	0	0	2	2
$P_1^{NC} - P_1$	7	1	2	yes	0	0	2	2
P_1 iso $P_2 - P_1$	9	1	2	yes	0	0	3	3
$P_2 - P_1$	9	1	2	yes	0	0	3	3

1.5.13 Summary of discrete frequencies

The previous results are summarized in Table 1.1, where n is the degree of the dispersion relation, and hence the total number of discrete frequencies, for the nine FE schemes examined in section 1.5.

We mention whether frequencies of type $O(1)$, corresponding to inertia-gravity modes, coincide with the analytical solution ω_{AN} in the limit as mesh parameter $h \to 0$. This is the case for all FE pairs, except for the $P_1^{NC} - P_0$ and $P_2 - P_0$ schemes, i.e., when n is even. We also notice the presence of spurious frequencies of type $O(\frac{1}{h})$ for the $P_1^{NC} - P_0$, $P_2 - P_0$, and RT_0 pairs, i.e., when a piecewise-constant representation of the surface elevation is employed. Such solutions have already been observed in [53], where the one-dimensional SW equations are discretized using the discontinuous Galerkin method. The slow mode corresponding to $\omega = 0$ is not present for the $P_2 - P_0$ pair; however, a mode of type $O(h)$ appears instead, while $\omega = 0$ is a double root for the $P_1^{NC} - P_0$ pair. Finally, the solutions $\pm f$ are propagating spurious inertial oscillations that have no particular spatial characteristics. They are a consequence of using more velocity nodes than surface-elevation nodes [52] and are present for all schemes except for the $P_1 - P_1$ and RT_0 pairs.

1.6 Analysis of the dispersion relations

We now analyze the computed frequencies obtained from the dispersion relations for the nine FE pairs examined in section 1.5. We first consider the gravity wave limit, where $f = 0$, and will then comment on the cases $f \neq 0$ and other modes.

1.6.1 Gravity wave limit

We let $f = 0$, and examine the computed frequencies, denoted by ω_{CP}, corresponding to pure gravity modes of type $O(1)$, as functions of kh and lh. From (1.13) and the results of section 1.5, we determine the analytical and computed phase speeds, denoted by c_{AN} and c_{CP}, respectively,

$$c_{AN} \equiv \frac{\omega_{AN}}{\sqrt{k^2 + l^2}} = \pm\sqrt{gH}, \quad c_{CP} \equiv \frac{\omega_{CP}}{\sqrt{k^2 + l^2}}.$$

We then obtain the phase speed ratio of the computed phase speed to the analytical one, denoted by r_{PH}, with

$$r_{PH} \equiv \frac{|c_{CP}|}{|c_{AN}|} = \frac{|\omega_{CP}|}{|\omega_{AN}|} = \frac{|\omega_{CP}|}{\sqrt{gH(k^2 + l^2)}}. \tag{1.25}$$

Note that we should have $r_{PH} = 1$ in the absence of numerical dispersion. We show r_{PH} as a surface function depending on kh and lh, and along the selected axes OX, OY, OD1, and OD2 (vertical section), in Figures 1.7 and 1.8, respectively, for the FE schemes previously examined. As in [25], the values of kh and lh vary over their complete domain $[-\pi, \pi]$. However, the phase speed ratio along selected axes is plotted only on $[0, \pi]$, and the solution is then deduced on $[-\pi, 0]$ by symmetry through the origin. The directions OX, OY, OD1, and OD2 are shown in Figure 1.6.

The directions OX and OY correspond to waves travelling in the x- and y-directions, for $lh = 0$ and $kh = 0$, respectively, while the directions OD1 and OD2 correspond to waves travelling along the diagonal axes, for $kh = lh$ and $lh = -kh$, respectively.

For the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs in Figure 1.7, r_{PH} is unexpectedly greater than 1 for small values of $|kh|$ and $|lh|$. Indeed, depending on the direction the wave numbers tend to zero, the phase speed ratio ranges from 1.2 to 1.7 for the $P_1^{NC} - P_0$ pair and from 1.8 to 2.5 for the $P_2 - P_0$ one. We notice that for both pairs r_{PH} decreases for increasing values of $|kh|$ and $|lh|$, implying that gravity modes propagate more rapidly for small wave numbers. Consequently, both pairs poorly represent the propagation of pure gravity waves.

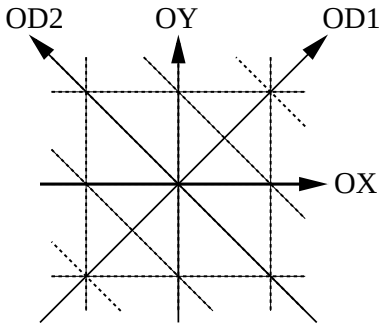


Figure 1.6: Definition of the directions OX, OY, OD1, and OD2 on a uniform mesh made up of biased right triangles.

For the $P_1 - P_1$ pair in Figure 1.8, we obtain $r_{PH} \leq 1$ for all values of kh and lh. Hence, the waves are travelling slower than expected, and this is particularly true for large values of $|kh|$ and $|lh|$. We have $r_{PH} = 0$ in the OX, OY, and OD1 directions at the following values of (kh, lh), $(0, \pm\pi)$, $(\pm\pi, 0)$, $(\pm\pi, \pm\pi)$ (waves of length $2h$), and also in the OD2 direction for $|kh| = |lh| = \frac{2\pi}{3}$ (waves of length $3h$). This is a consequence of having zero frequency for waves of length $2h$ and $3h$, as shown previously in section 3.8. Such waves do not propagate and describe oscillations of wavelengths $2h$ and $3h$. They are spurious surface-elevation modes introduced by the spatial discretization scheme. Hence, the $P_1 - P_1$ pair is usually not used to solve the SW system, unless a wave equation formulation or a stabilization procedure is employed [22, 39].

For the remaining FE pairs in Figures 1.7 and 1.8, we have r_{PH} close to 1 for small values of $|kh|$ and $|lh|$ in all directions. However, for increasing values of $|kh|$ and $|lh|$, corresponding to wavelengths ranging from $2h$ to $4h$, r_{PH} departs moderately from 1 in the OX, OY, and OD1 directions for the RT_0 element (waves decelerate) and the $P_0 - P_1$ pair (waves accelerate). In the OD2 direction, except for wavelengths ranging from $2h$ to $3h$, r_{PH} strongly departs from 1 for the $P_0 - P_1$, MINI, P_1 iso $P_2 - P_1$, and $P_2 - P_1$ pairs, and moderately for the RT_0 and $P_1^{NC} - P_1$ pairs, for increasing values of $|kh|$ and $|lh|$, and the waves will accelerate.

From (1.13) and the results of section 1.5, we now determine the analytical and computed group velocities, $\mathbf{c}_{g,AN}$ and $\mathbf{c}_{g,CP}$, respectively, i.e., the velocity at which the energy is carried by the waves, with

$$\mathbf{c}_{g,AN} = \left(\frac{\partial \omega_{AN}}{\partial k}, \frac{\partial \omega_{AN}}{\partial l} \right), \quad \mathbf{c}_{g,CP} = \left(\frac{\partial \omega_{CP}}{\partial k}, \frac{\partial \omega_{CP}}{\partial l} \right).$$

We also calculate the directional derivatives $\mathbf{c}_{g,AN} \cdot \mathbf{d}$ and $\mathbf{c}_{g,CP} \cdot \mathbf{d}$, where $\mathbf{d}$ is a unit vector in the OX, OY, OD1, or OD2 directions. For the continuous case and the nine FE pairs of section 1.5, the group velocities (left panels) and the directional derivatives

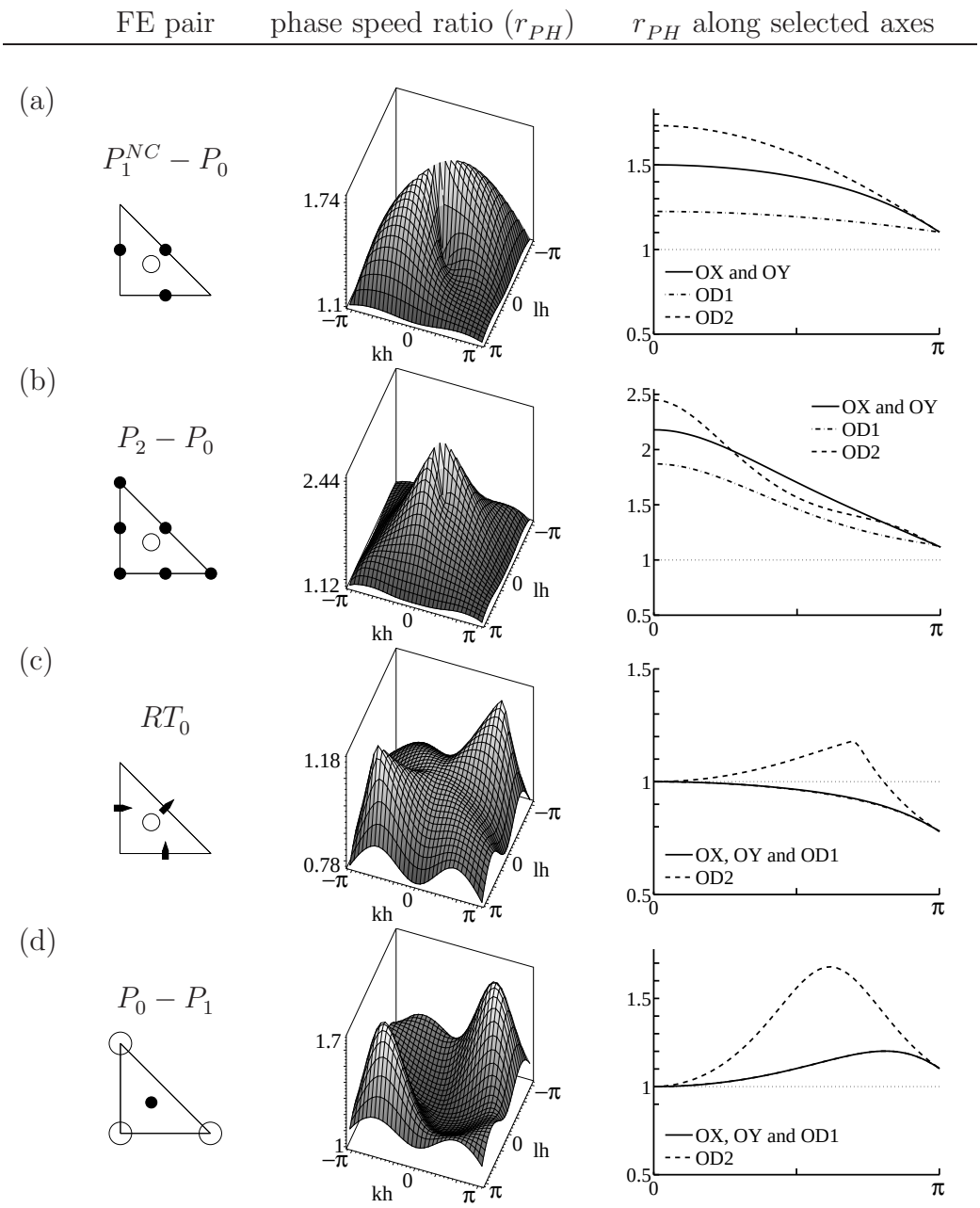


Figure 1.7: The phase speed ratio is plotted as a surface function and along selected axes (OX, OY, OD1, and OD2) for the (a) $P_1^{NC} - P_0$, (b) $P_2 - P_0$, (c) RT0, and (d) $P_0 - P_1$ pairs.

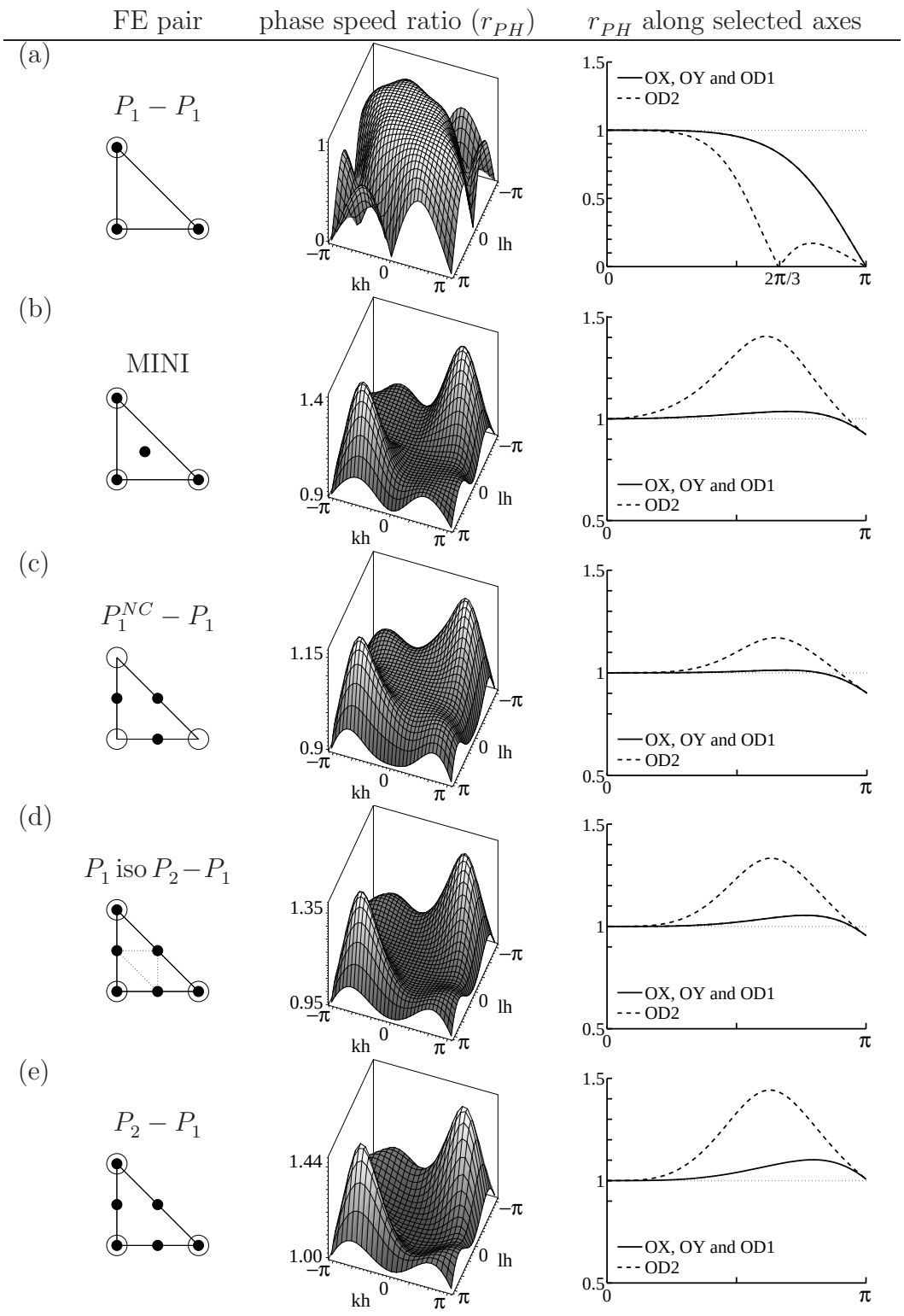


Figure 1.8: As for Figure 1.7 but for the (a) $P_1 - P_1$, (b) MINI, (c) $P_1^{NC} - P_1$, (d) P_1 iso $P_2 - P_1$, and (e) $P_2 - P_1$ pairs.

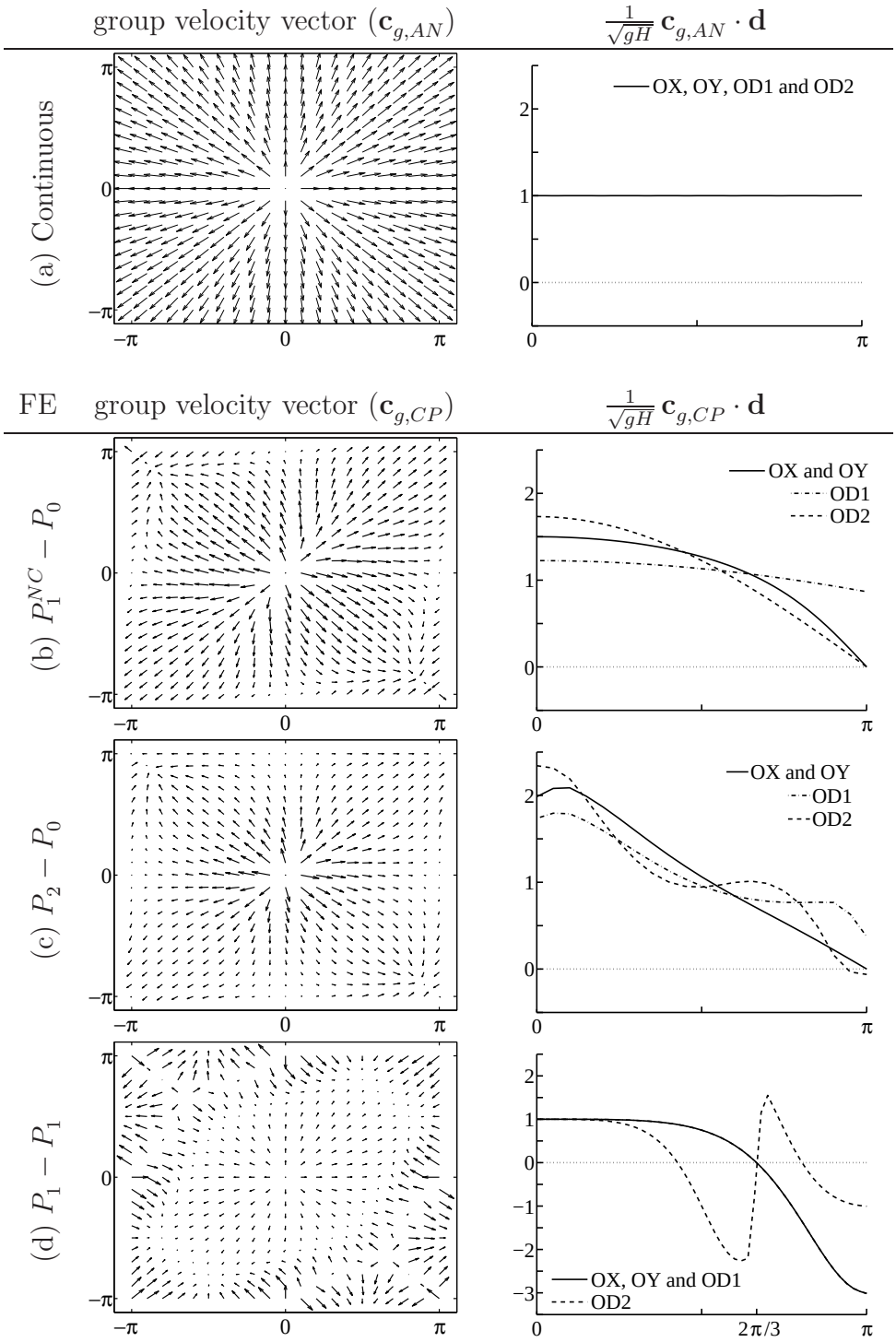


Figure 1.9: The group velocity vector (left panels) where the Cartesian coordinate axes are kh and lh, and the normalized directional derivative $\frac{1}{\sqrt{gH}}\,\mathbf{c}_g\cdot\mathbf{d}$ (right panels), where $\mathbf{d}$ is a unit vector in the OX, OY, OD1, or OD2 direction, for (a) the continuous case and the (b) $P_1^{NC}-P_0$, (c) P_2-P_0, and (d) P_1-P_1 pairs.

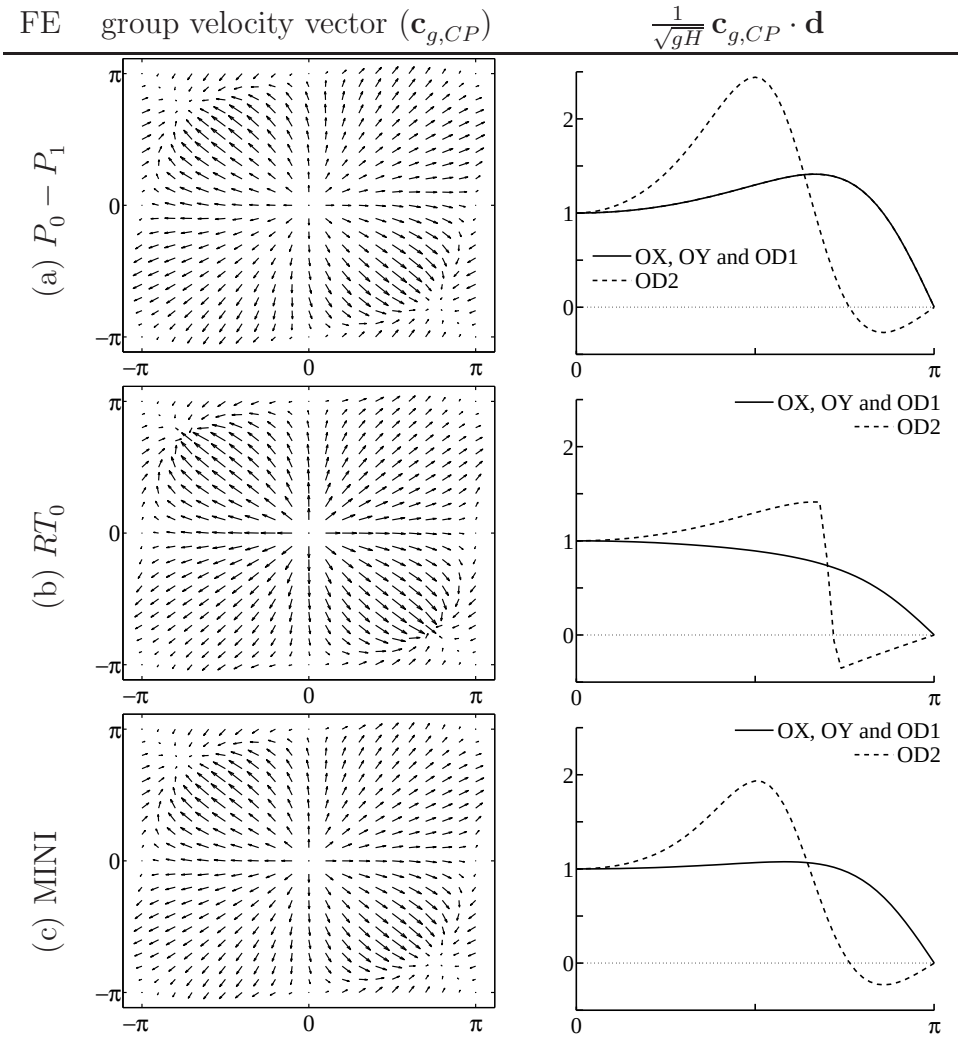


Figure 1.10: As for Figure 1.9 but for the (a) $P_0 - P_1$, (b) RT0, and (c) MINI pairs.

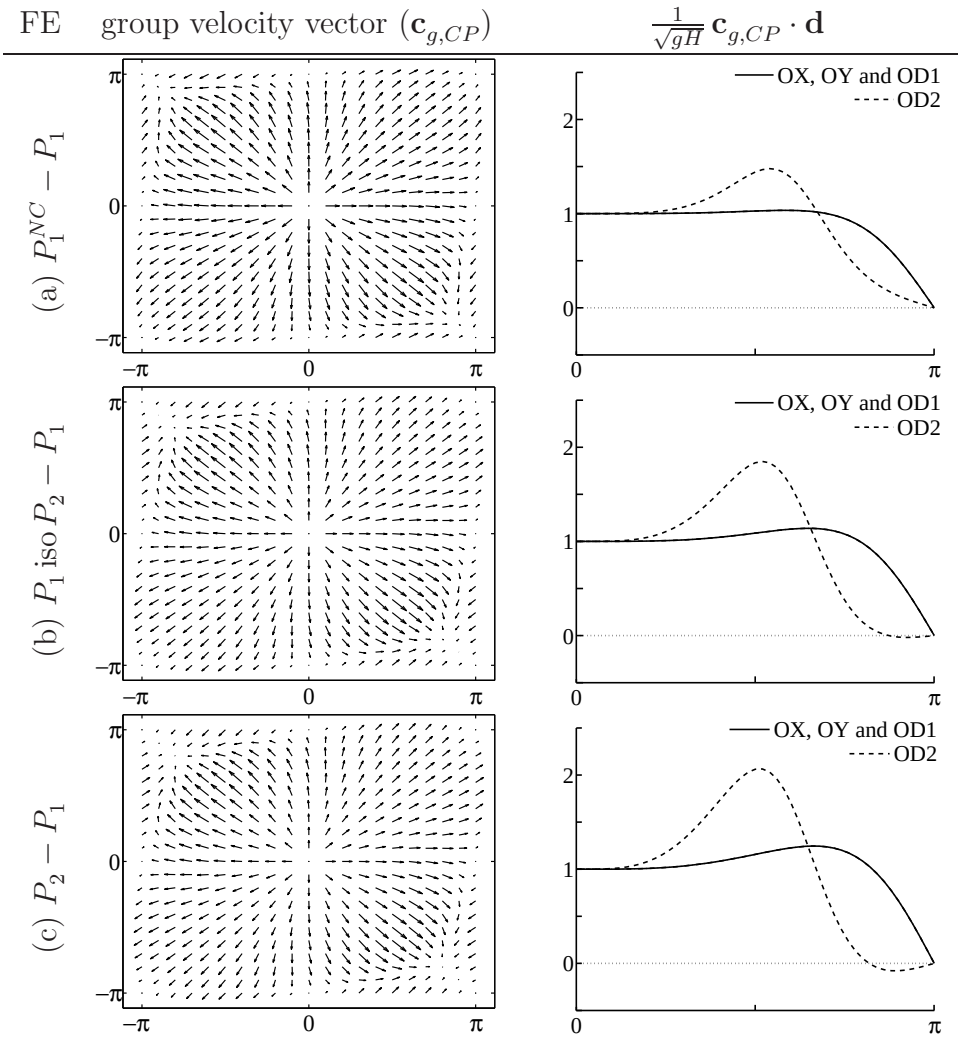


Figure 1.11: As for Figure 1.9 but for the (a) $P_1^{NC} - P_1$, (b) P_1 iso $P_2 - P_1$, and (c) $P_2 - P_1$ pairs.

normalized by the factor $\sqrt{gH}$ (right panels) are shown in Figures 1.9, 1.10, and 1.11.

For the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs in Figure 1.9, the normalized directional derivatives $\mathbf{c}_{g,CP} \cdot \mathbf{d}$ unexpectedly depart from 1 in all directions for small values of $|kh|$ and $|lh|$. Hence, $\mathbf{c}_{g,CP}$ is not consistent with the analytic solution for both pairs as kh and lh tend to zero, and large inaccuracies can be expected in the speed of energy propagation. For the $P_1 - P_1$ pair in Figure 1.9, the group velocity is zero for waves of length $3h$ in the OX, OY, and OD1 directions, and for wavelengths ranging from $2h$ to $3h$ the group velocity has the wrong sign. In the OD2 direction, because the group velocity is zero three times for waves longer than $2h$, it has alternatively right and wrong signs. This is a consequence of having zero frequency for waves of length $2h$ and $3h$, as shown in section 3.8.

For the remaining pairs of Figures 1.10 and 1.11, the group velocity is consistent with the analytic solution as kh and lh are small. However, at the following values of (kh, lh), $(0, \pm\pi)$, $(\pm\pi, 0)$, $(\pm\pi, \pm\pi)$ (waves of length $2h$), we have $\mathbf{c}_{g,CP} = \mathbf{0}$ contrarily to the analytic case, and hence although waves of length $2h$ are propagating, the associated energy is not. For the RT_0, $P_0 - P_1$, and MINI pairs in Figure 1.10 and for the $P_2 - P_1$ pair in Figure 1.11, we observe that $\mathbf{c}_{g,CP} \cdot \mathbf{d}$ has a negative (wrong) sign in the OD2 direction for wavelengths greater than $2h$ and smaller than $3h$. Finally, for all pairs in the OD2 direction, the magnitude of the group velocity reaches a peak ranging from 1.5 (for the $P_1^{NC} - P_1$ and RT_0 pairs) to 2.5 (for the $P_0 - P_1$ pair) times the analytical value for wavelengths close to $4h$, leading to inaccuracies if such waves are generated and sustained in a numerical model.

1.6.2 Inertia-gravity waves and other modes

We now assume $f \neq 0$ and note $\lambda \equiv \sqrt{gH}/f$, the Rossby radius of deformation. Within a numerical factor, it is the distance covered by a wave travelling at the speed $\sqrt{gH}$ during one inertial period $(2\pi/f)$. Using (1.13) and the results of section 1.5 we now obtain the phase speed ratio

$$r_{PH}^{f} \equiv \frac{|c_{CP}|}{|c_{AN}|} = \frac{|\omega_{CP}|}{|\omega_{AN}|} = \frac{f\sqrt{1 + \left(\frac{\lambda}{h}\right)^2 t(kh, lh)}}{f\sqrt{1 + \left(\frac{\lambda}{h}\right)^2 (k^2h^2 + l^2h^2)}}, \tag{1.26}$$

where ω_{CP} corresponds to an inertia-gravity mode of type $O(1)$, as in (1.25) but now with $f \neq 0$, and $t(kh, lh)$ is a trigonometric expression depending on kh and lh.

