Inertia-gravity wave generation: a WKB approach

Jonathan Maclean Aspden

Doctor of Philosophy
University of Edinburgh
2010

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Jonathan Maclean Aspden*)

Abstract

The dynamics of the atmosphere and ocean are dominated by slowly evolving, large-scale motions. However, fast, small-scale motions in the form of inertia-gravity waves are ubiquitous. These waves are of great importance for the circulation of the atmosphere and oceans, mainly because of the momentum and energy they transport and because of the mixing they create upon breaking. So far the study of inertia-gravity waves has answered a number of questions about their propagation and dissipation, but many aspects of their generation remain poorly understood.

The interactions that take place between the slow motion, termed balanced or vortical motion, and the fast inertia-gravity wave modes provide mechanisms for inertia-gravity wave generation. One of these is the instability of balanced flows to gravity-wave-like perturbations; another is the so-called spontaneous generation in which a slowly evolving solution has a small gravity-wave component intrinsically coupled to it.

In this thesis, we derive and study a simple model of inertia-gravity wave generation which considers the evolution of a small-scale, small amplitude perturbation superimposed on a large-scale, possibly time-dependent flow. The assumed spatial-scale separation makes it possible to apply a WKB approach which models the perturbation to the flow as a wavepacket. The evolution of this wavepacket is governed by a set of ordinary differential equations for its position, wavevector and its three amplitudes. In the case of a uniform flow (and only in this case) the three amplitudes can be identified with the amplitudes of the vortical mode and the two inertia-gravity wave modes. The approach makes no assumption on the Rossby number, which measures the time-scale separation between the balanced motion and the inertia-gravity waves.

The model that we derive is first used to examine simple time-independent flows, then flows that are generated by point vortices, including a point-vortex dipole and more complicated flows generated by several point vortices. Particular attention is also paid to a flow with uniform vorticity and elliptical streamlines which is the standard model of elliptic instability. In this case, the amplitude of the perturbation obeys a Hill equation. We solve the corresponding Floquet problem asymptotically in the limit of small Rossby number and conclude that the inertia-gravity wave perturbation grows with a growth rate that is exponentially small in the Rossby number. Finally, we apply the WKB approach to a flow obtained in a baroclinic lifecycle simulation. The analysis highlights the importance of the Lagrangian time dependence for inertia-gravity wave generation: rapid changes in the strain field experienced along wavepacket trajectories (which coincide with fluid-particle trajectories in our model) are shown to lead to substantial wave generation.

Contents

Abstract **vi**

List of figures **xii**

1 Introduction **1**
1.1 Geophysical fluid dynamics 1
1.2 Outline of thesis 3

2 Geophysical fluid dynamics **7**
2.1 Introduction 7
2.2 Rotation and stratification 7
2.2.1 Coriolis effect 7
2.2.2 Rossby number 12
2.2.3 The Brunt-Väisälä frequency 12
2.3 Governing equations 14
2.3.1 Introduction 14
2.3.2 Boussinesq approximation 15
2.3.3 The Boussinesq equations 16
2.4 Balance relations 17
2.5 Potential vorticity 18
2.6 Conclusion 20

3 Inertia-gravity waves **21**
3.1 Introduction 21
3.2 Time-scale separation 22
3.3 Dispersion relation 23
3.4 Generation mechanisms 25
3.4.1 Spontaneous generation 25
3.4.2 Generation through instabilities 26
3.5 Conclusion 28

4 WKB approach **29**
4.1 Introduction 29
4.2 Derivation of Equations 30
4.2.1 Adding a perturbation 30
4.2.2 WKB Theory 31

4.2.3 Applying the WKB theory . . . 32
4.2.4 Vorticity and divergence . . . 35
4.2.5 Potential vorticity . . . 36
4.2.6 Eliminating $\hat{\rho}_0$ . . . 37
4.2.7 Final equations . . . 39
4.2.8 Recovering the intrinsic frequency . . . 40
4.2.9 Solving the system . . . 41
4.3 Non-dimensionalising . . . 42
4.4 Removing the singularity at $m = 0$ . . . 42
4.5 Energy . . . 44
4.6 Conclusion . . . 45

5 Simple flows **47**
5.1 Introduction . . . 47
5.2 No Flow . . . 47
5.3 Pure Strain Field . . . 48
5.4 Transverse shear . . . 52
5.5 Strain and Shear . . . 54
5.6 Frontogenesis flow . . . 58
5.7 Conclusion . . . 61

6 Point-vortex model **63**
6.1 Introduction . . . 63
6.2 Point vortices . . . 64
6.3 Dipole . . . 66
6.3.1 Wavenumber and amplitude equations . . . 67
6.3.2 Non-dimensionalising . . . 69
6.4 Polarisation . . . 69
6.4.1 Eigensolution . . . 69
6.4.2 Finding A_{v} and $A_{\mathrm{g}^{\pm}}$ . . . 71
6.5 Initialisation . . . 73
6.6 Results . . . 74
6.7 Elliptical trajectories within a dipole . . . 76
6.8 Complex time dependent flows . . . 79
6.8.1 Introduction . . . 79
6.8.2 Initialisation . . . 82
6.8.3 Results . . . 83
6.9 Conclusion . . . 88

7 Elliptical instability **89**
7.1 Introduction . . . 89
7.2 Formulation . . . 91
7.3 WKB analysis . . . 96
7.4 The Stokes phenomenon . . . 98
7.5 Calculating M . . . 99
7.6 Using exponential asymptotics to calculate S . . . 102
7.7 Analysis of the α and β integrals . . . 107
7.7.1 The asymptotics of α for small and large values of μ . . . 107
7.7.2 The asymptotics of α for small and large values of ψ . . . 109
7.7.3 The effect of β . . . 110
7.8 Position and thickness of the instability bands . . . 110

7.9 Comparison with numerical results . . . 111
7.10 Justifying the hydrostatic approximation . . . 114
7.11 Conclusion . . . 115

8 Baroclinic lifecycle 117
8.1 Introduction . . . 117
8.2 Baroclinic instability . . . 118
8.3 Model setup . . . 120
8.4 Modifying the data . . . 121
8.5 Interpolation . . . 122
8.5.1 Smoothing the data fields . . . 123
8.6 Initialisation . . . 123
8.7 Results . . . 132
8.8 Conclusion . . . 138

9 Conclusion 139

A Change of coordinates 143

Bibliography 145

List of Figures

2.1 The set up used in the derivation of the Coriolis force. 9

4.1 The form of a wavepacket. 31

5.1 Inertia-gravity-waves in the case of no flow. 49
5.2 The streamlines and velocity field of a pure strain field. 50
5.3 The energy of a wavepacket in a pure strain field. 53
5.4 The velocity field of a transverse shear flow. 54

6.1 The streamlines and velocity field of a point vortex induced dipole. . . . 68
6.2 Trajectories in a flow generated by a quasi-geostrophic dipole. 75
6.3 The evolution of the wavenumbers as a wavepacket sweeps past a dipole. 76
6.4 Inertia-gravity waves generated as a wavepackets sweeps past a dipole. . 77
6.5 The final amplitudes of the inertia-gravity wave mode. 78
6.6 The elliptical trajectory of a wavepacket in close proximity to a dipole. . 79
6.7 The wavenumbers evolution on an elliptical trajectory in a dipole. . . . 80
6.8 The amplitudes evolution on an elliptical trajectory in a dipole. 81
6.9 The trajectory, wavenumbers, amplitudes, and local Rossby number of a wavepacket in a random strain flow. 84
6.10 The positions of the wavepacket and point vortices when growth is observed. 85
6.11 The trajectory, wavenumbers, amplitudes, and local Rossby number of a wavepacket in a random strain flow. 87

7.1 The stream lines and velocity field of an anticyclonic elliptical flow. . . . 92
7.2 The paths of the integrals used in the calculation of S. 105
7.3 Contours of the parameters α and β governing the maximum growth rates. 109
7.4 Numerical estimates of the local maxima of the growth rates. 112
7.5 Growth rates in anticyclonic flows. 113
7.6 Growth rates in cyclonic flows. 114
7.7 Effect of the hydrostatic approximation. 115

8.1 The contours of constant density in the atmosphere. 119
8.2 A diagram showing how the data is interpolated. 124
8.3 Contours of dU/dx demonstrating the sensitivity to smoothing. 125
8.4 The vertical velocity profile of the flow. 126

8.5 The first half of a wavepackets trajectory. 127
8.6 The second half of a wavepackets trajectory. 128
8.7 The first half of a wavepackets trajectory. 129
8.8 The second half of a wavepackets trajectory. 130
8.9 The evolution of a wavepacket's amplitudes for different levels of smoothing. 133
8.10 The evolution of two wavepacket's wavenumbers, amplitudes and local Rossby number. 134
8.11 The evolution of two wavepacket's wavenumbers, amplitudes and local Rossby number. 135
8.12 The trajectory, wavenumbers, amplitudes and local Rossby number of a wavepacket. 137

Chapter 1

Introduction

1.1 Geophysical fluid dynamics

Have you ever wondered how and why weather systems evolve, how weather forecasts are made, why large scale weather patterns seem to roll up into cyclonic shapes, what the gulf stream is and how it affects us, how El Niño or La Niña form and evolve and cause the sometimes drastic effects that they seem to bring, how the oceans large scale currents work, and what is happening in Jupiter's famous red spot? The understanding of all these common and, on the surface, seemingly simple phenomena, comes under the umbrella of geophysical fluid dynamics. Put as simply as possible, geophysical fluid dynamics in general is the dynamics of large scale, rotating, stratified flows. Closest to home this includes the large scale motions that occur in the earth's atmosphere and oceans, but in a wider context it can be used to study flows found on other planets which often have a lot of similarities to the flows found on our planet. If we were to add another layer of complexity to our so far very simple definition of geophysical fluid dynamics, it would be that it provides the fundamental principles and language for understanding geophysical flows without being suffocated by the overwhelming complexity of the real world. That is, its main focus is on large scale motions that have a big effect on the flow. These motions can be realistically modeled by simplified models that ignore the small scale motions that are less important to the flows evolution and that are prohibitively complex to model and solve for. These simplified models take into account that the earth's atmosphere and oceans are, in comparison to their width, very shallow.

Geophysical flows are natural large scale flows that are characterised by the fact

that they are dominated by large-scale motions of the order of thousands of kilometres in the atmosphere and hundreds of kilometres in the ocean. Although it may be hard to imagine behaviour at these scales, these motions are the driving forces behind the everyday, small scale, behaviours that we notice such as surface ocean waves and the last rain shower. In the oceans, large scale currents sweep round the oceans' basins causing mixing of water of different temperatures and, in the atmosphere, phenomenon such as the trade winds are caused by such flows.

From the wide range of examples that have already been given it is quite clear that understanding these flows and answering some of the questions that we posed initially is very important to our understanding of the planet on which we live. Although these questions appear quite straightforward on the surface, we find that as soon as we start to look for the answers everything seems to get prohibitively complex. This is due in part to the vast numbers of processes happening constantly at a range of scales in the evolution of the planet's atmosphere and oceans. As a result of these, we pretty soon realise that it is, and will be for many years, going to be impossible to understand completely and resolve all the phenomena that are happening at once and create a full picture, or a completely deterministic computer model, of our planet.

The fact that the earth is spinning on an axis and that its gravity causes fluids to stratify are two major considerations of geophysical fluid dynamics. In geophysical fluid dynamics we try to separate processes, so that we can study them individually through the use of simplified models that are easier to understand. While the effect of the earth spinning is fundamental to the behaviour of geophysical flows and can not be ignored, less important effects such as topography, moisture and density fluctuations can, in different circumstances, be ignored to create models that we can attempt to understand.

We can study these geophysical flows evolutions by setting up systems of equations, that take into account the relevant approximations, to model them. A striking feature that is found when doing this is that two different time scales of motion exist; slow, large scale motions and fast, short scale motions. The slow motions are solutions of hydrostatic and geostrophic balance and the fast motions take the form of inertia-gravity waves. A measure of the separation between these two modes of motion is given by the Rossby number which is a ratio of the time scale of the slow motions to the fast motions. When this number is small there is a large gap between the time scales, and when it is large the gap is small. The flows in the atmosphere and oceans have a

small Rossby number and hence a large time scale separation gap. The consequence of this is that the activity of the fast inertia-gravity wave modes is low and that there is only a weak coupling between these fast modes and the slow modes. Despite the low activity of inertia-gravity waves, they play an important part in the evolution of the atmosphere and oceans trough transporting energy and momentum, causing mixing and generating turbulence. This makes understanding their lifecycle important. Along with the slow motions, the evolution and dissipation of inertia-gravity waves is well understood but there still remains a lot of questions about their generation. Some of the mechanisms of generation are quite well understood, for instance inertia-gravity waves generated by topography and convection, but some are far less understood. One possible source mechanism, that has been a long standing subject in geophysical fluid dynamics, [30, 32], is the interactions between the fast and slow modes of motion. These interactions may lead to spontaneous generation of inertia-gravity waves. This is where inertia-gravity waves spontaneously emerge from the slow modes of the flow, no matter how well the system is initialised so that only slow motions are excited.

The aim of this thesis therefore is to develop new tools that will enable the study of the spontaneous generation mechanisms that are involved in creating inertia-gravity waves. This new approach, that is valid for arbitrary values of the Rossby number, is based on the analysis of small-scale wavepackets. These tools can then be used in models of slowly evolving flows to study the processes involved in the inertia-gravity wave generation. In this thesis we will study two mechanisms, spontaneous generation, [15, 36, 46, 52, 58, 65, 67, 68] and generation caused by unbalanced instabilities, [40, 42, 50, 59, 69].

1.2 Outline of thesis

The plan of this thesis is as follows. In chapter 2 we introduce the main concepts and equations of geophysical fluid dynamics that will set the stage for the rest of the thesis. We consider the ways in which the earth's rotation affects large scale flows and discuss their quantification. We then move on to analysing the other important aspect of geophysical fluid dynamics, stratification. We finish this chapter by introducing an important quantity when studying geophysical fluid dynamics, the potential vorticity.

We then move on, in chapter 3, to discuss the two types of motion that are possible in geophysical flows. We establish that along with slow motion, fast motion, in the

form of inertia-gravity waves, is also permitted. The implications of these fast motions are set out before moving on to establish that there is a large time scale separation between the two forms of motion, that is, the slow, large scale motion and the fast, small scale motion. Since the atmosphere and oceans are forced at low frequencies this gap leads to the slow motions being extremely dominant. Although this is the case, we establish that it is impossible to totally remove the fast motions from the flow, which will motivate the discussion of the generation mechanisms that can create the fast motions, that is, the inertia-gravity waves. This issue of inertia gravity wave generation will become the motivation for the rest of the thesis.

In chapter 4 we develop a new approach for studying inertia-gravity wave generation, valid for arbitrary Rossby number, based on the analysis of small-scale wavepackets. We do this by deriving a set of equations that describes the evolution of a small-scale, small-amplitude wavepacket that is placed in a geophysical flow. The first step in this derivation is to add a small perturbation to the flow fields in the primitive equations and to then use WKB techniques to search for an approximate solution. These equations are then simplified through the introduction of three new variables, namely the vertical component of vorticity, the divergence of the horizontal velocity and the potential vorticity. As a result of this simplification we are left with a closed set of equations that completely describes the wavepacket's evolution. This set of equations consists of two equations for the wavepacket's position, three for its wavenumbers and three for its amplitudes which can, for a uniform flow, be solved to recover the two fast inertia-gravity wave modes and the slow vortical mode.

In chapter 5 we use these equations to study a few simple time-independent flows. We begin with the case where there is no flow and steadily increase the complexity by working through flows that contain a horizontal strain, then a vertical shear flow before arriving at a case which is a combination of the two. In each case we use the equations we have derived to determine the evolution of the wavepacket's position, its wavenumbers and its amplitudes which will be useful in two ways. The first being to test the equations we have derived to see if they behave as expected by producing the same results as those that are already known for these flows and secondly to gain some intuition into how the wavepacket behaves in simple systems before tackling harder, more complex, systems.

In chapter 6 we build on the previous chapter by increasing the complexity of the flows to be studied. This is achieved by considering flows that are generated

by quasi-geostrophic point vortices and have a richer time dependence than those previously examined. The first of these flows is generated by a point vortex dipole. In this case the fact that the flow is uniform at large distances from the dipole is used to decompose the solutions to the wavepacket's evolution into three modes. By then examining the energy transfer between the vortical, balanced, mode of the wavepacket and its gravity wave modes, we can determine if there is any gravity wave generation. We then briefly consider the elliptical trajectories that lie close to the point vortices within the dipole, which will be used as a motivation for the next chapter. The second example that we consider is a complex time-dependent flow that is created by multiple quasi-geostrophic point vortices. By placing multiple point vortices of various strengths around the wave packet the effect will be like that of a random strain flow. Since this is a flow that has received some attention in the past, [24, 29], although in a different way without studying the amplitudes of a wavepacket, we can try to verify these results with our set up. We can also take these results further by using the decomposition that was done previously in this chapter to infer some results about the wavepacket's evolution.

In chapter 7 we study a flow, motivated by work in the previous chapter, with elliptical streamlines. We first formulate the problem and then derive the necessary equations in the small Rossby number regime. Following this, the instability problem can be posed as a Floquet problem. In the small Rossby number regime in which we are working the wave generation, and so the growth rate, through the elliptical instability is exponentially small. This growth rate can then be calculated by linking the growth of the solutions to the existence of a Stokes phenomenon, which we capture using a combination of matched asymptotics and WKB expansion. This growth rate is then compared to results obtained by numerically solving the Floquet multipliers problem. We also compare the results gained when the hydrostatic approximation is taken versus the results when it is not.

The final scenario we study, in chapter 8, is a simulation of an idealised baroclinic lifecycle taking place in the earth's atmosphere. We first adapt the data from the simulation to fit with our equations; the data then needs to be interpolated to the position of the wavepacket in the flow. We then study several trajectories that pass through areas where there are rapid changes in the velocity and divergences fields of the baroclinic lifecycle with different initial values of the wavevector. By doing this, a link can be established between the growth of the wavepacket to areas of the flow where there

is a Lagrangian transience. We will show that this transience causes sudden variations in the wavevector which has a knock on effect of causing the wavepacket's amplitudes to grow. A link will also be established between the growth of the wavepacket and a local value of the Rossby number.

To round off the thesis, chapter 9 contains a summary of the results and a discussion of the implications that they might have on our understanding of certain aspects of geophysical fluid dynamics. It also contains a brief discussion on the ways in which these results could be used in further research.

Chapter 2

Geophysical fluid dynamics

2.1 Introduction

The aim of this thesis is to study the generation of inertia-gravity waves in the context of geophysical fluid dynamics. Although this context can extend to other planets, we will restrict our study through the choice of parameters to flows that occur on the earth. It is worth noting that this restriction can easily be removed, and flows on other planets can be studied, by a simple change of parameters. To study the generation of inertia-gravity waves we first need to set the scene by laying down the fundamentals of geophysical fluid dynamics, the most important aspects of which are rotation and stratification. After we have these basics laid down we will be in a position to form a set of equations that will govern the dynamics taking place.

2.2 Rotation and stratification

The effect that plays a large part in determining the evolution of geophysical flows is the earth's rotation. The effect of the earth's rotation on large scale motion is to deflect the flow's direction. This is the Coriolis effect.

2.2.1 Coriolis effect

Since the earth is rotating it is easier and more convenient to study geophysical fluid dynamics in the rotating reference frame of an observer on the earth's surface. To do this we will consider a change of coordinates from an inertial system to one that is

rotating.

