THE UNIVERSITY OF READING

DEPARTMENT OF MATHEMATICS

The examination of balanced and unbalanced flow using Potential Vorticity in Atmospheric Modelling.

by M.A.Wlasak

Thesis submitted for the degree of

Doctor of Philosophy

January 2002

Abstract

Variational data assimilation (VAR) involves a minimisation of a cost functional with respect to a set of variables known as control variables. Within numerical weather prediction (NWP) VAR brings together observations and information from numerical models representing the atmosphere in a consistent way for a forecast to be made. It is considered desirable to define a set of control variables which separate the balanced and unbalanced parts of the flow. The current set of control variables used at the UK Met. Office represents the balanced control variable in terms of a streamfunction increment. Although this method is a good approximation to balance in high Burger regimes, it is not dependent on the Burger number and inadequately represents the balanced flow in low Burger regimes where the height increment is the balanced variable. We consider a new method in which the linearised potential vorticity increment represents the balanced part of the flow. The new set of control variables are dependent on the Burger number and should address the weaknesses in the current method. We compare the two methods in the context of the shallow water equations on a hemisphere.

We introduce a general framework to discuss various choices of control variables, looking at this from a dynamical perspective. In particular we consider various formulations in which to implement a potential vorticity control variable and propose a means with which to compare this new method with the current approach by looking at the divergence tendencies of the unbalanced variables.

Both the current Met Office method and the potential-vorticity-based method are implemented and tested numerically. The current method produces similar results to the potential vorticity method within high Burger regimes. This is due to the linearised potential vorticity increment approximating the vorticity in such regimes. Unlike the current method, however, the potential vorticity method is dependent on the Burger number and in low Burger regimes includes a substantial contribution from the height increment. The experiments suggest that the potential-vorticity-based method may be able to capture the balanced part of the flow better in low Burger regimes where the height increment is the balanced variable.

Acknowledgements

I wish to thank the many people which have made this thesis possible. My supervisors, Prof. Nancy Nichols, Dr. Ian Roulstone and Dr. Amos Lawless through their great wisdom, patience and expertise have made what was a challenging task enjoyable. I thank Prof. Mike Cullen for his stimulating advice and suggestions. I am glad that I chose to do this PHD in the warm friendly maths department.

I am indebted to my family for their constant support and encouragement. I am also thankful to all my friends who had to listen to my ramblings on the subject. Ann, Ray, Jo, Justin - I owe you big time!

Finally, I acknowledge the financial support for my studies from Engineering and Physical Sciences Research Council and from the U.K. Met. Office.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Mankind has attempted to predict the weather for millennia using subjective means. It is only in the last 60 years that independent, objective predictions have been possible. This is in part due to the increasing power of computers which allow numerical calculations that would take years to be evaluated by hand to be done within hours or even minutes. It is also a consequence of great strides in meteorology and in particular due to the increasing sophistication of numerical models representing of the atmosphere, the improved understanding of atmospheric and ocean processes and the development of data assimilation.

Data assimilation involves the bringing together of observations and information from numerical models in a consistent way. It is used in numerical weather prediction as a means of providing both a representation of the weather and also initial conditions for atmospheric models to give accurate forecasts [31].

There are various difficulties in the application of data assimilation to provide forecasts. There are both random and systematic errors with real observations as well as within the numerical models of the atmosphere. The atmosphere, due to its nonlinear behaviour, is increasingly less predictable with time. Within 6 hrs, the atmosphere can be accurately approximated by a a linear system. For forecasts of 1-2 days, non-linear interactions are evident and bifurcations in the possible solutions are much more likely. When we run forecasts of 1-2 weeks, the errors propagated in time are chaotic in nature and we are at the limit of predictability [5]. All these difficulties indicate the need for good data assimilation techniques to be determined to solve the problem of obtaining good initial conditions.

The general set of equations which describe the evolution of the atmosphere tends to be in terms of pressure, density, wind, temperature and humidity. Most of these variables act upon each other in a nonlinear fashion, varying in three-dimensional space and in time [26]. It is instructive to consider a simpler problem which retains many of the dynamical properties of the more general problem. We choose the nonlinear shallow water equations on a rotating sphere for this purpose. This set of equations describes the atmosphere as a thin layer of incompressible fluid, defined on a two-dimensional surface with varying height and horizontal velocities. It is the equation set which forms the kernel of many atmospheric models. In Chapter 2 we derive the shallow water equations from Newton's Second Law. We do this to show the limitations inherent with the dynamical system and its revelence to the more general three-dimensional situation.

The dynamics of the atmosphere is caused by the heating provided by the Sun and modified by the Earth's rotation. The atmosphere is continually adjusting itself towards an equilibrium state. This is never achieved however due to the effect of the rotating Earth which ensures that approximately 2/3 of the total energy remains in the form of potential energy [22]. A consequence of there being so much potential energy present is that the possible motions which the atmosphere can exhibit are constrained. Large scale features are forced to move slowly with the quantities involved being in a sense '*balanced*' with respect to each other. As the atmosphere adjusts towards equilibrium, fast waves with short wavelengths are typically generated. There is little interaction between the fast high frequency waves and the large slow waves. The fast waves have relatively little energy and are gradually eliminated through frictional processes. They have little effect on the general flow. For this reason numerical weather prediction has concerned itself with the modelling of large scale features as seen on weather maps.

A problem with the data assimilation of the atmosphere is that there are too few observations to apply in order to provide initial conditions at the required resolution for an operational forecast. This underdeterminancy is exacerbated by the uneven spread of observations with very few observations over the oceans. It is eliminated through using information prior to the moment in which data assimilation is performed. This additional information requires the evaluation of the balanced and unbalanced parts of the flow.

The present variables in which data assimilation is performed assumes that the full rotational part of the wind is balanced from which the remaining balanced quantities are derived [41], [31]. The actual pressure fields have no bearing on the balance. This is a good approximation to balance when variations in pressure contours are relatively small compared to full field. However, the approximation is poorer when the relative pressure variations are larger.

In this thesis, potential vorticity is considered as the variable used in data assimilation to represent the balanced part of the flow. Potential vorticity is often used as a quantity summarising dynamical information that is present within a flow [28] such as frontogenesis, cyclogenesis and key features in general circulation. It has the distinctive property that it is conserved for inviscid, isentropic flow and as such can be used to track parcels of air. It also uses both rotational wind and pressure in its evaluation and describes better the balanced part of the flow in regions where variations in pressure are important.

In Chapter 2 the shallow water equations (SWEs) are derived with corresponding linearisations needed for an analytical examination needed of the system's dynamic properties. In Chapter 3 we examine the dynamical properties of the atmosphere within the context of the shallow water equations. In particular the concept of balance is systematically described. Within this chapter we show a number of issues already known within the literature but which tend to be forgotten. We show that the divergence tendency, as defined in Section 3.7, in general does not always filter the unbalanced aspects of the flow. Additional conditions need to be enforced. We show how a simple potential vorticity inversion model takes contributions from the height and the rotational wind in a way dependent on the flow regime.

In Chapter 4 we systematically present and discuss the choice of variables in which the data assimilation is performed. These variables are called *control variables*. We provide a framework in which different sets of control variables are discussed. Such a systematic appraisal of control variables is not present within the current literature. We take the method used presently by the UK Met. Office as an example and discuss the strength and weaknesses of the current change of control variables. The properties of an 'idealised' set of control variables are considered from a dynamical perspective using the dynamical background presented in the previous chapter.

We present various formulations of control variables based on potential vorticity, discussing their respective advantages and disadvantages and how they vary in different regimes. We finally develop a means of approximating the balanced parts of the unbalanced variables.

Chapter 5 gives the numerical techniques used to calculate the present choice of control variables used at the UK Met. Office. These techniques are used in experiments in Chapter 7. The chapter also presents a Fourier-based technique which is used in Chapter 6 to develop a means of obtaining the potential vorticity-based set of control variables. To the author's knowledge, the coupled system of equations has not been previously solved in this way. In both chapters validatory tests are performed. In Chapter 7, we present various experiments to compare potential vorticity based set of control variables with current method, which illustrate the theory given in previous chapters. Finally in Chapter 8 we summarise the findings and detail possible avenues for future work.

Chapter 2

Shallow Water Equations

2.1 Introduction

We first derive the shallow water equations (SWEs) on a rotating sphere from Newton's Second Law. This is necessary so as to give an accurate representation of the approximations made to obtain the equations, so that the results may be compared to other studies which use different approximations to the equations of motion on a rotating sphere. The discussion draws mainly on the treatments given by Pedlosky [43] and Randall [45] and includes the derivation of the incompressible three-dimensional Euler equations as a stepping stone to obtaining the shallow water equations.

We then present properties of the shallow water equations on a rotating sphere, linearised about two linearisation states: a resting state and a general time-invariant state. The former state is useful due to its idyllic simplicity while the latter state gives a template to derive other SWEs about more restrictive linearisations states necessary for the development of Chapters 3 and 6.

In addition, versions of the SWEs which approximate the spherical geometry are introduced. These versions, such as the β-plane approximation, are useful as a means of making analytical studies of the SWEs more tractable.

2.2 Derivation of Shallow Water Equations on a Rotating Sphere from Newton's Second Law

Newton's Second Law of Motion states that the mass of an object multiplied by its absolute acceleration is equal to the total actual force acting on the object in a non-rotating co-ordinate system. When written for a fluid continuum, it is expressed in terms of density ρ, a three-dimensional wind $\mathbf{u}$, pressure p, the body force $\rho\nabla\phi$ and non-conservative force $\mathcal{F}$. In particular, ϕ is the potential field with which conservative body forces are represented and $\mathcal{F}$ is the frictional force. Newton's Second Law takes the form

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla p + \rho\nabla\phi + \mathcal{F}\left(\mathbf{u}\right), \tag{2.1}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla \tag{2.2}$$

is called the total material derivative with respect to time.

As stated, this law applies only on a stationary frame of reference. We wish to consider the momentum equation (2.1) for an observer in a uniformly rotating co-ordinate frame. We let the subscript I, represent a the non-rotating co-ordinate frame of reference and R are rotating one. We also denote the velocity under a rotating reference frame, $\mathbf{u}_R$, as the *relative velocity* and the velocity under a non-rotating reference frame, $\mathbf{u}_I$, as the *absolute velocity*. In this situation,

$$\begin{aligned} \left(\frac{D\mathbf{r}}{Dt}\right)_I &= \left(\frac{D\mathbf{r}}{Dt}\right)_R + \mathbf{\Omega} \times \mathbf{r}, & (2.3)\\ \mathbf{u}_I &= \mathbf{u}_R + \mathbf{\Omega} \times \mathbf{r}, & (2.4)\\ \left(\frac{D\mathbf{u}}{Dt}\right)_I &= \left(\frac{D\mathbf{u}}{Dt}\right)_R + \frac{D\mathbf{\Omega}}{Dt} \times \mathbf{r} + \mathbf{\Omega} \times \mathbf{u}_R, & (2.5) \end{aligned}$$

where $\mathbf{r}$ is a position vector of an arbitrary fluid element and $\mathbf{\Omega}$ is the angular velocity, with all vector quantities defined on a three-dimensional space.

The first step is to rewrite the acceleration in the non-rotating frame purely in terms of quantities in a rotating frame. When the material derivative operator (2.2) is applied to (2.4), we obtain

$$\left(\frac{D\mathbf{u_I}}{Dt}\right)_I = \left(\frac{D\mathbf{u}_I}{Dt}\right)_R + \mathbf{\Omega} \times \mathbf{u}_I.$$

The right hand side is further manipulated using (2.4), (2.5), with

$$\begin{aligned} \left(\frac{D\mathbf{u_I}}{Dt}\right)_I &= \left(\frac{D\mathbf{u}_R}{Dt}\right)_R + \frac{D\mathbf{\Omega}}{Dt} \times \mathbf{r} + \mathbf{\Omega} \times \mathbf{u}_R + \mathbf{\Omega} \times (\mathbf{u}_R + \mathbf{\Omega} \times \mathbf{r}) \\ &= \left(\frac{D\mathbf{u}_R}{Dt}\right)_R + 2\mathbf{\Omega} \times \mathbf{u}_R + \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}) + \frac{D\mathbf{\Omega}}{Dt} \times \mathbf{r}. & (2.6) \end{aligned}$$

The three additional terms on the right hand side are the Coriolis acceleration $2\mathbf{\Omega} \times \mathbf{u}_R$, the centripetal acceleration $\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r})$ and acceleration due to change in angular velocity. In the present application we assume that the angular velocity is constant in time, and as such, the $\frac{D\mathbf{\Omega}}{Dt} \times \mathbf{r}$ term vanishes.

The centripetal acceleration has a direction perpendicular to the rotational axis which allows it to be rewritten in terms of a perpendicular distance vector $\mathbf{r}_\perp$ as $\mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}_\perp)$. We then use the vector identities,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\,\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\,\mathbf{C},$$

where **A**, **B**, **C** are generic vectors. Thus,

$$\begin{aligned}\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{r}) &= \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{r}_\perp)\\ &= (\boldsymbol{\Omega}\cdot\mathbf{r}_\perp)\,\boldsymbol{\Omega}-|\boldsymbol{\Omega}|^2\mathbf{r}_\perp\\ &= 0-\nabla\frac{|\boldsymbol{\Omega}|^2|\mathbf{r}_\perp|^2}{2}\\ &= -\nabla\frac{|\boldsymbol{\Omega}\times\mathbf{r}_\perp|^2}{2}, \qquad (2.7)\end{aligned}$$

since Ω and $\mathbf{r}_\perp$ are orthogonal.

We now incorporate the centripetal acceleration with the other conservative terms in (2.1) by defining the apparent gravitational potential,

$$\Theta=\phi+\frac{|\boldsymbol{\Omega}\times\mathbf{r}_\perp|^2}{2}. \qquad (2.8)$$

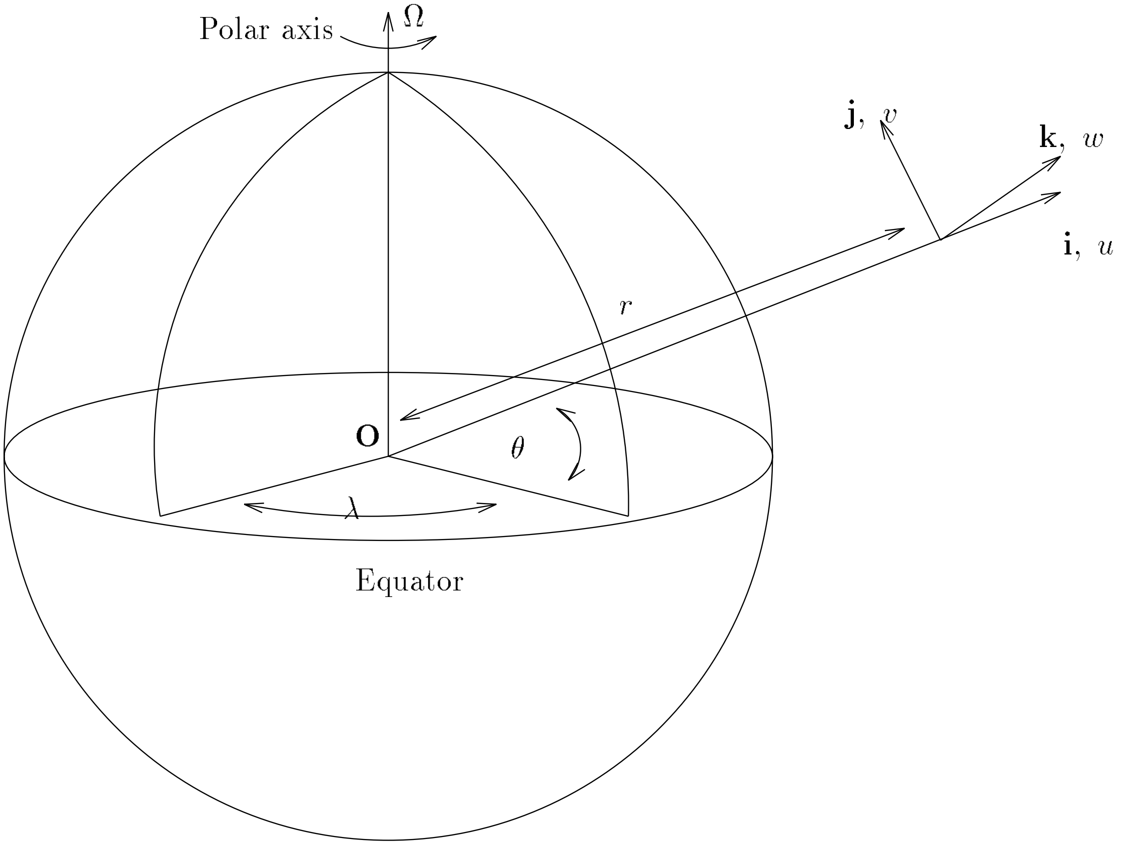
As the Coriolis acceleration $2\Omega\times\mathbf{u}_R$ cannot be further simplified, the momentum equation in a rotating co-ordinate frame is given by

$$\rho\left(\frac{D\mathbf{u}}{Dt}+2\boldsymbol{\Omega}\times\mathbf{u}\right)=-\nabla p+\rho\nabla\Theta+\mathcal{F} \qquad (2.9)$$

where all $\mathbf{u}_R$ are written as $\mathbf{u}$.

So far we have not been specific as to the three-dimensional space we are considering. We define $\nabla\Theta$ to define the apparent vertical direction $\mathbf{k}^+$, which is perpendicular to an oblately spheroidal geopotential surface. However since the centripetal acceleration is very small compared to the constant body force $\nabla\phi$, we let $\nabla\Theta=\nabla\phi=g\mathbf{k}$, where $\mathbf{k}$ is a unit vector pointing radially away from the centre of a sphere and g is the acceleration due to gravity. The neglection of the centripetal acceleration allows the oblate spheroidal surface to be approximated by a spherical surface with unit vectors $\mathbf{i}$,$\mathbf{j}$ denoting longitude and latitude directions. The spherical latitude/longitude co-ordinate system is shown in Figure 2.1, where any

Figure 2.1: A sphere rotating at a constant angular velocity of magnitude Ω. The diagram also shows the direction of the orthonormal unit vectors $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ for a spherical latitude θ, longitude λ co-ordinate system



non-polar point in the three-dimensional space is either represented by $\lambda\mathbf{i} + \theta\mathbf{j} + r\mathbf{k}$ or more simply as (λ, θ, r). The three-dimensional wind in terms of its components is given by $\mathbf{u} = (u\mathbf{i} + v\mathbf{j} + w\mathbf{k})$. Although the co-ordinate system is degenerate at the poles, it is the natural choice for problems involving a spherical geometry.

Then, ignoring the effects of friction, the momentum equation becomes

$$\rho\left(\frac{D\mathbf{u}}{Dt} + 2\mathbf{\Omega} \times \mathbf{u}\right) = -\nabla p - \rho g\mathbf{k}. \tag{2.10}$$

For the following derivation of the SWEs we need the mass conservation equation. This equation is written as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{u}) = 0 \quad \text{or} \quad \frac{D\rho}{Dt} + \rho\nabla \cdot \mathbf{u} = 0. \tag{2.11}$$

We now simplify (2.10), (2.11) to get the shallow water equations on a rotating sphere. Let us write the pressure and density as

$$p = \overline{p}(r) + p^*(\lambda, \theta, r, t) \quad \text{and} \quad \rho = \overline{\rho}(r) + \rho^*(\lambda, \theta, r, t), \tag{2.12}$$

where $\overline{p}(r)$ and $\overline{\rho}(r)$ are linearisation states which satisfy the hydrostatic equation,

$$\frac{d\overline{p}}{dr} = -g\overline{\rho}, \tag{2.13}$$

with p^*, ρ^* being departures from these linearisation states.

Substituting both (2.12) and (2.13) into the momentum equation (2.10) we obtain

$$\rho\left(\frac{D\mathbf{u}}{Dt} + 2\mathbf{\Omega} \times \mathbf{u}\right) = -\nabla p^* - \rho^* g\mathbf{k}. \tag{2.14}$$

We now assume that the fluid is incompressible with $\rho = \overline{\rho}$ and $\rho^* = 0$. This approximation discards the conservative body force in the momentum equation (2.14), giving

$$\rho\left(\frac{D\mathbf{u}}{Dt} + 2\mathbf{\Omega} \times \mathbf{u}\right) = -\nabla p^* \tag{2.15}$$

while reducing the mass conservation equation (2.11) to

$$\nabla \cdot \mathbf{u} = 0. \tag{2.16}$$

Equations (2.15) and (2.16) define the three-dimensional incompressible Euler equations on a rotating sphere.

To obtain the shallow water equations on a sphere, we need to reduce (2.15) and (2.16) onto a two-dimensional spherical surface. We choose to describe the flow by a shallow incompressible layer of fluid of height h and horizontal velocity field $\mathbf{v}$ that are functions of latitude/longitude and time.

Let $h(\lambda, \theta, t)$ be the height at the free surface and $h_s(\lambda, \theta)$ represent the bottom topography. The depth of the fluid is then given by $h^+ = h - h_s$. The height h is defined so that its material derivative is equal to the vertical velocity w, such that

$$\frac{Dh}{Dt} = w\left(\lambda, \theta, h, t\right) \tag{2.17}$$

and

$$\frac{Dh_s}{Dt} = \left(\mathbf{v} \cdot \nabla\right) h_s = w_s \tag{2.18}$$

since the topographic height is independent of time. The material derivative terms in equations (2.17), (2.18) involve advective terms $\mathbf{v} \cdot \nabla$ on the spherical surface. Vertical advection is set to zero as no vertical shear is allowed. The material derivative now defined as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \tag{2.19}$$

We integrate the mass conservation equation (2.16) with respect to r from $r = h_s$ to $r = h$ to give

$$w(\lambda, \theta, h, t) - w(\lambda, \theta, h_s, t) = -\left(h - h_s\right) \nabla \cdot \mathbf{v}, \tag{2.20}$$

with $\nabla \cdot \mathbf{v}$ representing the horizontal divergence. Equations (2.17), (2.18) are substituted into (2.20) to obtain

$$\frac{\partial h}{\partial t} + \mathbf{v} \cdot \nabla h - \mathbf{v} \cdot \nabla h_s = -\left(h - h_s\right) \nabla \cdot \mathbf{v}. \tag{2.21}$$

We rearrange this equation to give

$$\frac{\partial \left(h - h_s\right)}{\partial t} + \nabla \cdot \left[\mathbf{v}\left(h - h_s\right)\right] = 0 \quad \text{or} \quad \frac{\partial h^+}{\partial t} + \nabla \cdot \left[\mathbf{v} h^+\right] = 0, \tag{2.22}$$

which represents the mass conservation equation for the shallow water equations.

The momentum equation for the shallow water equations is derived by looking at the Coriolis, pressure gradient and the material derivative terms separately. Since we are considering a material derivative term on a surface, the material derivative term is simply $\frac{D\mathbf{v}}{Dt}$. As in equations (2.17), (2.18), the assumption of no vertical shear requires that there is no vertical advection to the horizontal momentum.

The pressure gradient term on the horizontal surface is obtained by integrating the hydrostatic equation (2.13) from some some arbitrary depth r within the fluid to the free surface h, giving

$$p(\lambda, \theta, h, t) - p(\lambda, \theta, r, t) = -g\rho(h - r), \tag{2.23}$$

with boundary conditions

$$p(\lambda, \theta, h, t) = p_f \qquad \text{and} \qquad p(\lambda, \theta, r, t) = g\rho\,(h - r) + p_f, \tag{2.24}$$

where p_f is the pressure at the free surface.

In equation (2.12) p is given in terms of a linearisation state $\overline{p}$ and perturbation p^*. Integration of the hydrostatic relation splits the pressure similarly, with

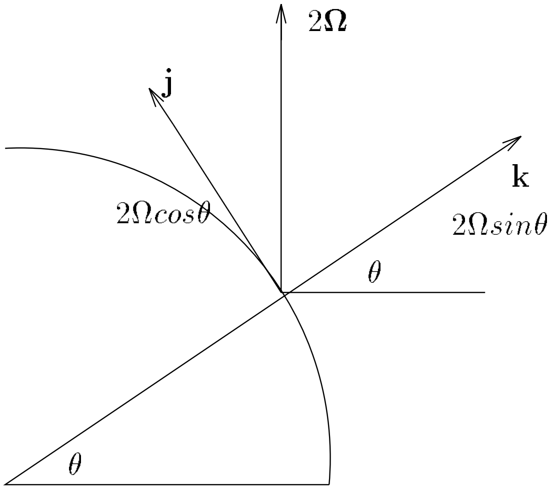
$$\overline{p} = -g\rho r \text{ and } p^* = g\rho h + p_f, \tag{2.25}$$

giving

$$\nabla p^* = g\rho\nabla h. \tag{2.26}$$

The $\mathbf{k}$ independent horizontal velocity reduces the Coriolis term in the momentum equation (2.14) from $2\boldsymbol{\Omega} \times \mathbf{u}$ to $f\mathbf{k} \times \mathbf{v}$. Figure 2.2 shows that the vector $\boldsymbol{\Omega}$ can be decomposed into $\mathbf{j}$ and $\mathbf{k}$ components such that $2\boldsymbol{\Omega} = 2\Omega\cos\theta\mathbf{j} + 2\Omega\sin\theta\mathbf{k}$, where θ is the latitude and Ω is the magnitude of the angular velocity. Hence we have

Figure 2.2: A diagram representing the decomposition of vector $\mathbf{\Omega}$ into $\mathbf{j}$ and $\mathbf{k}$ components.



$$
\begin{aligned}
2\mathbf{\Omega} \times \mathbf{u} &= (2\Omega \cos\theta \mathbf{j} + 2\Omega \sin\theta \mathbf{k}) \times \mathbf{u} \\
&= (2\Omega w \cos\theta - 2\Omega v \sin\theta)\, \mathbf{i} + (2\Omega u \sin\theta)\, \mathbf{j} + 2\Omega u \cos\theta \mathbf{k} \\
&= f\mathbf{k} \times \mathbf{v} + \textit{additional terms.} \qquad (2.27)
\end{aligned}
$$

The additional terms are ignored; the $\mathbf{k}$ term is discarded because it violates the hydrostatic relation and does not lie on the spherical surface. The $2\Omega w \cos\theta \mathbf{i}$ is removed for consistent energy conservation to occur. This approximation is called the *traditional approximation* [17].

We now use (2.26) write the horizontal momentum equation as

$$\frac{D\mathbf{v}}{Dt} + f\mathbf{k} \times \mathbf{v} = -g\nabla h. \qquad (2.28)$$

The advective part of the material derivative is not scalar invariant. The vector transformation $(\mathbf{v} \cdot \nabla)\,\mathbf{v} = \nabla\left[(\mathbf{v} \cdot \mathbf{v})/2\right] + (\nabla \times \mathbf{v}) \times \mathbf{v}$ is used to rewrite the momentum equation as

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla\left[(\mathbf{v} \cdot \mathbf{v})/2\right] + (\nabla \times \mathbf{v}) \times \mathbf{v} + f\mathbf{k} \times \mathbf{v} \;=\; -g\nabla h$$

$$
\begin{aligned}
\Rightarrow \frac{\partial \mathbf{v}}{\partial t} + \nabla\left[(\mathbf{v}\cdot\mathbf{v})/2\right] + ((\nabla\times\mathbf{v})\cdot\mathbf{k})\,\mathbf{k}\times\mathbf{v} + f\mathbf{k}\times\mathbf{v} &= -g\nabla h \\
\Rightarrow \frac{\partial \mathbf{v}}{\partial t} + \nabla\left[(\mathbf{v}\cdot\mathbf{v})/2\right] + (\zeta + f)\,\mathbf{k}\times\mathbf{v} &= -g\nabla h, \qquad (2.29)
\end{aligned}
$$

where ζ is the relative vorticity defined as the $\mathbf{k}\cdot\nabla\times\mathbf{v}$.

Equations (2.29), (2.22) are now scalar invariant. The standard vector operators on the sphere are obtained by introducing the scale factors $s_\theta = a$, $s_\lambda = a\cos\theta$, $s_r = 1$ into the general curvi-linear form of the vector operations on the 2D surface:

$$
\begin{aligned}
\nabla &= \frac{1}{s_\lambda}\frac{\partial}{\partial\lambda}\mathbf{i} + \frac{1}{s_\theta}\frac{\partial}{\partial\theta}\mathbf{j} \\
\nabla\cdot\mathbf{v} &= \frac{1}{s_\lambda s_\theta s_r}\left(\frac{\partial}{\partial\lambda}(s_\theta s_r u) + \frac{\partial}{\partial\theta}(s_\lambda s_r v)\right) \\
\zeta &= \frac{1}{s_\lambda s_\theta s_r}\left(\frac{\partial}{\partial\lambda}(s_\theta s_r v) - \frac{\partial}{\partial\theta}(s_\lambda s_r u)\right), \qquad (2.30)
\end{aligned}
$$

to give

$$
\begin{aligned}
\nabla &= \frac{1}{a\cos\theta}\frac{\partial}{\partial\lambda}\mathbf{i} + \frac{1}{a}\frac{\partial}{\partial\theta}\mathbf{j} \qquad (2.31) \\
\nabla\cdot\mathbf{v} &= \frac{1}{a\cos\theta}\left(\frac{\partial u}{\partial\lambda} + \frac{\partial}{\partial\theta}(\cos\theta v)\right) \qquad (2.32) \\
\zeta &= \frac{1}{a\cos\theta}\left(\frac{\partial v}{\partial\lambda} - \frac{\partial}{\partial\theta}(\cos\theta u)\right). \qquad (2.33)
\end{aligned}
$$

Equations (2.31), (2.32), (2.33) are substituted into (2.29), (2.22), from which we obtain

$$
\begin{aligned}
\frac{\partial u}{\partial t} + \frac{u}{a\cos\theta}\frac{\partial u}{\partial\lambda} + \frac{v}{a\cos\theta}\frac{\partial v}{\partial\lambda} \\
- \frac{v}{a\cos\theta}\left(\frac{\partial v}{\partial\lambda} - \frac{\partial}{\partial\theta}(\cos\theta u)\right) - fv &= -\frac{g}{a\cos\theta}\frac{\partial h}{\partial\lambda}, \qquad (2.34) \\
\frac{\partial v}{\partial t} + \frac{u}{a}\frac{\partial u}{\partial\theta} + \frac{v}{a}\frac{\partial v}{\partial\theta} + \frac{u}{a\cos\theta}\left(\frac{\partial v}{\partial\lambda} - \frac{\partial}{\partial\theta}(\cos\theta u)\right) + fu &= -\frac{g}{a}\frac{\partial h}{\partial\theta}, \qquad (2.35) \\
\frac{\partial h^+}{\partial t} + \frac{1}{a\cos\theta}\left(\frac{\partial u h^+}{\partial\lambda} + \frac{\partial}{\partial\theta}\left(\cos\theta v h^+\right)\right) &= 0 \qquad (2.36)
\end{aligned}
$$

Equation (2.36) is the expanded final form of the mass conservation equation for a shallow water model on a sphere. The momentum equations (2.34), (2.35) are rearranged, to give

$$\begin{aligned}
\frac{\partial u}{\partial t}+\frac{u}{a\cos\theta}\frac{\partial u}{\partial\lambda}+\frac{v}{a}\frac{\partial u}{\partial\theta}-(f+\frac{u}{a}tan\theta)v+\frac{g}{a\cos\theta}\frac{\partial h}{\partial\lambda} &= 0,\\
\frac{\partial v}{\partial t}+\frac{u}{a\cos\theta}\frac{\partial v}{\partial\lambda}+\frac{v}{a}\frac{\partial v}{\partial\theta}+(f+\frac{u}{a}tan\theta)u+\frac{g}{a}\frac{\partial h}{\partial\theta} &= 0,
\end{aligned} \tag{2.37}$$

where h is the height and (λ,θ) are the spherical longitude/latitude co-ordinates. The relative velocities on the sphere, (u,v) are in $\lambda\mathbf{i}$ and $\theta\mathbf{j}$ directions with $\boldsymbol{u}=(u,v)^T$. The acceleration due to gravity is denoted as g with the Coriolis parameter described as $f=2\Omega\sin\theta$, where Ω is the angular velocity of the earth.

2.3 Shallow-Water Equations Linearised About a Resting State on a Rotating Sphere

In this section we present some of the properties of a shallow-water equation model on a rotating sphere linearised about a resting state. We assume no bottom topography so that $h^+=h$ and choose perturbations u^*, v^*, h^* such that

$$\begin{aligned}
u(\lambda,\theta,t) &= u^*(\lambda,\theta,t)\\
v(\lambda,\theta,t) &= v^*(\lambda,\theta,t)\\
h(\lambda,\theta,t) &= H+h^*(\lambda,\theta,t)
\end{aligned} \tag{2.38}$$

where H is a constant value across the sphere.

Substituting (2.38) into (2.37), (2.36) and ignoring quadratic perturbation terms, gives:

$$\frac{\partial u^*}{\partial t}-fv^*+\frac{g}{a\cos\theta}\frac{\partial h^*}{\partial\lambda} = 0, \tag{2.39}$$

$$\frac{\partial v^*}{\partial t} + fu^* + \frac{g}{a}\frac{\partial h^*}{\partial \theta} = 0, \tag{2.40}$$

$$\frac{\partial h^*}{\partial t} + H\left(\frac{1}{a\cos\theta}\frac{\partial u^*}{\partial \lambda} + \frac{1}{a\cos\theta}\frac{\partial}{\partial \theta}\left(\cos\theta v^*\right)\right) = 0, \tag{2.41}$$

which are the shallow water equations linearised about a resting state and no bottom topography. The linearisation state for the height,H, has to be a constant, for the momentum equations to be consistent with the mass conservation equation.

We may rewrite the momentum equations in terms of vorticity and divergence. By applying the **k** component of the curl to the momentum equations we obtain the vorticity equation. It is given by

$$\frac{\partial \zeta^*}{\partial t} + f\zeta^* + \frac{2\Omega}{a}\cos\theta v^* = 0, \tag{2.42}$$

where the relative vorticity ζ^* is defined to be

$$\zeta^* = \frac{1}{a\cos\theta}\frac{\partial v^*}{\partial \lambda} - \frac{1}{a\cos\theta}\frac{\partial}{\partial \theta}\left(\cos\theta u^*\right). \tag{2.43}$$

In vector notation, the vorticity equation is written as

$$\frac{\partial \zeta^*}{\partial t} + \nabla\cdot\left(f\mathbf{v}^*\right) = 0, \tag{2.44}$$

where the divergence operator $\nabla\cdot$, and all subsequent spatial differential operators, lie on a spherical surface. The vector $\mathbf{v}^*$ denotes the wind on this surface and is equal to (u^*, v^*).

Application of the divergence operator to the momentum equations gives

$$\frac{\partial \delta^*}{\partial t} - f\zeta^* + 2\Omega\cos\theta u^* = -g\nabla^2 h^*, \tag{2.45}$$

where the divergence of the wind $\delta^* = \nabla\cdot\mathbf{v}^*$ is given by (2.32). In vector notation the divergence equation is given by

$$\frac{\partial \delta^*}{\partial t} - \mathbf{k}\cdot\nabla\times f\mathbf{v}^* = -g\nabla^2 h^*. \tag{2.46}$$

The divergence equation can be written in terms of two scalar fields called the *streamfunction* ψ^* and *velocity potential* χ^*, which are defined through the relations:

$$\begin{aligned} \nabla^2\psi^* &= \mathbf{k}\cdot(\nabla\times\mathbf{v}^*), & (2.47) \\ \nabla^2\chi^* &= \nabla\cdot\mathbf{v}^*, & (2.48) \\ \mathbf{v}^* &= \mathbf{k}\times\nabla\psi^* + \nabla\chi^*, & (2.49) \end{aligned}$$

and has the form

$$\frac{\partial\delta^*}{\partial t} + \nabla f\cdot(\mathbf{k}\times\nabla\chi^*) - \nabla\cdot(f\nabla\psi^*) = -g\nabla^2 h^*, \tag{2.50}$$

where

$$\nabla f\cdot(\mathbf{k}\times\nabla\chi^*) = \frac{1}{a^2\cos\theta}\frac{\partial f}{\partial\theta}\frac{\partial\chi^*}{\partial\lambda}. \tag{2.51}$$

2.3.1 Potential Vorticity

We now consider the evolution of the potential vorticity q, which we define to be either equivalent to the Rossby's quantity [48],[47],

$$q = \frac{f+\zeta}{h+H}, \tag{2.52}$$

or the linearisation of (2.52) about a linearisation state $\overline{q}(\theta) = \frac{f}{H}$,

$$\frac{q^*}{\overline{q}} = \frac{\zeta^*}{f} - \frac{h^*}{H} \tag{2.53}$$

where q^*, h^*, ζ^* are perturbations away from the respective linearisation states $\overline{q}$, H and f.

The evolution of the linearised potential vorticity (2.53) is easily deduced by subtracting the linearised vorticity equation (2.44) multiplied by $1/H$, from f/H^2

times the linearised mass conservation equation (2.41), giving

$$\frac{Dq^*}{Dt} = \frac{1}{H}\left(\frac{\partial}{\partial t}\left(\zeta^* - \frac{f}{H}h^*\right) + \nabla\cdot(f\mathbf{v}^*) - f\nabla\cdot\mathbf{v}^*\right) = 0 \tag{2.54}$$

and

$$\frac{\partial q^*}{\partial t} + (\mathbf{v}^*\cdot\nabla)\,\overline{q} = \frac{1}{H}\frac{\partial}{\partial t}\left(\zeta^* - \frac{f}{H}h^*\right) + \frac{\nabla f\cdot\mathbf{v}^*}{H} = 0. \tag{2.55}$$

It is apparent that the linearisation of the potential vorticity is conserved, with the advective term equal to the product of the gradient of the Coriolis term and the wind, divided by the constant height H.

2.4 Linearised SWE on a Rotating Sphere about a Time-Varying Linearisation State

Let us assume that the wind and the height, denoted by $\mathbf{v}$ and h, are each split into two components: a linearisation state, denoted by an overbar, which is dependent on latitude and longitude, and a perturbation component, denoted by a star. This can be written as:

$$\begin{aligned} \mathbf{v}(\lambda,\theta,t) &= \overline{\mathbf{v}}(\lambda,\theta,t) + \mathbf{v}^*(\lambda,\theta,t), \\ h(\lambda,\theta,t) &= \overline{h}(\lambda,\theta,t) + h^*(\lambda,\theta,t). \end{aligned} \tag{2.56}$$

Since the linearisation states vary with time, they satisfy the nonlinear SWEs:

$$\begin{aligned} \frac{\partial\overline{\mathbf{v}}}{\partial t} + \nabla\left[(\overline{\mathbf{v}}\cdot\overline{\mathbf{v}})/2\right] + ((\mathbf{k}\cdot(\nabla\times\overline{\mathbf{v}})) + f)\,\mathbf{k}\times\overline{\mathbf{v}} &= -g\nabla\overline{h} \\ \frac{\partial\overline{h}}{\partial t} + \overline{\mathbf{v}}\cdot\nabla\overline{h} + \overline{h}\nabla\cdot\overline{\mathbf{v}} &= 0. \end{aligned} \tag{2.57}$$

The linearised shallow-water equations under these linearisation states are given by substituting (2.56) into the nonlinear shallow water equations (2.36), (2.37), subtracting the linearisation state relations (2.57) and ignoring the quadratic perturbed terms. We obtain

$$\frac{\partial \mathbf{v}^*}{\partial t} + (\overline{\mathbf{v}}.\nabla)\,\mathbf{v}^* + (\mathbf{v}^*.\nabla)\,\overline{\mathbf{v}} + f\mathbf{k}\times\mathbf{v}^* = -g\nabla h^* \tag{2.58}$$

$$\frac{\partial h^*}{\partial t} + \nabla\cdot\left(\overline{h}\mathbf{v}^* + h^*\overline{\mathbf{v}}\right) = 0. \tag{2.59}$$

2.4.1 Vorticity and Divergence Equations

The vorticity equation and divergence equations are just generalisations of (2.42) and (2.46). The vorticity equation is

$$\frac{\partial \zeta^*}{\partial t} + \nabla\cdot\left(\left(\overline{\zeta} + f\right)\mathbf{v}^* + \zeta^*\overline{\mathbf{v}}\right) = 0, \tag{2.60}$$

while the divergence equation is

$$\frac{\partial \delta^*}{\partial t} - \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta} + f\right)\mathbf{v}^* + \zeta^*\overline{\mathbf{v}}\right) + \nabla^2\left(gh^* + \mathbf{v}^*\cdot\overline{\mathbf{v}}\right) = 0. \tag{2.61}$$

In later chapters, we use the divergence equation in terms of perturbations in streamfunction and velocity potential. Using (2.47), (2.49), the divergence equation (2.61) is re-written as

$$\begin{aligned}\frac{\partial \delta^*}{\partial t} + \nabla\cdot\left(f + \overline{\zeta}\right)\nabla\psi^* - \mathbf{k}\cdot\nabla\times\left(\left(f + \overline{\zeta}\right)\nabla\chi^* + \nabla^2\psi^*\overline{\mathbf{v}}\right)& \\ +\nabla^2\left(gh^* + \left(\mathbf{k}\times\nabla\psi^* + \nabla\chi^*\right)\cdot\overline{\mathbf{v}}\right) &= 0.\end{aligned} \tag{2.62}$$

2.4.2 Potential Vorticity

The linearisation of Rossby's potential vorticity (2.52) is similar to (2.53). The only difference is that we are linearising about a potential vorticity which is equal to

$$\overline{q} = \frac{f + \mathbf{k}\cdot\nabla\times\overline{\mathbf{v}}}{\overline{h}}. \tag{2.63}$$

The potential vorticity increment q^* associated with this linearisation state is

$$\frac{q^*}{\overline{q}} = \frac{\zeta^*}{\overline{\zeta}+f} - \frac{h^*}{\overline{h}} + \mathcal{O}((h^*)^2, h^*\nabla\times\mathbf{v}^*, (\nabla\times\mathbf{v}^*)^2). \tag{2.64}$$

where ζ^*, $\overline{\zeta}$ are defined by (2.43), using $\mathbf{v}^*$, $\overline{\mathbf{v}}$, respectively.

The evolution of the linearised potential vorticity q^* is given by

$$\frac{Dq^*}{Dt} \;=\; \frac{\partial q^*}{\partial t} + (\mathbf{v}^*\cdot\nabla)\,\overline{q} + (\overline{\mathbf{v}}\cdot\nabla)\,q^* \;=\; 0 \tag{2.65}$$

2.5 Approximations to the Geometry of SWE about a Rotating Sphere

Throughout the above description a latitude/longitude grid is used to represent the surface area of the sphere. This representation, with the exception of the poles, is a conformal mapping of the surface from a 3D Cartesian co-ordinate system to a flat rectangular domain. The spherical geometry manifests itself in a latitudinally dependent weighting with the introduction of $\cos\theta$ terms.