For the $P_1 - P_1$, $P_0 - P_1$, $P_1^{NC} - P_1$, P_1 iso $P_2 - P_1$, MINI, and $P_2 - P_1$ pairs, a realistic

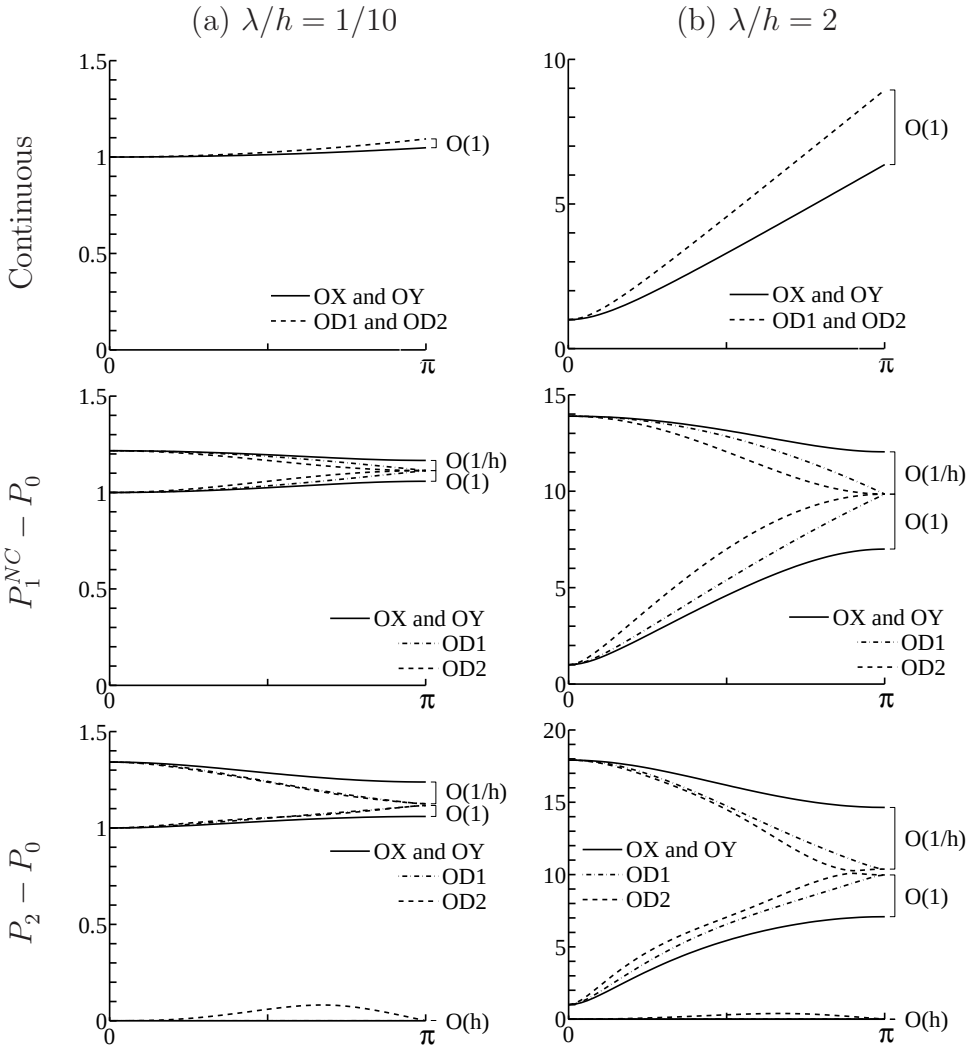


Figure 1.12: The nondimensional frequency ω_{CP}/f, corresponding to modes in $O(1)$, $O(\frac{1}{h})$, and $O(h)$ along selected axes (OX, OY, OD1, and OD2), for the continuous case and the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs. We note $\lambda \equiv \sqrt{gH}/f$, the radius of deformation (Rossby radius), and two cases are considered, namely, $\lambda/h = 1/10$ (coarse resolution) and $\lambda/h = 2$ (high resolution).

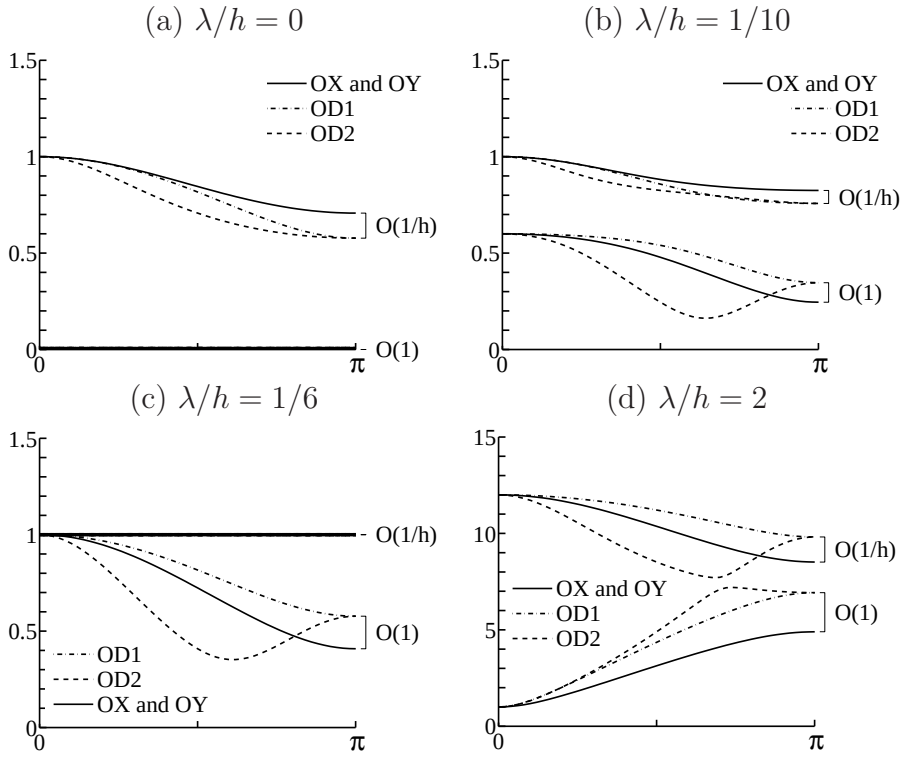


Figure 1.13: As for Figure 1.12 but for the RT0 element in the cases $\lambda/h = 0$, 1/10, 1/6, and 2.

positive value for f would have little effect on the plots of Figures 1.7 to 1.11 when $\lambda/h \gg 1$ (high resolution), as $t(kh, lh)$ is bounded in (1.26) for $0 \leq kh \leq \pi$ and $0 \leq lh \leq \pi$. Hence, we have $r^f_{PH} \simeq r_{PH}$ for $\lambda/h \gg 1$. In the case $\lambda/h \ll 1$ (coarse resolution), we deduce from (1.26) that $r^f_{PH} \simeq 1$. However, the results obtained for r_{PH} in Figures 1.7 to 1.11 are roughly preserved but at smaller scales.

The remaining $P_1^{NC} - P_0$, $P_2 - P_0$, and RT_0 pairs, as mentioned in Table 1.1, lead to discrete frequencies of type $O(\frac{1}{h})$, and additionally in $O(h)$ for the $P_2 - P_0$ pair. The nondimensional frequency ω_{CP}/f, corresponding to modes in $O(1)$, $O(\frac{1}{h})$, and $O(h)$ along selected axes (OX, OY, OD1, and OD2), is plotted for these pairs in Figures 1.12 and 1.13 and is compared with the continuous case. We consider λ/h ranging from 1/10 (coarse resolution) to 2 (high resolution).

For the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs in Figure 1.12, the mode of type $O(\frac{1}{h})$ is a spurious mode appearing as an approximate mirror image of the mode in $O(1)$. For the $P_2 - P_0$ pair, the mode in $O(h)$ is trapped close to zero, and this is especially true when the Rossby radius is resolved (high resolution). Because the $P_2 - P_0$ pair is the only one in Table 1.1 to have a mode in $O(h)$ and no mode corresponding to the geostrophic mode $\omega = 0$ in (1.12), we suggest considering the mode in $O(h)$ as a spurious one which coincides with the geostrophic mode as mesh spacing $h \to 0$. Recall that the geostrophic mode would correspond to the slow Rossby mode on a β-plane.

For the RT_0 pair in Figure 1.13 the mode in $O(\frac{1}{h})$ behaves quite differently from the corresponding ones for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs in Figure 1.12. Indeed, for values of λ/h below 0.2 (obtained numerically), the mode in $O(\frac{1}{h})$ is much closer to the continuous solution in Figure 1.13 than the $O(1)$ one. Moreover, the mode in $O(1)$ is uniformly zero for all kh and lh in the inertial limit ($g = 0$, i.e., $\lambda/h = 0$), but it tends to draw closer to the continuous solution as the values of λ/h progressively increase. However, in order for the mode in $O(1)$ to be monotonic in kh and lh, as for the continuous one, λ/h has to be chosen greater than 0.45 (obtained numerically). The modes in $O(\frac{1}{h})$ and in $O(1)$ have no common intersection except locally, when $\lambda/h = 1/6$ and $kh = lh = 0$, as shown in Figure 1.13(c). Note the mode in $O(\frac{1}{h})$ is uniformly one for all kh and lh in the case $\lambda/h = 1/6$, and as λ/h progressively increases (above 1/6), it appears as an approximate mirror image of the mode in $O(1)$, i.e., as a spurious mode. Finally, in the OD2 direction both the $O(\frac{1}{h})$ and $O(1)$ modes exhibit nonmonotonic curves for a wide range of λ/h.

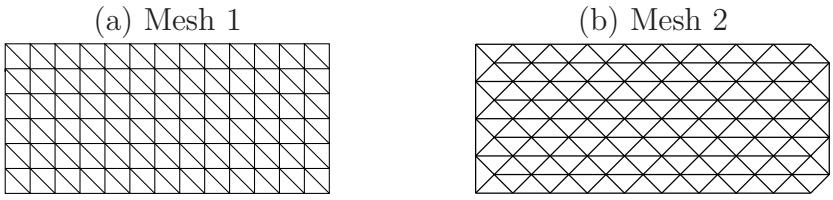


Figure 1.14: A window of (a) Mesh 1, made up of biased right isosceles triangles, and (b) Mesh 2, obtained from Mesh 1 by rotation.

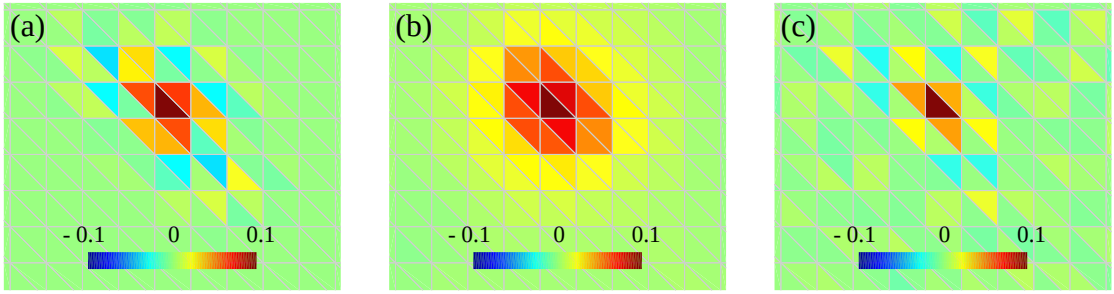


Figure 1.15: Window of the surface-elevation field around the point mass source (0.1 m) for the RT0 pair; after 1 day of simulation and (a) $\lambda/h = 1/6$ (coarse resolution), (b) $\lambda/h = 2$ (high resolution); (c) after 10 days of simulation and $\lambda/h = 1/12$ (coarse resolution).

1.7 Numerical results

The results of two tests using (1.1) and (1.2) are now presented. The first test is performed in [11] for FD grids. It is reproduced here to examine the consequences of the distortion observed in Figure 1.13 in the dispersion relation of the RT_0 element for small values of λ/h (coarse resolution).

The main point of interest is to verify whether or not the RT_0 discretization properly simulates the geostrophic adjustment process, in which the dispersion of the inertia-gravity waves leads to the establishment of a geostrophic balance.

The square domain extent is 5500 km $\times$ 5500 km, and it is discretized using biased right isosceles triangles (as in Figure 1.14(a)) with a uniform node spacing $h = 220$ km, leading to a 25×25 mesh. The fluid is initially at rest, and zero normal velocity is specified at the lateral boundaries. A point mass source and a point mass sink of 0.1 m and -0.1 m, respectively, are prescribed at fixed locations $12\,h$ apart at each time step; otherwise the initial surface elevation field is uniform. The Coriolis parameter is held constant and evaluated close to 45° N, with $f = 10^{-4}\,\mathrm{s}^{-1}$, and the time step is 450 s.

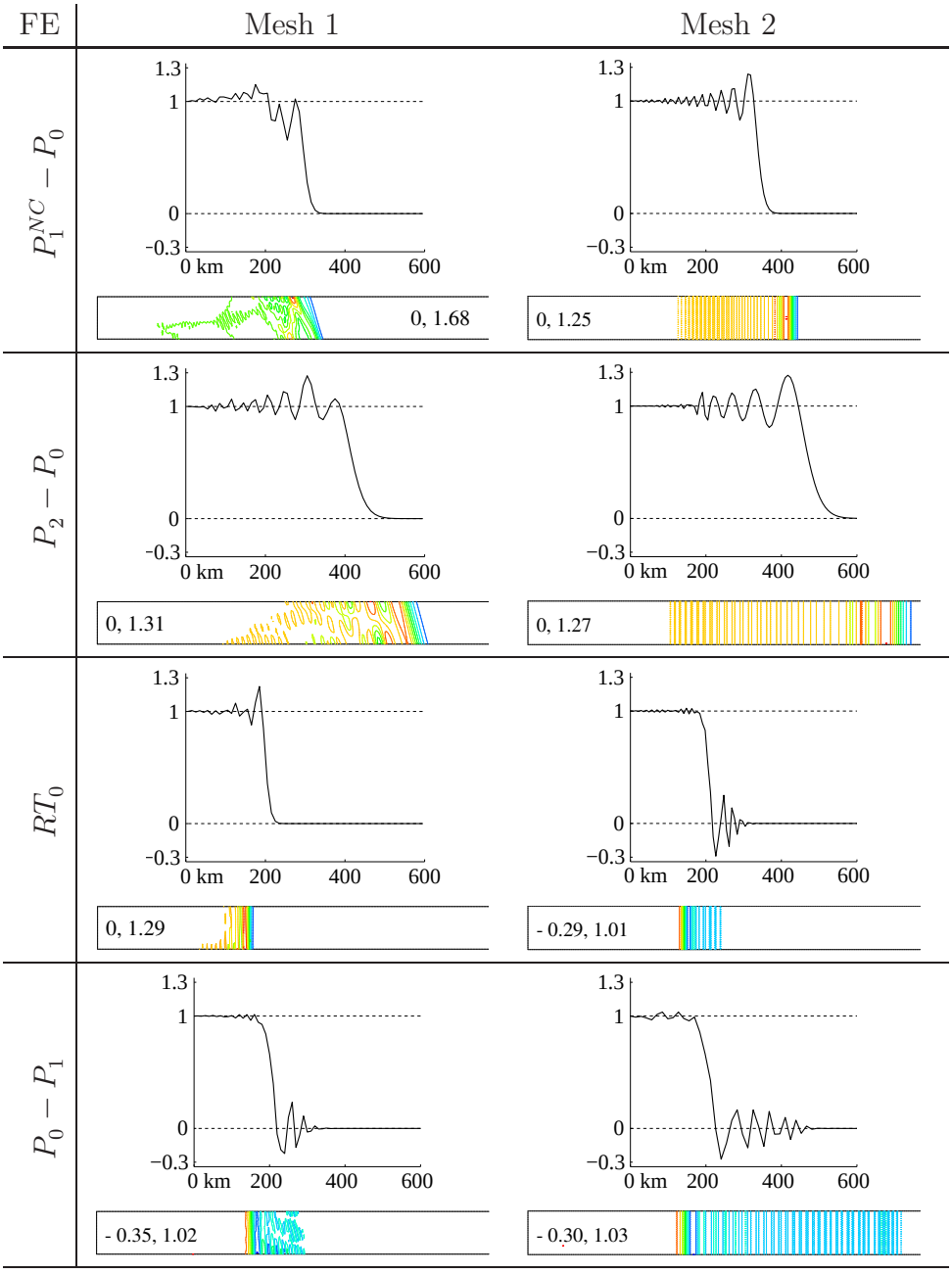


Figure 1.16: Vertical cross-sections (top) and isolines (bottom, over 2550 km) of the elevation field at times 2000 s and 10000 s of the propagation, respectively, for the $P_1^{NC} - P_0$, $P_2 - P_0$, RT0, and $P_0 - P_1$ pairs. For the isolines, the minimum and maximum values are specified in each panel.

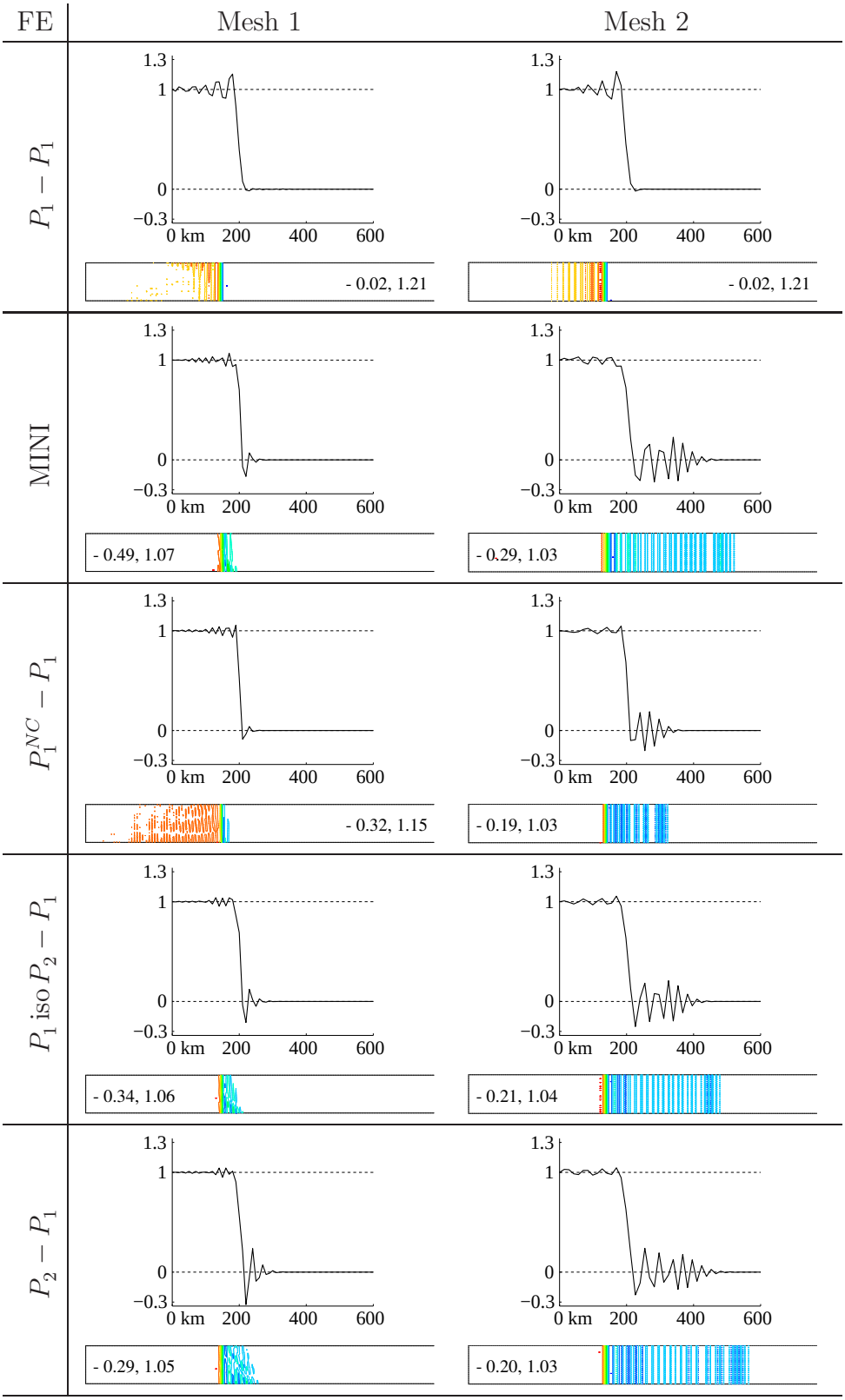


Figure 1.17: As for Figure 1.16 for the $P_1 - P_1$, MINI, $P_1^{NC} - P_1$, P_1 iso $P_2 - P_1$, and $P_2 - P_1$ pairs.

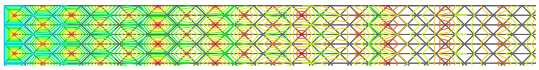


Figure 1.18: Isolines of the surface elevation on a window of Mesh 2 for the $P_1 - P_1$ pair at time 10000 s by using a linear bathymetry. As expected, a spurious mode of wavelength $3h$ is observed.

The surface-elevation field is shown after one day of simulation, which is approximately the time required for geostrophic adjustment to occur, in Figures 1.15(a) and (b). In Figure 1.15(a), for $\lambda/h = 1/6$ (coarse resolution) with $\sqrt{gH} = 11/3\,\mathrm{m\,s^{-1}}$, and also for $\lambda/h \leq 1/2$, the RT_0 scheme exhibits a checkerboard-like pattern of noise in the elevation field around the mass source and sink points, shown by an irregular alternation of positive and negative height areas (that are constant per triangle). Note that such a pattern is not observed for the other pairs. Due to the symmetry, only the result around the point mass source is shown. We believe the checkerboard-like appearance is probably due to the distortion observed in Figure 1.13 in the dispersion relation for small values of λ/h (coarse resolution). In Figure 1.15(b) for $\lambda/h = 2$ (high resolution) with $\sqrt{gH} = 44\,\mathrm{m\,s^{-1}}$, the surface-elevation distribution exhibits a regular and expected pattern as for the other pairs, except for the $P_1 - P_1$ one (due to the spurious surface-elevation mode). We also conducted the experiment over 10 days, and the result is shown in Figure 1.15(c) for $\lambda/h = 1/12$ (coarse resolution) with $\sqrt{gH} = 11/6\,\mathrm{m\,s^{-1}}$. The checkerboard-like pattern is more apparent than in Figure 1.15(a), particularly when considering the alternation of blue and yellow triangles surrounding the four centered ones (red and orange) in the OX, OY, and OD2 directions. Note that in all cases the gravitational Courant number is small (≤ 0.1) and the use of a forward Euler and a Crank–Nicolson time stepping schemes yields essentially the same results.

In the second test, (1.1) and (1.2) are solved using $f = 0$. The purpose of the experiment is to validate the analytical results obtained in section 4.1 on two meshes, Mesh 1 and Mesh 2, shown in Figure 1.14. Mesh 1 is made up of biased right triangles and Mesh 2 is obtained from Mesh 1 by a rotation of $\pi/4$. Such meshes are used in order to simulate gravity-wave propagation: along the OX and OY directions for Mesh 1 and along the OD2 direction for Mesh 2. The domain extent is 7000 km $\times$ 280 km for Mesh 1 and $5000\sqrt{2}$ km $\times$ $200\sqrt{2}$ km for Mesh 2. In both cases the resolution is $h = 10$ km, along the x- or y-axis for Mesh 1 and along the diagonal for Mesh 2. The fluid is initially at rest and zero normal velocity is specified at the lateral boundaries, except at the western one. The conditions $\eta = 1$ m and $u = 0.1\,\mathrm{m\,s^{-1}}$ (i.e., $u = \sqrt{\frac{g}{H}} \times 1$ m) are prescribed at each time step at the western boundary, with $H = 1000$ m and $g = 10\,\mathrm{m\,s^{-2}}$, i.e., $\sqrt{gH} = 100\,\mathrm{m\,s^{-1}}$. Note that the velocity has been imposed weakly at the western boundary for the $P_0 - P_1$ pair. The time step is set to 10 s, and the gravitational Courant number is thus 0.1.

In Figures 1.16 and 1.17 the vertical cross-sections (on the x, z-plane) and the isolines of the wave propagation are displayed for graphical convenience, after 2000 s and 10000 s, respectively, on the two meshes. For all pairs, except the $P_1^{NC} - P_0$ and $P_2 - P_0$ ones, the wave front propagates at the analytical speed $100\,\mathrm{m\,s^{-1}}$. This is because r_{PH} is close to 1 for small values of $|kh|$ and $|lh|$ in all directions as shown in Figures 1.7 and 1.8. However, the maximum values observed for r_{PH} in the OX and OD2 directions are 1.5 and 1.73, respectively, for the $P_1^{NC} - P_0$ pair, and 2.18 and 2.45, respectively, for the $P_2 - P_0$ one. For the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs, the wave front speeds are thus expected to be $150\,\mathrm{m\,s^{-1}}$ and $218\,\mathrm{m\,s^{-1}}$, respectively, in the OX direction, and $173\,\mathrm{m\,s^{-1}}$ and $245\,\mathrm{m\,s^{-1}}$, respectively, in the OD2 direction. These analytical values are in good agreement with the computed ones, observed in Figures 1.16 and 1.17. Indeed, for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs, the computed wave front speeds are $141\,\mathrm{m\,s^{-1}}$ and $209\,\mathrm{m\,s^{-1}}$, respectively, on Mesh 1, and $172\,\mathrm{m\,s^{-1}}$ and $242\,\mathrm{m\,s^{-1}}$, respectively, on Mesh 2, with a mesurement error of about $2\,\mathrm{m\,s^{-1}}$. The loss of symmetry observed in the front propagation for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs on Mesh 1 is likely due to a faster propagation in the OD2 direction. This explains the observed southeast orientation of the front and hence its partial propagation in the OD1 direction. Consequently, due to the results of Figure 1.7 in the OD1 direction, the computed wave speed is slightly smaller than the analytic one on Mesh 1. On Mesh 2, because the faster (dominant) propagation is along the OD2 direction, the front is not tilted and the analytical and computed values nearly coincide.

There is also a good agreement between the numerical dispersion observed in Figures 1.7 and 1.8 and the simulation results shown in Figures 1.16 and 1.17, except for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs due to the incorrect position of the front. Indeed, in the OX, OY, and OD1 directions we have $r_{PH} \simeq 1$ for the $P_1^{NC} - P_1$ pair in Figure 1.8(c), and a small amount of numerical dispersion is observed in Figure 1.17 on Mesh 1. For the MINI, P_1 iso $P_2 - P_1$, $P_2 - P_1$, and $P_0 - P_1$ pairs in the previous directions, r_{PH} progressively increases from 1.03 (MINI) up to 1.2 ($P_0 - P_1$) as observed in Figures 1.7 and 1.8. Hence, we have $|c_{CP}| \geq |c_{AN}|$ and increasing dispersion is noticeable upstream the front in Figures 1.16 and 1.17 for these pairs on Mesh 1. For the RT_0 and $P_1 - P_1$ pairs in the previous directions, we have $r_{PH} \leq 1$ in Figures 1.7 and 1.8, and hence $|c_{CP}| \leq |c_{AN}|$, leading to dispersion effects downstream the front as shown in Figures 1.16 and 1.17 on Mesh 1. In the OD2 direction, we still have $r_{PH} \leq 1$ for the $P_1 - P_1$ pair, and oscillations are again present downstream the front on Mesh 2 as shown in Figure 1.17. For the RT_0 pair we now have $r_{PH} \geq 1$ in the OD2 direction (except for small wavelengths), and the dispersion effects are noticeable upstream the front in Figure 1.16 on Mesh 2. For the $P_1^{NC} - P_1$, P_1 iso $P_2 - P_1$, MINI, $P_2 - P_1$, and $P_0 - P_1$ pairs in the OD2 direction, r_{PH} progressively increases from 1.15 ($P_1^{NC} - P_1$) up to 1.7 ($P_0 - P_1$) as observed in Figures 1.7 and 1.8. We thus have $|c_{CP}| \geq |c_{AN}|$, and increasing dispersion is noticeable

for these pairs upstream the front in Figures 1.16 and 1.17 on Mesh 2.

Spurious elevation modes are not observed in Figure 1.17 for the $P_1 - P_1$ pair. This is because a constant value of H is used. However, when $H = 1000$ m if $x \leq 400$ km, when $H = 800$ m if $x \geq 600$ km, and by assuming H is linear if 400 km $\leq x \leq 600$ km, a spurious surface elevation mode of wavelength $3h$ is triggered, as shown in Figure 1.18 on Mesh 2. This agrees well with the result predicted by theory in Figure 1.8(a) in the OD2 direction.

Finally, the second test has also been performed on unstructured meshes. Two main features are observed for all pairs: the numerical dispersion upstream the front is more significant than on Mesh 1 but less than on Mesh 2, and additional noise is generated downstream the front. Further, the loss of symmetry observed in the front propagation in Figure 1.16 for the $P_1^{NC} - P_0$ and $P_2 - P_0$ is only marginally present.

1.8 Conclusion

This appears to be the first study of the dispersion relation and spurious mode behavior for FE solutions of the two-dimensional linearized SW equations based on the examination of a variety of mixed FE pairs. For each pair the frequency wave number or dispersion relation is obtained and analyzed, and the dispersion properties are compared analytically and graphically with the continuous case to illustrate the main points of interest. Two numerical tests, concerning the propagation of gravity waves, are performed. In the first experiment, the simulation of the geostrophic adjustment process, the RT_0 element exhibits a checkerboard-like pattern of noise in the elevation field when the grid resolution is low relative to the Rossby radius of deformation. Such a pattern has not been observed for the other pairs. The second test illustrates the analytical dispersion results by simulating the propagation of pure gravity waves in a long channel. In particular, for the $P_1^{NC} - P_0$ and $P_2 - P_0$ pairs, the wave front propagates much faster than the analytical speed. These results can be compared with those in [58] for the geostrophic balance. We also notice the presence of spurious frequencies of type $O(\frac{1}{h})$ for the FE pairs having a piecewise-constant representation of the surface elevation. Such solutions have already been observed in [53], where the one-dimensional SW equations are discretized using the discontinuous Galerkin method, and they will be analyzed in a further study.

The preceding analysis illustrates how phase and group velocity accuracy can help in the selection of a spatial discretization scheme. In particular, the $P_1^{NC} - P_1$ and RT_0 pairs have been identified as a promising compromise for the discretization of the inviscid

linear SW equations, provided the grid resolution is high relative to the Rossby radius of deformation for the RT_0 element. However, because of its restrictive nature, due to the use of constant values for h and H, such an analysis should be only one step of the selection process. The use of nonconstant h and H will be analyzed in subsequent studies. Further, the implementation of a time stepping technique can change the relative accuracy and merits of two spatial discretization schemes.

2. KERNEL ANALYSIS OF THE DISCRETIZED FINITE DIFFERENCE AND FINITE ELEMENT SHALLOW-WATER MODELS

V. Rostand, D. Y. Le Roux, and G. F. Carey
Submitted to SIAM Journal on Scientific Computing (2007).
Keywords: shallow-water equations, finite element method, Rossby waves, vortex flows

2.1 Résumé

Une approche constructive basée sur l'algèbre linéaire est utilisée pour caractériser les noyaux des opérateurs discrets des équations de Saint-Venant. Trois conditions nécessaires sont identifiées afin que les modes stationnaires du système discret partagent les propriétés du système continu. Les noyaux sont calculés avec le logiciel MATLAB et sont utilisés pour déduire la présence, le nombre et la structure des modes parasitaires résultant des principaux schémas de discrétisation aux différences finies et d'éléments finis. Les noyaux sont également utilisés pour caractériser la forme des plus petis vortex discrets. Pour les éléments finis, les cas des maillages structurés et non-structuré sont analysés et comparés. Le problème du découplage des amplitudes est mis en évidence dans le problème de l'approximation de la balance géostrophique.

2.2 Abstract

A constructive linear algebra approach is developed to characterize the kernels of the discretized shallow-water equations. Three kernel relations are identified as necessary conditions for the discretized system to share the same stationary properties as the continuous system. This matrix kernel scheme is implemented symbolically using MATLAB and applied to investigate the presence, number, and structure of spurious modes aris-

ing in typical finite difference and finite element schemes. The kernel concept is then used to characterize the smallest representable vortices for several representative discrete finite difference and finite element schemes. Both uniform and unstructured mesh situations are considered and compared. Issues such as decoupling of vortex modes are also examined.