We begin by defining $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ to be the Cartesian basis vectors for the inertial reference frame given by (x, y, z), and $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$ to be the basis vectors for a rotating reference frame given by (x', y', z'). The rotating reference frame is rotating at a constant angular velocity of $\mathbf{\Omega}$ with respect to the inertial frame. If we let the two frames of reference share the same origin and orientation then $\mathbf{z}$, $\mathbf{z}'$ and $\mathbf{\Omega}$ will all share the same direction as shown in figure 2.1, [56].

If we place a particle in the domain, then its position vector $\mathbf{r}$ can be given in either set of coordinates as

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}'. \tag{2.1}$$

If we consider d/dt to be the rate of change of the particle measured in the inertial frame and d/dt' to be the rate of change in the rotating frame, then the particles velocity in the inertial frame is given by

$$\frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}. \tag{2.2}$$

It can also be expressed in terms of the rotating reference frame as

$$\frac{d\mathbf{r}}{dt} = \frac{dx'}{dt}\mathbf{i}' + \frac{dy'}{dt}\mathbf{j}' + \frac{dz'}{dt}\mathbf{k}' + x'\frac{d\mathbf{i}'}{dt} + y'\frac{d\mathbf{j}'}{dt} + z'\frac{d\mathbf{k}'}{dt}. \tag{2.3}$$

By considering figure 2.1 it is clear that

$$\frac{d\mathbf{i}'}{dt} = \mathbf{\Omega} \times \mathbf{i}', \tag{2.4}$$

$$\frac{d\mathbf{j}'}{dt} = \mathbf{\Omega} \times \mathbf{j}', \tag{2.5}$$

$$\frac{d\mathbf{k}'}{dt} = \mathbf{\Omega} \times \mathbf{k}' = 0. \tag{2.6}$$

Introducing these into (2.3) along with the fact that $dx'/dt = dx'/dt'$, with its counterparts, gives that

$$\frac{d\mathbf{r}}{dt} = \frac{dx'}{dt'}\mathbf{i}' + \frac{dy'}{dt'}\mathbf{j}' + \frac{dz'}{dt'}\mathbf{k}' + \mathbf{\Omega} \times (x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{j}'). \tag{2.7}$$

Now using the fact that

$$\frac{d\mathbf{i}'}{dt'} = \frac{d\mathbf{j}'}{dt'} = \frac{d\mathbf{k}'}{dt'} = 0, \tag{2.8}$$

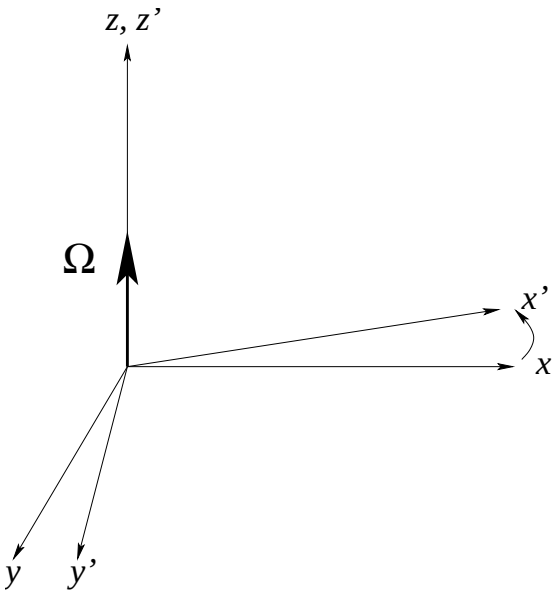

Figure 2.1: The set up used in the derivation of the Coriolis force which involves an inertial reference frame, given by (x, y, z), and a rotating reference frame, given by (x', y', z'), that is rotating with a constant angular velocity of $\mathbf{\Omega}$.

which stems from the definition of d/dt', we can rewrite (2.7) as

$$\frac{d\mathbf{r}}{dt} = \frac{d}{dt'}(x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{j}') + \mathbf{\Omega} \times (x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{j}') \tag{2.9}$$

$$= \frac{d\mathbf{r}}{dt'} + \mathbf{\Omega} \times r. \tag{2.10}$$

Differentiating with respect to t again to get the particles acceleration gives that the relation between the acceleration in the two reference frames is

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{d^2\mathbf{r}}{dt'^2} + 2\mathbf{\Omega} \times \frac{d\mathbf{r}}{dt'} + \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{r}). \tag{2.11}$$

The term on the left of this equation is the acceleration as seen in the inertial reference frame and the first term on the right is the acceleration of the particle as seen in the rotating frame. The second and third terms on the right are then the Coriolis force and the centrifugal force, [49]. It is worth noting that neither of these are actual forces, but may be thought of as quasi-forces that can be seen to act on a body and affect its motion when it is observed from a rotating frame of reference.

A key point to note here is that the centrifugal force can be combined with the gravitational force to create a single force. This force does not just act on geophysical fluids to affect their evolution but it also acts on the earth's surface. Since the surface

of the earth is elastic, its natural state of equilibrium is normal to this combined force. In contrast to this, if the earth's surface was not elastic, objects would not stay at rest but would drift towards the equator. Hence, as a result of the elasticity of the earth's surface, the centrifugal force does not appear explicitly in the equations governing a geophysical flow. The earth's rotation does not then just introduce apparent forces that are associated with the rotation of the reference frame, but it is also the cause of a genuine physical effect, central to much of the dynamics of the atmosphere and ocean.

The Coriolis force is named after Gaspard Gustave Coriolis, $(1792 - 1843)$, who discovered it during his study of rotating mechanical systems. This force has many far reaching consequences for geophysical fluid dynamics but its basic properties can be summarised as, [56]

1. there is no Coriolis force acting on bodies that are stationary in a rotating frame,
2. it acts to deflect moving bodies at right angles to their direction of travel,
3. it does no work on a body as it acts perpendicular to the velocity of the body and so $\mathbf{v} \cdot (\boldsymbol{\Omega} \times \mathbf{v}) = 0$.

We can work out the value of the Coriolis force on earth by first considering the expression for its force per unit mass which is given by

$$\mathbf{F} = -2\boldsymbol{\Omega} \times \mathbf{v}. \tag{2.12}$$

If we consider this equation from the position of an observer standing on the earth's surface, rotating at speed $\Omega = 2\pi/\text{day}$ and a latitude of ϕ, and set up a local coordinate system around them so that the x-axis is due east, the y-axis due north and the z-axis straight up then we can write $\boldsymbol{\Omega}$ and $\mathbf{v}$ as

$$\boldsymbol{\Omega} = \omega \begin{pmatrix} 0 \\ \cos\phi \\ \sin\phi \end{pmatrix}, \tag{2.13}$$

and

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \tag{2.14}$$

Introducing these into (2.12) gives

$$\mathbf{F} = -2\Omega \begin{pmatrix} v_z \cos\phi - v_y \sin\phi \\ v_x \sin\phi \\ -v_x \cos\phi \end{pmatrix}. \tag{2.15}$$

When considering the earth's atmosphere or oceans, the vertical component of the velocity is very small compared to the horizontal components and the vertical component of the Coriolis force is small compared to gravity. This means that we can use the traditional approximation and restrict this expression to the horizontal plane giving that

$$F = \begin{pmatrix} v_y \\ -v_x \end{pmatrix} f, \tag{2.16}$$

where

$$f = 2\Omega \sin\phi \tag{2.17}$$

is known as the Coriolis parameter. This enables us to quantify the effect that the Coriolis force has on moving bodies on the earth's surface at different latitudes. It is important to note that the value of this parameter increases as the latitude increases, i.e. closer to the poles, and decreases and actually vanishes at the equator. Although this might at first seem quite surprising, it is in fact quite intuitive. At the equator the earth's rotation vector is parallel to the earth's surface and is therefore applying a force which is perpendicular to the earth's surface. This force is directly opposed by the gravitational force and so there is no net motion. In contrast, when not at the equator, the gravitational force and the force perpendicular to the earth's rotation axis are no longer in opposing directions, but an angle has formed between them. This change in angle between the forces gives rise to the variation, across different latitudes, of the Coriolis force.

The complications that this variation in the Coriolis force can create can be approximated out under the right conditions. Although the earth's rotation is central to many geophysical fluid dynamics problems, the fact that the earth is near spherical is often not. This is particularly the case when studying flows that have a scale which is smaller than global. In these situations it becomes really awkward to use spherical coordinates and so finding a way to use a local Cartesian system becomes important. This is done by defining a tangent plane to the earth's surface at the latitude that

we are interested in and then taking the value of the Coriolis parameter, (2.17), as a constant over the whole tangent plane. This approximation is called the f-plane and it works well for any flows that are limited in their latitudinal extent so that the effects of the spherical nature of the earth are unimportant. This approximation can greatly simplify the study of these flows, as the value of the Coriolis parameter, f, is now a constant and the work can be undertaken in Cartesian coordinates rather than spherical coordinates.

2.2.2 Rossby number

Now that we have quantified the effect that the earth's rotation has on moving bodies, we need a way to determine whether that rotation has any effect on the phenomenon we are studying. To do this we define a dimensionless number by

$$\epsilon = \frac{U}{fL}, \tag{2.18}$$

where U and L are the characteristic velocity and length scales of the phenomenon, respectively, and f is the Coriolis parameter. This number is called the Rossby number, after Carl-Gustav Arvid Rossby, and is essentially a ratio of magnitude of the relative acceleration to the Coriolis acceleration, [37]. When the Rossby number is small the effects of rotation are important and when it is large they are not. For example two people throwing a ball in a park may have $U = 30$ ms^{-1} with $L = 40$ m and so have a Rossby number of $\epsilon = 7500$, where we have taken $f = 10^{-4}$ which is a reasonable value on earth. By contrast, an intercontinental missile with $U = 100$ ms^{-1} and $L = 10000$ km which has $\epsilon = 0.1$. This shows that while playing catch in a park you do not need to worry about the ball deflecting to the right, in the northern hemisphere, or left in the southern hemisphere, whereas the effect of the earth's rotation will affect the path of the missile causing it to miss its target quite considerably.

2.2.3 The Brunt-Väisälä frequency

Since we now have a handle on how the earth's rotation affects geophysical flows we need to move our attention to how the other defining feature of these flows, the stratification, affects them. Fluids on earth naturally settle under gravity so that the denser particles are at the bottom and the lighter ones at the top. When this occurs the fluid is said to be stratified. This occurs in the atmosphere and the oceans and so plays a large

part in geophysical fluid dynamics by acting on a particle that gets perturbed from its natural position of equilibrium.

We can derive the effect that the stratification has on a perturbed particle by considering a fluid parcel of density ρ that is placed in an incompressible fluid, that is, a fluid that conserves the density of a particle. If we now perturb this particle adiabatically then it is going to be surrounded by fluid that is of a different density than itself. If the vertical density profile of the fluid is given by $\bar{\rho}(z)$ and the particle in question has been moved from its initial height of z, where it had density $\bar{\rho}(z)$, up to a new height of $z+\delta z$, where it still has the same density, then its density will differ from its surroundings by

$$\begin{aligned} \delta\rho &= \rho(z+\delta z) - \bar{\rho}(z+\delta z), && (2.19)\\ &= \bar{\rho}(z) - \bar{\rho}(z+\delta z), && (2.20)\\ &= -\frac{\partial\bar{\rho}}{\partial z}\delta z. && (2.21) \end{aligned}$$

If $\partial\bar{\rho}/\partial z < 0$, then at this new height the particle will be heavier than its surroundings and so there will be a restoring force acting on it to bring it back to its original height. On the other hand if $\partial\bar{\rho}/\partial z > 0$ then the particle will be lighter than its surroundings and so the displacement will be unstable and the particle will continue to rise.

In the first case we expect the restoring force to cause the particle to move back down again towards its position of equilibrium. As it reaches this position it will not just stop there but its momentum will cause it to continue past this point. There will now be an upwards restoring force that will force the particle back up and so on causing the particle to oscillate around its position of equilibrium. We can calculate the frequency of this oscillation by first writing the force per unit volume on the displaced particle as

$$F = -g\delta\rho = g\frac{\partial\bar{\rho}}{\partial z}\delta z. \tag{2.22}$$

We can now use Newton's second law of motion to derive the equation of motion of the particle which gives

$$\frac{\partial^2\delta z}{\partial t^2} = \frac{g}{\rho_b}\frac{\partial\bar{\rho}}{\partial z}\delta z, \tag{2.23}$$

where we have approximated $\bar{\rho}$ by ρ_b, a reference density, in the denominator. Solving

this equation for δz gives that

$$\delta z = A\cos(Nt) + B\sin(Nt) \tag{2.24}$$

where A and B are constants and N, defined by

$$N^2 = -\frac{g}{\rho_b}\frac{\partial \bar{\rho}}{\partial z}, \tag{2.25}$$

is the buoyancy frequency. This buoyancy frequency is called the Brunt-Väisälä frequency after David Brunt and Vilho Väisälä and it gives the frequency at which a vertically displaced particle oscillates in a stably stratified fluid. From the expression for N^2, we can see that if $N^2 > 0$ then the upwardly displaced particle will be heavier than its surroundings and so will experience a restoring force. This force causes the particle to oscillate around its starting position with frequency N. Conversely if $N^2 < 0$, then the density profile of the fluid is unstable, as heavier fluid particles are resting on top of lighter ones. In this case the particle will be surrounded by heavier particles and so it will continue to rise in a process called convection.

2.3 Governing equations

2.3.1 Introduction

Now that we have an understanding of the basic principles that underpin geophysical fluid dynamics the next step is to derive a set of equations so that we can start to study the phenomena that take place in this setting. It is worth noting here that we will assume, in all the derivations and equations that follow, that the flows we deal with are inviscid. The validity of this assumption is guaranteed because we are dealing with large scale flows and so boundary effects can be ignored. The full derivation of the primitive equations, that is, the Eulerian equations of motion of a fluid in terms of the fluid's velocity field, can be found in all textbooks on geophysical fluid dynamics, [18, 37, 49, 56], so here we will just give an outline.

There are five equations that are needed to describe the evolution of a stratified fluid on an f-plane in a rotating environment. They are a momentum equation in each of the three Cartesian directions, a density equation and finally a continuity equation. These five equations form a closed system that enables us to study these flows.

In their full form these equations are very complicated and hard to handle, therefore justified approximations have been devised that simplify the equations, without losing a significant amount of the detail. In the case of the atmosphere and the oceans, a very useful approximation is the Boussinesq approximation.

2.3.2 Boussinesq approximation

A simplification of the primitive equations for a geophysical flow can be achieved by using the Boussinesq approximation. In geophysical systems the density of the fluid varies very slightly around a mean value, depending on position and temperature. As an example, the mean temperature of the ocean is 4℃ and the mean salinity is 3.47% which combines to give a mean density of 1028 kgm^{-3}. Within one ocean basin these numbers are so stable that the variations in density rarely exceed ± 3 kgm^{-3} from the mean value, which is a very small percentage change, [11]. Intuitively, we may think that this is not the case with the atmosphere, since the air gets more rarefied with altitude. However, the altitude range that we are interested in is the range where all the weather patterns are confined to. This region is known as the troposphere and it is the first atmospheric region above the earth's surface. It contains approximately 75% of the atmosphere's mass and 99% of its water vapour. The depth of the troposphere changes with latitude with its depth being greater in the tropical regions, up to 18 km, and shallower near the poles, about 7 km in summer and nearly indistinct in winter, [3].

We have pointed out that it is justifiable to expect that the fluid density, ρ, will not vary very much from its mean value, which we call the reference density, ρ_b, and so we can write the buoyancy as

$$\rho_{\text{total}} = \rho_b \left(1 + \frac{1}{g}\rho(x, y, z, t)\right). \tag{2.26}$$

Here $\rho(x, y, z, t)$, which has been scaled by the reference density over the gravitational constant g, is the variation in density that is induced by a change in position or time. It is very important to note that in this formulation the perturbation term is a lot smaller than the mean term, that is $\rho(x, y, z, t)/g \ll 1$. Neglecting $\rho(x, y, z, t)$ and so fixing ρ to be ρ_b in all density terms, except those multiplied by the gravitational acceleration g, greatly simplifies the governing equations. This is the Boussinesq approximation. In essence this approximation says that the difference in inertia is negligible but gravity

is sufficiently strong to make the specific weight appreciably different.

2.3.3 The Boussinesq equations

After applying the Boussinesq approximation to the governing equations, the five equations that govern a stratified, rotating, inviscid fluid, on an f-plane, are given by

$$\frac{Du}{Dt} - fv = -\frac{\partial p}{\partial x}, \tag{2.27}$$

$$\frac{Dv}{Dt} + fu = -\frac{\partial p}{\partial y}, \tag{2.28}$$

$$\frac{Dw}{Dt} + \rho = -\frac{\partial p}{\partial z}, \tag{2.29}$$

$$\frac{D\rho}{Dt} = 0, \tag{2.30}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{2.31}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla, \tag{2.32}$$

is the material derivative, that is, the rate of change of a property of a particular infinitesimal particle of the fluid. This operator is derived by considering the rate of change of a property, ϕ say, of a fluid that has velocity field $\mathbf{u}$. Since the value of this property is changing with time and space, the chain rule is used to write

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial t}\frac{\partial t}{\partial t} + \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial \phi}{\partial z}\frac{\partial z}{\partial t}, \tag{2.33}$$

$$= \frac{\partial \phi}{\partial t} + \frac{\partial \mathbf{x}}{\partial t} \cdot \nabla \phi, \tag{2.34}$$

$$= \frac{\partial \phi}{\partial t} + \mathbf{u} \cdot \nabla \phi, \tag{2.35}$$

which is the same as (2.32). For a full derivation of this set of governing equations see [11]. In this system, (2.27)-(2.31), u, v and w are the components of the velocity in the x, y and z Cartesian directions respectively and the pressure, p, has been scaled so that $p = \tilde{p}/\rho_b$, where $\tilde{p}$ is the actual pressure. We will take this set of equations as the basis for all the analysis and discussion that follows.

These equations are almost the same, notation aside, as the widely used, primitive equations that are defined using a different vertical coordinate by McWilliams and Gent, [38]. The main difference between these sets of equations is that McWilliams and Gent

have taken the hydrostatic approximation, detailed in section 2.4. This is not a problem because the large-scale motions considered and the inertia-gravity-waves excited are hydrostatic and so the difference will be negligible. To go from the set of primitive equations that we defined above to the set defined by McWilliams and Gent, we need to replace p by ϕ, the geopotential, and $-\rho$ by θ, the potential temperature. Here, the geopotential is the potential energy that a particle has due to the earth's gravitational field and the potential temperature is the temperature that a particle would gain if moved adiabatically to a reference pressure, which is usually 1000 millibars. The vertical coordinate is now the pressure-like coordinate that was defined by Hoskins and Bretherton, [26], which may be thought of as the geometric height in shallow layers. This discussion is expanded further in appendix A.

2.4 Balance relations

There are two fundamental balances in geophysical fluid dynamics: the hydrostatic balance and the geostrophic balance. The corresponding states of hydrostasy and geostrophy are very rarely exactly realised but their approximate satisfaction has profound consequences on the behaviour of the atmosphere and oceans.