There are instances where it is convenient to simplify the SWEs by simplifying the geometry. One choice is to approximate the latitude and longitude co-ordinate system by a two-dimensional Cartesian grid, which sets the distance $\mathbf{r}_\perp$, perpendicular to the line bisecting the poles, to be constant. This, in effect, enforces a change in the scale factors applied to the generalised curvilinear formulation of the vector operations (2.30), from $s_\theta = a$, $s_\lambda = a\cos\theta$, $s_r = 1$ to $s_\theta = a$, $s_\lambda = a\cos\overline{\theta}$, $h_r = 1$ where $\overline{\theta}$ is the central latitude value for the approximation. In addition, if the rescaling,

$$x \;=\; \lambda a\cos\overline{\theta},$$

$$\begin{aligned} y &= a\left(\theta-\overline{\theta}\right), \\ \frac{\partial}{\partial x} &= \frac{1}{a\cos\overline{\theta}}\frac{\partial}{\partial\lambda}, \\ \frac{\partial}{\partial y} &= \frac{1}{a}\frac{\partial}{\partial\theta}, \end{aligned} \tag{2.66}$$

is used, the standard Cartesian formulation is obtained.

However, an additional concern arises in the treatment of the effects of the rotating co-ordinate system. As shown in the derivation of the SWE on a rotating sphere, the effects of a rotating, relative co-ordinate system results in the addition of Coriolis terms into the momentum equations. The approximation for the Coriolis terms needs to be consistent with the 2D Cartesian surface. Two approximations are used: the f-plane and the β-plane assumptions; the f-plane assumes that Coriolis parameter is constant and equal to $f_0 = 2\Omega\sin\overline{\theta}$ while the β-plane approximates the term by $f_0 + \beta y$, where

$$\beta = \left.\frac{df}{dy}\right|_{\overline{\theta}} = 2\Omega\cos\overline{\theta}. \tag{2.67}$$

Both approximations to f are just truncations of a Taylor series expansion of f about $\overline{\theta}$.

A typical example of a β-plane Cartesian approximation about a mid-latitude range of between 30 deg $-$60 deg is now presented for vorticity, divergence and mass conservation equations whose height and wind fields are linearised about a constant height H and constant wind u^+, where $\overline{\mathbf{u}} = (u^+, 0)$. The equations are obtained by subsituting the linearisation states and β-plane approximation into (2.60) , (2.61) and (2.59) , giving

$$\frac{\partial\zeta^*}{\partial t} + u^+\frac{\partial\zeta^*}{\partial x} + (f_0+\beta y)\left(\frac{\partial u^*}{\partial x}+\frac{\partial v^*}{\partial y}\right) + \beta v^* = 0, \tag{2.68}$$

$$\frac{\partial \delta^*}{\partial t} + u^+ \frac{\partial \delta^*}{\partial x} - (f_0 + \beta y)\left(\frac{\partial v^*}{\partial x} - \frac{\partial u^*}{\partial y}\right) + \beta u^* + g\left(\frac{\partial^2 h^*}{\partial x^2} + \frac{\partial^2 h^*}{\partial y^2}\right) = 0, \tag{2.69}$$

$$\frac{\partial h^*}{\partial t} + u^+ \frac{\partial h^*}{\partial x} + H\left(\frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y}\right) = 0, \tag{2.70}$$

These equations in streamfunction and velocity potential formulation are

$$\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x}\right)\nabla^2 \psi^* + (f_0 + \beta y)\nabla^2 \chi^* + \beta\left(\frac{\partial \psi^*}{\partial x} + \frac{\partial \chi^*}{\partial y}\right) = 0, \tag{2.71}$$

$$\begin{aligned}\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x}\right)\nabla^2 \chi^* - (f_0 + \beta y)\nabla^2 \psi^* + \beta\left(-\frac{\partial \psi^*}{\partial y} + \frac{\partial \chi^*}{\partial x}\right)& \\ +g\left(\frac{\partial^2 h^*}{\partial x^2} + \frac{\partial^2 h^*}{\partial y^2}\right) &= 0,\end{aligned} \tag{2.72}$$

$$\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x}\right) h^* + H\nabla^2 \chi^* = 0, \tag{2.73}$$

using the relations (2.47), (2.48), (2.49) and assuming

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \tag{2.74}$$

To apply simple analysis methods to the β-plane model, it is beneficial to apply a further approximation to the Coriolis terms, so that the partial differential equations have only constant coefficients. To do this in a energetically consistent way, we present an argument similar to one given by White [58]. If the terms $\frac{\partial \chi^*}{\partial y}$ from (2.71) and $\frac{\partial \psi^*}{\partial y}$ from (2.72) are omitted, and f is set to f_0, we obtain

$$\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x}\right)\nabla^2 \psi^* + f_0 \nabla^2 \chi^* + \beta\left(\frac{\partial \psi^*}{\partial x}\right) = 0, \tag{2.75}$$

$$\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x}\right)\nabla^2 \chi^* - f_0 \nabla^2 \psi^* + \beta\left(\frac{\partial \chi^*}{\partial x}\right) + g\nabla^2 h^* = 0. \tag{2.76}$$

The removal of the two terms are necessary to imply conservation of kinetic energy which occurs with the non-approximated version of the SWEs (2.29), (2.22).

2.6 Why Shallow Water Equations?

The SWEs on a rotating sphere are a compromise between simplicity and dynamical realism invoked by 3D models describing the three-dimensional fluid flow over the sphere.

In a general three-dimensional situation the lines of constant density and pressure are not always coincident. The flow is defined to be *baroclinic* and can exhibit vertical shearing which cause meteorological phenonomena such as fronts. Dynamical behavior in the upper troposhere can influence the flow at the surface. Temperature and humidity affects the dynamical behavior of the atmosphere.An example is the vertical shearing of winds that generally occurs in the atmosphere in the presence of horizontal temperature gradients.

Hydrostatic balance (2.13) is assumed in most three-dimensional models of the atmosphere and is a generally a good approximation for the gross atmospheric features within global forecast models. However in SWEs we integrate this assumption and make the flow incompressible setting the density to a constant value. In reality, density does vary within the troposphere and is not incompressible. However for purely horizontal flow the atmosphere behaves as if it were an incompressible fluid.

Baroclinity is not supported in the SWEs. Instead, contours of constant pressure trivially equate to contours of constant density. Models that exhibit such behaviour are defined to be *barotropic* and are restricted in the type of motions which they can give. The SWEs are also inviscid and without friction. In reality the atmosphere is viscous and there is friction occuring with the surface. Realistic atmospheric models have additional boundary layer mechanisms.

As shown in the derivation in Section 2.2, the SWEs assume there is no vertical shear and that the implied vertical velocity is given by equation (2.17). Such an approximation assumes that the fluid is *shallow* with the range of heights considered being small compared to wavelengths in the horizontal direction. It also expects the fluid to have weak vertical motion. The lack of vertical shear is the most serious limitation of the SWEs as an atmospheric model.

The SWEs are not a viable atmospheric model as it has far too many limitations to its behaviour. Even so, the SWEs on a rotating sphere have many dynamical mechanisms which are revelent to the more general problem. The spherical geometry has in itself made interesting features with quite specialised boundary conditions. The effects of a rotating sphere are considered by a variable Coriolis term f. The SWEs exhibit non-trivial solutions due their nonlinearity in the advective term and have slow and fast aspects to the flow which behave very differently from each other. These properties are examined in the following chapter, where the concepts of *balance* and *geostrophic adjustment* are introduced.

Chapter 3

Dynamical Behaviour of the SWEs

3.1 Introduction

In Chapter 2 we derived the SWEs. We now present the dynamical heart of the SWEs. In particular we present the concepts of characteristic scales and regimes in Section 3.2, wave solutions, balance in Section 3.3 and geostrophic adjustment in Section 3.6. We show that divergence tendency, as defined in Section 3.7, in general does not always filter the 'noise' aspects of the flow: additional conditions also need to be established. We present an example of a wave 'in balance' - the Rossby–Haurwitz wave in Section 3.4. We examine the departure from linear balance and the divergence and show that dynamically, for a simplified problem, they are propagated by a linear combination of the eigenmodes of the dynamical system. We then finally look at in Section 3.9 the behaviour of perturbations satisfying a linear balance relationship and linearised potential vorticity under different Burger regimes.

3.2 Characteristic Scales, Regimes: Rossby and Burger numbers

Atmospheric dynamics typically involve an interaction of waves with various wavelengths and amplitudes. A technique to identify the relative importance of one term in a set of equations over another is to non-dimensionalise the problem and assume that the flow is characterised by a typical velocity U and typical height H. We assume that the height and winds are harmonic and the flow is identified by a single typical wavelength λ. The characteristic horizontal length scale L is then equal to $L = \lambda/(2\pi)$ [16]. The corresponding time-scale for the height and wind fields is set to L/U.

Using the scaling defined above the non-dimensional quantities, denoted by $\breve{}$, are

$$\begin{aligned} \mathbf{v} &= U\breve{\mathbf{v}}, \\ h &= H\breve{h}, \\ (\lambda, \theta, t) &= \left(L\breve{\lambda}, L\breve{\theta}, \frac{L}{U}\breve{t}\right) \end{aligned} \tag{3.1}$$

where h, $\mathbf{v}$ are the height and wind fields on the spherical surface, defined at the end of Section 2.2. The latitude, longitude co-ordinates are given by λ, θ and t denotes the time.

Introducing equations (3.1) into the momentum equation part of the full non-linear SWE on the rotating sphere (2.29) gives

$$\frac{U^2}{L}\left(\frac{\partial\breve{\mathbf{v}}}{\partial\breve{t}} + \nabla\left[\left(\breve{\mathbf{v}}\cdot\breve{\mathbf{v}}\right)/2\right] + \left(\nabla\times\breve{\mathbf{v}}\right)\times\breve{\mathbf{v}}\right) + fU\mathbf{k}\times\breve{\mathbf{v}} = -\frac{gH}{L}\nabla\breve{h}. \tag{3.2}$$

where, for this section, the vector operators are applied to scaled $\left(\breve{\lambda}, \breve{\theta}, \breve{t}\right)$.

Dividing equation (3.2) by fU gives

$$\begin{aligned} \frac{U}{fL}\left(\frac{\partial\breve{\mathbf{v}}}{\partial t} + \nabla\left[\left(\breve{\mathbf{v}}\cdot\breve{\mathbf{v}}\right)/2\right] + \left(\nabla\times\breve{\mathbf{v}}\right)\times\breve{\mathbf{v}}\right) & \\ +\mathbf{k}\times\breve{\mathbf{v}} &= -\frac{gH}{fUL}\nabla\breve{h}, \\ \Rightarrow \frac{U}{fL}\left(\frac{\partial\breve{\mathbf{v}}}{\partial t} + \nabla\left[\left(\breve{\mathbf{v}}\cdot\breve{\mathbf{v}}\right)/2\right] + \left(\nabla\times\breve{\mathbf{v}}\right)\times\breve{\mathbf{v}}\right) & \\ +\mathbf{k}\times\breve{\mathbf{v}} &= -\left(\frac{gH}{f^2L^2}\right)\left(\frac{fL}{U}\right)\nabla\breve{h}, \quad (3.3) \end{aligned}$$

which by setting the non-dimensional numbers

$$R_o = \frac{U}{fL}, \quad (3.4)$$

$$B_u = \frac{\sqrt{gH}}{fL}, \quad (3.5)$$

gives

$$R_o\left(\frac{\partial\breve{\mathbf{v}}}{\partial t} + \nabla\left[\left(\breve{\mathbf{v}}\cdot\breve{\mathbf{v}}\right)/2\right] + \left(\nabla\times\breve{\mathbf{v}}\right)\times\breve{\mathbf{v}}\right) + \mathbf{k}\times\breve{\mathbf{v}} = -\frac{B_u^2}{R_o}\nabla\breve{h}. \quad (3.6)$$

The non-dimensional numbers, R_o and B_u, are called the Rossby number and the Burger number, respectively. In the situation, where R_o is made large ($R_o >> 1$) and B_u is arbitrarily fixed ($B_u \approx 1$) the equation (3.6) is approximated by

$$R_o\left(\frac{\partial\breve{\mathbf{v}}}{\partial t} + \nabla\left[\left(\breve{\mathbf{v}}\cdot\breve{\mathbf{v}}\right)/2\right] + \left(\nabla\times\breve{\mathbf{v}}\right)\times\breve{\mathbf{v}}\right) = 0. \quad (3.7)$$

In contrast, when R_o is made small ($R_o << 1$) and B_u is arbitrarily fixed, U and $\frac{D\mathbf{v}}{Dt}$ are reduced in size with equation (3.6) approximated by

$$\mathbf{k}\times\breve{\mathbf{v}} = -\frac{B_u^2}{R_o}\nabla\breve{h}. \quad (3.8)$$

Thus when $R_o << 1$ and $B_u^2 << R_o$, the size of $\nabla\breve{h}$ is larger in magnitude than the velocity $\breve{\mathbf{v}}$. In contrast, when $R_o << 1$ and $B_u^2 >> R_o$, the size of $\nabla\breve{h}$ is smaller in magnitude than the velocity $\check{\mathbf{v}}$. The consequences of this are examined in Sections 3.6 and 3.9.

The three-dimensional atmosphere tends to have large horizontal length scales and relatively small vertical length scales and can be approximated by being considered as a number of layers of fluid on top of each other. A fluid with this property is said to be stably stratified. The Burger number describes the relative importance of the effects of stratification and rotation. When this number is larger than one the layers are stable with respect to changes in the interfaces between them ; for Burger number much smaller than one the rotation dominates the flow.

The Burger number is described in numerous ways dependent on the source. Pedlosky [43] defines for two-dimensional horizontal flow the non-dimensional number as $gH/(f^2L^2)$, the square of the quantity described here. Haltiner et al. [23] defines the Burger number as the ratio between the the Rossby radius of deformation, defined as

$$L_r = \frac{\sqrt{gH}}{f} \tag{3.9}$$

and the characteristic length scale L, which is identical to the definition given in equation (3.5).

As described in Chapter 2, the SWEs are defined on a two-dimensional surface and consist of a single layer of fluid. A non-trivial interface is considered when topography is included. In this situation, stable stratification manifests itself when the single layer is not affected greatly by topography and flows over obstacles, (compared to going around them). A pseudo vertical length scale exists which is given by the Rossby radius (3.9).

3.3 Wave Solutions and Balance

SWEs are well known for having more than one type of wave solution progressing at various time-scales. Some of these waves are more important than others. In this section we distinguish between the various waves where we generalise the treatments given by Haltiner and Williams [23], White [58]. We also refer to comments given by Thompson [55].

Let us consider an approximation to the β-plane SWE model, linearised about a steady wind $\overline{\mathbf{v}} = (u^+, 0)$ under a Cartesian geometry, as described by equations (2.75), (2.76) and (2.73). Since all three equations are partial differental equations (PDEs) with constant coefficients, it is easy to reduce the system into a single higher order PDE with one variable. This is achieved by applying $\left(\frac{\partial}{\partial t}\nabla^2 + u^+\frac{\partial}{\partial x}\nabla^2 + \beta\frac{\partial}{\partial x}\right)\nabla^2$ (2.76) + $f_0\nabla^4$ (2.75) to eliminate the streamfunction perturbation ψ^*, to give

$$\begin{aligned}
n_1 \left(\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x} \right) \nabla^2 + \beta \frac{\partial}{\partial x} \right)^2 \nabla^2 \chi^* \\
+ \left(\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x} \right) \nabla^2 + \beta \frac{\partial}{\partial x} \right) \left(-f_0 \nabla^4 \psi^* + g \nabla^4 h^* \right) \\
+ f_0 \nabla^4 \left(\left(\frac{\partial}{\partial t} + u^+ \frac{\partial}{\partial x} \right) \nabla^2 + \beta \frac{\partial}{\partial x} \right) \psi^* + f_0^2 \nabla^6 \chi^* &= 0, \\
\Rightarrow \left(n_1 \left(\frac{\partial}{\partial t} \nabla^2 + u^+ \frac{\partial}{\partial x} \nabla^2 + \beta \frac{\partial}{\partial x} \right)^2 + f_0^2 \nabla^4 \right) \nabla^2 \chi^* \\
+ g \left(\frac{\partial}{\partial t} \nabla^2 + u^+ \frac{\partial}{\partial x} \nabla^2 + \beta \frac{\partial}{\partial x} \right) \nabla^4 h^* &= 0, \quad (3.10)
\end{aligned}$$

with n_1 being a tracing label, identifying the contribution given by terms $\left(\left(\frac{\partial}{\partial t} + u^+\frac{\partial}{\partial x}\right)\nabla^2 + \beta\frac{\partial}{\partial x}\right)\nabla^2\chi^*$, from the divergence equation (2.76).

We now replace $\nabla^2\chi^*$ for $-\frac{1}{H}\left(\frac{\partial}{\partial t} + u^+\frac{\partial}{\partial x}\right)h^*$ in equation (3.10) using (2.73), to obtain

$$\begin{aligned}\Bigg[\left(n_1\left(\frac{\partial}{\partial t}\nabla^2+u^+\frac{\partial}{\partial x}\nabla^2+\beta\frac{\partial}{\partial x}\right)^2+f_0^2\nabla^4\right)\left(\frac{\partial}{\partial t}+u^+\frac{\partial}{\partial x}\right)&\\ -gH\left(\frac{\partial}{\partial t}\nabla^2+u^+\frac{\partial}{\partial x}\nabla^2+\beta\frac{\partial}{\partial x}\right)\nabla^4\Bigg]h^* &= 0.(3.11)\end{aligned}$$

We assume that the height perturbations have a harmonic structure

$$h^*=\hat{h}e^{i(k_1x+k_2y-\sigma t)} \tag{3.12}$$

where $\hat{h}$ is a complex coefficient associated with the both wavenumbers, k_1 and k_2, and the angular frequency σ. The symbol i denotes the imaginary number satisfying $i^2=-1$. Introducing (3.12) into (3.11), produces a cubic polynomial in σ

$$\begin{aligned}\left[f_0^2K^4-n_1\left(\sigma K^2-u^+k_1K^2+\beta k_1\right)^2\right]\left(-i\sigma+u^+ik_1\right)&\\ -gH\left(i\sigma K^2-u^+ik_1K^2+\beta ik_1\right)K^4 &= 0\\ \Rightarrow\left[f_0^2K^4-n_1\left(\sigma K^2-u^+k_1K^2+\beta k_1\right)^2\right]\left(\sigma-u^+k_1\right)&\\ +gH\left(\sigma K^2-u^+k_1K^2+\beta k_1\right)K^4 &= 0 \qquad (3.13)\end{aligned}$$

where $K^2=k_1^2+k_2^2$.

For $n_1=1$, the three roots of the cubic equation give the typical frequencies of three wave solutions. The two largest roots of the cubic (3.13), define *inertio − gravity waves*. In the non-rotating case, the dispersion relation would just be just $\sigma-k_1u^+=\pm K\sqrt{gh}$; in the rotating case we expect the roots to be in the neighbourhood of $\sigma-k_1u^+=\pm K\sqrt{gh}$. If we consider the f-plane case, we can ignore the β terms in (3.13) and rearrange. The dispersion relation is then given by

$$\sigma-k_1u^+=\pm K\sqrt{gH+\frac{f_0^2}{K^2}}. \tag{3.14}$$

[A dispersion relation for the β-plane can be obtained by retaining the β terms in (3.13), and solving the resulting cubic using Vieta's subsitution [57]. We leave such details as they serve as a distraction to the discussion given.]

The smallest root is approximated by setting the tracer n_1 to zero. In this situation, the cubic polynomial (3.13) reduces to a linear equation in σ, where

$$\sigma = k_1 u^+ - \frac{gH\beta k_1}{gHK^2 + f_0^2} = k_1 u^+ - \frac{\beta k_1}{K^2 + \frac{f_0^2}{gH}}. \qquad (3.15)$$

The dispersion relation (3.15) defines the angular frequency of what is identified as a *Rossby wave.*

Setting the tracer n_1 to zero is equivalent to setting the total material time derivative term and $\beta\frac{\partial\chi}{\partial x}$ of the divergence equation (2.76) to zero. As observed in [55], the condition that is necessary and sufficient for the elimination of inertio-gravity waves of the form (3.14) requires $\left(\sigma K^2 - u^+ k_1 K^2 + \beta k_1\right)^2 = 0$. However it is necessary for the existence of solutions of the type (3.15) that $(\sigma K^2 - u^+ k_1 K^2 + \beta k_1)$ does not vanish, justifying the need to set the tracer to zero. This is called the *generalised filtering approximation.* The remaining part of the divergence equation is given by the linear balance equation,

$$\nabla \cdot f_0 \nabla \psi = g \nabla^2 h. \qquad (3.16)$$

The main balance relation used in this thesis resembles (3.16). However, we generally consider a spherical domain and allow the Coriolis parameter f to vary with latitude. The resulting equation (3.37) is described in Section 3.5 and in Chapter 4.

When the inertio-gravity waves are no longer present, the fluid is considered to be in *balance.* This occurs when the dispersion relations related to the two largest roots of the cubic equation (3.13) are not exhibited by the flow in question; the motion of the fluid only is described by the dispersion relation defined by the smallest root of the frequency equation. Models which propagate only Rossby waves are called *balanced models*. There are a number of techniques to approximate balance and produce balanced models, obtained from using semi-geostrophic or quasi-geostrophic theories [21]. They all share the property that provided we consider linearised SWEs with constant coefficients, the dispersion relation defined by the smallest root of the SWEs cubic frequency equation is equivalent to the linear dispersion relation of the respective balanced model.

Typically, in a mid-latitude region the characteristic height H is approximately equal to 10 km with the inertio-gravity waves and Rossby waves having speeds around $300ms^{-1}$ and $10ms^{-1}$, respectively. This shows the large separation in timescales between the two types within the mid-latitudes. Pairs of inertio-gravity waves with same angular frequency and amplitude move in opposite directions to each other. The Rossby wave propagates westwards which is in the direction perpendicular to the potential vorticity gradient relative to the mean flow.

Given the low angular frequency of the Rossby wave, the wave phase speed is expected to be slow. This is true for linearised equations. However, when the nonlinear advective term $(\mathbf{v} \cdot \nabla)\,\mathbf{v}$ is present, slow Rossby waves interact with each other to give waves that are slower or faster. Instead of there being a clear distinction between the timescales of fast inertio-gravity waves and slow Rossby waves, the non-linear interactions produce Rossby waves with a wide range of angular frequencies. The amplitude and energy present within these waves diminish with increased angular frequency. However in a non-linear description of balanced flow all Rossby waves need to be considered. Potential vorticity is a good variable to choose in this respect as it can in practice exhibit a wide range of frequencies. The effects of Rossby wave interaction is left to the numerical experiments in Chapter 7; the theory presented in this thesis limits itself to properties present in linearisations of the SWEs.

About the equator there exists additional types of waves localised to the region. Gill [22] examines these waves by examining a Cartesian equatorial β plane model, where $f = \beta y$. The waves are called equatorial *Kelvin waves* and *mixed Rossby – gravity waves* and have timescales which lie between those of inertio-gravity waves and Rossby waves, complicating the dynamical situation. Thus, for the purpose of simplicity, the emphasis of this thesis is concerned with motions away from the equator.

3.4 Example of balanced flow: Rossby–Haurwitz Wave

The Rossby Haurwitz wave is an example of a balanced flow. First identified by Haurwitz [24], it is a wave which moves without changing its profile for the two-dimensional Euler equations. Under a Cartesian mid-latitude β-plane assumption, the two-dimensional Euler equations are given by

$$\begin{aligned}
\frac{\partial u}{\partial t} + u^{+}\frac{\partial u}{\partial x} - (f_0 + \beta y)v + g\frac{\partial h}{\partial x} &= 0 \qquad (3.17)\\
\frac{\partial v}{\partial t} + u^{+}\frac{\partial v}{\partial x} + (f_0 + \beta y)u + g\frac{\partial h}{\partial y} &= 0 \qquad (3.18)\\
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= 0. \qquad (3.19)
\end{aligned}$$

This system of equations occurs when the SWEs are constrained to have no divergent wind, which approximately occurs when the variations in the height field h^* are small compared to the magnitude of the height field H itself. We now present a derivation of a Rossby–Haurwitz wave, similar to the treatment given by Dutton [16].

If the waves are assumed to propagate only in the x direction, we can let the solutions take the form

$$\begin{aligned} u &= \hat{u}e^{(k_1x-\sigma t)i}, & (3.20)\\ v &= \hat{v}e^{(k_1x-\sigma t)i}, & (3.21)\\ h &= \hat{h}e^{(k_1x-\sigma t)i}, & (3.22)\end{aligned}$$

and substitute (3.20)-(3.22) into (3.17)-(3.19), to obtain

$$\begin{aligned} i\left(-\sigma+k_1u^+\right)\hat{u}-(f_0+\beta y)\hat{v}+igk_1\hat{h} &= 0 & (3.23)\\ i\left(-\sigma+k_1u^+\right)\hat{v}+(f_0+\beta y)\hat{u}+g\frac{\partial\hat{h}}{\partial y} &= 0 & (3.24)\\ ik_1\hat{u}+\frac{\partial\hat{v}}{\partial y} &= 0. & (3.25)\end{aligned}$$

Taking $k_1\frac{\partial}{\partial y}$(3.23) - ik_1^2(3.24), eliminates $\hat{h}$, giving

$$\left(-\sigma+k_1u^+\right)\left(ik_1\frac{\partial\hat{u}}{\partial y}+k_1^2\hat{v}\right)-k_1\beta\hat{v}-k_1(f_0+\beta y)\left(ik_1\hat{u}+\frac{\partial\hat{v}}{\partial y}\right)=0. \qquad (3.26)$$

The $\hat{u}$ is eliminated using condition (3.25) to produce

$$\begin{aligned} \left(-\sigma+k_1u^+\right)\left(-\frac{\partial^2\hat{v}}{\partial y^2}+k_1^2\hat{v}\right)-k_1\beta\hat{v} &= 0, \\ \Rightarrow \frac{\partial^2\hat{v}}{\partial y^2}+\left(\frac{\beta}{u^+-\frac{\sigma}{k_1}}-k_1^2\right)\hat{v} &= 0 & (3.27)\end{aligned}$$

provided $\frac{\sigma}{k_1}$ is not equal to u^+. (When $\frac{\sigma}{k_1}=u^+$, $\hat{v}$ is equal to zero and the wave moves at constant speed u^+ in the x direction.). The general solution of this second order ODE with constant coefficients is given by

$$\hat{v}=C_1\cos\left[\left(\frac{\beta}{u^+-\frac{\sigma}{k}}-k_1^2\right)^{\frac{1}{2}}y\right]+C_2\sin\left[\left(\frac{\beta}{u^+-\frac{\sigma}{k}}-k_1^2\right)^{\frac{1}{2}}y\right], \qquad (3.28)$$

where C_1, C_2 are constants which are fixed by appropriate boundary conditions. If the boundary conditions are such that the domain considered is a channel of width D for which $|\hat{v}|$ is at a maximum at $y=0$ and zeros at $y=\pm D/2$, then

$$\begin{aligned} C_2 &= 0 \\ \left(\frac{\beta}{u^+ - \frac{\sigma}{k_1}} - k_1^2\right)^{\frac{1}{2}} \frac{D}{2} &= \frac{\pi l}{2} \quad \text{for} \quad l = \pm 1, \pm 3, \dots \quad . \end{aligned} \tag{3.29}$$

The dispersion relation is obtained by rearranging (3.29) into

$$\sigma = k_1 u^+ - \frac{\beta k_1}{k_1^2 + \left(\frac{l\pi}{D}\right)^2}. \tag{3.30}$$

The relationship between the Rossby wave within the SWEs and the Rossby–Haurwitz wave under two-dimensional Euler equations are readily seen by letting $H \to \infty$ in (3.15) and $D \to \infty$ in (3.30). In both situations, the Rossby formula reduces to

$$\sigma = k_1 u^+ - \frac{\beta k_1}{K^2}. \tag{3.31}$$

where, in this case, $K^2 = k_1^2$.

As defined, the Rossby–Haurwitz wave consists of a balanced wind field for the equation set (3.17)-(3.19); for it to be used in SWEs context a balanced height also needs to be calculated that is consistent with the SWEs. For this purpose a Charney balance condition is used. This balance condition, first defined by Charney [8] [6], assumes that the balanced wind is defined by the rotational part of the flow, with

$$\mathbf{v} = \mathbf{k} \times \nabla \psi. \tag{3.32}$$

Applying the divergence operator to the momentum equations and using (3.32), produces the balance condition

$$g\nabla^2 h = \nabla \cdot (f_0 + \beta y) \nabla \psi + 2\left(\left(\frac{\partial^2 \psi}{\partial x \partial y}\right)^2 - \frac{\partial^2 \psi}{\partial x^2}\frac{\partial^2 \psi}{\partial y^2}\right). \tag{3.33}$$

Solving this Monge-Ampere equation provides the appropriate balanced height. It is important to note that the Rossby–Haurwitz wave in SWEs context does not produce a balanced flow that stays balanced when propagated in time. At best, under a high Burger regime, the Rossby–Haurwitz wave produces SWEs solutions over 12 hours with relatively little divergence [6] which are close to the balanced flow given by the two-dimensional Euler equations. As such, it is used as an initial solution which produces solutions over a $12hrs - 24hrs$ timescale that is close to balance.

3.4.1 The Rossby–Haurwitz wave on a Sphere

In practice, throughout this thesis the initial height and wind field relating to a Rossby–Haurwitz wave is defined over a sphere [44]. This wave is characterised by parameters a, g, Ω, R, h_0, ω and K, where a is the radius of the sphere, g is the acceleration due to gravity, R is the wave number and h_0 is the height at the poles. The strength of the underlying zonal wind from west to east is given by ω and K controls the amplitude of the wave. As in Chapter 2 the latitude and longitude co-ordinates are represented by θ and λ.

The initial velocity field is defined as,

$$\begin{aligned} u &= a\omega\cos\theta + aK\cos^{R-1}\theta(R\sin^2\theta - \cos^2\theta)\cos R\lambda, \\ v &= -aKR\cos^{R-1}\theta sin\theta sin R\lambda. \end{aligned} \tag{3.34}$$

The initial height field is defined as,

$$h = h_0 + \frac{a^2}{g}[A(\theta) + B(\theta)\cos R\lambda + C(\theta)\cos(2R\lambda)], \tag{3.35}$$

where the variables $A(\theta),\ B(\theta),\ C(\theta)$ are given by

$$\begin{aligned} A(\theta) &= \frac{\omega}{2}(2\Omega+\omega)\cos^2\theta + \frac{1}{4}K^2\cos^{2R}\theta[(R+1)\cos^2\theta \\ &\quad +(2R^2-R-2) - 2R^2\cos^{-2}\theta], \\ B(\theta) &= \frac{2(\Omega+\omega)K}{(R+1)(R+2)}\cos^R\theta[(R^2+2R+2) \\ &\quad -(R+1)^2\cos^2\theta], \\ C(\theta) &= \frac{1}{4}K^2\cos^{2R}\theta[(R+1)\cos^2\theta-(R+2)]. \end{aligned} \tag{3.36}$$

3.5 Linear Balance Equation

In Section 3.3 we derived the linear balance equation (LBE) by applying the general filtering approximation to SWEs defined on a Cartesian mid-latitude β-plane (2.73), (2.75) and (2.76). More generally LBE is defined over the sphere, where

$$g\nabla^2 h - \nabla\cdot f\nabla\psi = 0 \tag{3.37}$$

and ψ, the streamfunction is defined by (2.47). In subsequent chapters this balance relation is compared with another which conserves potential vorticity. Consequently, the properties of this balance condition need to be described.

The LBE is viewed in more than one way. Burger [7] considers the LBE as a simple generalisation of geostrophic balance over the whole sphere for waves of planetary length scale $L \approx a$. By applying scaling arguments with this length scale to the divergence equation (2.61) about mid-latitudes, the terms in (3.37) are found to be ten times larger than the other terms in the divergence equation.

It is also a linear non-divergent mass-wind law that naturally takes into account the latitudinal variation of the Coriolis parameter and is useful when length scales

$L \approx 10^6 m$ are considered [12]. However, balanced divergent parts to a wind do exist for the SWEs on a sphere [49] and are 'invisible' to this balance condition. This is seen when appropriate equations are added to LBE to produce a closed energically-consistent dynamical system [58]. This requires not only a modified vorticity equation but also a thermodynamic equation. The kinetic energy of this particular dynamical system comes from only the rotational part of the flow and no divergent contribution exists.

Two problems need to be considered: the calculation of a balanced height field from the streamfunction

$$g\nabla^2 h = \nabla \cdot f \nabla \psi \tag{3.38}$$

and the backward relation

$$\nabla \cdot f \nabla \psi = g\nabla^2 h \tag{3.39}$$

where the streamfunction is determined by the height.

The calculation of a balanced height field from the streamfunction is straightforward; the existence, uniqueness and boundary conditions are the same as those needed to invert a Poisson equation on a sphere and are given in Section 5.5.

The reverse transformation, the transformation from height field to a streamfunction, is a little more complex. The majority of the attempts to solve (3.39) set the problem in terms of spherical harmonics [37], [1]. Daley [14] shows that solutions to the reverse transformation can become singular about the equator. The problem is worsened by the sensitivity of the solution to the height field localised about the equator; small errors in the height field in this region can trigger the presence of these singular solutions.

In [12] Daley considers the practical use of (3.39) with height fields including error and shows that the solution tends to be erroneous within 20 degrees latitude of the equator. A means to get around this difficulty is to modify the linear balance equation so that these singularities do not exist [14] using an approach that uses the singular value decomposition method. This procedure has the effect of making the reverse LBE as accurate as the forward LBE outside the tropics as well as removing possible singularies in the streamfunction around the equator. A consequence of these considerations is that the reverse transformation is not used.

3.6 Geostrophic Balance and Geostrophic Adjustment

In Section 3.2, we showed that when R_o is small, the acceleration term $D\mathbf{v}/Dt$ is negligible compared to the Coriolis term with equation (3.8) being approximately satisfied for non-dimensional $\hat{\mathbf{v}}$, $\hat{h}$. If this relationship holds exactly, then the height and wind fields are said to be in *geostrophic balance*. If the height and wind fields are defined on a sphere of radius a then geostrophic balance is written as:

$$f\mathbf{k} \times \mathbf{v} - g\nabla h = 0 \tag{3.40}$$

Similarly the *geostrophic wind*, $\mathbf{v}_g$ is defined to be

$$\mathbf{v}_g = \frac{g}{f}\mathbf{k} \times \nabla h. \tag{3.41}$$

Most wind and height fields considered do not satisfy balance conditions unless enforced to do so. It is beneficial to define the discrepancy in how far the winds and height fields depart from such relations. The *geostrophic departure* $\boldsymbol{r}^d$ is given by

$$\boldsymbol{r}^d = f\mathbf{k} \times \mathbf{v} - g\nabla h, \tag{3.42}$$

and is a measure of how discrepant the wind and height fields are from geostrophic balance. We also define the departure from the linear balance relation (3.16) as

$$r = \nabla \cdot f\nabla\psi - \nabla^2 h. \tag{3.43}$$

We notice that the departure r is related to $\boldsymbol{r}^d$, just as the divergence equation is related to the momentum equations.

The geostrophic balance relationship is intimately linked to the mechanism of *geostrophic adjustment.* Haltiner and Williams [23] and Schoenstadt [50] showed that for SWEs linearised about a resting state on a f-plane under a Cartesian geometry, the Rossby wave is stationary. The transient flow is described by the inertio-gravity waves which decays as it moves along the x-axis. The Rossby waves are dependent on whether the characteristic horizontal scale is larger or smaller than the Rossby radius of deformation L_r.

The stationary Rossby wave has a form highly dependent on the Burger number. When $|B_u| >> 1$, the final wind is equal to the initial wind and the final height in geostrophic balance with the initial wind. When $|B_u| << 1$, the final height is determined by the initial height field with the final wind similarly in geostrophic balance with the initial height. On a sphere the effect of geostrophic adjustment is similar to the above situation, provided viscous terms are added to SWEs to eliminate the inertio-gravity waves before they travel the circumference of the sphere [60].

3.7 Divergence Tendency

A justification for using the Charney balance condition to derive a balanced height field as an initial condition for the SWEs is due to the observation that introducing a purely rotational wind into the divergence equation sets the local change in the divergent wind, the *divergence tendency* $\partial\nabla^2\chi/\partial t$, to be zero [8].

Let us consider SWEs linearised about a resting state on an f-plane. The setting of the divergence tendency to zero reduces the cubic frequency equation to a linear one and gives the dispersion relationship for the balanced fields. In this situation the inertio-gravity waves are fully filtered out of the system. The divergence equation reduces to (3.16). However when the SWEs are linearised about a constant velocity u^+ and a β-plane approximation is used, setting the divergence tendency to zero only reduces the cubic frequency equation to a quadratic; the different linearisation state gives a set of perturbed equations which no longer have the symmetry that exists in the f-plane case and the inertio-gravity waves are no longer fully filtered.

We can see this again, by considering an f-plane model of the SWEs, linearised about a resting state

$$\begin{aligned} \frac{\partial\zeta}{\partial t} + f_0\delta &= 0 && (3.44)\\ \frac{\partial\delta}{\partial t} - f_0\zeta &= -g\nabla^2 h && (3.45)\\ \frac{\partial h}{\partial t} + H\delta &= 0 && (3.46)\\ & && (3.47) \end{aligned}$$

where f_0 and H are constant values. Setting the partial time derivative $\frac{\partial\delta}{\partial t}$ to zero, enforces the divergence δ to be constant. If we take the remaining terms of the divergence equation and apply the partial time derivative operator we get

$$f_0\frac{\partial\zeta}{\partial t} - g\frac{\partial\nabla^2 h}{\partial t} = 0 \qquad (3.48)$$

$$\Rightarrow gH\nabla^2\delta - f_0^2\delta = 0, \qquad (3.49)$$

substituting the partial time derivative for the divergence using the vorticity and continuity equations above. This Helmholz equation is trivally satisfied as the laplacian of the divergence is zero, making the solution, the divergence, equal also to zero. The unbalanced part of the SWEs has been filtered.

We can use a similar technique when SWEs are linearised about a general time-varying state, but due to the additional terms we need to set not only $\frac{\partial\delta}{\partial t}$ to zero, but also a higher local time derivative to be zero. This is because in this more general situation two time constraints are needed to eliminate the two of the three local time derivatives dependencies present within the SWEs. The time constraints have to be chosen such that it is the fast inertio-gravity waves are eliminated. In Section 4.6 we set both $\frac{\partial\delta}{\partial t}$, $\frac{\partial^2\delta}{\partial^2 t}$ to find balanced parts of our control variables.

3.8 Relationship Between Potential Vorticity, Geostrophic Departure and Divergence

Consider a mid-latitude f-plane Cartesian approximation to the SWEs linearised about a resting state

$$\begin{aligned}
\frac{\partial u^*}{\partial t} - f_0 v^* + g\frac{\partial h^*}{\partial x} &= 0,\\
\frac{\partial v^*}{\partial t} + f_0 u^* + g\frac{\partial h^*}{\partial y} &= 0,\\
\frac{\partial h^*}{\partial t} + H\left(\frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y}\right) &= 0. \qquad (3.50)
\end{aligned}$$

We apply the spatial Fourier decomposition

$$\hat{\boldsymbol{X}} = \begin{pmatrix} \hat{u}_{k_1,k_2}(t) \\ \hat{v}_{k_1,k_2}(t) \\ \hat{h}_{k_1,k_2}(t) \end{pmatrix} = \int_{y=-\infty}^{y=\infty} \int_{x=-\infty}^{x=\infty} \begin{pmatrix} u^*(x,y,t) \\ v^*(x,y,t) \\ h^*(x,y,t) \end{pmatrix} e^{-i(k_1 x + k_2 y)}\, dxdy, \qquad (3.51)$$

in the xy plane to transform the coupled system of PDE's into a coupled system of ODE's which for each k_1, k_2 is represented by

$$\hat{\boldsymbol{X}}_t + L\hat{\boldsymbol{X}} = 0, \qquad (3.52)$$

with

$$L = \begin{pmatrix} 0 & -f_0 & igk_1 \\ f_0 & 0 & igk_2 \\ iHk_1 & iHk_2 & 0 \end{pmatrix}. \qquad (3.53)$$

We now calculate the eigenvalues and eigenvectors of the system. The determinant of the characteristic equation is given by

$$det\,|L - i\lambda_+ I| = i\lambda_+ \left(\lambda_+^2 + \sigma^2\right), \qquad (3.54)$$

where $\lambda_+ = (O, -\sigma, \sigma)$,$\sigma = (f_0^2 + gHK_w^2)^{\frac{1}{2}}$ and $K_w^2 = k_1^2 + k_2^2$. The eigenvalues $i\Lambda$ and eigenvectors $E = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ are related by

$$LE = iE\Lambda \qquad (3.55)$$

with

$$E = (\boldsymbol{e}_1\ \boldsymbol{e}_2\ \boldsymbol{e}_3) = \begin{pmatrix} igk_2 & -igk_1 - \frac{gk_2 f}{\sigma} & \frac{gk_2 f}{\sigma} \\ igk_1 & -igk_2 + \frac{gk_1 f}{\sigma} & -igk_2 - \frac{gk_1 f}{\sigma} \\ f & i\left(\sigma - \frac{f^2}{\sigma}\right) & -i\left(\sigma - \frac{f^2}{\sigma}\right) \end{pmatrix} \qquad (3.56)$$

and

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sigma & 0 \\ 0 & 0 & \sigma \end{pmatrix}. \tag{3.57}$$

The inverse of the matrix of eigenvectors, E^{-1}, is given by

$$E^{-1} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \boldsymbol{f}_3 \end{pmatrix} = \frac{g}{2i\sigma g^2 K_w^2} \begin{pmatrix} -\frac{2gHk_1K_w^2}{\sigma} & \frac{2gHk_1K_w^2}{\sigma} & \frac{2igfK_w^2}{\sigma} \\ -k_1\sigma - ik_2f & -k_2\sigma + ik_1f & gK_w \\ -k_1\sigma + ik_2f & -k_2\sigma - ik_1f & -gK_w \end{pmatrix} \tag{3.58}$$

and is calculated by taking the complex conjugate of the matrix of co-factors of E, divided by the determinant of E.

Now we are in a position to apply a similarity transformation to uncouple the system concerned. Let

$$E\hat{\boldsymbol{Y}} = \hat{\boldsymbol{X}}. \tag{3.59}$$

Applying (3.59) to (3.52) gives

$$\hat{\boldsymbol{Y}}_t + E^{-1}LE\hat{\boldsymbol{Y}} = \hat{\boldsymbol{Y}}_t + i\Lambda\hat{\boldsymbol{Y}} = 0, \tag{3.60}$$

using (3.55). The beauty of this particular similarity transformation is that each variable of $\hat{\boldsymbol{Y}}$ corresponds to an eigenvalue of (3.52).