2.3 Introduction

Numerical oscillations associated with spurious solution modes are a familiar occurrence in computer simulations of shallow-water (SW) and Navier Stokes problems. These oscillatory modes are usually associated with the choice of discretization model. In coupled mixed formulations, the modes arise out of inconsistencies in the discrete model relative to the continuous problem. Some familiar examples, for both the finite difference (FD) and finite element (FE) discretizations, are pressure oscillations in viscous flow calculations and elevation or velocity oscillations in the shallow-water problem [14, 34, 35, 58, 65]. There are several means for investigating the consistency of the discrete schemes. A familiar strategy for shallow-water schemes is to apply dispersion analysis [6, 56, 47, 65]. In the present work, we develop a related kernel analysis and constructive computational approach for the shallow-water system in which the properties of the kernel of the associated discretized problem are used to determine the presence and number of different types of spurious modes. The kernel characterization may also be used to determine the form of the smallest representative velocity vortex structures that can be represented in a given discretization. This, in turn, enables one to examine the nature of different modes and also sheds light on modal decoupling and related issues.

The treatment is developed as follows: The model equations and stationary solutions are presented in Section 2.4. The discretization of the stationary solutions is performed in Section 2.5 and the kernels of the discretized equations are characterized and analysed in Section 2.6 . Numerical experiments to validate the analytic results are performed in Section 2.7 . Some concluding remarks complete the study.

2.4 Governing equations and stationary solutions

The inviscid linear SW equations in a 2D domain Ω may be expressed as [37]

$$\frac{\partial \mathbf{u}}{\partial t} + f\mathbf{k} \times \mathbf{u} + g\nabla\eta = 0, \tag{2.1}$$

$$\frac{\partial \eta}{\partial t} + H\nabla \cdot \mathbf{u} = 0. \tag{2.2}$$

where $\mathbf{u}(\mathbf{x}, \mathbf{t}) = (u, v)$ is the horizontal velocity field with $\mathbf{x} = (x, y)$, η is the surface elevation with respect to the reference level $z = 0$, g is the gravitational acceleration, and $\mathbf{k}$ is a unit vector in the vertical direction. The Coriolis parameter f and the mean depth H are assumed constant. Note that η plays a role similar to that of pressure in the Navier–Stokes equations. The governing equations (2.1) and (2.2) describe a first order hyperbolic system. Initial conditions at $t = t_0$ and boundary conditions complete the mathematical statement of the problem. Here we consider periodic and contained flows. For instance, in a contained flow, the no-normal flow condition on boundary Γ is simply

$$\mathbf{u}(\mathbf{x}, t) \cdot \mathbf{n} = 0, \quad \text{on } \Gamma, \forall t \geq t_0, \tag{2.3}$$

where $\mathbf{n}$ is the outward pointing normal to the boundary.

The stationary solutions of (2.1) and (2.2) satisfy

$$f\mathbf{k} \times \mathbf{u} + g\nabla\eta = 0, \tag{2.4}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{2.5}$$

and are of central importance for large scale models, particularly in atmospheric and ocean modelling. For instance, equation (2.4) represents geostrophic equilibrium, a balance between the Coriolis and the pressure gradient terms. Such a balance is very important for both atmospheric and oceanic flows. When this near-equilibrium is disturbed, oceanic and atmospheric flows adjust to a near-geostrophic state. This takes place as a rapid adjustment process via inertia-gravity waves accompanied by a much slower adjustment through the propagation of planetary (Rossby) waves whose existence is due to the variation of the Coriolis parameter with latitude. With respect to cyclone-scale motions in extra-tropical latitudes, the free atmosphere frequently approaches a state of geostrophic equilibrium. Oceanic flows are also largely geostrophic.

We deduce from (2.4) that

$$\mathbf{u} = \frac{g}{f}\mathbf{k} \times \nabla\eta = \frac{g}{f}\mathbf{rot}\,\eta, \tag{2.6}$$

with $\mathbf{rot}\,\eta \equiv (-\eta_y, \eta_x)$, and it follows that (2.5) is automatically satisfied. Equation (2.6) permits us to associate a stationary solution to a given elevation field.

On the other hand, if $(\mathbf{u}, \eta)$ is a stationary solution then, $\mathbf{u}$ satisfies (2.5) and, applying the fundamental theorem of vector calculus, can be expressed as

$$\mathbf{u} \;=\; \mathbf{rot}\,\zeta, \tag{2.7}$$

where ζ is a scalar field. Substitution of (2.7) in (2.4) leads to

$$\nabla\eta = -\frac{f}{g}\mathbf{k} \times \mathbf{rot}\,\zeta = \frac{f}{g}\nabla\zeta, \tag{2.8}$$

and hence

$$\eta = \frac{f}{g}\zeta + c, \tag{2.9}$$

where c is a constant. That is, η is determinable within an arbitrary constant c. It follows that for every such elevation field η, there is one and only one stationary solution $\mathbf{u}$ such that (2.4) and (2.5) are satisfied. The constant c corresponds to the physical hydrostatic mode. This also implies that a specified admissible elevation field η can be used to construct the associated velocity field and hence a stationary solution pair for the shallow-water equations.

Finally, the boundary condition (2.3) implies that such elevation fields η have to satisfy

$$\mathbf{rot}\,\eta \cdot \mathbf{n} = \nabla\eta \cdot \mathbf{t} = 0 \quad \text{on } \Gamma, \tag{2.10}$$

where $\mathbf{t}$ is the tangent vector to Γ.

2.5 *Discretization of the stationary SW equations*

Equations (2.4) - (2.5) are now spatially discretized. The finite-difference approximation on a regular mesh is obtained by approximating derivatives as difference quotients using Taylor series expansions at the grid points in the standard manner. The finite-element discretization on structured or unstructured grids is obtained by introducing finite element approximations in a weak integral formulation. We describe the finite element formulation in more detail below since it will be utilised later for different formulations with various velocity and elevation fields.

2.5.1 Weak formulation

Weak formulations of (2.4) and (2.5) may be constructed by projecting the associated residuals against admissible test functions $\boldsymbol{\phi}$ and θ:

$$\int_\Omega f(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\phi}\,d\mathbf{x} + g\int_\Omega \nabla\eta\cdot\boldsymbol{\phi}\,d\mathbf{x} = 0, \tag{2.11}$$

$$\int_\Omega \nabla\cdot\mathbf{u}\,\theta\,d\mathbf{x} = 0. \tag{2.12}$$

Depending on the regularity of functions $\mathbf{u}$ and η, the terms containing derivatives in (2.11) and (2.12) may be integrated by parts to obtain

$$\int_\Omega f(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\phi}\,d\mathbf{x} - g\int_\Omega \eta\,\nabla\cdot\boldsymbol{\phi}\,d\mathbf{x} = 0, \tag{2.13}$$

and

$$\int_\Omega \mathbf{u}\cdot\nabla\theta\,d\mathbf{x} = 0, \tag{2.14}$$

where the boundary integral contributions arising from the integration by parts procedure have been set to zero to accommodate either periodic boundary conditions (with respect to opposite sides) or the natural boundary conditions $\mathbf{u}\cdot\mathbf{n} = 0$ on Γ. Approximation of both integral forms will be considered later. Let $\mathbf{V}$ and W, denote the velocity and elevation spaces for a given weak formulation with the test functions chosen accordingly; e.g. the square-integrable space $L^2(\Omega)$ or the Hilbert space $H^1(\Omega)$ of functions in $L^2(\Omega)$ whose first derivatives belong to $L^2(\Omega)$.

2.5.2 Finite-element discretization

Consider a standard finite-element partition (e.g. a triangulation) $\mathcal{T}_h$, of the polygonal domain Ω, where h denotes a representative meshlength parameter. For element (triangle or quadrilateral) $K\in\mathcal{T}_h$, let $P_r(K)$ denote the space of polynomials of degree r on K.

Then the approximate solutions $\mathbf{u}_h$ and η_h sought belong to finite-dimensional spaces $\mathbf{V}_h$ and W_h, respectively, whose restrictions on K belong to $P_r(K)\times P_r(K)$ for $\mathbf{u}_h$ and

to $P_s(K)$ for η_h. The components of $\mathbf{u}_h$ and η_h are represented over an element K by interpolating polynomials of degree r and degree s, respectively:

$$\mathbf{u}_h = \sum_{i \in S_K^{\mathbf{u}}} \mathbf{u}_i \, \phi_i, \quad \eta_h = \sum_{i \in S_K^{\eta}} = \eta_i \, \theta_i, \tag{2.15}$$

where i represents a node of K and $S_K^{\mathbf{u}}$ and S_K^{η} denote the set of nodes of K.

The global piecewise polynomial expansions of $\mathbf{u}_h$ and η_h over the whole domain Ω follow in the standard way from (2.15) applied over the union of all triangles K of the triangulation $\mathcal{T}_h$. Setting the previous two forms of integral statements on the finite element subspaces leads to corresponding finite element approximate formulations with $\mathbf{u}, \eta$ replaced by the finite element trial functions $\mathbf{u}_h, \eta_h$ and ϕ, θ replaced by the corresponding finite-element test functions.

For FE pairs having a discontinuous representation of velocity, the term $\nabla \cdot \mathbf{u}$ appearing in (2.12) is integrated by parts to avoid computing u and v derivatives, and (2.14) is used instead. For FE pairs having a discontinuous representation of pressure, an integration by parts of $\nabla\eta$ in the left hand side of (2.11) is performed and hence (2.13) is employed.

In the discrete formulation arising from (2.11) and (2.12), the element matrix entries are

$$c_{i,j} = \int (\mathbf{k} \times \boldsymbol{\phi_i}) \cdot \boldsymbol{\phi_j} \, d\Omega, \quad g_{i,j} = \int \nabla\theta_i \cdot \boldsymbol{\phi_j} \, d\Omega, \quad d_{i,j} = \int \nabla \cdot \boldsymbol{\phi_i} \, \theta_j \, d\Omega, \tag{2.16}$$

for the respective the Coriolis, gradient and divergence terms. When (2.13) is employed instead of (2.11) the element gradient matrix is then $-d_{j,i}$, and if (2.14) is used the element divergence matrix becomes $-g_{j,i}$.

Assembling the element matrices leads to a global linear system of the form

$$\begin{pmatrix} C & G \\ D & O \end{pmatrix} \begin{pmatrix} \mathbf{u}_i \\ \eta_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{2.17}$$

where we have written the solution vector in the form $(\mathbf{u}_i, \eta_j)^T$, with $i = 1, ..., p$ and $j = 1, ..., q$ to indicate the order of nodal variables. The integers p and q represent the number of velocity and surface-elevation nodes, respectively. C, G, D are the global Coriolis, gradient and divergence submatrices and are of dimension $p \times p$, $p \times q$ and $q \times p$, respectively; O denotes the null matrix of dimension $q \times q$. Note that periodic or no-normal flow boundary conditions are assumed in (2.17).

2.5.3 Kernel properties

The global matrix on the left hand side of (2.17) is termed CDG in the remaining part of this paper. Let

$$CD \equiv \begin{pmatrix} C \\ D \end{pmatrix}, \qquad CG \equiv \begin{pmatrix} C & G \end{pmatrix}, \qquad DO \equiv \begin{pmatrix} D & O \end{pmatrix}. \tag{2.18}$$

The discrete stationary modes belong to the kernel of CDG as they are the solutions of (2.17). Further, the kernel of CDG includes the kernel of CD and G in the following sense: if $\mathbf{u}$ belongs to $\ker(CD)$ and if η belongs to $\ker(G)$ then

$$\begin{pmatrix} \mathbf{u} \\ 0 \end{pmatrix} \in \ker(CDG) \qquad \text{and} \qquad \begin{pmatrix} 0 \\ \eta \end{pmatrix} \in \ker(CDG). \tag{2.19}$$

The solutions of (2.19) are named CD-modes and G-modes, respectively. Note that the G-modes always include the physical hydrostatic mode which corresponds to the mode with constant elevation and zero velocity. Additional G-modes would be spurious and analogous to spurious pressure modes arising from inconsistent choice of bases in the Stokes and Navier-Stokes discretized systems.

The kernel of CDG can also be written as

$$\ker(CDG) = \ker(CG) \ \cap \ \ker(DO), \tag{2.20}$$

where

$$\ker(DO) = \{ \ (\mathbf{u}, \eta)^t \mid \mathbf{u} \in \ker(D) \ , \ \eta \in \mathbf{R}^q \ \}. \tag{2.21}$$

In the previous section it was shown that an analytical solution (u, η) of (2.4) was a solution of (2.5). We would like to verify if such a property also holds for the discrete solution (u_i, η_j) of (2.17); that is, if

$$\ker(CG) \subset \ker(DO). \tag{2.22}$$

On the other hand, it was shown that a solution of (2.5) also satisfies (2.4). Again, it is important to verify that the discrete analogue holds. That is,

$$\ker(D) \subset \{\mathbf{u}| \ \exists \eta \text{ such that } (\mathbf{u}, \eta) \in \ker(CG)\}, \tag{2.23}$$

is true. Further, it was demonstrated by construction that the SW equations admit a stationary solution for every given admissible elevation field. At the discrete level, this would imply:

$$\forall \eta, \ \exists \mathbf{u} \text{ such that } (\mathbf{u}, \eta) \in \ker(CDG). \tag{2.24}$$

Verifying (2.22), (2.23) and (2.24) is crucial in order to determine if the discretization schemes share the stationary properties of the continuous system. We investigate this question in the following section.

2.5.4 The FD grids and FE pairs

In subsequent tests we consider several popular choices of finite difference stencils and finite element pairs proposed in shallow-water simulations. Note, however, that our matrix kernel approach applies more generally to other element choices and to similar difference schemes.

FD grids

The three candidate finite-difference grids, shown in Figure 2.1, are the classical B and C grids [41] and the C-D grid [1]. The symbols • and black arrows indicate the location of velocity and normal velocity nodes, respectively. The arrow points in the direction of the chosen normal. The symbol ○ indicates nodes for surface-elevation. For all schemes the elevation field is carried at the center of the computational cells, and the elevation and velocity variables are staggered in space. For the B grid the two components of the velocity field are located at the cell corners, while the normal velocities are carried at cell edge midpoints for the C grid. To circumvent the Coriolis modes of the C grid in the inertial limit, due to the spatial averaging of the Coriolis terms, the C-D grid employs a combination of the C and D grids such that the two components of the velocity field are located at cell edge midpoints.

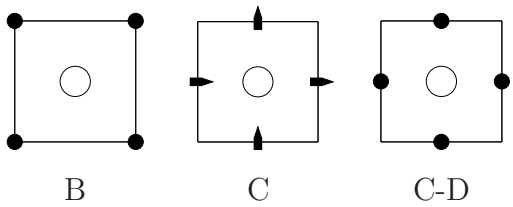


Figure 2.1: The three FD grids employed in this study. The symbols • and black arrows indicate the location of velocity and normal velocity nodes, respectively. The symbol ○ indicates nodes for surface-elevation.

FE pairs

Seven FE velocity / surface-elevation pairs are considered here for the spatial discretization and are shown in Figure 2.2. Conventional finite-element terminology is used to describe the FE pairs. The nomenclature $P_r - P_s$ means that velocity components and

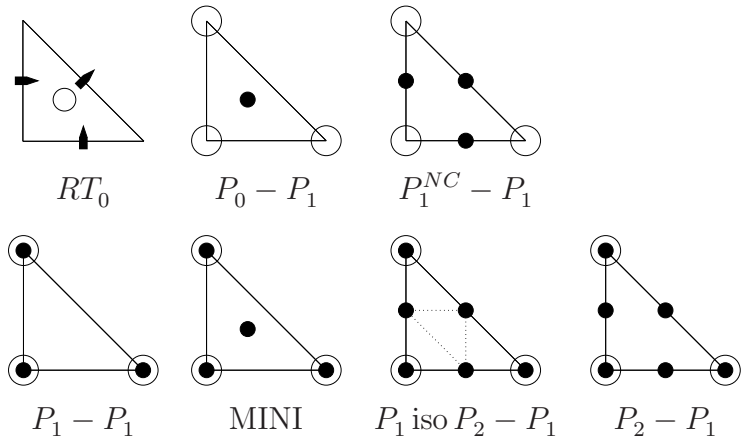


Figure 2.2: As for Figure 2.1 but for the seven FE pairs employed in this study.

surface elevation are respectively represented as piecewise-defined polynomials of degree r and s. Common to the last six FE pairs is a piecewise-linear continuous representation of surface elevation and they differ from one another in their representation of velocity. The $P_0 - P_1$, $P_1 - P_1$ and $P_2 - P_1$ pairs have constant, linear and quadratic velocities, respectively. The $P_1 \,\mathrm{iso}\, P_2 - P_1$ pair [12] has piecewise-linear basis functions for velocity on a refined triangulation obtained by dividing each triangle into four similar subtriangles using the midpoints of triangle sides. There are thus 6 velocity nodes over each unrefined triangle, the same as for a quadratic approximation of velocity, termed P_2. The designation P_1 in $P_1 \,\mathrm{iso}\, P_2$ denotes linear velocity elements on subtriangles, whereas $\mathrm{iso}\, P_2$ indicates that the nodal placement is that associated with quadratic elements on unrefined triangles. The MINI element [8] also has continuous piecewise-linear basis functions at the vertices but bubble functions are added at the barycenters for velocity in order to stabilize the pair. The $P_1^{NC} - P_1$ pair [34, 52, 56], has velocity nodes at triangle edge midpoints and linear basis functions are used to approximate the two velocity components on the pair of elements adjacent to the associated edge. Since this particular representation of velocity is only continuous across triangle boundaries at midpoint nodes, and discontinuous everywhere else around a triangle boundary, this element is termed non-conforming (NC) in the finite-element literature. Finally, RT_0 corresponds to the low order Raviart-Thomas element [48]. It has normal velocity components at triangle midedge points and a discontinuous piecewise-constant representation of surface-elevation.

Table 2.1: Dimension of the discrete operator kernels on an $m \times n$ grid made up of biased triangles for the FD grids and FE pairs with periodic boundary conditions.

	p	q	C	G	D	CD	CG	CDG
B-grid	$2mn$	nm	0	β	$nm+\beta$	0	nm	nm
C-grid	$2mn$	nm	2α	1	$nm+1$	α	$nm+\alpha$	nm
C-D-grid	$4mn$	nm	0	1	$3nm+1$	0	nm	nm
RT_0	$3nm$	$2nm$	nm	1	$nm+1$	δ	$2nm+\delta$	nm
P_0-P_1	$4nm$	nm	0	1	$3nm+1$	0	nm	nm
$P_1^{NC}-P_1$	$6nm$	nm	0	1	$5nm+1$	0	nm	nm
P_1-P_1	$2nm$	nm	0	γ	$nm+\gamma$	0	nm	nm
MINI	$6nm$	nm	0	1	$5nm+1$	0	nm	nm
P_1 iso P_2-P_1	$8nm$	nm	0	1	$7nm+1$	0	nm	nm
P_2-P_1	$8nm$	nm	0	1	$7nm+1$	0	nm	nm

2.6 Numerical stationary solutions

In this set of analytical studies of local kernel structure for various basis choices, we consider a rectangular mesh bisected consistently to form biased right triangle pairs. The underlying cartesian mesh resolution is m by n, where m and n are the number of segments in the x and y directions, respectively, so the total number of triangles is $2nm$.

2.6.1 Kernel dimension for periodic boundary conditions

First consider the case of periodic boundary conditions. The SW equations are discretized using physical parameters that represent classical oceanic parameters at mid-latitudes. However, the precise choice of these parameters does not affect the following results as long as the parameters do not degenerate to zero. The grid resolution is set such that m and n range between 2 and 12 and the computations were performed under the MATLAB environment. On each grid, with $m, n = 2, 3, ..., 12$ the dimension of the discrete operator kernels C, G, D, CD, CG and CDG are obtained and the extrapolated results for any m and n are given in Table 2.1. For the B and C-D grids and the last six FE pairs in Table 2.1 $\dim(\ker(C))$ is always zero. This is because both components of the velocity field are computed at velocity nodes for these schemes. Indeed, by ordering the x-components first and then the y-components, the Coriolis matrix has

the form

$$C = \begin{pmatrix} 0 & -M \\ M & 0 \end{pmatrix}, \tag{2.25}$$

where M is the velocity mass matrix. Since M is full-rank we deduce from (2.25) that C is also full-rank. This implies that $\dim(\ker(C)) = 0$. However, the same reasoning does not apply for the C grid and the RT_0 element because only one component of the velocity field (the normal component) is computed at velocity nodes for these schemes. Hence, a spatial averaging of the Coriolis terms is needed to evaluate the second component of the velocity field (the tangential one) and this leads to a spurious Coriolis mode, that is also termed an f-mode in [66]. The results of Table 2.1 for the C grid show that $\dim(\ker(C)) = 2\alpha$, with

$$\alpha = \begin{cases} n+m-1 & \text{if } 2 \mid n \text{ and } 2 \mid m. \\ m & \text{if } 2 \mid n \text{ and } 2 \nmid m, \\ n & \text{if } 2 \nmid n \text{ and } 2 \mid m, \\ 0 & \text{if } 2 \nmid n \text{ and } 2 \nmid m, \end{cases} \tag{2.26}$$

where the symbols $\mid$ and $\nmid$ imply the integers are divisible by 2 (even) or not divisible (odd), respectively. For the RT_0 element, we have $\dim(\ker(C)) = mn$, i.e. exactly $p/3$.

As expected, the dimension of the gradient matrix G is always equal to or greater than one. One mode is the hydrostatic mode, which can be simply considered as a constant of integration associated with the solution of the governing equations. For the continuous mathematical problem the hydrostatic mode is the only one present but this may not be the case in the discretized problem. For example, in the case of the B grid and the $P_1 - P_1$ pair, as shown in Table 2.1, additional modes may occur. The number of such modes for the B grid is

$$\beta = \begin{cases} 2 & \text{if } 2 \mid n \text{ and } 2 \mid m, \\ 1 & \text{otherwise.} \end{cases} \tag{2.27}$$

and is given in Table 2.2 for the $P_1 - P_1$ scheme. The structure of the latter modes is periodic with a periodicity of $2h$ or $3h$ [56]. These modes are small-scale artifacts introduced by the spatial discretization scheme. They do not propagate but are trapped within the grid, and associated with zero frequency. If these spurious solutions are left undamped, they can cause an accumulation of energy in the smallest-resolvable scale, leading to noisy approximations. The occurrence of such spurious modes has been observed in a variety of finite-difference [66] and finite-element [58, 51, 57, 65] approximations to the SW equations.

Table 2.2: Dimension γ of the gradient operator kernel on an $m \times n$ grid made up of biased triangles with periodic boundary conditions for the $P_1 - P_1$ pair.

	$2 \nmid n$ $3 \nmid n$	$2 \mid n$ $3 \nmid n$	$2 \nmid n$ $3 \mid n$	$2 \mid n$ $3 \mid n$
$2 \nmid m$ and $3 \nmid m$	1	2	1	2
$2 \mid m$ and $3 \nmid m$	2	4	2	4
$2 \nmid m$ and $3 \mid m$	1	2	3	4
$2 \mid m$ and $3 \mid m$	2	4	4	6

The discrete gradient and divergence operators are linked by the following relation

$$D = -G^t, \tag{2.28}$$

and using the rank theorem we deduce

$$\dim(\ker(D)) = p - q + \dim(\ker(G)). \tag{2.29}$$

The value $\dim(\ker(D))$ only depends on p and q if spurious pressure modes are not present. In this case, those discretization schemes having identical p and q numbers also have the same value for $\dim(\ker(D))$, as seen for the $P_1^{NC} - P_1$ and MINI pairs in Table 2.1, for example.

Further, since the Coriolis matrix is skew-symmetric, equation (2.28) leads to

$$CG = -(CD)^t, \tag{2.30}$$

and applying the rank theorem we obtain

$$\dim(\ker(CG)) = q + \dim(\ker(CD)). \tag{2.31}$$

The kernel of the CD matrix is the intersection of the C and D matrix kernels. Hence, there is no CD-mode for any of the six last finite-element pairs in Table 2.1 because the C matrix is full rank for these pairs. We then deduce that

$$\dim(\ker(CG)) = q. \tag{2.32}$$

For the RT_0 element, the situation is different as shown in Table 2.1. Indeed, the kernel of the C matrix has a non-trivial intersection with the kernel of D on particular grids and

$$\delta = \begin{cases} 2 & \text{if } 3 \mid n \text{ and } 3 \mid m, \\ 0 & \text{otherwise.} \end{cases} \tag{2.33}$$

Table 2.1 shows that $\dim(\ker(CG))$ is precisely mn for all schemes that are considered here, except for the C grid (when $\alpha \neq 0$) and the RT_0 element, which have Coriolis modes. This number mn exactly coincides with $\dim(\ker(CG)) = q$; i.e. the CG matrix is full rank and hence, the inclusion (2.22) holds. However, for the C grid (when $\alpha \neq 0$) and the RT_0 element we have $\dim(\ker(D))$ smaller than $\dim(\ker(CG))$, which immediately negates the desired inclusion (2.22).

The stationary modes of (2.17) correspond to the kernel of the CDG matrix which is the intersection of the CG and DO matrix kernels. Table 2.1 shows that $\dim(\ker(CDG)) = mn = p$ for all schemes, and hence (2.24) holds, except for the RT_0 pair since $mn = p/2$.

At this stage of the argument it can be concluded that only the B and C grids (if $\alpha = 0$) and the $P_1 - P_1$ FE pair satisfy the three properties (2.22), (2.23) and (2.24). For those schemes, $p = 2q$ and hence there are the same number of equations for the discrete variables u, v and η, as for the continuous system.

2.6.2 Kernel dimension using no-normal flow boundary condition

The no-normal flow boundary condition (2.3) implies that η has to satisfy (2.10). Here, we investigate the numerical influence of this boundary condition at the discrete level and its impact on the stationary modes. As previously, biased right triangles are again used and the dimension of the discrete operators is computed on grids of size $m \times n$, with $m, n = 2, 3, ..., 14$. The extrapolated results for any m and n are given in Tables 2.3 and 2.4. Note that for the B and C-D grids, no slip boundary conditions are enforced to avoid the appearance of spurious boundary modes that result from the use of no normal flow conditions alone. Further, for the finite element pairs where both component of the velocity are computed at boundary nodes the condition (2.3) is enforced and solved on both parts of the momentum equation (2.11) such that only the tangential component of the flow remains. In Table 2.3 we observe that the C grid and RT_0 element are still subject to Coriolis modes while $\dim(\ker(C)) = 0$ for the other schemes and hence the C matrix is full rank. The B grid now exhibits the classical spurious pressure mode when no slip boundary conditions are used. In Table 2.1, with periodic boundary conditions, the B grid was found free of spurious modes for m or n odd integers. We also notice that $P_1 - P_1$ has no spurious pressure mode when no-normal flow boundary conditions are employed. For all other FD and FE schemes $\ker(G)$ only contains the hydrostatic mode. Finally, the values of $\dim(\ker(D))$ in Table 2.3 verify (2.29) for all FD and FE schemes.

Contrary to what has been observed in Table 2.1 for periodic boundary conditions, we

Table 2.3: Dimension of discrete operator kernels on a $m \times n$ grid made up of biased triangles for the B and C-D grids with no slip boundary condition and for the C grid and the FE pairs with no normal flow boundary condition.

	p	q	C	G	D
B grid	$2(m-1)(n-1)$	mn	0	2	$p-q+2$
C grid	$2mn-(m+n)$	mn	$2q-p-2$	1	$p-q+1$
C-D grid	$4mn-2(m+n)$	mn	0	1	$p-q+1$
RT_0	$3mn-m-n$	$2mn$	$p-q+2$	1	$p-q+1$
P_0-P_1	$4mn$	$(m+1)(n+1)$	0	1	$p-q+1$
$P_1^{NC}-P_1$	$6mn$	$(m+1)(n+1)$	0	1	$p-q+1$
P_1-P_1	$2mn-2$	$(m+1)(n+1)$	0	1	$p-q+1$
MINI	$6mn-2$	$(m+1)(n+1)$	0	1	$p-q+1$
P_1 iso P_2-P_1	$8mn-2$	$(m+1)(n+1)$	0	1	$p-q+1$
P_2-P_1	$8mn-2$	$(m+1)(n+1)$	0	1	$p-q+1$

now have $\dim(\ker(CDG)) < \dim(\ker(CG))$ in Table 2.4 for all FD and FE schemes considered, due to the use of the no-normal flow condition (and no slip for the B and C-D grids). Consequently, equation (2.24) is no longer valid. For the RT_0 element, $\dim(\ker(CDG)) = \dim(\ker(D)) + 1$. Thus, equation (2.23) holds and all numerically divergence free velocity fields can be associated with a stationary solution.

In Table 2.4 the CDG stationary modes are split into the smallest representable vortices (SRV), i.e. the elements of the CDG kernel with minimum support, and other additional modes (others). The properties of such SRV modes are investigated in the next section. The other modes do not appear to have any specific local structure and their representation varies with the grid dimension and element type. The number of such modes is specified by the positive integer $\sigma_i(m,n)$ ranging from 1 to i, and depends on the values of m and n with $i = 1, 2, 3, ...$ Note that in Table 2.1 we have $\dim(\ker(CDG)) = mn$ for all the FD and FE schemes.

For the C grid and the RT_0 element, with $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs as in Table 2.4 the number of SRV is $(m-1)(n-1)$ and thus exactly coincides with the number of surface-elevation nodes, excluding nodes lying on the domain boundary. This result is consistent with the observation that stationary modes must have a constant elevation on boundary pressure nodes. For the remaining schemes in Table 2.4 the number of SRV is only $(m-i)(n-i)$, with $i = 2, 3, 5$. Further comments are given in the next section.

Table 2.4: As for Table 2.3. The CDG stationary modes are split into the smallest representable vortices (SRV) and other additional modes (others). The integer $\sigma_i(m,n)$ ranges from 1 to i, depending on the values of m and n with $i = 1, 2, 3, ...$, and β is defined in (2.27).