To find the hydrostatic balance, we consider the relative sizes of the terms in the vertical momentum equation, (2.29). This gives,

$$\begin{array}{cccc} \dfrac{\partial w}{\partial t} + \mathbf{u}\cdot\nabla w & = & -\dfrac{\partial p}{\partial z} - \rho, \\ W/T \quad UW/L & & \end{array} \tag{2.36}$$

with the term's scales given underneath in terms of U, W, L and T which are the characteristic horizontal velocity, vertical velocity, length scale and time scale respectively. For most large-scale motion in the atmosphere and oceans, the terms on the right-hand side of this equation are orders of magnitude larger than the terms on the left-hand side, that is, the vertical accelerations are small compared to the gravitational acceleration and therefore they must be approximately equal to each other, i.e.

$$\frac{\partial p}{\partial z} = -\rho, \tag{2.37}$$

where ρ is still a buoyancy rather that the density. This equation is known as the

hydrostatic balance relation and, when it holds, it implies that the pressure at any point in the fluid is only due to the weight of the fluid above it.

The other balance relation is geostrophic balance. As stated above, the Rossby number is the ratio of the magnitude of the relative acceleration to the Coriolis acceleration. This can be seen by examining the terms involved in the horizontal momentum equations, (2.27) and (2.28). After expanding the material derivative, the horizontal momentum equation in the x direction, (2.27), becomes

$$\begin{array}{ccccc} \dfrac{\partial u}{\partial t} & + \mathbf{u}\cdot\nabla u & - fv & = & -\dfrac{\partial p}{\partial x}, \\ & U^2/L & fU & & \end{array} \tag{2.38}$$

where the scales of the relative and Coriolis acceleration terms have been placed below their terms, confirming our expression for the Rossby number, (2.18). If the Rossby number is sufficiently small, then it is clear that the rotation term will dominate the nonlinear advection term. The rotation term also dominates the local time derivative term if the time can be scaled as L/U, [62]. When this is the case, the only term that can balance the rotation term is the pressure term. This means that

$$fu \approx -\frac{\partial p}{\partial y}, \tag{2.39}$$

$$fv \approx \frac{\partial p}{\partial x}. \tag{2.40}$$

This balance of terms is called the geostrophic balance and when it occurs, the fluid flows parallel to the lines of constant pressure. Although in practice geostrophic balance is rarely achieved in the atmosphere and oceans, outside of the tropics they are close to being in geostrophic balance and so it is a very valuable first approximation.

2.5 Potential vorticity

No introduction to geophysical fluid dynamics would be complete without introducing a quantity of great importance to the study of this area, that was introduced by Carl-Gustaf Rossby in the 1930's. This quantity is the potential vorticity and is defined by

$$q = \frac{1}{\rho}\zeta_a \cdot \nabla\theta, \tag{2.41}$$

where ρ is the full density, θ is the potential temperature and ζ_a is the absolute vorticity vector, that is, the curl of the three-dimensional velocity field viewed in an inertial frame. The defining feature of the potential vorticity that makes it so useful is that it is materially conserved in an unforced dissipationless flow, that is, if we denote the potential vorticity by q, then

$$\frac{Dq}{Dt} = 0. \tag{2.42}$$

For such a flow, we also have that

$$\frac{D\theta}{Dt} = 0. \tag{2.43}$$

This material conservation of the potential temperature gave Rossby the idea of creating a new quantity from the vorticity by using the same process that creates the potential temperature from the temperature. This idea led to the creation of the potential vorticity.

In more rigorous terms, the potential vorticity is a conservation law that builds on Kelvin's circulation theorem, which states that the circulation around a material fluid parcel is conserved, or in another way, the circulation is conserved following the flow, [35]. As it is, we will not be able to use Kelvin's circulation theorem in this study as it only applies when certain conditions are met. A condition of this theorem that we fail to meet is that the flow must be barotropic. This restriction means that the pressure depends only on the density and vice versa, that is $p = p(\rho)$. This is problematic for geophysical fluid dynamics, since the flows that are dealt with in geophysical fluid dynamics are rarely barotropic.

This problem with Kelvin's circulation theorem is what motivated Rossby, and then in a more general way Ertel, to search for a new quantity which obeys a conservation law, which led to the idea of potential vorticity. The underlying principle of this is to use a scalar field that is being advected by the flow to encode all the information about the fluids' evolution. Using the equation for this scalars evolution along with the vorticity equation then gives a scalar conservation equation which is the potential vorticity. For a rigorous derivation of the potential vorticity equation in a variety of different circumstances see [62].

When Rossby first defined potential vorticity he used a few multiplicative constants in his definition so that it would have the same units as the vorticity, in the same way as the potential temperature has the same units as the temperature. This convention has since been replaced with the convention of 'PV units' where one potential vorticity unit

is defined as 10^{-6} K m^2 kg^{-1} s^{-1}, as implied by (2.41). As an interesting aside, potential vorticity can be used to determine where the tropopause, the boundary between the troposphere that we have already defined and the next layer of the earth's atmosphere called the stratosphere, lies in non tropical areas. The tropopause is also the point at which air ceases to cool with height and becomes completely dry. It turns out that cross sections of the earth's atmosphere at the tropopause have a potential vorticity value of close to 2 PV units.

The potential vorticity is also an extremely important quantity because it has an inversion principle. When a flow is balanced, it satisfies the potential vorticity invertibility principle. By balanced we mean that the inertia-gravity waves are eliminated and the flow satisfies the balance relations as described above in section 2.4, or more formally, a flow is balanced when there exists a function that relates the three-dimensional velocity field to the spatial distribution of mass throughout the fluid, and the mass under each isentropic surface, that is, a surface with constant potential temperature, is known. For a fuller definition and a complete discussion of balanced flows see [34]. This inversion principle states that if the potential vorticity distribution is known on all isentropic surfaces, then all the remaining dynamical information about the flow is implicitly contained within the data. To retrieve the rest of the flow's data, that is, the pressure, density, potential temperature and velocity fields, the potential vorticity distribution is put into the inversion operator. Hence the potential vorticity is a very powerful tool that can also encode a lot of information about the flow. Again, for a more in-depth discussion of this see [35].

2.6 Conclusion

In this chapter we have set out the main features of geophysical fluid dynamics that we will need in our study of inertia-gravity wave generation. We have also set out the governing equations of a geophysical flow that we will use as a basis for our study. The next step is to consider the different types of motions the can occur in geophysical fluid dynamics.

Chapter 3

Inertia-gravity waves

3.1 Introduction

All the flows and phenomena that we have mentioned so far have been large scale motions. In fact, the atmosphere and the oceans are dominated by such large scale motions. These motions are typically low frequency and balanced, that is, nearly hydrostatically and geostrophically balanced. Although the majority of the flow's energy is taken up in these slow, large scale motions, fast, short scale motions also exist. These motions take the form of inertia-gravity waves that can be generated through interactions between the fast and slow modes.

Inertia-gravity waves play a substantial part in determining the circulation of the atmosphere and oceans and so understanding how they generate, evolve and then dissipate is very important. These waves are important for a number of reasons. They transport energy and momentum as they propagate either vertically or horizontally. This causes a transfer of energy and momentum from the source of the waves to their sink, where they dissipate. This transfer is of great importance to the momentum balance of the upper atmosphere. The dissipation of energy and momentum when the waves break is also important because it can cause mixing. Inertia-gravity waves are known to contribute to the generation of turbulence and to influence the thermal structure of the middle atmosphere, [17], and in the oceans they contribute to mixing and are a possible energy source for the thermohaline circulation. Clearly then, understanding the behaviour of inertia-gravity waves is essential if we are to try and understand the dynamics of the earth's middle atmosphere and its oceans.

Out of the three phases of the life of an inertia-gravity wave, generation, propagation and dissipation, the latter two are quite well understood but there are still a lot of questions that need answering regarding their generation. In this thesis we will focus our attention on this aspect of their lifecycle. Specifically, we examine mechanisms of generation that are associated with the unavoidable coupling between balanced motion and inertia-gravity waves. At a linear level, there is a complete separation between the two modes of motion but when nonlinear effects are taken into account the two modes become inextricably coupled. This implies that no matter how close the motion is to being perfectly balanced, inertia-gravity waves will always be present. A very simple example that can illustrate this is a flow passing over topography such as a mountain. When a flow passes over a mountain, the flow is disrupted, causing fast waves, called mountain waves, to form as the flow passes. The issue of separating balanced motion and inertia-gravity waves has been of long-standing interest, particularly in the area of weather forecasting.

While there are many mechanisms of inertia-gravity wave generation, such as convection and topography, the two mechanisms that we will concentrate on in this thesis are spontaneous generation and generation through unbalanced instabilities. The first one occurs in balanced flows where a slowly evolving solution has a small gravity-wave component intrinsically coupled to it. In this case, energy is transferred from the balanced, slow mode of the system to the unbalanced, fast modes. In the other case, inertia-gravity waves are generated through the instability of balanced flows to gravity-wave-like perturbations.

3.2 Time-scale separation

There is a large time-scale separation in the atmosphere and the oceans and so we can split the motions that occur there into two categories. First, there is the slow balanced part of the motion, the vortical mode, that evolves on a time scale that is roughly L/U, where U and L are the characteristic velocity and length scales of the flow respectively. Secondly, there is the fast part of the motion, the inertia-gravity wave modes, that evolve on a time scale that is equal to or less than f, the Coriolis parameter. Since this is the case, the Rossby number, (2.18), can be used as a gauge of the scale separation. The larger the Rossby number, the smaller the scale separation is and the smaller the Rossby number, the larger the scale separation.

The atmosphere and the oceans primarily have a small Rossby number, $\epsilon \approx 0.1$ in the atmosphere and $\epsilon \approx 0.01$ in the oceans. Although most of the atmosphere and oceans subscribe to this regime, the Rossby number can also be $O(1)$ locally. This forms part of the motivation for the approach we take to studying inertia-gravity wave generation, which we will set out in the next chapter. The approach, while being able to study the $\epsilon \leq 1$ regime that is most relevant to the atmosphere and oceans, will also be able handle $O(1)$ Rossby numbers.

When the two modes of motion in a system have a large time scale separation they are only weakly coupled. The geophysical flows that we consider in the atmosphere and the oceans are examples of this. These flows are dominated by large scale, slow motions and so the activity of small scale, fast motions is very low. This has led to balanced models, that is, models that do not include the fast parts of the motions, being used to represent the dynamics to a high level of accuracy. These models filter out inertia-gravity waves entirely and base all their evolution on what has been called a slow manifold. These are sub-manifolds of the state space in which the fast motion is very weak, [31, 70]. To use this slow-manifold, a balance relation, such as geostrophic balance, see section 2.4, is used to couple the fast dynamical variables to the slow ones. This is taken as the definition of the slow-manifold while the dynamics that take place in the manifold are given by a balance model, which is often taken as the quasi-geostrophic model, [64]. This model approximates horizontal flows with what they would be under geostrophic balance and ignores all vertical momentum fluxes. As a balanced model it is ideal for describing motions on a large scale when the Rossby number is less than unity, but cannot describe small fast motions.

3.3 Dispersion relation

Inertia-gravity waves are found by examining a linearised version of the governing equations, (2.27)-(2.31), in the absence of a background flow;

$$\frac{\partial u}{\partial t} - fv = -\frac{\partial p}{\partial x}, \tag{3.1}$$

$$\frac{\partial v}{\partial t} + fu = -\frac{\partial p}{\partial y}, \tag{3.2}$$

$$\frac{\partial w}{\partial t} + \rho = -\frac{\partial p}{\partial z}, \tag{3.3}$$

$$\frac{\partial \rho}{dt} - N^2 w = 0, \tag{3.4}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \tag{3.5}$$

If we now look for small scale perturbation solutions to this system that take the form of a wave that can be represented by

$$\begin{pmatrix} u \\ v \\ w \\ \rho \\ p \end{pmatrix} = \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\rho} \\ \hat{p} \end{pmatrix} e^{i(kx+ly+mz-\omega t)}, \tag{3.6}$$

where the hatted quantities are constants, (k, l, m) are the wavenumbers in the (x, y, z) direction and ω is the frequency of the wave, [18]. Substituting this expression for the perturbation into the governing equations gives that

$$-i\omega\hat{u} - f\hat{v} + ik\hat{p} = 0, \tag{3.7}$$

$$-i\omega\hat{v} + f\hat{u} + il\hat{p} = 0, \tag{3.8}$$

$$i\omega\hat{w} + \hat{\rho} + im\hat{p} = 0, \tag{3.9}$$

$$i\omega\hat{\rho} + N^2\hat{w} = 0, \tag{3.10}$$

$$ik\hat{u} + il\hat{v} + im\hat{w} = 0. \tag{3.11}$$

Using the last equation to get an expression for $\hat{w}$, eliminating $\hat{p}$ using the third equation and writing the system in matrix form gives that

$$\begin{pmatrix} -i\omega(1+\frac{k^2}{m^2}) & -f - \frac{ikl\omega}{m^2} & -\frac{k}{m} \\ f - \frac{ikl\omega}{m^2} & -i\omega(1+\frac{l^2}{m^2}) & -\frac{l}{m} \\ \frac{N^2k}{m} & \frac{N^2l}{m} & -i\omega \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{\rho} \end{pmatrix} = 0. \tag{3.12}$$

This system admits the trivial solution $\hat{u} = \hat{v} = \hat{w} = 0$, unless its determinant vanishes. Hence for a non-trivial solution, that is, for waves to occur, we need that either

$$\omega = 0 \quad \text{or} \quad \omega = \pm\sqrt{\frac{m^2 f^2 + N^2(k^2 + l^2)}{k^2 + l^2 + m^2}}. \tag{3.13}$$

The first case corresponds to the large scale, slowly evolving, steady vortical mode of the flow. Balanced motion can then be thought of as the nonlinear evolution of this mode. The other two roots correspond to the fast motions of the flow that take the form of travelling waves called inertia-gravity waves. This expression for the frequency of the fast modes is the inertia-gravity waves dispersion relation. It relates the waves' frequency to the spatial scales of the waves, that is, their wavenumbers, and to the flow's background properties, the Coriolis parameter, f, and the Brunt-Väisälä frequency, N^2. It is important to note that inertia-gravity waves have no potential vorticity, that is, $q = 0$. Also of interest is that for vertically propagating waves, that is, when k, l and m are all real, the frequency of the inertia-gravity waves is confined to the range $N^2 > \omega^2 > f^2$, [17].

3.4 Generation mechanisms

3.4.1 Spontaneous generation

It is now of interest to ask whether these slow-manifolds can be completely without fast motions, that is, inertia-gravity waves. If this is the case then we would not expect to see any spontaneous generation of these waves and so we would have to look for a different generation mechanism. It actually turns out that this is not the case. In the case of large scale atmospheric or oceanic flows, there are not any exactly invariant slow manifolds, [64]. Therefore, no matter how well the system is initialised, that is, the initial conditions are picked so that only slow motions are excited, fast motions in the form of inertia-gravity waves will always emerge. This is called spontaneous generation, where the word spontaneous is used since all inertia-gravity wave activity that can be eliminated by suitable initialisation has been eliminated. Along with showing that balanced models have limitations, in the sense that they can never be truly balanced, this also shows that there is a mechanism for generating inertia-gravity waves from balanced flows.

Although we have just stated that these slow manifolds can never be truly balanced,

it has been shown in some models that they can be balanced up to an error that is exponentially small in the Rossby number, [33]. This error, which can be thought of as the inertia-gravity wave amplitude, has the form $\exp(-\beta/\epsilon)$ for some $\beta > 0$. This result depends on the time scale separation of the flow's two modes remaining large for all spatial scales, that is, the Rossby number remaining small. It is worth noting here that it is not enough to only have a strong stratification, as in that regime the time scale separation does not hold for all spatial scales. The fact that the balanced modes of the system can only be balanced to an exponentially small accuracy has an implication for what we can expect in terms of inertia-gravity waves that are spontaneously generated. Since the unbalanced parts of the flow are the error terms in the initial balancing procedure, which we have already shown to be exponentially small in the Rossby number, we can deduce that the inertia-gravity waves that are spontaneously generated must also be exponentially small in the Rossby number.

To try and quantify how important spontaneous generation is as a source of inertia-gravity waves in the small Rossby number regime and to try and understand the mechanisms involved more fully, two lines of study have been carried out. While one of these lines of study has focused on running high-resolution models of three-dimensional flows, the other line has focused on theoretical analysis of simplified models. In the first case, the work has been carried out by studying models such as a baroclinic lifecycle, [50, 52, 58]. Although progress has been made on the generation processes involved in these cases, no link has yet been made between the Rossby number and the amplitude of the waves generated. However, in the theoretical case the exponential smallness of the generated inertia-gravity waves has been established by using exponential asymptotics to gain an asymptotic estimate for the inertia-gravity wave amplitude, [63, 65].

In this thesis we will visit both these lines of study, that is, using the data from high resolution models and the theoretical analysis of highly simplified models, in a context that has not been examined before, with the aim of developing simple models that can diagnose inertia-gravity wave generation.

3.4.2 Generation through instabilities

In the last section, we discussed the generation of inertia-gravity waves in flows that are initially balanced, that is, there are no inertia-gravity waves present at $t = 0$. In this section, we will introduce the concept of a generation mechanism that occurs

through the instability of balanced flows to small unbalanced perturbations. In this situation, significant levels of inertia-gravity wave activity are generated by instabilities in the balanced flow. These instabilities amplify the inertia-gravity wave-like, or more generally unbalanced, perturbations that are initially present in the flow. Hence, in contrast to spontaneous generation the flow must initially be very slightly unbalanced. This compromising act of having a balanced flow that has very small unbalanced perturbations is achieved by taking the initial potential vorticity of the perturbations to be zero since this is characteristic of inertia-gravity waves, [64]. This ensures that all but a very small amount of the flow's energy is initially in the balanced modes. As we discussed above, this is always the case down to an exponentially small scale for balanced flows. This implies that if the instability mechanism was truly unbalanced it would be eliminated from the balanced flow and so any instability that does exist must result in a growth rate that is also exponentially small in the Rossby number. The instabilities that have been studied so far have been in time-independent flows. In that case, the aim is to look for exponentially growing perturbations.

In previous studies of unbalanced instabilities, several mechanisms have been identified. Most instabilities involve the interaction between two wave modes. The resonance between an edge wave and an inertia-gravity wave was exploited in [42, 50]. In contrast, the instability in a horizontally sheared flow was created by the resonance between either two Kelvin waves, two inertia-gravity waves or a combination of a Kelvin wave with an inertia-gravity wave in [39, 66, 72].

Since this generation mechanism and the mechanism of spontaneous generation have a similar effect, that is, they both produce inertia-gravity waves that are exponentially small in the Rossby number, it may often be difficult to distinguish between each mechanism, particularly in the case of time-dependent complex flows.

It is also worth noting here that another type of inertia-gravity wave generation mechanism has been studied, [40]. This study involved taking the standard model of elliptical instability, [27], and creating a rotating stratified version of it. Here, resonances between the rotation of the system and the rotation of the wavepacket around the elliptical streamlines create an instability that amplifies the wavepacket.

3.5 Conclusion

In this chapter we have discussed how that even though the atmosphere and the oceans are dominated by slow, large scale motions there still exists fast, small scale motions, in the form of inertia-gravity waves, that have a large influence on the evolution of their surroundings. After introducing these fast motions, we have established that although their propagation and dissipation is generally well understood, there are lots of questions surrounding their generation. We have also derived an expression for their dispersion relation. This led to a discussion of the mechanisms in which these waves can be generated and which has laid down some of the motivations for the work we carry out in this thesis.