So far, the system has been uncoupled into variables without any physical significance. We now construct the potential vorticity increment, the departure from linear balance and the divergence out of the uncoupled variables. This is achieved by setting $\hat{\boldsymbol{Z}} = BP\hat{\boldsymbol{Y}} = BPE^{-1}\hat{\boldsymbol{X}}$, with

$$B = \begin{pmatrix} \frac{-\sigma^2}{H^2} & 0 & 0 \\ 0 & gK_w^2 & 0 \\ 0 & 0 & \frac{i\sigma gK_w^2}{f} \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}. \tag{3.61}$$

Evaluating $\hat{\boldsymbol{Z}} = BP\hat{\boldsymbol{Y}} = BPE^{-1}\hat{\boldsymbol{X}}$ gives

$$\hat{\boldsymbol{Z}} \;=\; \begin{pmatrix} \hat{q}_{k_1,k_2} \\ \hat{d}_{k_1,k_2} \\ \hat{r}_{k_1,k_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{H}\left(-ik_2\hat{u}_{k_1,k_2} + ik_1\hat{v}_{k_1,k_2}\right) - \frac{f}{H^2}\hat{h}_{k_1,k_2} \\ ik_1\hat{u}_{k_1,k_2} + ik_2\hat{v}_{k_1,k_2} \\ -ik_2\hat{u}_{k_1,k_2} + ik_1\hat{v}_{k_1,k_2} + \frac{gK_w^2}{f}\hat{h}_{k_1,k_2} \end{pmatrix} \tag{3.62}$$

The variables $\hat{q}_{k_1,k_2}$, $\hat{d}_{k_1,k_2}$, $\hat{r}_{k_1,k_2}$ are the linearised potential vorticity increment, the divergence and the departure from linear balance, respectively within Fourier space. Applying the inverse Fourier transform

$$\boldsymbol{X} \;=\; \begin{pmatrix} u^*(x,y,t) \\ v^*(x,y,t) \\ h^*(x,y,t) \end{pmatrix} = \frac{1}{4\pi^2}\int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty} \begin{pmatrix} \hat{u}_{k_1,k_2}(t) \\ \hat{v}_{k_1,k_2}(t) \\ \hat{h}_{k_1,k_2}(t) \end{pmatrix} e^{i(k_1x+k_2y)}\; dxdy,$$

$$\boldsymbol{Z} \;=\; \begin{pmatrix} q^*(x,y,t) \\ d^*(x,y,t) \\ r^*(x,y,t) \end{pmatrix} = \frac{1}{4\pi^2}\int_{y=-\infty}^{y=\infty}\int_{x=-\infty}^{x=\infty} \begin{pmatrix} \hat{q}_{k_1,k_2}(t) \\ \hat{d}_{k_1,k_2}(t) \\ \hat{r}_{k_1,k_2}(t) \end{pmatrix} e^{i(k_1x+k_2y)}\; dxdy, \tag{3.63}$$

leaves the linearised potential vorticity q^*, the divergence d^* and the departure from linear balance r^* perturbations, with:

$$q^* \;=\; \frac{1}{H}\left(\frac{\partial v^*}{\partial x} - \frac{\partial u^*}{\partial y}\right) - \frac{f}{H^2}h^*, \tag{3.64}$$

$$d^* \;=\; \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y}, \tag{3.65}$$

$$r^* \;=\; f_0\left(\frac{\partial v^*}{\partial x} - \frac{\partial u^*}{\partial y}\right) - g\left(\frac{\partial^2 h^*}{\partial x^2} + \frac{\partial^2 h^*}{\partial y^2}\right). \tag{3.66}$$

As the variables in vector $\hat{\boldsymbol{Y}}$ are uncoupled with respect to each other, each quantity evolves in time separately. The solutions in terms of

$$\hat{\boldsymbol{Y}}(t) = \begin{pmatrix} \hat{y}^{(1)}_{k_1,k_2}(t) \\ \hat{y}^{(2)}_{k_1,k_2}(t) \\ \hat{y}^{(3)}_{k_1,k_2}(t) \end{pmatrix}, \tag{3.67}$$

are given by

$$
\begin{aligned}
\hat{y}^{(1)}_{k_1,k_2}(t) &= \hat{y}^{(1)}_{k_1,k_2}(0), \\
\hat{y}^{(2)}_{k_1,k_2}(t) &= e^{-i\sigma t}\hat{y}^{(2)}_{k_1,k_2}(0), \\
\hat{y}^{(3)}_{k_1,k_2}(t) &= e^{i\sigma t}\hat{y}^{(3)}_{k_1,k_2}(0).
\end{aligned} \tag{3.68}
$$

Since $\hat{\boldsymbol{Z}} = BP\hat{\boldsymbol{Y}}$, the solution $\hat{q}_{k_1,k_2}(t)$, $\hat{d}_{k_1,k_2}(t)$, $\hat{r}_{k_1,k_2}(t)$ at time t in terms of $\hat{y1}_{k_1,k_2}(0)$, $\hat{y2}_{k_1,k_2}(0)$, $\hat{y3}_{k_1,k_2}(0)$ is

$$
\begin{aligned}
\hat{q}_{k_1,k_2}(t) &= \frac{-\sigma^2}{H^2}\hat{y}^{(1)}_{k_1,k_2}(0), \\
\hat{d}_{k_1,k_2}(t) &= gK_w^2\left(e^{i\sigma t}\hat{y}^{(2)}_{k_1,k_2}(0) + e^{-i\sigma t}\hat{y}^{(3)}_{k_1,k_2}(0)\right), \\
\hat{r}_{k_1,k_2}(t) &= \frac{i\sigma gK_w^2}{f}\left(-e^{i\sigma t}\hat{y}^{(2)}_{k_1,k_2}(0) + e^{-i\sigma t}\hat{y}^{(3)}_{k_1,k_2}(0)\right).
\end{aligned} \tag{3.69}
$$

We see, unsurprisingly, that the linearised potential vorticity increment is uncoupled from the departure from linear balance and divergence. It is representing a stationary Rossby wave. In contrast, the departure from linear balance and the divergence are the variables describing the movement of the fast inertio-gravity waves. These two variables are coupled with respect to each other, each being a linear combination of vectors $\boldsymbol{f}_2$ and $\boldsymbol{f}_3$. In the situation where both $\hat{y}^{(2)}_{k_1,k_2}(0) \approx \hat{y}^{(3)}_{k_1,k_2}(0)$ and the wave is represented by a single wavenumber for k_1 and k_2, the divergence $d_{k_1,k_2}(t)$ moves as a cosine wave while $r_{k_1,k_2}(t)$ becomes $2\frac{\sigma g K_w^2}{f}\hat{y}^{(3)}_{k_1,k_2}(0)\sin\sigma t$. We also notice that both waves are moving with frequency σ.

3.9 Relative Contributions to Scaled Potential Vorticity Perturbations

We now consider properties of height, vorticity and potential vorticity perturbations that satisfy both the linearised potential vorticity relationship (3.64) and the linear balance equation (3.16) when the Coriolis term f_0 is constant. It is valid to consider relative vorticity perturbations ζ^*_{rel} since the linear balance equation (3.16) for constant f_0 is equal to

$$f_0 \zeta^*_{rel} = g \nabla^2 h^*. \tag{3.70}$$

We describe how the ratio between scaled perturbations in height and absolute vorticity changes with the Burger number. The change in this ratio is equivalent to a change in the relative contribution of these terms as needed to produce scaled potential vorticity peturbations. This is because the scaled potential vorticity perturbation is defined to be the sum of the scaled perturbations in height and absolute vorticity. To show this mathematically, we define perturbations of any quantity, as in Chapter 2, to be the difference between the true value and its respective linearisation state. By letting perturbations satisfy both the linearised potential vorticity equation and the linear balance equation (3.70), an equation is found that shows how the potential and absolute vorticity perturbations are inextricably linked to the Burger number. Next, we present an equation that links the potential vorticity perturbation to both the height perturbation and the Burger number.

Let us first consider the velocity and height to be on a two dimension Cartesian grid with standard axes (x, y). Suppose that both the linearisation states and the perturbations in the velocity and height are known. The relative vorticity linearisation state $\overline{\zeta_{rel}}$ and the relative vorticity ζ^*_{rel} perturbation are calculated using

$$\overline{\zeta_{rel}} = \frac{\partial \overline{v}}{\partial x} - \frac{\partial \overline{u}}{\partial y}, \qquad \zeta^*_{rel} = \frac{\partial v^*}{\partial x} - \frac{\partial u^*}{\partial y}. \tag{3.71}$$

The full nonlinear potential vorticity q and linearisation state $\overline{q}$ are defined as

$$q = \frac{1}{h}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} + f_0\right) \equiv \frac{\zeta_{rel} + f_0}{h} \tag{3.72}$$

$$\overline{q} = \frac{1}{\overline{h}}\left(\frac{\partial \overline{v}}{\partial x} - \frac{\partial \overline{u}}{\partial y} + f_0\right) \equiv \frac{\overline{\zeta_{rel}} + f_0}{\overline{h}}, \tag{3.73}$$

where (3.73) is the Cartesian f-plane form of (2.63). By linearising the nonlinear potential vorticity equation (3.72), the perturbations in potential vorticity q^*, height h^* and absolute vorticity ζ^*_{rel} are connected by the equation

$$\frac{q^*}{\overline{q}} = \frac{\zeta^*_{rel}}{\overline{\zeta_{rel}} + f_0} - \frac{h^*}{\overline{h}}. \tag{3.74}$$

We use the relationship (3.70) between the perturbation in absolute vorticity and the height to derive a relationship between the potential vorticity perturbation and the height. We consider perturbations in the height and the velocity that take the form $h^* = \hat{h}e^{i(k_1x+k_2y-\sigma t)}$, $u^* = \hat{u}e^{i(k_1x+k_2y-\sigma t)}$, $v^* = \hat{v}e^{i(k_1x+k_2y-\sigma t)}$, where k_1 is the wave number in the x direction, k_2 is the wave number in the y direction and σ is the frequency as in previous sections. Also, we assume that the perturbations satisfy (3.70). Using these two assumptions,

$$\zeta^*_{rel} = -\frac{(k_1^2 + k_2^2)gh^*}{f_0}. \tag{3.75}$$

If the characteristic length scale L is considered to be equal to $(k_1^2 + k_1^2)^{-\frac{1}{2}}$, then the Burger number is equal to $(k_1^2 + k_1^2)^{\frac{1}{2}}(g\overline{h})^{\frac{1}{2}}/f_0$, where the characteristic height scale H is considered to be equal to $\overline{h}$. By using (3.74), (3.75), two separate relationships can be determined: one defines scaled perturbations in potential vorticity in terms of scaled perturbations in height, the other shows how perturbations in scaled relative vorticity perturbations are related to scaled perturbations in potential vorticity. These relationships are given by

$$\frac{q^*}{\overline{q}} = -N\frac{h^*}{\overline{h}} \qquad \left(1 - \frac{1}{N}\right)\frac{q^*}{\overline{q}} = \frac{\zeta^*_{rel}}{\overline{\zeta_{rel}} + f_0} \tag{3.76}$$

with

$$N = 1 + \frac{f_0\ B_u^2}{\overline{\zeta_{rel}} + f_0}. \tag{3.77}$$

As the Burger number is always greater than zero, for any given perturbation, N is always greater than 1. For a fixed $q^*/\overline{q}$ and $N >> 1$, $h^*/\overline{h}$ will not contribute much to the scaled potential vorticity perturbations; the potential vorticity perturbations are sensitive to the absolute vorticity perturbations with $q^*/\overline{q} \ \approx \zeta^*_{rel}/(\overline{\zeta_{rel}} + f_0)$. Moreover, the greater the value of N, the more sensitive $q^*/\overline{q}$ will be to $\zeta^*_{rel}/(\overline{\zeta_{rel}}+f_0)$. The equation (3.77) shows that a number of conditions can make N large. One possible way, assuming $(\overline{\zeta_{rel}} + f_0)$ to be constant, is to produce a large Burger number. A large Burger number will be obtained when $\overline{h}$ is large or when f_0 is small. In summary, it is expected that for large Burger number $q^*/\overline{q}$ will be dominated by changes in $\zeta^*_{rel}/(\overline{\zeta_{rel}} + f_0)$. The equations (3.76) and (3.77) can also be written as

$$\left(1 - \frac{1}{P}\right)\frac{q^*}{\overline{q}} = -\frac{h^*}{\overline{h}} \qquad \frac{q^*}{\overline{q}} = P\frac{\zeta^*_{rel}}{\overline{\zeta_{rel}} + f_0} \tag{3.78}$$

with

$$P = 1 + \frac{\overline{\zeta_{rel}} + f_0}{f_0 B_u^2}. \tag{3.79}$$

It is clear that for a small Burger number, $P >> 1$ with $q^*/\overline{q} >> \zeta^*_{rel}/(\overline{\zeta_{rel}} + f_0)$ and $q'/\overline{q} \approx h'/\overline{h}$. In this situation it is the scaled height perturbations $h'/\overline{h}$

which will dominate $q'/\overline{q}$. Small Burger number regimes will occur where f_0 is not small as in the mid-latitudes and where $\overline{h}$ is small. It is in these regions that the height perturbations will most resemble the potential vorticity perturbations. The linearisation of potential vorticity may not be legitimate. Nonlinear features of the potential vorticity may suppress the relationships suggested above. Thus an important question which this study wishes to examine is whether this analysis transfers to the full non-linear potential vorticity transformation on the sphere.

Chapter 4

Change in control variables: Theoretical Aspects

4.1 Introduction

As mentioned in Chapter1, data assimilation brings together observations and information from a forecast model in some consistent manner. The current means of achieving this at the UK Met. Office involves a formulation of the problem called incremental 3D Variational Data Assimilation (3DVAR). In this chapter we wish to make precise the description of a change of 'control variables' for this formulation of the data assimilation problem. To this end, we provide in Section 4.2 a description of the incremental 3DVAR problem.

In Section 4.3, we describe the change in 'control variables' in the background term of the cost functional in 3DVAR and describe a general framework for change in variables in Section 4.4. We take the method used presently by the UK Met. Office as an example and discuss the strengths and weaknesses of the current change of control variables in relation to what an ‘idealised’ set of control variables should be like.

In Section 4.5, we discuss the advantages and disadvantages of choosing potential vorticity as the balanced control variable with the departure from linear balance and divergence as the two other unbalanced variables. We describe a method to evaluate control variables with a description of the boundary conditions needed. The method readily presents five variables of which three are needed as control variables. We discuss the various choices for the three variables.

In Section 4.6, using the ideas in McIntyre and Norton’s paper on balanced models that conserve potential vorticity [36], we can evaluate a higher order approximation to the balanced part of the flow at a given time. From this higher order approximation we can find an estimate of the balanced parts of the unbalanced variables. We adapt this theory and propose a method to evaluate balanced corrections for various sets of control variables, identifying the associated divergence tendency of each set. This allows a comparision to be made in Chapter 7 between the present set of control variables and the new potential vorticity-based set.

As shown in Section 3.9, the height and wind fields behave differently under various Burger regimes. This is also true with control variables. In Section 4.7 we identify how the potential vorticity control transform behaves under various regimes and show that the solution given by the current control variable set and the new control variable sets are the same in high Burger regimes and vary in low Burger regimes.

4.2 Data Assimilation

The majority of linear data assimilation methods can be considered in terms of a prototype data assimilation problem, using least squares estimation. We specify this problem by showing how it relates to the 3DVAR formulation. We then present incremental 3DVAR as a technique to deal with operators which are weakly nonlinear.

Suppose we have a linear operator H which maps variables $\boldsymbol{x}$ from a forecast into quantities and positions where observations $\boldsymbol{y}$ exist, assuming that any error $\boldsymbol{\epsilon}$ between $\boldsymbol{y}$ and $H\boldsymbol{x}$ is unbiased, i.e.

$$\boldsymbol{y} = H\boldsymbol{x} + \epsilon \qquad\qquad < \epsilon >= 0, \tag{4.1}$$

where <> denotes the statistical expectation operator. We leave basic statistical definitions to textbooks such as [51].

Let us also assume that the background state $\boldsymbol{x}_b$, produced from a forecast, is unbiased with a background covariance matrix B. The Best Linear Unbiased Estimator (BLUE) is defined as

$$\boldsymbol{x}_a = \boldsymbol{x}_b + K(\boldsymbol{y} - H\boldsymbol{x}) \tag{4.2}$$

with

$$K = BH^T(HBH^T + O)^{-1}. \tag{4.3}$$

where K is called the gain matrix and O is a matrix describing the error in the observations.

The analysis $\boldsymbol{x}_a$ is the solution to the data assimilation problem. It is optimal when its difference from the true state is minimised under a L_2 norm. In practice the covariances B and O are too large to calculate explicitly and data assimilation methods are needed to curtail this difficulty. The Optimal Interpolation method assumes that only a few observations are important in calculating each analysis increment, and so only considers observations in local proximity to model variables. A typical variational method avoids the calculation of K. This is written as:

Minimise $\mathcal{J}$ with respect to $\boldsymbol{x}$, where

$$\begin{aligned} \mathcal{J} &= (\boldsymbol{x}-\boldsymbol{x}_b)^T B^{-1}(\boldsymbol{x}-\boldsymbol{x}_b) + (\boldsymbol{y}-H\boldsymbol{x})^T O^{-1}(\boldsymbol{y}-H\boldsymbol{x}), \qquad (4.4) \\ &= \mathcal{J}_b(\boldsymbol{x}) + \mathcal{J}_o(\boldsymbol{x}). \end{aligned}$$

The variational method and (BLUE) are equivalent to each other. This is easily checked by calculating the first variation of $\mathcal{J}$ and setting this to zero at optimal $\boldsymbol{x}_a$. To make the BLUE problem optimal, it is assumed that the matrix K is linear. In practice, the observational operator H tends not be linear, which in turn will produce a nonlinear K. However, if H is weakly nonlinear, we may linearise about the background state $\boldsymbol{x}_b$, using Taylor series expansion and ignoring variations greater than order one. Formally, we need

$$\boldsymbol{y} - H\boldsymbol{x} = \boldsymbol{y} - \mathbf{H}(\boldsymbol{x}-\boldsymbol{x}_b) - H\boldsymbol{x}_b + \mathcal{O}||\boldsymbol{x}-\boldsymbol{x}_b||_2^2 \qquad (4.5)$$

for all values of $\boldsymbol{x}$ used in the analysis procedure and all trial values in the minimisation of $\mathcal{J}(\boldsymbol{x})$, where $\mathbf{H}$ is the Jacobian of H at point $\boldsymbol{x}_b$. Thus we need the difference between $H(\boldsymbol{x}) - H(\boldsymbol{x}_b)$ and $\mathbf{H}(\boldsymbol{x}-\boldsymbol{x}_b)$ to be smaller than the observation errors (defined by matrix O) for all model state perturbations $\boldsymbol{x}-\boldsymbol{x}_b$ with size and structure consistent with both typical background errors and amplitudes of the analysis increments $\boldsymbol{x}_a - \boldsymbol{x}_b$. The incremental 3DVAR approach presently used at the UK Met. Office, applies a low resolution correction to a high resolution background. The low resolution incremental problem is a inner loop of the minimisation procedure and is solved for each update of the full high resolution problem [32]. The method is described by minimising the objective functional $J(\boldsymbol{w})$, where

$$\mathcal{J}(\boldsymbol{w}) = \boldsymbol{w}^T B^{-1} \boldsymbol{w} + (\boldsymbol{d} - \mathbf{H}\boldsymbol{w})^T O^{-1} (\boldsymbol{d} - \mathbf{H}\boldsymbol{w}) \tag{4.6}$$

and

- the variable increment is $\boldsymbol{w} = \boldsymbol{x} - \boldsymbol{x}_b$,
- $\boldsymbol{d} = \boldsymbol{y} - H(\boldsymbol{x}_b)$ are the observation increments. $\boldsymbol{y}$ are the full observation values. The nonlinear function H is being used to interpolate the background field to the position of the data points. The linearisation of $H(\boldsymbol{x})$ gives $H(\boldsymbol{x}) = H(\boldsymbol{x}_b) + \mathbf{H}\boldsymbol{w}$.

The variables in which the objective function (4.6) is minimised are called control variables. The cornerstone of this thesis is to examine 'different changes between control variables' in light of dynamical theory established in Chapters 2 and 3. In the following section we describe this phrase within the context of data assimilation.

4.3 Change between Control Variables

3DVAR incremental formulation involves the variational minimisation of a functional (4.6). We wish to consider a transformation of the original variable increments into control variables $\boldsymbol{\tau}$ by applying a linear operator on $\boldsymbol{w}$. i.e. $\boldsymbol{\tau} = T(\boldsymbol{w})$. We assume that T can be represented by a nonsingular matrix with an inverse $U = T^{-1}$. Under this control transform a new objective function is minimised:

$$\mathcal{J}(\boldsymbol{\tau}) = \boldsymbol{\tau}^T B_\tau^{-1} \boldsymbol{\tau} + (\boldsymbol{d} - \mathbf{H}(U\boldsymbol{\tau}))^T O_\tau^{-1} (\boldsymbol{d} - \mathbf{H}(U\boldsymbol{\tau})). \tag{4.7}$$

with $\boldsymbol{d} = \boldsymbol{y} - H(\boldsymbol{x}_b)$ and $B_\tau^{-1} = U^T B U$.

The error covariance matrices, B and O, are not readily known in real life situations. The background term $\boldsymbol{\tau}^T B_\tau^{-1} \boldsymbol{\tau}$ is the dominant term in equation (4.7). The matrix B_τ^{-1} greatly affects the solution of (4.7).

The change to a new set of control variables is intended to precondition the problem by a similarity transformation so that the control variables become uncorrelated with each other with B_τ^{-1} and B_τ being turned into diagonal matrices. (For this linear problem the diagonalisation can be achieved by applying a Cholesky decomposition to the symmetric positive definite matrix B^{-1} to obtain a lower triangular matrix L such that $B^{-1} = LL^T$. The diagonalisation of B_τ^{-1} is given by a singular value decomposition of L. This is not done because this technique is prohibitively expensive to achieve for the resolution needed. In practice we need to choose control variables that approximate this diagonalisation.)

The present statistical method assumes that the control variables are uncorrelated with one another. In practice they are not. In making them more uncorrelated, an iterative minimisation procedure applied to (4.7) will take fewer iterations to get an approximation to the solution of required tolerance.

In this thesis we take an dynamical approach to the problem and consider the dynamical properties that control variables should have. One such property is to separate the unbalanced and balanced parts of the flow. In unpublished work [42] Tim Payne, by assuming that the balanced and unbalanced parts of the flow move on different time-scales, shows that the balanced and unbalanced parts of the flow are uncorrelated with one another. We attempt to find a set of control variables which distinguishes between balanced and unbalanced parts more efficiently. In order to do this we establish in the next section a general framework for describing changes between sets of control variables.

4.4 A General Framework to Examine the Change in Control Variables

Consider the transform T as a series of matrix operations $T_1, T_2, \cdots T_m$, which when applied to the original variables stored in a vector $\boldsymbol{x}^0$ of size $n \times 1$, produces a vector $\boldsymbol{y}^0$ of size $n \times 1$ which contains the control variables. This is given by

$$\boldsymbol{y}^0 = T_m \cdots T_2 T_1 \boldsymbol{x}^0. \tag{4.8}$$

We denote the reverse transformation U by a series of operations $U_1, U_2, \cdots U_m$ from $\boldsymbol{y}^0$ to $\boldsymbol{x}^0$, with

$$\boldsymbol{x}^0 = U_m \cdots U_2 U_1 \boldsymbol{y}^0. \tag{4.9}$$

Unlike the full transforms U and T, the matrix operations need not be non-singular and are allowed to project or restrict the variables concerned.

We can consider each operation in turn setting

$$\begin{aligned} \boldsymbol{x}^i &= T_i \boldsymbol{x}^{i-1}, \\ \boldsymbol{y}^i &= U_i \boldsymbol{y}^{i-1}, \qquad \text{for} \quad i = 1, \cdots, m, \end{aligned} \tag{4.10}$$

with

$$\boldsymbol{y}^0 = \boldsymbol{x}^m,$$

$$\boldsymbol{x}^0 = \boldsymbol{y}^m. \tag{4.11}$$

Thus we can relate any $\boldsymbol{x}^i$ with any vector $\boldsymbol{y}^j$ by

$$\boldsymbol{x}^i = \left(\prod_{l=1}^{l=i} T_l\right)\left(\prod_{k=j+1}^{k=m} U_k\right)\boldsymbol{y}^j, \tag{4.12}$$

for $i = 1, \cdots, m$ and $j = 0, \cdots, m-1$.

In particular

$$\boldsymbol{x}^1 = T_1\boldsymbol{x}^0 = T_1 U_m \boldsymbol{y}^{m-1}. \tag{4.13}$$

and if $\boldsymbol{x}^1 = \boldsymbol{y}^{m-1}, \forall \boldsymbol{y}^{m-1}$, then T_1 is the inverse of U_m.

We now formalise our definitions T_1, U_m in a way similar to a treatment of control variables given by unpublished work of Tim Payne [42]. His description is in terms of general infinite dimensional operators. We present the change between control variables as a finite dimensional problem. Thus the linear differential operators presented in the control variable transforms in Sections 4.4.1, 4.5, 4.7, are finite dimensional approximations to the true analytic differential operators.

Let us denote the model variable increments as

$$\boldsymbol{x}^0 = \begin{pmatrix} u' \\ v' \\ h' \end{pmatrix}, \tag{4.14}$$

with the new set of control variables denoted as

$$\boldsymbol{x}^1 = \boldsymbol{y}^{m-1} = \begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}. \tag{4.15}$$

Each of the model variables, u', v', h', and control variables y_1', y_2', y_3' are considered to be discrete and represented by vectors of size $s \times 1$.

We set three projections from the control variable space to the model space which we define as

$$U_m^i : \Re^s \to \Re^{3s} \qquad \text{for } i = 1, \cdots, 3 \tag{4.16}$$

such that

$$\boldsymbol{x}^0 = U_m^1 y_1' + U_m^2 y_2' + U_m^3 y_3'. \tag{4.17}$$

The balance part of the flow is spanned by U_m^1.

To obtain the control variables y_i a dual basis is needed, defined as

$$T_1^i : \Re^{3s} \to \Re^s \qquad \text{for } i = 1, 2, 3, \tag{4.18}$$

such that

$$\begin{aligned} T_1^i U_m^j &= 0 \qquad && \text{if } i \neq j, \\ T_1^i U_m^j &= I \qquad && \text{if } i = j. \end{aligned}$$

Thus, the full matrix differential operators T_1, U_m are given by

$$T_1 = \begin{pmatrix} T_1^1 \\ T_1^2 \\ T_1^3 \end{pmatrix} \qquad U_m = \left(U_m^1, \; U_m^2, \; U_m^3 \right). \tag{4.19}$$

Our interest lies in the choice of T_1 and U_m within the setting of a nonlinear SWE model on a rotating sphere. At this point it is instructive to consider what properties an ideal set of control variables should have from a dynamical perspective. Let us consider data assimilation that is applied to only the balanced part of the flow. This would be most easily applicable for flows where the predominant balanced part of the flow is notably slower that the inertio-gravity waves. If the data assimilation is applied so that the resulting full adjusted fields are on the part of the phase space of the SWEs that evolves on a slow timescale, we would no longer need an initialisation step between the data assimilation and the running of the forecast model. An initialisation step, such as normal mode initialisation, typically adjusts the fields so that the first and second local derivatives of the divergence are zero [53],[54], limiting the generation of spurious inertio-gravity waves in the subsequent forecast. Also having to deal with fewer variables would substantially reduce the computational demands that data assimilation poses. The underlying theory describing balance needs to be developed before this becomes a viable option. In particular the treatment of boundary conditions (A.White, personal communication) and understanding the balance in situations where balanced and unbalanced parts of the flow have similar timescales needs to be improved.

Because our understanding of balance is limited we consider the total number of control variables equal to the number of dynamical variables of our atmospheric model. For our SWE model this means that we have 3 control variables, one that represents the balanced part of the flow and the other two representing the unbalanced contributions consistent with the dynamical nature of the SWE model. In the following example we describe a choice of U_m, T_1 transforms which are analogous to those used by the UK Met. Office, where the rotating wind is used to describe the balanced part of the flow.

4.4.1 Example

The transform T_1 changes the height and wind increments, h' $\mathbf{u}'$, into a different set of physical variables in which the data assimilation is performed. In addition to having a set of full fields in h, $\mathbf{u}$ we also have a set of background (base) fields denoted by h_{base}, $\mathbf{u}_{base}$. We subtract the same linearisation state from both sets of full fields to give the respective perturbed fields $\mathbf{u}_p^*$, h_p^*, $\mathbf{u}_{base}^*$, h_{base}^* where

$$\begin{aligned}
\mathbf{u}_p^* &= \mathbf{u} - \overline{\mathbf{u}}, \\
h_p^* &= h - \overline{h}, \\
\mathbf{u}_{base}^* &= \mathbf{u} - \overline{\mathbf{u}}, \\
h_{base}^* &= h_{base} - \overline{h}_{base}. \qquad (4.20)
\end{aligned}$$

The height and wind increments are defined by the difference between base height and wind perturbations h_{base}^*, $\mathbf{u}_{base}^*$, and the full height and wind perturbations h_p, $\mathbf{u}_p$. They are given by

$$\begin{aligned}
\mathbf{u}' &= \mathbf{u}_p^* - \mathbf{u}_{base}^*, \\
h' &= h_p^* - h_{base}^*. \qquad (4.21)
\end{aligned}$$

The properties of these linearisation states is given in Section 2.4 while the means in which they are calcluated is left to Section 5.3.2.

The change in control variables transforms the height and wind increments into the streamfunction ψ', velocity potential χ' and unbalanced height h'_{ub}. The full streamfunction increment is considered balanced. The unbalanced height is the difference between the full height increment and the balanced height increment, obtained from the streamfunction using the linear balance equation.

We set $\mathbf{u}' = u'\mathbf{i} + v'\mathbf{j}$ with $\mathbf{i}, \mathbf{j}$ being orthonormal vectors on the surface of the sphere and $\mathbf{k}$ being a unit vector pointing radially away from the centre of the sphere. We write

$$T_1^1 : \quad \psi' = \nabla^{-2}\left(\mathbf{k} \cdot \nabla \times \mathbf{u}'\right), \qquad (4.22)$$

$$\begin{aligned} T_1^2 : \; h'_{ub} &= h' - \frac{1}{g}\nabla^{-2}\nabla\cdot f\nabla\psi', \\ &= h' - \frac{1}{g}\nabla^{-2}\nabla\cdot f\nabla\left(\nabla^{-2}\left(\mathbf{k}\cdot\nabla\times\mathbf{u}'\right)\right) \\ T_1^3 : \; \chi' &= \nabla^{-2}\nabla\cdot\mathbf{u}', \end{aligned} \tag{4.23}$$

with ∇^{-2} being a discrete approximation to the inverse laplacian operator. For T_1^1 , the inverse laplacian ∇^{-2} as used to represent the Poisson problem, $\nabla^2\psi' = (\mathbf{k}\cdot\nabla\times\mathbf{u}')$. The appropriate boundary conditions to the Poisson equation are described in Section 5.5.1.

The inverse transform U_m is

$$\begin{aligned} U_m^1 : \; (\mathbf{u}_1'^T, h_1') &= \left((\mathbf{k}\times\nabla\psi), \frac{1}{g}\nabla^{-2}\nabla\cdot f\nabla\psi'\right) \\ U_m^2 : \; (\mathbf{u}_2'^T, h_2') &= (0, h'_{ub}) \\ U_m^3 : \; (\mathbf{u}_3'^T, h_3') &= (\nabla\chi', 0) \end{aligned} \tag{4.24}$$

with

$$\begin{pmatrix} \mathbf{u}' \\ \\ h' \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1' \\ \\ h_1' \end{pmatrix} + \begin{pmatrix} \mathbf{u}_2' \\ \\ h_2' \end{pmatrix} + \begin{pmatrix} \mathbf{u}_3' \\ \\ h_3' \end{pmatrix} \tag{4.25}$$

The symbol ∇ is representing a discrete approximation to the gradient operator.

The American National Meteorological Centre [41] was the first to use a change of physical control variables which approximates the balanced part of the flow as rotational and the unbalanced flow as divergent and a departure from linear balance. Their set of control variables are slightly different; instead of choosing the streamfunction and velocity potential they use the vorticity and divergence instead.

It is clear that the change in control variable is kinematic in nature where the $2D$ Helmholtz decomposition is being used to separate the rotational and divergent aspects of the wind and represent them as the scalar potential fields ψ' and χ'. The evaluation of the balanced height h_1' is dependent purely on the rotational part of the flow with the unbalanced height h_2' holding the rest of the height. Thus h_3' is zero; we assume that the velocity potential increment does not contribute to the unbalanced height.

The choice of control variables which are constrained by ψ are not unique; we could be perverse and choose $y_1 = h_b$, by calculating the streamfunction and then applying the linear balance equation. Thus,

$$T_1^1 : \quad h_b' \quad = \quad \frac{1}{g}\nabla^{-2}\nabla \cdot f\nabla\nabla^{-2}\left(\mathbf{k} \cdot \nabla \times \mathbf{u}'\right) \tag{4.26}$$

and

$$U_m^1 : \quad (\mathbf{u}_1', h_1) \quad = \quad \left(\mathbf{k} \times \nabla\left(\left(\nabla \cdot f\nabla\right)^{-1}\nabla^2 g\ h_b\right), h_b\right), \tag{4.27}$$

with the other variables and operators the same as before.

There are two reasons why this is not used. In data assimilation the U transform is applied every time during the minimisation procedure. The procedure to calculate the winds from the balanced height is computationally more costly taking more cpu time to evaluate. Also, as noted in Section 3.5, it is noticeably less accurate compared to using just the streamfunction because we need to use the reverse linear balance equation.

There are additional issues to consider when using LBE in a data assimilation context. The dynamical behaviour of the winds in the tropics is not captured by the LBE. Thus observations in the tropics are going to be inconsistent with the dynamical behaviour of LBE. The observations for the winds are also comprehensive on horizontal surfaces, while their vertical structure is less well known. Meanwhile, height/ pressure observations are more comprehensive in the vertical direction. This is a secondary reason why (4.22) is prefered to (4.26).

4.5 Transformations to and from Control Variables Based on Potential vorticity

In meteorology potential vorticity is considered to be a balanced quantity due to being able to capture key dynamical features in the atmospheric flow such as frontogenesis, cyclogenesis and the general circulation . Cullen [10] proposes a potential vorticity based change in control variables for the three-dimensional incompressible Navier Stokes equations on an f-plane. This first variable involves an inversion of potential vorticity to produce a balanced height. The second variable describes the ageostrophic part of the horizontally non-divergent wind and pressure and is defined in terms of a streamfunction variable. The third variable, the horizontal divergence, completes the description. This change in variables requires a constant f and a linearisation state in potential temperature that does not vary in the vertical.

In this thesis we consider the equivalent change in variable for the shallow water equations on a sphere. We represent the potential vorticity with a linearisation of potential vorticity (2.64) about time-varying height and wind fields that are a function of latitude only. We choose the linearisation state as a compromise between accuracy and numerical complexity. A constant linearisation state would be far too inaccurate; a linearisation state dependent on both longitude and latitude is too complex for an initial investigation.

In order to use potential vorticity to define the balanced part of the flow, we need an additional balance constraint. We choose this to be the LBE. This potential vorticity inversion is a somewhat cruder version of McIntyre and Norton's [36] first-order direct inversion, which uses a Charney balance condition instead of the LBE. The LBE is used in the inversion, because this allows a direct comparison to be made between using the rotational wind to define the balance and using potential vorticity. We do this by solving the equations representing the linear balance and linear potential vorticity increments simultaneously for ψ_b and h_b, as

$$\nabla \cdot f \nabla \psi_b - g \nabla^2 h_b = 0, \tag{4.28}$$

$$\nabla^2 \psi_b - \overline{q} h_b = \mathbf{k} \cdot (\nabla \times \mathbf{u}') - \overline{q} h'. \tag{4.29}$$

The height increment and wind increments and the potential vorticity linearisation state are known before application of this potential vorticity inversion. The height and winds increments, h', $\mathbf{u}'$, are defined in equation (4.21) and the potential vorticity linearisation state $\overline{q}$ is given by (2.63).

From this coupled system we obtain a 'balanced' height h_b and a 'balanced' wind, defined by $u_b = \mathbf{k} \times \nabla \psi_b$. The 'balanced' wind increment is non-divergent, and approximates the full rotational wind increment for high Burger number regimes. The rest of the rotational wind is described as having no potential vorticity increment and conserving a departure from linear balance. This can be obtained in one of two ways, either by subtracting the balanced wind and height from the full rotational wind and height, or by explicitly solving the simultaneous system

$$\nabla \cdot f \nabla \psi_{ub} - g \nabla^2 h_{ub} = \nabla \cdot f \nabla \psi' - \nabla^2 h' \tag{4.30}$$

$$\nabla^2 \psi_{ub} - \overline{q} h_{ub} = 0, \tag{4.31}$$

where the unbalanced rotational wind is defined to be $u_{rub} = \mathbf{k} \times \nabla \psi_{ub}$ and the unbalanced height is denoted by h_{ub}. The equivalence of the two methods to calculate the unbalanced height and unbalanced rotational wind is readily seen by adding equation (4.28) to (4.30) and (4.29) to (4.31), to give

$$\begin{aligned}
\nabla \cdot f\nabla\left(\psi_b + \psi_{ub}\right) - g\nabla^2\left(h_b + h_{ub}\right) &= \nabla \cdot f\nabla\psi' - \nabla^2 h' && (4.32)\\
\nabla^2\left(\psi_b + \psi_{ub}\right) - \overline{q}\left(h_b + h_{ub}\right) &= \mathbf{k}\cdot\left(\nabla \times \mathbf{u}'\right) - \overline{q}h', && (4.33)
\end{aligned}$$

The third variable contains the remaining information, namely the divergent part of the wind and is stored in the velocity potential.

The above description produces five different variables, ψ_b', ψ_{ub}', χ', h_b', h_{ub}', which together give the original height and wind fields. From these five variables we choose three control variables, from which the ignored part of the height and wind fields is easily recovered. To mimic the dynamic behaviour of the shallow water equations we choose the control variables so that one is balanced and two others are unbalanced.

There are four possible choices to obtain such a control set. Each method we now discuss.

4.5.1 Method 1: ψ_b', ψ_{ub}', χ'

$$\begin{aligned}
T_1^1: \quad \psi_b' &= \left[\nabla \cdot f\nabla - g\nabla^2\left(\frac{1}{\overline{q}}\nabla^2\right)\right]^{-1}\\
&\qquad \left(-g\nabla^2\left(\frac{1}{\overline{q}}\left(\mathbf{k}\cdot\nabla\times\mathbf{u}'\right)\right) + g\nabla^2 h'\right) && (4.34)\\
T_1^2: \quad \psi_{ub}' &= \left[\nabla \cdot f\nabla - g\nabla^2\left(\frac{1}{\overline{q}}\nabla^2\right)\right]^{-1}\\
&\qquad \left(\nabla \cdot f\nabla\nabla^{-2}\left(\mathbf{k}\cdot\nabla\times\mathbf{u}'\right) - g\nabla^2 h'\right) && (4.35)\\
T_1^3: \quad \chi' &= \nabla^{-2}\nabla\cdot\mathbf{u}' && (4.36)
\end{aligned}$$

The transformation T_1^1 into ψ_b' is obtained simply by multiplying (4.28) by $-g\nabla^2\frac{1}{\overline{q}}$

and substituting $g\nabla^2 h_b$ from (4.28) with $\nabla \cdot f\nabla\psi_b$. Likewise T_1^2 is given by substituting into (4.30) $\frac{1}{\overline{q}}\nabla^2\psi$ using (4.31) for h_b.

$$
\begin{aligned}
U_m^1 : \quad (\mathbf{u}_1'^T, h_1) &= (\mathbf{k} \times \nabla\psi_b', \frac{1}{g}\nabla^{-2}\nabla \cdot f\nabla\psi_b') && (4.37)\\
U_m^2 : \quad (\mathbf{u}_2'^T, h_2) &= (\mathbf{k} \times \nabla\psi_{ub}', \frac{1}{\overline{q}}\nabla^2\psi_{ub}') && (4.38)\\
U_m^3 : \quad (\mathbf{u}_3'^T, h_3) &= (\nabla\chi', 0) && (4.39)
\end{aligned}
$$

The U_m transform is self-evident and recovers the full height and wind with

$$
(\mathbf{u}'^T, h') = (\mathbf{u}_1'^T, h_1) + (\mathbf{u}_2'^T, h_2) + (\mathbf{u}_3'^T, h_3). \tag{4.40}
$$

The winds are obtained by a generalisation of Helmholtz decomposition with h_1 defined by the linear balance relation applied to ψ_b and h_2 defined by a zero potential vorticity increment.