	CD	CG	CDG		
			SRV		others
B-grid	0	q	$(m-2)(n-2)$	+	2
C-grid	$\beta-1$	$q+\beta-1$	$(m-1)(n-1)$	+	1
C-D-grid	0	q	$(m-2)(n-2)$	+	1
RT_0	0	q	$(m-1)(n-1)$	+	1
$P_0 - P_1$	0	q	$(m-1)(n-1)$	+	$\sigma_2(m,n)$
$P_1^{NC} - P_1$	0	q	$(m-1)(n-1)$	+	1
$P_1 - P_1$	0	q	$(m-3)(n-3)$	+	$\sigma_5(m,n)$
MINI	0	q	$(m-3)(n-3)$	+	$\sigma_4(m,n)$
P_1 iso $P_2 - P_1$	0	q	$(m-5)(n-5)$	+	$\sigma_8(m,n)$
$P_2 - P_1$	0	q	$(m-5)(n-5)$	+	$\sigma_8(m,n)$

2.6.3 Smallest representable vortices

The purpose of this section is to find a simple linear basis for the stationary solutions of (2.17). The results of the preceding section indicate a strong link between the number of triangle vertices and the dimension of the stationary SRV basis. Hence, it seems natural to associate each vertex with one element of the basis. We define $\eta^{[i]} = (0, ..., 0, 1, 0, ..., 0)$ the elevation field with zero components except at node i where the elevation is one. Then, if the Coriolis matrix is full rank we have

$$\mathbf{u} \equiv -C^{-1}G\eta^{[i]}, \tag{2.34}$$

and we need to verify that the velocity field is divergence free; i.e.

$$D\mathbf{u} = 0. \tag{2.35}$$

For example, with the $P_1^{NC} - P_1$ pair, the orthogonality of the velocity basis functions [62], implies that the mass and Coriolis matrices are diagonal. Hence the SRV, from $\mathbf{u}$ in (2.34) is obtained by applying the discrete operator C^{-1}; i.e. a rotation of $-\pi/2$, to $-G\eta^{[i]}$. Since column i of $-G$ corresponds to row i of D, the SRV for the $P_1^{NC} - P_1$ pair follows as in Figure 2.3 by a rotation of $-\pi/2$ from the divergence stencil. This procedure only works for the B and C-D grids and for the six last finite-element pairs in Table 2.1 with periodic boundary conditions. For the C grid and RT_0 element the C

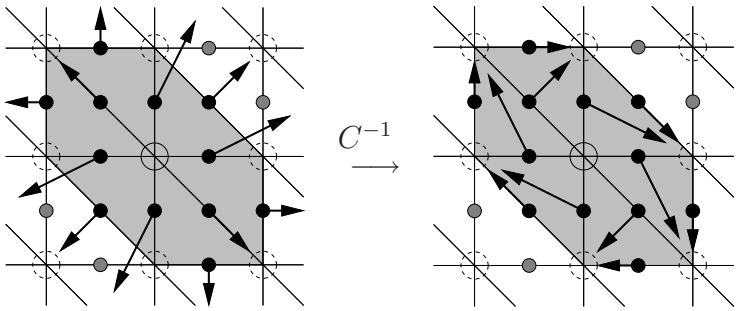


Figure 2.3: For the $P_1^{NC} - P_1$ pair, the SRV (right) is obtained from the divergence stencil (left) by a rotation of $-\pi/2$ corresponding to the C^{-1} operator.

matrix is not invertible and hence $C\mathbf{u} = -G\eta^{[i]}$ cannot be solved. When the no-normal flow boundary condition is used, we have $D\mathbf{u} \neq 0$ in the vicinity of boundary nodes for the FD and FE schemes of Table 2.3 so (2.35) is not satisfied at boundary nodes for such schemes. Further, if the C matrix is not diagonal which is the case for the C grid and all the FE pairs, except $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs, the solution of (2.34) and (2.35) has a very large support covering the whole domain. This is why a second and improved procedure is used in the following. It consists of computing a matrix K such that the range $\mathrm{R}(K)$ of K satisfies

$$\ker(CDG) = \mathrm{R}(K), \tag{2.36}$$

and where the matrix K^T is in a row reduced echelon form. This procedure allows us to find members of the searched basis, i.e. the columns of K, with a maximum number of zero components. Such members have a minimal support and their corresponding velocity fields take the form of local vortices centered around mesh nodes. These smallest representable vortices (SRV) are illustrated in Figures 2.4 and 2.5 for the FD and FE schemes, respectively. The support of the elevation field is shaded. The velocity, normal velocity and elevation nodes are identified using the symbols of Figure 2.1 and 2.2. When the velocity and elevation node symbols are represented in grey and dotted circles, respectively, this means that velocity and elevation are zero at those nodes.

The SRV diagram and x-component of the velocity field are shown in Figure 2.4 for the B, C and C-D grids. For the B grid, the SRV is centered around a pressure node and flows clockwise in a circular motion. This diagram is close to the smallest vortex in [26, Figure 5] for the $Q_1 - P_0$ pair. This is because both schemes lead to similar G and D discrete operators on a regular and uniform grid. Note that in [26] the SRV is defined as the solution of $D\mathbf{u} = 0$ only. However, from (2.23) it coincides with the definition given above in (2.34) and (2.35). The number of SRV is hence $(m-2)(n-2)$, which coincides

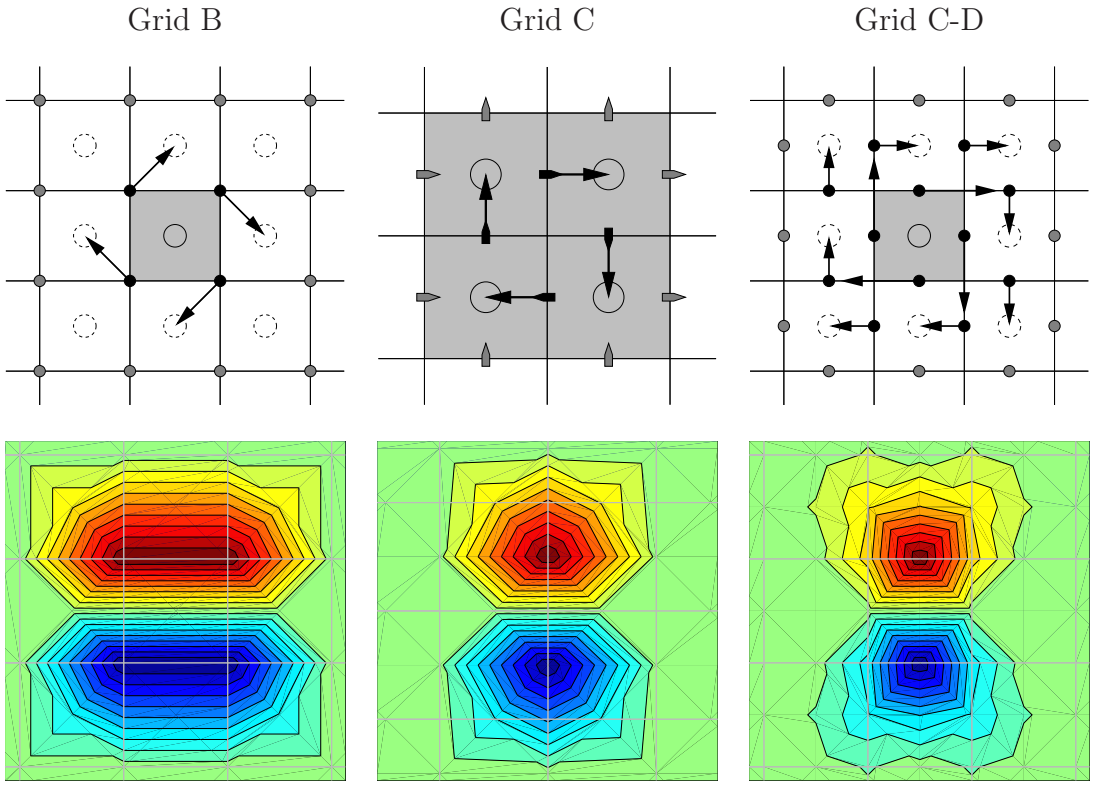


Figure 2.4: SRV diagram (above) and x-component of the SRV normalized velocity field using linear interpolation (bottom) for the B, C and C-D grids.

with the number of pressure nodes not adjacent to the boundary, when the no-normal flow boundary condition is used. The dimension of CDG in Table 2.4 is obtained by adding the two other modes corresponding to the hydrostatic and the spurious pressure modes when no slip boundary conditions are used. For the C grid the SRV is centered in a 2×2 patch where the four surface-elevation nodes have the same positve values, and thus $(m-1)(n-1)$ SRV are present as indicated in Table 2.4. Further, if m and n are even, the domain can be tiled (partitioned) by 2×2 patches (colored here in white and grey for graphical purpose) as shown in Figure 2.6, and a spurious mode corresponding to constant surface-elevation and non zero velocity is obtained. This spurious mode is a solution of $C\mathbf{u} = 0$ and $D\mathbf{u} = 0$, and accordingly is termed a CD-mode in (2.19). The other mode in Table 2.4 is the hydrostatic mode. The SRV diagram for the C-D grid is centered around a pressure node, as in the B grid, and flows clockwise as shown in Figure 2.4. The number of SRV also corresponds to the B grid case, except for the spurious pressure mode which is not present here. In Figure 2.4 the x-component of the SRV normalized velocity field using linear interpolation is given for all grids. The x- and y-components are symmetric and thus the latter is not shown.

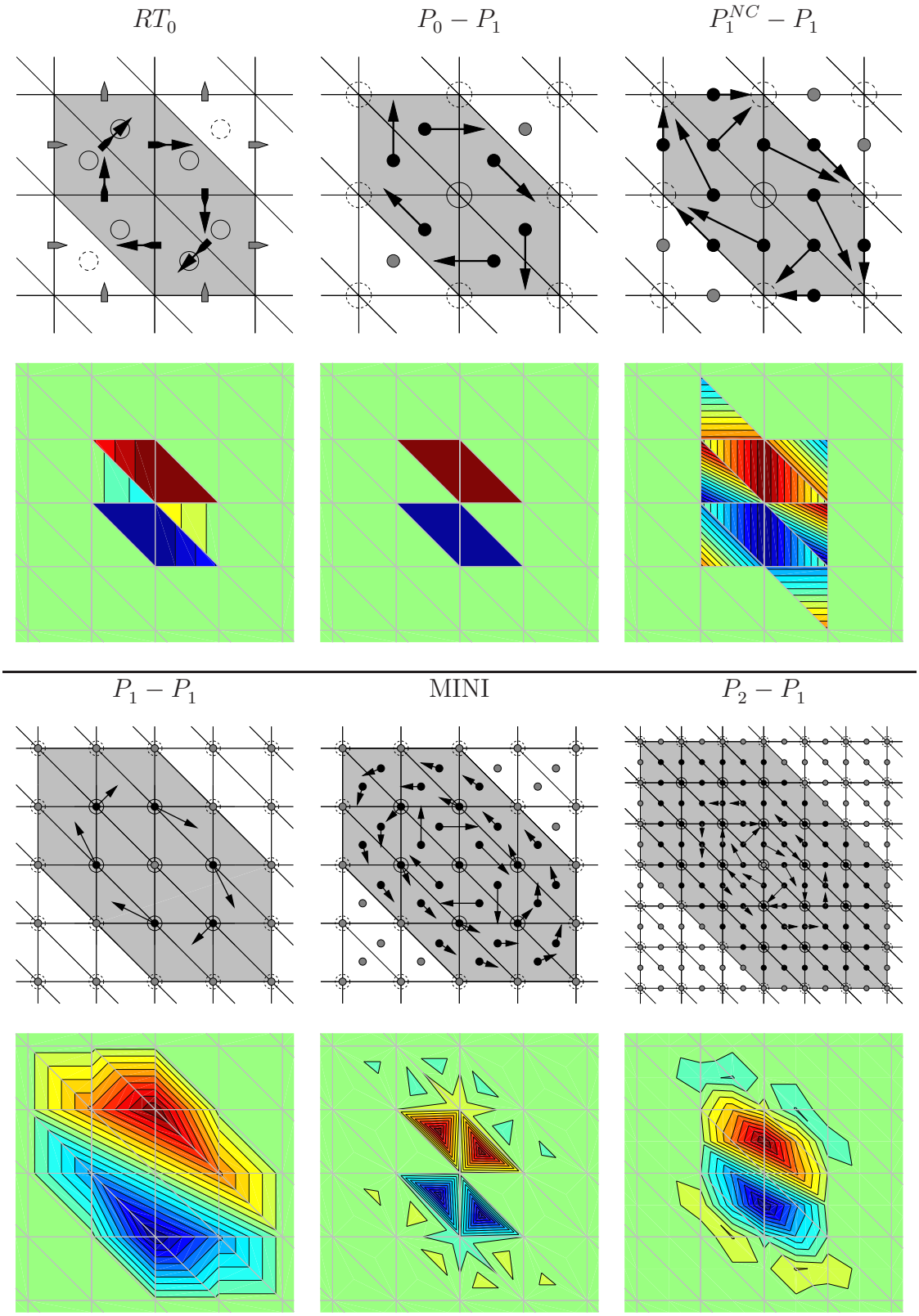


Figure 2.5: As in Figure 2.4 for the FE pairs.

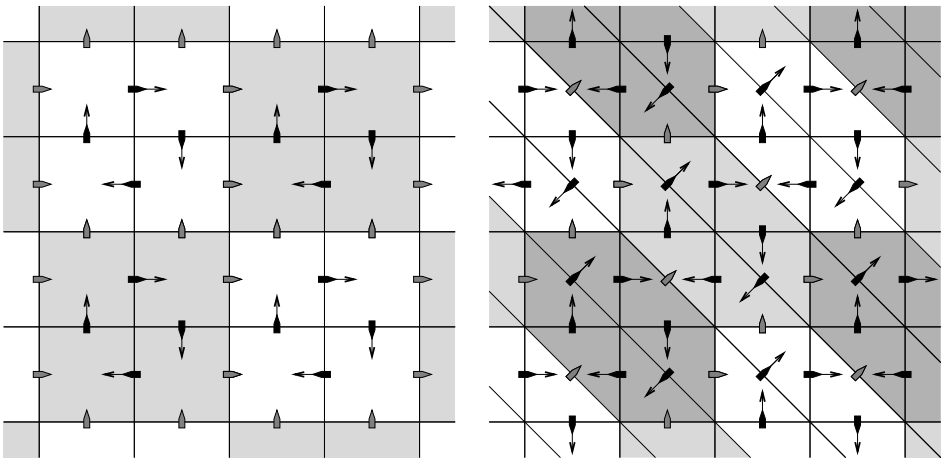


Figure 2.6: CD-Mode of the C grid (left) and RT_0 element (right).

Figure 2.5 shows the SRV diagrams and the x-component of the SRV normalized velocity field for the seven FE pairs considered here. For the RT_0 element, the elevation field is constant over a patch of six elements adjacent to a central vertex node. This patch corresponds to the hexagonal shaded area in Figure 2.5. The support of the SRV velocity field is made up of the six interior edges of the shaded patch. Because the elevation is constant over the hexagonal patch, one can tile the domain with SRV having the same elevation field if the domain is periodic and m and n are both divisible by three. In this case we obtain a non zero velocity mode with constant elevation, which has been defined above as a CD mode and is shown in Figure 2.6. As for the C grid, white and grey colored areas are introduced to graphically highlight SRV support. When the no-normal flow boundary condition is employed, the total number of SRV is $(m-1)(n-1)$ as shown in Table 2.4 because each SRV lies on a 2×2 patch. The P_0-P_1 and $P_1^{NC}-P_1$ pairs have SRV similar to the RT_0 scheme, except that the elevation is now non zero at the vortex center node only, as shown in Figure 2.5. The number of SRV is again $(m-1)(n-1)$ as for the RT_0 element when the no-normal flow boundary condition is used (Table 2.4), because the SRV patch size is still 2×2.

The SRV patch for the P_1-P_1 pair is larger than for the three previous pairs. The elevation field is again centered around a central vertex node but it is now non-zero at that node and also at the six adjacent vertices. The support of the elevation field is shaded in Figure 2.5. The SRV lies on a 4×4 patch and hence we have $(m-3)(n-3)$ SRV modes when no-normal flow boundary conditions are used as shown in Table 2.4.

For the MINI pair, the SRV also lies on a 4×4 patch, leading to $(m-3)(n-3)$ SRV in Table 2.4, as for the P_1-P_1 pair. However the velocity field is different and now includes bubble nodes which affect the vortex structure. The vortex amplitude is

essentially located at the bubble nodes of the six triangles sharing the vertex center and the circulation is clockwise, as shown in Figure 2.5. A much weaker counter-clockwise circulation is also present on the remaining vertex and bubble nodes of the 4×4 patch. Indeed, the arrow length has been increased for the counter clockwise circulation for graphical purposes.

The P_1 iso $P_2 - P_1$ and $P_2 - P_1$ pairs behave similarly for both the number and structure of SRV and hence only the results for the $P_2 - P_1$ pair are shown in Figure 2.5. The SRV lies on a 6×6 patch, leading to $(m-5)(n-5)$ SRV modes when no-normal flow boundary conditions are used as shown in Table 2.4. The vortex structure is mainly located at the mid-side nodes of the six triangles sharing the vertex center and the circulation is clockwise, as shown in Figure 2.5. As was the case above for MINI, an additional counter-clockwise and much weaker circulation again is present in the near vicinity of the clockwise circulation. The absence of arrows at the velocity nodes of the shaded area implies the velocity amplitude is very small.

The x-component of the normalized SRV velocity field is given for all the FE pairs in Figure 2.5. The solutions are plotted according to the nature of the FE interpolation basis. This is why discontinuities appear in the representation of the first component of the velocity field in Figure 2.5 for the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs. Such a discontinuous representation naturally leads to 2×2 SRV supports and this advantage is illustrated in the next section. Note that the SRV velocity field for the $P_1^{NC} - P_1$ pair lies outside the 2×2 patch due to the presence of non zero velocities on the boundary edges of the grey area. However, such velocities only have tangential components (contrary to the B grid in Figure 2.4) and consequently, the $P_1^{NC} - P_1$ SRV behavior is similar to that of the RT_0 and $P_0 - P_1$ SRV, as shown later.

Finally, there are mn SRV on periodic domains for all FD and FE formulations, as shown in Table 2.1. This is because there are mn triangle vertices and each of them is the center of a SRV. Note that if we sum all SRV of a periodic domain, the velocity arrows add up to zero by symmetry on the whole domain and the elevation field is constant. We then obtain the hydrostatic mode as a linear combination of all SRV. This is why in Table 2.1 the hydrostatic mode is not counted separately, in contrast to Table 2.4.

2.6.4 Kernel dimension for an unstructured mesh

A key advantage of the FE method over FD schemes is the ease with which unstructured meshes may be used for discretizing complex domains, to promote mesh grading and

for local adaptive mesh refinement. Consequently, we now investigate the impact on the stationary modes of the FE schemes and unstructured meshes.

Table 2.5: Dimension of discrete operators kernels on Mesh 2 with periodic boundary condition.

FE Pair	p	q	C	G	D	CD	CG	CDG		
								SRV		others
RT_0	342	228	102	1	115	0	228	114	+	0
$P_0 - P_1$	456	114	0	1	343	0	114	114	+	0
$P_1^{NC} - P_1$	684	114	0	1	571	0	114	114	+	0
$P_1 - P_1$	228	114	0	1	115	0	114	0	+	2
MINI	684	114	0	1	571	0	114	0	+	2
P_1 iso $P_2 - P_1$	912	114	0	1	799	0	114	0	+	2
$P_2 - P_1$	912	114	0	1	799	0	114	0	+	2

Table 2.6: As for Table 2.5 but with the no-normal flow boundary condition.

FE Pair	p	q	C	G	D	CD	CG	CDG		
								SRV		others
RT_0	322	228	84	1	95	0	228	95	+	1
$P_0 - P_1$	456	135	0	1	322	0	135	95	+	2
$P_1^{NC} - P_1$	684	135	0	1	550	0	135	95	+	1
$P_1 - P_1$	226	135	0	1	92	0	135	0	+	1
MINI	682	135	0	1	548	0	135	0	+	1
P_1 iso $P_2 - P_1$	910	135	0	1	776	0	135	0	+	1
$P_2 - P_1$	910	135	0	1	776	0	135	0	+	1

In the following, an unstructured Mesh 2 has been generated for kernel comparison studies with the previous structured 10×10 Mesh 1 shown in Figure 2.8. The respective meshes are approximately of the same nodal density and the same equi-spaced nodes are used on the boundary of Mesh 2 to easily impose periodic boundary conditions. Tables 2.5 and 2.6 show the dimension of the discrete operator kernels on Mesh 2 for the seven FE pairs examined here, in the case of periodic and no-normal flow boundary conditions, respectively. For the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs in Table 2.5 the number of SRV correspond to the number of mesh vertices, i.e. 114. This is still the case for Table 2.6 since there are 40 boundary nodes and 95 mesh vertices inside the

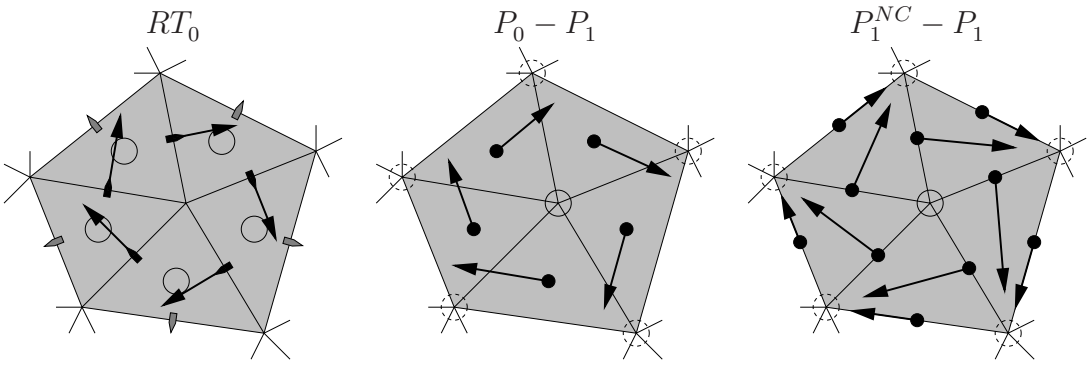


Figure 2.7: As for Figure 2.5 for the SRV diagrams of the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs but on unstructured meshes.

domain. Note that the hydrostatic mode is included in the SRV space for these pairs in the case of periodic conditions as shown in Table 2.5. This is not true in Table 2.6 when the no-normal flow condition is used. The SRV structures for these pairs are shown in Figure 2.7 and they behave similarly to the corresponding SRV on structured meshes. We note that (2.23) is still satisfied on unstructured meshes for the RT_0 element with the no-normal flow boundary condition. For periodic boundary conditions, equations (2.22) and (2.24) hold for the $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs, as previously observed for structured meshes. For the $P_1 - P_1$, MINI, P_1 iso $P_2 - P_1$ and $P_2 - P_1$ pairs, the SRV are dramatically absent from $\ker(CDG)$. Indeed, $\ker(CDG)$ only contains other modes and no longer SRV. Since the hydrostatic mode always belongs to the other mode set, whatever the choice of boundary conditions for those pairs, we deduce that (2.22) is no longer satisfied on unstructured meshes.

In order to better understand the behaviour of the stationary modes on unstructured meshes, we compute the spectrum of the CDG matrix for the $P_1 - P_1$ and $P_2 - P_1$ pairs on Meshes 1 and 2. The results, shown in Figure 2.8, represent the eigenvalues of the CDG operator in the complex plane. On Mesh 1, the eigenvalues are mostly concentrated near the origin for both pairs, and correspond to SRV solutions. However, on Mesh 2 the eigenvalues extend along the imaginary axis and hence the corresponding modes are no longer stationary.

2.7 Numerical representation of vortices

In the previous section we determined the kernel SRV structure for the FD and FE schemes of Section 3.3. Here, our purpose is to conduct two numerical experiments in

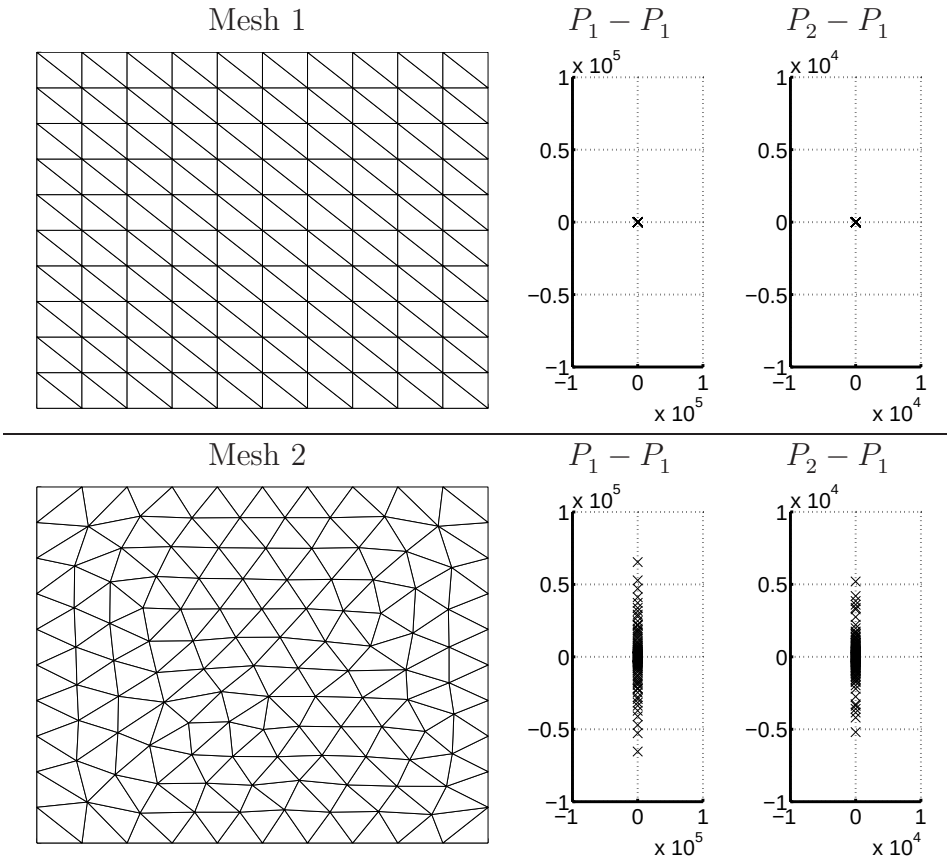


Figure 2.8: Regular Mesh 1 and unstructured Mesh 2 (left) and eigenvalues of the *CGD* matrix for the $P_1 - P_1$ and $P_2 - P_1$ pairs on Mesh 1 and Mesh 2 (right).

order to examine how those stationary SRV combine together to represent realistic and larger scale vortices. In the first test, a large stationary vortex is computed in a similar manner to that in Section 4.3 using the no-normal flow boundary condition. In the second test, the time-dependent SRV are approximated for the case of a non-constant Coriolis parameter.

2.7.1 Representation of large stationary vortices

On the structured FD grids and FE Mesh 1, the square domain extent is 600 km ×600 km. A 25×25 regular cartesian grid is used for the FD schemes to filter the spurious CD-mode of the C grid; i.e. $\beta = 1$ in Table 2.4 and hence m or n are odd in (2.27). To exhibit the CD mode on a C grid a 24×24 regular grid is employed. For the FE element pairs two meshes are used: the regular 24×24 Mesh 1 and the unstructured Mesh 3 (on a larger square domain of dimension $1\,200 \times 1\,200$ km) shown in Figure 2.9.

At the center of Mesh 3 the high resolution area has approximately the same resolution as the 24×24 Mesh 1.

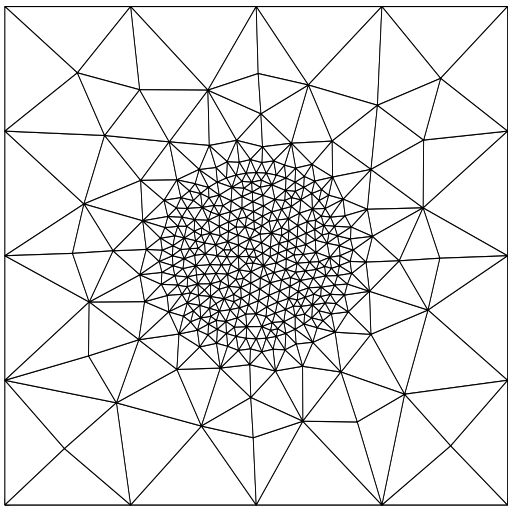


Figure 2.9: The unstructured Mesh 3

A Gaussian distribution of the surface elevation, centered in the domain, is prescribed

$$\eta(x, y) = A\mathrm{e}^{-(x^2+y^2)/r^2}, \tag{2.37}$$

where $r = 1.3 \times 10^5$ m. The gravity and Coriolis parameters are $g = 9.81\ \mathrm{m\,s^{-2}}$ and $f = 6.16 \times 10^{-5}\ \mathrm{s^{-1}}$, with a resolution $h = 2.5 \times 10^4$ m. The e-folding radius of the Gaussian (the distance from the center for which $\eta(x, y) == Ae^{-1}$) is 130 km and it is resolved by 5 mesh nodes. By setting $A = 0.95$ m, the initial maximum surface azimuthal velocity is $1\,\mathrm{m\,s^{-1}}$. These values are chosen to match observations of the eddies in the Gulf of Mexico [38] and they are also representative of the oceanic circulation at mid-latitude. The elevation field is mainly concentrated in the refined area of Mesh 3 and it is numerically zero at the boundaries, avoiding boundary effects for the SRV.

The velocity field is obtained by solving (2.17) rewriten as

$$C\,\mathbf{u} = -G\,\eta, \tag{2.38}$$

$$D\,\mathbf{u} = 0. \tag{2.39}$$

The C matrix in (2.38) is invertible for the FE schemes examined earlier, except for the RT_0 element and (2.24) does not hold. In order to solve (2.38) for the RT_0 scheme

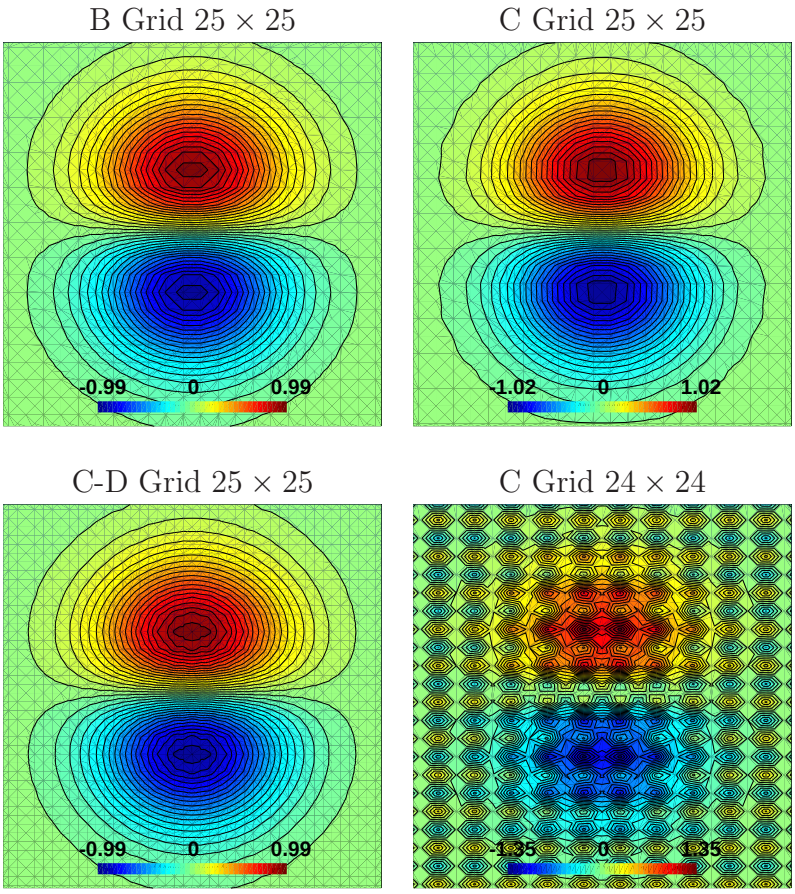


Figure 2.10: x-components of the large stationary vortex for the FD grids. The scale legend shows the minimum and maximum values in $\mathrm{m\,s^{-1}}$.

we ensure that $G\,\eta$ belongs to the range of C by setting the elevation on each element to be the mean value of η at the element mesh vertices obtained from (2.37). Such a procedure is used in the following and (2.38) and (2.39) then admit a solution in the SRV basis for the RT_0 element.