Previous studies of spontaneous generation and unbalanced instabilities have involved studying very specific flows. For spontaneous generation, they have been flows for which the fluid equations can be solved explicitly; for unbalanced instabilities, they have been steady flows for which the standard normal-mode method can be applied. In the next chapter, we develop a new method to examine inertia-gravity wave generation in essentially arbitrary balanced flows.

Chapter 4

WKB approach

4.1 Introduction

Now that we have set out the basics of geophysical fluid dynamics and have introduced the set of primitive equations that we can use to study certain phenomena in this field, we can start to build a model that we can use to look for inertia-gravity wave generation. The aim of this chapter is to derive a set of equations that model the evolution of a small-scale, small-amplitude wavepacket that is placed in a geophysical flow. This model is formulated with the goal of learning more about the instabilities that can arise in these types of flow.

The geophysical flows model that we will use here to try to achieve these aims, consists of five partial differential equations to which we will apply the WKB (Wentzel-Kramers-Brillouin) approximation. This involves using the exponential nature of the equations' behaviour to find a global approximation to the equations' solution. This is done by seeking an exponential solution to the equations, where the exponential exponent and the exponential multiplier of this solution can be expanded in a power series. This makes it possible to describe the evolution of the small-scale, small-amplitude wavepacket in terms of ordinary differential equations, thus making the system a lot simpler to solve and hence understand. The vertical component of the vorticity and the divergence of the horizontal velocity, along with the potential vorticity, can then be introduced into the system. This simplifies the system further by reducing it to a set of three ray equations. This approach is closely related to the WKB approach to stability, reviewed in [16], which has recently been applied to

rotating, stratified flows in [22], where equations equivalent to the ones we now derive have been obtained.

4.2 Derivation of Equations

4.2.1 Adding a perturbation

Our aim is to study what happens to small perturbations that are placed in a time-dependent flow, with an emphasis on looking for growth. To do this we take the small perturbation in the form of a small amplitude wavepacket, whose form can be seen in figure 4.1. The scale of this wavepacket is much smaller than that of the surrounding flow.

To study this we can add the wavepacket to the Boussinesq equations that were stated and discussed in section 2.3.3, by introducing

$$u = U + u', \tag{4.1}$$

$$v = V + v', \tag{4.2}$$

$$w = W + w', \tag{4.3}$$

$$p = P + p', \tag{4.4}$$

$$\rho = R + \rho', \tag{4.5}$$

where the capital letters denote the background flow of the fluid and the dashed quantities denote small perturbations that are caused by the wavepacket. In this study, we have assumed that the background flow is purely horizontal and therefore from here on we will assume that $W = 0$ and so $w = w'$.

Using the above identities, the x-momentum equation, (2.27), becomes

$$\partial_t U + \partial_t u' + \mathbf{u}' \cdot \nabla U + \mathbf{U} \cdot \nabla u' + \mathbf{u}' \cdot \nabla u' + \mathbf{U} \cdot \nabla U - fV - fv' = -\partial_x P - \partial_x p', \tag{4.6}$$

were ∂_x stands for $\partial/\partial x$. The terms with just capital letters can be removed from this equation as they satisfy the original governing equation, (2.27), for the background flow. Since we are dealing with a small-scale, small amplitude wavepacket, we can linearise these equations, so after neglecting the $\mathbf{u}' \cdot \nabla u'$ term, we are left with

$$\partial_t u' + u'\partial_x U + v'\partial_y U + w'\partial_z U + U\partial_x u' + V\partial_y u' - fv' = -\partial_x p'. \tag{4.7}$$

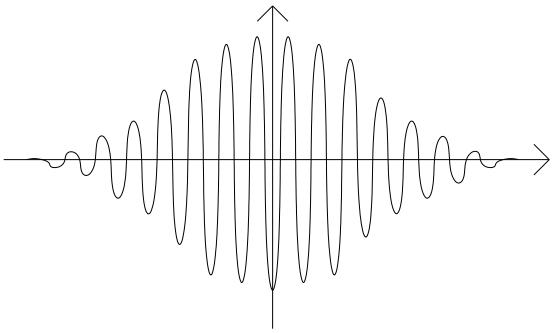

Figure 4.1: The form of the wavepackets that will be used to add a perturbation to the geophysical flow.

After applying the same process to the rest of the governing equations, (2.28)-(2.31), the system becomes

$$\partial_t u' + u'\partial_x U + v'\partial_y U + w'\partial_z U + U\partial_x u' + V\partial_y u' - fv' = -\partial_x p', \tag{4.8}$$

$$\partial_t v' + u'\partial_x V + v'\partial_y V + w'\partial_z V + U\partial_x v' + V\partial_y v' + fu' = -\partial_y p', \tag{4.9}$$

$$\partial_t w' + U\partial_x w' + V\partial_y w' + \rho' = -\partial_z p', \tag{4.10}$$

$$\partial_t \rho' + U\partial_x \rho' + V\partial_y \rho' + u'\partial_x R + v'\partial_y R + w'\partial_z R = 0, \tag{4.11}$$

$$\partial_x u' + \partial_y v' + \partial_z w' = 0. \tag{4.12}$$

These equations look very similar to the original set of equations, just recast in terms of the perturbation instead of the background flow.

4.2.2 WKB Theory

The exact solution to this set of differential equations is prohibitively complicated and so WKB theory is used to find an approximate solution. The theory is named after Wentzel, Kramers and Brillouin who popularised, it although in truth credit should also be given to many others including Jefferys and Rayleigh who developed a lot of the early theory. WKB theory is a very powerful tool that is used to find a global approximation to the solution of a differential equation, whose highest derivative is multiplied by a small parameter. For a detailed motivation and a rigorous derivation of this technique see [8].

Briefly, the aim of the WKB approximation is to seek a solution to a set of

differential equations that is of the form

$$u = \hat{u}(x, y, z, t)e^{i\theta(x,y,z,t)/\mu}. \tag{4.13}$$

Here the functions $\hat{u}$ and θ can be expanded in powers of μ, where $0 < \mu \ll 1$, and the phase, θ, is defined by

$$\nabla\theta = \mathbf{k} \quad \text{and} \quad \frac{\partial\theta}{\partial t} = -\omega, \tag{4.14}$$

where $\mathbf{k}$ is the wavevector and ω is the frequency. These two quantities are connected through a dispersion relation which has the form

$$\omega = \omega(k, l, m, x, y, z). \tag{4.15}$$

We derived the explicit expression of the dispersion relation for inertia-gravity waves in section 3.3, although it is worth noting that this expression did not include the Doppler shift that is induced by the fluid's flow. This adds a term of the form $\mathbf{U} \cdot \mathbf{k}$ to the expression for the dispersion relation.

4.2.3 Applying the WKB theory

The WKB approximation described above is implemented by setting

$$\begin{pmatrix} u' \\ v' \\ w' \\ \rho' \\ p' \end{pmatrix} = \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \\ \hat{\rho} \\ \hat{p} \end{pmatrix} e^{\frac{i\theta(\mathbf{x},t)}{\mu}}, \tag{4.16}$$

where the real part is always taken. Here the hatted quantities can all be expanded in powers of μ, the scale separation parameter, ie $\hat{u} = \hat{u}_0 + \mu\hat{u}_1 + \ldots$ and similarly with $\hat{v}$, $\hat{w}$, etc. As above, θ is defined so that $\nabla\theta = \mathbf{k}$ and $\partial_t\theta = -\omega$, where $\mathbf{k}$ is the wavevector and ω is the frequency. Substituting (4.16) into (4.8)-(4.12) gives that, at $O(1/\mu)$,

$$-i\omega\hat{u}_0 + iUk\hat{u}_0 + iVl\hat{u}_0 = -ik\hat{p}_0, \tag{4.17}$$

$$-i\omega\hat{v}_0 + iUk\hat{v}_0 + iVl\hat{v}_0 = -il\hat{p}_0, \tag{4.18}$$

$$-i\omega\hat{w}_0 + iUk\hat{w}_0 + iVl\hat{w}_0 = -im\hat{p}_0, \tag{4.19}$$

$$-i\omega\hat{\rho}_0 + iUk\hat{\rho}_0 + iVl\hat{\rho}_0 = 0, \tag{4.20}$$

$$ik\hat{u}_0 + il\hat{v}_0 + im\hat{w}_0 = 0. \tag{4.21}$$

By simplifying (4.20) we obtain that the frequency is given by

$$\omega = \mathbf{U} \cdot \mathbf{k}. \tag{4.22}$$

It is interesting to note that this expression only includes the Doppler shift part of the frequency and that the intrinsic frequency, as defined in section 3.3, is not present. This is because it does not feature at this order in this formulation and so we will have to look at the next order for the full expression for the frequency.

Substituting this expression for ω into the three momentum equations, (4.17)-(4.19), in turn, gives that

$$\hat{p}_0 = 0, \tag{4.23}$$

since we know that $k \neq 0$ for all time.

An expression for the wavenumbers can be derived from (4.22). This is done by first using the definitions of $\nabla\theta$ and $\partial_t\theta$, that were given above, in $D_t\theta$ to acquire that

$$\frac{D\theta}{Dt} = 0. \tag{4.24}$$

If we then take the gradient of this expression, we find that

$$\frac{\partial\nabla\theta}{\partial t} + (\nabla\mathbf{U})^T \cdot \nabla\theta + \mathbf{U} \cdot \nabla^2\theta = 0, \tag{4.25}$$

which simplifies to

$$\frac{D\mathbf{k}}{Dt} = -(\nabla\mathbf{U})^T \cdot \mathbf{k}. \tag{4.26}$$

This is a differential equation for the evolution of the wavenumbers, which can be solved to give each wavenumber's evolution in terms of gradients of the flow's velocity field. This expression can also be written as

$$\frac{Dk_i}{Dt} = -\frac{\partial U_j}{\partial x_i}k_j, \tag{4.27}$$

where summation is understood. A useful analogy to this equation that can offer an

insight into its meaning, is the equivalent equation from the field of passive scalars which governs a scalar's concentration. The two equations are identified with each other by letting θ be the concentration of the passive scalar, which gives that $\mathbf{k}$ is analogous to the gradients of the concentration. It is then possible to think of the concentration of the wavecrests, which is described by the wavenumber, as the same as the concentration of the passive scalar. This helps to add some intuition to the evolution of the wavevector, particularly in the case of a simple contraction or expansion field.

A useful result concerning the evolution of the wavenumbers in a time-independent velocity field can be derived by considering the material derivative of $\mathbf{k} \cdot \mathbf{U}$ which gives

$$\frac{D}{Dt}(\mathbf{k} \cdot \mathbf{U}) = \frac{D\mathbf{k}}{Dt} \cdot \mathbf{U} + \mathbf{k} \cdot \frac{D\mathbf{U}}{Dt}. \tag{4.28}$$

By using (4.26) and the fact that since we have taken $\mathbf{U}$ to be time independent, we have that

$$\frac{D\mathbf{U}}{Dt} = \nabla\mathbf{U} \cdot \mathbf{U}, \tag{4.29}$$

which reduces the original equation to

$$\frac{D}{Dt}(\mathbf{k} \cdot \mathbf{U}) = 0. \tag{4.30}$$

This shows that $\mathbf{k} \cdot \mathbf{U}$ is materially conserved, i.e. it is conserved along a flow's streamlines. This can be used to either give an intuitive feel as to how the wavenumbers will evolve or as a check that the evolution that a calculation predicts is reasonable.

Now that we have established the first order behaviour of the wavepacket we need to move up an order and look at (4.10) at $O(1)$. This gives that

$$\frac{D\hat{w}_0}{Dt} + \hat{\rho}_0 = -im\hat{p}_1. \tag{4.31}$$

This expression can be rearranged in terms of $\hat{p_1}$ and then used to eliminate it from the other equations in the set at this order. After doing this and substituting in (4.22), the governing equations at $O(1)$ become

$$D_t\hat{u}_0 + \hat{u}_0\partial_x U + \hat{v}_0\partial_y U + \hat{w}_0\partial_z U - f\hat{v}_0 = \frac{k}{m}\left(\frac{D\hat{w}_0}{Dt} + \hat{\rho}_0\right), \tag{4.32}$$

$$D_t\hat{v}_0 + \hat{u}_0\partial_x V + \hat{v}_0\partial_y V - \hat{w}_0\partial_z V + f\hat{u}_0 = \frac{l}{m}\left(\frac{D\hat{w}_0}{Dt} + \hat{\rho}_0\right), \tag{4.33}$$

$$D_t\hat{\rho}_0 + \hat{u}_0\partial_x R + \hat{v}_0\partial_y R + \hat{w}_0\partial_z R = 0, \tag{4.34}$$

where $\hat{w}_0 = -(l\hat{v}_0 + k\hat{u}_0)/m$, which was obtained by rearranging (4.21). At this order we now have a system of three equations, instead of the original five, but there are still a lot of variables in these equations and so it would be advantageous if we could eliminate some of them.

4.2.4 Vorticity and divergence

The system can now be made simpler through the introduction of ζ, the vertical component of the vorticity, and δ, the divergence of the horizontal velocity. These quantities are defined as

$$\hat{\zeta}_0 = ik\hat{v}_0 - il\hat{u}_0 \quad \text{and} \quad \hat{\delta}_0 = ik\hat{u}_0 + il\hat{v}_0, \tag{4.35}$$

with dimensions that are inverse times. Taking the material derivative of the expression for $\hat{\zeta}_0$ gives that

$$D_t\hat{\zeta}_0 = ikD_t\hat{v}_0 - ilD_t\hat{u}_0 + i\hat{v}_0D_tk - i\hat{u}_0D_tl. \tag{4.36}$$

Equations (4.26), (4.32) and (4.33) can now be used to eliminate the material derivatives on the right hand side of this expression to give, after simplification, that

$$D_t\hat{\zeta}_0 = (\partial_y U - \partial_x V - \frac{l}{m}\partial_z U + \frac{k}{m}\partial_z V)\hat{\delta}_0 - (\partial_y V + \partial_x U)\hat{\zeta}_0 - f\hat{\delta}_0. \tag{4.37}$$

Since we are dealing with an incompressible fluid, which means that $\partial_x U + \partial_y V = 0$, the previous expression can be simplified to

$$D_t\hat{\zeta}_0 = -\left(f + \Omega + \frac{l}{m}\partial_z U - \frac{k}{m}\partial_z V\right)\hat{\delta}_0, \tag{4.38}$$

where $\Omega = \partial_x V - \partial_y U$.

In the same manner the expression for $\hat{\delta}_0$ can be manipulated to produce

$$\begin{aligned} D_t\hat{\delta}_0 &= i\left(\frac{D\hat{w}_0}{Dt} + \hat{\rho}_0\right)\frac{k^2+l^2}{m} + f\hat{\zeta}_0 - 2ik\hat{u}_0\partial_x U - 2ik\hat{v}_0\partial_y U \\ &- 2il\hat{u}_0\partial_x V - 2il\hat{v}_0\partial_y V + \left(\frac{k}{m}\partial_z U + \frac{l}{m}\partial_z V\right)\hat{\delta}_0. \end{aligned} \tag{4.39}$$

This expression still has four hatted terms that we would like to eliminate. To start with, $\hat{u}_0$ and $\hat{v}_0$ can be eliminated by combining and rearranging the expressions for $\hat{\delta}_0$ and $\hat{\zeta}_0$ given in (4.35) to obtain

$$\hat{u}_0 = \frac{i(l\hat{\zeta}_0 - k\hat{\delta}_0)}{k^2+l^2} \quad \text{and} \quad \hat{v}_0 = -\frac{i(k\hat{\zeta}_0 + l\hat{\delta}_0)}{k^2+l^2}. \tag{4.40}$$

Substituting in these identities along with the fact that $\hat{w}_0 = i\hat{\delta}_0/m$ reduces (4.39) to

$$\begin{aligned} D_t\hat{\delta}_0 &= i\hat{\rho}_0\frac{m(k^2+l^2)}{\kappa^2} + \frac{fm^2}{\kappa^2}\hat{\zeta}_0 + \left(\frac{m^2-k^2-l^2}{m\kappa^2}\right)(k\partial_z U + l\partial_z V)\hat{\delta}_0 \\ &+ \frac{2m^2\hat{\zeta}_0}{\kappa^2(k^2+l^2)}\big(kl(\partial_x U - \partial_y V) + l^2\partial_x V - k^2\partial_y U\big) \\ &- \frac{2m^2\hat{\delta}_0}{\kappa^2(k^2+l^2)}\big(kl(\partial_x V + \partial_y U) + l^2\partial_y V + k^2\partial_x U\big), \end{aligned} \tag{4.41}$$

where $\kappa^2 = k^2+l^2+m^2$. The governing equations have now been reduced to a system of two equations which are given in terms of the wavepacket's wavenumbers, the gradients of the flow's velocity fields and $\hat{\rho}_0$. We would now like to find a way of eliminating $\hat{\rho}_0$ from these equations, since we do not have a way of calculating it, to give a complete closed set of equations that we can solve.