There are various forms in which the projections T_1^1, T_1^2 can be written. The equations (4.34), (4.35) are the simplest versions. They are difficult to solve as they have singularities around the equator due to $\overline{q} \to 0$ as $\theta \to 0$. This is overcome by assuming $f\nabla\psi_b = g\nabla h_b$ instead of (4.28), and $f\nabla\psi_b - g\nabla h_b = f\nabla\psi' - g\nabla h'$ such that $\mathbf{k} \cdot (\nabla \times f\nabla\psi_b) = 0$, $\mathbf{k} \cdot (\nabla \times f\nabla\psi') = 0$ instead of (4.30) to give

$$
\begin{aligned}
T_1^1 : \quad \psi_b' &= \left[\overline{q}\nabla^4 - \left[\nabla^2\overline{q} + \frac{\overline{q}^2 f}{g}\right]\nabla^2 - \nabla\left(\frac{\overline{q}^2 f}{g}\right)\cdot\nabla\right]^{-1} \\
&\qquad \left(\left[\overline{q}\nabla^2 - \left(\nabla^2\overline{q}\right)\right](\mathbf{k}\cdot(\nabla\times\mathbf{u}')) - \left[\overline{q}^2\nabla^2 + \nabla\overline{q}^2\cdot\nabla\right]h'\right) && (4.41)\\
T_1^2 : \quad \psi_{ub}' &= \left[\overline{q}\nabla^4 - \left[\nabla^2\overline{q} + \frac{\overline{q}^2 f}{g}\right]\nabla^2 - \nabla\left(\frac{\overline{q}^2 f}{g}\right)\cdot\nabla\right]^{-1} \\
&\qquad \left(\left[\frac{\overline{q}^2 f}{g} - \nabla\left(\frac{\overline{q}^2 f}{g}\right)\cdot\nabla\nabla^{-2}\right](\mathbf{k}\cdot(\nabla\times\mathbf{u}')) + \right. \\
&\qquad\qquad \left.\left[\nabla\overline{q}^2\cdot\nabla + \overline{q}^2\nabla^2\right]h'\right). && (4.42)
\end{aligned}
$$

Equation (4.41) is given by first applying the gradient operator to (4.29). The terms

$\nabla\left(\overline{q}h_b\right)$ and $\nabla\left(\overline{q}h'\right)$ are expanded to $\overline{q}\nabla h_b + h_b\nabla\overline{q}$ and $\overline{q}\nabla h' + h'\nabla\overline{q}$ such that

$$\nabla\nabla^2\psi_b - h_b\nabla\overline{q} - \overline{q}\nabla h_b = \nabla\nabla^2\psi' - h'\nabla\overline{q} - \overline{q}\nabla h'. \tag{4.43}$$

The terms ∇h_b and h_b in equation (4.43) are substituted for $\frac{f}{g}\nabla\psi_b$ and $\left(-\nabla^2\psi_b + \nabla^2\psi' - \overline{q}h'\right)/\overline{q}$, respectively, separating the unknown terms ψ_b to the left side of the equation and the known terms h', ψ' to the right. We apply $\nabla\cdot(\overline{q}$ to the the result to give

$$\begin{aligned}
&\nabla\cdot\left(\overline{q}\nabla\nabla^2\psi_b\right) - \nabla\cdot\left(\nabla\overline{q}\left(\nabla^2\psi_b\right)\right) - \nabla\cdot\left(\frac{\overline{q}f}{g}\nabla\psi_b\right) = \\
&\nabla\cdot\left(\overline{q}\nabla\nabla^2\psi'\right) - \nabla\cdot\left(h'\overline{q}\nabla\overline{q'}\right) - \nabla\cdot\left(\overline{q}^2\nabla h'\right) \\
&+\nabla\cdot\left(\left(-\nabla^2\psi' + \overline{q}h'\right)\nabla\overline{q}\right).
\end{aligned} \tag{4.44}$$

We now use the vector identity for generic vector $\mathbf{A}$ and generic scalar α:

$$\nabla\cdot\alpha\mathbf{A} = \nabla\alpha\cdot\mathbf{A} + \alpha\nabla\cdot\mathbf{A}, \tag{4.45}$$

to expand (4.44) to

$$\begin{aligned}
&\overline{q}\nabla^4\psi_b + \nabla\overline{q}\cdot\nabla\nabla^2\psi_b - \nabla\overline{q}\cdot\nabla\nabla^2\psi_b - \nabla^2\overline{q}\nabla^2\psi_b - \nabla\left(\frac{\overline{q}^2 f}{g}\right)\cdot\nabla\psi_b - \frac{\overline{q}^2 f}{g}\nabla^2\psi_b = \\
&\overline{q}\nabla^4\psi' + \nabla\overline{q}\cdot\nabla\nabla^2\psi' - \nabla\cdot\left(\overline{q}\nabla\overline{q}\right)h' - \overline{q}\nabla\overline{q}\cdot\nabla h' - \nabla\overline{q}^2\cdot\nabla h' - \overline{q}^2\nabla^2 h' \\
&-\nabla\overline{q}\cdot\nabla\nabla^2\psi' - \nabla^2\overline{q}\nabla^2\psi' + \nabla\cdot\left(\overline{q}\nabla\overline{q}\right)h' + \overline{q}\nabla\overline{q}\cdot\nabla h'
\end{aligned} \tag{4.46}$$

which reduces to (4.41). Equation (4.42) is obtained through a similar procedure.

4.5.2 Method 2: h'_b, h'_{ub}, χ'

The transformation into velocity potential and back is the same as in Choice 1 and is given by (4.36), (4.39). Also, the following expressions for T_1^1, T_1^2 are not unique.

A difficulty lies in the calculation of the balanced height. Within equations (4.28), (4.29) not only the second order derivatives in ψ_b need to be substituted but also the first order derivatives as well. We present an approach to alleviate this problem. We calculate the balanced streamfunction and then use the linear balance relation to obtain the balanced height.

$$T_1^1 : \quad h_b' = \frac{1}{g}\nabla^{-2}\nabla \cdot f\nabla\psi_b' \tag{4.47}$$

where ψ_b' is given by either (4.34).

Similarly, we may obtain h_{ub} from ψ_{ub}, using the fact that the linearised potential vorticity increment conserved by these two variables is zero, giving

$$h_{ub}' = \frac{1}{\bar{q}}\nabla^2\psi_{ub}' \tag{4.48}$$

where ψ_{ub}' is given by (4.35).

The associated U_m^1 transformation uses reverse linear balance to derive the balanced part of the rotational wind.

$$\begin{aligned} U_m^1 : \quad (\mathbf{u}_1'^T, h_1) &= (\mathbf{k} \times \nabla \left(\nabla \cdot f\nabla\right)^{-1} \nabla^2 g h_b', h_b') & (4.49) \\ U_m^2 : \quad (\mathbf{u}_2'^T, h_2) &= (\mathbf{k} \times \nabla \left(\nabla^{-2} \left(\overline{q} h_{ub}'\right)\right), h_{ub}') & (4.50) \end{aligned}$$

Equation (4.49) shows that in order to calculate the balanced wind from the balanced height, the balanced streamfunction needs to be evaluated using the reverse linear balance equation. In Section 3.5 we described a number of studies in which showed that the solution to the reverse linear balance equation is very sensitive to the height fields about the equator. A similar problem exists for equations (4.49), (4.50).

The two remaining methods are the control variables sets h'_b, ψ'_{ub}, χ' and ψ'_b, h'_{ub}, χ'. The equations relating to these control variables are given above.

In a shallow water context on a doubly periodic f-plane we showed in Section 3.8 that the slow mode relating to the zero eigenvalue is in geostrophic balance and is described by a linearised potential vorticity increment. The other two variables, the divergence and departure from geostrophic balance, are linear combinations of the unbalanced eigenvalues of the system. On this f-plane the balanced variable is independent from the unbalanced part. Ideally, we would wish the unbalanced components to represent the eigenvalues of the unbalanced part. Unfortunately this in not the case.

In summary, the changes in control variable described in this section do have certain difficulties: namely the problem with the linearisation of the potential vorticity going to zero at the equator and the problem of constraining the control transforms (4.41) by $\mathbf{k} \cdot \nabla \times f \nabla \psi_b = 0$ and (4.42) by $\mathbf{k} \cdot \nabla \times f \nabla \psi_{ub} = 0$. Balanced and unbalanced ψ variables seem better than the corresponding balanced and unbalanced height, due to the need to solve the reverse LBE equation in the corresponding U transform. For these reasons we solve for the balanced variables in their original formulation given in (4.28), (4.29).

4.5.3 Conditions for Solving the Potential Vorticity-based Change in Control Variables

The boundary conditions for solving the simultaneous system (4.28), (4.29) comes from the consideration of the solution of the linear balance equation and the existence and uniqueness conditions, necessary for the solution of a Poisson equation on a sphere. We are considering solutions over the hemisphere. This is done by solving over the sphere and making the right-hand side of (4.29) anti-symmetric about the equator. This enforces ψ_b to be antisymmetric and h_b to be symmetric about the equator. This is equivalent to setting

$$\begin{aligned} \frac{\partial h_b}{\partial \theta}(0,\lambda) &= 0, \\ \psi_b(0,\lambda) &= 0. \end{aligned} \tag{4.51}$$

The standard boundary conditions used to solve Poisson equations on the sphere apply at the poles. We set Neumann boundary conditions there to make the ψ_b, h_b single valued at these points.

The final condition needed is for the balanced height to conserve the mass present in the height increment. In each of our datasets the total mass, represented by the summation of h_f, is the same. Since we subtract the same linearisation state $\overline{h}$ from all the full heights, the perturbation fields, all have the same mass. Similarly, since we are dealing with differences to these perturbations, the total mass of these increments should be zero, with

$$\int h_b dS = 0. \tag{4.52}$$

4.6 Divergence tendency within control variables

We wish to identify the divergence tendency of the control variables to ascertain how well the balanced and unbalanced parts are represented by the balanced and unbalanced control variables. A good representation of balance gives divergence tendencies that are small.

Let us first consider control variable increments with linearisation states $\overline{\mathbf{u}} = 0$ and $\overline{h} = H$, where H is constant. If we consider the linearisation of the shallow water equations about such states and introduce only the balanced variables, from whatever method, into the respective divergence equation (2.50), we get a divergence tendency of zero. This is due to the divergence of advective term in the SWEs momentum equation (2.56) being comprised of squared perturbations. If we, however, consider the linearised divergence equation linearised about a time-varying state and apply the balanced variables, the divergence of advective terms of (2.37) remains. The divergence tendency of the balanced control variable is then given by

$$\begin{aligned}\frac{\partial \delta_r}{\partial t} &= -\nabla^2 \left[(\mathbf{v_r}' \cdot \overline{\mathbf{v}}') \right] + \mathbf{k} \cdot \nabla \times \left(\overline{\zeta} \mathbf{v_r}' + \zeta' \overline{v} \right), \quad \text{LB method} && (4.53)\\ \frac{\partial \delta_p}{\partial t} &= -\nabla^2 \left[(\mathbf{v_{pv}}' \cdot \overline{\mathbf{v}}') \right] + \mathbf{k} \cdot \nabla \times \left(\overline{\zeta} \mathbf{v_{pv}}' + \zeta' \overline{v} \right) \quad \text{PV method} && (4.54)\end{aligned}$$

where the LB method is the standard change in control variables described by (4.22) and the PV method is the name we give to the change in control variables based of potential vorticity inversion described by (4.29), (4.28) in Section 4.5, $\mathbf{v_r}$ is the rotational wind and $\mathbf{v_{pv}}$ is the balanced rotational wind derived using the PV method.

The advective term is not accounted for due to both control variable transforms using a linear balance relationship intrinsic in their definitions. In some data assimilation systems which use the LBE, the balanced advective parts of the flow are presented in some data assimilation schemes by using approximations to the divergence tendency as an additional observation [46]. Another way to account for the advective term is to consider a higher order potential vorticity inversion scheme. In McIntyre and Norton's paper [36] the 2nd order direct potential vorticity inversion by definition sets both the divergence tendency of the nonlinear divergence equation and the second order partial time derivative of the divergence to zero. By using a similar concept we derive in the next subsection a correction to the balanced wind and heights given by the LB and PV methods which sets the divergence tendency of the divergence equation linearised about a time-varying state to be zero. This method has the added benefit that the correction identifies balanced parts in unbalanced variables.

4.6.1 Approximation of Divergence Tendencies of Balanced Corrections to Unbalanced Variables

In the direct second-order potential vorticity inversion model in McIntyre and Norton's paper [36], one of the equations which close their dynamical system enforces the second partial time derivative of the divergence to be set to zero. We derive a similar equation which uses a balanced correction to find the divergent wind. This is done by first applying the local partial time operator to the divergence equation (2.61) linearised about time-varying linearisation states $\overline{h}, \overline{\mathbf{v}}, \overline{\zeta}$ as defined in Section 2.4, giving

$$\begin{aligned}\frac{\partial^2 \delta'}{\partial^2 t} - \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\frac{\partial \mathbf{v}'}{\partial t} + \frac{\partial\left(\overline{\zeta}+f\right)}{\partial t}\mathbf{v}' + \frac{\partial \zeta'}{\partial t}\overline{\mathbf{v}} + \zeta'\frac{\partial \overline{\mathbf{v}}}{\partial t}\right)&\\ +\nabla^2\left(g\frac{\partial h'}{\partial t} + \frac{\partial \mathbf{v}'}{\partial t}\cdot\overline{\mathbf{v}} + \frac{\partial \overline{\mathbf{v}}}{\partial t}\cdot\mathbf{v}'\right) &= 0, \quad (4.55)\end{aligned}$$

where

$$\frac{\partial \mathbf{v}'}{\partial t} = -\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right) - \left(\overline{\zeta}+f\right)\mathbf{k}\times\mathbf{v}' - \zeta'\mathbf{k}\times\overline{\mathbf{v}} - g\nabla h', \quad (4.56)$$

$$\frac{\partial \zeta'}{\partial t} = -\nabla\cdot\left(\left(\overline{\zeta}+f\right)\mathbf{v}' + \zeta'\overline{\mathbf{v}}\right), \quad (4.57)$$

$$\begin{aligned}
\frac{\partial h'}{\partial t} &= -\overline{\mathbf{v}}\cdot\nabla h' - \mathbf{v}'\cdot\nabla\overline{h} - h'\nabla\cdot\overline{\mathbf{v}} - \overline{h}\nabla\cdot\mathbf{v}'. && (4.58)\\
\frac{\partial \overline{\mathbf{v}}}{\partial t} &= -\nabla\left(\frac{\overline{\mathbf{v}}\cdot\overline{\mathbf{v}}}{2}\right) - \left(\overline{\zeta}+f\right)\mathbf{k}\times\overline{\mathbf{v}} - g\nabla\overline{h}, && (4.59)\\
\frac{\partial \overline{\zeta}}{\partial t} &= -\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right), && (4.60)\\
\frac{\partial \overline{h}}{\partial t} &= -\overline{\mathbf{v}}\cdot\nabla\overline{h} - \overline{h}\nabla\cdot\overline{\mathbf{v}}. && (4.61)
\end{aligned}$$

Incorporating (4.56) - (4.61) into (4.55) and setting the second partial time derivative of the divergence to zero leaves on simplifying

$$\begin{aligned}
&\mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right)\right)\right) + \left(\overline{\zeta}+f\right)^2\nabla\cdot\mathbf{v}' + \nabla\left(\overline{\zeta}+f\right)^2\cdot\mathbf{v}'\\
&\quad+\mathbf{k}\cdot\nabla\times\left[\left(\overline{\zeta}+f\right)\left(\zeta'\mathbf{k}\times\overline{\mathbf{v}}\right)\right] + g\mathbf{k}\cdot\nabla\times\left[\left(\overline{\zeta}+f\right)\nabla h'\right]\\
&\quad+\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right)\right)\mathbf{v}'\right)\\
&\quad+\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\mathbf{v}'+\zeta'\overline{\mathbf{v}}\right)\right)\overline{\mathbf{v}}\right)\\
&\quad+\mathbf{k}\cdot\nabla\times\left(\zeta'\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right)\right)\right)\\
&\quad-g\nabla^2\left(\overline{\mathbf{v}}\cdot\nabla h' + \mathbf{v}'\cdot\nabla\overline{h} + h'\nabla\cdot\overline{\mathbf{v}} + \overline{h}\nabla\cdot\mathbf{v}'\right)\\
&\quad-\nabla^2\left(\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right) + \left(\overline{\zeta}+f\right)\mathbf{k}\times\mathbf{v}' + \zeta'\mathbf{k}\times\overline{\mathbf{v}} + g\nabla h'\right)\cdot\overline{\mathbf{v}}\right)\\
&\quad-\nabla^2\left(\left(\nabla\left(\frac{\overline{\mathbf{v}}\cdot\overline{\mathbf{v}}}{2}\right) + \left(\overline{\zeta}+f\right)\mathbf{k}\times\overline{\mathbf{v}} + g\nabla\overline{h}\right)\cdot\mathbf{v}'\right) = 0. \qquad (4.62)
\end{aligned}$$

We now apply a scale analysis to equation (4.62) and assume that the flow is characterised in the mid-latitudes with: the characteristic length scales $L \approx 10^6 m$, the characteristic height linearisation state $\overline{H} \approx 10^3 m$, the characteristic velocity linearisation state $\overline{U} \approx 10^1 ms^{-1}$, the height increments $H \approx 10^2 m$, the wind increments $U \approx 10^1 ms^{-1}$, $f \approx 10^{-4} s^{-1}$ and $g \approx 10$. In this scaling the characteristic vorticity state is approximately $10^{-5} ms^{-1}$ and is smaller than the Coriolis parameter at $45°\theta$. The results from the scale analysis are presented in Table 4.6.1. Two of the terms of the greatest order of magnitude, $\left(\overline{\zeta}+f\right)^2\nabla\cdot\mathbf{v}'$, $g\nabla^2\left(\overline{h}\nabla\cdot\mathbf{v}'\right)$ are increments involving divergence and are rewritten as $\left(\overline{\zeta}+f\right)^2\delta'$, $g\nabla^2\left(\overline{h}\delta'\right)$. We wish to make the divergence increment the variable to be solved for, so we keep those two terms on the left and place the rest on the right hand side.

This gives a modified Helmholtz equation to be solved of the form

$$\nabla^2\left(g\overline{h}\delta_b\right)-\kappa\left(g\overline{h}\delta_b\right)=\varpi \tag{4.63}$$

where κ is given by

$$\kappa=\frac{(\zeta+f)^2}{g\overline{h}} \tag{4.64}$$

and ϖ is equal to

$$\begin{aligned}
\varpi \quad = \quad & \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right)\right)\right)+\nabla\left(\overline{\zeta}+f\right)^2\cdot\mathbf{v}' \\
& +\mathbf{k}\cdot\nabla\times\left(\overline{\zeta}+f\right)\left(\zeta'\mathbf{k}\times\overline{\mathbf{v}}\right)+g\mathbf{k}\cdot\nabla\times\left(\overline{\zeta}+f\right)\nabla h' \\
& +\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right)\right)\mathbf{v}'\right) \\
& +\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\mathbf{v}'+\zeta'\overline{\mathbf{v}}\right)\right)\overline{\mathbf{v}}\right) \\
& +\mathbf{k}\cdot\nabla\times\left(\zeta'\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right)\right)\right) \\
& -g\nabla^2\left(\overline{\mathbf{v}}\cdot\nabla h'+\mathbf{v}'\cdot\nabla\overline{h}+h'\nabla\cdot\overline{\mathbf{v}}\right) \\
& -\nabla^2\left(\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right)+\left(\overline{\zeta}+f\right)\mathbf{k}\times\mathbf{v}'+\zeta'\mathbf{k}\times\overline{\mathbf{v}}+g\nabla h'\right)\cdot\overline{\mathbf{v}}\right) \\
& -\nabla^2\left(\left(\nabla\left(\frac{\overline{\mathbf{v}}\cdot\overline{\mathbf{v}}}{2}\right)+\left(\overline{\zeta}+f\right)\mathbf{k}\times\overline{\mathbf{v}}+g\nabla\overline{h}\right)\cdot\mathbf{v}'\right).
\end{aligned} \tag{4.65}$$

The terms $h_b, \mathbf{v_b}, \zeta_b$ are obtained either from the LB or PV method. The linearisation states are set to be in geostrophic balance. The solution of the modified Helmholtz equation gives $g\overline{h}\delta_b$. Since we know the value of $g\overline{h}$, the balanced divergence is easily obtained.

The PV or LB method give balanced height and balanced rotational winds. The modified Helmholtz equation gives a balanced divergence. The remaining balanced

Table 4.1: The scaling of various terms in equation (4.62) ,with the characteristic length scales $L \approx 10^3 m$, the characteristic height linearisation state $\overline{H} \approx 10^3 m$, the characteristic velocity linearisation state $\overline{U} \approx 10^1 ms^{-1}$, the height increments $H \approx 10^2 m$, the wind increments $U \approx 10^1 ms^{-1}$, $f \approx 10^{-4} s^{-1}$ and $g \approx 10$

Terms considered	Dimensional scaling	Size of term
$\left(\overline{\zeta}+f\right)^2 \nabla\cdot\mathbf{v}'$, $\nabla\left(\overline{\zeta}+f\right)^2\cdot\mathbf{v}'$	$\frac{f^2U}{L}$	10^{-13}
$g\nabla^2\left(\mathbf{v}'\cdot\nabla\overline{h}\right)$, $g\nabla^2\left(\overline{h}\nabla\cdot\mathbf{v}'\right)$	$\frac{g\overline{H}U}{L^3}$	10^{-13}
$\mathbf{k}\cdot\nabla\times\left[\left(\overline{\zeta}+f\right)\left(g\nabla h'\right)\right]$	$\frac{gfH}{L^2}$	10^{-14}
$\mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right)\right)\right)$	$\frac{fU\overline{U}}{L^2}$	10^{-14}
$\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left[\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right]\right)\mathbf{v}'\right)$	$\frac{fU\overline{U}}{L^2}$	10^{-14}
$\nabla^2\left(\left(\left(\overline{\zeta}+f\right)\mathbf{k}\times\mathbf{v}'\right)\cdot\overline{\mathbf{v}}\right)$	$\frac{fU\overline{U}}{L^2}$	10^{-14}
$\mathbf{k}\cdot\nabla\times\left[\left(\overline{\zeta}+f\right)\left(\zeta'\mathbf{k}\times\overline{\mathbf{v}}\right)\right]$	$\frac{fU\overline{U}}{L^2}$	10^{-14}
$\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\mathbf{v}'\right)\right)\overline{\mathbf{v}}\right)$	$\frac{fU\overline{U}}{L^2}$	10^{-14}
$\nabla^2\left(\left(g\nabla h'\right)\cdot\overline{\mathbf{v}}\right)$	$\frac{gH\overline{U}}{L^3}$	10^{-14}
$g\nabla^2\left(\overline{\mathbf{v}}\cdot\nabla h'\right)$, $g\nabla^2\left(h'\nabla\cdot\overline{\mathbf{v}}\right)$	$\frac{gH\overline{U}}{L^3}$	10^{-14}
$\mathbf{k}\cdot\nabla\times\left(\left(\nabla\cdot\left(\zeta'\overline{\mathbf{v}}\right)\right)\overline{\mathbf{v}}\right)$	$\frac{U\overline{U}^2}{L^3}$	10^{-15}
$\nabla^2\left(\left(\nabla\left(\overline{\mathbf{v}}\cdot\mathbf{v}'\right)\right)\cdot\overline{\mathbf{v}}\right)$	$\frac{U\overline{U}^2}{L^3}$	10^{-15}
$\nabla^2\left(\left(\nabla\left(\overline{\mathbf{v}}\cdot\overline{\mathbf{v}}\right)\right)\cdot\mathbf{v}'\right)$	$\frac{U\overline{U}^2}{L^3}$	10^{-15}
$\nabla^2\left(\left(\zeta'\mathbf{k}\times\overline{\mathbf{v}}\right)\cdot\overline{\mathbf{v}}\right)$	$\frac{U\overline{U}^2}{L^3}$	10^{-15}
$\nabla^2\left(\left(\left(\overline{\zeta}+f\right)\mathbf{k}\times\overline{\mathbf{v}}\right)\cdot\mathbf{v}'\right)$	$\frac{fU\overline{U}}{L^2}$	10^{-15}
$\mathbf{k}\cdot\nabla\times\left(\zeta'\left(\nabla\cdot\left(\left(\overline{\zeta}+f\right)\overline{\mathbf{v}}\right)\right)\right)$	$\frac{fU\overline{U}}{L^3}$	10^{-20}

part of the flow lies in the departure from linear balance. Suppose we consider the divergence equation, linearised about a time-varying linearisation state (2.61). We presume that the total balanced parts of the flow satisfies this equation and has a divergence tendency of zero.

Since we are considering a linearisation of the divergence equation, it is possible to separate the balanced part of the divergence equation that we know, from those parts that we do not. Thus the balanced contribution to the divergence equation is given by

$$\begin{aligned} -\mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\mathbf{v_1}+\zeta_1\overline{\mathbf{v}}\right)+\nabla^2\left(gh_1+\mathbf{v_1}\cdot\overline{\mathbf{v}}\right) &= \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\mathbf{v_2}+\zeta_2\overline{\mathbf{v}}\right) \\ &\quad -\nabla^2\left(gh_2+\mathbf{v_2}\cdot\overline{\mathbf{v}}\right) \qquad (4.66) \end{aligned}$$

where $\mathbf{v_1}$, $\mathbf{h_1}$ is the balanced velocity and height contribution the variable describing a departure from linear balance. The wind $\mathbf{v_2}$ is defined as

$$\mathbf{v_2}=\mathbf{v_{pv}}+\mathbf{v_{bd}}, \qquad h_2=h_{pv} \qquad (4.67)$$

or

$$\mathbf{v_2}=\mathbf{v_{lb}}+\mathbf{v_{bd}}, \qquad h_2=h_{lb}, \qquad (4.68)$$

where $\mathbf{v_{bd}}$ is the balanced divergent wind given by $\nabla\left(\nabla^{-2}\delta_b\right)$.

Equation (4.66) has two unknowns, a balanced height and a balanced velocity. An additional equation is needed to produce a closed system in which the two unknowns are able to be determined. As we have determined the balanced divergence, we assume that the balanced velocity obtained form a departure from linear balance is non-divergent and expressible in terms of a streamfunction ψ_1. In Section 4.5, in equations (4.30), (4.31) we assume that the departure from linear balance also lies in the kernel of the linearised potential vorticity increment, as we assumed in the control variables that unbalanced flow contained no linearised potential vorticity. Thus the closed dynamical system is given by

$$\begin{aligned}-\nabla\cdot\left(f+\overline{\zeta}\right)\nabla\psi_1-\mathbf{k}\cdot\nabla\times\left(\nabla^2\psi_1\overline{\mathbf{v}}\right)&\\ +g\nabla^2 h_1+\nabla^2\left(\overline{\mathbf{v}}\cdot\left(\mathbf{k}\times\nabla\psi_1\right)\right) &= \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\mathbf{v_2}+\zeta_2\overline{\mathbf{v}}\right)\\ &\quad -\nabla^2\left(gh_2+\mathbf{v_2}\cdot\overline{\mathbf{v}}\right) \qquad (4.69)\\ \nabla^2\psi_1-\overline{q}h_1 &= 0, \qquad (4.70)\end{aligned}$$

where (4.69) is (4.66) with the wind on the left hand side written in terms of a streamfunction and (4.70) is a restatement of (4.31).

We are now in a position to evaluate the divergence tendency of the balanced corrections. The divergence tendency $\frac{\partial\delta_{bd}}{\partial t}$ is given by the balanced divergence

$$\frac{\partial\delta_{bd}}{\partial t}=+\mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\mathbf{v_{bd}}+\left(\mathbf{k}\times\mathbf{v_{bd}}\right)\overline{\mathbf{v}}\right)-\nabla^2\left(\mathbf{v_{bd}}\cdot\overline{\mathbf{v}}\right) \qquad (4.71)$$

where $\mathbf{v_{bd}}=\nabla\left(\nabla^{-2}\delta_b\right)$.

Similarly, the balanced correction from the departure from linear balance $\frac{\partial\delta_{blb}}{\partial t}$, takes the form

$$\begin{aligned}\frac{\partial\delta_{blb}}{\partial t} &= -\nabla^2\left[\left(\left(\mathbf{k}\cdot\nabla\psi_1\right)\cdot\overline{\mathbf{v}}\right)\right]+\\ &\quad \mathbf{k}\cdot\nabla\times\left(\left(\overline{\zeta}+f\right)\left(\left(\mathbf{k}\cdot\nabla\psi_1\right)+\left(\nabla^2\psi_1\right)\overline{v}\right)\right)+g\nabla^2 h_1. \quad (4.72)\end{aligned}$$

In principle the balanced corrections could be incorporated into the transformation to and from control variables and give a more accurate balanced variable which would include a balanced divergence. However in a data assimilation context the additional numerical cost needed to implement this scheme would outweigh the benefit of increased accuracy in the balanced variable. For this reason the above procedure is used only for comparisons between the PV and LB methods.

4.7 Control Variables and Burger Regimes

The PV method approximates the LB method under high Burger regimes since the majority of the linearised potential vorticity is held in the rotational wind in such regimes. For very small Burger number regimes for which the linearised potential vorticity is predominantly composed of a weighted height increment, $\overline{q}h'$, the PV method approximates the control variables set in which the height increment is considered to be the balanced variable. This is written as

$$\begin{aligned}
T_1^1: \quad h' &= h', \\
T_1^2: \quad \psi'_{ub} &= \nabla^{-2}\left(\mathbf{k}\cdot(\nabla\times\mathbf{u}')\right) - \left(\nabla\cdot(f\nabla)\right)^{-1}\nabla^2 g h', \\
T_1^3: \quad \chi' &= \nabla^{-2}\left(\nabla\cdot\mathbf{u}'\right). \qquad (4.73)
\end{aligned}$$

The inverse transform U_m is

$$\begin{aligned}
U_m^1: \quad (\mathbf{u}_1'^T, h_1) &= \left(\left(\nabla\cdot(f\nabla)\right)^{-1}\nabla^2 g h', h'\right) \\
U_m^2: \quad (\mathbf{u}_2'^T, h_2) &= (\mathbf{k}\times\nabla\psi_{ub}, 0) \\
U_m^3: \quad (\mathbf{u}_3'^T, h_3) &= (\nabla\chi', 0) \qquad (4.74)
\end{aligned}$$

The above change in control variables involves solving the reverse LBE which we have already mentioned as being problematic. Though the PV method is approximating this change in control variables, it may not have the same problems around the equator. This is because around the equator, due to $f \to 0$, the Burger number gets increasingly large and the PV method is going to approximate the LB method in these regions.

We need to examine the performance of LB and PV methods for a wide range of Burger numbers. Though it seems that there should be better results using the PV method, compared to the LB method for low Burger number, it is unclear as to the cutoff value in the Burger number when these improvements are noticable. The cutoff value should be close to 1, but this is to be checked in the experiments in Chapter 7.

4.8 Conclusion

In this chapter we have defined a change in control variables in terms of the data assimilation problem. A framework is proposed in which control variables can be examined. It is used to view the current change in control variables at the UK Met. Office when applied to SWEs on a sphere. A change in control variables is defined which conserves a potential vorticity increment. The dynamical properties of an ideal set of control variables are discussed. We then consider the relationship between control variables and divergence tendency and present a means to establish the respective performance, by finding a way to evaluate the divergence tendency in the unbalanced variables. The chapter concludes with an examination of how the control variables from the PV method vary in behaviour with Burger number.

Chapter 5

Numerical Background

5.1 Introduction

In this chapter we present the numerical details of the algorithms used in Chapter 7. The first part of this chapter relates to the shallow water equations. Section 5.2 gives some of the numerical details of the UK Met. Office's shallow water equation model on a sphere. In Section 5.3 we describe two different initial conditions: one described by the Rossby-Haurwitz wave on the sphere (3.34) - (3.36) and the other representing a realistic atmospheric situation.

Throughout Chapters 2, 4 we use linearisation states $\overline{h}$, $\overline{\mathbf{u}}$, $\overline{q}$. In Section 5.4 the experimental details of various types of linearisation states are presented that are used in the experiments in Chapter 7. The third part of this chapter considers the Poisson equation. Understanding the numerical properties of the Poisson equation not only gives the tools to solve the linear balance equation (LBE) and obtain the streamfunction and velocity potential from the vorticity and divergence - it also presents a method of solution which we use later in Chapter 6 to solve a control

Figure 5.1: Arakawa C grid

$$\begin{array}{ccc}
h_{i,j} & u_{i,j} & h_{i,j+1} \\
 & & \\
v_{i,j} & & v_{i,j+1} \\
 & & \\
h_{i+1,j} & u_{i+1,j} & h_{i+1,j+1}
\end{array}$$

variable transform based on conserving potential vorticity (4.29). The numerical evaluation of the Poisson equation is discussed in Section 5.5 while in Section 5.5.1 we consider the question of existence and uniqueness of solutions to this problem.

In the last two Sections, various finite difference approximations are presented, for which we give validatory evidence of their correct evaluation.

5.2 Shallow Water Equations: General Experimental Details

The numerical model approximating the shallow water equations calculates the height and wind fields on a staggered mesh called an Arakawa C grid. The relative positions of h, u, v are given in Figure 5.1.

The code which solves the SWEs is that used within the UK's Met. Office numerical weather prediction model (Unified Model). It is a semi-Lagrangian, semi-implicit, predictor-corrector scheme. The wind field is predicted for the next time step and the difference between the present time step and the next is calculated and stored. The continuity equation is used to define a variable coefficient Helmholtz equation for height differences between the two time levels. Once the Helmholtz equation is solved using a multigrid procedure, the height field at the new time level is known and the winds are corrected giving the final values for all the variables. Further details concerning its performance are presented in [35].

5.3 Some Initial Conditions for the Shallow-Water Equations

The initial conditions used in this thesis are constructed from standard tests that are normally used to compare the performance of various SWEs models. We use test cases 6 and 7 as defined in [30], [59], [25]. Case 6 uses the Rossby-Haurwitz wave on the sphere. Test case 7*c* uses observed 500mb height and wind fields from 0000 GMT January 9, 1979. Each case we now describe.

5.3.1 The Rossby-Haurwitz wave on a Sphere

As shown in Section 3.4, the Rossby-Haurwitz wave is a solution to the nonlinear non-divergent barotropic vorticity equation,

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot (f + \zeta) = 0, \tag{5.1}$$

which propagates a non-divergent solution. The Rossby-Haurwitz wave on the sphere, as presented in Section 3.4.1, is initially in Charney balance. In a shallow water model the unbalanced parts of the flow are created as time progresses.

Provided we consider a high Burger regime, the unbalanced parts remain small compared to the balanced parts.

Since Philips [44], investigators have been using this wave to provide initial conditions for the SWEs. Throughout this thesis we choose the wavenumber to be equal to 4, as we want the field to be stable. Following Hoskins [27], Rossby-Haurwitz waves with zonal wavenumbers less than or equal to 5 are commonly considered to be stable while those greater than 5 are considered unstable. Recently in papers [3], [56], the stability of the Rossby-Haurwitz wave for zonal wavenumber $R = 4$ has been put into question, due to significant differences between between various 5-day model simulations using different numerical models. Also certain numerical models provide solutions which disrupt both the basic symmetry of the U, H fields about the equator and also the antisymmetry of the V field. The numerical techniques we use to evaluate the change in control variables rely on these symmetries. Fortunately, the Met Office Semi-Lagrangian SWE model which we use preserves the symmetries needed.

5.3.2 Real initial conditions

Case $7c$ was used originally by Baumhefner and Bettge [4]. It involves a strong zonal flow. The observed height and (u, v) wind fields come from an FGGE analysis which was initialised using normal mode initialisation [13] at the UK Met. Office. The mean height field was set to 10 km. The balanced data is unformatted and represents the height and winds on a Arawaka C grid at a medium resolution with $\Delta\theta = \pi/65$ and $\Delta\lambda = \pi/96$ and gives an adequately detailed representation of the original observations. We also use the uninitialised fields held in file called

*VDG*7.11.*cdf* found in *ftp : //ftp.cgd.ucar.edu/pub/jet/shallow/nminit/* with mean depth 10.0 km and $T106$ spectral resolution.

From the initialised and uninitialised data fields we construct the fields for use in the numerical experiments of Chapter 7 by making the height h and u component of the wind symmetric about the equator while setting the v component of the wind to be anti-symmetric about the equator. This is done by taking the height and wind fields on the North Hemisphere and copying them onto the Southern Hemisphere. To enforce antisymmetry to the v component of the wind, we apply a sign change to the modified v field in the Southern Hemisphere. Since the original mass is not equally split between the hemispheres, we need to adjust the mean height level so that the total mass of new height field is the same as the original height field.

An additional set of height and wind fields are obtained similarly by taking the data values in the Southern Hemisphere, overwriting them onto the Northern Hemisphere, applying the sign change in the v field and adjust the mean height. Taking the average of the two sets of height and wind fields produces sensible base states. Appropriate latitudinally dependent linearisation states that are in geostrophic balance are constructed using the base states, following the procedure in Section 5.4.

5.4 Linearisation States

Whether we are dealing with perturbations or increments, linearisation states need to be defined. We look at two linearisation states: one set about a resting state as in Section 2.3 and another about a balanced latitudinally varying linearisation state. Both are used in the experiments in Chapter 7. In this section we present linearisation states in terms of perturbations, mentioning any differences to using increments along the way.

When a linearisation state about a resting state is used, the linearisation state H is a constant. We choose H and the perturbation h^* so that

$$\begin{aligned} h &= H + h^* \\ \int h^* dS &= 0 \end{aligned} \tag{5.2}$$

This is a unique choice due to h^* being symmetric about the equator. For height increments the surface integral is automatically true.

Since the wind perturbations are defined about a state of rest (2.38) $\overline{\psi}$ is equal to a constant and the full streamfunction perturbation is given by

$$\nabla^2 \psi^* = \mathbf{k} \cdot (\nabla \times \mathbf{v}) . \tag{5.3}$$

Given (2.38), (5.2), (5.3), we see that full increments can be defined in a similar way for use in various changes in control variables.

We define a balanced latitudinally dependent linearisation state using the PV method. We apply the linearisations about a resting state as a first approximation using (2.38), (5.2), (5.3) and $\overline{q} = \frac{f}{H}$. The coupled system,

$$\nabla \cdot f \nabla \psi_l - g \nabla^2 h_l = 0, \tag{5.4}$$

$$\nabla^2 \psi_l - \overline{q} h_l = \int_0^{2\pi} \nabla^2 \psi' - \overline{q} h' d\lambda, \tag{5.5}$$

is then solved for ψ_l and h_l. The new latitudinally varying linearisation states are defined as

$$\overline{u}(\theta) = -\frac{1}{a} \frac{\partial \psi_l}{\partial \theta} \tag{5.6}$$

$$\overline{h}(\theta) = h_l \tag{5.7}$$

$$\overline{v} = 0. \tag{5.8}$$

Since $\overline{h}q^* \approx \nabla^2\psi^* - \overline{q}h^*$ to order $O((h^*)^2, h^*\nabla^2\psi^*)$, the linearised potential vorticity perturbation is only accurate provided $\overline{h}q^* >> O(\overline{q}(h^*)^2/\overline{h}, h^*\nabla^2\psi^*)$. When $\overline{h}$ is small, as for low Burger regimes, there is a greater opportunity for this criterion to be violated, especially when q^*, h^* and ψ^* involve relatively large departures from respective linearisation states. This is why a latitudinally-varying linearisation state is considered, as the departures from the linearisation states are going to be much smaller compared to when $\overline{u} = 0$, $\overline{v} = 0$ and $\overline{h} = H$.

The effect of choosing latitudinally varying states that satisfy the PV method gives linearisation states that are in geostrophic balance. A linearisation state that satisfies a balance condition is necessary in Section 4.6 to identify correctly the divergence tendency in the unbalanced variables.

5.5 Poisson Equation

The Poisson equation on the sphere is given by

$$\frac{1}{a^2\cos^2\theta}\frac{\partial^2\phi}{\partial\lambda^2} + \frac{1}{a^2\cos\theta}\frac{\partial}{\partial\theta}\left[\cos\theta\frac{\partial\phi}{\partial\theta}\right] = F(\lambda,\theta), \tag{5.9}$$

where a is the radius of the sphere, λ and θ are the longitude and latitude, ϕ is the solution and F is a known forcing term.

The solution of the Poisson equation over a sphere is required for the calculation of the streamfunction, velocity potential and the calculation of balanced height from the streamfunction using the Linear Balance equation. The first step to solving the Poisson equation is to apply discrete fast Fourier transforms (DFFTs) in the longitudinal direction. This provides second order ordinary differential equations (ODEs) for each zonal wavenumber. These ODEs are discretised using a finite volume approach that is equivalent to a 2^{nd} order centered finite difference method. They are solved by applying a standard tri-diagonal solver. The solution to the Poisson equation is obtained by applying an inverse fast Fourier transform at every value of latitude considered. As the method of solution is similar to that given by Moorthi [38] we present only the key implementational details and leave the definition of the DFFT and its inverse to that paper.

Consider a regular latitude/longitude grid over a sphere, with the grid spacing in the latitude and longitude denoted respectively by

$$\Delta\theta = \frac{\pi}{N-1} \qquad \text{and} \qquad \Delta\lambda = \frac{2\pi}{M}, \tag{5.10}$$

where N is the number of points in the co-latitudinal direction, indexed as i, going from the north pole to the south pole. Similarly, M is the number of grid points in the longitudinal direction. The longitudinal dimension to the grid points, denoted by index j, are numbered positively in an anti-clockwise direction around the north pole. We stipulate that $M = 2^n 3^n$ where m, n are integers. This choice makes the Fourier transform and its inverse efficient to use. We also set N to be odd. This allows grid values with index $i = \frac{N+1}{2}$ to lie on the zero latitude.

We are solving the discrete problem on a sphere where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\lambda \in [0, 2\pi]$ and θ, λ are the latitude and the longitude. Periodic boundary conditions are assumed for the longitude. The problem is scaled onto the surface of a sphere of unit radius to eliminate unneccesary rounding error. Once the solution to the scaled problem is known, the rescaling back onto a spherical Earth is performed. Thus, for the following discussion, we consider the Poisson equation on a sphere of unit radius.

Application of the DFFT gives a tridiagonal system to be solved for each wavenumber considered. For $i = 2, \cdots, N-1$ and generic wavenumber k, the tridiagonal system takes the form

$$\begin{aligned} -\frac{k^2\pi^2\tilde{\phi}_i}{\cos(\theta_i)}+ & \\ \frac{\cos(\theta_{i-1/2})(\tilde{\phi}_{i-1}-\tilde{\phi}_i)+cos(\theta_{i+1/2})(\tilde{\phi}_{i+1}-\tilde{\phi}_i)}{\Delta\theta^2} &= \cos(\theta_i)\tilde{f}_i, \end{aligned} \tag{5.11}$$

where $\tilde{\phi}_i$, $\tilde{f}_i$ are the complex coefficients of the solution and right-hand side for latitude circle i. The grid and half grid values of the latitude are defined as:

$$\begin{aligned} \theta_i &= \frac{\pi}{2}-(i-1)\Delta\theta, && \text{for } i = 1, \cdots, N, \\ \theta_{i+1/2} &= \frac{\pi}{2}-(i-\frac{1}{2})\Delta\theta, && \text{for } i = 1, \cdots, N-1, \end{aligned} \tag{5.12}$$

To complete the description of the tridiagonal system, we need equations for the polar values. The equation at each pole is derived using the integral, finite volume approach as presented by Barros [2]. The equation is derived as follows:

$$\begin{aligned} \int_0^{2\pi}\int_{\frac{\pi}{2}-\frac{\Delta\theta}{2}}^{\frac{\pi}{2}} \nabla^2\phi\cos\theta d\theta d\lambda &= \sum_{l=1}^{M}\int_{(l-1)\Delta\lambda}^{l\Delta\lambda}\int_{\frac{\pi}{2}-\frac{\Delta\theta}{2}}^{\frac{\pi}{2}} \nabla^2\phi\cos\theta d\theta d\lambda \\ &= \sum_{l=1}^{M}\int_{(l-1)\Delta\lambda}^{l\Delta\lambda} -\nabla\phi\cdot\hat{\theta}\Big|_{\frac{\pi}{2}-\frac{\Delta\theta}{2}} d\lambda \\ &= \sum_{l=1}^{M}\Delta\lambda\cos(\theta_{\frac{3}{2}})\frac{\phi_{2,l}-\phi_{1,l}}{\Delta\theta} \\ &= \int_0^{2\pi}\int_{\frac{\pi}{2}-\frac{\Delta\theta}{2}}^{\frac{\pi}{2}} F\cos\theta\Delta\theta\Delta\lambda \\ &= F(P_N)V_N = F(P_N)\pi\frac{\Delta\theta}{2}\cos(\theta_{\frac{3}{2}}). \end{aligned} \tag{5.13}$$

The two-dimensional divergence theorem and the midpoint rule are used to evaluate

(5.13) where F is the scaled right-hand side value at the north pole P_N with the associated surface area V_N. The surface area V_N is the surface area of the spherical cap above latitude $\theta = (\pi - \Delta\theta)/2$.