We define the relative residual norm of (2.17) as

$$R \equiv \frac{\|CDG\,(\mathbf{u},\eta)^t\|}{\|CDG\|\,\|(\mathbf{u},\eta)^t\|}. \tag{2.40}$$

The values of R are shown in Table 2.7 for all FE pairs on Meshes 1 and 3. These results demonstrate that (2.38) and (2.39) are satisfied on the regular Mesh 1 for all pairs, and only for the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs on the unstructured Mesh 3. For the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$, the computation of (2.38) - (2.39) is straightforward due to the presence of SRV on both regular and unstructured meshes. This is also the case for the $P_1 - P_1$, MINI, $P_1\,\mathrm{iso}\,P_2 - P_1$ and $P_2 - P_1$ pairs on regular meshes. However,

Table 2.7: Values of R on Meshes 1 and 3 for the FE schemes.

FE	Mesh 1	Mesh 3
RT_0	3.31×10^{-15}	5.09×10^{-16}
$P_0 - P_1$	3.26×10^{-15}	2.29×10^{-16}
$P_1^{NC} - P_1$	3.11×10^{-15}	4.37×10^{-16}
$P_1 - P_1$	1.42×10^{-15}	7.62×10^{-5}
MINI	2.06×10^{-15}	3.56×10^{-6}
P_1 iso $P_2 - P_1$	2.53×10^{-15}	8.73×10^{-6}
$P_2 - P_1$	4.82×10^{-15}	9.72×10^{-6}

for the last four pairs no SRV are present on unstructured meshes, as shown previously in Section 4.4 and the system (2.38) - (2.39) is over-constrained. In order to compute $\mathbf{u}$, we proceed as follows: first (2.39) is solved and (2.38) is then approximated in the least-square sense. Consequently, we obtain a divergence-free velocity field where the solution $(\mathbf{u}, \eta)$ is not numerically stationary as shown in Table 2.7.

In Figure 2.10 the x-components of the computed velocity field for the finite difference schemes are shown. The C-grid result exhibits strong oscillations on the 24×24 grid due to the presence of the periodic CD-mode (Figure 2.6). However, this mode does not exist on the 25×25 grid of Figure 2.10 and the behaviour of the solution is smooth. Similar results are obtained for the B-Grid and the C-D-Grid.

The x-components of the computed velocity field for the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs are shown in Figure 2.11 on Meshes 1 and 3. The results are displayed in a similar manner to those in Figure 2.5; i.e. according to the nature of the FE basis (constant for the $P_0 - P_1$ pair and discontinuous for the RT_0 and $P_1^{NC} - P_1$ pairs). For those three pairs, the results are consistent with the analytical solution.

Finally, the x-components of the computed velocity field are shown in Figure 2.12 for the remaining FE pairs. Smooth results are obtained for the $P_1 - P_1$ pair on both meshes. However, for MINI, strong oscillations appear on meshes 1 and 3, due to a much higher amplitude of the solution at bubble nodes compared to vertex nodes. For the $P_1 \text{iso} P_2 - P_1$ and $P_2 - P_1$ pairs oscillations are again present for both pairs on Mesh 1 and they are amplified on the unstructured Mesh 3.

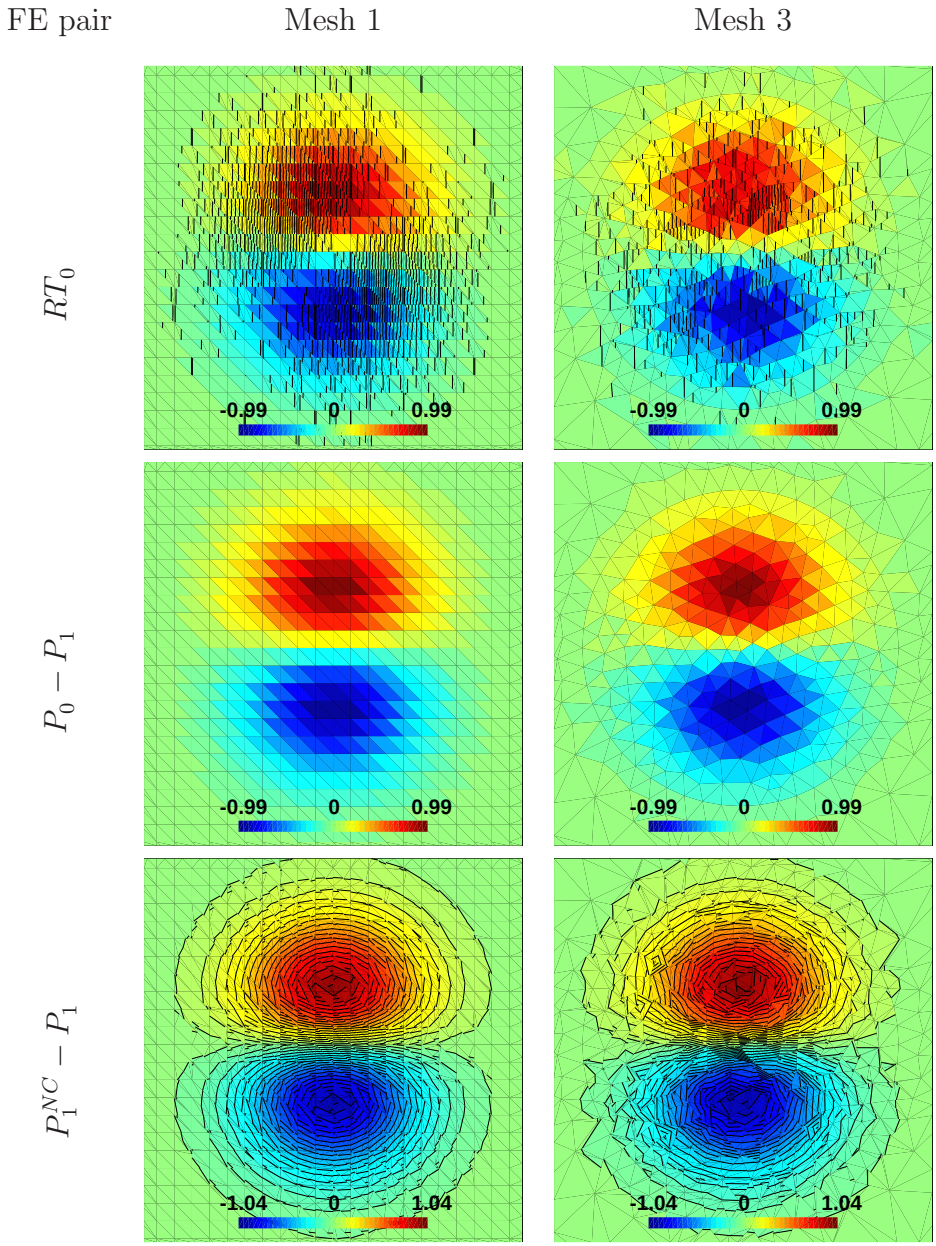


Figure 2.11: As in Figure 2.10 for the RT_0, $P_0 - P_1$ and $P_1^{NC} - P_1$ pairs on Meshes 1 and 3.

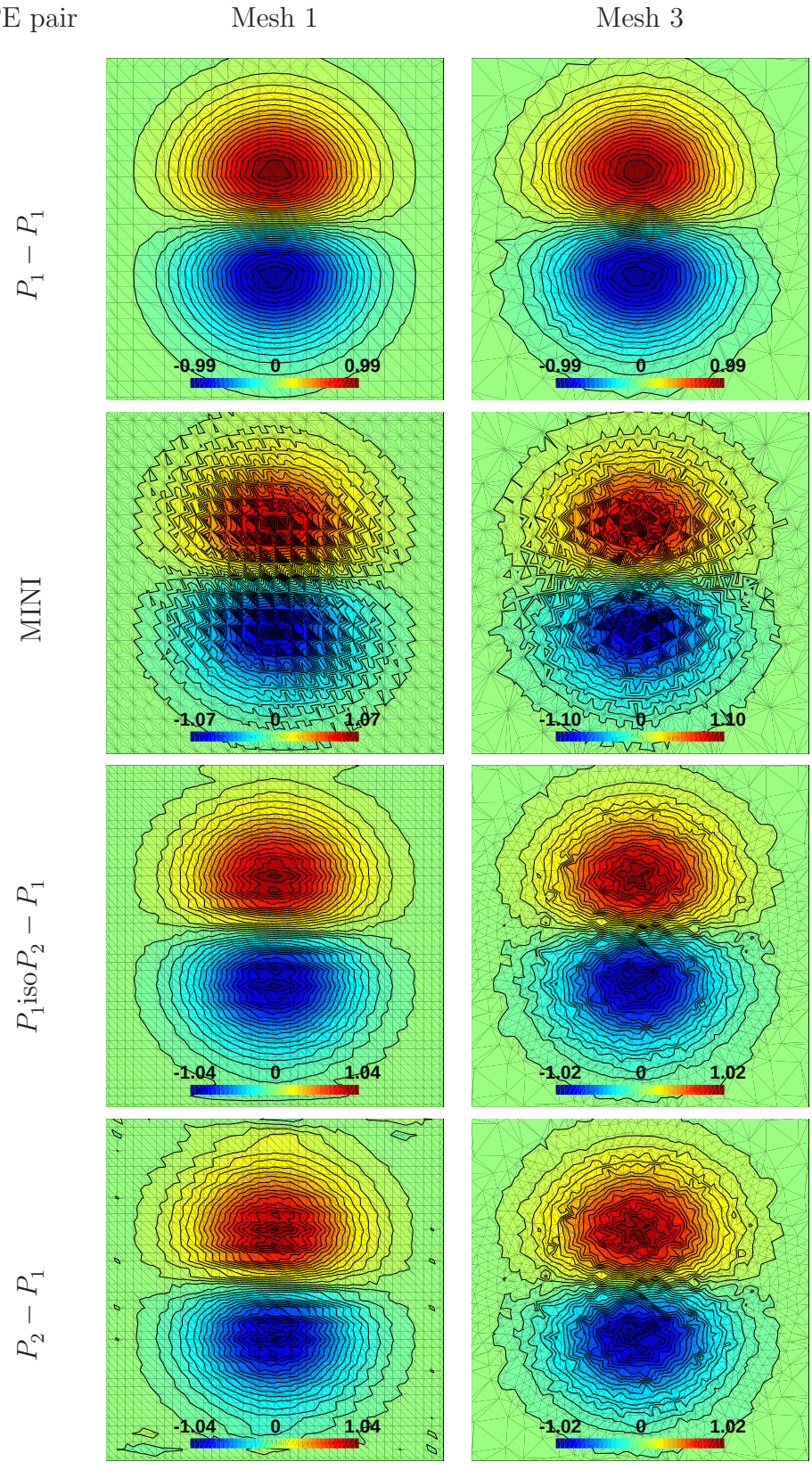


Figure 2.12: As in Figure 2.11 for the $P_1 - P_1$,MINI, $P_1\text{iso}P_2 - P_1$ and $P_2 - P_1$ pairs.

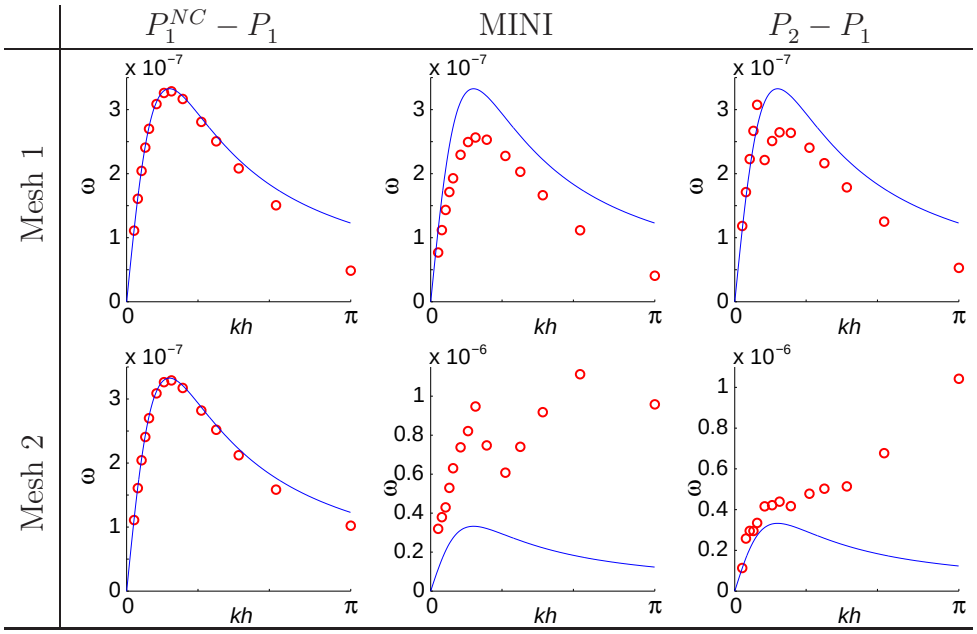


Figure 2.13: Analytic (solid line) and computed (circles) dispersion relations for the $P_1^{NC} - P_1$, MINI and $P_2 - P_1$ pairs on Mesh 1 and Mesh 2. The domain is a 4800 km $\times$ 280 km rectangular basin.

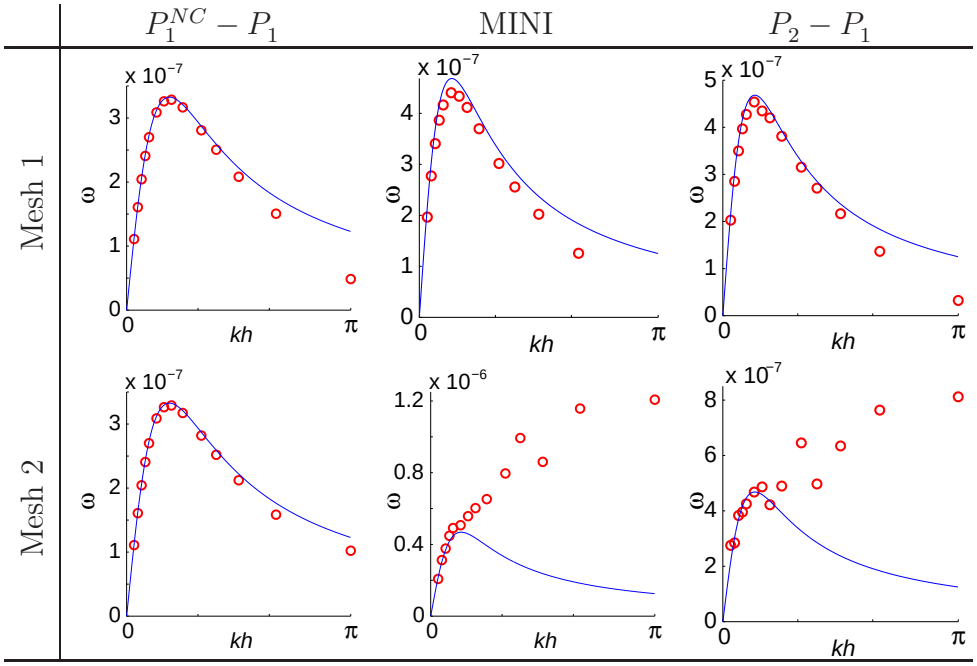


Figure 2.14: As for Figure 2.13 but the domain is now a 4800 km $\times$ 840 km rectangular basin.

2.7.2 Representation of time-dependent vortices

The domain is an idealized $L_x \times L_y$ rectangular basin discretized using the vertex node spacing $h = 40$ km. We let $L_x = 4\,800$ km and hence the mesh has 120 elements in the $x-$direction. The initial solution

$$\bar{\eta} = \cos(kx)\,\sin\left(\frac{\pi y}{L_y}\right), \tag{2.41}$$

$$\bar{u} = -\frac{\pi g}{f L_y}\,\cos(kx)\,\cos\left(\frac{\pi y}{L_y}\right), \tag{2.42}$$

$$\bar{v} = -\frac{gk}{f}\,\sin(kx)\,\sin\left(\frac{\pi y}{L_y}\right), \tag{2.43}$$

is in geostrophic balance and satisfies the no-normal flow boundary condition on the horizontal boundaries. Periodic conditions are applied along the eastern and western boundaries. The β-plane approximation, $f = f_0 + \beta\, y$, is used where $f_0 = 10^{-4}\,\mathrm{s}^{-1}$, $\beta = 10^{-11}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$ and the choice $H = 1.631$ m results in a phase speed for gravity waves of $10\,\mathrm{m\,s}^{-1}$. The radius of deformation at midbasin is thus $R_d = 100$ km.

The wave number k considered in the following study (the wave moves in the $x-$direction) is such that $k = 2\pi n/L_x$, for $n = 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 24, 30, 40, 60$. At each time step the elevation and velocity approximations are computed. The time evolution of the computed elevation at the center of the domain $(L_x/2, L_y/2)$ is used to deduce the period T and the frequency $\omega = 2\pi/T$ for each value of k. The timestep is 2×10^4 s and the duration of the simulation is 1×10^4 time steps. A Crank-Nicolson time integration scheme is used for the Coriolis, gradient and divergence operators. A smoothing procedure is applied here in order to filter the small scale noise induced by the inertia-gravity waves.

In Figure 2.13 the analytic (solid line) and computed (circles) dispersion relations are shown for the $P_1^{NC} - P_1$, MINI and $P_2 - P_1$ pairs on Mesh 1 and Mesh 2. The domain is a 4800 km $\times$ 280 km rectangular basin. For the $P_1^{NC} - P_1$ pair the agreement between the two curves is quite good on both meshes, but for the other pairs the match is poor even on Mesh 1. This is probably due to a larger SRV for the two last pairs and a larger L_y is needed.

In Figure 2.14 the domain is now a 4800 km $\times$ 840 km rectangular basin. The results are unchanged for the $P_1^{NC} - P_1$ pair but for the MINI and $P_2 - P_1$ pairs the results are considerably improved on Mesh 1. This is likely because L_y has increased. However,

on Mesh 2 significant errors are still observed for the two last pairs. This is probably because no SRV are present for the MINI and $P_2 - P_1$ pairs on unstructured meshes.

Finally, we note that the RT_0 and $P_0 - P_1$ pairs lead to results that are similar to those for the $P_1^{NC} - P_1$ pair. Likewise, the $P_1 - P_1$ and P_1 iso $P_2 - P_1$ pairs give similar results to the $P_2 - P_1$ pair [55].

2.8 Conclusion

An algebraic kernel approach has been developed for analysis of spurious modes arising in shallow-water simulation and several finite difference and finite element discretization schemes have been investigated to demonstrate the approach and key ideas. Three key properties of the continuous formulation lead to corresponding tests on the associated kernel representations for the discrete form. Of particular interest is the use of the kernel to characterize the "smallest representable vortices" (SRV) of the discrete models as in Table 2.4, to compare variations in their size and structure on local mesh patches, and to interpret results concerning modal decoupling. The effects of "no-normal flow" and periodic boundary conditions on the dimension of the discrete operator kernels are also investigated and compared in Tables 2.1, 2.3 and 2.4. The numerical studies with, for instance, different element pairs for approximating elevation reveal some unexpected and interesting features. For example, as seen in the studies with the $P_2 - P_1$ and MINI pairs (see Tables 2.4, 2.5, 2.6, 2.7 and Figure 2.12) the use of high order approximation is not necessarily better since this may lead to larger SRV diameters or to more oscillatory and noisy approximations. Moreover, for the unstructured mesh tests SRV are not evident. This suggests that low order schemes may be preferable since they are less computationally costly for the same cell scale and can characterize smaller SRV structures efficiently and with less oscillatory behavior. For instance, based on the numerical evidence from the present kernel characterization and supporting tests, the $P_1^{NC} - P_1$, RT_0 and $P_0 - P_1$ pairs are appealing choices since they all yield SRV on unstructured meshes as seen in Tables 2.5, 2.6, 2.7 and Figure 2.7, 2.11 and are computationally inexpensive. Since unstructured meshes do not appear to pose the same difficulty in characterizing SRV, these three low-order schemes are more amenable to practical simulations that ideally utilize mesh grading or adaptive unstructured mesh refinement. This latter point is a topic of study in our ongoing research collaboration.

3. RAVIART-THOMAS AND BREZZI-DOUGLAS-MARINI FINITE ELEMENT APPROXIMATIONS OF THE SHALLOW-WATER EQUATIONS

V. Rostand and D. Y. Le Roux
Submitted to International Journal for Numerical Methods in Fluids (2007).
Keywords: shallow-water equations, finite element method, dispersion relations, Raviart-Thomas, Brezzi-Douglas-Marini

3.1 Résumé

Une étude des équations de Saint-Venant discrétisées par la méthode des éléments finis utilisant les espaces de type Raviart-Thomas et Brezzi-Douglas-Marini est présentée. Les équations discrètes sont utilisées pour calculer les relations de dispersion afin de quantifier la dispersion causée par la discrétisation sur les ondes d'inertie-gravité. Deux maillages sont considérés; le premier est constitué de triangles équilatéraux et le second de triangles biaisés. Une approche basée sur l'algèbre linéaire est aussi utilisée pour déterminer la présence possible de modes numériques parasitaires. La balance géostrophique est examinée par le calcul des plus petits vortex discrets. Les résultats de deux tests numériques reproduisant les ondes de gravité et de Rossby corroborent l'étude.

3.2 Abstract

An analysis of the discrete shallow-water equations using the Raviart-Thomas and Brezzi-Douglas-Marini finite elements is presented. For inertia-gravity waves, the discrete formulations are obtained and the dispersion relations are computed in order to quantify the dispersive nature of the schemes on two meshes made up of equilateral

and biased triangles. A linear algebra approach is also used to ascertain the possible presence of spurious modes arising from the discretization. The geostrophic balance is examined and the smallest representable vortices are characterised on both structured and unstructured meshes. Numerical solutions of two test problems to simulate gravity and Rossby modes are in good agreement with the analytical results.

3.3 Introduction

The finite-element (FE) method is attractive for problems of environmental engineering due to the flexibility of triangulation for the representation of irregular boundaries and for local mesh refinement [15, 22, 34, 45, 58, 31, 39]. The method is widely applied to shallow-water (SW) simulation for a variety of environmental problems including groundwater, coastal regions, atmospheric, and oceanic flows [15, 34, 20, 54, 68]. The SW equations describe the behavior of a shallow homogeneous incompressible and inviscid fluid layer. They are derived from the Navier-Stokes equations under Boussinesq and hydrostatic pressure assumptions.

One of the key issues that arises with mixed formulations is the possible presence of spurious modes, i.e. small scale artifacts, introduced by the spatial discretization scheme. This difficulty with mixed methods is encountered in both the SW and Navier-Stokes formulations. The appearance of spurious solutions is mainly due to an inappropriate choice of approximation function spaces for the FE method. The spurious modes usually take the form of pressure or surface-elevation, velocity and/or Coriolis modes. They do not propagate but are trapped within the model grid. The spurious solutions usually cause aliasing and an accumulation of energy in the smallest-resolvable scale, leading to noisy solutions. Improvements have been achieved through the use of a variety of mixed-order FE interpolation schemes [34, 58, 54, 2, 35, 70], and a wave equation formulation [39, 36].

Dispersion analysis of the discretized form of the linear SW equations is a useful tool to define the relationships between frequency and wave number. It also permits to explicitly ascertain the presence and determine the form of spurious solutions as well as the dispersive/dissipative nature of a FE formulation [53, 52, 56, 55, 65]. However, such an approach is restricted to uniform meshes and periodic solutions. In order to study the existence and the behaviour of stationary spurious modes associated with zero frequency on both uniform and unstructured meshes, a linear algebra approach may be employed [57, 50]. It consists in determining the properties of the kernel of

the associated discretized problem. The kernel characterization may also be used to determine the form of the smallest representative velocity vortex structures that can be represented in a given discretization.

The dispersion analyses and linear algebra approach performed in [56, 55, 50] suggest that three FE schemes, the $P_1^{NC} - P_1$ and $P_0 - P_1$ pairs and the low order Raviart-Thomas element RT_0 can been identified as promising schemes in terms of dispersion properties for the discretization of the SW equations. The aim of this paper is to conduct such analyses for higher order Raviart-Thomas and Brezzi-Douglas-Marini (BDM_1) FE approximation spaces. The BDM_1 element has been introduced in [13] for solving elliptic problems and to our knowledge it is analysed here for the first time in the context of inviscid SW flows. Two formulations are considered: the $BDM_1 - P_0$ and the $BDM_1 - P_1$ pairs.

The paper is organized as follows. We first present the governing equations and the spatial discretization in Sections 2 and 3, respectively. A dispersion analysis is performed in Section 4 for the inertia-gravity waves. In Section 5, a linear algebra approach based on the kernel properties of the discretized problem is used for Rossby waves. The analytical computations are followed by numerical simulations. Some concluding remarks complete the study.

3.4 Governing equations

Let Ω be a 2D domain with boundary Γ. The inviscid linear equations are expressed in Cartesian coordinates [37] as

$$\frac{\partial \mathbf{u}}{\partial t} + f\mathbf{k} \times \mathbf{u} + g\nabla\eta = 0, \tag{3.1}$$

$$\frac{\partial \eta}{\partial t} + H\nabla \cdot \mathbf{u} = 0, \tag{3.2}$$

where $\mathbf{u} = (u, v)$ is the velocity field, η is the surface elevation with respect to the reference level $z = 0$, $\mathbf{k}$ is the unit vector pointing in the vertical direction, g denotes the gravitational acceleration, and the mean depth H and the Coriolis parameter f are assumed constant, unless stated otherwise. The velocity is subject to the no-normal flow boundary condition $\mathbf{u} \cdot \mathbf{n} = 0$ on Γ, where $\mathbf{n}$ is the outward pointing vector at the boundary.

For the analyses performed in this paper, time is assumed continuous and we seek periodical solutions of (3.1)-(3.2) of the from

$$\mathbf{u}(x,y,t) = \mathbf{u}(x,y)\exp(i\omega t), \qquad \eta(x,y,t) = \eta(x,y)\exp(i\omega t), \tag{3.3}$$

where ω is the angular frequency.

By substituting (3.3) in (3.1) and (3.2) we obtain

$$i\omega\,\mathbf{u} + f\mathbf{k}\times\mathbf{u} + g\nabla\eta = 0, \tag{3.4}$$

$$i\omega\,\eta + H\nabla\cdot\mathbf{u} = 0. \tag{3.5}$$

Equations (3.4) and (3.5) are now discretized in space.

3.5 Spatial discretization

3.5.1 The weak formulation

We assume $\mathbf{u}$ and η belong to the spaces V and W, with V and W are in the square-integrable space $L^2(\Omega)$. The weak formulation is obtained by multiplying (3.4) and (3.5) by test functions φ and ψ belonging to V and W, respectively, and by integrating over the whole domain

$$\int_\Omega i\omega\,\mathbf{u}\cdot\boldsymbol{\varphi}\,d\mathbf{x} + \int_\Omega f\,(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\varphi}\,d\mathbf{x} + \int_\Omega g\,\nabla\eta\cdot\boldsymbol{\varphi}\,d\mathbf{x} = 0, \tag{3.6}$$

$$\int_\Omega i\omega\,\eta\,\psi\,d\mathbf{x} + \int_\Omega H\,\nabla\cdot\mathbf{u}\,\psi\,d\mathbf{x} = 0, \tag{3.7}$$

where $d\mathbf{x}$ is the area element with $\mathbf{x} = (x,y)$. The terms containing gradient and divergence operators in (3.6) and (3.7) may be integrated by parts using the Green's theorem, depending on the regularity of spaces V and W.

3.5.2 Galerkin finite-element discretization

The Galerkin method approximates the solution of (3.6) and (3.7) in finite-dimensional subspaces. Consider a representative meshlength parameter h that measures resolution

and a FE triangulation $\mathcal{T}_h$ of the polygonal domain Ω.

The discrete solutions $\mathbf{u}_h$ and η_h sought belong to finite-dimensional spaces $\mathbf{V}_h$ and W_h, respectively, with $\dim(\mathbf{V}_h) = p$ and $\dim(W_h) = q$. The spaces $\mathbf{V}_h$ and W_h are defined as a set of piecewise polynomial functions over the triangulation $\mathcal{T}_h$. The degree and continuity order of these polynomial functions are specified in the sequel for the FE schemes investigated in this paper.

The components of $\mathbf{u}_h$ and η_h are represented over a triangle K by interpolating functions $\varphi(x, y)$ and $\psi(x, y)$, belonging to $\mathbf{V}_h$ and W_h, respectively. We thus have

$$\mathbf{u}_h = \sum_{j=1}^{p} \mathbf{u}_j \varphi_j, \qquad \eta_h = \sum_{k=1}^{q} \eta_k \psi_k. \tag{3.8}$$

Note that for the RT_0 and BDM_1 elements, the normal velocites are expressed in terms of the vector interpolating function $\boldsymbol{\varphi}(x, y)$.