4.2.5 Potential vorticity

The conservation law for the potential vorticity, as discussed in section 2.5, can now be used to take us forward. In a Boussinesq fluid the scaled potential vorticity, q, is given by

$$q = -\frac{g}{\rho_b}(f\hat{\mathbf{z}} + \text{curl } \mathbf{u}) \cdot \nabla\rho_{\text{total}}, \tag{4.42}$$

whose dimensions are an inverse time cubed. The potential vorticity is a really useful quantity to introduce to the system because it obeys a simple conservation law that can be written as

$$\frac{Dq}{Dt} = 0. \tag{4.43}$$

In the same manner as the other flow fields were defined earlier, the potential vorticity can be written as $q = Q + q'$. Using this and $\mathbf{u} = \mathbf{U} + \mathbf{u}'$ means that, once we have

added our perturbation into it, the conservation law becomes

$$\left(\frac{\partial}{\partial t} + \mathbf{U}\cdot\nabla + \mathbf{u}'\cdot\nabla\right)(Q+q') = 0. \tag{4.44}$$

The background potential vorticity, Q, is materially conserved, that is, $D_tQ = 0$, and so after using this fact and linearising we are left with

$$\frac{Dq'}{Dt} = -\mathbf{u}'\cdot\nabla Q. \tag{4.45}$$

4.2.6 Eliminating $\hat{\rho}_0$

The potential vorticity can now be used to eliminate $\hat{\rho}_0$ from (4.41). To do this we must first calculate all the terms in the definition of the potential vorticity, (4.42). Since $\mathbf{u} = \mathbf{U} + \mathbf{u}'$,

$$\text{curl}\ \mathbf{u} = -\partial_z V\hat{\mathbf{x}} + \partial_z U\hat{\mathbf{y}} + (\partial_x V - \partial_y U)\hat{\mathbf{z}} + \text{curl}\ \mathbf{u}', \tag{4.46}$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are the standard Cartesian unit vectors. Using (4.5) to modify (2.26) gives

$$\rho_{\text{total}} = \rho_b\left(1 + \frac{1}{g}(R + \rho'(x,y,z,t))\right), \tag{4.47}$$

where $1 \gg R/g, \rho'/g$ and so by taking the gradient of this expression we determine that

$$\nabla\rho_{\text{total}} = \left(\nabla R + \nabla\rho'\right)\frac{\rho_b}{g}. \tag{4.48}$$

Using these two derived identities in the definition of q gives that

$$Q = \partial_z V\frac{\partial R}{\partial x} - \partial_z U\frac{\partial R}{\partial y} - (f+\Omega)\frac{\partial R}{\partial z}, \tag{4.49}$$

and

$$q' = \partial_z V\frac{\partial \rho'}{\partial x} - \partial_z U\frac{\partial \rho'}{\partial y} - (f+\Omega)\frac{\partial \rho'}{\partial z} - \text{curl}\ \mathbf{u}'\cdot\nabla R. \tag{4.50}$$

As expected from the definition of the potential vorticity, these definitions involve the density of the flow and the density of the perturbation respectively. To simplify the expression for the perturbation's potential vorticity, we need to try to remove the $\mathbf{u}'$

term. By using (4.16) again, with q' added, we find that

$$\text{curl } \mathbf{u}' \cdot \nabla R = \begin{pmatrix} \partial_y \hat{w} + il\hat{w}/\mu - \partial_z \hat{v} - im\hat{v}/\mu \\ \partial_z \hat{u} + im\hat{u}/\mu - \partial_x \hat{w} - ik\hat{w}/\mu \\ \partial_x \hat{v} + ik\hat{v}/\mu - \partial_y \hat{u} - il\hat{u}/\mu \end{pmatrix} \cdot \begin{pmatrix} \partial_x R \\ \partial_y R \\ \partial_z R \end{pmatrix}, \tag{4.51}$$

which, if we expand each variable in powers of μ as done previously, can be used to find that at $O(\mu^{-1})$,

$$\begin{aligned} \hat{q}_0 = & - (f+\Omega) im\hat{\rho}_0 + ik\hat{\rho}_0 \partial_z V - il\hat{\rho}_0 \partial_z U - \partial_z R\hat{\zeta}_0 \\ & - \partial_x R(il\hat{w}_0 - im\hat{v}_0) - \partial_y R(im\hat{u}_0 - ik\hat{w}_0). \end{aligned} \tag{4.52}$$

We can now remove all the hatted terms from this expression, except $\hat{\rho}_0$ and $\hat{q}_0$, by using $\hat{w}_0 = i\delta/m$ and the identities derived in (4.40) to obtain that, after simplification,

$$\begin{aligned} \hat{q}_0 = & - (f+\Omega) im\hat{\rho}_0 + ik\hat{\rho}_0 \partial_z V - il\hat{\rho}_0 \partial_z U - \partial_z R\hat{\zeta}_0 \\ & + \partial_x R\left(\frac{l\hat{\delta}_0}{m} + \frac{m(k\hat{\zeta}_0 + l\hat{\delta}_0)}{k^2+l^2}\right) + \partial_y R\left(\frac{m(l\hat{\zeta}_0 - k\hat{\delta}_0)}{k^2+l^2} - \frac{k\hat{\delta}_0}{m}\right). \end{aligned} \tag{4.53}$$

This expression can now be rearranged in terms of $\hat{\rho}_0$ and used to eliminate it from (4.41). This is of great use to us since this expression now only contains terms that involve gradients of the background flow, the wavepacket's wavenumbers and the wavepacket's amplitudes, all of which we can calculate. Since this is the case, all that remains to do now to have a complete closed set of equations for the wavepacket's amplitudes, is to modify (4.45) so that it is in terms of $\hat{q}_0$ instead of q'. This is done by using (4.16), with q added. This gives that

$$\frac{D\hat{q}}{Dt} + \hat{\mathbf{u}} \cdot \nabla Q + \frac{i}{\mu}(Uk + Vl - \omega) = 0, \tag{4.54}$$

where we can use the expressions for $\hat{\mathbf{u}}$ that were found in (4.40). This equation can be greatly simplified expanding $\hat{u}$ and $\hat{q}$ in powers of μ as done previously. From earlier we know that the leading order behaviour of $\hat{q}$ is at $O(1/\mu)$ and the leading order behaviour of $\hat{u}$ is at 0(1). This means that at leading order the term containing $\hat{u}$ drops out of the above equation. The final term of the above expression also disappears when we

substitute in $\omega = \mathbf{U} \cdot \mathbf{k}$, which we derived earlier. This leaves us with

$$\frac{D\hat{q}_0}{Dt} = 0. \tag{4.55}$$

4.2.7 Final equations

The system can now be written in closed form in terms of $\hat{\delta}_0$, $\hat{\zeta}_0$, and $\hat{q}_0$, which from now on will be denoted by δ, ζ and q, as

$$\begin{aligned}
\frac{D\delta}{Dt} &= \left(\frac{m^2 f}{\kappa^2} + \frac{2m^2}{\kappa^2(k^2+l^2)}(kl(\partial_x U - \partial_y V) + l^2\partial_x V - k^2\partial_y U)\right)\zeta \\
&+ \left(\frac{m}{\alpha\kappa^2}(k\partial_x R + l\partial_y R) - \frac{k^2+l^2}{\alpha\kappa^2}\partial_z R\right)\zeta \\
&+ \left(\frac{m^2-k^2-l^2}{m\kappa^2}(k\partial_z U + l\partial_z V) + \frac{l\partial_x R - k\partial_y R}{\alpha m}\right)\delta \\
&- \left(\frac{2m^2}{\kappa^2(k^2+l^2)}(kl(\partial_x V + \partial_y U) + l^2\partial_y V + k^2\partial_x U)\right)\delta \\
&- \frac{k^2+l^2}{\alpha\kappa^2}q, \qquad (4.56) \\
\frac{D\zeta}{Dt} &= -\alpha\delta, \qquad (4.57) \\
\frac{Dq}{Dt} &= 0, \qquad (4.58)
\end{aligned}$$

where

$$\alpha = f + \Omega + \frac{l}{m}\partial_z U - \frac{k}{m}\partial_z V. \tag{4.59}$$

The equations governing the system have now been greatly reduced in number and complexity.

An interesting feature of these equations is that they depend on the angle of the wavevector, $\mathbf{k}/|\mathbf{k}|$, rather than on the magnitude of its components. This means that the scale of $\mathbf{k}$ becomes irrelevant and so when deciding on the initial conditions when running simulations, we can ignore the magnitude of the wavenumber and just decide on the angle of the wavevector. This observation can also be used to change the equation for the evolution of the wavenumbers, (4.26), to one that only tracks the evolution of the angle of the wavevector. This is done by defining $n = \mathbf{k}/|\mathbf{k}|$ and then introducing it into (4.26). This gives an equation for $D_t n$ which can then be transformed by writing n in terms of spherical polar coordinates. This means that instead of tracking k, l and m, this equation will now track the two angles involved in the spherical polar coordinates. Although this sounds like it might simplify the solving of this system of equations, this

route has not been taken here since it turns out that it is not any more convenient than what we already have.

With these equations we can look at two different modes of inertia-gravity wave generation. These two different modes are accessed by either setting $q \neq 0$ or $q = 0$. The first mode corresponds to spontaneous generation of waves that occur when energy is transferred from the vortical modes of the system to the inertia-gravity wave modes. The second mode of generation corresponds to waves being created through unbalanced instabilities of the flow. Here the waves are generated through properties of the flow that are unbalanced and lead to the growth of the wavepacket.

We can contrast these equations, (4.56)-(4.58), to the equations that are gained through standard WKB procedures that have been carried out before on the same initial set of equations, i.e. in the case of shallow water. The first difference that appears between the two is that the intrinsic frequency is found as a second order behaviour here rather than as a leading order behaviour as is found in more standard models. This is understandable since in this derivation we have assumed that the horizontal wavevectors are large and so in the limit of $k \to \infty$ the full frequency reduces to $\omega = \mathbf{U} \cdot \mathbf{k}$, as was found at the first order. This is different to the shallow water case where in the limit of large $\mathbf{k}$ the intrinsic frequency term remains in the expression for the full frequency. The other main difference is that at leading order, standard WKB techniques yield three decoupled modes: a vortical mode and two inertia-gravity wave modes which can be used to solve for one amplitude. This amplitude satisfies wave action conservation, which is a conservation law for the action of a wave that is given by the wave's energy over its intrinsic frequency. This is clearly different to the above method which results in three coupled modes at the second order, which can only be uncoupled in a uniform flow which will be demonstrated later.

4.2.8 Recovering the intrinsic frequency

It is useful to note that the intrinsic frequency of inertia-gravity waves, as derived in section 3.3, can be recovered from these equations. This is done by assuming there is no flow, i.e. by letting $U = V = 0$ and $R = -N^2 z$ and then setting

$$\begin{pmatrix} \delta \\ \zeta \\ q \end{pmatrix} = \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \\ \tilde{q} \end{pmatrix} e^{i\omega t}, \tag{4.60}$$

where ω is the frequency. This greatly simplifies the system which can now be written as

$$\begin{pmatrix} i\omega & -\frac{m^2f^2+N^2(k^2+l^2)}{\kappa^2} & \frac{k^2+l^2}{f\kappa^2} \\ f & i\omega & 0 \\ 0 & 0 & i\omega \end{pmatrix} \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \\ \tilde{q} \end{pmatrix} = 0. \tag{4.61}$$

Solving this in the same way as we did in the original derivation, that is solving for ω such that the determinant of the matrix vanishes, gives that

$$\omega = 0 \quad \text{or} \quad \omega = \pm\sqrt{\frac{m^2f^2 + N^2(k^2+l^2)}{\kappa^2}}. \tag{4.62}$$

This is the same as was derived earlier, (3.13), which is a nice simple check of the above equations.

4.2.9 Solving the system

Our WKB approach has made it possible to examine the growth of perturbations to solutions of partial-differential equations, by solving ordinary differential equations. See [45, 60] for an alternative approach that uses the pressureless approximation to also lead to a set ordinary differential equations. There are three steps to solving this system. The first step is to find the evolution of the position of the wavepackets involved which is done by solving

$$\frac{D\mathbf{x}}{Dt} = \mathbf{U}, \tag{4.63}$$

where $\mathbf{x}$ is the position vector of the particle and $\mathbf{U}$ is the velocity field. The next step is to solve (4.26) for the evolution of the wavenumbers along the wavepacket's trajectories and the final step is to solve (4.56)-(4.58) for the wavepacket's amplitudes.

In general, this set of ordinary differential equations needs to be solved numerically but some general comments about the behaviour of their solutions can nonetheless be made. It is clear from the potential vorticity's conservation equation that the perturbation's potential vorticity has a constant amplitude, $q(t) = q(0)$. This means that the equations (4.56)-(4.57) for δ and ζ are equivalent to the equations governing a linear oscillator with a time-dependent frequency and time-dependent forcing. In the original system there are two fast modes in the form of the momentum equations for x and y and one slow mode which is a combination of the z momentum equation with

the density equation. These modes correspond to the two fast modes here, δ and ζ and the slow mode given by q, [70].

4.3 Non-dimensionalising

To non-dimensionalise the amplitude equations we re-scale time by taking $t = f^{-1}t^*$. We also take $\mathbf{x} = L_{\text{ref}}\mathbf{x}^*$, $\mathbf{U} = U_{\text{ref}}\mathbf{U}^*$, $\delta = f\delta^*$, $\zeta = f\zeta^*$ and $q = f^3q^*$, where L_{ref} is a reference length, U_{ref} is a reference speed and the starred quantities are non-dimensional. This choice of f^{-1} as a time scale means that the time scale of the waves, and so the background flow, will vary slowly but the reference length and reference velocity are characteristic of the background flow. To non-dimensionalise R we let $R = -N^2z + U_{ref}NR^*$ where N is the Brunt-Väisälä frequency that was defined in chapter 2.2.3. This encapsulates that the vertical component of R is much larger than the horizontal components and that it is still present even in the absence of a flow. Using these definitions, (4.56)-(4.58), after dropping the stars, become

$$\begin{aligned}
\frac{D\delta}{Dt} &= \left(\frac{m^2}{\kappa^2} + \frac{2m^2\epsilon}{\kappa^2(k^2+l^2)}(kl(\partial_x U - \partial_y V) + l^2\partial_x V - k^2\partial_y U)\right)\zeta \\
&+ \left(\frac{m\epsilon}{\alpha\kappa^2}(k\partial_x R + l\partial_y R) + \frac{k^2+l^2}{\alpha\kappa^2}(S - \epsilon\partial_z R)\right)S\zeta \\
&+ \left(\frac{m^2-k^2-l^2}{m\kappa^2}(k\partial_z U + l\partial_z V) + S\frac{l\partial_x R - k\partial_y R}{\alpha m}\right)\epsilon\delta \\
&- \left(\frac{2m^2}{\kappa^2(k^2+l^2)}(kl(\partial_x V + \partial_y U) + l^2\partial_y V + k^2\partial_x U)\right)\epsilon\delta \\
&- \frac{k^2+l^2}{\alpha\kappa^2}q, \qquad (4.64)\\
\frac{D\zeta}{Dt} &= -\alpha\delta, \qquad (4.65)\\
\frac{Dq}{Dt} &= 0, \qquad (4.66)
\end{aligned}$$

where $S = N/f$, which is Prandtl's ratio, $\epsilon = U_{\text{ref}}/fL_{\text{ref}}$, which is the Rossby number and α is now $1 + \epsilon\left(\Omega + \frac{l}{m}\partial_z U - \frac{k}{m}\partial_z V\right)$. It is worth stressing at this point that we have implicitly assumed that $\mu \ll \epsilon$ where μ is the WKB scale separation parameter.

4.4 Removing the singularity at $m = 0$

Since m can pass through 0, the above equations need to be modified to remove the singularity that is located there. As m approaches 0 the WKB approximation actually

breaks down because the scales of the system that have been assumed change. The WKB approximation relies on the scale of the flow being much larger than the WKB wavescale, but as m nears 0 the vertical wavelength of the WKB wave increases towards infinity and so the approximation breaks down. Even though this is the case, we would still like to have equations that will work as m passes through 0 to use in simulations.

The singularity can be removed by introducing a new variable, Γ, defined as

$$\Gamma = \frac{\delta}{m}. \tag{4.67}$$

Taking the material derivative of this definition gives that

$$\frac{D\Gamma}{Dt} = \frac{1}{m}\frac{D\delta}{Dt} - \frac{\Gamma}{m}\frac{Dm}{Dt}, \tag{4.68}$$

in which we can use, from (4.26),

$$\frac{Dm}{Dt} = -\frac{\partial U}{\partial z}k - \frac{\partial V}{\partial z}l. \tag{4.69}$$

Using this along with (4.56) gives that the system, after non-dimensionalising in the same manner as above, is now given by

$$\begin{aligned}
\frac{D\Gamma}{Dt} &= \left(\frac{m}{\kappa^2} + \frac{2m\epsilon}{\kappa^2(k^2+l^2)}(kl(\partial_x U - \partial_y V) + l^2\partial_x V - k^2\partial_y U)\right)\zeta \\
&+ \left(\frac{m\epsilon}{\nu\kappa^2}(k\partial_x R + l\partial_y R) + \frac{k^2+l^2}{\nu\kappa^2}(S - \epsilon\partial_z R)\right)S\zeta \\
&+ \left(\frac{2m}{\kappa^2}(k\partial_z U + l\partial_z V) + S\frac{l\partial_x R - k\partial_y R}{\nu}\right)\epsilon\Gamma \\
&- \left(\frac{2m^2\epsilon}{\kappa^2(k^2+l^2)}(kl(\partial_x V + \partial_y U) + l^2\partial_y V + k^2\partial_x U)\right)\Gamma \\
&- \frac{k^2+l^2}{\nu\kappa^2}q, && (4.70)\\
\frac{D\zeta}{Dt} &= -\nu\Gamma, && (4.71)\\
\frac{Dq}{Dt} &= 0. && (4.72)
\end{aligned}$$

where $\nu = m\alpha = m + \epsilon\,(m\Omega + l\partial_z U - k\partial_z V)$. It is clear that there is now no longer a problem with using these equations in a simulation that has m passing through 0. Although this is the case it is worth noting that sudden jerks in the amplitudes can still occur as m passes through 0. This happens when the remaining terms in the

expression for ν are small, hence making ν also close to zero. Since ν only appears in the denominator in the above equations this can cause sudden, undesired, growth. It is also interesting to note that these equations look very similar to how they did before the change of variable, apart from the factor in front of $(k\partial_z U + l\partial_z V)$ on the third line of (4.70), which appears to have undergone some neat cancelation that has greatly simplified it. This cancelation comes from the last term of (4.68) since the expression for $D_t m$ is the same as the factor involved here. This term in front of the factor in question then becomes

$$\frac{m^2 - k^2 - l^2}{m\kappa^2} + \frac{1}{m} = \frac{m^2 - k^2 - l^2 + k^2 + l^2 + m^2}{m\kappa^2} = \frac{2m}{\kappa^2}. \tag{4.73}$$

4.5 Energy

In some situations it is useful to have a measure of the energy of the wavepacket. This can be defined as

$$E = \lim_{L\to\infty} \frac{1}{8L^3} \iiint_{-L}^{L} \left(u'^2 + v'^2 + w'^2 + \rho'^2/N^2\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.74}$$

where

$$u' = \mathrm{Re}\left(\hat{u}e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}\right). \tag{4.75}$$

This integral can be tackled in sections by looking at it term by term. Evaluating the first term gives that

$$\lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} u'^2 \,\mathrm{d}x = \lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} \left(\hat{u}^2 e^{2ikx} + \hat{\bar{u}}^2 e^{-2ikx} + 2|\hat{u}|^2\right) \mathrm{d}x. \tag{4.76}$$

By integrating by parts we obtain that the first two parts of this integral vanish as $L \to \infty$, and so we are left with

$$\lim_{L\to\infty} \frac{1}{2L} \int_{-L}^{L} u'^2 \,\mathrm{d}x = 2|\hat{u}|^2, \tag{4.77}$$

and hence integrating over the full domain gives that

$$\lim_{L\to\infty} \frac{1}{8L} \iiint_{-L}^{L} u'^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 8|\hat{u}|^2. \tag{4.78}$$

Applying this procedure to the other terms in the original energy integral gives that

$$E = 8\left(|\hat{u}|^2 + |\hat{v}|^2 + |\hat{w}|^2 + \frac{|\hat{\rho}|^2}{N^2}\right), \tag{4.79}$$

where, from earlier in this chapter we have that

$$\hat{u} = \frac{i(l\zeta - k\delta)}{k^2 + l^2}, \tag{4.80}$$

$$\hat{v} = -\frac{i(k\zeta + l\delta)}{k^2 + l^2}, \tag{4.81}$$

$$\hat{w} = \frac{i\delta}{m}, \tag{4.82}$$

$$\hat{\rho} = -iS\frac{\partial_x R\left(\frac{l\delta}{m} + \frac{m(k\zeta+l\delta)}{k^2+l^2}\right) + \partial_y R\left(\frac{m(l\zeta-k\delta)}{k^2+l^2} - \frac{k\delta}{m}\right) + (\frac{S}{\epsilon} - d_z R)\zeta - \frac{\hat{q}}{S\epsilon}}{(1/\epsilon + \Omega)m + l\partial_z U - k\partial_z V}. \tag{4.83}$$

4.6 Conclusion

In this chapter we have derived a set of equations that will enable us to study the evolution of a wavepacket that has been introduced into a geophysical flow. We have also tried to get a general feel for how the equations behave. The next step will be to apply these equations to very simple 'toy' models of flows to both get a feel for the results they give and also to check that they do not give any unexpected results. After analysing very simple flows we will start to increase the complexity of the flows involved and also look at some time-dependent flows. These equations will be used to try to give an insight into the generation of inertia-gravity waves in these situations.