We observe that in (5.13), a longitudinal mean is taken for both the grid points at the north pole and for $i = 2$. Since $\tilde{\phi}_1$ and $\tilde{\phi}_2$ represent longitudinal means when $k = 0$, equation (5.13) can be rewritten as

$$-\frac{4}{\Delta\theta^2}\tilde{\phi}_1 + \frac{4}{\Delta\theta^2}\tilde{\phi}_2 = \tilde{f}(0)_1. \tag{5.14}$$

For non-zero values of k we stipulate Dirichlet boundary conditions $\tilde{\phi}_1 = 0$, $\tilde{\phi}_N = 0$ and solve over the whole sphere. These boundary conditions enforce single values at the poles and give tridiagonal systems of full rank.

When $k = 0$ additional conditions are needed for a unique solution. We consider the problem when the right hand side is either symmetric or anti-symmetric about the equator. When the right hand side is anti-symmetric about the equator there is no difficulty in obtaining a solution. We solve over the upper hemisphere setting a zero Dirichlet boundary condition at the equator. The solution is copied onto the other hemisphere and a sign change is applied.

When $k = 0$ and the right hand side is symmetric about the equator a uniqueness condition needs to be satisfied. Such a problem occurs in this thesis when velocity potential $\tilde{\chi}_i$ needs to be evaluated from the divergence. This condition is

$$\int_{\theta=-\frac{\pi}{2}}^{\theta=\frac{\pi}{2}} \tilde{\chi}(\theta)\cos\theta d\theta \quad = \quad 0. \tag{5.15}$$

We solve over a hemisphere, where the equation at the equator is overwritten by

$$\frac{1}{a^2}\sum_{i=1}^{\frac{N+1}{2}} V_i\tilde{\chi}_i = 0 \tag{5.16}$$

where V_i is given by

$$\begin{aligned} V_1 &= 2\pi a^2 \frac{\Delta\theta}{2} \cos\theta_{\frac{1}{2}}, \\ V_i &= 4\pi a^2 \Delta\theta \cos\theta_i \qquad \text{for} \quad i = 2, \ldots, \frac{N-1}{2}. \\ V_{\frac{N+1}{2}} &= 2\pi a^2 \Delta\theta \end{aligned} \tag{5.17}$$

The solution to the tri-diagonal sytem is then copied to the other hemisphere before IDFFT's are used to get the solution to the Poisson equation.

A few additional notes have to be made. The computational procedure for the numerical solution of the Poisson equation is written in the Matlab language. It uses subroutines from the FFTW library [19],[20] that are well known and reliable. The right hand side of the Poisson equation has to satisfy the compatibility condition for the solution to exist. Further details are given in Section 5.5.1.

5.5.1 Compatibility condition

The compatibility condition is a necessary and sufficient constraint [52] that needs to be applied to the Poisson equation on the surface of a sphere in order for a solution to exist. It states that the right-hand side f has to satisfy

$$\int_{\Omega} f = 0. \tag{5.18}$$

Then the solution is unique up to the value of a constant.

The discrete problem preserves this property. A solution exists to the discrete problem provided that the right hand side is orthogonal, under the Euclidean inner product, to the kernel of the adjoint of the discretisation matrix. By multiplying both sides of the equation (5.9) by $\cos(\theta)$ the Poisson problem is rewritten in a self-adjoint form for which the constant function spans the kernel.

The self-adjointness manifests itself in the discretisation of $\cos(\theta)$ times (5.9). When a uniform grid and centered finite differences are used the associated discretisation matrix is symmetric.

Suppose that the $i = 2, \ldots, N-1$ rows of this matrix, A, are given by equation (5.11). The first row is equal to $\frac{1}{4}\cos(\theta_{\frac{3}{2}})$ times (5.14). The last row representing the equation at the south pole is defined in a similar fashion.

The discrete problem needs to satisfy a discrete version of (5.18). This discrete condition is equivalent to setting the sum of all the entries of the new right-hand side to zero.

If the discrete new right-hand side does not satisfy this discrete version of (5.18), then it needs to be perturbed so that the solution to the resulting consistent system is a least squares solution to the unperturbed one. This is achieved by

Let $e = (1, \ldots, 1)^T$, $f_i = \tilde{f}(0)_i$ for $i = 1, \ldots, N$,

$$h^T = (\frac{1}{4}\cos(\theta_{3/2}), h_1^T, \frac{1}{4}\cos(\theta_{\frac{2N-1}{2}})) \tag{5.19}$$

and

$$h_1^T = (\cos(\theta_2), \cos(\theta_2), \ldots, \cos(\theta_{N-1})) \tag{5.20}$$

We define the discretisation matrix B by $1/(cos(\theta_i))$ times equation (5.11) for rows $i = 2, \ldots, N-1$ with the first row given by $\frac{1}{4}$ times (5.14). The last row represents the equation at the south pole and is defined in a similar fashion to the north pole.

We define an inner product between two vectors f and g

$$(f, g)_H = f^T H g, \tag{5.21}$$

where $H = diag(h)$ and

$$||u||_h = u^T H u. \tag{5.22}$$

If we adjust the right-hand side to g, such that

$$g = f - \frac{(e, f)_H}{(e, e)_H} e, \tag{5.23}$$

the solution to the system $Bx = g$ is consistent. Also x minimises $||Bx - f||_H$.

In fact, any matrix norm can be used to perturb the right-hand side. However it is better to perturb f by a vector e that has no functional dependence as the functional dependence of the errors in f will not normally be known. If the L_2 norm is used, then the perturbation is no longer constant but equal to h. It also approximates a function which is not regular at the poles.

5.5.2 Right hand sides

As mentioned at the beginning of this section, a Poisson equation needs to be solved to obtain streamfunction, velocity potential and balanced height. Each of the necessary right hand sides are approximated using a second-order centered finite difference discretisation. Due to the periodicity of the sphere,

$$\begin{aligned} f_{i,M+1} &= f_{i,1} && \qquad (5.24) \\ f_{i,0} &= f_{i,M} && \qquad i = 1, \cdots, N \end{aligned}$$

The polar values, denoted as f_N and f_S, are taken as the mean average of the values of f on the latitudinal rings $i = 2$ and $i = N - 1$.

The streamfunction ψ is calculated by solving

$$\nabla^2 \psi = \mathbf{k} \cdot (\nabla \times \mathbf{v}) \tag{5.25}$$

where $\mathbf{k} \cdot (\nabla \times \mathbf{v})$ is the relative vorticity. The winds are evaluated on an Arakawa C grid and the relative vorticity given at grid points coincident with height values.

The relative vorticity is evaluated to $\mathcal{O}(\Delta\theta^2, \Delta\lambda^2)$ by

$$\begin{aligned}
f_{i,j} &= (\mathbf{k}.\,(\nabla \times \mathbf{v}))_{i,j} \qquad (5.26) \\
&= \frac{1}{a\ \cos\theta}\left(\frac{\partial v}{\partial \lambda} + \frac{\partial}{\partial \theta}(\cos\theta u)\right)_{i,j} \\
&= \frac{1}{a\ \cos\theta_i}\Bigg(\frac{v_{i,j+1} - v_{i,j-1} + v_{i-1,j+1} - v_{i-1,j-1}}{4\Delta\lambda} \\
&\quad - \frac{\cos\theta_{i-1}(u_{i-1,j} + u_{i-1,j-1}) - \cos\theta_{i+1}(u_{i+1,j} + u_{i+1,j-1})}{4\Delta\theta}\Bigg),
\end{aligned}$$

for $i = 2, \ldots, N-1$, $j = 1, \ldots, M$. The Poisson equation is first solved over the unit sphere. To do this the right-hand side is multiplied by the radius of the earth. The solution is multiplied by a to give the value on the surface on the Earth.

The velocity potential χ is calculated by solving

$$\nabla^2\chi = \nabla \cdot \mathbf{v} \qquad (5.27)$$

where $\nabla\cdot\mathbf{v}$ is the horizontal divergence of the wind-field. It is evaluated to $\mathcal{O}(\Delta\theta^2, \Delta\lambda^2)$ at points coincident with the height on an Arakawa C grid. The divergence is given by

$$\begin{aligned}
f_{i,j} &= (\nabla \cdot \mathbf{v})_{i,j} \qquad (5.28) \\
&= \frac{1}{a\ \cos\theta}\left(\frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \theta}(\cos\theta v)\right)_{i,j} \\
&= \frac{1}{a\ \cos\theta_i}\Bigg(\frac{u_{i,j} - u_{i,j-1}}{\Delta\lambda} + \\
&\quad \frac{\cos\theta_{i-1} v_{i-1,j} - \cos\theta_i v_{i,j}}{\Delta\theta}\Bigg),
\end{aligned}$$

for $i = 2 \ldots N-1$, $j = 1, \ldots M$. The same scaling is used as with the streamfunction.

The right-hand side to the LBE (3.37) is approximated to $\mathcal{O}(\Delta\theta^2, \Delta\lambda^2)$, using

$$\begin{aligned}
f_{i,j} &= (\nabla \cdot f\nabla\psi)_{i,j} \qquad (5.29) \\
&= \frac{1}{a^2\ \cos\theta_i}\Bigg(\frac{f_i(\psi_{i,j+1} - 2\psi_{i,j} - \psi_{i,j-1})}{\cos\theta\Delta\lambda^2} +
\end{aligned}$$

$$\frac{f_{i-\frac{1}{2}}cos\theta_{i-\frac{1}{2}}\psi_{i-1,j} - \left(f_{i-\frac{1}{2}}\cos\theta_{i-\frac{1}{2}} + f_{i+\frac{1}{2}}\cos\theta_{i+\frac{1}{2}}\right)\psi_{i,j} + f_{i+\frac{1}{2}}\cos\theta_{i+\frac{1}{2}}\psi_{i+1,j}}{\Delta\theta^2}\Bigg),$$

for $i = 2 \ldots N-1$, $j = 1, \ldots, M$. For this approximation, the right-hand side is multiplied by a^2 to produce the correct scaling. The solution obtained on the unit sphere is the same as on the surface of the Earth.

5.6 Additional Numerical Schemes and Numerical Tests

Since Poisson inversion is used throughout this thesis, it is expedient to present validation tests for the right hand sides within Section 5.5.2 and the Poisson inversion used to calculate the streamfunction, velocity potential from the vorticity and divergence. We also validate the LBE where balanced height fields are obtained from the streamfunction.

The winds we use in the tests are

$$\begin{aligned} U &= a\omega\cos\theta + aK\left(\cos^{R-1}\theta\left(R\sin^2\theta - \cos^2\theta\right)\cos R\lambda\right) \\ &\quad -aKR\cos^{R-1}\theta\sin R\lambda, && (5.30) \\ V &= -aKR\sin\theta\cos^{R-1}\theta\sin R\lambda \\ &\quad -2a\omega\sin\theta\cos\theta - aKR\left(\sin\theta\cos^{R-1}\theta\cos R\lambda\right), && (5.31) \end{aligned}$$

where $R = 4$, $K = \omega = 7.847 \times 10^{-7}\ s^{-1}$, $a = 6371220$m. The rotational part of the winds are equivalent to a Rossby Haurwitz wave. The streamfunction and velocity potential of these winds are

$$\psi = -a^2\omega\sin\theta + a^2K\cos^R\theta\sin\theta\cos R\lambda, \qquad (5.32)$$

$$\chi = a^2\omega\cos^2\theta + a^2K\cos^R\theta\cos R\lambda \tag{5.33}$$

and the vorticity and streamfunction are given by

$$\zeta = 2\omega\sin\theta - K\sin\theta\cos^R\theta\left(R^2+3R+2\right)\cos R\lambda, \tag{5.34}$$

$$\begin{aligned}\delta &= -2\omega\left(\cos^2\theta - 2\sin^2\theta\right)\\ &\quad +K\cos^{R-2}\theta\left(-R^2 - R\cos^2\theta + R^2\sin^2\theta\right)\cos R\lambda.\end{aligned} \tag{5.35}$$

The right hand side of LBE is

$$\begin{aligned}\nabla\cdot f\nabla\psi &= -2\Omega R^2K\cos^{R-2}\theta\sin^2\theta\cos R\lambda\\ &\quad +2\Omega\left[2\omega\sin^2\theta - \omega\cos^2\theta\right.\\ &\quad +K\cos R\lambda\left(R^2\cos^{R-2}\theta\sin^4\theta - (4R+2)\cos^R\theta\sin^2\theta\right.\\ &\quad \left.\left.+\cos^{R+2}\theta\right)\right].\end{aligned} \tag{5.36}$$

We use a normalised L_2 vector norm to get estimates of the error between the numerical approximations and the true analytic solutions. The error estimate ε is of the form

$$\varepsilon = \frac{\left(\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\alpha_{i,j}^t - \alpha_{i,j}^a\right)^2\right)^{\frac{1}{2}}}{\left(\sum_{i=1}^{N}\sum_{j=1}^{M}\left(\alpha_{i,j}^t\right)^2\right)^{\frac{1}{2}}} \tag{5.37}$$

for a generic true field α^t and generic approximation α^a.

The third and fourth columns of Table 5.1 show the L_2 error in ζ, δ between the analytic solutions (5.34), (5.35) and the the numerical approximations (5.26), (5.29) applied to the analytic winds (5.30), (5.31) . The far right column of Table 5.1 shows the L_2 error between analytically defined $\nabla\cdot f\nabla\psi$ given by (5.36) and the numerical approximated value given by the numerical discretisation (5.30) using values defined by (5.33). In all three cases, a doubling of the resolution results in

Table 5.1: L_2 error between numerically and analytically defined ζ, δ and $\nabla \cdot f\nabla\psi$ using equation (5.37)

M	N	L_2 error in ζ	L_2 error in δ	L_2 error in $\nabla \cdot f\nabla\psi$
48	33	0.0494	0.0110	0.0197
96	65	0.0126	0.0027	0.0049
192	129	0.0032	6.72×10^{-4}	0.0012

a four-fold reduction in the relative error, as expected for second order accurate schemes.

The third and fourth columns of Table 5.2 shows the decrease the relative L_2 error with increased resolution of ψ, χ . This is a test of the Poisson inverter where the right hand sides have been evaluated using second order finite difference schemes (5.26), (5.29) applied to the analytically defined wind (5.30), (5.31). The final column shows the decrease the relative L_2 error of the laplacian of balanced height h_b. The laplacian of the balanced height is calculated by first solving a Poisson equation to obtain a streamfunction using the analytically derived winds (5.30), (5.31). Then we solve the LBE using the finite difference approximation (5.36) to retrieve the balanced height h_b. Then we apply a discretised second order centered finite difference approximation of the laplacian to the balanced height of the form

$$\begin{aligned}\nabla^2 h_{b\ i,j} \quad = \quad & \frac{h_{b\ i,j+1} - 2h_{b\ i,j} + h_{b\ i,j+1}}{a\cos^2\theta_i(\Delta\lambda)^2} + \\ & \frac{\cos\theta_{i-\frac{1}{2}} h_{b\ i-1,j} - \left(\cos\theta_{i-\frac{1}{2}} + \cos\theta_{i-\frac{1}{2}}\right) h_{b\ i,j} + \cos\theta_{i-\frac{1}{2}} h_{b\ i+1,j}}{a\cos\theta_i(\Delta\theta)^2} \end{aligned} \tag{5.38}$$

Table 5.2: L_2 error between analytical defined streamfunction ψ (5.33), velocity potential χ (5.33) and $\nabla \cdot f\nabla\psi$ 5.36 and the respective numerically approximated fields, given by solving a Poisson equation with the respective right hand sides and evaluated using schemes (5.26), (5.29), (5.30) applied to winds (5.30), (5.31)

M	N	L_2 error in ψ	L_2 error in χ	L_2 error in $\nabla^2 h_b$
48	33	0.0043	0.0131	0.0759
96	65	0.0010	0.0032	0.0188
192	129	2.48×10^{-4}	7.74×10^{-4}	0.0047

and compare with the analytically defined $\nabla \cdot f\nabla\psi$, (5.36). This is done to show that the error overall after repeated Poisson inversion remains second order accurate. In each situation the relative L_2 error decreases with $O((\Delta\theta)^2, (\Delta\lambda)^2)$ as expected.

The other test that is applied to the Poisson solver is to compare an analytically defined wind field (5.30), (5.31) with a numerically approximated one. The numerically approximated wind is obtained by first converting the original wind fields (5.30), (5.31) into vorticity and divergence using (5.26), (5.29). The Poisson inverter is used to obtain both the velocity potential and streamfunction. The wind fields are recalculated using the Helmholtz relations

$$\begin{aligned} u &= u_r + u_d \\ &= -\frac{1}{a}\frac{\partial\psi}{\partial\theta} + \frac{1}{a\ \cos\theta}\frac{\partial\chi}{\partial\lambda} \qquad (5.39) \\ v &= v_r + v_d \\ &= \frac{1}{a\ \cos\theta}\frac{\partial\psi}{\partial\lambda} + \frac{1}{a}\frac{\partial\chi}{\partial\theta} \qquad (5.40) \end{aligned}$$

where (u_r, u_d) is the rotational part of the wind and (v_r, v_d) is the divergent part.

Second order centered finite difference approximations are used, where the ψ and χ fields coincident with the height positions of the Arakawa C grid, giving wind fields consistent with the considered staggered grid.

$$\begin{aligned}
u_r(i,j) &= -\frac{1}{a}\frac{\psi_{i-1,j} - \psi_{i+1,j} + \psi_{i-1,j+1} - \psi_{i+1,j+1}}{4\Delta\theta} \\
v_r(i,j) &= \frac{1}{a \ \cos\theta_{i+\frac{1}{2}}}\frac{\psi_{i,j+1} - \psi_{i,j-1} + \psi_{i+1,j+1} - \psi_{i+1,j-1}}{4\Delta\lambda} \\
u_d(i,j) &= \frac{1}{a \ \cos\theta_i}\frac{\chi_{i,j+1} - \chi_{i,j}}{\Delta\lambda} \\
v_d(i,j) &= \frac{1}{a}\frac{\chi_{i,j} - \chi_{i+1,j}}{\Delta\theta} \qquad (5.41)
\end{aligned}$$

for $i = 1, \ldots, N$ and $j = 1, \ldots, M$ with ghost points (i,j) indexed as

$$(i,0) = (i,M), \qquad (i,M+1) = (i,1) \quad \text{for} \quad i = 1, \ldots, N,$$

due to periodicity in longitudinal direction. In addition, for $i = 1, N$, the stencil of the discretisation goes over the pole such that ghosts point are:

$$\begin{aligned}
(0,j) &= (1, j+M/2) \quad \text{if } j \leq M/2, \qquad (1, j-M/2) \quad \text{if } j > M/2; \\
(N+1,j) &= (N, j+M/2) \quad \text{if } j \leq M/2, \qquad (N, j-M/2) \quad \text{if } j > M/2;
\end{aligned}$$

remembering that M is chosen to be an even integer. Since the spherical co-ordinate system is degenerate at the poles the values of the u field at such points are set to zero.

Table 5.3 shows the difference in the error between the original and the numerically approximated wind fields using grid-points that do not reside at the pole. The relative error decreases as the resolution is increased consistent with an accuracy

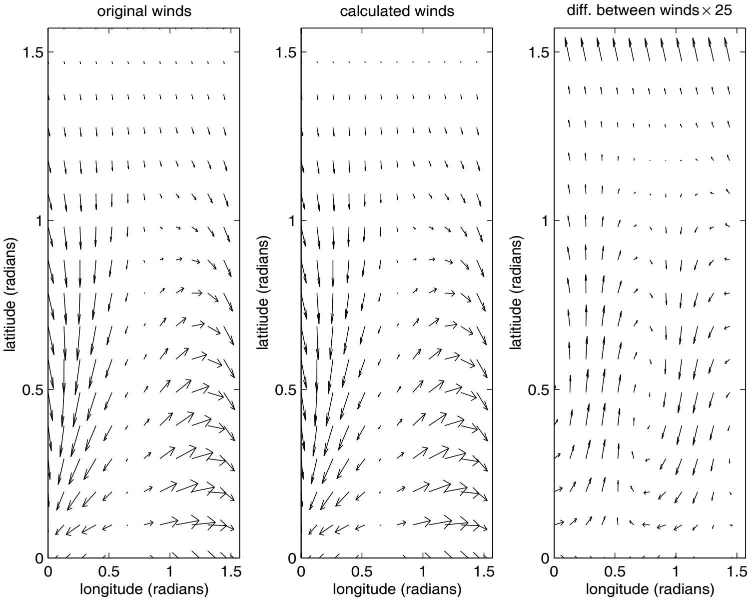
Table 5.3: Relative L_2 error in the (u,v) components of the wind field at various resolutions

N	M	Relative L_2error in u component of wind field	Relative L_2error in v component of wind field
33	48	0.0163	0.0384
65	96	0.0043	0.0114
129	192	0.0012	0.0036

just under the formal 2^{nd} order of the numerical schemes. The original wind field (left) are compared with the derived wind fields (centre) in Figure 5.2 for $N = 33$, $M = 96$ and $\theta \in [0, \frac{\pi}{2}]$, $\lambda \in [0, \frac{\pi}{2}]$. Even at this low resolution the wind fields are visually very similar to each other. The difference between the two fields is shown on the right multiplied by 25. The greatest error is at the pole. This is due to difficulty in evaluating the winds from the streamfunction and velocity potential where the co-ordinate system is degenerate.

The validation tests above show that the numerical technique used to solve the Poisson equation on the sphere produces overall good results. We have presented results that show that we can solve the LBE. By being able to solve LBE, we are not only in a position to reproduce the current change in control variables as it relates to the SWEs but also ready to describe the subject of the following chapter - the numerical implementation of a new change in control variables whose balanced variable conserves potential vorticity.

Figure 5.2: Original wind field (left) and calculated wind field (centre) and diffrence between them ×25 for $N = 33$, $M = 96$ and $\theta \in [0, \frac{\pi}{2}]$, $\lambda \in [0, \frac{\pi}{2}]$



Chapter 6

A potential vorticity control variable transform

6.1 Introduction

In Chapter 4, we derive balanced height and balanced rotational wind increments from a potential vorticity increment. We show how these height and wind differences are consistent with the process of geostrophic adjustment. In this chapter we present the numerical details for solving the coupled system of equations (4.28), (4.29). The method we use is an extension of the procedure used in Chapter 5 to solve the Poisson equation. In Section 6.2, as in the previous chapter, we describe the problem from a continuous standpoint. In this way the essence of the method is revealed. Then the numerical details are presented: first by showing appropriate rescaling of the problem in Section 6.3 and then, in Section 6.4, displaying the numerical discretisation used.

The overall procedure to obtain the new control variables is shown in Section

6.5, where we briefly describe how the height and wind fields are transformed into control variables of increments. We finish the chapter with a description of a number of experiments to verify the implementation of the method.

The original proposal, as suggested by Mike Cullen, solves the coupled system of PDE's, by eliminating the height variable from the equations to produce a fourth order boundary value problem with a leading biharmonic term. Difficulties have arisen however in the numerical implementation of the appropriate boundary conditions. Thus, we solve instead the coupled system directly. The advantage of solving the problem directly is that the associated boundary conditions are analogous to boundary conditions of problems that are well researched in the literature, namely the solution of the Poisson equation on the sphere and the solution of the forward and reverse linear balance equation.

6.2 The Continuous Problem

Consider the coupled system of partial differential equations over a sphere (4.28), (4.29) where $\overline{h}$ and $\overline{q}$ are known linearisation fields dependent on latitude only and h_b, ψ_b are increments in height and streamfunction to be found over a sphere.

By stipulating that the linearisation states $\overline{h}$, $\overline{q}$, $\overline{\psi}$ are functions of latitude only, we make the coupled system of PDE's separable. We can solve the system by first applying a Fourier transformation to the RHS of (4.29) for each latitude circle, to produce an ODE for each wavenumber from the north pole to the south across latitudes. An inverse Fourier transform is used to obtain balanced streamfunction and height.

The assumption that the potential vorticity linearisation state is a function of latitude only is a reasonable assumption when the data is coming from a global shallow water model. Except for around the equator, the major contributor to the absolute vorticity is the Coriolis parameter, which is a function of the latitude only. Also the change in height field at any given latitude seems to vary between 5 % and 20 % of its average value.

This method has a number of advantages. Since the ODE's to be solved for each wavenumber are independent of each other, they can be solved in parallel, making this method quite efficient. Memory requirements are relatively low. Obtaining correct boundary conditions at the poles is well documented in the literature [2]. The equation should be less sensitive to error as we are solving coupled systems of ODEs, instead of a highly sensitive fourth order PDE.

Let us assume that the fields h_b, ψ_b, $\overline{h}q'$ can be described by a discrete inverse fast Fourier transform (DIFFT) in the longitudinal direction, such that

$$\begin{aligned}
\mathcal{Q}'(\lambda,\theta) &= \frac{1}{I}\sum_{k=-I/2}^{k=I/2} \tilde{\mathcal{Q}}'^{k}(\theta)e^{ik\lambda}, \qquad (6.1)\\
h_b(\lambda,\theta) &= \frac{1}{I}\sum_{k=-I/2}^{k=I/2} \tilde{h}^{k}(\theta)e^{ik\lambda},\\
\psi_b(\lambda,\theta) &= \frac{1}{I}\sum_{k=-I/2}^{k=I/2} \tilde{\psi}^{k}(\theta)e^{ik\lambda},
\end{aligned}$$

where I is an integer setting a truncation limit to the Fourier approximation, k is the wavenumber, i is equal to $i = \sqrt{-1}$ and

$$\begin{aligned}
\mathcal{Q}' &= \overline{h}q' \qquad (6.2)\\
\tilde{\mathcal{Q}}' &= \widetilde{\overline{h}q'} = \mathcal{Q}'^{r} + i\mathcal{Q}'^{i},
\end{aligned}$$

$$\begin{aligned}\widetilde{h} &= h^r + ih^i,\\ \widetilde{\psi} &= \psi^r + i\psi^i,\end{aligned}$$

with h^r, h^i, ψ^r, ψ^i, $\mathcal{Q}'^r$, $\mathcal{Q}'^i$ being real. Subsitution of (6.1) into (4.28) and (4.29), produces a coupled system of second order ODE's in θ to be solved for variables $\widetilde{h}$, $\widetilde{\psi}$. The system for a generic value of k is given by

$$-\frac{k^2}{a^2\cos^2\theta}[-g\widetilde{h} + f\widetilde{\psi}] + \frac{1}{a^2\cos\theta}\frac{\partial}{\partial\theta}[-\cos\theta\frac{\partial g\widetilde{h}}{\partial\theta} + f\cos\theta\frac{\partial\widetilde{\psi}}{\partial\theta}] = 0 \qquad (6.3)$$

$$-\frac{k^2}{a^2\cos^2\theta}[\widetilde{\psi}] + \frac{1}{a^2\cos\theta}\frac{\partial}{\partial\theta}[\cos\theta\frac{\partial\widetilde{\psi}}{\partial\theta}] - \overline{q}\widetilde{h} = \widetilde{\mathcal{Q}}'. \qquad (6.4)$$

To solve this system we obtain $\widetilde{\mathcal{Q}'}$ using a discrete fast Fourier transform (DFFT). Once the system (6.3)-(6.4) is solved, we use DIFFTs to recover the required fields h_b, ψ_b.

To solve (6.4) we need boundary conditions. As with the Poisson equation, the coefficients $\widetilde{h}$ and $\widetilde{\psi}$ are set to zero at the poles for all non-zero wavenumbers in order to enforce single-values at these points.

For $k = 0$ we solve (6.4) over the sphere and enforce a zero value at the equator for the anti-symmetric balanced streamfunction increment. A global uniqueness condition is used of the form,

$$\int_{\theta=-\frac{\pi}{2}}^{\theta=\frac{\pi}{2}} \widetilde{h}(\theta)\cos\theta d\theta = 0. \qquad (6.5)$$

The global uniqueness condition applied to $\widetilde{\psi}$,

$$\int_{\theta=-\frac{\pi}{2}}^{\theta=\frac{\pi}{2}} \widetilde{\psi}(\theta)\cos\theta d\theta = 0. \qquad (6.6)$$

is automatically satisfied due to the imposed anti-symmetric nature of the right hand side.

6.3 Scaling

The actual numerical procedure used to solve (6.3), (6.4) is slightly different to the simplified explanation given in the previous section. A scaling is introduced to make terms in the discretised operator of approximately the same size. Scaling of equations is important so as to eliminate unnecessary sensitivity in the problem due to numerical error. We wish to choose a scaling which will produce a coupled system of ODE's that can be represented by a positive definite matrix. Such a matrix guarantees a unique solution to the discrete problem.

We look at the scaling of the problem by first examining the effect of a control volume approach and then apply an additional scaling to weigh the terms appropriately. To consider the control volume technique the discrete problem has to be considered in so far as defining the grid points and surface areas used. The grid spacing $\Delta\theta$, $\Delta\lambda$ is the same as (5.10) as well as the indexing, with θ_i, $\theta_{i+\frac{1}{2}}$, $\theta_{i-\frac{1}{2}}$ defined by (5.12). M and N denote the number of grid points in the longitudinal and latitudinal directions. The Discrete Fourier transformation applied on latitude rings produces a partially discretised version of both (6.3), (6.4), giving for a general wavenumber k

$$\begin{aligned}
-\frac{k^2}{a^2\cos^2\theta_i}[-g\widetilde{h}_i + f_i\widetilde{\psi}_i] - \frac{g}{a^2\cos\theta_i}\left(\frac{\partial}{\partial\theta}[\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{h}_i & & \\
+\frac{1}{a^2\cos\theta_i}\left(\frac{\partial}{\partial\theta}[f\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{\psi_i} &= 0 & (6.7) \\
-\frac{k^2}{a^2\cos^2\theta_i}\widetilde{\psi_i} + \frac{1}{a^2\cos\theta_i}\left(\frac{\partial}{\partial\theta}[\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{\psi}_i - \overline{q}_i\widetilde{h}_i &= \tilde{\mathcal{Q}}'_i. & (6.8)
\end{aligned}$$

with $i = 1, \ldots, N$, representing the position in latitude, $k = -\frac{M}{2}, \ldots \frac{M}{2}$ representing the Fourier coefficient and the operators at each point i not yet discretised. When the control volume approach is used, both equations are multiplied by the surface area of the segment V_i represented by

$$\begin{aligned} V_1 &= 2\pi a^2 \frac{\Delta\theta}{2}\cos\theta_{\frac{1}{2}}, & V_N = 2\pi a^2 \frac{\Delta\theta}{2}\cos\theta_{N-\frac{1}{2}}, \\ V_i &= 2\pi a^2 \Delta\theta \cos\theta_i & \text{for } \ i = 2,\ldots,N-1. \end{aligned} \tag{6.9}$$

This is an $\Delta\theta^2$ approximation to the true surface area of a spherical segment whose value is equal to $4\pi a^2 \cos\theta_i \sin\Delta\theta$ for $i = 2,\ldots,N-1$.

The effect of this eliminates a^2 explicitly from the left hand side of the coupled system of equations

$$\begin{aligned} 2\pi\Delta\theta \left(-\frac{k^2}{\cos\theta_i}[-g\widetilde{h}_i + f_i\widetilde{\psi}_i] - g\left(\frac{\partial}{\partial\theta}[\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{h}_i \right. & \\ \left. + \left(\frac{\partial}{\partial\theta}[f\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{\psi}_i \right) &= 0 \qquad (6.10) \\ 2\pi\Delta\theta \left(\left(\frac{\partial}{\partial\theta}[\cos\theta\frac{\partial}{\partial\theta}]\right)_i \widetilde{\psi}_i - \frac{k^2}{\cos\theta_i}\widetilde{\psi}_i - a^2\cos\theta_i \overline{q}_i \widetilde{h}_i \right) &= 4\pi a^2 \Delta\theta \cos\theta_i \widetilde{\mathcal{Q}}'_i. \end{aligned}$$

The scaling at the poles is similar but involves both the discretisation of the operators concerned and the boundary conditions considered and as such is left to Section 6.4.

The proposed additional scaling is

$$\begin{aligned} \widetilde{\mathcal{H}}_i &= g\widetilde{h}_i, \\ \widetilde{\Psi}_i &= 2\Omega\widetilde{\psi}_i, \end{aligned} \tag{6.11}$$

where Ω is the Earth's angular velocity. When (6.11) is introduced into both (6.10), $2\Omega\times$ (6.12), the following ODE system is obtained:

$$\begin{aligned} -\frac{k^2\Delta\theta}{\cos\theta_i}[-\widetilde{\mathcal{H}}_i + \sin\theta_i\widetilde{\Psi}_i] - \Delta\theta\left(\frac{\partial}{\partial\theta}\cos\theta\frac{\partial}{\partial\theta}\right)_i \widetilde{\mathcal{H}}_i & \\ +\Delta\theta\left(\frac{\partial}{\partial\theta}\sin\theta\cos\theta\frac{\partial}{\partial\theta}\right)_i \widetilde{\Psi}_i &= 0 \qquad (6.12) \\ -\frac{k^2\Delta\theta}{\cos\theta_i}\widetilde{\Psi}_i + \Delta\theta\left(\frac{\partial}{\partial\theta}[\cos\theta_i\frac{\partial}{\partial\theta}]\right)_i \widetilde{\Psi}_i - Q_i\widetilde{\mathcal{H}}_i &= R_i \qquad (6.13) \end{aligned}$$

where $Q_i = 2\Delta\theta\Omega\mathbf{Q}_i/g$ and $R_i = 2\Omega a^2\Delta\theta\cos\theta_i\widetilde{\mathcal{Q}}'_i$. The constant 2π, present in all terms of equations (6.10), is ignored in (6.12) and (6.13). Since $2\Omega \geq f$ the $\mathcal{H}$ terms will be slightly greater than the Ψ terms in (6.13). Also Q is smaller than the matrix coefficients representing ∇^2, making Ψ dominate over $\mathcal{H}$ in (6.13). However, the size of these differences are not too large to make the equations 'numerically unbalanced'; the solutions Ψ and $\mathcal{H}$ are roughly of the same magnitude.

6.4 Numerical Implementation of Coupled ODE System

We are finally in a position to describe the numerical discretisation of the coupled ODE system. As with the Poisson equation, we use the same staggered grid, evaluating the solution and the right hand side of (6.12), (6.13) on a Arakawa C grid (Figure 5.1) coincident with the height values. As before, we apply discrete fast Fourier transforms and the inverse fast Fourier transforms longitudinally. All that remains is the description of the numerical solution to the second order latitudinally-dependent coupled ODE which we now describe.

Let

$$\mathbf{x}_i = \begin{pmatrix} \widetilde{\mathcal{H}}_i \\ \widetilde{\Psi}_i \end{pmatrix}, \qquad \text{for } \ i = 1, \cdots, N \tag{6.14}$$

$$\mathbf{b}_i = \begin{pmatrix} 0 \\ 2\Omega a^2\Delta\theta\cos\theta_i\widetilde{\mathcal{Q}}'_i \end{pmatrix}, \quad \text{for } \ i = 1, \cdots, N, . \tag{6.15}$$

where the superscript k is dropped for clarity purposes and it is assumed as before that $\mathbf{x}_i$ and $\mathbf{b}_i$ are chosen for some generic value of k.

The systems of ODE's for wavenumbers $k \neq 0$ are solved over the whole sphere and are approximated by a second order centred finite difference approximation.

Since $\widetilde{\Psi}$ and $\widetilde{H}$ are scalar values at the poles, they have no longitudinal dependency. Thus $\widetilde{\Psi}$, $\widetilde{H}$, $\mathbf{x}_1^k$ and $\mathbf{x}_N^k$ are zero at these points for $|k| \geq 1$ and act as Dirichlet boundary conditions. The system has a block tridiagonal structure and can be represented in the form $A\mathbf{x} = \mathbf{b}$ with A being a $(2N-4) \times (2N-4)$ block tri-diagonal nonsingular matrix and is solved using the QR method. The tridiagonal system that is solved has the form:

$$\begin{aligned} S_{2,2}\mathbf{x}_2 + S_{2,3}\mathbf{x}_3 &= \mathbf{b}_2, \\ S_{i,i-1}x_{i-1} + S_{i,i}x_i + S_{i,i+1}\mathbf{x}_{i+1} &= \mathbf{b}_i \qquad \text{for} \quad i = 3, \cdots, N-2, \\ S_{N-1,N-2}\mathbf{x}_{N-2} + S_{N-1,N-1}\mathbf{x}_{N-1} &= \mathbf{b}_{N-1}. \end{aligned} \tag{6.16}$$

Each block S is a 2×2 matrix defined as

$$S_{i,i} = \begin{pmatrix} \frac{k^2 \Delta\theta}{\cos\theta_i} + \frac{(\cos\theta_{i-\frac{1}{2}} + \cos\theta_{i+\frac{1}{2}})}{\Delta\theta} & -\frac{k^2\Delta\theta}{\cos\theta_i}\frac{f_i}{2\Omega} - \frac{(f_{i-\frac{1}{2}}\cos\theta_{i-\frac{1}{2}} + f_{i+\frac{1}{2}}\cos\theta_{i+\frac{1}{2}})}{2\Omega\Delta\theta} \\ -\Delta\theta\ Q_i & -\frac{k^2\Delta\theta}{\cos\theta_i} - \frac{(\cos\theta_{i-\frac{1}{2}} + \cos\theta_{i+\frac{1}{2}})}{\Delta\theta} \end{pmatrix} \tag{6.17}$$

$$S_{i,i-1} = \begin{pmatrix} -\frac{\cos\theta_{i-\frac{1}{2}}}{\Delta\theta} & \frac{f_{i-\frac{1}{2}}\cos\theta_{i-\frac{1}{2}}}{2\Omega\Delta\theta} \\ 0 & \frac{\cos\theta_{i-\frac{1}{2}}}{\Delta\theta} \end{pmatrix}, \tag{6.18}$$

$$S_{i,i+1} = \begin{pmatrix} -\frac{\cos\theta_{i+\frac{1}{2}}}{\Delta\theta} & \frac{f_{i+\frac{1}{2}}\cos\theta_{i+\frac{1}{2}}}{2\Omega\Delta\theta} \\ 0 & \frac{\cos\theta_{i+\frac{1}{2}}}{\Delta\theta} \end{pmatrix}. \tag{6.19}$$

It is clear that the matrix operator representing the discretisation of the coupled ODE is defined by (6.15)-(6.19) for $i = 2, \cdots, N-1$ with $S_{2,1}$ and $S_{N-1,N}$ ignored.

For $k = 0$ the situation is a little more complex. At the poles we set Neumann boundary conditions

$$\frac{\partial\Psi}{\partial\theta} = 0 \qquad \frac{\partial H}{\partial\theta} = 0, \tag{6.20}$$

which fix the Ψ and $\mathcal{H}$ fields to the addition of a constant. Additional boundary conditions are applied to both Ψ and $\mathcal{H}$. The value of Ψ must be zero at the equator as the balanced streamfunction is anti-symmetric about this value and is assumed to be a continuous smooth function. This additional piece of information is achieved by removing the $\widetilde{\Psi}_{\frac{N+1}{2}}$ from the system considered, by extracting the $N+1^{th}$ row and column of the system. A consequence of doing this, is that it removes one of the equations that need to be satisfied at the equator, specifically

$$\frac{\partial^2 \Psi}{\partial \theta^2}_{i=\frac{N+1}{2}} = 0, \tag{6.21}$$

and is resolved by adding the coupled equations at the equator together to give

$$-\frac{\partial^2 \mathcal{H}}{\partial \theta^2}_{i=\frac{N+1}{2}} + \frac{\partial \Psi}{\partial \theta}_{i=\frac{N+1}{2}} + \frac{\partial^2 \Psi}{\partial \theta^2}_{i=\frac{N+1}{2}} = 0. \tag{6.22}$$

The simplicity of the equations (6.21), (6.22) is due to $\overline{q} = q' = 0$ at this value of latitude and that the spherical scaling in the equations approximate those on a Cartesian co-ordinate system.

As stated previously, an antisymmetric solution in ψ_b enforces a symmetric solution in h_b and there is no need to enforce $\frac{\partial h_b}{\partial \theta} = 0$. Instead mass conservation is enforced, so that the balanced height field represents the same mass as the full height increment. This is achieved by adding an additional equation to the system, namely

$$\frac{1}{a^2}\sum_{i=1}^{N} V_i \mathcal{H}_i = \frac{1}{a^2}\sum_{i=1}^{N}\sum_{j=1}^{M} \frac{V_i h'_{i,j}}{M} \tag{6.23}$$

which enforces the discrete approximation to the mass consevation where V_i as before by (6.9). A discrete approximation to the mass and not the exact continuous value is used for reasons of consistency. In fact, the total mass of the increment h' is set to zero, leaving the right hand side of equation (6.23) equal to zero.