Let $\underline{M}$ and $\underline{N}$ be the velocity and elevation mass matrices, respectively, $\underline{C}$ the Coriolis matrix, and $\underline{G}$ and $\underline{D}$ the gradient and divergence matrices, which are obtained from the elementary matrices

$$\begin{aligned}
&\underline{M}_{j_1,j_2} = i\omega \int_\Omega \boldsymbol{\varphi}_{j_1} \cdot \boldsymbol{\varphi}_{j_2}, \qquad \underline{N}_{k_1,k_2} = i\omega \int_\Omega \psi_{k_1}\, \psi_{k_2}, \\
&\underline{C}_{j_1,j_2} = \int_\Omega f(\mathbf{k} \times \boldsymbol{\varphi}_{j_1}) \cdot \boldsymbol{\varphi}_{j_2}, \quad \underline{D}_{k_1,j_1} = \int_\Omega H(\nabla \cdot \boldsymbol{\varphi}_{j_1})\, \psi_{k_1}, \\
&\underline{G}_{j_1,k_1} = \int_\Omega g \nabla \psi_{k_1} \cdot \boldsymbol{\varphi}_{j_1},
\end{aligned} \tag{3.9}$$

where $j_1, j_2 = \{1, \cdots, p\}$ and $k_1, k_2 = \{1, \cdots, q\}$. By using (3.8) and (3.9), equations (3.6) and (3.7) are rewritten in matrix form as

$$\begin{pmatrix} \underline{M} + \underline{C} & \underline{G} \\ \underline{D} & \underline{N} \end{pmatrix} \begin{pmatrix} \underline{\mathbf{u}} \\ \underline{\eta} \end{pmatrix} = 0, \tag{3.10}$$

where $\underline{\mathbf{u}} = (u_1, \cdots, u_p)^t$ and $\underline{\eta} = (\eta_1, \cdots, \eta_q)^t$.

3.5.3 The BDM_1 element

The BDM_1 element has been introduced by F. Brezzi, J. Douglas and L.D. Marini [13, 14] in 1986 to solve elliptic problems. It can be seen as an enriched version of the lowest order Raviart-Thomas element denoted by RT_0. The RT_0 element is based on

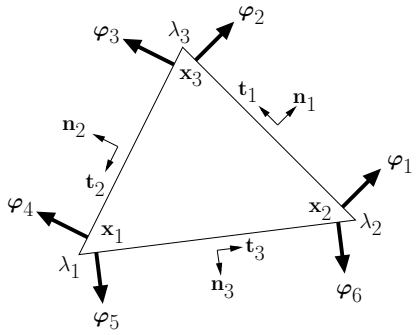


Figure 3.1: Elementary basis functions for the BDM_1 space.

flux conservation on element edges and it has continuous normal components at triangle midedge points. For the BDM_1 element, the normal component is continuous through triangle edges, as for the RT_0 element, but the corresponding basis functions are linear instead of being constant. Consequently, the BDM_1 element has two degrees of freedom over each triangle side instead of one for the RT_0 element.

Consider a triangle K of $\mathcal{T}_h$, and let $\mathbf{x}_i = (x_i, y_i),\ i = 1, 2, 3$, the coordinates of the three vertices of K. At each side of K the tangential and normal vectors are defined as

$$\begin{aligned} \mathbf{t}_1 &:= \mathbf{x}_3 - \mathbf{x}_2, \quad \mathbf{n}_1 := -\mathbf{k} \times \mathbf{t}_1, \\ \mathbf{t}_2 &:= \mathbf{x}_1 - \mathbf{x}_3, \quad \mathbf{n}_2 := -\mathbf{k} \times \mathbf{t}_2, \\ \mathbf{t}_3 &:= \mathbf{x}_2 - \mathbf{x}_1, \quad \mathbf{n}_3 := -\mathbf{k} \times \mathbf{t}_3. \end{aligned} \tag{3.11}$$

Let $\lambda_i,\ i = 1, 2, 3$ be the barycentric functions of K and define for each permutation (ijk) of (123) the basis functions $\boldsymbol{\varphi}$ associated with the BDM_1 element

$$\boldsymbol{\varphi}_{(ijk)} := \frac{\lambda_j \mathbf{t}_k}{\mathbf{t}_k \cdot \mathbf{n}_i} = \frac{\lambda_j \mathbf{t}_k}{2Area(K)} (-1)^{\mathrm{sgn}(ijk)}, \tag{3.12}$$

where $Area(K)$ denotes the area of triangle K. By using (3.12), one can verify that the normal component $\boldsymbol{\varphi}_{(ijk)} \cdot \mathbf{n}$ is linear on edge i and zero on edges j and k. Further, on edge i the normal component takes the value 1 at vertex j and 0 at vertex k. In order to simplify the notation, the functions $\boldsymbol{\varphi}$ are numbered from 1 to 6

$$\begin{aligned} \boldsymbol{\varphi}_1 &:= \boldsymbol{\varphi}_{(123)}, \quad \boldsymbol{\varphi}_2 := \boldsymbol{\varphi}_{(132)}, \quad \boldsymbol{\varphi}_3 := \boldsymbol{\varphi}_{(231)}, \\ \boldsymbol{\varphi}_4 &:= \boldsymbol{\varphi}_{(213)}, \quad \boldsymbol{\varphi}_5 := \boldsymbol{\varphi}_{(312)}, \quad \boldsymbol{\varphi}_6 := \boldsymbol{\varphi}_{(321)}, \end{aligned} \tag{3.13}$$

in terms of the permutations, as shown in Figure 3.1. Note that on each triangle edge of K the sum of the two basis functions $\boldsymbol{\varphi}$ exactly corresponds to the basis function associated with the RT_0 element on that edge.

3.5.4 Finite-element pairs

As previously mentioned, one of the issues associated with mixed methods is the possibility of spurious modes and anomalous dispersion in the representation of waves. The choice of FE pairs to approximate the velocity field $\mathbf{u}$ and the surface elevation η is thus delicate. In [56, 55, 50] three FE pairs, namely the $P_1^{NC} - P_1$ and $P_0 - P_1$ pairs and the low order Raviart-Thomas element RT_0 have been identified as promising schemes in terms of dispersion properties for the discretization of the SW equations. Common to the first two FE velocity/surface-elevation pairs is a piecewise-linear continuous representation of surface elevation, and they differ from one another in their representation of velocity. The $P_1^{NC} - P_1$ pair [34, 52] has velocity nodes at triangle edge midpoints, and linear basis functions are used to approximate the two velocity components on the element's two-triangle support. Since this particular representation of velocity is continuous only across triangle boundaries at midpoint nodes, and discontinuous everywhere else around a triangle boundary, this element is termed nonconforming (NC) in the FE literature. The $P_0 - P_1$ pair has a piecewise-constant representation of velocity. In [56, 55, 50] the RT_0 element has a discontinuous piecewise-constant representation of surface elevation and such a FE pair is thus named $RT_0 - P_0$ in the sequel.

The purpose of this paper is to study the dispervive properties and eventually to detect the presence of spurious modes of three other FE pairs. Those are the $RT_0 - P_1$, $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs. The $RT_0 - P_1$ pair has RT_0 basis functions for velocity and a piecewise-linear continuous representation of surface elevation. Common to the last two pairs is a BDM_1 representation of velocity and the corresponding basis functions for elevation are piecewise-constant and piecewise-linear, respectively.

3.6 Inertia-gravity waves

3.6.1 Analytical frequencies

In the continuum case the free modes of (3.4) and (3.5) are examined by perturbing about the basic state $u = v = \eta = 0$. We seek periodic solutions of the form

$$(\mathbf{u}(x, y),\, \eta(x, y)) = (\tilde{\mathbf{u}},\, \tilde{\eta})\, \exp\left(\, i(kx + ly)\,\right), \tag{3.14}$$

where k and l are the wave numbers in the x- and y-directions, respectively. Substitution into (3.4) and (3.5) leads to a square matrix system for the amplitudes $\tilde{\mathbf{u}}$ and $\tilde{\eta}$. For a

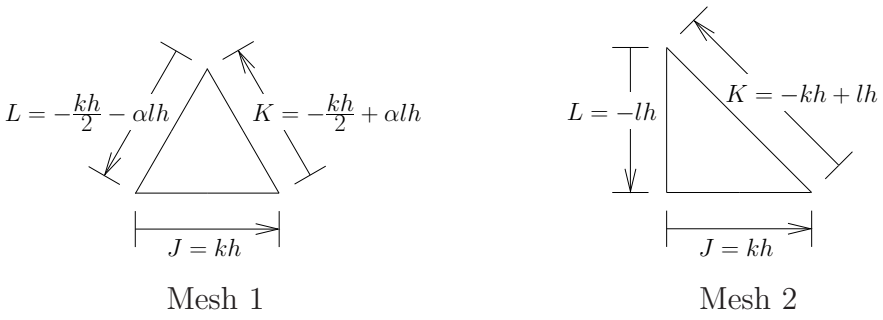


Figure 3.2: Elementary displacements are represented on the reference triangle on Meshes 1 and 2.

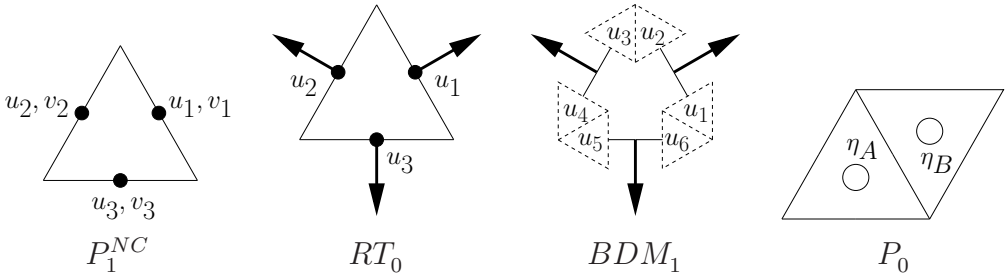


Figure 3.3: Discrete amplitude numbering for the P_1^{NC}, RT_0, BDM_1 and P_0 elements.

nontrivial solution to exist, the 3×3 determinant of the matrix must equal zero, and this constraint leads to the following dispersion relation for the frequency

$$\omega \left(\omega^2 - f^2 - gH\sqrt{k^2 + l^2} \right) = 0. \tag{3.15}$$

The first solution $\omega = 0$ is the geostrophic mode and would correspond to the slow Rossby mode on a β-plane, while the other two solutions

$$\omega_{AN} = \pm\sqrt{f^2 + gH(k^2 + l^2)}, \tag{3.16}$$

correspond to the free-surface gravitational modes with rotational correction. Since ω is purely real, all modes are neutrally stable and neither amplify nor decay.

3.6.2 Discrete frequencies

The dispersion analysis is performed on two types of regular meshes: Mesh 1 and Mesh 2. Mesh 1 corresponds to equilateral triangles and Mesh 2 is made up of biased right isosceles triangles as shown in Figure 3.10. The mesh spacing, denoted by h, is defined as the triangle side length for Mesh 1 and the shortest triangle side length for Mesh 2.

Nodal unknowns are located on typical nodal sets, e. g. faces, vertices, and barycenters. In this analysis only selected discrete equations are retained and they correspond to each type of nodes. For example, three discrete momentum equations are considered for the $RT_0 - P_0$ pair on the three possible types of faces, and two discrete continuity equations are considered at the two possible types of barycenters, e. g. upward and downward pointing triangles as shown in Figure 3.3 for Mesh 1. Consequently, selected discrete amplitudes $\tilde{u}_j$ and $\tilde{\eta}_j$ are considered for each nodal set. Assembling the elementary matrices defined in (3.9) leads to the stencils of Figures 3.4, 3.5, 3.6, and 3.7 for the $P_1^{NC} - P_1$, $RT_0 - P_0$, $RT_0 - P_1$, $BDM_1 - P_0$, and $BDM_1 - P_1$ pairs. Substitution of (3.14) in these reference stencils leads to the linear system

$$\begin{pmatrix} M + C & G \\ D & N \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{u}} \\ \tilde{\eta} \end{pmatrix} = 0, \tag{3.17}$$

where $\tilde{\mathbf{u}}$ and $\tilde{\eta}$ are the amplitude vectors. Matrices M, N, C, D, and G are the contributions of velocity mass, surface-elevation mass, Coriolis, divergence, and gradient stencils, respectively, for the selected equations. As for the continuous case, the determinant of the matrix in the left hand side of (3.17) must vanish to admit non-trivial amplitude solutions, and this leads to the dispersion relation. Note that we have

$$G = -\frac{g}{H}\,\overline{D^t}, \tag{3.18}$$

where $\overline{D^t}$ is the complex conjugate transpose of D. Consequently, only divergence stencils are displayed in the following. To simplify the notation we define the quantities J, K, and L as elementary displacements as shown in Figure 3.2 on Meshes 1 and 2. Note that J, K, and L are different on both meshes. Further we let I_n be the $n \times n$ identity matrix and $\alpha = \sqrt{3}/2$.

We now examine the dispersion relations corresponding to the FE pairs considered in this study.

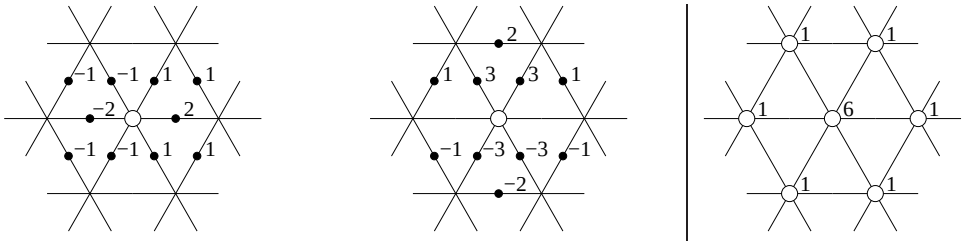
The $P_1^{NC} - P_1$ pair

The discrete velocity amplitudes are considered on three typical faces (see Figure 3.3) and ordered as

$$\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{v}_1, \tilde{v}_2, \tilde{v}_3)^t. \tag{3.19}$$

Because the P_1^{NC} basis functions are orthogonal in L^2, the velocity mass and Coriolis matrices are block diagonal. For Mesh 1, we obtain from (3.9)

$$M = i\omega\,\frac{\alpha h^2}{3} I_6, \qquad C = f\,\frac{\alpha h^2}{3} N \begin{pmatrix} 0 & -I_3 \\ I_3 & 0 \end{pmatrix}. \tag{3.20}$$



Divergence $-x$ $\left(\times\frac{\alpha}{6}\,h\right)$ Divergence $-y$ $\left(\times\frac{1}{12}\,h\right)$ | Mass $\left(\times\frac{\alpha}{12}\,h^2\right)$

$P_1^{NC} - P_1$ | P_1

Figure 3.4: $P_1^{NC} - P_1$ divergence and P_1 mass stencils.

For the surface elevation, only one discrete amplitude is required for $\tilde{\eta}$ at mesh vertices. The stencils of the divergence operator in $-x$ and $-y$ directions are shown in Figure 3.4 and the corresponding matrix is

$$D = H\,\frac{h}{6}\begin{pmatrix} D_1 & D_2 & D_3 & D_4 & D_5 & D_6 \end{pmatrix}, \tag{3.21}$$

where

$$\begin{aligned} D_1 &= 2i\alpha\left(\sin\left(\tfrac{J-L}{2}\right)+\sin\left(\tfrac{-K}{2}\right)\right), & D_4 &= \;i\left(\sin\left(\tfrac{J-L}{2}\right)+3\sin\left(\tfrac{K}{2}\right)\right), \\ D_2 &= 2i\alpha\left(\sin\left(\tfrac{J-K}{2}\right)+\sin\left(\tfrac{-L}{2}\right)\right), & D_5 &= \;i\left(\sin\left(\tfrac{J-K}{2}\right)+3\sin\left(\tfrac{L}{2}\right)\right), \\ D_3 &= 4i\alpha\sin\left(\tfrac{J}{2}\right), & D_6 &= 2i\sin\left(\tfrac{K-L}{2}\right). \end{aligned} \tag{3.22}$$

The mass stencil for the P_1 elevation is shown in Figure 3.4 and the corresponding matrix is

$$\begin{aligned} N &= i\omega\,\frac{\alpha h^2}{12}\left(6+e^{iJ}+e^{iL}+e^{iK}+e^{-iJ}+e^{-iL}+e^{-iK}\right) \\ &= i\omega\,\frac{\alpha h^2}{6}\Big(3+\cos(J)+\cos(K)+\cos(L)\Big). \end{aligned} \tag{3.23}$$

Vanishing the 7×7 determinant vanishes leads to

$$\omega_{1,2} = \pm\omega_{AN} + O(h^4), \qquad \omega_3 = 0, \qquad \omega_{4,5,6,7} = \pm f \text{ (double root)}.$$

The first two roots $\omega_{1,2}$ correspond to inertia-gravity waves and they coincide with the analytical solution (3.16) in the limit as mesh spacing $h \to 0$. The third root ω_3 corresponds to the geostrophic mode and $\omega_{4,5,6,7}$ represent spurious propagating inertial oscillations that have no particular spatial characteristics [52, 56]. The dispersion relation on Mesh 2 is given in [52, 56], and the frequency corresponding to inertia-gravity waves is found to be the same order $O(h^4)$.

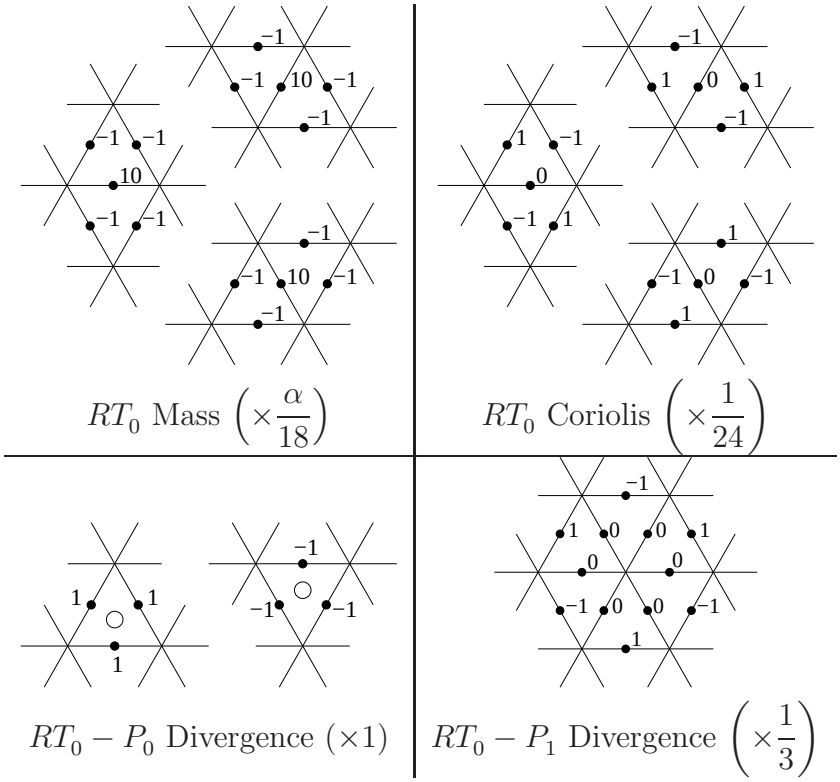


Figure 3.5: Stencils for the $RT_0 - P_0$ and $RT_0 - P_1$ pairs.

The $RT_0 - P_0$ pair

The dispersion relation for the $RT_0 - P_0$ pair on Mesh 2 is given in [56] and the results on Mesh 1 are computed here. By using the notations of Figure 3.3 the following discrete amplitudes are considered

$$\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \tilde{u}_3)^t, \qquad \tilde{\eta} = (\tilde{\eta}_A, \tilde{\eta}_B)^t, \tag{3.24}$$

where A and B represent typical barycenters of upward and downward pointing triangles. The velocity mass, Coriolis, and divergence stencils are shown in Figure 3.5, and

the corresponding matrices are

$$M = i\omega\frac{\alpha}{9}\begin{pmatrix} 5 & -\cos\left(\frac{J}{2}\right) & -\cos\left(\frac{L}{2}\right) \\ -\cos\left(\frac{J}{2}\right) & 5 & -\cos\left(\frac{K}{2}\right) \\ -\cos\left(\frac{L}{2}\right) & -\cos\left(\frac{K}{2}\right) & 5 \end{pmatrix}, \tag{3.25}$$

$$C = f\frac{1}{3}\begin{pmatrix} 0 & -\cos\left(\frac{J}{2}\right) & \cos\left(\frac{L}{2}\right) \\ \cos\left(\frac{J}{2}\right) & 0 & -\cos\left(\frac{K}{2}\right) \\ -\cos\left(\frac{L}{2}\right) & \cos\left(\frac{K}{2}\right) & 0 \end{pmatrix}, \tag{3.26}$$

$$D = H\begin{pmatrix} E\left(\frac{J-L}{6}\right) & E\left(\frac{K-J}{6}\right) & E\left(\frac{L-K}{6}\right) \\ -E\left(\frac{L-J}{6}\right) & -E\left(\frac{J-K}{6}\right) & -E\left(\frac{K-L}{6}\right) \end{pmatrix}, \tag{3.27}$$

$$N = i\omega\frac{\alpha h^2}{2}I_2, \tag{3.28}$$

where $E(X) = \exp(iX)$. For example, the mass stencil for the RT_0 element is shown in Figure 3.5 and at node 1 (see Figure 3.3) we obtain

$$\frac{\alpha}{18}\left(10\tilde{u}_1 - \left(e^{iJ/2} + e^{-iJ/2}\right)\tilde{u}_2 - \left(e^{iL/2} + e^{-iL/2}\right)\tilde{u}_3\right),$$

which corresponds to the first line of M in (3.25).

For the 5×5 determinant to vanish we obtain

$$\omega_{1,2} = \pm\omega_{AN} + O(h^2), \qquad \omega_3 = 0, \qquad \omega_{4,5} = O(h^{-1}),$$

for infinitesimal mesh spacing.

The $RT_0 - P_1$ pair

Common to the $RT_0 - P_0$ and $RT_0 - P_1$ is a RT_0 representation of the velocity and they differ from one another in their representation of elevation. Consequently, the M and C matrices are given by (3.25) and (3.26), respectively. As for the $P_1^{NC} - P_1$ pair only one discrete amplitude is required for $\tilde{\eta}$ at mesh vertices and hence N is given by (3.23). The divergence stencil on Mesh 1 is shown in Figure (3.5) and the D matrix is obtained as

$$D = H\frac{2i}{3}\begin{pmatrix}\sin\left(\frac{J-L}{2}\right) & \sin\left(\frac{K-J}{2}\right) & \sin\left(\frac{L-K}{2}\right)\end{pmatrix}. \tag{3.29}$$

Vanishing the 4×4 determinant leads to

$$\omega_{1,2} = \pm\omega_{AN} + O(h^2), \qquad \omega_{3,4} = 0 \quad \text{(double root)},$$

for infinitesimal mesh spacing. For Mesh 2, M and N are found in [56] and C and D are deduced from (3.26) and (3.29) by using $J = kh$, $K = -kh + lh$, and $L = -lh$, as written in Figure 3.2. We obtain $\omega_{3,4} = 0$ and $\omega_{1,2}$ again coincide with ω_{AN} for infinitesimal mesh spacing.

The $BDM_1 - P_0$ pair

By using the notations of Figure 3.3 the following discrete amplitudes are considered

$$\tilde{\mathbf{u}} = (\tilde{u}_1, \cdots, \tilde{u}_6)^t, \qquad \tilde{\eta} = (\eta_A, \eta_B)^t. \tag{3.30}$$

After long and tedious algebra, we obtain from stencils of Figures (3.6) and (3.7) on Meshes 1 and 2

$$M = i\omega \frac{\alpha}{36} \begin{pmatrix} M_1(K) & M_2(J,K,L) & M_2(L,K,J) \\ M_2(J,L,K) & M_1(L) & M_2(K,L,J) \\ M_2(L,J,K) & M_2(K,J,L) & M_1(J) \end{pmatrix}, \tag{3.31}$$

$$C = f \frac{1}{24} \begin{pmatrix} 0 & -C_1(K,L) & C_1(K,J) \\ C_1(L,K) & 0 & -C_1(L,J) \\ -C_1(J,K) & C_1(J,L) & 0 \end{pmatrix}, \tag{3.32}$$

$$D = H \frac{1}{2} \begin{pmatrix} D_1(J,K,L) & D_1(K,L,J) & D_1(L,J,K) \end{pmatrix}, \tag{3.33}$$

where

$$\begin{aligned} M_1(J) &= \begin{pmatrix} 8 & 2E(J) \\ 2E(-J) & 8 \end{pmatrix}, \\ M_2(J,K,L) &= \begin{pmatrix} -E(K) - E(-L) & 2 - 2E(-J) \\ 2 - 2E(J) & -E(-K) - E(L) \end{pmatrix}, \\ C_1(K,L) &= \begin{pmatrix} E(K) + E(-L) & 2 \\ 2 & E(-K) + E(L) \end{pmatrix}, \\ D_1(J,K,L) &= \begin{pmatrix} -E\left(\frac{L-K}{3}\right) & -E\left(\frac{K-J}{3}\right) \\ E\left(\frac{J-K}{3}\right) & E\left(\frac{K-L}{3}\right) \end{pmatrix}. \end{aligned}$$

For example, the mass stencil for the BDM_1 element is shown in Figure 3.6 and at node 1 (see Figure 3.3) we obtain

$$\begin{aligned} \tfrac{\alpha}{36} \big(8\tilde{u}_1 + 2e^{iK}\tilde{u}_2 + \left(-e^{iK} - e^{-iL}\right)\tilde{u}_3 + \left(2 - 2e^{-iJ}\right)\tilde{u}_4 \\ + \left(-e^{-iJ} - e^{iK}\right)\tilde{u}_5 + \left(2 - 2e^{-iL}\right)\tilde{u}_6 \big) \end{aligned}$$

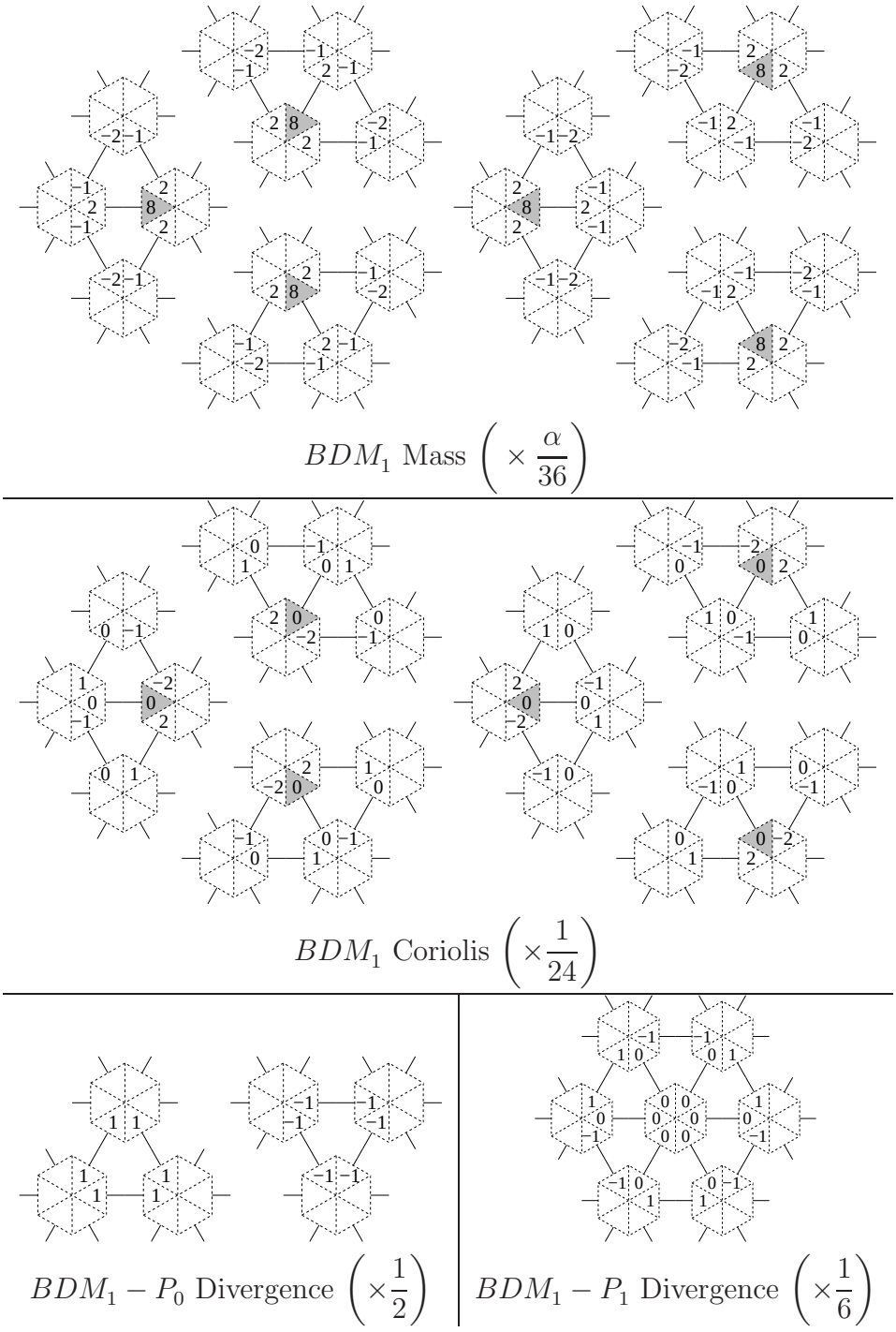


Figure 3.6: Stencils for the $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs on Mesh 1.

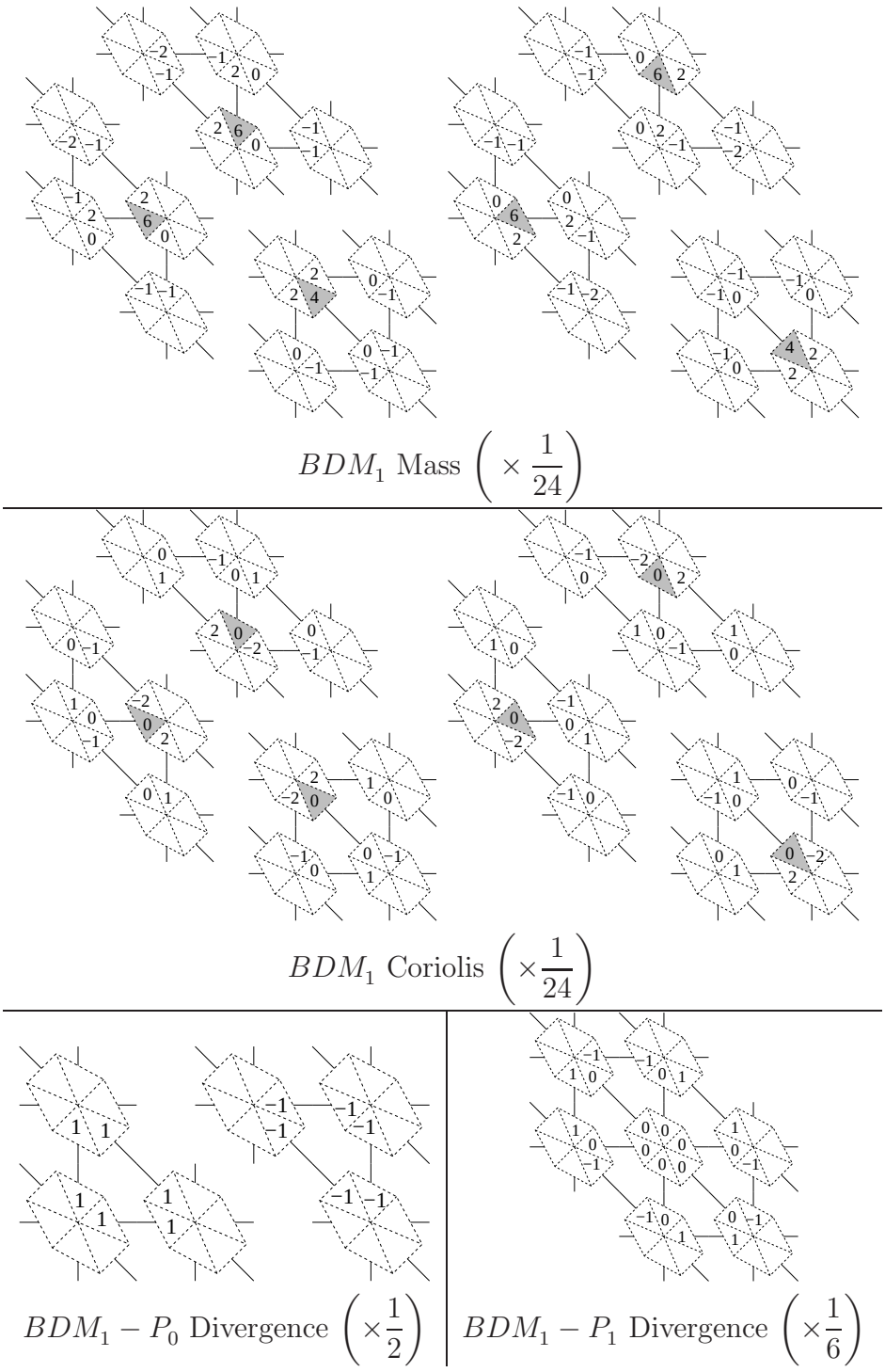


Figure 3.7: As for Figure 3.6 but on Mesh 2.

which corresponds to the first line of M in (3.31).