Chapter 5

Simple flows

5.1 Introduction

Now that we have derived the equations that govern the evolution of a small amplitude wavepacket placed in a geophysical flow, we can use them to study inertia-gravity wave generation in various flows. In this chapter we consider highly idealised, time independent flows starting with the case where there is no flow and then gradually building up the complexity. The simplicity of these flows greatly reduces the number of terms in the derived equations, which in turn greatly simplifies the analysis of the system. As the complexity of the flow increases, more and more terms are added back into the system, hence adding more and more complexity to the analysis.

5.2 No Flow

First we consider the simplest case possible, that of no flow, for which

$$U = 0, \tag{5.1}$$

$$V = 0, \tag{5.2}$$

$$R = -N^2 z. \tag{5.3}$$

In this situation it is clear that the gradients of the flow's velocity field are zero and hence from (4.26) that the wavenumbers remain constant. If for simplicity we set $q = 0$,

then the non-dimensional system, (4.65)-(4.66), becomes

$$D_t\delta = \left(\frac{m^2 + S^2(k^2+l^2)}{\kappa^2}\right)\zeta, \tag{5.4}$$

$$D_t\zeta = -\delta, \tag{5.5}$$

$$D_t q = 0, \tag{5.6}$$

where $S = N/f$. This has solution

$$\delta(t) = \zeta(0)\sigma\sin\sigma t + \delta(0)\cos\sigma t, \tag{5.7}$$

$$\zeta(t) = \frac{1}{\sigma}\left(\zeta(0)\sigma\cos\sigma t - \delta(0)\sin\sigma t\right), \tag{5.8}$$

$$q(t) = 0, \tag{5.9}$$

where

$$\sigma = \pm\sqrt{\frac{m^2 + S^2(k^2+l^2)}{\kappa^2}} \quad \text{or} \quad \sigma = 0. \tag{5.10}$$

These solutions, which are shown in figure 5.1, are constant amplitude inertia-gravity waves that are periodic, with frequency σ. As a check of the above solutions we note that the frequency of these inertia-gravity waves, σ, matches up with the intrinsic frequency for inertia-gravity waves, after scaling by f, that was derived in section 3.3.

A quick inspection of the energy equation reveals that the wavepacket's energy is conserved, which is what is to be expected when there are no outside forces acting on it.

5.3 Pure Strain Field

We now consider the evolution of a wavepacket in a pure strain flow. This is a flow that has been extensively studied in the context of wave capture and has been used to show that a wavepacket can behave like a pair of oppositely signed vortices, [7, 10]. The velocity field of this flow is given by

$$U = \beta x, \tag{5.11}$$

$$V = -\beta y, \tag{5.12}$$

where β is a positive constant. The streamlines and velocity field for this flow are shown in figure 5.2.

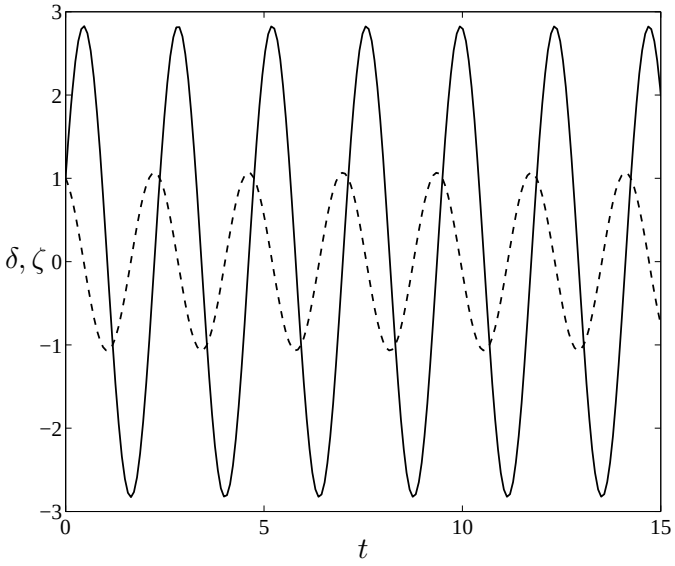

Figure 5.1: Inertia-gravity-waves that can be seen in the amplitudes of δ, solid line, and ζ, dashed line.

The fact that there are no vertical gradients of the horizontal velocity field implies that there are no horizontal gradients of R and hence the expression for R reduces to $R = -N^2 z$. An expression for the pressure, P, can now be derived from the original Boussinesq equations, (2.27)-(2.31). This is done by introducing the expressions for U and V into (2.27) and (2.28) giving that

$$\beta^2 x + f\beta y = -\partial_x P, \tag{5.13}$$

$$\beta^2 y + f\beta x = = -\partial_y P. \tag{5.14}$$

Solving this set of equations yields

$$P = -f\beta xy - \frac{\beta^2}{2}(x^2 + y^2) + g(z), \tag{5.15}$$

where $g(z)$ is an arbitrary function in terms of z. From (2.29), $R = -\partial_z p$, meaning that we need $g(z) = N^2 z^2/2$, which gives that

$$P = -f\beta xy - \beta^2 \left(\frac{x^2}{2} + \frac{y^2}{2}\right) + \frac{N^2 z^2}{2}. \tag{5.16}$$

As can be seen in figure 5.2, or intuitively interpreted from the flow's velocity equations, (5.11)-(5.12), in this flow the x-axis is an extension axis and the y-axis is a

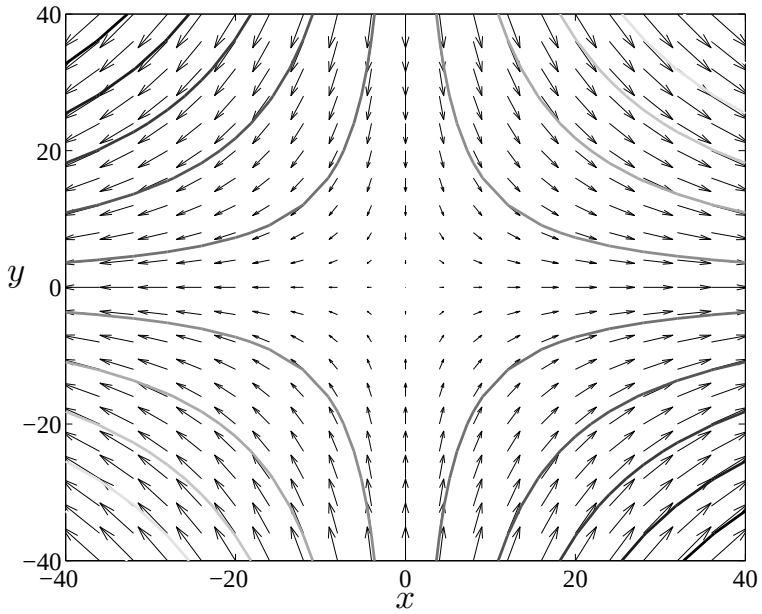

Figure 5.2: The streamlines and velocity field of the pure strain field (5.11)-(5.12).

contraction axis. This intuitively suggests that the wavepacket's evolution is such that the wavevector will align itself with the contraction axis. This particular case, when the contraction and expansion axes are perpendicular to each other, is an example of a general theory about the direction of the wavevector in a situation where both a contraction axis and expansion axis exist. It has been shown that for the general case of this scenario, where the stream function is given by $\psi = -\beta xy + 0.5\gamma(x^2 + y^2)$, that if $\gamma > 0$ then the axis of extension is turned counterclockwise by $0.5 \arcsin(\gamma/\alpha)$ whilst the contraction axis is turn clockwise by the same angle, [10]. Our example is recovered by setting $\gamma = 0$ in this stream function. In this case it has been shown that the growing horizontal wavevector always aligns perpendicularly to the extension axis. It is also worth linking this theory to the brief discussion of the analogy to passive tracers in the previous chapter, section 4.2.3. There it was suggested that we could think of the phase of the wave packet, θ, as being the concentration of a passive scalar and so the wavenumber, $\mathbf{k}$, behaves like its gradient. In this pure strain flow we can use our intuition to imagine that the passive scalar would get squeezed to a filament like structure along the expansion axis, which would again lead to the conclusion that the wavevector will align itself with the contraction axis.

As the wave crests move with the flow in the direction of the expansion axis and are

stretched, their wavenumber, k, will decrease. In contrast, the wavenumber, l, of the waves travelling along the contraction axis will increase. This can be seen analytically from the wavenumber equation, (4.26), which in this case becomes

$$\begin{pmatrix} \dot{k} \\ \dot{l} \\ \dot{m} \end{pmatrix} = \begin{pmatrix} -\beta & 0 \\ 0 & \beta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} k \\ l \end{pmatrix}. \tag{5.17}$$

Solving this system of equations gives that

$$k = k(0)e^{-\beta t}, \tag{5.18}$$

$$l = l(0)e^{\beta t}, \tag{5.19}$$

$$m = m(0), \tag{5.20}$$

which confirms the discussion above.

Now that we know how the wavenumbers behave we can look at how the wavepacket's amplitudes behave. For the pure strain field given by (5.11)-(5.12) the equations for the wavepacket's amplitudes, after taking $q = 0$, reduce to

$$\begin{aligned} \frac{D\delta}{Dt} &= \left(\frac{m^2 + S^2(k^2 + l^2)}{\kappa^2} + \frac{4\epsilon\beta m^2 kl}{\kappa^2(k^2 + l^2)} \right) \zeta \\ &- \left(\frac{2\epsilon m^2 \beta}{\kappa^2(k^2 + l^2)} (k^2 - l^2) \right) \delta, \end{aligned} \tag{5.21}$$

$$\frac{D\zeta}{Dt} = -\delta. \tag{5.22}$$

Since k is exponentially decreasing and l is exponentially increasing, these equations reduce to

$$D_t\delta \sim S^2\zeta, \tag{5.23}$$

$$D_t\zeta \sim -\delta, \tag{5.24}$$

as $t \to \infty$. The solution to these equations can be written as

$$\begin{pmatrix} \delta \\ \zeta \end{pmatrix} = e^{\sigma t} \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix}, \tag{5.25}$$

where σ is the growth rate. Introducing this equation into (5.23) and (5.24) gives the eigenvalue problem

$$\sigma \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix} = \begin{pmatrix} 0 & S^2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix}, \tag{5.26}$$

which when solved yields

$$\sigma = \pm iS. \tag{5.27}$$

This shows that as t gets large the wavepacket's amplitudes, δ and ζ, become periodic with a constant amplitude. This also shows that a pure strain field increases the intrinsic frequency of the wavepacket, from that of the intrinsic frequency of an inertia-gravity wave, derived in section 3.3, to S.

Although we have just shown that in the long time limit the solutions of the wavepacket's amplitude equations are constant amplitude periodic functions, initially the solutions are periodic with growing amplitudes. These solutions are buoyancy oscillations that are created because $\sqrt{k^2 + l^2}/m$ increases and causes a displacement of the fluid, which in turn initiates the oscillations. See section 2.2.3 for a fuller discussion. The fact that in this situation we have the horizontal wavenumber increasing while m remains constant is a consequence of having no shear present in the flow. The amplitudes of these solutions grow until t is sufficiently large before saturating to become constant amplitudes. This in turn leads to the energy initially growing and then levelling out to a constant. An example of this is shown in figure 5.3 where the wavepacket was initially placed at $(1, 1, 1)$ with initial conditions $k(0) = l(0) = -m(0) = 1$, $\delta(0) = \zeta(0) = 1$, $S = 10$, $\epsilon = 1$ and $\beta = 0.01$.

5.4 Transverse shear

It is also of interest to consider the effect of a transverse shear on a wavepacket. In this case, the velocity field is given by

$$U = \Sigma z, \tag{5.28}$$

$$V = 0, \tag{5.29}$$

where Σ is a constant. This flow is illustrated in figure 5.4. By using (2.28) we find that the pressure is given by $P = -f\Sigma zy$ which can be rearranged to give $\Sigma = -\partial_{yz}P/f$. We also have from the definition of the flow that $\Sigma = \partial_z U$ and so combining these two

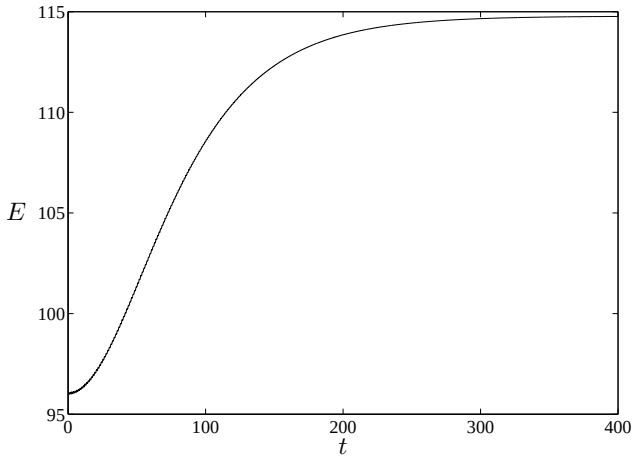

Figure 5.3: An example of the temporary growth of the energy of a wavepacket in a pure strain field.

equations and using the fact that $R = -\partial_z p$ gives that $\partial_z U = R_y/f$. From this we derive that $R = f\Sigma y + g(z)$ where $g(z)$ is an arbitrary function in terms of z. We know from the previous flow that $g(z)$ is given by $-N^2 z$, and so $R = f\Sigma y - N^2 z$.

Since the velocity field is independent of x or y, it is clear that the horizontal wavenumbers are going to remain constant. Solving for the vertical wavenumber, as done previously, gives that

$$m = -\Sigma k(0)t + m(0). \tag{5.30}$$

By using what we have found so far, the equations for the amplitudes δ and ζ, (4.56) and (4.57), can now be simplified to

$$\begin{aligned} \frac{D\delta}{Dt} &= \left(\frac{m^2 f}{\kappa^2} + \frac{k^2+l^2}{\alpha\kappa^2}N^2 + \frac{mlf\Sigma}{\alpha\kappa^2}\right)\zeta \\ &+ \left(\frac{m^2-k^2-l^2}{m\kappa^2}k\Sigma - \frac{kf\Sigma}{\alpha m}\right)\delta, \qquad (5.31) \\ \frac{D\zeta}{Dt} &= -\alpha\delta, \qquad (5.32) \end{aligned}$$

where $\alpha = f + l\Sigma/m$ and we have again taken $q = 0$. As $t \to \infty$, $\alpha \to f$ and $m \to -\infty$ and therefore

$$\begin{aligned} D_t\delta &\sim f\zeta, \qquad (5.33) \\ D_t\zeta &\sim -f\delta. \qquad (5.34) \end{aligned}$$

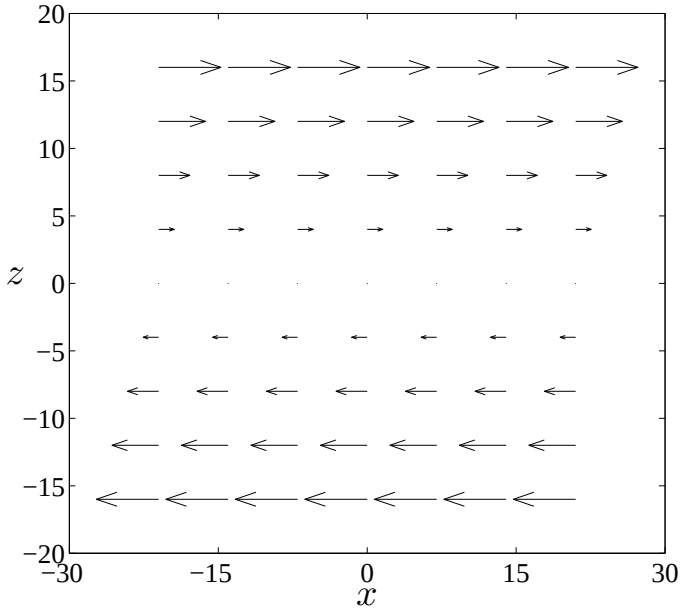

Figure 5.4: The velocity field of a transverse shear flow.

Solving these equations yields $2\pi/f$ periodic solutions of constant amplitude. Hence, as in the pure strain field, there may be some initial growth of the wavepacket placed in this flow as the frequency rises, but it will be short lived as the wavepacket will reach a stable equilibrium after a long enough time.

5.5 Strain and Shear

The next level of complexity is achieved by combining the previous two flows to produce a strain and shear flow. This is a flow that was used to study the effect of vertical scale reduction caused by quasi-horizontal stirring in a large-scale flow and the implications that this has on dissipation and mixing, [24]. In this case the velocity field is given by

$$U = \beta x, \tag{5.35}$$

$$V = -\beta y + \Sigma z, \tag{5.36}$$

where β and Σ are positive constants.

Following a similar procedure to the one followed in the pure strain case, we can derive expressions for the pressure, P, and the density, R, from the original Boussinesq equations, (2.27)-(2.31). This is done by putting the expressions for U and V into

(2.27) and (2.28), giving that

$$\beta^2 x + f\beta y - f\Sigma z = -\partial_x P, \tag{5.37}$$

$$\beta^2 y + f\beta x - \beta\Sigma z = -\partial_y P. \tag{5.38}$$

Solving this set of equations gives that

$$P = -f\beta xy - \frac{\beta^2}{2}(x^2 + y^2) + z\Sigma(fx + \beta y) + g(z), \tag{5.39}$$

where $g(z)$ is an arbitrary function in terms of z. From (2.29), $R = -\partial_z P$, and so

$$R = -\Sigma(fx + \beta y) + h(z), \tag{5.40}$$

where $h(z)$ is also an arbitrary function in terms of z. Since we know that the term solely in terms of z must be $-N^2 z$, we have that $h(z) = -N^2 z$ and so $g(z) = N^2 z^2/2$. Combining these equations gives that

$$P = z\Sigma(fx + \beta y) - f\beta xy - \frac{\beta^2}{2}(x^2 + y^2) + \frac{N^2 z^2}{2}, \tag{5.41}$$

$$R = -\Sigma(fx + \beta y) - N^2 z. \tag{5.42}$$

As we would hope, these equations are the same as the pure strain case when we remove the shear by setting Σ to zero.

Solving (4.26) for the flow's wavenumbers gives that

$$k(t) = k(0)e^{-\beta t}, \tag{5.43}$$

$$l(t) = l(0)e^{\beta t}, \tag{5.44}$$

$$m(t) = \frac{\Sigma l(0)(1 - e^{\beta t})}{\beta} + m(0). \tag{5.45}$$

It is worth noting that this is the same as the pure strain case, which can be obtained by taking $\Sigma \to 0$, apart from the added exponential growth term in the vertical wavenumber's equation, which has been added by the shear.

We are now in a position to study the evolution of the amplitudes, δ, ζ and q. Using (5.35)-(5.36) and (5.41)-(5.42), the equations of these amplitudes, originally given in

(4.65)-(4.66), become

$$\begin{aligned}
\frac{D\delta}{Dt} &= \left(\frac{m^2}{\kappa^2} + \frac{4\epsilon m^2 kl\beta}{\kappa^2(k^2+l^2)} - \frac{\epsilon m S\Sigma}{\alpha\kappa^2}(kf+l\beta) + \frac{S(k^2+l^2)}{\alpha\kappa^2}(S+\epsilon N^2)\right)\zeta \\
&+ \left(\frac{m^2-k^2-l^2}{m\kappa^2}l\Sigma + S\frac{k\Sigma\beta - lf\Sigma}{\alpha m} - \frac{2m^2\beta}{\kappa^2(k^2+l^2)}(k^2-l^2)\right)\epsilon\delta \\
&- \frac{k^2+l^2}{\alpha\kappa^2}q, \qquad (5.46)\\
\frac{D\zeta}{Dt} &= -\alpha\delta, \qquad (5.47)\\
\frac{Dq}{Dt} &= 0, \qquad (5.48)
\end{aligned}$$

where $\alpha = 1 - \epsilon\Sigma k/m$.