The coupled system is solved using 2^{nd} order centered differences and 4^{th} order centered differences for the equation at the equator. If 2^{nd} order differences are used throughout, spurious linear solutions are obtained about the equator. The coupled system is described by

$$\begin{aligned}
S_{1,1}\mathbf{x}_1 + S_{1,2}\mathbf{x}_2 &= \mathbf{b}_1, \\
\text{for } \; i = 2, \ldots, \frac{N-3}{2}, \qquad S_{i,i-1}\mathbf{x}_{i-1} + S_{i,i}\mathbf{x}_i + S_{i,i+1}\mathbf{x}_{i+1} &= \mathbf{b}_i, \\
S_{\frac{N-1}{2},\frac{N-3}{2}}\mathbf{x}_{i-1} + S_{\frac{N-1}{2},\frac{N-1}{2}}\mathbf{x}_{\frac{N-1}{2}} + A_{\frac{N-1}{2},\frac{N+1}{2}}\mathbf{y} &= \mathbf{b}_{\frac{N-1}{2}}, \\
\left(\sum_{i=1}^{i=\frac{N-1}{2}} D_{\frac{N+1}{2},i}\mathbf{x}_i\right) + D_{\frac{N+1}{2},\frac{N+1}{2}}\mathbf{x}_{\frac{N+1}{2},\frac{N+1}{2}} + \left(\sum_{i=\frac{N+3}{2}}^{i=N} D_{\frac{N+1}{2},i}\mathbf{x}_i\right) &= \mathbf{o}, \\
A_{\frac{N+3}{2},\frac{N+1}{2}}\mathbf{y} + S_{\frac{N+3}{2},\frac{N+3}{2}}\mathbf{x}_{\frac{N+3}{2}} + S_{\frac{N+3}{2},\frac{N+5}{2}}\mathbf{x}_{\frac{N+5}{2}} &= \mathbf{b}_{\frac{N+5}{2}}, \\
\text{for } \; i = \frac{N+5}{2}, \ldots, N-1, \qquad S_{i,i-1}\mathbf{x}_{i-1} + S_{i,i}\mathbf{x}_i + S_{i,i+1}\mathbf{x}_{i+1} &= \mathbf{b}_i, \\
S_{N,N-1}\mathbf{x}_{N-1} + S_{N,N}\mathbf{x}_N &= \mathbf{b}_N. \quad (6.24)
\end{aligned}$$

where each block $S_{i,i-1}$, $S_{i,i}$, $S_{i,i+1}$ is a 2×2 matrix defined by (6.18)-(6.19) for $i = 2, \ldots, N-1$.

The discretisation at the poles is given using the integral method first used by Barros [2], resulting in at the north pole

$$S_{1,1} = \begin{pmatrix} \frac{\cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} & -\frac{\sin(\frac{\pi}{2} - \frac{\Delta\theta}{2} \cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} \\ \frac{\Delta\theta}{2}\ Q_1 & \frac{\cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} \end{pmatrix} \tag{6.25}$$

$$S_{1,2} = \begin{pmatrix} -\frac{\cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} & \frac{\sin(\frac{\pi}{2} - \frac{\Delta\theta}{2}) \cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} \\ 0 & -\frac{\cos(\frac{\pi}{2} - \frac{\Delta\theta}{2})}{\Delta\theta} \end{pmatrix} \tag{6.26}$$

with

$$b_1 = \begin{pmatrix} 0 \\ \Omega a^2 \Delta\theta \cos\left(\frac{\pi}{2} - \frac{\Delta\theta}{2}\right) \end{pmatrix}. \tag{6.27}$$

At the south pole

$$S_{N,N} = \begin{pmatrix} \frac{\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta} & -\frac{\sin(-\frac{\pi}{2}+\frac{\Delta\theta}{2})\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta} \\ \frac{\Delta\theta}{2}\ Q_N & \frac{\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta} \end{pmatrix}, \tag{6.28}$$

$$S_{N,N-1} = \begin{pmatrix} -\frac{\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta} & \frac{\sin(-\frac{\pi}{2}+\frac{\Delta\theta}{2})\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta}) \\ 0 & -\frac{\cos(-\frac{\pi}{2}+\frac{\Delta\theta}{2})}{\Delta\theta} \end{pmatrix}, \tag{6.29}$$

$$b_N = \begin{pmatrix} 0 \\ \Omega a^2 \Delta\theta \cos\left(\frac{-\pi}{2}+\frac{\Delta\theta}{2}\right) \end{pmatrix}. \tag{6.30}$$

The discretisation at the equator is given by

$$\mathbf{x}_{\frac{N+1}{2},\frac{N+1}{2}} = \begin{pmatrix} \widetilde{\mathcal{H}}_{\frac{N+1}{2}} \\ \widetilde{\Psi}_{\frac{N+1}{2}} \end{pmatrix} \qquad \mathbf{o} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.31}$$

with the 2×2 matrix blocks $D_{\frac{N+1}{2},i}$ equal to

$$\begin{aligned} D_{\frac{N+1}{2},i} &= \begin{pmatrix} 0 & 0 \\ \frac{V_i}{a^2} & 0 \end{pmatrix} \quad \text{for } i = 1, \ldots, \frac{N-5}{2} \\ & \qquad\qquad\qquad\quad \text{for } i = \frac{N+7}{2}, \ldots, N \end{aligned} \tag{6.32}$$

and

$$\begin{aligned} D_{\frac{N+1}{2},\frac{N-3}{2}} &= \begin{pmatrix} 0 & -\frac{1}{12\Delta\theta} - \frac{1}{12} \\ \frac{V_{\frac{N-3}{2}}}{a^2} & 0 \end{pmatrix} \quad D_{\frac{N+1}{2},\frac{N-1}{2}} = \begin{pmatrix} \frac{1}{\Delta\theta} & +\frac{16}{12\Delta\theta} + \frac{8}{12} \\ \frac{V_{\frac{N-1}{2}}}{a^2} & 0 \end{pmatrix} \\ D_{\frac{N+1}{2},\frac{N+1}{2}} &= \begin{pmatrix} -\frac{2}{\Delta\theta} & -\frac{30}{12\Delta\theta} \\ V_{\frac{N+1}{2}} & 0 \end{pmatrix} \\ D_{\frac{N+1}{2},\frac{N+3}{2}} &= \begin{pmatrix} \frac{1}{\Delta\theta} & +\frac{16}{12\Delta\theta} - \frac{8}{12} \\ \frac{V_{\frac{N+3}{2}}}{a^2} & 0 \end{pmatrix} \quad D_{\frac{N+1}{2},\frac{N+5}{2}} = \begin{pmatrix} 0 & -\frac{1}{12\Delta\theta} + \frac{1}{12} \\ \frac{V_{\frac{N+5}{2}}}{a^2} & 0 \end{pmatrix}. \end{aligned} \tag{6.33}$$

The entries V_i are given by equations (6.9) which together with $\mathcal{H}_i$ enforce mass conservation. The $(1,1)$ entries of submatrices $D_{\frac{N+1}{2},\frac{N-1}{2}}$, $D_{\frac{N+1}{2},\frac{N+1}{2}}$, $D_{\frac{N+1}{2},\frac{N+1}{2}}$holds a 2^{nd} order difference approximation of $\mathcal{H}$ term in equation (6.22). In contrast, the $(1,2)$ entries contain coefficients of a fourth order centred discretisation of the Ψ terms for equation (6.22).

The boundary condition $\Psi = 0$ is enforced by the submatrices A, where

$$A_{\frac{N-1}{2},\frac{N+1}{2}} = \begin{pmatrix} \frac{\cos\theta_{\frac{N-1}{2}+\frac{1}{2}}}{\Delta\theta} & 0 \\ 0 & 0 \end{pmatrix} \qquad A_{\frac{N+3}{2},\frac{N+1}{2}} = \begin{pmatrix} \frac{\cos\theta_{\frac{N+3}{2}-\frac{1}{2}}}{\Delta\theta} & 0 \\ 0 & 0 \end{pmatrix}. \tag{6.34}$$

6.5 Overall Procedure to obtain balanced / unbalanced streamfunction using Coupled System Method

We now give a summary of the overall procedure to obtain the new set of control variables.

Given full height and wind fields on a Arakawa C grid, we calculate the height, winds and potential vorticity linearisation states ($\overline{h}$,$\overline{\mathbf{v}} = (\overline{u}, \overline{v})$, $\overline{q}$) on grid positions coincident with the full height field. The linearisation states are calculated by taking the mean average of the full height and wind fields for each latitudinal ring, averaged again over all D datasets considered. This is written as

$$\begin{aligned} \overline{h}_i &= \frac{1}{D}\sum_{d=1}^{d=D}\left(\frac{1}{M}\sum_{j=1}^{j=M} h_{f\ i,j,d}\right) \qquad (6.35) \\ \overline{u}_i &= \frac{1}{D}\sum_{d=1}^{d=D}\left(\frac{1}{M}\sum_{j=1}^{j=M} u_{f\ i,j,d}\right) \\ \overline{v}_i &= \frac{1}{D}\sum_{d=1}^{d=D}\left(\frac{1}{M}\sum_{j=1}^{j=M} v_{f\ i,j,d}\right) \end{aligned}$$

where $h_{f\ i,j,d}$, $u_{f\ i,j,d}$, $v_{f\ i,j,d}$ are the full height and wind fields h_f, u_f, v_f evaluated at grid position i, j for the d^{th} dataset. Since the Arawaka C grid is staggered, the position of $\overline{v}_i$ will be $\Delta\theta/2$ away from $\overline{h}_i$ and $\overline{u}_i$.

Calculation of the potential vorticity linearisation state $\overline{q}_i$ involves the evaluation of the relative vorticity for positions coincident with $\overline{h}_i$. We achieve this using the discretisation (5.26) applied to the full wind fields for all the datasets, followed by the same averaging process as in (6.35). The Corolis parameter f that is needed is simply $f_i = 2\Omega sin\theta_i$ with $\overline{q}_i$ given by the ratio of the absolute vorticity $f_i+(\mathbf{k}\cdot\nabla\times\overline{\mathbf{u}})_i$ to the height linearisation state $\overline{h}_i$.

Perturbations from the height and wind linearisation states are calculated (h_p, u_p, v_p) for a number of datasets using

$$\begin{aligned}\mathbf{v}_{p\ i,j,d} &= \mathbf{v}_{f\ i,j,d} - \overline{\mathbf{v}}_i \\ h_{p\ i,j,d} &= h_{f\ i,j,d} - \overline{h}_i \end{aligned} \tag{6.36}$$

for each grid point (i, j) and every dataset d. The potential vorticity perturbation needs the evaluation of the relative vorticity perturbations using the discretisation (5.26) with perturbation wind fields (u_p, v_p). All subsequent quantities are calculated on the Arakawa C grid where the height grid points are located. The product of the potential vorticity perturbations and the height linearisation state are evaluated by

$$\mathcal{Q}_{p\ i,j,d} = \overline{h}_i q_{p\ i,j,d} = \mathbf{k}\cdot(\nabla\times\mathbf{v}_{p\ i,j,d}) - \overline{q}_i h_{p\ i,j,d} \tag{6.37}$$

with

$$\overline{q}_i = \frac{f_i + \mathbf{k}\cdot(\nabla\times\overline{\mathbf{v}_i})}{\overline{h}_i} \tag{6.38}$$

where (6.38) and (6.37) are the discrete versions of (2.64) and (2.63). One of these datasets define a base state. The potential vorticity perturbations defining the base state is subtracted from the perturbations defined by the other datasets to give increments and are denoted using a dash.

The change in control variables needs $\mathcal{Q}'$ and $\overline{q}$ in order to formulate the appropriate right hand side and variable coefficient needed to solve (6.12), (6.13). We apply the DFFT to $\mathcal{Q}'$ to produce the complex coefficients $\tilde{\mathcal{Q}}'$. We rescale the right hand side and $\overline{q}$ to give $2\Omega a^2 cos\theta_i \Delta\theta\tilde{\mathcal{Q}}'$, Q and solve the second order ODE for all wavenumbers using the discretisation presented in Section 6.4. We apply the DIFFT algorithm to $\tilde{\Psi}_i$ and $\widetilde{\mathcal{H}}_i$ and rescale to give h_b and ψ_b.

Three control variables are either chosen from or derived from ψ_b, ψ_{ub}, h_b, h_{ub} and the velocity potential χ. The calculation of the velocity potential of the wind increments (u', v') is described comprehensively in Section 5.5 and Section 5.5.2. The unbalanced height h_{ub} is defined as the difference between h' and h_b.

6.6 Validation Tests

We first apply a 1 dimensional problem to test this coupled system. The base states are chosen to be

$$\begin{aligned} \overline{h} &= \frac{1}{g} \\ \overline{\psi} &= 0 \\ \overline{q} &= 2g\omega \sin\theta \end{aligned} \tag{6.39}$$

with the scaled potential vorticity increment being,

$$q' = -\frac{2g^2 \sin\theta}{a^2} + g^2\omega^2 \sin\theta \left(\cos 2\theta - \frac{1}{3}\right) \tag{6.40}$$

so that the analytic balanced streamfunction ψ_b and balanced height are given by

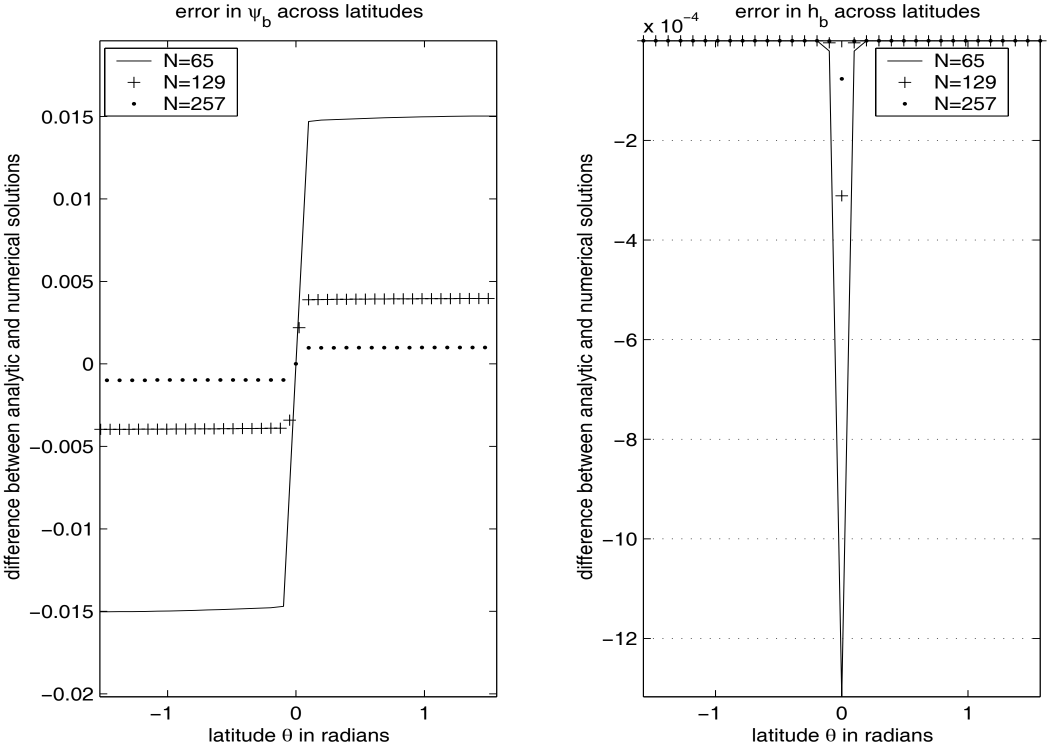
$$\begin{aligned}\psi_b &= g\sin\theta \\ h_b &= -\frac{\omega\cos 2\theta}{2}+\frac{\omega}{6}.\end{aligned} \tag{6.41}$$

This is a good problem which examines the longitudinally independent part of a general solution to the 2 dimensional problem, for the Fourier coefficents relating to $k = 0$.

In Figure 6.1 we show the error profile between the analytic solutions and the experimental results, when $(N = 33, 65, 129)$ across latitudes. The error in the balanced height is slightly larger at the equator than elsewhere. This is expected due to the need to solve equation (6.22) with finite differencing (6.33). In contrast, the error in the streamfunction is the most at the poles. The decrease in error with increasing resolution moves to 2^{nd} order accuracy as the resolution is increased, as shown in Tables 6.1, 6.2. Table 6.1 shows the decrease in error in the balanced height and streamfunction, where the error for each M considered, is taken as the L_2 vector norm of the error at each latitudinal point, divided by M. The order of accuracy is given in Table 6.2.

The 'coupled system' method is tested using full height and wind fields, $h, \mathbf{v}$, which satisfy a Rossby Haurwitz wave (3.36), (3.34) with parameters $R = 4$, $K = \omega = 7.848 \times 10^{-7}s^{-1}$ and $h_0 = 8000m$. The linearisation states are defined to be about a resting state $\overline{\mathbf{v}} = 0$ and a constant height field $\overline{h} = 8000m$. The base state perturbation is also defined to be stationary with a constant height field h_{base} chosen such that the surface integral of this quantity is equal to the surface integral of $h - \overline{h}$. Thus the full increment is defined by $R = 4$, $K = \omega = 7.848 \times 10^{-7}s^{-1}$ and

Figure 6.1: Error in balanced streamfunction and height across various latitudes and resolutions



N	normalised L_2 error of ψ_b,	normalised L_2 error of h_b coupled system
65	2.11×10^{-3}	2.58×10^{-4}
129	5.59×10^{-4}	7.07×10^{-5}
257	1.40×10^{-4}	1.76×10^{-5}

Table 6.1: The averaged L_2 error in h_b and ψ_b at different resolutions

Comparision of resolution	Order of convergence of ψ_b	Order of convergence of h_b
$129-65$	1.43	1.69
$257-129$	2.00	2.00

Table 6.2: Order of convergence for h_b and ψ_b at different resolutions

$h_0 = -h_{base}$ and always chosen so that the surface intergal over the sphere is zero.

Figure 6.2 shows the full height and streamfunction increments, for the high Burger regime defined by $h_0 = 8000\ m$. In contrast, the balanced streamfunction and balanced height obtained from the 'coupled system' are given in Figure 6.3 for $N = 65$ and $M = 96$. The full and balanced streamfunction increments are very similar as expected in a high Burger regime.

When $h_0 = 50\ m$ is chosen the increments are in a low Burger regime with $B_u \approx 0.2$ and $R_o \approx 0.01$ around 45° latitude. The balanced height increments presented in Figure 6.5 is similar to the full height increments shown in Figure 6.4, where $\overline{h} = h_0 = 50\ m$ for the full field h, and $h_0 = -h_{base}$ for the increment in height, h'. This is also consistent with theory present in Chapter 3. Further experimental results concerning geostrophic adjustment are presented in Chapter 7.

A number of experiments are needed to verify the methods that have been outlined.

The first of these is to compare a balanced height field, produced by applying the linear balance equation to the coupled system's balanced streamfunction, with a balanced height field derived directly from coupled system. A Poisson equation

Figure 6.2: Full height increment (left) and full streamfunction increment (right) in a Rossby Haurwitz increment $K = 7.848 \times 10^{-7} s^{-1}$, $h_0 = -h_{base}$ and $R = 4$ with linearisation and base states $\overline{h} = 8000\ m$, $\overline{u} = 0$, $\overline{v} = 0$

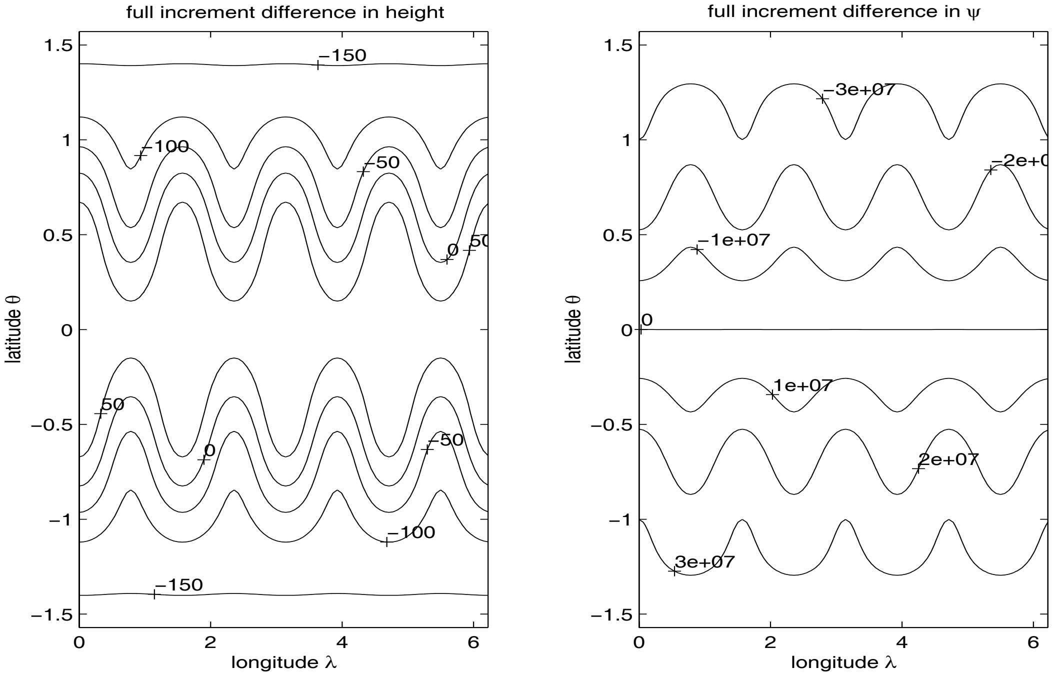


Figure 6.3: Balanced height increment (left) and balanced streamfunction increment (right) for Rossby Haurwitz increment defined by $K = 7.848 \times 10^{-7} s^{-1}$, $h_0 = -h_{base}$ and $R = 4$ with linearisation and base states $\overline{h} = 8000\ m$, $\overline{u} = 0\ ms^{-1}$, $\overline{v} = 0\ ms^{-1}$

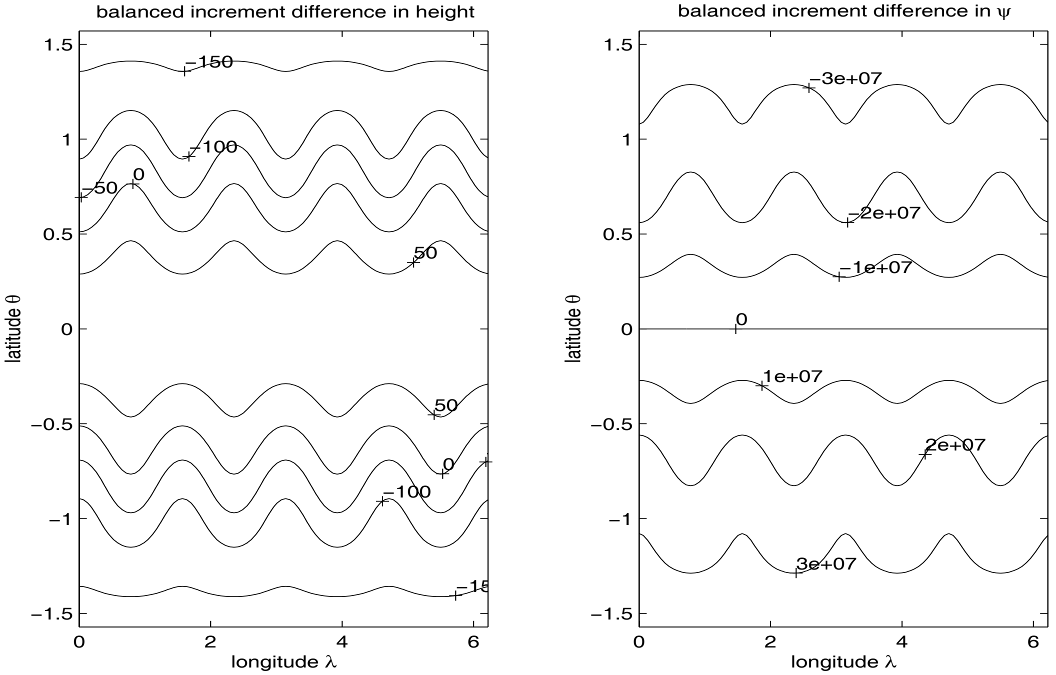


Figure 6.4: Full height increment (left) and full streamfunction increment (right) for Rossby Haurwitz increment defined by $K = 7.848 \times 10^{-7} s^{-1}$, $h_0 = -h_{base}$ and $R = 4$ with linearisation and base states $\overline{h} = 50\ m$, $\overline{u} = 0\ ms^{-1}$, $\overline{v} = 0\ ms^{-1}$

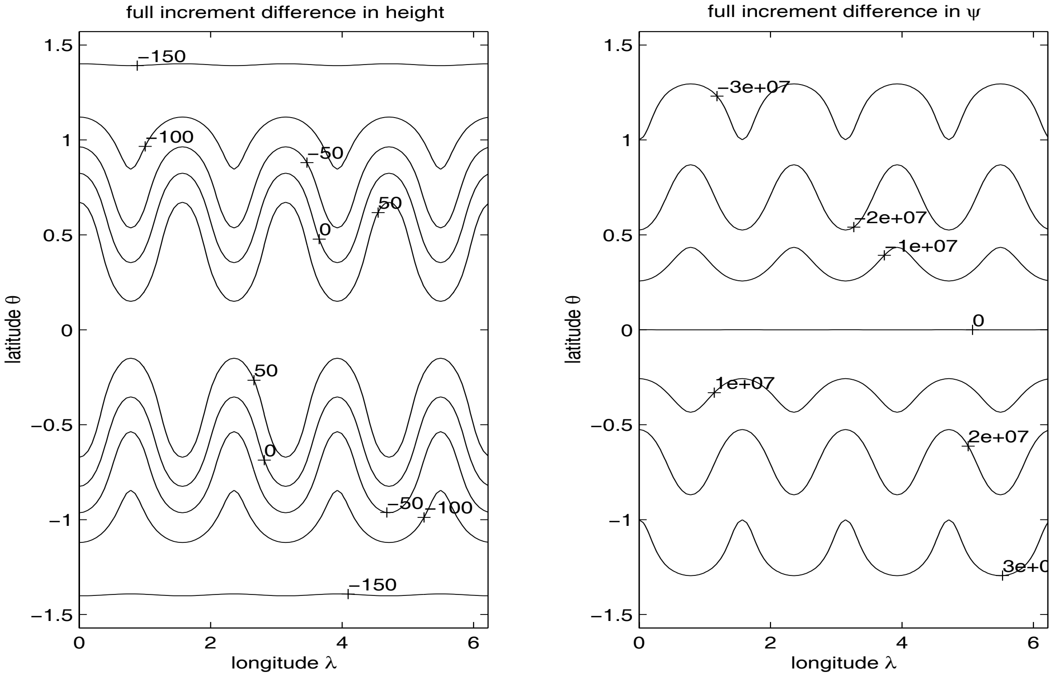


Figure 6.5: Balanced height increment (left) and balanced streamfunction increment (right) for Rossby Haurwitz increment defined by $K = 7.848 \times 10^{-7} s^{-1}$, $h_0 = -h_{base}$ and $R = 4$ with linearisation and base states $\overline{h} = 50\ m$, $\overline{u} = 0\ ms^{-1}$, $\overline{v} = 0\ ms^{-1}$

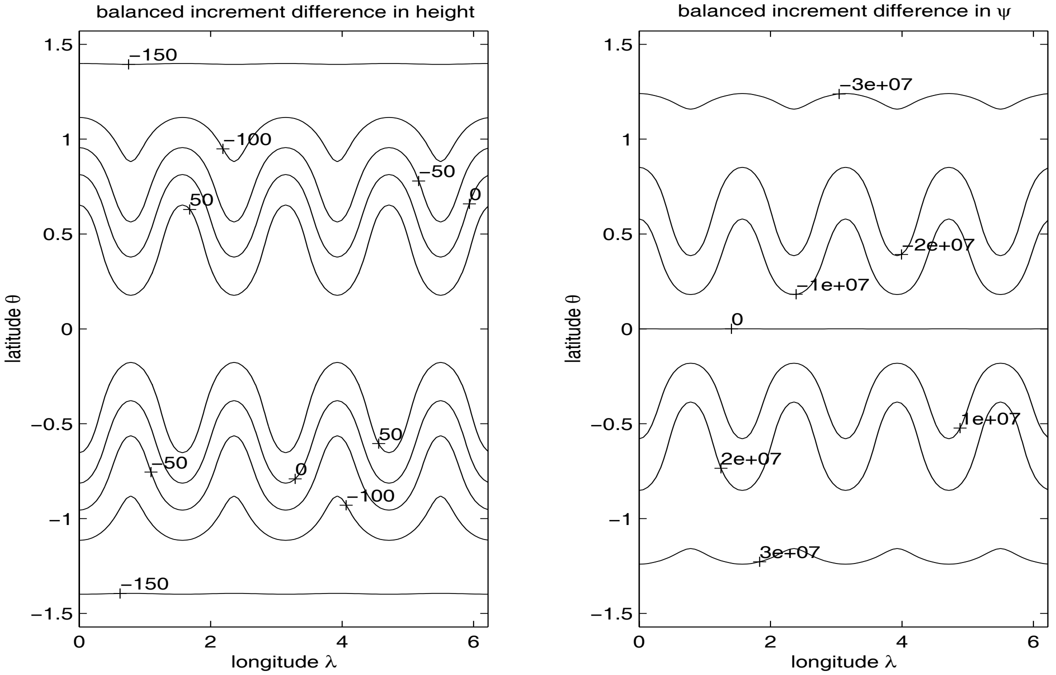
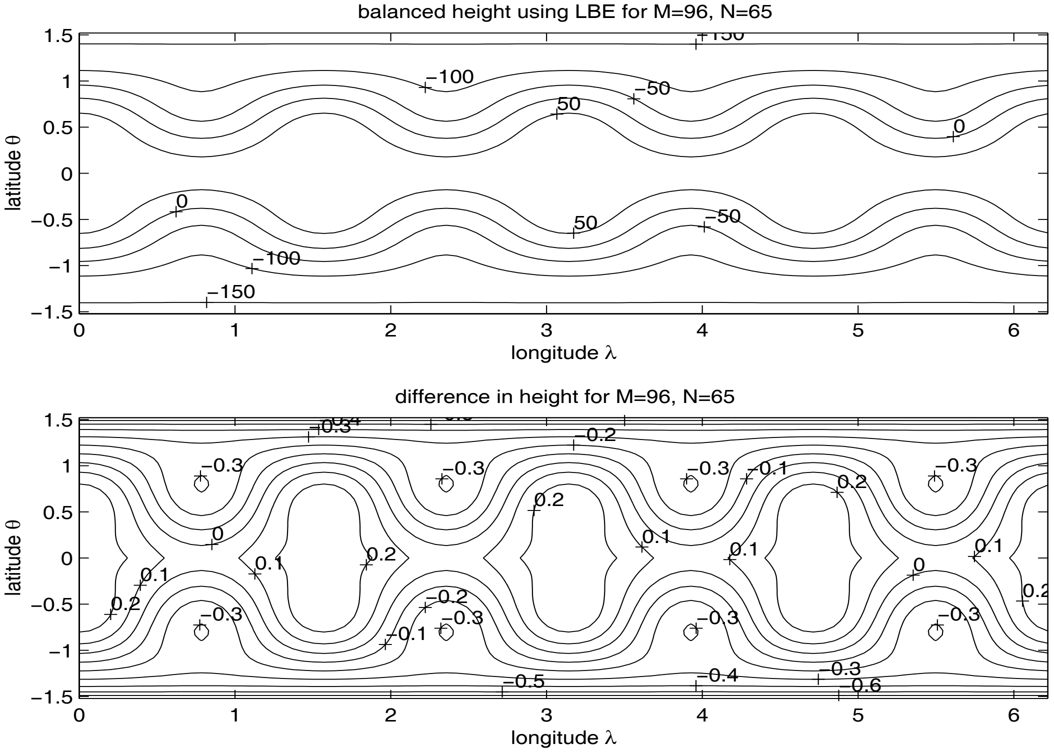


Figure 6.6: Error between balanced height increments derived from LBE using ψ_b from Figure 6.3 and from the coupled system for M=96 and N=65



is needed to be solved to get the balanced height from the balanced streamfunction and is achieved using the method described in Section 5.5. Figure 6.6 shows the error between the two height fields when the streamfunction field, shown in Figure 6.3, is chosen.

Table 6.3 shows the L_2 difference between the two balanced heights at different resolutions. We see a 2^{nd} order decrease in the difference with increased resolution.

The 'coupled system' method can be further checked by computing $\overline{h}q'$ from the balanced streamfunction and height and compare it to the original $\overline{h}q'$ under different grid resolutions. This will confirm that the linearised potential vorticity is being conserved by the balanced streamfunction and height.

Figure 6.7 shows the $\mathcal{Q}'$ calculated from the full increments shown in Figure 6.2

Table 6.3: (top)L_2 error in h_b, divided by $M(N-2)+2$, under different resolutions, (bottom) order of convergence of h_b under different resolutions

N	M	L_2 difference in balanced height using coupled system
33	48	2.81×10^{-1}
65	96	3.20×10^{-2}
129	192	7.27×10^{-3}

N	M	
$65-33$	$96-48$	3.13
$129-65$	$192-96$	2.14

and the difference between this field and $\mathcal{Q}'$ obtained by applying a second order discrete Laplacian operator to ψ_b and subtracting $\overline{q}h_b$. We see that at this resolution the computed error is approximately a hundredth of the value of the original field.

Table 6.6 shows a similar decrease in the difference of these fields with increasing grid resolution. In this table, the discrete integral L_2 norm is used, whose form is given by

$$\mathcal{Q}'_{err} = \sqrt{(}\sum_{i=1}^{i=N}\sum_{j=1}^{j=M}\left(\frac{V_i\mathcal{Q}'_{diff\ i,j}}{M}\right)^2. \tag{6.42}$$

where V_i is defined in equation (6.9), $\mathcal{Q}'_{diff}$ is the difference between the the full increment and balanced increment evaluations of $\mathcal{Q}'$ at grid position i, j.

Figure 6.7: (left) $\mathcal{Q}'$ as calculated from the full increments; (right) error in $\mathcal{Q}'$, derived from full increments and from balanced increments

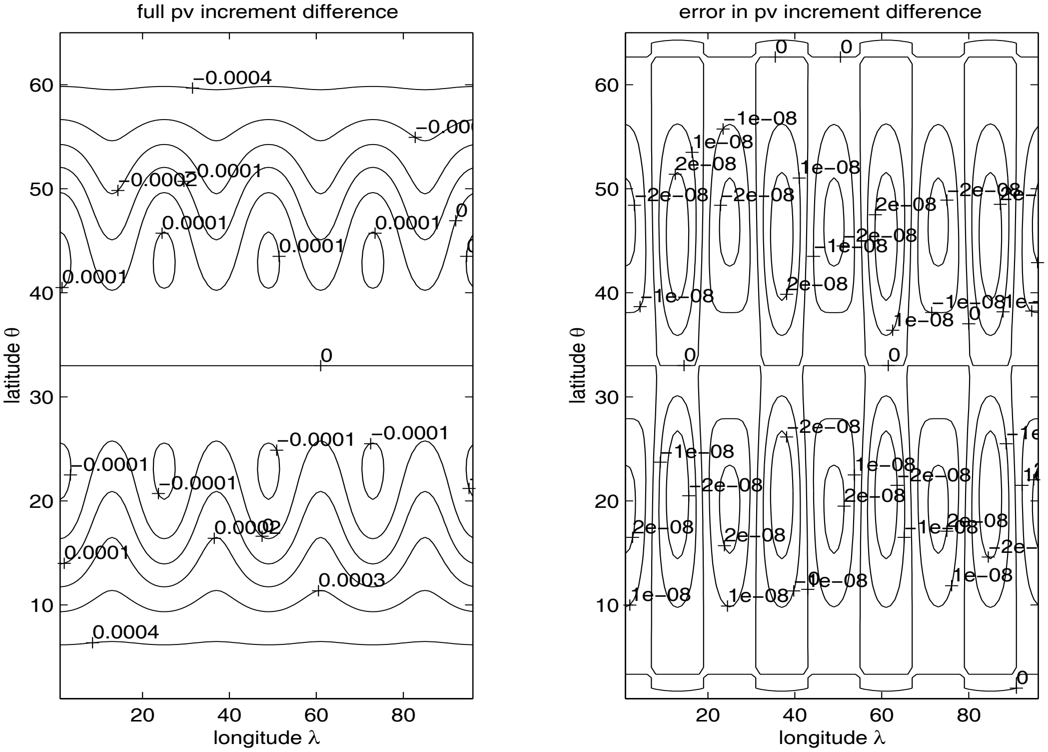


Table 6.4: Integral L_2 error in $\mathcal{Q}'$ under different resolutions

N	M	$\mathcal{Q}'_{err}$
33	48	1.8021×10^{-8}
65	96	2.2665×10^{-9}
129	192	2.8375×10^{-10}

6.7 The inverse transform

In this section we describe the opposite transformation from old and new control variables to the associated wind and height fields. In Chapter 4, various changes into and out of control variables are discussed. The changes out of control variables into height and wind fields use at some point variants of a simple Helmholtz decomposition (5.40). The numerical details of the decomposition are already given in Section 5.6.

A test of internal consistency of the T and U transforms is to calculate ψ_b ψ_{ub} and χ, as defined in Section 4.5 and reconstruct the full wind increments, comparing them to the original values using (5.41). Figure 6.8 shows the approximated component of the winds and the difference between them and the full wind increments. This error decreases with resolution as shown in Table 6.7, and has overall third order accuracy which is surprising as second order accuracy is expected. This increase in accuracy may be consequence of the smoothness of fields being used.

In this chapter we have presented a numerical method to solve a coupled system of equations which gives the balanced variable of the PV method. We have provided tests which validate this numerical technique and are in a position to compare in the following chapter the properties of the LB and PV methods.

Figure 6.8: (upper) Approximated U and V components of the wind calculated by applying the T and P transforms for $M = 96$, $N = 65$; (lower) the difference in the approximated and the true U and V components for $M = 96$, $n = 65$

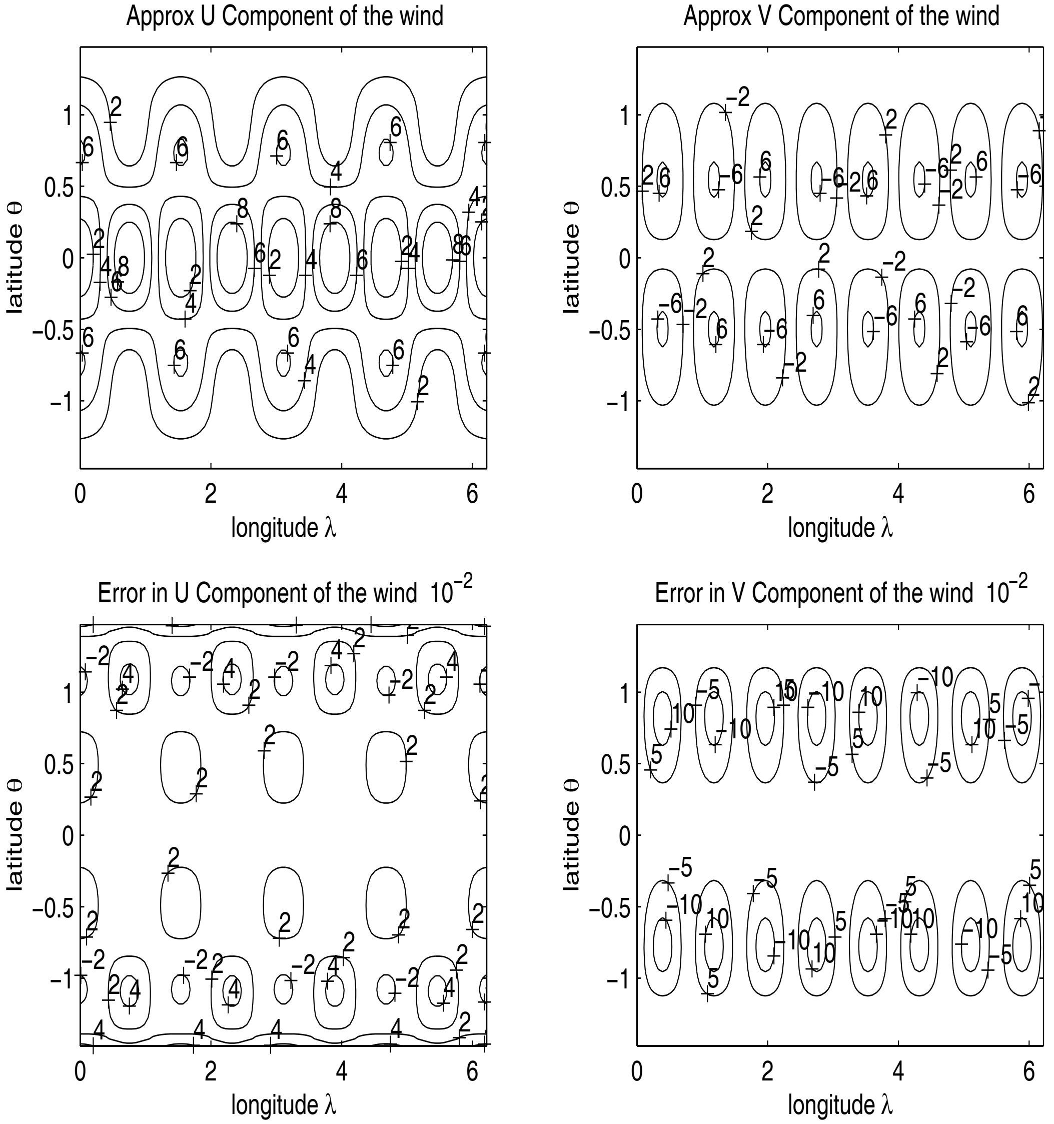


Table 6.5: Discrete L_2 integral error in U and V under different resolutions

N	M	L_2 error in U	L_2 error in V
33	48	2.39×10^{-2}	7.02×10^{-2}
65	96	3.08×10^{-3}	9.07×10^{-3}
129	192	3.89×10^{-4}	1.14×10^{-3}

N	M	Order accuracy in U	Order accuracy in V
$65 - 33$	$96 - 48$	2.96	2.96
$129 - 65$	$192 - 96$	2.99	2.99

Chapter 7

Numerical Results

7.1 Introduction

In this chapter, we provide various experiments for testing the PV method as described in Chapter 4, comparing the results to the LB method. Before this is done we examine the relative sensitivity of potential vorticity, height and absolute vorticity perturbations at different Burger regimes. This is to confirm that for a low Burger regime, the height variable represents the flow. At high Burger regimes the winds dominate and the vorticity is the appropriate variable. For this purpose we consider perturbations to a Rossby–Haurwitz wave in Section 7.2 and subjectively compare the height, potential vorticity and vorticity fields in Section 7.2.1. Section 7.3 shows that the PV method behaves in the way suggested in Sections 3.9 and 4.7. In high Burger regimes it is approximating the solution given by using the LB method. At low Burger number it is acting as if the full height increment/perturbation is determining the balanced flow.

In the final section we examine the divergence tendencies of control variables for both LB and PV methods. We use divergence tendency as an indicator as to how 'balanced' the balanced control variables are. A small value indicates good performance. Two experiments are performed; one with the balanced Rossby Haurwitz wave and the other with unbalanced increments. The experiment with the Rossby–Haurwitz wave shows slightly worse results in the low Burger number. The experiment using unbalanced increments shows that the PV method produces the lowest divergence tendency when applied to the original Burger regime from which the increments originate. In this regime it is doing better than the LB method. However the PV method is performing slightly worse in low Burger regimes. This may be due to the method trying to approximate full unbalanced height increments.