As for the $RT_0 - P_0$ the matrix N is given by (3.28). On Mesh 2 we obtain

$$M = i\omega \frac{1}{24} \begin{pmatrix} M_3 & M_5(K,J) & M_5(K,L) \\ M_5(-K,J) & M_4(L) & M_6(J,K,L) \\ M_5(-K,L) & M_6(L,K,J) & M_4(J) \end{pmatrix}, \tag{3.34}$$

where

$$M_3 = 4I_2,$$
$$M_4(L) = \begin{pmatrix} 6 & 2E(L) \\ 2E(-L) & 6 \end{pmatrix},$$
$$M_5(K,J) = \begin{pmatrix} -E(K) & 2 - E(-J) \\ 2 - E(J) & -E(-K) \end{pmatrix},$$
$$M_6(J,K,L) = \begin{pmatrix} -E(L) - E(-J) & -2E(-K) \\ -2E(K) & -E(-L) - E(J) \end{pmatrix}.$$

For Mesh 2, C and D are deduced from (3.32) and (3.33) by using the values of J, K, and L as in Figure 3.2. For the 8×8 determinant to vanish we obtain

$$\omega_{1,2} = \pm\omega_{AN} + O(h^2), \qquad \omega_{3,4,5,6} = 0 \quad \text{(quadruple root)}, \qquad \omega_{7,8} = O(h^{-1}),$$

for both meshes and infinitesimal mesh spacing.

The $BDM_1 - P_1$ pair

Common to the $BDM_1 - P_0$ and $BDM_1 - P_1$ is a BDM_1 representation of the velocity and they differ from one another in their representation of elevation. Consequently, the M and C matrices are given by (3.31), (3.34), and (3.32). As for the $P_1^{NC} - P_1$ pair only one discrete amplitude is required for $\tilde{\eta}$ at mesh vertices and hence N is given by (3.23). The divergence stencils on Mesh 1 and Mesh 2 are found in Figures 3.6 and 3.7 and the D matrix is written as

$$D = \frac{1}{6h} \begin{pmatrix} D_1(J,L) & D_1(K,J) & D_1(L,K) \end{pmatrix}, \tag{3.35}$$

where

$$D_1(J,L) = \begin{pmatrix} E(J) - E(L) & E(-L) - E(-J) \end{pmatrix}.$$

For both meshes, vanishing the 7×7 determinant leads to

$$\omega_{1,2} = \pm\omega_{AN} + O(h^4), \qquad \omega_{3,4,5} = 0 \quad \text{(triple root)}, \qquad \omega_{6,7} = O(h^2).$$

We note that similarly to the $P_1^{NC} - P_1$ pair the discrete inertia-gravity wave frequency are computed very accurately ($O(h^4)$) for infinitesimal mesh spacing.

3.6.3 Summary of discrete frequencies

The previous results are summarized in Table 3.1, where n is the dimension of the linear system (3.17), and hence the degree of the dispersion relation for the five FE pairs examined here. We mention the multiplicity of the discrete frequency types and the order of accuracy for inertia-gravity waves. For infinitesimal mesh spacing ($h \to 0$) we have $\omega_{1,2} \to \omega_{AN}$ for all schemes and hence discrete inertia-gravity frequencies are consistent with the continuous case. The slow mode corresponding to $\omega = 0$ is present for all pairs. However for the $RT_0 - P_1$, $BDM_1 - P_0$, and $BDM_1 - P_1$ pairs the multiplicity of this mode is greater than 1. Further, a $O(h^2)$ mode exists for the $BDM_1 - P_1$ pair. We also observe the presence of solutions of the form $\omega = \pm f$ for the $P_1^{NC} - P_1$ pair. These modes usually arise when the discrete scheme involves more velocity nodes than surface elevation nodes [52, 56]. They are propagating spurious inertial oscillations that have no particular spatial characteristics. Finally, the FE pairs having a discontinuous representation of surface elevation have spurious frequencies of type $O(h^{-1})$ as in [53, 56].

Table 3.1: Multiplicity of the discrete frequencies obtained from the dispersion relations for the five FE pairs examined in Section 3.6.2.

FE pair	n	$\omega = \pm\omega_{AN}$	$\omega = 0$	$\omega = \pm f$	$\omega = O(h^{-1})$	$\omega = O(h^2)$
$P_1^{NC} - P_1$	7	$2 \;:\; O(h^4)$	1	4	0	0
$RT_0 - P_0$	5	$2 \;:\; O(h^2)$	1	0	2	0
$RT_0 - P_1$	4	$2 \;:\; O(h^2)$	2	0	0	0
$BDM_1 - P_0$	8	$2 \;:\; O(h^2)$	4	0	2	0
$BDM_1 - P_1$	7	$2 \;:\; O(h^4)$	3	0	0	2

3.6.4 Gravity wave limit of discrete frequencies

We now analyze the computed frequencies $\omega_{CP} \equiv \omega_{1,2}$ for the five FE pairs examined in section 3.6.2. Because the Coriolis factor does not have a significant impact on the propagation of gravity waves we let $f = 0$ to obtain the gravity wave limit. From (3.16) and the results of section 3.6.2, we determine the analytical and computed phase speeds, denoted by c_{AN} and c_{CP}, respectively,

$$c_{AN} \equiv \frac{\omega_{AN}}{\sqrt{k^2 + l^2}} = \pm\sqrt{gH}, \qquad c_{CP} \equiv \frac{\omega_{CP}}{\sqrt{k^2 + l^2}}.$$

The phase speed ratio, denoted by r_{PH}, is then computed as the ratio of the computed phase speed to the analytical one, with

$$r_{PH} \equiv \left|\frac{c_{CP}}{c_{AN}}\right| = \left|\frac{\omega_{CP}}{\omega_{AN}}\right| = \frac{|\omega_{CP}|}{\sqrt{gH(k^2+l^2)}}. \tag{3.36}$$

Note that we should have $r_{PH} = 1$ in the absence of numerical dispersion. We show r_{PH} as a surface function depending on the normalized wave numbers kh and lh for Meshes 1 and 2 in Figure 3.8. The values of kh and lh for Mesh 1 vary between $\pm 4\pi/3$ and $\pm\pi/\alpha$, respectively, in order to include the six points defined later in (3.37). For Mesh 2, kh and lh vary between $\pm\pi$ as in [56, 25]. The phase advance (r_{PH}) is also plotted in Figure 3.9 along the selected axes OE and OT for Mesh 1 and OX, OD1, and OD2 for Mesh 2. The directions of these axes are defined in Figure 3.8 and Table 3.2. For symmetrical reasons the phase advance surfaces are only shown for positive wave numbers in Figure 3.9.

Table 3.2: Normalized wave numbers relations for the selected axes defined in Figure 3.8.

Mesh 1		Mesh 2	
OE :	$lh = 0$ or $kh = \pm\frac{2\alpha}{3}lh$	OX :	$lh = 0$ or $kh = 0$
OT :	$kh = 0$ or $kh = \pm 2\alpha\, lh$	OD1:	$kh = lh$
		OD2:	$kh = -lh$

For the $RT_0 - P_0$ pair we obtain $r_{PH} \leq 1$ on Mesh 1 for all values of kh and lh and the waves do not accelerate on Mesh 1. On Mesh 2, we have $r_{PH} \leq 1$ in the OX and OD1 directions but this is not the case in the OD2 direction since the maximum value observed for r_{PH} is 1.18. The $P_1^{NC} - P_1$ pair behaves similarly although a weak acceleration ($r_{PH} = 1.02$) is noticeable in the OT, OX, and OD1 directions. For larger wave numbers the $RT_0 - P_0$ and $P_1^{NC} - P_1$ pairs present less dispersion effects than the three other ones examined in Section 3.6.2 because r_{PH} is closer to 1 for those pairs.

For the $RT_0 - P_1$ and $BDM_1 - P_1$ pairs we have $r_{PH} \leq 1$ for all values of kh and lh on both meshes. Further we have $r_{PH} = 0$ for

$$(kh, lh) = \left(\pm\frac{4\pi}{3}, 0\right), \left(\pm\frac{2\pi}{3}, \pm\frac{\pi}{\alpha}\right), \quad \text{on Mesh 1,} \tag{3.37}$$

$$(kh, lh) = \pm\left(\frac{2\pi}{3}, -\frac{2\pi}{3}\right), \quad \text{on Mesh 2.} \tag{3.38}$$

The wave numbers in (3.37) and (3.38) correspond to spurious surface-elevation modes introduced by the spatial discretization scheme and they are represented in Figure 3.10.

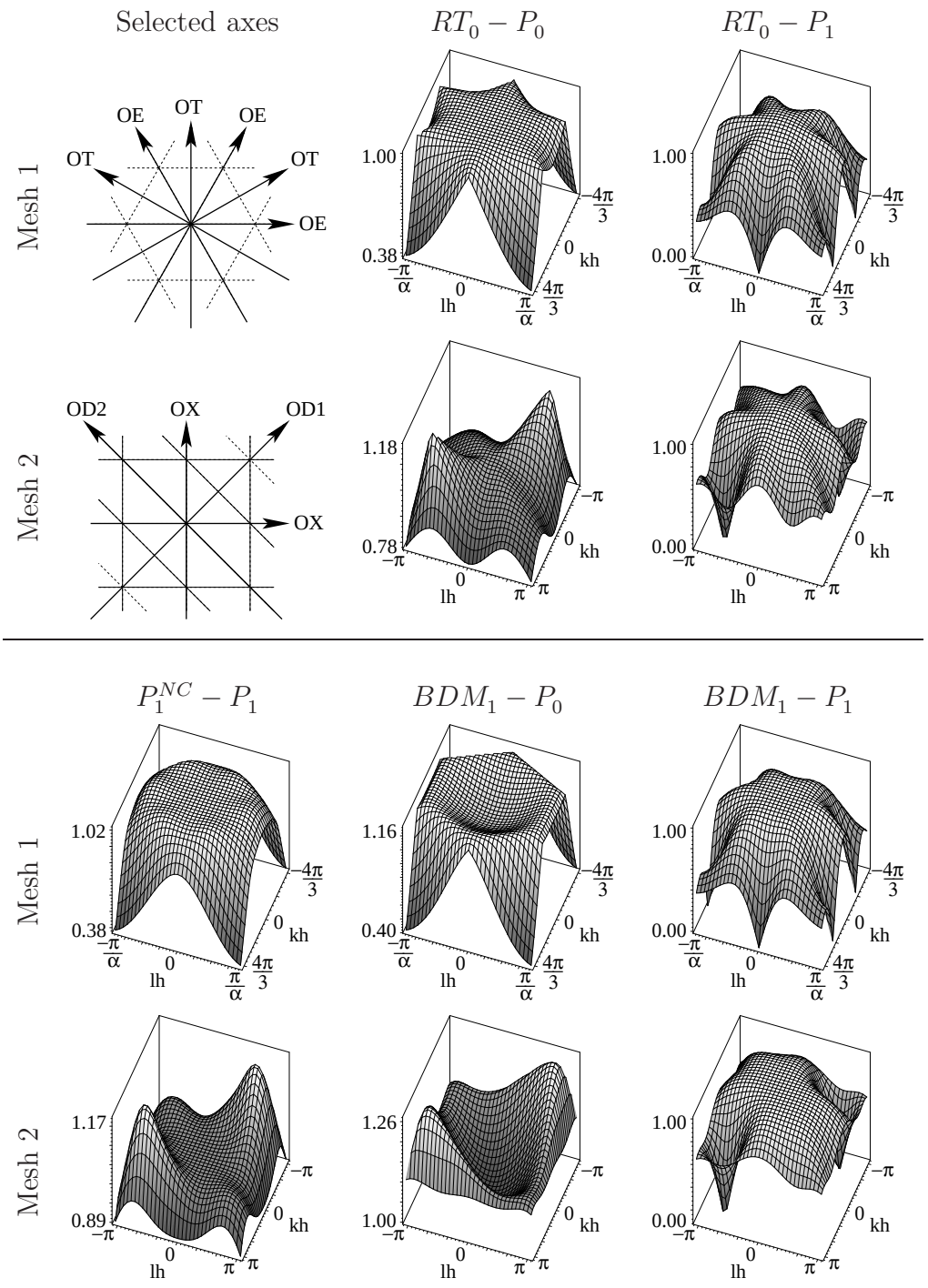


Figure 3.8: Definition of selected axes and the phase speed ratio (r_{PH}) as a surface function on Meshes 1 and 2 for the $RT_0 - P_0$, $RT_0 - P_1$, $P_1^{NC} - P_1$, $BDM_1 - P_0$, and $BDM_1 - P_1$ pairs.

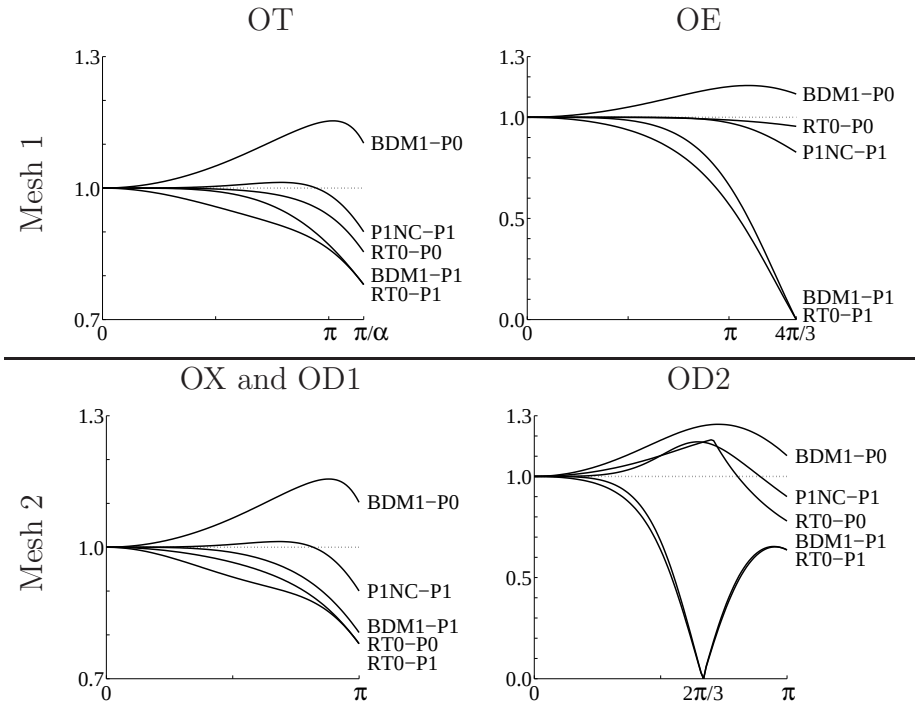


Figure 3.9: The phase speed ratio (r_{PH}) along selected axes in function of the normalized wave number.

These modes do not propagate but are trapped within the model grid which usually leads to noisy solutions. For the $BDM_1 - P_0$ we observed that $r_{PH} \geq 1$ on Mesh 1 for values of kh and lh inside the hexagonal area defined by the six points in (3.37). On Mesh 2 we have $r_{PH} \geq 1$ for all kh and lh values. Consequently, the gravity waves always accelerate with such a discretization scheme.

3.6.5 Canal simulation

In this test, equations (3.1) and (3.2) are solved using $f = 0$. The purpose of the experiment is to validate the analytical results obtained in Section 3.6.4 on Mesh 2 in the OX and OD2 directions. For the OX direction, Mesh 2 is considered while in the OD2 direction we employ a mesh obtained from Mesh 2 by a rotation of $\pi/4$ as in [56, Figure 5.1]. The domain extent is 2 000 km × 280 km. The resolution is set to $h = 10$ km. The fluid is initially at rest and zero normal velocity is specified at the boundaries, except at the western one. On the western side, the fluid velocity is set to 0.1 $\mathrm{m\,s^{-1}}$ inbound and $\eta = 1$ m, with $H = 1\,000$ m and $g = 10\,\mathrm{m\,s^{-2}}$, i.e. the phase speed of gravity waves is $\sqrt{gH} = 100\,\mathrm{m\,s^{-1}}$. The Crank-Nicolson scheme is used for the time discretization. The time step is set to 10 s and the gravitational Courant number is

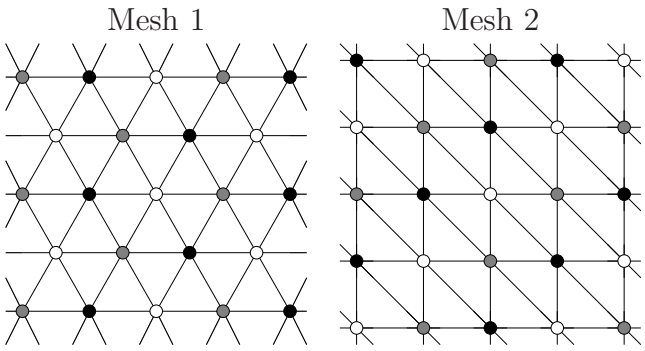


Figure 3.10: Zero frequency surface-elevation modes for the $RT_0 - P_1$ and $BDM_1 - P_1$ pair s: Mesh 1 (left) and Mesh 2 (right).

thus 0.1, such that the time step has no significant impact on the wave dispersion. The duration of the simulation is 1 000 time steps and the wave front should be located at midbasin, i.e. at 1 000 km from the western boundary, at the end of the simulation.

In Figure 3.11 the numerical solutions obtained for all FE pairs are shown for both the OX and OD2 directions. The results for the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs are identical to those displayed in [56] and they are reproduced here for comparison purposes with the other pairs. The $RT_0 - P_1$ pair exhibits the strongest oscillations trailing behind the front in both directions, while the region ahead of the front is free of oscillations. This result is in good agreement with the analysis conducted in Section 3.6.4 where $r_{PH} \leq 1$. Small oscillations appearing at the beginning of the simulation travel about two times as fast as the front and they are observed at the eastern part of the domain in Figure 3.11. The presence of these oscillations is beyond the scope of our analysis. The results for the $BDM_1 - P_1$ and $RT_0 - P_1$ pairs are similar but the amplitude of the oscillations behind the front is slightly smaller for the $BDM_1 - P_1$ pair. The results for the $BDM_1 - P_0$ pair show oscillations travelling faster than the wavefront while the region behind the front is noise free. Again these results are in good agreement with those of Section 3.6.4 because $r_{PH} \geq 1$ in the OX and OD2 directions for all kh and lh.

The experiment has also been performed on Mesh 1 in the OE and OT directions but the results are not presented here as they are very similar to those obtained on Mesh 2.

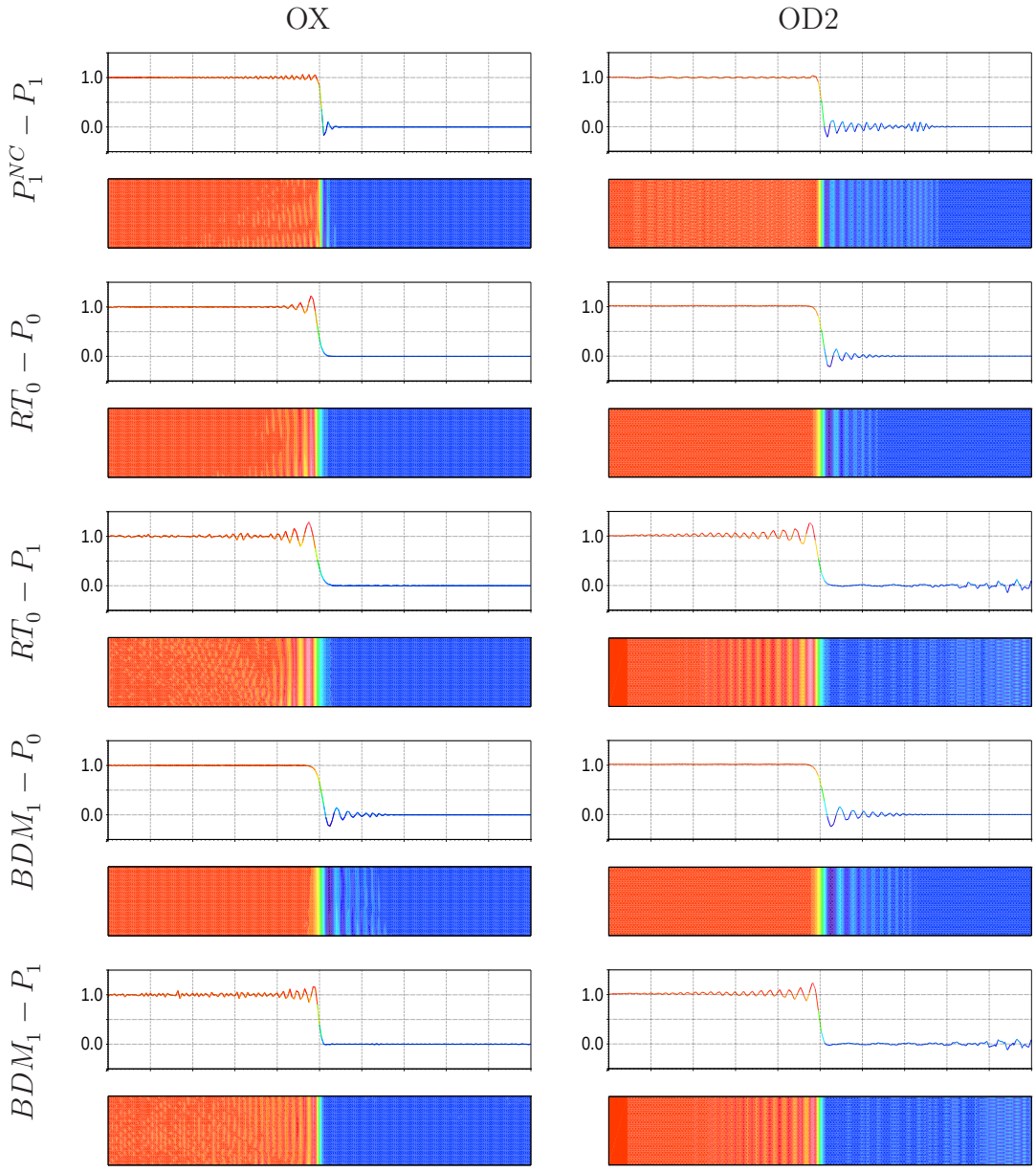


Figure 3.11: Surface elevation for the canal test after 10 000 s of simulation on Mesh 2.

Table 3.3: Dimension of the discrete operator kernels and Λ on a $n \times n$ regular Mesh 2 with no normal flow boundary condition.

	$P_1^{NC} - P_1$	$RT_0 - P_0$	$RT_0 - P_1$	$BDM_1 - P_0$	$BDM_1 - P_1$
p	$6\ n^2$	$3(n-\frac{1}{3})^2 - \frac{1}{3}$	$3(n-\frac{1}{3})^2 - \frac{1}{3}$	$6(n-\frac{1}{3})^2 - \frac{2}{3}$	$6(n-\frac{1}{3})^2 - \frac{2}{3}$
q	$(n+1)^2$	$2\ n^2$	$(n+1)^2$	$2\ n^2$	$(n+1)^2$
$\underline{C}$	0	$(n-1)^2+1$	$(n-1)^2+1$	$2(n-1)^2+2$	$2(n-1)^2+2$
$\underline{G}$	1	1	3	1	3
$\underline{D}$	$5(n-\frac{1}{5})^2 - \frac{1}{5}$	$(n-1)^2$	$2(n-1)^2$	$4(n-\frac{1}{2})^2$	$5(n-\frac{3}{5})^2 + \frac{1}{5}$
$\underline{CD}$	0	0	$(n-1)^2$	$2(n-1)^2+1$	$2(n-1)^2+1$
$\underline{CG}$	$(n+1)^2$	$2\ n^2$	$2\ n^2+2$	$4(n-\frac{1}{2})^2+2$	$3(n-\frac{1}{3})^2 + \frac{11}{3}$
$\underline{CDG}$	$(n-1)^2+1$	$(n-1)^2+1$	$2(n-1)^2+3$	$4(n-\frac{1}{2})^2+1$	$3(n-\frac{1}{3})^2 + \frac{8}{3}$
Λ	$(n-1)^2$	$(n-1)^2$	$(n-1)^2$	$2\ n^2-2$	$(n+1)^2-4$

3.7 Geostrophic balance

3.7.1 Kernel analysis

In this analysis we seek stationary solutions to the SW system (3.4) and (3.5) as done in [50] for the P_0-P_1, $P_1^{NC}-P_1$, P_1-P_1, MINI, $P_1\text{iso}P_2-P_1$, and P_2-P_1 pairs. Here, we examine the RT_0-P_1, BDM_1-P_0, and BDM_1-P_1 pairs. Stationary solutions are obtained by substituting $\omega = 0$ in (3.9) and (3.10) and this leads to

$$\begin{pmatrix} \underline{C} & \underline{G} \\ \underline{D} & 0 \end{pmatrix} \begin{pmatrix} \underline{\mathbf{u}} \\ \underline{\eta} \end{pmatrix} = 0. \tag{3.39}$$

The stationary solutions thus belong to the kernel of the matrix in the left hand side of (3.39) denoted by $\underline{CDG}$. We also define

$$\underline{CD} \equiv \begin{pmatrix} \underline{C} \\ \underline{D} \end{pmatrix}, \qquad \underline{CG} \equiv \begin{pmatrix} \underline{C} & \underline{G} \end{pmatrix}. \tag{3.40}$$

We consider a square domain with no-normal flow boundary condition and the regular $n \times n$ Mesh 2. The dimension of the $\underline{C}, \underline{D}, \underline{G}, \underline{CD}, \underline{CG}$, and $\underline{CDG}$ matrix kernels are computed numerically using MATLAB for $n = 3, \cdots, 14$, and the results are extrapolated for any integer n and given in Table 3.3. We observe that the $\underline{C}$ matrix has a non-trival kernel for the RT_0 and the BDM_1 elements. For the latter the kernel dimension is twice bigger than the RT_0 one. The rank deficiency for the $\underline{C}$ matrix is a consequence of using only one velocity component at the velocity nodes as for the

Arakawa C-grid [1, 41]. For the $P_1^{NC} - P_1$ pair both velocity components are used at velocity nodes and the $\underline{C}$ matrix is thus full rank.

We note that the dimension of the $\underline{G}$ matrix kernel is equal to one for the $P_1^{NC} - P_1$, $RT_0 - P_0$, and $BDM_1 - P_0$ pairs. This solution corresponds to the hydrostatic surface-elevation mode, i.e. the solution with constant elevation and zero velocity. It can be simply considered as a constant of integration associated with the solution of the governing equations. However, for the $RT_0 - P_1$ and $BDM_1 - P_1$ pairs the dimension of the $\underline{G}$ matrix kernel is equal to 3 and hence 2 spurious surface-elevation modes are present.

The dimension of the $\underline{D}$ matrix kernel can be deduced from (3.18) and by using the rank theorem we obtain

$$\dim(\ker(\underline{D})) = p - q + \dim(\ker(\underline{G})). \tag{3.41}$$

The $\underline{CD}$ matrix kernel is the intersection of the $\underline{C}$ and $\underline{D}$ matrix kernels. For the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs this intersection is empty. The dimension of the $\underline{CD}$ matrix kernel is $(n-1)^2$ for the $RT_0 - P_1$ pair while it is twice greater for the $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs as shown in Table 3.3.

The modes lying in the $\underline{CD}$ matrix kernel, and named here CD-modes, behave similarly to the spurious surface-elevation modes but they belong to the velocity space instead. The fact that the number of CD-modes is $O(n^2)$ compared to $O(1)$ for the spurious surface-elevation modes, suggests that the CD-modes may appear locally on the mesh which is usually not the case for the elevation modes.