Using (5.43)-(5.45) and noting that as $t \to \infty$, $\alpha \to 1$, the wavepacket's amplitude equations in the limit of $t \to \infty$ become

$$\begin{aligned}
D_t\delta &\sim A\delta + B\zeta - Cq, \qquad (5.49)\\
D_t\zeta &\sim -\delta, \qquad (5.50)\\
D_t q &\sim 0, \qquad (5.51)
\end{aligned}$$

where

$$A = \beta, \quad B = \frac{\Sigma^2 + S^2\beta^2}{\beta^2+\Sigma^2} \quad \text{and} \quad C = \frac{\beta^2}{\Sigma^2+\beta^2}, \qquad (5.52)$$

are constants. An interesting point to note here is that if we remove the shear in the above expression for $D_t\delta$, that is, take $\Sigma \to 0$, it does not become the same as the one given in the case of the pure strain flow. This is surprising since the pure strain flow is the same as the strain and shear flow that we are dealing with here in the limit of $\Sigma \to 0$. The reason for this is that here we have taken the limit $t \to \infty$ and then the limit $\Sigma \to 0$. When taking the limits this way round, the term arising from $\partial_z V$ in the flow with the shear contributes to the asymptotics arising from taking $t \to \infty$, rather than when the limits are taken the other way round, where it is removed before taking $t \to \infty$ and so makes no contribution.

If we consider the case where $q = 0$ the solution can be written as

$$\begin{pmatrix} \delta \\ \zeta \end{pmatrix} = e^{\sigma t}\begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix}, \qquad (5.53)$$

where σ is the growth rate. Introducing this into (5.49) and (5.50) gives the eigenvalue

problem

$$\sigma \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix} = \begin{pmatrix} A & B \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\delta} \\ \tilde{\zeta} \end{pmatrix}, \tag{5.54}$$

which when solved yields

$$\sigma = \frac{A \pm \sqrt{A^2 - 4B}}{2}. \tag{5.55}$$

To take this further and check whether the energy of the wavepacket can grow, we can use the equations of the wavenumbers, (5.43)-(5.45), to see that as t increases, the equations given in (4.80)-(4.83) become

$$|\hat{u}|^2 \sim \left| \frac{\zeta}{e^{\beta t}} \right|^2, \tag{5.56}$$

$$|\hat{v}|^2 \sim \left| \frac{\delta}{e^{\beta t}} \right|^2, \tag{5.57}$$

$$|\hat{w}|^2 \sim \left| \frac{\delta}{e^{\beta t}} \right|^2, \tag{5.58}$$

$$|\hat{\rho}|^2 \sim \left| \frac{\zeta}{e^{\beta t}} \right|^2. \tag{5.59}$$

In each case it is clear that for growth we need either $\delta > e^{\beta t}$ or $\zeta > e^{\beta t}$. From (5.53) this is equivalent to requiring that $\mathrm{Re}\,\sigma > \beta$ which in turn, from (5.55), is equivalent to requiring

$$\mathrm{Re}\,\frac{A \pm \sqrt{A^2 - 4B}}{2} > \beta. \tag{5.60}$$

If $A^2 - 4B$ is negative then the real part of this is $\beta/2$, which is clearly never greater than β since β is always positive. On the other hand if $A^2 - 4B$ is positive, we now need

$$4\beta^2 - 4\beta A + 4B < 0, \tag{5.61}$$

but from (5.52), $A = \beta$ and so this simplifies to

$$4B < 0. \tag{5.62}$$

Since β and σ are always both positive, B is also always positive, so this inequality will never hold. This means that although there might be some initial growth in the energy, it will eventually vanish as $t \to \infty$ and so there is no long time growth in the system.

We may find by measuring the energy of the waves that there is no growth but this is

not necessarily the most physically important measure of the waves' amplitudes. Even when the energy of the waves is small they can still overturn and break if there are large vertical gradients of density. When the waves overturn and break they release energy into the surrounding fluid and create a mixing effect. We can find a condition that will enable us to know when to expect the waves to overturn. From $\partial_z(-N^2 z + \rho) = -N^2 + \partial_z \rho$ we see that if $\partial_z \rho$ grows larger than N^2 then the effect of the stratification will be reversed, which will cause the waves to break. To determine when this will happen, we note that

$$\partial_z \rho \sim m\rho \sim \frac{m\zeta}{e^{\beta t}} \sim m e^{(\sigma-\beta)t}, \tag{5.63}$$

with the second step coming from (5.59). From (5.45) it is clear that as $t \to \infty$, $m \to \pm e^{\beta t}$ and so

$$m e^{(\sigma-\beta)t} \to e^{\sigma t} \to \pm\infty. \tag{5.64}$$

This shows that if $m > 0$, and hence $m \to \infty$, then $\partial_z \rho$ will become larger than N^2, causing the waves to overturn. This shows that there is an instability of some kind in this system that occurs when $m > 0$.

5.6 Frontogenesis flow

The pure strain field can be taken further in a different way by applying a slightly different form of vertical shear. The procedure that we have applied a few times above is applied to a flow that is given by

$$U = -\beta x + \Sigma z e^{-\beta t}, \tag{5.65}$$

$$V = \beta y - 2\Sigma \frac{\beta}{f} z e^{-\beta t}. \tag{5.66}$$

This flow is similar to the pure strain flow that was studied earlier but with the contraction and expansion axes swapped around. The main difference in the two flows comes from an exponentially decaying shear term that has been added. This is a flow that has been extensively studied in relation to the generation of inertia-gravity waves accompanying frontogenesis, [19, 20], which is the formation of a new atmospheric front or the regeneration of an old one. Frontogenesis is a balanced model with no inertia-gravity waves built into the model, but in the following analysis we will see what happens when we introduce some.

As a first step we can find expressions for P and R by first using the original Boussinesq equations. Putting this flow into (2.27) gives the same as for the pure strain flow, whilst putting it into (2.28) gives that

$$\beta^2 y - f\beta x + f\Sigma z e^{-z\beta t} = -\partial_y P. \tag{5.67}$$

Noticing that the first two terms of this expression are also the same as the pure strain field lets us write P as $P = P_1 + P_2$, where P_1 are the terms that are the same as the pure strain case and P_2 are the extra terms supplied by the final term above. Integrating the above equation gives that $P_2 = -f\Sigma z y e^{-\beta t}$ and so

$$P = f\beta xy - \beta^2\left(\frac{x^2}{2} + \frac{y^2}{2}\right) - f\Sigma z y e^{-\beta t} + A(z,t), \tag{5.68}$$

where $A(z,t)$ is an unknown function in z and t. Using (2.29) can now give that

$$R = f\Sigma y e^{-\beta t} - N^2 z + \partial_z A, \tag{5.69}$$

which can be used in (2.30) to get that

$$\partial_{tz} A = 2\Sigma^2 \beta z e^{-2\beta t}. \tag{5.70}$$

After integrating, we find that

$$A(z,t) = -\frac{\Sigma^2}{2} z^2 e^{-2\beta t}, \tag{5.71}$$

and hence we have that

$$P = f\beta xy - \beta^2\left(\frac{x^2}{2} + \frac{y^2}{2}\right) - f\Sigma z y e^{-\beta t} - \frac{\Sigma^2}{2} z^2 e^{-2\beta t}, \tag{5.72}$$

$$R = f\Sigma y e^{-\beta t} - \Sigma^2 z e^{-2\beta t} - N^2 z. \tag{5.73}$$

Solving the differential equations that govern the evolution of the wave numbers for this flow gives that

$$k(t) = k(0)e^{\beta t}, \tag{5.74}$$

$$l(t) = l(0)e^{-\beta t}, \tag{5.75}$$

$$m(t) = -\Sigma k(0)t + \frac{\Sigma l(0)}{f}\left(1 - e^{-2\beta t}\right) + m(0), \tag{5.76}$$

where again the similarities can be seen with previous flows in this chapter.

The expressions that have been derived above can now be used to study the behaviour of δ and ζ. Introducing the expressions found above into the dimensional amplitude equations, (4.56)-(4.58), gives that

$$\begin{aligned}\frac{D\delta}{Dt} &= \left(\frac{m^2 f}{\kappa^2} - \frac{4m^2 kl\beta}{\kappa^2(k^2+l^2)} + \frac{k^2+l^2}{\alpha\kappa^2}(\Sigma^2 e^{-2\beta t} + N^2) + \frac{mlf\Sigma e^{-\beta t}}{\alpha\kappa^2}\right)\zeta \\ &+ \left(\frac{m^2-k^2-l^2}{m\kappa^2}\Sigma e^{-\beta t}(k - 2\beta l/f) - \frac{kf\Sigma e^{-\beta t}}{\alpha m} - \frac{2m^2\beta}{\kappa^2(k^2+l^2)}(l^2-k^2)\right)\delta \\ &- \frac{k^2+l^2}{\alpha\kappa^2}q, \end{aligned} \tag{5.77}$$

$$\frac{D\zeta}{Dt} = -\alpha\delta, \tag{5.78}$$

$$\frac{Dq}{Dt} = 0, \tag{5.79}$$

where $\alpha = f + \Sigma e^{-\beta t}(l + 2k\beta/f)/m$. A quick inspection of these equations reveals that if the vertical shear is removed by setting $\Sigma = 0$, then they are the same as the equations found for the pure strain case with just a few minus signs changed due to the swapping of the contraction and expansion axes.

From (5.74)-(5.76) it is clear that as $t \to \infty$, $|k| \to \infty$ exponentially, $|l| \to 0$ and $|m| \to \infty$ linearly. Applying these to the above equations for δ and ζ gives that

$$D_t\delta \to N^2/f, \tag{5.80}$$

$$D_t\zeta \to -f\delta, \tag{5.81}$$

in the limit of $t \to \infty$. It is worth noting here that these equations are the same as those that were found for the pure strain flow as $t \to \infty$. This can be understood by realising that the vertical shear applied to this system is only temporary. From (5.65) and (5.66) it is clear that the magnitude of the shear is decreasing exponentially in time, so it falls out of the equations when we look at large t. Solving this system for δ and ζ for large t gives that

$$\delta(t) \sim \frac{bN}{f}\sin(Nt) + a\cos(Nt), \tag{5.82}$$

$$\zeta(t) \sim -\frac{af}{N}\sin(Nt) + b\cos(Nt), \tag{5.83}$$

$$q(t) \sim q(0), \tag{5.84}$$

where a and b are constants related to the initial values of δ and ζ. This shows that in the long time limit ζ and δ will settle down to periodic functions with constant amplitudes, while q will remain a constant.

We can use the same technique as we used for the case of the strain and shear flow to check if the waves in this flow will overturn. By using the expressions for the wavenumbers and R that we have derived above and introducing them into (4.83) we find that as $t \to \infty$, $\rho \to \zeta/m$. Hence from the first few parts of (5.63) we have that $\partial_z \rho \to \zeta$, which from above tends to a periodic function with a constant amplitude. This means that the waves will never overturn in this flow.

5.7 Conclusion

We have looked at flows from the very simple to the slightly more complex and analysed them using the equations derived in the previous chapter. This has shown that in these fairly basic flows there is no growth in the wavepacket's amplitudes in the long time limit or in the energy of the wavepacket. Although this is the case we have found a few situations where we have observed interesting physical effects such as transient growth and wave overturning. This leads us to conclude that unless there is some interesting feature of the flow then there are unlikely to be any instabilities that lead to growth in the long term but, given the right conditions, there can be other interesting features of the system that lead to different phenomena occurring with instabilities forming as a result.

Since all these model flows have just been 'toy' models, the analysis has been straightforward and so we have always been able to carry it out by hand. The next step is to increase the complexity of the flows which will in turn increase the complexity of the analysis and possibly yield a situation where we will find some long term growth. This will be considered in the next chapter by first studying a flow that involves a sharp change in trajectory and examining the effect that this has on the wavepacket, before

looking at a more complex, time-dependent, flow.

Chapter 6

Point-vortex model

6.1 Introduction

The aim of this chapter is to build on the results of the previous chapter by illustrating spontaneous generation of inertia-gravity waves in flows of a more complex nature that have a richer time dependence. This is achieved by considering non-uniform flows. These flows have a time-dependent Lagrangian strain and hence time-dependent coefficients in the equations for the wavenumbers and the wavepacket's amplitudes. In this chapter, we use quasi-geostrophic point vortices to generate crude models of flows that interest us. We derive the theory and equations for the evolution of the point vortices and then study the spontaneous growth in two distinct systems.

The first system to be studied is a flow created by a point vortex dipole. As done for previous flows, this involves tracking the evolution of a wavepacket as it moves through the flow and following the solution to its amplitude equations. The unique part of this analysis uses the fact that at large distances from the dipole, the flow's velocity field is uniform. This allows us to decompose the solutions to the equations for the wavepacket's amplitudes into three modes, namely the vortical mode and the two gravity-wave modes. This enables the transfer of energy from the vortical mode to the gravity wave modes to be monitored. The system is initialised with the wavepacket placed far enough away from the dipole that the flow around it is uniform. The wavepacket's initial amplitudes are chosen so that only the vortical mode is excited, i.e. there are no inertia-gravity waves present. The moving dipole will then approach and pass the wavepacket leaving it once more in a uniform flow. By studying the three

modes of the wavepacket we can see if any energy has transferred between them. If energy has transferred from the vortical mode to the gravity-wave modes then we have a reliable indication of spontaneous inertia-gravity wave generation.

As well as studying trajectories that pass around the point vortex dipole, it is also of interest to study the closed elliptical trajectories that exist within the dipole. We will briefly examine the effect that these periodic, almost elliptical, streamlines have on a wavepacket to see if there are any signs of spontaneous growth. Although we will not delve into this scenario in any great depth here, the results can be used to motivate further study of elliptical flows.

The second system created by point vortices that we will study, in considerably less detail than the first system, is that of a time-dependent flow generated by multiple point vortices. These flows are set up with point vortices of different strengths and circulation directions surrounding the wavepacket in three dimensional space. This will have the effect of creating a seemingly random, time dependent strain field around the wavepacket which we expect to lead to a growth in the wavepacket's wavenumbers and amplitudes. By initialising the system so that only one of the two gravity wave modes is excited and then running the simulations over a long time period, we expect to be able to see unbalanced instabilities leading to a growth of the amplitudes and we hypothesise a link between this growth and the frequency of the wavepacket.

6.2 Point vortices

When the non-zero vorticity in a system is confined to a finite number of points and the vorticity is zero everywhere else, understanding the dynamics of the system reduces to understanding the evolution of these discrete points called point vortices.

To understand the evolution of a point vortex over time in a three-dimensional, quasi-geostrophic system, we derive the equations of motion for the velocity in the x direction, U, the y direction, V, and in the z direction, W. It is worth noting here that this flow is actually only a solution of the quasi-geostrophic fluid equations in the limit of large f and N but it can be used as an approximate model of a flow. This is because the qualitative properties of the wavepacket's evolution will not be affected by the details of the flow. This effectively means that we are using a flow that does not quite fit the situation to move the wavepacket around without it having an effect on the wavepacket's intrinsic properties.

These point vortices are studied under the quasi-geostrophic approximation by first introducing the stream function, $\psi(x, y, z, t)$, [49], which is defined so that

$$U = \frac{\partial \psi}{\partial y}, \tag{6.1}$$

$$V = -\frac{\partial \psi}{\partial x}, \tag{6.2}$$

$$R = -f\frac{\partial \psi}{\partial z} - N^2 z, \tag{6.3}$$

where R is the scaled density that was defined in (4.5). Note that in the quasi-geostrophic approximation the vertical velocity is taken as zero and so it is not included here.

There are two important quantities that must be considered when studying point vortices under these conditions. The first of these quantities is the potential vorticity which, under the quasi-geostrophic approximation, is expressed after scaling

$$\frac{N}{f} z \to z, \tag{6.4}$$

[56], by

$$q = \nabla^2 \psi, \tag{6.5}$$

where ∇^2 is the three-dimensional Laplacian.

The second of these important quantities is the point vortex strength. The potential vorticity is concentrated on a point, so we can write that

$$q = \kappa \delta(x - X(t))\delta(y - Y(t))\delta(z - Z(t)), \tag{6.6}$$

where κ is the strength of the point vortex and δ is the Dirac delta function. In this equation the capital letters are the position of the point vortex. It is clear from this equation that by integrating over the whole domain we can express κ as

$$\kappa = \iiint q \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z. \tag{6.7}$$

Since the potential vorticity is a point source it makes solving the Poisson equation, (6.5), for the stream function relatively easy. A Poisson equation usually has the integral of a Green's function as a solution, but here the point source nature of the potential vorticity means that the solution we require is just a Green's function. Solving for this

Green's function over an infinite domain gives that

$$\psi = -\frac{1}{4\pi}\frac{\kappa}{\sqrt{(x-X(t))^2+(y-Y(t))^2+(z-Z(t))^2}} = -\frac{1}{4\pi}\frac{\kappa}{r}, \tag{6.8}$$

where r is the distance from the point vortex. Since the Poisson equation is linear, we can apply this solution to multiple point vortices, resulting in the i^{th} point vortices' stream function being given by

$$\psi_i = -\frac{1}{4\pi}\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{\kappa_j}{\sqrt{(x_i-x_j(t))^2+(y_i-y_j(t))^2+(z_i-z_j(t))^2}}, \tag{6.9}$$

where n is the number of point vortices involved.

We are now in a position to determine the horizontal velocity field associated with this system. Differentiating (6.8) according to (6.1)-(6.2) gives that

$$U = -\frac{\kappa}{4\pi}\frac{y-Y(t)}{r^3}, \tag{6.10}$$

$$V = \frac{\kappa}{4\pi}\frac{x-X(t)}{r^3}. \tag{6.11}$$

A similar procedure provides the scaled density given in (6.3). After taking into account the scaling applied to z in (6.4), the scaled density is expressed by

$$R = N\left(\frac{\kappa}{4\pi}\frac{z-Z(t)}{r^3}+fz\right). \tag{6.12}$$

As with the stream function, these expressions can be adjusted in the same way to provide expressions for systems where there are multiple point vortices.

The evolution of a system of point vortices in the three-dimensional, quasi-geostrophic approximation is now governed by a system of ordinary differential equations in terms of the point vortices' strengths and their positions.

6.3 Dipole

An interesting, steady, quasi-geostrophic flow to study is one that is induced by a dipole of potential vorticity, [6]. In this case we are studying the system with the aim of finding spontaneous generation of inertia-gravity waves, that is with $q \neq 0$. The flow is created by two point vortices of the same strength that are rotating in opposite directions. For

the sake of simplicity, but without loss of generality, we will consider the case where both point vortices are on the x-y plane, i.e. with $z = 0$. We also initially place the two point vortices on the x-axis with separation d, with their positions mirroring each other in the y-axis. This has the effect of creating a dipole that moves through the fluid with a velocity proportional to the strength of the point vortices. Figure 6.1 shows this configuration for $d = 1$. It also shows the streamlines around the point vortices and the velocity field that the vortices are creating.