7.2 Experiment 1: Relative Contribution of Height and Absolute Vorticity to Potential Vorticity

We want to show the linearised potential vorticity perturbation captures the dynamical flow over different regimes.

In Section 3.9 we show that the scaled height

$$\vartheta = -\frac{h^*}{\overline{h}} \tag{7.1}$$

and scaled vorticity

$$\iota = \frac{\zeta^*}{\overline{\zeta}} \tag{7.2}$$

contribute to the scaled potential vorticity

$$\kappa = \frac{q^*}{\overline{q}} \tag{7.3}$$

in such a way that is dependent on the Burger number (3.5) : for a high Burger regime, ι contributes most towards κ, while for low Burger regimes the ϑ is the dominant part of κ. Since the theory is carried out on an f-plane, it is prudent to perform an experiment to see whether the theory is satisfied in practice within a more general setting where the Coriolis parameter that varies with latitude.

In order the evaluate ϑ, ι and κ, the linearisation states and the respective perturbations need to be defined. The linearisation states, $\overline{h}$, $\overline{\zeta}$ are defined to be the respective mean average values of the Rossby–Haurwitz wave at a given latitude. They are given by

$$\begin{aligned} \overline{h}_i &= \frac{1}{M} \sum_{j=1}^{j=M} h_{i\ j}, \\ \overline{\zeta}_i &= \frac{1}{M} \sum_{j=1}^{j=M} \zeta_{i\ j}, \quad \text{for} \quad i = 1, \ldots, N, \ j = 1, \cdots, M, \end{aligned} \tag{7.4}$$

For balanced fields this longitudinal averaging is a reasonable approximation to geostrophic balance and is sufficient for the experiment in this section. The absolute vorticity $\zeta_{i\ j}$ is defined spatially on grid points coincident with the height values and is calculated using equation (5.26). As in Chapter 5, the height and wind fields are defined on a Arakawa C grid. The potential vorticity linearisation state is equal to

$$\overline{q}_i = \frac{\overline{\zeta}_i + f_i}{\overline{h}_i}, \quad \text{for} \quad i = 1, \ldots, N. \tag{7.5}$$

The perturbations h^*, ζ^* and q^* are taken to be equal to the maximum departure of the full fields from the respective linearisation states at a given latitude.

The Burger number, as defined by equation (3.5), needs known values for a characteristic height scale and characteristic length scale. At a given latitude the linearisation state $\overline{h}$ is used as the characteristic depth while the characteristic length scale L is defined as half the distance from a peak to a trough.

The full velocity and height fields satisfy the Rossby-Haurwitz wave over a sphere (3.34), (3.35), (3.36). Throughout the experiment, the radius of the sphere a, the acceleration due to gravity g, the angular rotation rate Ω and wave number R are kept constant, with $a = 6371km$, $g = 9.81ms^{-2}$, $\Omega = 7.292 \times 10^{-5}s^{-1}$, $R = 4$.This gives the wave a characteristic length $L = (2\pi\ a\ cos\theta)/(4R)$.

As described in Section 3.4.1, the remaining parameters which define the wave over the sphere are given by h_0, K and ω. We choose these parameters such that we have a wide range of Burger regimes: some that are predominately described by Burger numbers lower than one; others that are completely identified with Burger numbers larger than one and the rest that are described by a mixture of high and low Burger numbers. A high Burger regime is always present around the equator. As one approaches the equator the Coriolis parameter tends to zero and the Burger number tends to infinity. Meanwhile Burger numbers lower than 1 may occur away from the equator as f is larger, provided $\overline{h}$ is small.

We choose K to be either equal to $7.848 \times 10^{-7}\ s^{-1}$ or $7.8480 \times^{-6}\ s^{-1}$ and arbitrarily set $\omega = K$. When $K = \omega = 7.848 \times 10^{-6}\ s^{-1}$ the height at the equator is approximately 2400 m higher than at the poles, while when $K = \omega = 7.848 \times 10^{-7}\ s^{-1}$ the difference in height between the equator and the poles is approximately 300 m. The height field at the poles, h_0, is varied greatly from 50 m to 8000 m. The various Burger regimes for latitudes $\theta = 10°$, $30°$, $45°$, $60°$, $h_0 = 8000\ m, 850\ m, 600\ m, 400\ m, 200\ m, 50\ m$ and $K = 7.848 \times 10^{-7}\ s^{-1}$ are given in Figure 7.1. We see that for the latitudes considered the Burger number stays beneath 1 only when h_0 is equal to 50; at all other values of h_0, the Burger number is larger than 1 for $\theta = 10°$. Conversely when h_0 is equal to 8000 m, the

Burger number is larger than 1 for all latitudes considered. Thus we identify: the wave with parameters $h_0 = 50m$, $K = \omega = 7.848 \times 10^{-7}\ s^{-1}$ as representing a low Burger regime, the wave with parameters $h_0 = 8000\ m$, $K = \omega = 7.848 \times 10^{-7}\ s^{-1}$ representing a high Burger regime, and the rest in between representing a mixture of high and low Burger regimes dependent on latitude.

Figure 7.1 also shows the Burger values at different latitudes when $K = \omega = 7.848 \times 10^{-6}\ s^{-1}$. As for $K = \omega = 7.848 \times 10^{-7}\ s^{-1}$, the Burger number increases with increasing height h_0 and decreasing latitude. The Burger number at $\theta = 10°$ is much larger than at the other latitudes considered. This is due to the Coriolis parameter tending to zero as it approaches the equator. When $h_0 = 8000\ m$, $K = \omega = 7.848 \times 10^{-6}\ s^{-1}$ the Burger number is greater than 1 and a high Burger regime is said to be present. However, for other values of h_0 that we consider there are regions described by Burger numbers larger and smaller than 1. The waves in such cases are said to have a mixture of high and low Burger regimes. Compared to the fields where $K = \omega = 7.848 \times 10^{-7}\ s^{-1}$, these fields have larger regions for which the Burger number is greater than 1.

The scaled perturbations were calculated for all values of K and h_0 and all latitudes considered in the examination of the Burger number in Figures 7.1. In Section 3.9 it is predicted that the influence of the scaled height perturbation on the scaled potential vorticity perturbation decreases with increasing Burger number. This is confirmed with Figures 7.2, 7.3, which compare the relative contributions of scaled absolute vorticity ι and height ϑ to the potential vorticity κ for $K = 7.848 \times 10^{-7}\ s^{-1}$, $7.848 \times 10^{-6}\ s^{-1}$ for different latitudes and h_0. When $h_0 = 8000\ m$ and the latitude is $\theta = 30°$, the Burger number is large and the absolute vorticity

Figure 7.1: Burger values at different latitudes and h_0 when $K = 7.848 \times 10^{-7}\ s^{-1}$ (left), $K = 7.848 \times 10^{-6}\ s^{-1}$ (right)

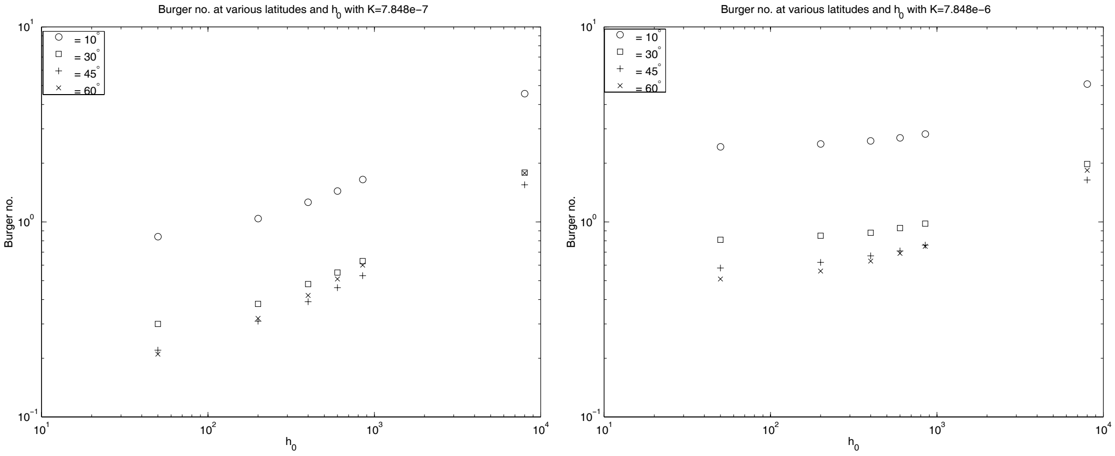
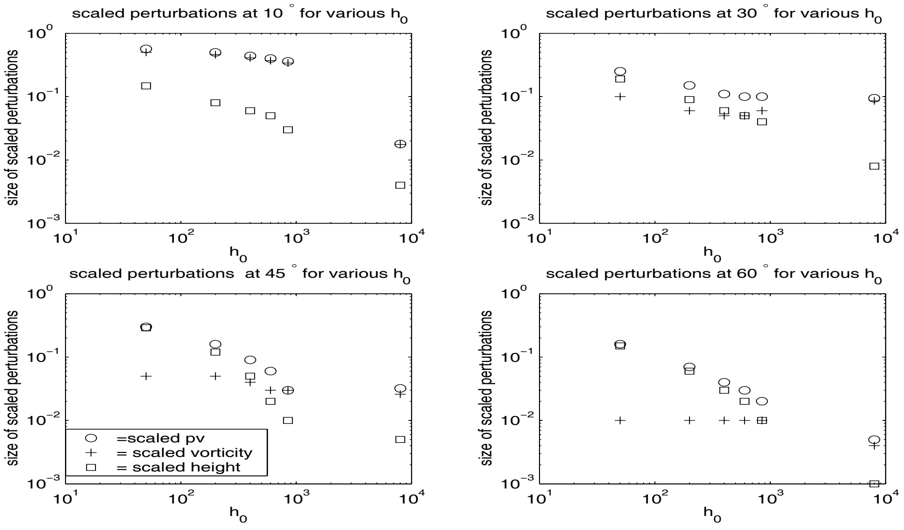


Figure 7.2: Relative contributions of scaled absolute vorticity ι and height ϑ to the potential vorticity κ for $K = 7.848 \times 10^{-7}\ s^{-1}$ with different latitudes and h_0. Sensitivity is defined by the magnitude of the scaled perturbation in question



contribution ι is similar in size to the scaled potential vorticity perturbation κ. Conversely, the height contributes most for low Burger numbers. This is seen when $h_0 = 50$ and $\theta = 60°$.

Also, for the most part, the size of the scaled perturbation in the absolute vorticity ι does not change greatly when h_0 is changed, unlike the scaled perturbation in height ϑ which increases in size with smaller h_0.

7.2.1 Experiment 2: Visual Comparison of Fields

We wish to visually compare the potential vorticity, absolute vorticity and height fields produced from different time integrations of the shallow-water equations over a rotating sphere (2.37), (2.36), with the Rossby-Haurwitz wave acting as an initial condition. We expect that the trends seen in the relative contributions to the potential vorticity for different Burger regimes will also exist in the full fields. For high Burger number, the potential vorticity fields should be similar to the absolute vorticity while at low Burger number the potential vorticity field should resemble the height field. We use the same selection of parameters, h_0, K and ω, as in the previous study. However, only 3 choices of h_0, K and ω are described; the other studied values of h_0, K and ω give results which lie between the three extreme cases picked. The first example, with $h_0 = 8000\ m$, $\omega = K = 7.847 \times 10^{-7}\ s^{-1}$, produces a high Burger regime within this region. The second example, with $h_0 = 50\ m$, $\omega = K = 7.847 \times 10^{-7}\ s^{-1}$, gives a low Burger regime in the mid-latitudes. The final example shows a number of vortices that appear when $h_0 = 8000\ m$, $\omega = K = 7.847 \times 10^{-6}\ s^{-1}$. High Burger regimes will exist within these vortices.

Example 1: $h_0 = 8000\ m$, $\omega = K = 7.847 \times 10^{-7}\ s^{-1}$ **(Figure 7.4)**

In this high Burger regime the height field for the most part keeps the shape of its initial condition; the whole wave just moves at a common angular velocity. The contours of constant potential vorticity vary far less latitudinally than the respective height contours. The phase relationship between the potential vorticity, the absolute vorticity and the height remains approximately constant. The potential vorticity and the absolute vorticity fields are far more similar to each other than the corresponding height field. This is shown explicitly in Figure 7.4 with height, absolute and potential vorticity fields produced by integrating the shallow water equations for 2 days.

Example 2: $h_0 = 50\ m$, $\omega = K = 7.847 \times 10^{-7}\ s^{-1}$ **(Figure 7.5)**

A low Burger regime is produced in the mid-latitude region. In Figure 7.5, the absolute vorticity, height and potential vorticity fields are presented after the model has run for 2 days. We can see that in the mid-latitudes, where the Burger number is low, the potential vorticity has a similar structure to the height. In the equatorial region, where the Burger number is high, the potential vorticity closely resembles the absolute vorticity field.

Example 3: $h_0 = 8000\ m$, $\omega = K = 7.847 \times 10^{-6}\ s^{-1}$ **(Figure 7.6)**

When the value of K is increased, the speed of propagation of the Rossby Haurwitz wave is more influenced by the average zonal height at each latitude; the shallower waves in the mid-latitudes and in the polar regions are moving more slowly than the deeper waves in the equatorial region. This is seen when $K =$

$7.848 \times 10^{-6}\ s^{-1}$. Over the first five days the wave shape becomes increasingly asymmetrical, especially between $\theta = 50°$ and $\theta = 10°$ latitude. At $\theta = 45°$ latitude, the height drops far more steeply to the east of local maxima. On going west from these maxima, the height decreases more gradually. The differences in the propagation in the wave are so acute that lobes are formed in potential and absolute vorticity fields. As the wind shear increases, vortices are created. The steepest gradients are seen to exist around these vorticies in the absolute vorticity fields. The potential vorticity is the next most sensitive field to this vortex. The height field is the least sensitive in that it shows little change in these regions. This is analogous to a shear instability. The production of vortices seems to occur quickly when the base height h_0 is small compared to the height at the equator; it is an effect of an interaction of Burger regimes which are both greater and less than 1. Figure 7.6 shows the vortices that are produced in the mid-latitudes after running the shallow water equations with $h_0 = 8000\ m$, $\omega = K = 7.847 \times 10^{-6}\ s^{-1}$ for 5 days. Within these vortices the Burger number is high as the characteristic length scale is significantly reduced. The potential vorticity more closely resembles the absolute vorticity field than the height field.

7.2.2 Discussion

We have shown that if the potential vorticity perturbation satisfies both the f-plane linear balance equation and the linearised potential vorticity equation, then for high Burger number the scaled absolute vorticity perturbations are similar in magnitude to scaled potential vorticity perturbations. Similarly if the Burger number is small, scaled height perturbations are roughly equal in size to scaled potential vorticity perturbations. Experiment 1 also concurs with this finding. In Experiment 2 we have shown that for small Burger numbers, the height fields contribute more strongly to the potential vorticity. For Burger numbers larger than 1, the greater contribution to the potential vorticity comes from the absolute vorticity, an example of which occurs when $h_0 = 8000\ m$. In this region the potential vorticity mirrors the absolute vorticity field. However there is a loss in detail in these regions. When vortices are produced in the mid-latitudes, they involve sharp changes in velocity and a substantial decrease in the characteristic length scale. The smaller characteristic length scale results in the formation of a high Burger number regime. Again, the absolute vorticity resembles the potential vorticity more accurately than the height. However, the potential vorticity field is less detailed than the absolute vorticity. Hence the theory developed in Section 3.9 for the f-plane appears to hold on the sphere.

7.3 Comparsion of balanced with full fields at high and low Burger number

In order to investigate whether the coupled system of equations provides a better representation of balanced and unbalanced control variables it is necessary to check whether this system has the properties that we expect it to have. We have shown that scaled absolute vorticity perturbations contribute more to potential vorticity perturbations than height perturbations when the Burger number is large and both the linearised potential vorticity equation and geostrophic balance equations are being satisfied. Similarly, we have seen that at low Burger number, scaled

Figure 7.3: Relative contributions of the absolute vorticity ι and height ϑ to the potential vorticity κ for $K = 7.848 \times 10^{-6}\ s^{-1}$ with different latitudes and h_0. Sensitivity is defined by the magnitude of the scaled perturbation in question

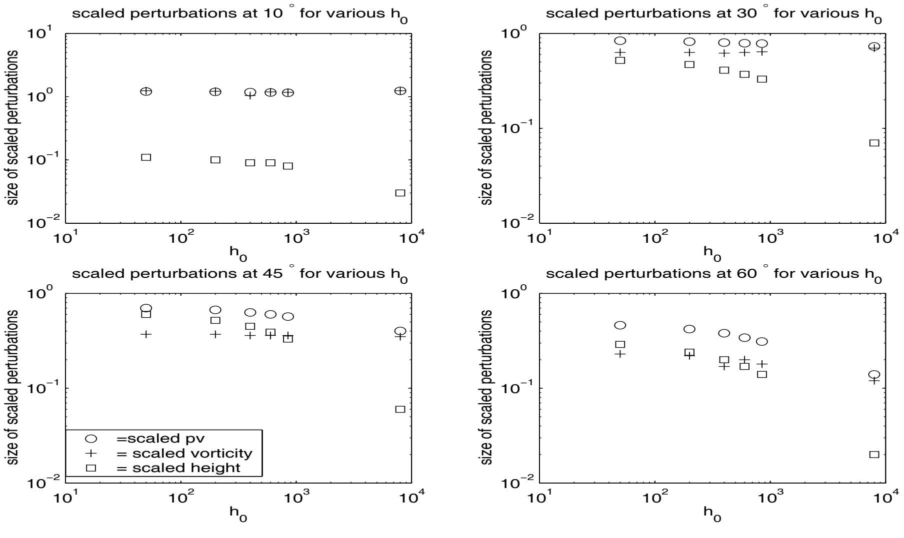


Figure 7.4: Potential vorticity, Absolute vorticity and height fields when $K = 7.848 \times 10^{-7} s^{-1}$ and $h_0 = 8000\ m$ after 2 days

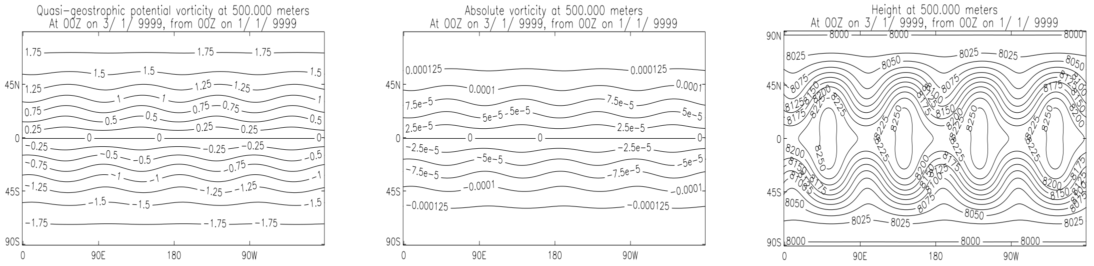


Figure 7.5: Potential vorticity, Absolute vorticity and height fields when $K = 7.848 \times 10^{-7} s^{-1}$ and $h_0 = 50\ m$ after 2 days

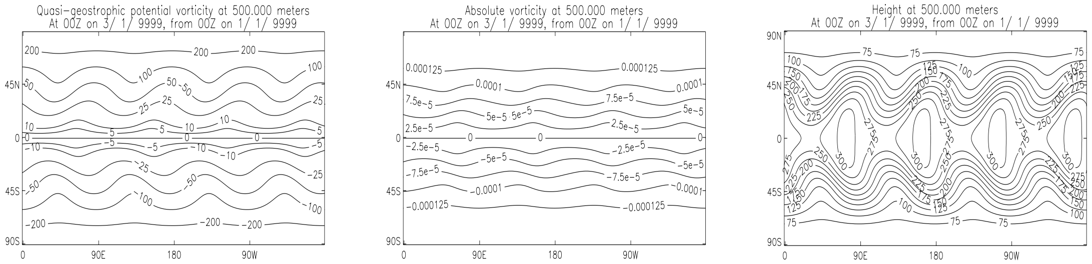
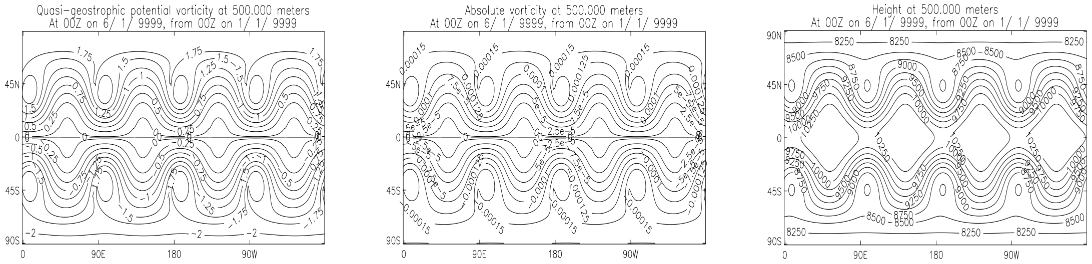
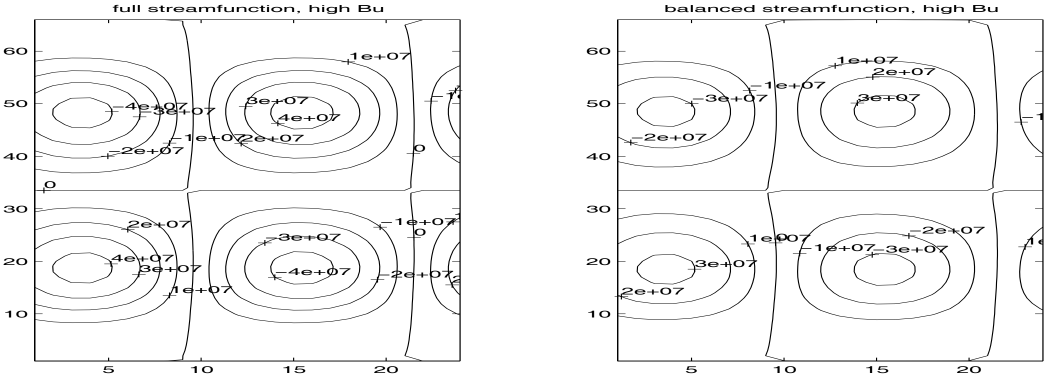


Figure 7.6: Potential vorticity, Absolute vorticity and height fields when $K = 7.848 \times 10^{-6} s^{-1}$ and $h_0 = 8000\ m$ after 5 days



height perturbations contribute more to the potential vorticity perturbations. If the coupled system is behaving properly, then in high Burger regimes the streamfunction ψ should be similar to the balanced streamfunction ψ_b. Similarly, at low Burger regimes the balanced height should resemble the full height field. A Rossby Haurwitz wave (3.34), (3.35), (3.36) (RH wave) is used as an initial condition to a global SWE (2.37), (2.36) model. The defining parameters are $R = 4$, $K = \omega = 7.847e^{-7}s^{-1}, h_0 = 8000m$. Such values produce a high Burger regime across the whole globe, with the Burger number, $B_u = 1.78$ at $60°\theta$, $B_u = 1.55$ at $45°\theta$ and $B_u = 4.55$ at $10°\theta$. The global SWE model was run for 24 hrs, with a timestep of 0.5 hr at medium spatial resolution with grid spacing $\Delta\theta = \pi/64$, and $\Delta\lambda = \pi/48$. The coupled balanced method was used to produce balanced height and streamfunction by applying the procedure detailed in Section 6.5 and excluding the final calculation of increments. Figure 7.7 compares the balanced streamfunction to the respective full field over the area $[(\theta \in [\pi/2, -\pi/2]) \times (\lambda \in [0, \pi/2])]$.

Figure 7.7: Balanced ψ (left) and full ψ (right) for RH wave propagated 1 day at high Burger number, for $(\theta \in [\pi/2, -\pi/2]) \times (\lambda \in [0, \pi/2])$ (scale denotes grid points)



The two fields are qualitatively similar in shape. The balanced streamfunction has an amplitude approximately 80% of the full ψ field. Also the balanced field

Figure 7.8: Balanced height perturbations (left) and full height perturbations (right) for RH wave propagated 1 day at high Burger number, with $(\theta \in [\pi/2, -\pi/2]) \times (\lambda \in [0, \pi/2])$ (scale denotes grid points)

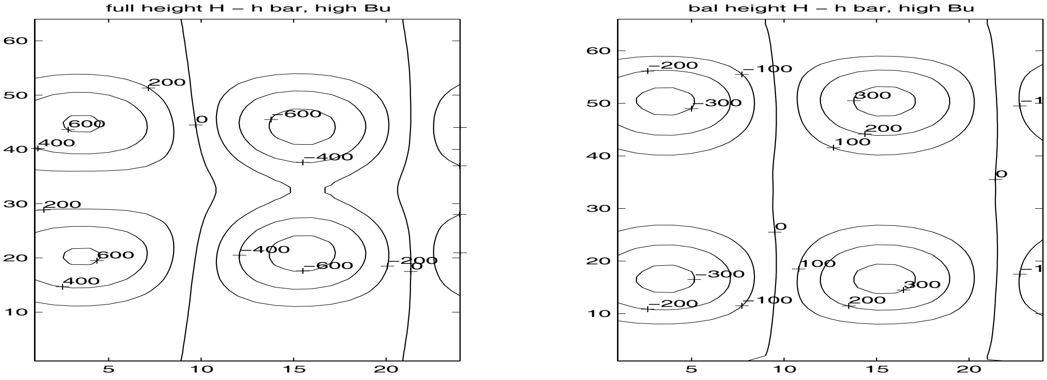
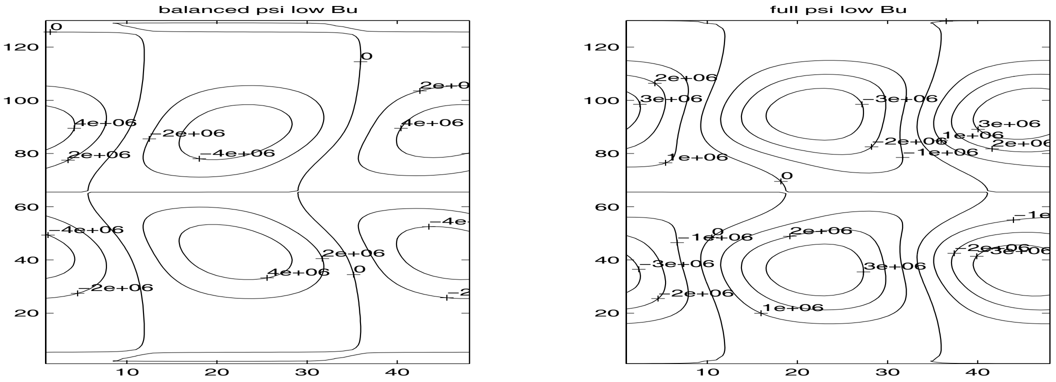


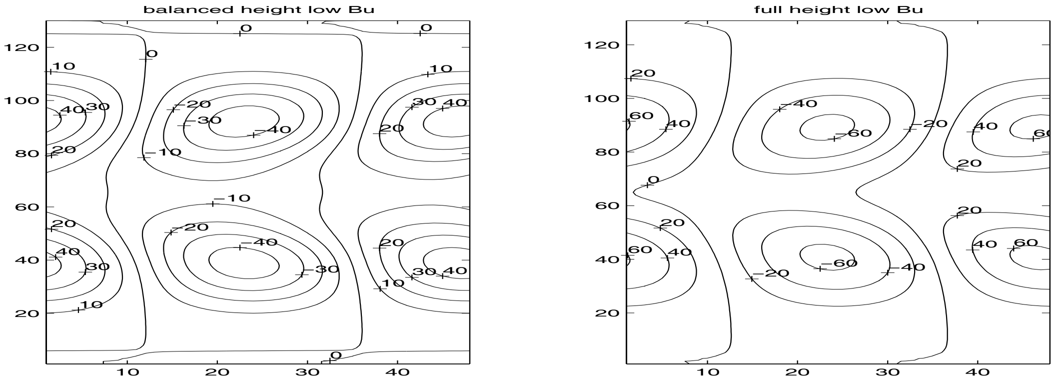
Figure 7.9: Balanced ψ (left) and full ψ (right)for RH wave propagated 1 day at low Burger number, with $(\theta \in [\pi/2, -\pi/2]) \times (\lambda \in [0, \pi/2])$ (scale denotes grid points)



is slightly more diffuse. In comparision, the balanced height and the full height perturbations in Figure 7.8 are notably different in shape.

The above experiment was repeated with RH wave parameters set to: $R = 4$, $K = w = 7.847 \times 10^{-7}\ s^{-1}$ and $h_0 = 50\ m$. This produces a low Burger regime in the mid-latitudes. The value for a Burger number at $\theta = 60°$ is 0.21 with $Bu = 0.22$ at $\theta = 45°$ and $Bu = 0.84$ at $\theta = 10°$. A high spatial resolution is used with $\Delta\theta = \pi/128$, $\Delta\lambda = \pi/192$. Other model parameters were kept the same. As seen in Figure 7.10, regions such as the mid-latitudes where Bu is low the balanced height

Figure 7.10: Balanced height (left) and full height (right) for RH wave propagated 1 day at low Burger number, with $(\theta \in [\pi/2, -\pi/2]) \times (\lambda \in [0, \pi/2])$ (scale denotes grid points)


balanced height low Bu

full height low Bu

perturbations resemble the full height perturbations. The full streamfunction and its respective balanced field are far less similar to each other. This is seen in Figure 7.9.

The Rossby-Haurwitz wave is an idealistic wave to consider; the wave is smooth and is analytically defined. We now consider a more realistic situation by applying the methods to increments derived from a real data set *INI7C*. These are produced by taking the initialised fields used at the UK Met. Office and subtracting them from the corresponding uninitialised field. The uninitialised field is obtained from a spherical harmonic description of the observed fields at T106 resolution from a NETCDF file *VDG7.13.cdf* kept at NCAR. This experiment shows the strengths and weaknesses of the control variables that we have developed.

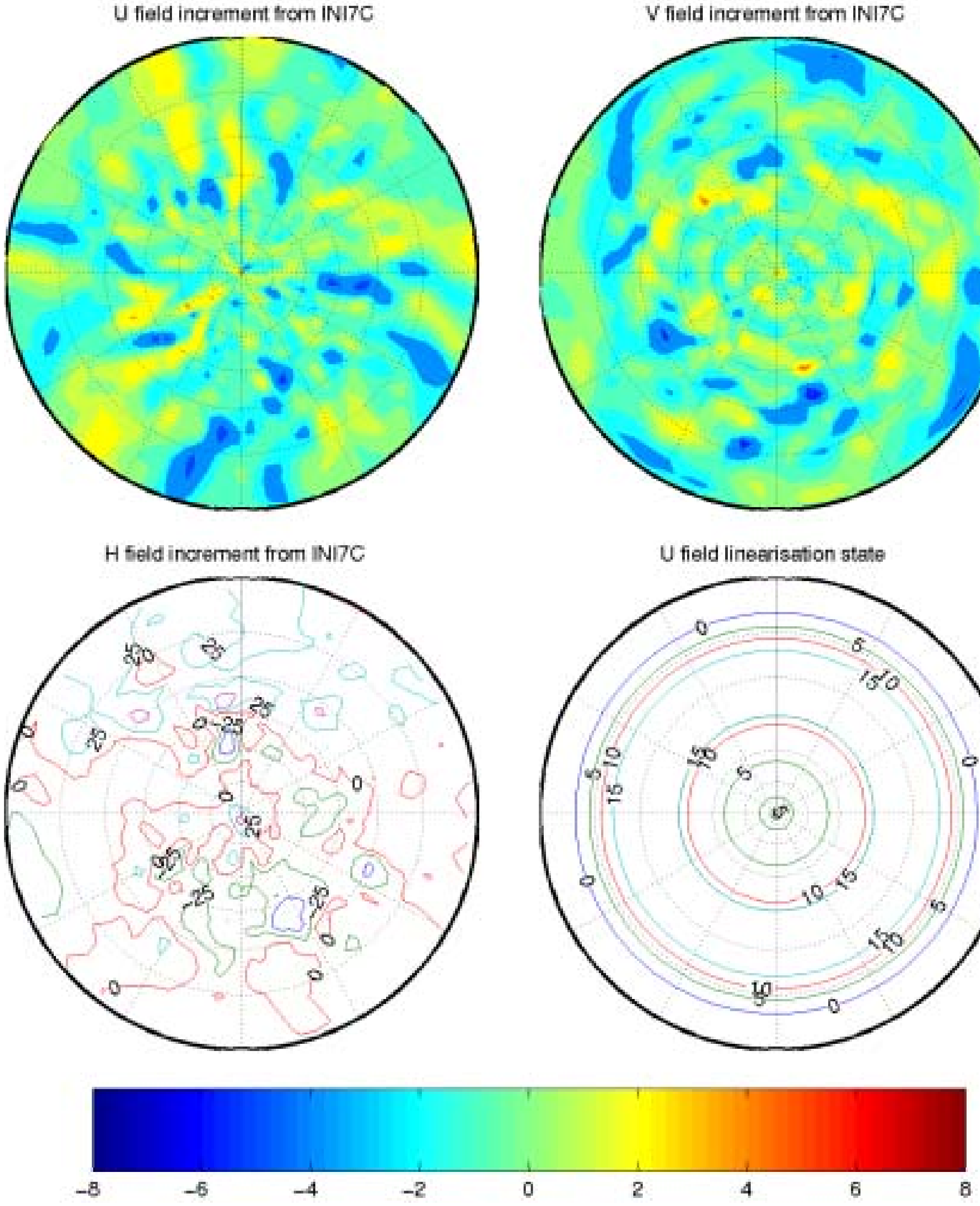
Figure 7.11 shows the height and wind increments used to test the control transforms. A stereographic projection is used centred on the North Pole. The increments are composed of many different waves on a wide range of length scales. The wind increments are typically between $-8ms^{-1}$ and $8ms^{-1}$ and the height increments vary between $-60m$ and $60m$. It is also clear from the figure that there is great variability in the flow with waves of both short and long wavelengths present.

If the initialisation is perfect then the increments consist of just the unbalanced flow. A perfect set of control variables would apportion the flow into the two unbalanced variables. In practice this will not occur. Since the balanced condition used by both sets of control variables holds for the f-plane SWEs linearised about a state of rest, it only can approximate the true balanced and unbalanced parts of the nonlinear flow. In fact, provided that the normal mode initialisation is perfect, the performance of the balanced control variables is determined by their relative size; the balanced control variable that corresponds to the smallest balanced height and wind increments identifies the better set.

The linearisation states are determined using the technique described in Section 5.3.2 under a medium resolution with $\Delta\theta = \pi/64$ and $\Delta\lambda = \pi/48$. The U linearisation state is shown in the bottom right corner of Figure 7.11. It is calculated by applying the PV method to longitudinally averaged uninitialised fields. This method also calculates a height field which is used as a latitudinally varying linearisation state. The linearisation states for low and high Burger regimes are shown in Figure 7.12. The high Burger regime has a $B_u \approx 3.33$ at $45°\theta$. The mean height of the linearisation state is $11km$. The low Burger regime is given by reducing the mean height of the linearisation state to $41m$ at the poles. This gives a Burger number less than 1 above $\theta = 45°$ and makes the sum of the increment and the linearisation states non-negative. The linearisation state of the winds are unchanged.

Figures 7.13 and 7.14 show the balanced height and wind increments produced by the LB and PV methods in the high Burger regime. At a first glance they are

Figure 7.11: (Top) U and V wind increments produced using test case *INI7C*, (bottom right) height increment using test case *INI7C*, (bottom left) U field linearisation state



somewhat surprising. The balanced winds from the two methods are dissimilar. The balanced winds from the PV method are much smaller. This is because the scaled potential vorticity increment $\overline{h}q'$ does not resemble the full vorticity increment ζ'. There is cancellation between ζ' and $\overline{q}h$ as they are of the small magnitude throughout the fields. This makes the scaled linearised potential vorticity increments $\overline{h}q'$ a factor of ten smaller than the vorticity increment ζ'. This also shows that the PV method is performing better than the LB method at producing balanced fields.

To check that the balanced wind produced by the PV method moves towards the balanced wind from the LB method when Burger number is increased, we choose a mean height of $101km$. We see that in Figure 7.16 the balanced winds from the two methods are quantitively similar; the U component of the balanced wind is positive about the pole and swaps direction in the mid-latitudes. In the low Burger regime the balanced wind increments produced by the PV method are more pronounced due to the balanced height increment approaching the full height increment. This is readily seen in Figures 7.15 and 7.13.

The theory presented in Sections 3.9, 4.7 concur with these findings. However they presume that for low Burger regimes the full height increments represents the balance in the system and that the vorticity is the key balanced variable when the Burger number is large. However in this experiment both the full height and wind increments are unbalanced. When the PV method is at very high Burger number, the scaled potential vorticity increment represents the full unbalanced vorticity increment. Conversely, at low Burger regimes the scaled potential vorticity increment resembles the full unbalanced height increment.

It is interesting to note that around the equator, the LB and PV methods are not producing similar results even though a high Burger regime is always present within this region. This is due to the unbalanced height increment still contributing to the scaled potential vorticity within these regions.

In Section 4.6 we propose a means of obtaining balanced corrections to the unbalanced variables. This involves solving a modifield Helmholtz problem given in equation (4.63) to obtain the balanced divergence. It also involves solving a coupled system to obtain the balanced contribution to the departure from linear balance. This coupled system is given by equations (4.69) (4.70) simultaneously. Both are numerically evaluated using a similar technique to those used to solve the Poisson equation, the LBE and the PV method. An Inverse Fourier transform is used zonally and second order centered finite differences are applied to the Fourier components. An Fourier transform is used reconstitute the solutions. The linearisation states are those used as before and are in geostrophic balance. The same grid resolution is used as before. This may not provide a particularly accurate solution as in some regions the linearisation states have the same order of magnitude as the increments. Nevertheless it is good enough to give rough estimates.

We take the L_2 vector norm of the balanced divergence. We see that the LB method produces a value of $5.4times10^{-8}s^{-1}$. The PV method has a smaller value of $3.0 \times 10^{-9}s^{-1}$ at high Burger number. This is due to the much smaller balanced rotational wind in the balanced control variable. In contrast, at low Burger number the PV method produces a slightly larger balanced divergence of $8.7 \times 10^{-8}s^{-1}$ that is of the same order of magnitude as the LB method. This is due to the balanced control variable approximating the full height increment and producing relatively larger winds.

Similar findings are found in the balanced correction to the departure from linear balance. The L_2 vector norm is also used to compare the different increments. We describe this correction in terms of a streamfunction and see that the LB method gives a value of 3.4×10^6 m^2s^{-1} . Again the PV method has a smaller value of 2.3×10^5 m^2s^{-1} at high Burger number. In contrast, at low Burger number the PV method produces a larger value of 1.8×10^7 m^2s^{-1}.

This section supports much of the theory given in Sections 3.9, 4.7. The experiments with RH wave have shown that the balanced height and wind produced by the PV method vary with Burger number as expected. The application of unbalanced increments to both methods shows the importance in how increments are generated. The PV method at low Burger number produces 'balanced' increments that are similar to the full increments, even when the full increments are 'unbalanced'. We have shown that at high Burger number the PV method is performing better than the LB method, with there being far less balanced flow found in the control variables. At low Burger number the balanced divergence obtained from the PV and LB method are of the same order of magnitude.

7.4 Divergence Tendency

The divergence tendency is a good measure with which to compare different sets of control variables. Ideally, we wish the divergence tendency of a set of variables to be small. The first series of experiments we present considers only the divergence tendency of the balanced variables from different control sets. We want to show that the divergence tendencies from new choices of control variables are comparable in magnitude to the present version.

We choose extreme examples of the RH wave at high and low Burger regimes to provide the full fields. The high Burger regime is determined by choosing the defining parameters to be ($h_0 = 8000\ m$, $K = \omega = 7.686 \times 10^{-7}\ s^{-1}$) and giving a Burger number of approximately 1.55 at $\theta = 45°$. The low Burger regime uses ($h_0 = 0\ m$, $K = \omega = 7.686 \times 10^{-7}\ s^{-1}$) and has $B_u \approx 0.30$ at $\theta = 45°$. The RH wave is by definition in Charney balance and has a divergence tendency of zero. When the LB and PV methods are applied to the RH wave, the divergence tendency from the balanced variables is no longer zero but given by (4.53), (4.54). As the winds of the RH wave are rotational the LB method considers the full wind perturbation as balanced. The divergence tendency in this case is just equal to minus the divergence of the advective term of the shallow water momentum equations.

We apply the PV method to calculate perturbations about a resting state and a constant height H using the full height and winds from the high Burger regime which we describe in Section 5.4 and equations (5.2), (5.3). The corresponding low Burger regime is produced by only changing the value of the constant height H. The height and wind perturbations are kept the same as in the high Burger regime. The value of H in the low Burger regime is chosen such that the sum of the height perturbation and H gives a value of zero about the poles. This is done so that the full height field is non-negative. Also keeping the perturbations and the other linearisation states the same allows comparisons to be made easily.

In the high Burger regime the balanced wind perturbations from the PV method approximate the full perturbations. Since the divergence tendency is determined solely from the balanced winds, we expect the divergence tendencies to be similar.

This is clearly seen in Figure 7.17 where the L_2 norm of the divergence tendencies about each latitude ring is given. The results from the LB and PV methods at high Burger number are denoted by circles and crosses, respectively.

In the low Burger regime the divergence tendency from the LB method remains unchanged as the rotational wind perturbation is not varied with Burger number. The results from the PV method do change. The norm of the divergence tendencies of the PV method are noticably worse around the equator in between $\theta = 0°$ and $\theta = 10°$. This may be due to a possible inconsistency between the linearised potential vorticity perturbation and its associated balance condition. This could be rectified by using a balance condition which is more applicable to the tropics. In the mid-latitudes the results from the PV method are mixed. There are regions within the mid-latitudes in which the PV method is performing better. Likewise there are regions where LB method is superior. Overall the PV method may be performing slightly worse than the LB method. However the differences between the two methods are small with the divergence tendencies being of comparable magnitude.

Figure 7.18 presents the L_2 norm of the full height and winds and balanced height and wind fields for the PV method for high Bu and low regimes about each latitude ring. The figure shows that for a low Burger number the balanced height perturbation obtained from the PV method approaches the full height perturbation. At high Burger number the balanced wind perturbations approach full wind perturbations. This is interesting as the PV method uses a linear balance condition while the full height and wind perturbations satisfy the nonlinear Charney balance; the dependence on Burger number of the PV method compensates in part for the simple balance condition used.