The $\underline{C}$ matrix is skew-symmetric and by using (3.18) and the rank theorem we deduce

$$\dim(\ker(\underline{CG})) = q + \dim(\ker(\underline{CD})). \tag{3.42}$$

The $\underline{CDG}$ matrix kernel contains all stationary modes by definition, including the hydrostatic mode, possible spurious surface-elevation and CD-modes. It also contains other modes that are solution of the discrete geostrophic balance, i.e. the balance between the Coriolis and pressure gradient operators. The number of such modes is given by

$$\Lambda \equiv \dim(\ker(\underline{CDG})) - \dim(\ker(\underline{G})) - \dim(\ker(\underline{CD})), \tag{3.43}$$

and it is mentioned in the last line of Table 3.3 for all pairs. For the three first pairs in Table 3.3 we obtain $\Lambda = (n-1)^2$ which corresponds to the number of mesh vertices that dot not lie on the boundary. For the two last pairs we have $\Lambda = 2n^2 - 2$ and

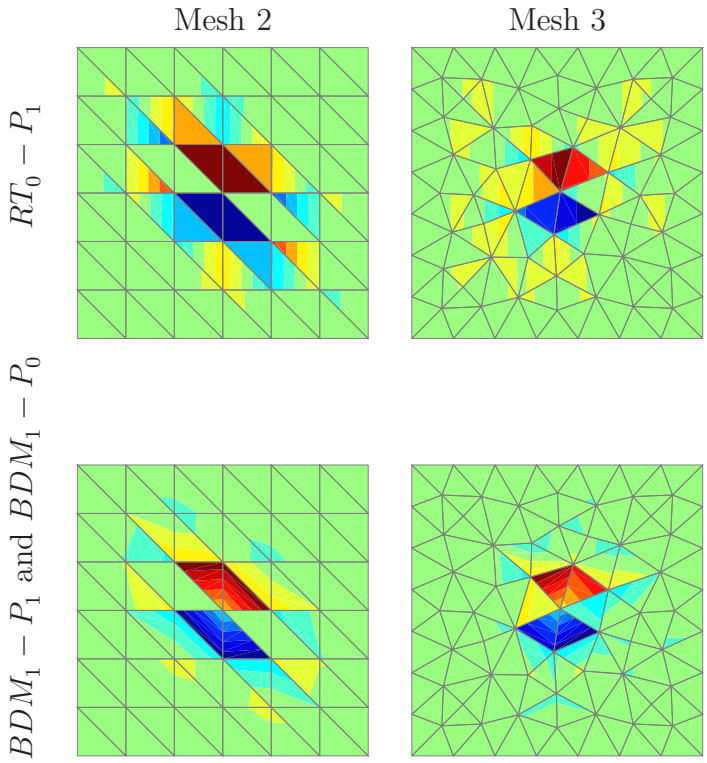


Figure 3.12: x-component of the velocity for a typical $RT_0 - P_1$, $BDM_1 - P_0$, and $BDM_1 - P_1$ smallest representable vortex.

$(n+1)^2 - 4$, respectively. For the three first pairs, each vertex can be associated with a smallest representable vortex (SRV) defined in [50] as the stationary solution of (3.39) with minimal support. As in [50], the SRV form a basis for the discrete geostrophic balance. A typical SRV is presented in Figure 3.12 for the $RT_0 - P_1$, $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs on Mesh 2 and Mesh 3, an unstructured mesh with smoothing. We see that the SRV structure is larger for these three pairs than for the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs examined in [50] on both meshes. Note that SRV with more complex structures also exist near the boundary for the $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs.

3.7.2 Propagating eddy simulation

In this experiment our purpose is to validate the results obtained in Table 3.3. The slowly propagating Rossby modes are simulated in the case of the evolution of a typical anticyclonic eddy at midlatitudes. The domain is a 2 000 km $\times$ 1 200 km rectangular basin and the triangulation as a resolution of 20 km. The β-plane approximation, $f = f_0 + \beta y$, is used where $f_0 = 6.1634 \times 10^{-5}$ s^{-1}, $\beta = 2.0746 \times 10^{-11}$ m^{-1} s^{-1}

and the choice $g = 9.81$ m s^{-2}, $H = 1.6309$ m results in a phase speed for gravity waves of $\sqrt{gH} = 4\,\text{m s}^{-1}$. The radius of deformation at midbasin is $\sqrt{gH}/f_0 = 65$ km. A Gaussian distribution centered in the domain is prescribed at initial time for the elevation and the initial velocity is in geostrophic balance

$$\mathbf{u} = -2\frac{g}{f}\frac{A}{r^2}\exp\left(-\frac{|\mathbf{x}|^2}{r^2}\right)\mathbf{k}\times\mathbf{x}, \tag{3.44}$$

$$\eta = A\exp\left(-\frac{|\mathbf{x}|^2}{r^2}\right), \tag{3.45}$$

where $r = 1.3\times10^5$ m and $A = 0.95$ m. The Crank-Nicolson scheme is again used and the time step is set to 1 800 s. The vortex moves slowly westward as predicted by the Rossby wave dynamics. Figures 3.13 and 3.14 show the surface-elevation and flow speed field after 5 weeks of simulation on Mesh 2 and the unstructured Mesh 3, respectively. Note that the solutions have been linearly interpolated to represent continuous isolines.

On Mesh 2, the solution is very smooth for both the elevation and flow speed field. The solutions are nearly identical for the five FE pairs and minor differences are barely observed. On Mesh 3, the surface elevation and flow speed field are very similar to those obtained on Mesh 2 in Figure 3.13 for the $P_1^{NC} - P_1$, $RT_0 - P_0$ and $RT_0 - P_1$ pairs. However, for the $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs the situation is very different. Indeed, the flow speed field exhibits severe oscillations that lead to unstable results. The observed oscillations appear early in the simulation after only few time steps and gradually grow in time. Although the oscillations in the flow speed field rapidly increase, the surface elevation remains coherent with the presence of mild oscillations. We suspect that the presence of CD-modes in Table 3.3 might be responsible for the oscillations observed in Figure 3.14 for the $BDM_1 - P_0$ and $BDM_1 - P_1$ pairs. In that case the CD-modes, although present, are not triggered by the use of an unstructured mesh for the $RT_0 - P_1$ pair.

3.8 Conclusion

An analysis of the Raviart-Thomas and Brezzi-Douglas-Marini FE pairs is presented to determine the ability of these pairs in solving the SW equations. An inertia-gravity wave dispersion analysis is performed on meshes made up of equilateral and right biased triangles. Spurious surface elevation modes are observed for the $RT_0 - P_1$ and $BDM_1 - P_1$ pairs. The analysis also permits to show that the $RT_0 - P_0$ and $P_1^{NC} - P_1$ pairs have

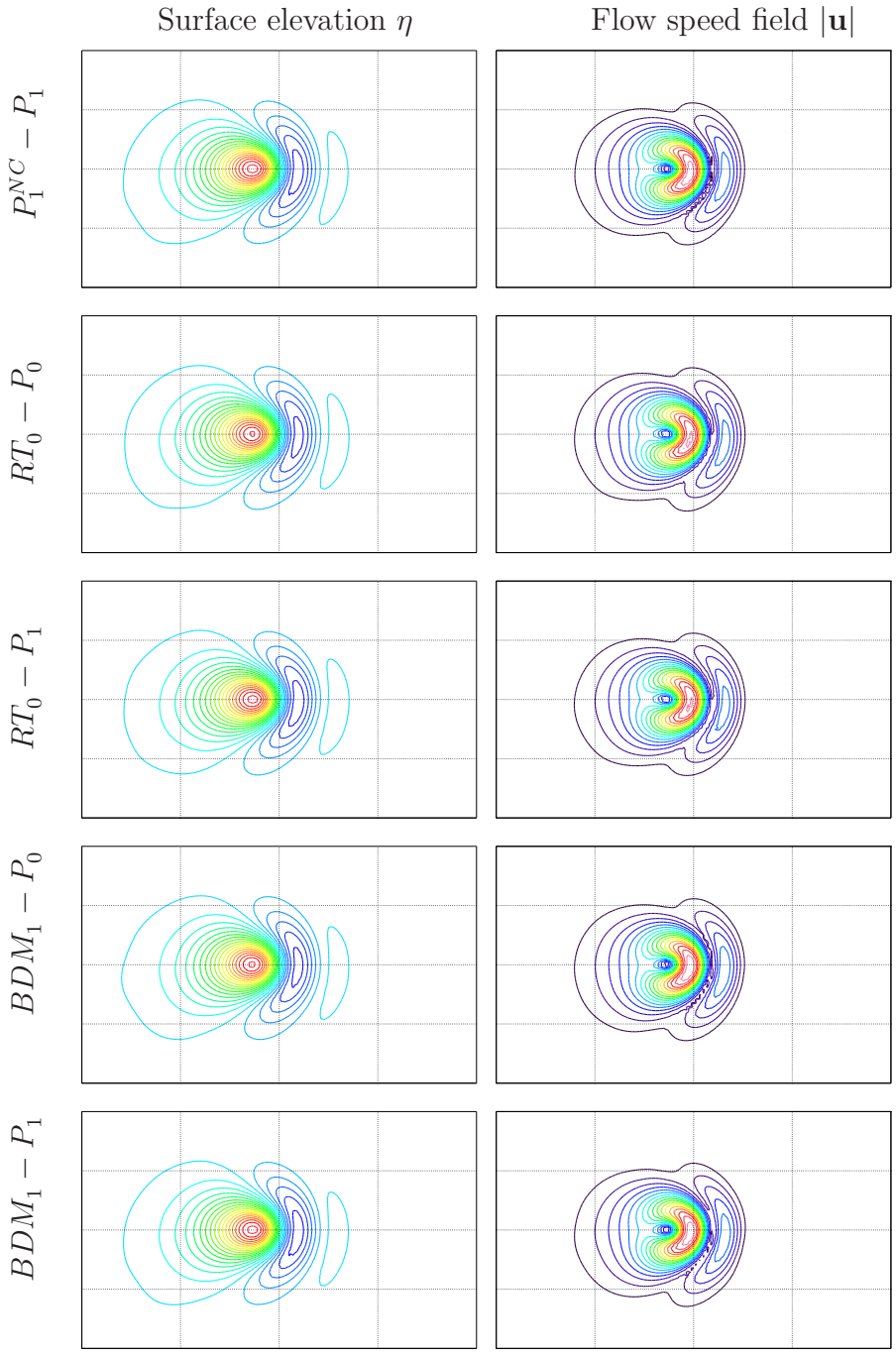


Figure 3.13: Surface elevation and flow speed field after 5 weeks of simulation on Mesh 2.

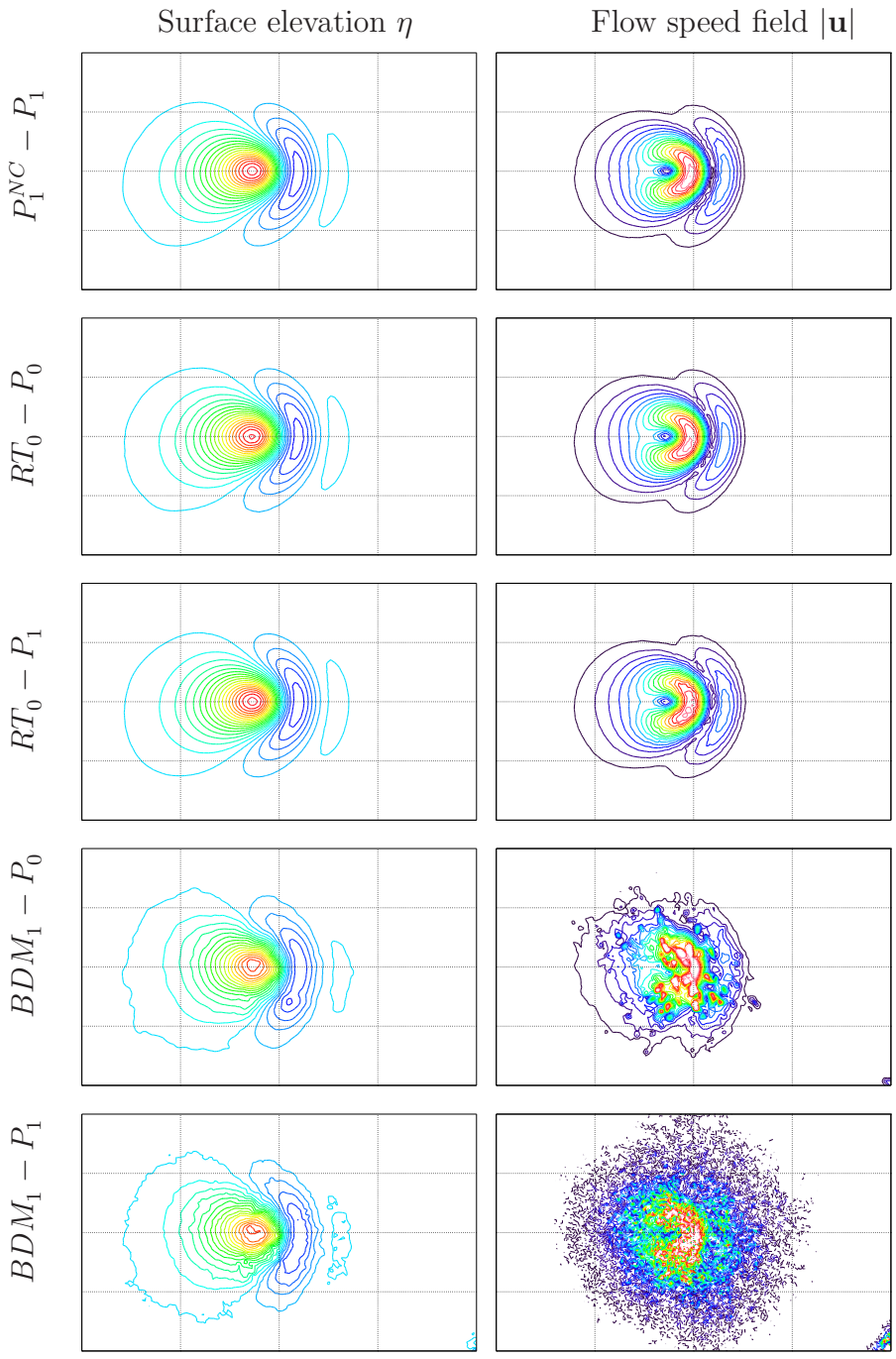


Figure 3.14: As for figure 3.13, but on Mesh 3.

the lowest amount of dispersion on both meshes compared to the $RT_0 - P_1$, $BDM_1 - P_0$, and $BDM_1 - P_1$ ones. We note that solutions on equilateral meshes present less dispersion than on meshes made up of right biased triangles. The simulation results of a gravity wave propagating in a canal are in good agreement with the analytical computations. The geostrophic equilibrium is investigated through a linear algebra kernel computation approach. Such an analysis shows the presence of spurious CD-modes in the velocity space for the RT_0-P_1, BDM_1-P_0, and BDM_1-P_1 pairs. Smallest representable vortex structures have been computed for all pairs and compared with the $RT_0 - P_0$ and $P_1^{NC} - P_1$ ones on both structured and unstructured meshes. Numerical solutions of a propagating eddy at midlatitudes give smooth elevation and flow speed field on the uniform mesh for all pairs. Those results still hold for the $P_1^{NC} - P_1$, $RT_0 - P_0$, and $RT_0 - P_1$ pairs on the unstructured triangulation while unstable results are obtained for the $BDM_1 - P_0$, and $BDM_1 - P_1$ pairs in the representation of the flow speed field only. We suspect the CD-modes to be responsible for this behavior.

4. CONCLUSION

In Chapter 1, we presented the dispersion analyses of nine finite element pairs on the regular mesh made up of right biased triangles. We observed that the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs are significantly less dispersive for inertia-gravity waves, provided that the deformation radius of Rossby is well resolved for the $RT_0 - P_0$ pair. We noted the presence of spurious $O(h^{-1})$ modes for the pairs with discontinuous surface elevation representation.

In Chapter 2, a linear algebra approach was developed to characterize the kernels of the discrete operators. Three kernel relations were identified as necessary conditions for the discretized system to share the same stationary properties as the continuous system. The discrete kernels were investigated to ascertain the presence, number and structure of spurious modes. We introduced the concept of the smallest representable vortex (SRV) to characterize the set of stationary modes. On a structured mesh, we observed that the SRV of the $P_1^{NC} - P_1$, $RT_0 - P_0$, and $P_0 - P_1$ pairs are smaller than those of the $P_1 - P_1$, MINI, $P_2 - P_1$, and $P_1 iso P_2 - P_1$ pairs. Surprisingly, on a unstructured mesh the SRV were only observed for the $P_1^{NC} - P_1$, $RT_0 - P_0$, and $P_0 - P_1$ pairs. A simulation showed that the absence of SRV on unstructured grids could lead to numerical instabilities for time dependent vortex flows. Another numerical test showed that strong modal oscillations are present in the geostrophic balance approximation for the MINI, $P_2 - P_1$, and $P_1 iso P_2 - P_1$ pairs. This phenomenon is also observed for the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs but with lesser amplitude. No modal decoupling is present for the $P_0 - P_1$ and $P_1 - P_1$ pairs.

In Chapter 3, we analysed the Raviart-Thomas and Brezzi-Douglas-Marini finite element spaces. Dispersion analyses were performed on meshes made up of equilateral and right biased triangles. We noted the presence of spurious surface elevation modes for the $BDM_1 - P_1$ and $RT_0 - P_1$ pairs. The discrete frequencies for inertia-gravity waves of the $BDM_1 - P_1$, $BDM_1 - P_0$, and $RT_0 - P_1$ pairs are consistent with the continuous frequencies. However, we observed that the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs are less dispersive. The presence of SRV was noted on both structured and unstructured meshes. We also observed a large number of spurious CD-velocity modes for the

$BDM_1 - P_1$, $BDM_1 - P_0$, and $RT_0 - P_1$ pairs. On the uniform mesh, the numerical solutions of a propagating eddy at midlatitudes gave smooth elevation and flow speed field for all pairs. On the unstructured mesh, those results still hold for the $P_1^{NC} - P_1$, $RT_0 - P_0$, and $RT_0 - P_1$ pairs while unstable results are obtained for the $BDM_1 - P_1$ and $BDM_1 - P_0$ pairs in the representation of the velocity field. We suspect the CD-modes to be responsible of this behavior.

In conclusion to this study, we highly recommend the $P_1^{NC} - P_1$ and $RT_0 - P_0$ finite element pairs for solving the SWE. The numerical dispersion of inertia-gravity waves for these pairs is minimal when compared to the other pairs. The $P_1^{NC} - P_1$ pair also presents the advantage of a diagonal velocity mass matrix due to the orthogonality of P_1^{NC} base functions. This feature can be used to greatly reduce the numerical workload required to solve the system. Note that similar results can be achieved for the $RT_0 - P_0$ pair with a mass lumping procedure. On unstructured meshes, the presence of SRV structures for the $P_1^{NC} - P_1$ and $RT_0 - P_0$ pairs gives them the upper hand in the approximation of geostrophic flows although some oscillations due to modal decoupling may be visible. To avoid the modal decoupling problem one may consider the $P_0 - P_1$ pair. The approximation of the geostrophic flows is better with this pair but this scheme is more disperse for inertia-gravity waves.

BIBLIOGRAPHY

[1] A.J. Adcroft, C.N. Hill, and J.C. Marshall, *A new treatement of the coriolis term in C-grid models at both high and low resolutions*, Monthly Weather Review, 127 (1999), pp. 1928–1936.

[2] V. Agoshkov, E. Ovtchinnikov, V. Pennati, D. Ambrosi, and F. Saleri, *Finite element, finite volume, and finite differences approximation to the shallow water equations*, in Finite Elements in Fluids, K. Morgan, E. Oñate, J. Periaux, J. Peraire, and O.C. Zienkiewicz, eds., Pineridge Press, Barcelona, Spain, 1993, pp. 1001–1009.

[3] V. Aizinger and C.N. Dawson, *A discontinuous Galerkin method for two-dimensional flow and transport in shallow water*, Adv. in Water Res., 25 (2002), pp. 67–84.

[4] F. Alcrudo and P. Garcia-Navarro, *A high-resolution Godunov-type scheme in finite volumes for the 2D shallow water equations*, Int. J. Numer. Methods Fluids, 16 (1993), pp. 489–505.

[5] K. Anastasiou and C.T. Chan, *Solution of the 2D shallow water equations using the finite volume method on unstructured triangular meshes*, Int. J. Numer. Methods Fluids, 24 (1997), pp. 1225–1245.

[6] A. Arakawa and V.R. Lamb, *Computational design of the basic dynamical processes of the UCLA general circulation model*, Methods Comput. Phys., 17 (1977), pp. 173–265.

[7] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[8] D.N. Arnold, F. Brezzi, and M. Fortin, *A stable finite element for the Stokes equations*, Calcolo, 21 (1984), pp. 337–344.

[9] J.H. Atkinson, J.J. Westerink, and J.M. Hervouet, *Similarities between the wave equation and the quasi-bubble solutions to the shallow water equations*, Int. J. Numer. Methods Fluids, 45 (2004), pp. 689–714.

[10] J.H. Atkinson, J.J. Westerink, and Jr. R.A. Luettich, *Two dimensional dispersion analyses of finite element approximations to the shallow water equations*, Int. J. Numer. Methods Fluids, 45 (2004), pp. 715–749.

[11] M.L. Batteen and Y.J. Han, *On the computational noise of finite-difference schemes used in ocean models*, Tellus, 33 (1981), pp. 387–396.

[12] M. Bercovier and O. Pironneau, *Error estimates for the finite element method solution of the Stokes problem in the primitive variables*, Numer. Math., 33 (1979), pp. 211–224.

[13] F. Brezzi, J. Douglas, and L.D. Marini, *Recent results on mixed finite element methods for second order elliptic problems*, in Vistas in Applied Math., Numerical Analysis, Atmospheric Sciences, Immunology, Balakrishanan, Dorodnitsyn, and Lions, eds., Optimization Software Publications, 1986.

[14] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, Berlin, 1991.

[15] G.F. Carey, ed., *Finite Element Modeling of Environmental Problems*, John Wiley and Sons, Chichester, UK, 1995.

[16] S. Chippada, C.N. Dawson, M.L. Martinez, and M.F. Wheeler, *A Godunov-type finite volume method for the system of shallow water equations*, Comput. Methods Appl. Mech. Engrg., 151 (1998), pp. 105–129.

[17] B. Choi, M. Iskandarani, J.C. Levin, and D.B. Haidvogel, *A spectral finite-volume method for the shallow-water equations*, Mon. Wea. Rev., 132 (2004), pp. 1777–1791.

[18] P.G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[19] B. Cockburn, E. Karniadakis, and C.W. Shu, eds., *Discontinuous Galerkin Methods-Theory, Computation, and Applications*, Lect. Notes Comput. Sci. Engrg. 11, Springer-Verlag, Berlin, 2000.

[20] J. Côté and A. Staniforth, *An accurate and efficient finite-element global model of the shallow-water equations*, Mon. Wea. Rev., 118 (1990), pp. 2707–2717.

[21] B. Cushman-Roisin, *Introduction to Geophysical Fluid Dynamics*, Prentice Hall, 1994.

[22] S. Danilov, G. Kivman, and J. Schröter, *A finite-element ocean model: Principles and evaluation*, Ocean Modelling, 6 (2004), pp. 125–150.

[23] C.N. Dawson and S. Chippada, *Numerical modeling of shallow water flows with wetting and drying boundaries by a finite volume method*, in Proceedings of the 1998 Annual Conference on Mission Earth: Modeling and Simulation of the Earth System, Western MultiConference, 1998, pp. 9–14.

[24] C.N. Dawson and J. Proft, *Coupled discontinuous and continuous Galerkin finite element methods for the depth-integrated shallow-water equations*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 289–318.

[25] M.G.G. Foreman, *A two-dimensional dispersion analysis of selected methods for solving the linearized shallow-water equations*, J. Comput. Phys., 56 (1984), pp. 287–323.

[26] M. Fortin, *Old and new finite elements for incompressible flows*, Int. J. Numer. Methods Fluids, 1 (1981), pp. 347–364.

[27] L.P. Franca and C. Farhat, *On the limitation of bubble functions*, Comput. Methods Appl. Mech. Engrg., 117 (1994), pp. 225–230.

[28] V. Girault and P.A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer Ser. Comput. Math. 5, Springer-Verlag, Berlin, 1986.

[29] C.M. Gossard and R.L. Kolar, *Phase behavior of a finite volume shallow water algorithm*, in Proceedings of the CMWR XIII Volume 2: Computational Methods, Surface Water Systems, and Hydrology, L.R. Bentley et al., ed., Rotterdam, 2000, pp. 921–928.

[30] E. Hanert, *Towards a Finite Element Ocean Circulation Model*, PhD thesis, UCL, 2004.

[31] E. Hanert, V. Legat, and E. Deleersnijder, *A comparison of three finite elements to solve the linear shallow water equations*, Ocean Modelling, 5 (2003), pp. 17–35.

[32] P. Hood and C. Taylor, *Navier-Stokes equations using mixed interpolation*, in Finite Elements in Flow Problems, J.T. Oden, O.C. Zienkiewicz, R.H. Gallagher, and C. Taylor, eds., Huntsville, AL, 1974, pp. 121–132.

[33] M. Horritt, *Development and testing of a simple 2D finite volume model of sub-critical shallow water flow*, Int. J. Numer. Methods Fluids, 44 (2004), pp. 1231–1255.

[34] B.L. Hua and F. Thomasset, *A noise-free finite-element scheme for the two-layer shallow water equations*, Tellus, 36 (1984), pp. 157–165.

[35] M. Iskandarani, D.B. Haidvogel, and J. Boyd, *A staggered spectral finite-element model for the shallow-water equations*, Int. J. Numer. Methods Fluids, 20 (1995), pp. 393–414.

[36] I.P. Kinnmark and W.G. Gray, *A two-dimensional analysis of the wave equation model for finite element tidal computations*, Int. J. Numer. Methods Engrg., 20 (1984), pp. 369–383.

[37] P.H. LeBlond and L.A. Mysak, *Waves in the Ocean*, Elsevier, Amsterdam, 1978.

[38] J.K. Lewis and A.D. Kirwan, *Genesis of a Gulf of Mexico ring as determined from kinematic analyses*, J. Geophys. Res., 92 (1987), pp. 11727–11740.

[39] D.R. Lynch and W.G. Gray, *A wave-equation model for finite-element tidal computations*, Comput. Fluids, 7 (1979), pp. 207–228.

[40] D.R. Lynch, J.T.C. Ip, C.E. Naimie, and F.E. Werner, *Comprehensive coastal circulation model with application to the Gulf of Maine*, Continent. Shelf Res., (1996), pp. 875–906.

[41] F. Mesinger and A. Arakawa, *Numerical methods used in atmospheric models*, Garp Publications, 1976.

[42] R. Mullen and T. Belytschko, *Dispersion analysis of finite element semidiscretizations of the 2-dimensional wave equation*, Int. J. Numer. Methods Engrg., 18 (1982), pp. 11–29.

[43] J. Pedlosky, *Geophysical Fluid Dynamics*, Springer-Verlag, 1986.

[44] G.W. Platzman, *Some response characteristics of finite element tidal models*, J. Comput. Phys., 40 (1981), pp. 36–63.

[45] C. Le Provost and P. Vincent, *Finite element for modeling ocean tides*, in Tidal Hydrodynamics, B. Parker, ed., New York, 1991, John Wiley and Sons, pp. 41–60.

[46] S. Rahmstorf, *Thermohaline ocean circulation*, in Encyclopedia of Quaternary Sciences, S.A. Elias, ed., Elsevier, Amsterdam, 2006.

[47] D.A. Randall, *Geostrophic adjustment and the finite-difference shallow-water equations*, Monthly Weather Review, 122 (1994), pp. 1371–1377.

[48] P.A. Raviart and J.M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Mathematical Aspects of the Finite Element Methods, Lecture Notes in Math. 606, I. Galligani and E. Magenes, eds., Berlin, 1977, Springer-Verlag, pp. 292–315.

[49] ——, *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, Paris, 1983.

[50] V. Rostand, D.Y. Le Roux, and G.F. Carey, *Kernel analysis of the discretized finite difference and finite element shallow-water models*, SIAM J. on Sci. Comput., (submitted).

[51] D.Y. Le Roux, *A new triangular finite-element with optimum constraint ratio for compressible fluids*, SIAM J. Sci. Comput., 23 (2001), pp. 66–80.

[52] ——, *Dispersion relation analysis of the* $\mathrm{P}_1^{\mathrm{NC}} - \mathrm{P}_1$ *finite element pair in shallow-water ocean models.*, SIAM J. Sci. Comput., 27 (2005), pp. 394–414.

[53] D.Y. Le Roux and G.F. Carey, *Stability/dispersion analysis of the discontinuous Galerkin linearized shallow-water system*, Int. J. Numer. Methods Fluids, 48 (2005), pp. 325–347.

[54] D.Y. Le Roux, C.A. Lin, and A. Staniforth, *A semi-implicit semi-Lagrangian finite-element shallow-water ocean model*, Monthly Weather Review, 128 (2000), pp. 1384–1401.

[55] D.Y. Le Roux and B. Pouliot, *Analysis of numerically-induced oscillations in 2D finite-element shallow-water models, Part II: Free planetary waves*, SIAM J. Sci. Comput., (submitted).

[56] D.Y. Le Roux, V. Rostand, and B. Pouliot, *Analysis of numerically-induced oscillations in 2D finite-element shallow-water models, Part I: Inertia-gravity waves*, SIAM J. Sci. Comput., 29 (2007), pp. 331–360.

[57] D.Y. Le Roux, A. Sène, V. Rostand, and E. Hanert, *On some spurious mode issues in shallow-water models using a linear algebra approach*, Ocean Modelling, 10 (2005), pp. 83–94.

[58] D.Y. Le Roux, A. Staniforth, and C.A. Lin, *Finite element for shallow-water equation ocean models*, Monthly Weather Review, 126 (1998), pp. 1931–1951.

[59] D. Schwanenberg and J. Köngeter, *A discontinuous Galerkin method for the shallow-water equations with source terms*, in Discontinuous Galerkin Methods-Theory, Computation, and Applications, Lect. Notes Comput. Sci. Engrg. 11, B. Cockburn, E. Karniadakis, and C.W. Shu, eds., Berlin, 2000, Springer-Verlag, pp. 419–424.

[60] C.M. Szpilka and R.L. Kolar, *Numerical analogs to Fourier and dispersion analysis: Development, verification, and application to the shallow water equations*, Adv. in Water Res., 26 (2003), pp. 649–662.

[61] M. Taylor, J. Tribbia, and M. Iskandarani, *The spectral element method for the shallow water-equations on the sphere*, J. Comput. Phys., 130 (1997), pp. 92–108.

[62] F. Thomasset, *Implementation of Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, Berlin, 1981.

[63] C.B. Vreugdenhill, *Numerical Methods for Shallow-Water Flow*, Water Science and Technology Library, Kluwer Academic Publisher, 1994.

[64] R.A. Walters, *A three-dimensional finite-element model for coastal and estuarine circulation*, Continent. Shelf Res., 12 (1992), pp. 83–102.

[65] R.A. Walters and G.F. Carey, *Analysis of spurious oscillation modes for the shallow water and Navier-Stokes equations*, Computer and Fluids, 11 (1983), pp. 51–68.

[66] ——, *Numerical noise in ocean and estuarine models*, Advances in Water Resources, 7 (1984), pp. 15–20.

[67] R.A. Walters and V. Casulli, *A robust, finite-element model for hydrostatic surface water flows*, Commun. Numer. Meth. Engrg., 14 (1998), pp. 931–940.

[68] R.A. Walters and R.T. Cheng, *Accuracy of an estuarine hydrodynamic model using smooth elements*, Water Resources Res., 16 (1980), pp. 187–195.

[69] L. White, E. Deleersnijder, and V. Legat, *A three-dimensional unstructured mesh finite element marine model, with application to the flow around a shallow-water island*, J. Geophys. Res., (submitted).

[70] R.T. Williams and O.C. Zienkiewicz, *Improved finite-element forms for the shallow-water wave equations*, Int. J. Numer. Methods Fluids, 1 (1981), pp. 81–97.