So far we have considered only the divergence tendency of balanced variables from LB and PV methods when the initial field field is in Charney balance. We wish to consider increments derived from subtracting the uninitialised fields in data set *VDG7.13.cdf* from the respective initialised fields and calculate overall divergence tendencies for high and low Burger number. In particular we present the L_2 norm of the linearised divergence tendency increments for not only the balanced control variables but also the balanced corrections to the control variables. The linearised divergence tendencies are as defined in equations (4.53),(4.54), (4.71), (4.72). They are approximated using second order centered finite differences. The remaining experimental details are the same as in Section 7.3.

Table 7.1 shows the L_2 vector norm of the linearised divergence tendencies increment of the balanced control variable increments and balanced corrections to unbalanced control variables. We see that the PV method at high Burger number performs the best for balanced control variable increments and balanced corrections to unbalanced control variables. If the PV method is set at even higher Burger numbers, the divergence tendencies would move to those given by the LB method. For low Burger number the PV method gives the divergence tendencies that are significantly worse. This is due to the method approximating the full unbalanced height increments.

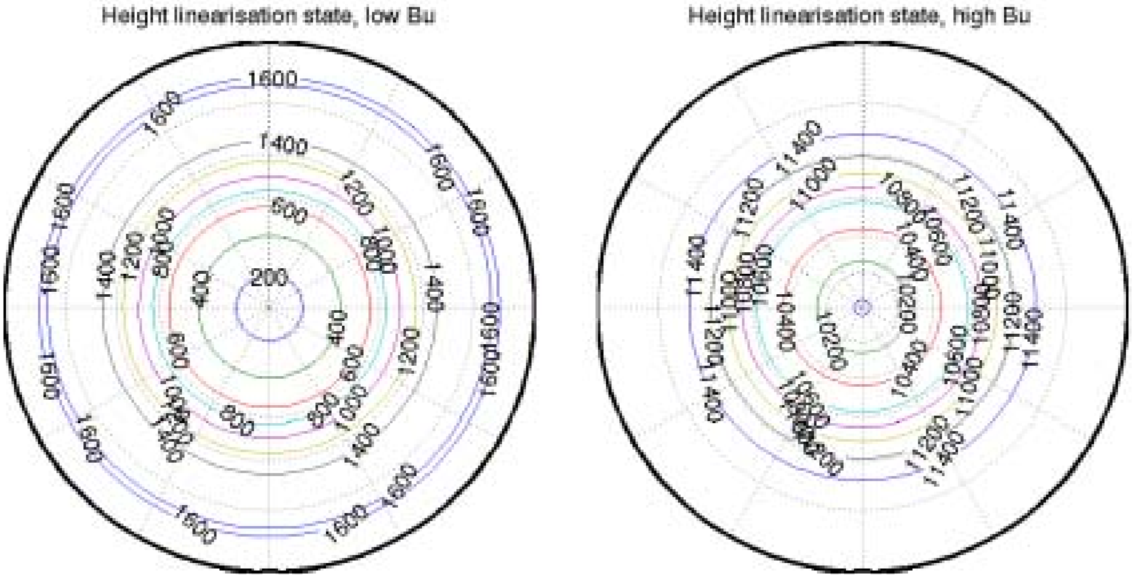
In conclusion, we see that a regime dependent set of control variables given by the PV method gives overall results that are promising. Sections 7.2, 7.2.1 show the relationship between scaled potential vorticity, height and absolute vorticity perturbations at different Burger regimes for a Rossby-Haurwitz wave. Section 7.3 has shown that the PV method behaves as expected when a Rossby–Haurwitz wave is

Table 7.1: L_2 vector norm of the linearised divergence tendencies increment of balanced control variable increments and balanced corrections to unbalanced control variables: (a) L_2 norm of linearised divergence tendency increment from the balanced control variable (m^2s^{-2}); (b)L_2 norm of linearised divergence tendency increment from the balanced divergence increment (m^2s^{-2}); (c) L_2 norm of linearised divergence tendency increment from the balanced correction to the control variable increment describing departure from linear balance (m^2s^{-2})

	(a)	(b)	(c)
LB method	2.0×10^{-10}	3.2×10^{-12}	4.0×10^{-10}
PV method high Bu	1.5×10^{-10}	1.8×10^{-13}	1.3×10^{-10}
PV method low Bu	2.0×10^{-8}	5.7×10^{-12}	8.9×10^{-8}

used; in high Burger regimes it is approximating the solution given by using the LB method and at low Burger number the balanced height is determined by the height perturbation. When unbalanced increments are used in a high Burger regime the PV method performs better than the LB method. We then perform two experiments considering the divergence tendencies of the control variables. The experiment with the RH wave shows slightly worse results in the low Burger number. The experiment using unbalanced increments, shows that the PV method produces the lowest divergence tendency when applied to the original Burger regime from which the increments originate. The PV method is better in this situation at capturing the unbalanced part of the flow. However the PV method works slightly worse than expected when the same increments are introduced into different regime. However it still remains to be seen whether the PV method will perform better with height and wind increments that are mainly balanced and also have a small unbalanced part.

Figure 7.12: H field linearisation states for low Burger regime (left) and high Burger regime (right)


Height linearisation state, low Bu

Height linearisation state, high Bu

Figure 7.13: (Top left) Height increment produced using test case *INI7C* . (Top right) balanced height increment produced by LB method. (Bottom left) Balanced height increment using PV method at low Bu. (Bottom right) Balanced height increment using PV method at high Bu

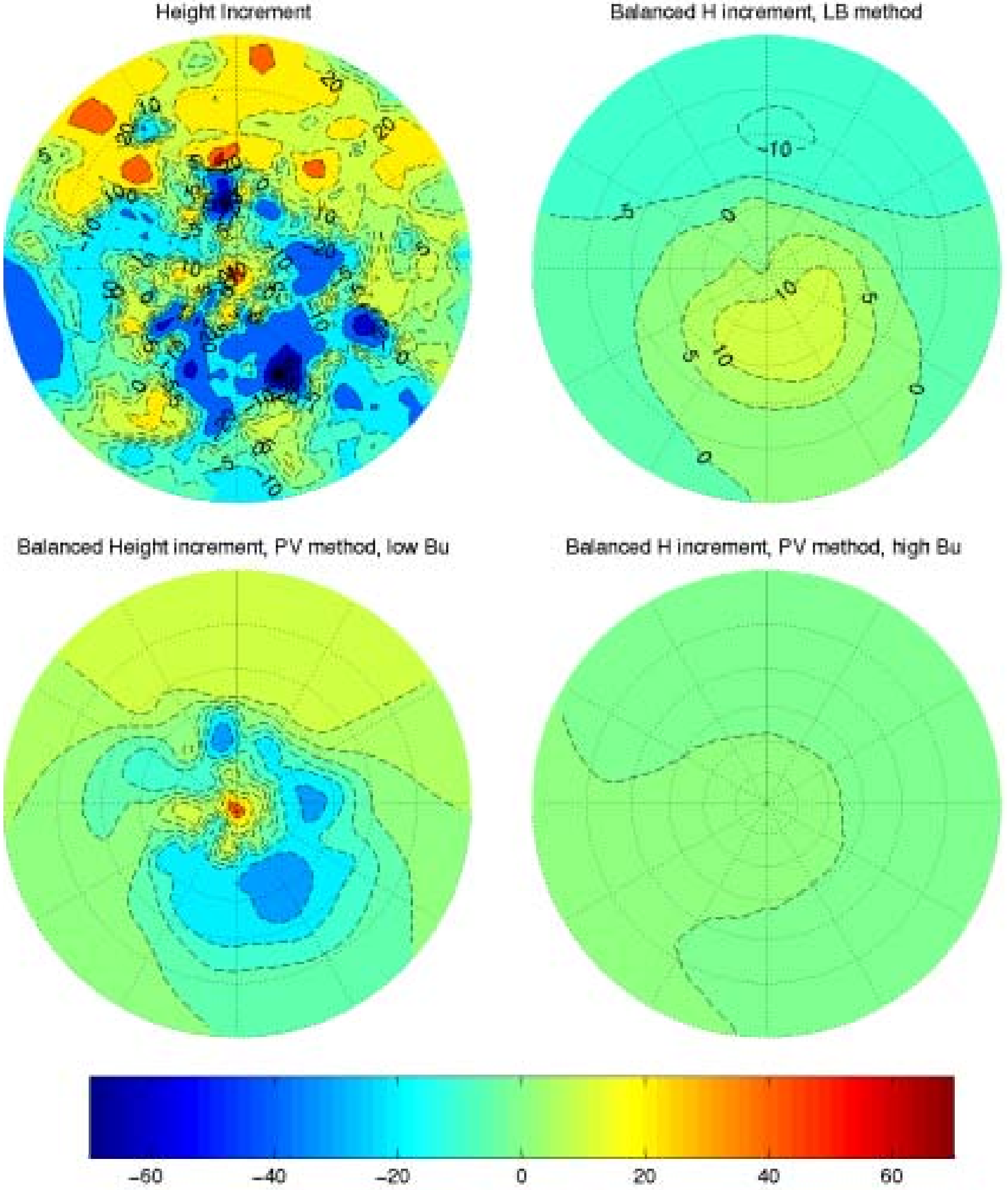


Figure 7.14: Balanced wind increments produced by using the LB and PV methods at high Bu (mean height $H \approx 11km$)

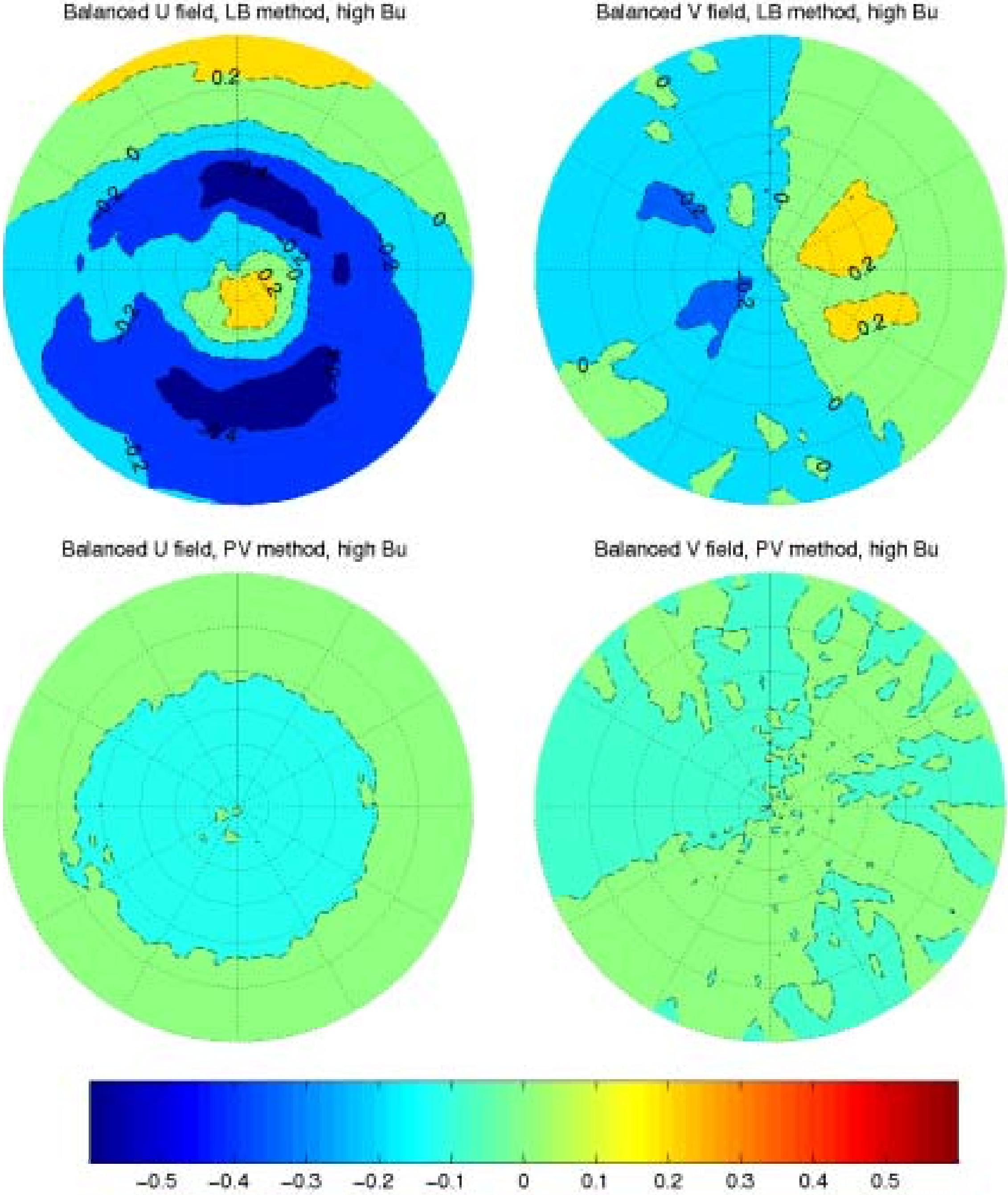
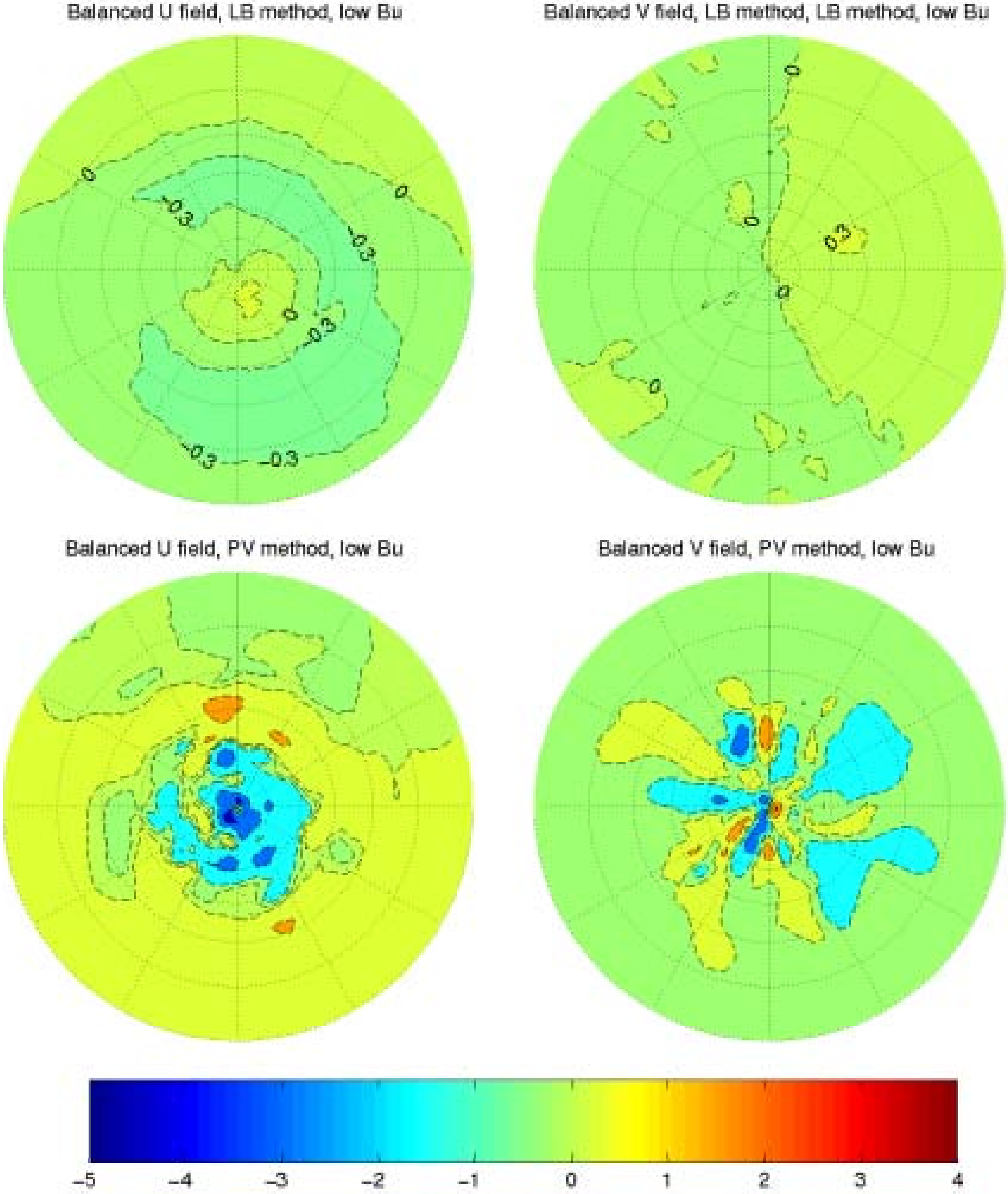


Figure 7.15: Balanced wind increments produced by using the LB and PV methods at low Bu (mean height $H \approx 1km$)

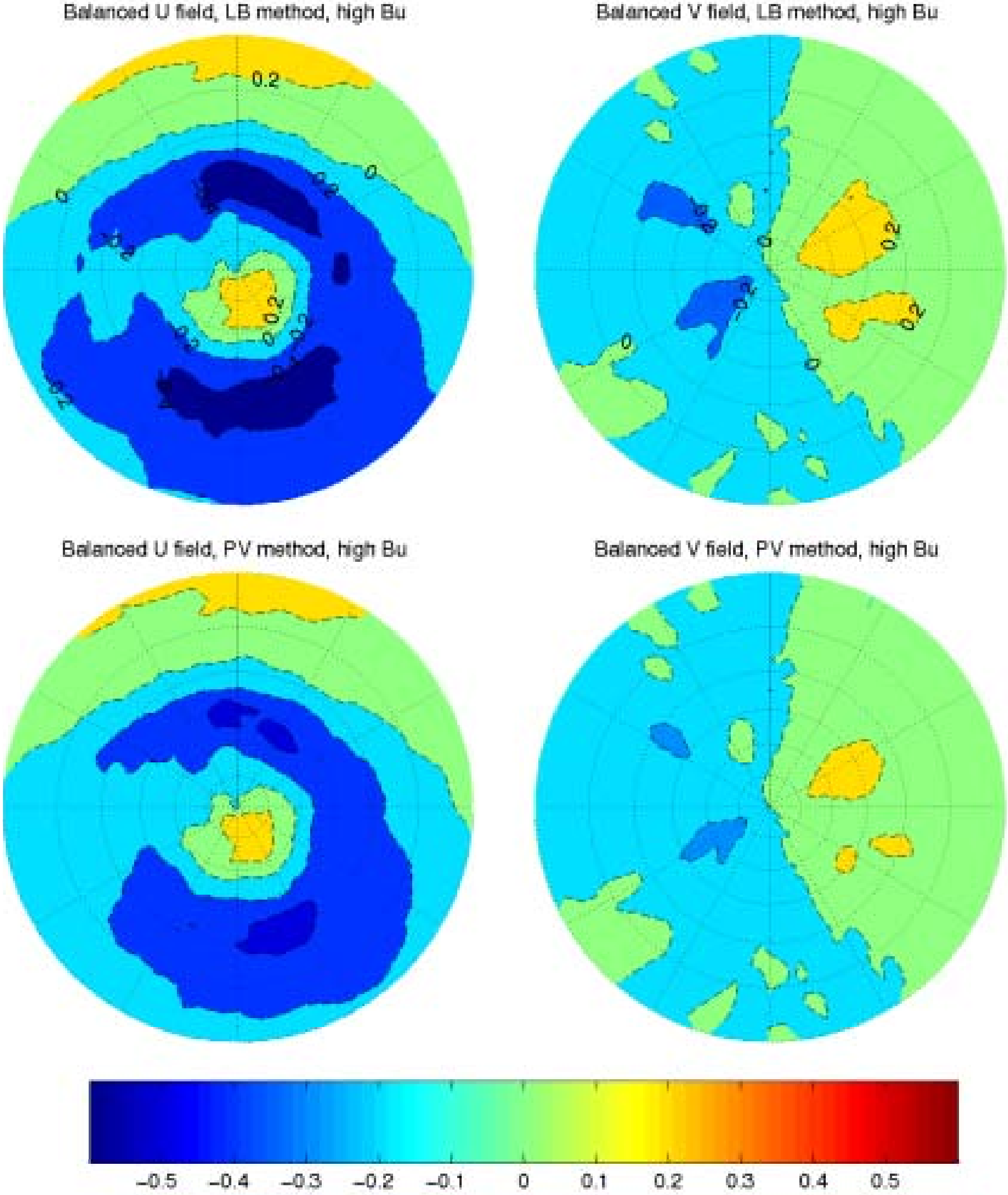

Balanced U field, LB method, low Bu

Balanced V field, LB method, LB method, low Bu

Balanced U field, PV method, low Bu

Balanced V field, PV method, low Bu

Figure 7.16: Balanced wind increments produced by using the LB and PV methods at very hi Bu (mean height $H \approx 100km$), (latitudinally varying linearisation states)



Balanced U field, LB method, high Bu

Balanced V field, LB method, high Bu

Balanced U field, PV method, high Bu

Balanced V field, PV method, high Bu

Figure 7.17: L_2 norm of divergence tendencies RH waves at high and low Bu using PV and LB methods

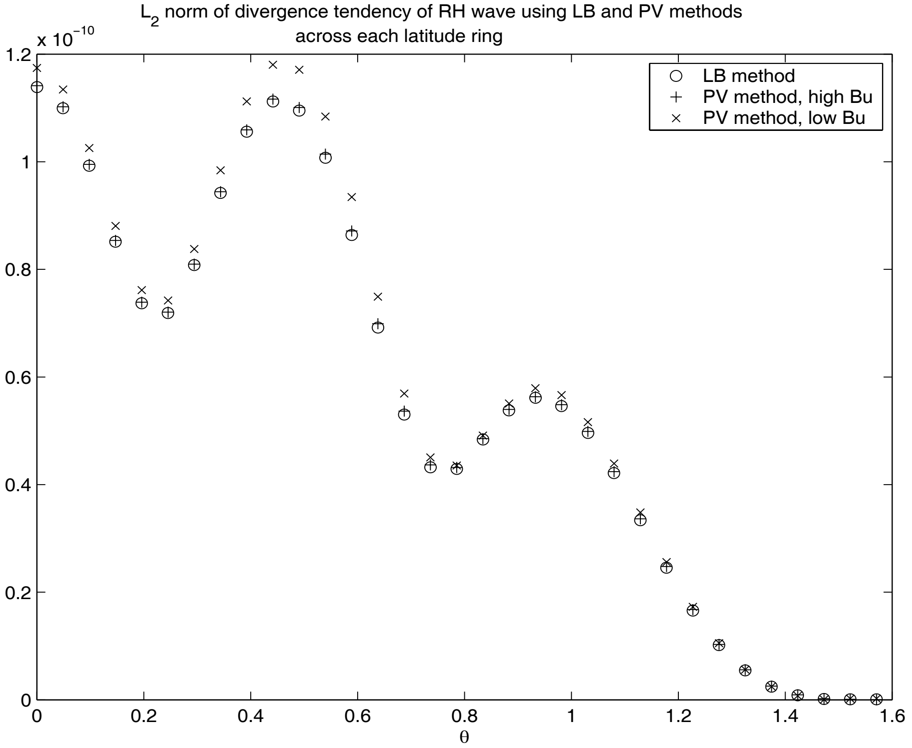
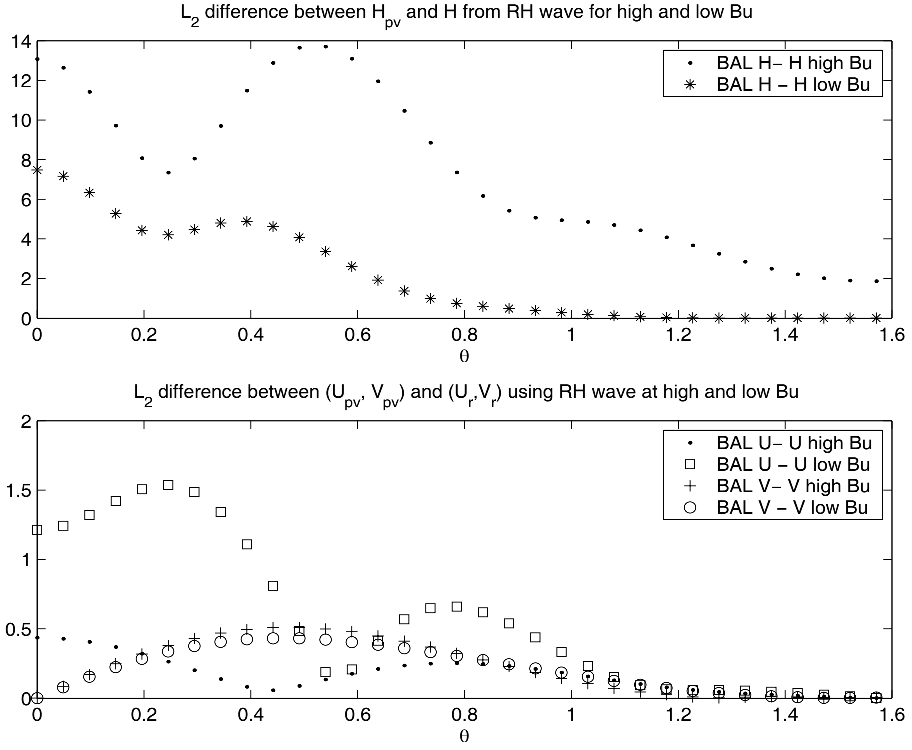


Figure 7.18: L_2 difference between balanced height and winds perturbations from the PV method and full perturbations from RH wave for high and low Bu



Chapter 8

Conclusion

Throughout this thesis we have considered the use of potential vorticity as a control variable. To this end in Chapter 3 we have given a clear mathematical description of balance. A number of issues have come to light that tend to get forgotten. Setting the divergence tendency to zero eliminates the unbalanced inertio-gravity waves only when we consider the f-plane SWEs linearised about a resting state. For more general SWEs an additional condition is needed. Also balance is not just represented by the full rotational wind. At low Burger numbers the height field becomes the balanced variable. Using potential vorticity as the control variable allows the balanced flow to be more accurately portrayed for a wider range of Burger regimes than is done with the present method.

There is only a small body of literature that specifically deals with the problem of control variables in the context of data assimilation within numerical weather prediction. There is no published framework to examine a change of control variables. We present such a framework in Chapter 4 and describe from a dynamical perspective the ideal set of properties that control variables should have. The balanced and unbalanced control variables should mirror the dynamical structure present within the the model concerned. This is a problem in choosing potential vorticity, departure from linear balance and divergence as control variables. Even when the SWEs are considered linearised about a state of rest and on a f-plane, the inertio-gravity waves are coupled in being a linear combination of the eigenmodes of the system. Andy White (personal communication) has suggested that the unbalanced wind components could be chosen to represent the unbalanced control variables. This would represent the the eigenmodes of the system better but would mean that each control variable no longer contributes a portion of the model variables on their own. Nevertheless this choice for the unbalanced variable should still be examined.

The means of identifying approximations of balanced corrections to unbalanced variables presented in Section 4.6 is new. Although Mike Cullen proposed the original idea,[10], [11], the fine detail needed to implement the idea had to be established. The same is true for the PV method. In Chapter 4 we propose various formulations of the idea and then conclude that solving a coupled system simultaneously is the best way to obtain the balanced control variable. We notice in Chapter 6 that provided a latitudinally varying linearisation states are used, the method we adopt is ideal for the task. We use a Fourier method to decompose the zonal component into wavenumber contributions and then apply second order finite differences to each wavenumber component. This numerical technique is a fast, relatively cheap and accurate way to obtain a solution.

In Chapter 4 we also have shown how the control variables produced by the PV method are influenced by Burger number. At high Burger number the LB method and PV method should be similar. In Section 4.7 we have shown that the PV method will approximate a height constrained set of control variables described by (4.73) and (4.74).

The experiments in Chapter 7 for the most part substantiate the theory in the previous chapters. The PV method at high Burger number produce control variables that are similar to those produced by the streamfunction constrained LB method. Similarly at low Burger number the PV method produces control variables in which the full height increments/perturbations dictate the balanced height and wind fields. A difficulty arises in using a linearised potential vorticity increment at very low Burger number. The smaller the height linearisation state, the less accurate the linearisation to the potential vorticity increment becomes. For this reason very low Burger regimes cannot be examined by this technique.

The experiment applying the LB and PV methods to unbalanced increments give interesting results as they complement the studies produced with balanced perturbations. It shows that in low Burger regimes the PV method gives a balanced control variable which is predominantly determined by the full height increment. This occurs even when the height increment is unbalanced. The PV is unable to distinguish between a full balanced increment and a full unbalanced one. From the experiments presented in Section 7.2 we observe that the height perturbations and potential vorticity perturbations hold the most of the information about the balanced flow at low Burger regimes. Similarly wind perturbations and the potential vorticity perturbations represent the balanced part of the flow for high Burger number. The improved performance at high Burger regime for unbalanced increments shows a level of consistency between the PV method and the initialisation procedure used.

8.1 Further Work

More experiments need to be performed in order to be certain that the potential vorticity-based set of control variables is better than the current method. One way would be to apply the method to height and wind increments generated using the NMC method ([41]), ([29]) from a multi-layered barotropic model approximating the atmosphere. As the full increments should be predominantly balanced, the PV method should produce smaller divergence tendencies in the balanced control variable than the LB method.

Obtaining the balanced control variable of the PV method is essentially a 'poor man's' version of the 1st order direct potential vorticity inverter described by McIntyre [36]. Instead of using a Charney balance condition, the LBE is used instead. It would be interesting to use the Charney balance condition instead and compare the results. This should be not difficult to achieve as it would need minor changes to be applied to the code used to produce balanced corrections to the departure from linear balance. A linearisation of the Charney balance equation about latitudinally varying states would be used. The coupled system to be solved on a hemisphere is

$$\begin{aligned} -\nabla\cdot\left(f+\overline{\zeta}\right)\nabla\psi_b-\mathbf{k}\cdot\nabla\times\left(\nabla^2\psi_b\overline{\mathbf{v}}\right) & \\ +g\nabla^2 h_b+\nabla^2\left(\overline{\mathbf{v}}\cdot\left(\mathbf{k}\times\nabla\psi_b\right)\right) &= 0, \\ \nabla^2\psi_b-\overline{q}h_b &= \nabla^2\psi'-\overline{q}h', \end{aligned} \tag{8.1}$$

where the boundary conditions are the same as those used in the PV method.

So far we have considered the control variable transformations on a hemisphere. To consider the control variable transformations in a more realistic context we need to generalise the work for flows whose variables are neither symmetric nor antisymmetric about the equator. If we use the same technique as before, we would need four equations and four variables. In addition to solving for only symmetric height and an antisymmetric streamfunction we would need to evaluate an antisymmetric height field and a symmetric streamfunction. In principle this can be achieved by solving the system

$$\begin{aligned}
\nabla \cdot f\nabla\psi_a - g\nabla^2 h_s &= 0 && (8.2)\\
\nabla^2\psi_a - \overline{q}_a h_s - \overline{q}_s h_a &= \nabla^2\psi'_a - \overline{q}_a h'_s - \overline{q}_s h'_a && (8.3)\\
\nabla \cdot f\nabla\psi_s - g\nabla^2 h_a &= 0 && (8.4)\\
-\overline{q}_a h_s + \nabla^2\psi_s - \overline{q}_s h_a &= \nabla^2\psi'_s - \overline{q}_s h'_s - \overline{q}_a h'_a && (8.5)
\end{aligned}$$

simultaneously, where ψ_a and h_a are the balanced height and streamfunction parts which are antisymmetric about the equator, ψ_s and h_s are the balanced symmetric contributions. The variables h'_s, ψ'_s, h'_a, h'_s are the symmetric and antisymmetric parts of the full height and streamfunction increments such that

$$\begin{aligned}
h' &= h'_a + h'_s\\
\nabla^2\psi' &= \nabla^2\psi'_a + \nabla^2\psi'_s && (8.6)
\end{aligned}$$

Similarly, the balanced streamfunction increment ψ'_b and balanced height increment h' are equal to the sum of symmetric and antisymmetric balanced parts, with

$$\begin{aligned}
h'_b &= h_a + h_s,\\
\mathbf{k} \times \nabla\psi'_b &= \mathbf{k} \times \nabla\psi_a + \mathbf{k} \times \nabla\psi_s. && (8.7)
\end{aligned}$$

The scaled potential vorticity increment $\overline{h}q'$ and linearisation state $\overline{q}$ are similarly split into symmetric and antisymetric parts, with

$$\overline{h}q' = \left(\nabla^2\psi'_a - \overline{q}_a h'_s - \overline{q}_s h'_a\right) + \left(\nabla^2\psi'_s - \overline{q}_s h'_s - \overline{q}_a h'_a\right),$$

$$\overline{q} \quad = \quad \overline{q}_a + \overline{q}_s. \tag{8.8}$$

We set the surface integral of the symmetric part of the scaled potential vorticity increment over the sphere to sum to zero. In addition to Neumann boundary conditions at the poles, and the respective symmmetric and antisymmetric conditions implicitly set by the right hand sides of (8.3) and(8.5), we need two additional conditions. We assume that the total mass of the system held in the increments is equal to the balanced mass of the system. Also we take the total momentum of the rotational part of the balanced flow to be equal to the total momentum held by the rotational part of the full wind and height increments.

The above formulation allows there to be non-zero potential vorticity increments and non-zero values for the balanced streamfunction about the equator. Since we are using a linear balance condition, the solutions will not give realistic results about the equator. Using a linearised Charney balance condition about a latitudinally varying state is an option. The above formulation given in equations (8.2)-(8.8) is easily adapted to include this balance condition. However around the equator the linearisation states in the wind are frequently of the same order of magnitude as the increments and the approximation given by linearisation of the Charney balance condition will be poor. A means around this problem would be to use linearisation states that are a function of latitude and longitude. For this to be done either an iterative method or a spherical harmonic method would be more appropriate.

Around the tropics balanced divergence is an important quantity to diagnose. A means by which this can be achieved is through using a higher order direct potential vorticity inversion. The second order direct potential vorticity inversion sets the divergence tendency and the second order partial time derivative of the balanced divergence to zero. This method gives a time invariant balanced divergence. In the third order direct inversion the second and third order partial time derivatives of the balanced divergence are set to zero. This would allow a time varying balanced divergence to be obtained. The use of this third order inversion would accurately represent key dynamical features present in the tropics. It would also be the limit in which potential vorticity inversion is useful [18].

References

[1] F. Baer. The spectral balance equation. *Tellus*, 29:107–115, 1977.

[2] S.R.M. Barros. Multigrid methods for two- and three-dimensional Poisson-type equations on the sphere. *J. Comput. Phys.*, 92:313–348, 1991.

[3] J.R. Bates, Y. Li, A. Brandt, S.F. McCormick, and Ruge. A global shallow-water numerical model based on the semi-lagrangian advection of potential vorticity. *Quart. J. R. Met. Soc.*, 121:1981–2005, 1995.

[4] D.P. Baumhefner and T.W. Bettge. Characteristics of atmospheric circulation and associated model forecast skill during fgge case studies selected by wgne. In *International Conference on Early Results og FGGE and Larger-Scale Aspects of its Monsoon Experiments*, pages 2–26. WMO,Geneva, 1981.

[5] Onno Bokhove. *On Balanced Models in Geophysical Fluid Dynamics: Slowest Invariant Manifolds, Slaving Principles, and Hamiltonian Structure.* PhD thesis, Physics, University of Toronto, 1996.

[6] Bolin. Numerical forecasting with the barotropic model. *Tellus*, 7:27–49, 1955.

[7] A.P. Burger. Scale consideration of planetary motions of the atmosphere. *Tellus*, 10:195–205, 1958.

[8] J. Charney. The use of the primitive equations of motion in numerical prediction. *Tellus*, 7:22–26, 1955.

[9] J.W. Cooley and J.W. Tukey. An algorithm for the machine computation of the complex Fourier series. *Mathematics of Computation*, 19:297–301, 1965.

[10] M.J.P. Cullen. New mathematical developments in atmosphere and ocean dynamics, and their application to computer simulations. Technical report, UKMO: Forecasting Research, 1997.

[11] M.J.P. Cullen. "approximation of atmospheric flows by balanced states at low Burger numbers. to be published, UKMO, 1999.

[12] R. Daley. Linear non-divergent mass-wind laws on the sphere. *Tellus*, 35A:17–27, 1983.

[13] R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, 1991.

[14] R. Daley. Generation of global multivariate error covariances by singular-value decomposition of the linear balance equation. *Mon. Weather Rev.*, 124:2574–2587, 1996.

[15] P. Duhamel and M. Vetterli. Fast Fourier transforms: A tutorial review and a state of the art. *Signal Processing*, 19:259–299, 1990.

[16] Dutton. *Dynamics of Atmospheric Motion*. Dover, 1986.

[17] Eckart. *Hydrodynamics of Oceans and Atmosphere*. Pergamon Press, 1960.

[18] R. Ford, M.E. McIntyre, and W.A. Norton. Balance and the slow quasimanifold:some explicit results. *Journal of Atmospheric Sciences*, 57(9):1236–1254,

2000.

[19] M. Frigo and S.G. Johnson. FFTW (http://www.fftw.org).

[20] M. Frigo and S.G. Johnson. FFTW: An adaptive software architecture for the FFT. *Proceedings of the International Conference on Acoustic, Speech, and Signal Processing*, 3:1381–1384, 1998.

[21] P.R. Gent. Balanced models in isentropic coordinates and the shallow water equations. *TellusA*, 36A:166–171, 1984.

[22] A. E. Gill. *Atmosphere-Ocean Press.* Academic Press, 1982.

[23] Haltiner and Williams. *Numerical Prediction and Dynamical Meteorology.* Wiley, 2 edition, 1980.

[24] Haurwitz. *Compendium of Meteorology*, chapter The Perturbation Equations of the Atmosphere, pages 416–417. AMS, 1951.

[25] R. Heikes and D.A.Randall. The shallow water equations on a spherical geodesic grid. Technical Report 524, Department of Atmospheric Science, 1993.

[26] J.R Holton. *An Introduction to Dynamic Meteorology.* Academic Press Inc., second edition, 1979.

[27] B.J. Hoskins. Stability of the Rossby-Haurwitz wave. *Quart. J. R. Met. Soc.*, 99:723–745, 1973.

[28] B.J. Hoskins, M.E. McIntyre, and A.W. Robertson. Use and significance of isentropic potential vorticity maps. *Quart. J. R. Met. Soc.*, 1992.

[29] N.B. Ingleby. The statistical structure of forecast errors and its representation in The Met. Office Global Variational Data Assimilation Scheme. *Quart. J. R. Met. Soc.*, 571A:209–222, 2001.

[30] R. Jacob, J. Hack, and D. Williamson. Solutions to the shallow water test set using the spectral transform method. Technical report, NCAR, 1993.

[31] A.C. Lorenc. Analysis methods for numerical weather predition. *Quart. J. R. Met. Soc.*, 111:1777–1194, 1986.

[32] A.C. Lorenc, S.B. Ballard, R.S. Bell, N.B. Ingleby, P.L. Andrews, D.M. Barker, J.R. Bray, A.M.Clayton, T. Dalby, D.Li, T.J. Payne, and F.W. Saunders. The met. office 3-dimensional variational data assimilation scheme. *Q.J.R. Meteorol. Soc.*, 126:2991–3012, 2000.

[33] E.N. Lorenz. On the existence of a slow manifold. *J. Atmos. Sci.*, 43:1547–1557, 1986.

[34] E.N. Lorenz. The slow manifold -what is it? *J. Atmos. Sci.*, 49:2449–2451, 1992.

[35] A. Malcolm. Evaluation of the proposed new unified model scheme vs the current unified model scheme on the shallow water equations. Numerical Analysis Report 1/96, University of Reading, 1992.

[36] M. McIntyre and W. Norton. Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, 57:1214–1235, 2000.

[37] P.E. Merilees. On the linear balance equation in terms of spherical harmonics. *Tellus*, 20:200–202, 1968.

[38] Moorthi, Shrinivas, and R. Higgins. Application of fast Fourier transforms to the direct soluiton of a class of two-dimensional separable elliptic equations. *Mon. Wea. Rev.*, 121(1):290–296, 1993.

[39] Moura. The eigensolutions of the linearised balance equations over the sphere. *J. Atmos. Sci.*, 33:887–907, 1976.

[40] A.V. Oppenheim and R.W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, 1989.

[41] D.F. Parrish and J.C. Derber. The national meteorological centre's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, 120:1747–1763, 1992.

[42] T. Payne. Advantages of pv as principal control variable. document under development.

[43] Joseph Pedlosky. *Geophysical Fluid Dynamics*. Springer, 2 edition, 1986.

[44] N Phillips. Numerical integration of the primitive equations on the hemisphere. *Mon. Wea. Rev.*, 87:333–345, 1953.

[45] Randall. The shallow water equations. Lecture Notes.

[46] S.R Rizvi. Documentation of the spectral statistical interpolation scheme. ECMWF Documentation, 1994.

[47] C G Rossby. Planetary flow patterns in the atmosphere. *Quart. J. R. Met. Soc*, 66:68–87, 1940.

[48] C G Rossby and collaborators. Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *Journal of Marine Research*, 2:38–55, 1939.

[49] I. Roulstone and M.J. Sewell. The mathematical structure of theories of semi-geostrophic type. *Phil. Trans. R. Soc. Lond.*, A355:2489–2517, 1997.

[50] A.L. Schoenstadt. The effect of spatial discretisation on the steady state and transient behaviour of a dispersive wave equation. *J. Comp. Physics*, pages 364–379, 1977.

[51] J.G. Smith and A.J.Duncan. *Elementary Statistics and Applications: Fundamentals of the theory of statistics*. McGraw-Hill, 1944.

[52] P.N. Swarztrauber. The direct solution of the discrete Poisson equation on the surface of a sphere. *J. Comp. Phys.*, 1974.

[53] C. Temperton. Implicit normal mode initialisation. *Mon. Wea. Rev.*, 116:1013–1031, 1988.

[54] C. Temperton. Implicit normal mode initialisation for spectral models. *Mon. Wea. Rev.*, 117:436–451, 1989.

[55] P.D. Thompson. A theory of large-scale disturbances in non-geostrophic flow. *J. Met*, 13:251–261, 1956.

[56] J. Thurburn and Yong Li. Numerical simulations of Rossby-Haurwitz waves. *Tellus*, 52A:181–189, 2000.

[57] Eric W. Weisstein. *CRC Concise Encyclopedia of Mathematics*. CRC Press, 1998.

[58] A. White. A view of the equations of meteorological dynamics and various approximations. Technical report, The Met. Office, 2000.

[59] D.L. Williamson and J.B. Drake. A standard test set for numerical approximations to the shallow water equations in spherical geometry. *J. Comp. Phys.*, 102:211–224, 1992.

[60] A. Winn-Nielsen. On geostrophic adjustment on the sphere. *Beit. Phys. Atmos.*, 49:254–271, 1976.