If we label the point vortex on the left in figure 6.1 as vortex 1 and the point vortex on the right as vortex 2, then the stream function of vortex 1 is given by

$$\psi_1 = -\frac{\kappa_2}{4\pi} \frac{1}{\sqrt{(x_1 - x_2(t))^2 + (y_1 - y_2(t))^2}}, \tag{6.13}$$

where the subscripts 1 and 2 denote the two point vortices. From the expression for the horizontal velocity, (6.1), or by interpreting (6.10), we find that

$$U_1 = \frac{\kappa_2}{4\pi} \frac{y_1 - y_2}{((x_1 - x_2(t))^2 + (y_1 - y_2(t))^2)^{3/2}}. \tag{6.14}$$

The way we have initialised the dipole means that it will propagate as a unit along the y-axis with both point vortices moving at the same speed. Therefore $y_1 = y_2$ for all time and so $U_1 = 0$ as expected. A similar procedure gives that $U_2 = 0$ which means that the separation of the two point vortices will remain constant. An analogous method can be used to obtain V resulting in

$$V_1 = -\frac{\kappa_2}{4\pi} \frac{x_1 - x_2}{((x_1 - x_2(t))^2 + (y_1 - y_2(t))^2)^{3/2}}, \tag{6.15}$$

which can be simplified using $x_1 - x_2 = -d$ along with the identities used above, to determine that

$$V_1 = \frac{1}{4\pi} \frac{\kappa_2}{d^2}. \tag{6.16}$$

Since we have already established that $y_1 = y_2$ it is clear that $V_1 = V_2$ and so this is the speed at which the dipole propagates.

6.3.1 Wavenumber and amplitude equations

We will launch a wavepacket upstream from the point vortex dipole and track the evolution of the wavepacket's wavenumbers and amplitudes. We expect to see some

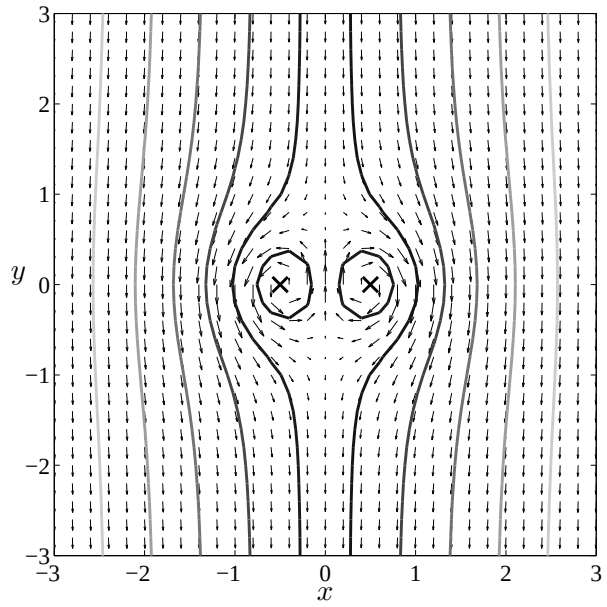

Figure 6.1: The streamlines and velocity field of a point vortex induced dipole. The point vortices are marked by crosses.

growth of the wavepacket's amplitudes caused by the transient activity of the dipole sweeping past the wavepacket. To see this we will first need to find the equations of these amplitudes.

In this scenario the explicit expressions for the wavenumbers' equations, of the form derived in (4.26), follow very easily from (6.10) and (6.11) giving that

$$\dot{k} = -\frac{3\kappa}{4\pi}\frac{(x-X(t))(y-Y(t))}{r^5}k - \frac{\kappa}{4\pi r^3}\left(1-\frac{3(x-X(t))^2}{r^2}\right)l \tag{6.17}$$

$$\dot{l} = \frac{\kappa}{4\pi r^3}\left(1-\frac{3(y-Y(t))^2}{r^2}\right)k + \frac{3\kappa}{4\pi}\frac{(x-X(t))(y-Y(t))}{r^5}l \tag{6.18}$$

$$\dot{m} = -\frac{3\kappa}{4\pi}\frac{(y-Y(t))(z-Z(t))}{r^5}k + \frac{3\kappa}{4\pi}\frac{(x-X(t))(z-Z(t))}{r^5}l. \tag{6.19}$$

To complete the set of equations required for this system we need to include the amplitude equations for δ, ζ and q. The form of these can be found in (4.56)-(4.58). Since there is a possibility that m could pass through 0, the equations that were derived in section 4.4, (4.70)-(4.72), will be used in the code for this flow.

6.3.2 Non-dimensionalising

In the same way that the amplitude equations were non-dimensionalised in section 4.3 we have to non-dimensionalise the position and wavenumber equations. As before, this involves using f, the Coriolis parameter, as an inverse time scale. Here we also use d, the dipole separation, as a length scale and take the reference velocity, U_{ref}, to be the dipole translation speed. This gives that the Rossby number, ϵ, is given by

$$\epsilon = \frac{U_{ref}}{fL} = \frac{1}{4\pi}\frac{\kappa}{fd^3}. \tag{6.20}$$

By using these identities in (6.17)-(6.19) we find that the relationship between the dimensional and nondimensional (starred) versions of the wavevector equations is

$$\frac{d\mathbf{k}^*}{dt^*} = \frac{1}{\epsilon}\frac{d\mathbf{k}}{dt}. \tag{6.21}$$

The explicit forms of the nondimensional expressions for the positions of the point vortices and their associated nondimensional wavevector equations can all now be easily derived from (6.10)-(6.12) and (6.17)-(6.19).

6.4 Polarisation

6.4.1 Eigensolution

Since our aim is to show that energy is transferred from the vortical mode of the system to the gravity wave modes, it is clear that we need to find a way of decomposing the equations into the three modes so that we can track them. It is worth noting that doing this will only make strict sense when the flow around the wavepacket is completely uniform. This means that the decoupling of the three modes will only make sense when the wavepacket is far enough away from the dipole for it not to feel any of its effects.

To achieve this decoupling we begin by writing the system as

$$\begin{aligned} D_t\delta &= A\delta + B\zeta + Cq, \\ D_t\zeta &= -\alpha\delta, \\ D_tq &= 0, \end{aligned} \tag{6.22}$$

where the expressions for A, B, C and α are clearly identified in (4.64)-(4.66). When the flow is uniform, A, B, C and α are time independent, and so a solution of this system can be written as

$$\begin{pmatrix} \delta \\ \zeta \\ q \end{pmatrix} = e^{i\omega t} \begin{pmatrix} \hat{\delta} \\ \hat{\zeta} \\ \hat{q} \end{pmatrix}. \tag{6.23}$$

Introducing (6.22) into (6.23) gives that

$$M \begin{pmatrix} \hat{\delta} \\ \hat{\zeta} \\ \hat{q} \end{pmatrix} = \begin{pmatrix} A & B & C \\ -\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\delta} \\ \hat{\zeta} \\ \hat{q} \end{pmatrix} = i\omega \begin{pmatrix} \hat{\delta} \\ \hat{\zeta} \\ \hat{q} \end{pmatrix}, \tag{6.24}$$

which is of the form of an eigenvalue problem. Solving the characteristic equation of the matrix gives that its eigenvalues are

$$\omega = 0 \qquad \text{and} \qquad \omega = -i/2(A \pm \sqrt{A^2 - 4\alpha B}). \tag{6.25}$$

These eigenvalues correspond to the vortical mode and two gravity wave modes of the system respectively. It is interesting to note here that the form of these eigenvalues looks similar to the form of the dispersion relation found in (3.13).

Now that we have the system in this form we can write its solution as the sum of the three modes by expressing it as

$$\begin{pmatrix} \delta \\ \zeta \\ q \end{pmatrix} = A_{\mathrm{v}}\mathbf{e}_{\mathrm{v}} + A_{\mathrm{g}^+}\mathbf{e}_{\mathrm{g}^+} + A_{\mathrm{g}^-}\mathbf{e}_{\mathrm{g}^-}, \tag{6.26}$$

where $\mathbf{e}_{\mathrm{v}}$, $\mathbf{e}_{\mathrm{g}^+}$ and $\mathbf{e}_{\mathrm{g}^-}$ are the right eigenvectors of M; A_{v}, A_{g^+} and A_{g^-} are functions of time and v, g^+ and g^- refer to the vortical mode and the plus and minus gravity wave modes respectively. For a uniform flow this is an exact solution with

$$A_{\mathrm{v}} = A_{\mathrm{v}}(0) \quad \text{and} \quad A_{\mathrm{g}^\pm} = A_{\mathrm{g}^\pm}(0)e^{-i\omega_{\mathrm{g}^\pm}t}. \tag{6.27}$$

For non-uniform flows we can write (6.26) with time dependent A_{v} and $A_{\mathrm{g}^\pm}$.

6.4.2 Finding A_{v} and $A_{\mathrm{g}\pm}$

We need to find expressions for A_{v} and $A_{\mathrm{g}\pm}$ to be able to track the transfer of energy between the modes of the system. This will enable us to project a given (δ, ζ, q) on to the right eigenvectors, as in (6.26). If we let

$$\mathbf{x}(t) = \begin{pmatrix} \delta(t) \\ \zeta(t) \\ q(t) \end{pmatrix}, \tag{6.28}$$

then (6.26) reads as

$$\mathbf{x}(t) = \sum_{i=1}^{3} A_i \mathbf{e}_i. \tag{6.29}$$

Since M is not self-adjoint the derivation of this projection will require the use of the left eigenvectors, which satisfy $\mathbf{e}_j^+ M = \mathbf{e}_j^+ \lambda_j$, as well as the more standard right eigenvectors, which satisfy $M\mathbf{e}_i = \lambda_i \mathbf{e}_i$.

Multiplying the definition for the right eigenvectors by $\mathbf{e}_j^+ M$ on the right gives that

$$M\mathbf{e}_i \cdot \mathbf{e}_j^+ M = \lambda_i \mathbf{e}_i \cdot \mathbf{e}_j^+ M. \tag{6.30}$$

Following in a similar manner and multiplying the definition of the left eigenvectors on the left by $M\mathbf{e}_i$ gives that

$$M\mathbf{e}_i \cdot \mathbf{e}_j^+ M = M\mathbf{e}_i \cdot \mathbf{e}_j^+ \lambda_j. \tag{6.31}$$

By taking (6.30) away from (6.31) we are left with

$$0 = (\lambda_i - \lambda_j)\mathbf{e}_i \cdot \mathbf{e}_j^+, \tag{6.32}$$

which leads to the conclusion that $\mathbf{e}_i \cdot \mathbf{e}_j^+ = 0$ unless $i = j$.

Going back to (6.26) and multiplying it on the left by the sum of the left eigenvectors gives that

$$\mathbf{e}_j^+ \cdot \mathbf{x}(t) = \sum_{i=1}^{3} A_i \, \mathbf{e}_j^+ \cdot \mathbf{e}_i. \tag{6.33}$$

Using the above eigenvector result in this expression gives that for the right hand side

mode	eigenvalue	left eigenvector	right eigenvector
vortical	0	$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -\frac{C}{B} \\ 1 \end{pmatrix}$
gravity$^+$	$\frac{A+\sqrt{A^2-4\alpha B}}{2}$	$\begin{pmatrix} \frac{A+\sqrt{A^2-4\alpha B}}{2B} \\ 1 \\ \frac{C}{B} \end{pmatrix}$	$\begin{pmatrix} -\frac{A+\sqrt{A^2-4\alpha B}}{2\alpha} \\ 1 \\ 0 \end{pmatrix}$
gravity$^-$	$\frac{A-\sqrt{A^2-4\alpha B}}{2}$	$\begin{pmatrix} \frac{A-\sqrt{A^2-4\alpha B}}{2B} \\ 1 \\ \frac{C}{B} \end{pmatrix}$	$\begin{pmatrix} -\frac{A-\sqrt{A^2-4\alpha B}}{2\alpha} \\ 1 \\ 0 \end{pmatrix}$

Table 6.1: The eigenvalues and the left and right eigenvectors that correspond to the three modes of the dipole system.

to be non-zero we need to let $i = j$ and so we find

$$\mathbf{e}_i^+ \cdot \mathbf{x}(t) = A_i \, \mathbf{e}_i^+ \cdot \mathbf{e}_i, \tag{6.34}$$

which can be rearranged to give

$$A_i = \frac{\mathbf{e}_i^+ \cdot \mathbf{x}(t)}{\mathbf{e}_i^+ \cdot \mathbf{e}_i}. \tag{6.35}$$

It is clear from this expression that if the flow is uniform then the A_i's are constant, as was expected.

After calculating the left and right eigenvectors, which can be found in table 6.1,

we find that

$$A_v = q(t), \tag{6.36}$$

$$A_{g^+} = \frac{1}{\Lambda_-}\left(\frac{A+\sqrt{A^2-4\alpha B}}{2B}\delta(t)+\zeta(t)+\frac{C}{B}q(t)\right), \tag{6.37}$$

$$A_{g^-} = \frac{1}{\Lambda_+}\left(\frac{A-\sqrt{A^2-4\alpha B}}{2B}\delta(t)+\zeta(t)+\frac{C}{B}q(t)\right), \tag{6.38}$$

where

$$\Lambda_\pm = 1-\frac{(A\pm\sqrt{A^2-4\alpha B})^2}{4\alpha B}. \tag{6.39}$$

Tracking these constants throughout the evolution of the system gives an appropriate measure of any spontaneous generation that may occur.

6.5 Initialisation

The system was initially set up as described in section 6.3.1, which results in the dipole propagation speed being given by (6.16). This expression allows us to determine the correct values to use for the strength of the point vortices, κ, from our desired values of d and the dipole propagation speed V. In this study we have chosen the speed of the dipole to be 10 ms^{-1} and the point vortex separation, d, to be 250 km. These values were chosen so that the model would be quite typical of what could be found in the earth's atmosphere. In the non-dimensional case this corresponds to setting $d = 1$ and $\kappa = \pm 1$.

The system was studied in a reference frame that follows that dipole. This is achieved by subtracting the speed of the dipole from the velocity in the y-direction of the two point vortices and the wavepacket, where the wavepacket is modelled as a point vortex with zero strength. This can be thought of as the dipole being arrested by a uniform flow. The other relevant parameters of the flow are chosen to be $f = 10^{-4}$ s^{-1} and $N = 10^{-2}$ s^{-1}, giving that $S = 100$.

The wavepacket was placed with a wavevector that satisfied $m/\sqrt{k^2+l^2} \approx 10$. This may seem different to elsewhere in this thesis but because of the scaling of z, (6.4), that has been employed it is in fact consistent.

Since $D_t q = 0$, we can choose that $q = 1$ without loss of generality. This is because

the whole system of equations is linear, therefore if $q = 2$ the system is scaled by a factor of 2, which means there is no loss of generality using $q = 1$.

Picking the initial values of δ and ζ requires a bit more thought. To see the exchange of energy from the vortical mode to the gravity wave modes, we want to set up the system so that initially only the vortical mode is excited. We achieve this by placing the wavepacket far enough in front of the dipole so that it is unaffected by the dipole's presence. This means that all the gradients of the velocity field around the wavepacket can initially be ignored because they are almost zero. We then set the values of δ and ζ to be such that their material derivatives are zero. This amounts to choosing $A_{g^\pm} = 0$ at $t = 0$ which ensures that initially only the vortical mode is excited. We expect that this will be the case until the wavevector undergoes a transient change which will happen as the wavepacket sweeps past the dipole.

From the values of the eigenvectors, the expressions given in (6.22) and the form of (6.26), the only way to ensure that $D_t\zeta = 0$ is to let $\delta(0) = 0$. Now using the fact that initially all the gradients of the flow fields are zero, which means that $\alpha = 0$, and the fact that we have set $\delta = 0$, gives us that

$$D_t\delta = \left(\frac{m^2}{\kappa^2} + \frac{k^2 + l^2}{\kappa^2}S^2\right)\zeta - \frac{k^2 + l^2}{\kappa^2}q. \tag{6.40}$$

Hence to get $D_t\delta(0) = 0$ we need to set

$$\zeta(0) = \frac{(k^2 + l^2)q}{m^2 + S^2(k^2 + l^2)}, \tag{6.41}$$

where all the quantities involved are taken as their initial values.

6.6 Results

Figure 6.2 shows the trajectories in the (x, y) plane corresponding to three wavepackets located at distances $D = 50$, 100 and 150 km from the dipole axis. The wavepackets trace out a straight path towards the dipole and as they get closer to the dipole they start to react to it, sweep past it and rejoin their original trajectories. Intuitively, one might expect this transient behaviour to lead to gravity-wave generation.

Since initially the wavepacket does not feel the effects of the dipole, and so the gradients of the velocity fields are almost zero, we expect the wavenumbers of the wavepacket to remain unchanged, which is the case. As the wavepacket passes the

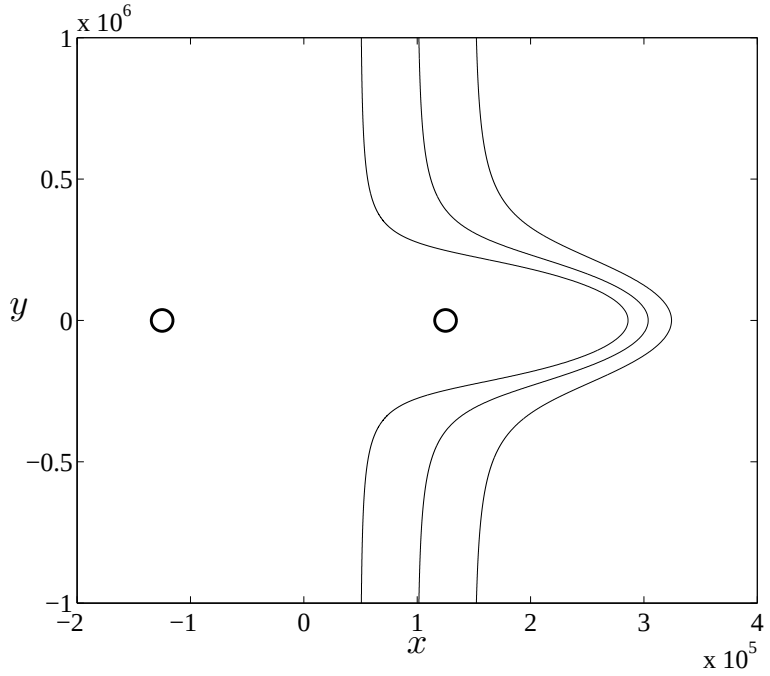

Figure 6.2: Trajectories of three wavepackets in a flow generated by a quasi-geostrophic dipole in three dimensions. The locations of the two point vortices are indicated by circles and are placed at $x = \pm 125$ km. The wavepackets are placed in the plane $z = 0$ along with the dipole and are characterised by their distance $D = 50$, 100 and 150 km from the axis of the dipole as $t \to \pm\infty$

dipole there is a transient change in the wavevector, as seen in figure 6.3, then, as the velocity field settles down, so do the wavenumbers. This can be understood by using the result found in (4.30), which shows that $\mathbf{k} \cdot \mathbf{U}$ is conserved along the streamlines. However, it is worth stressing that they may have changed as a result of the encounter, which is clear from the fact that since U is zero, apart from during the transient motion, $\mathbf{k} \cdot \mathbf{U} \approx lV$. Since $V(-\infty) = V(\infty)$, it is clear that we also need to have $l(-\infty) = l(\infty)$. This not the case for k since $U \approx 0$ and so k can change without affecting the value of $\mathbf{k} \cdot \mathbf{U}$.

The transient growth of the wavevector as it passes the dipole is passed on to the amplitudes of the wavepacket and hence on to the three modes that they are decomposed into. Figure 6.4 shows how the A_{g^+} mode evolves. This figure clearly demonstrates that the gravity mode is initially not excited and that fast oscillations appear following the transient behaviour associated with the encounter with the dipole. This is spontaneous generation of inertia-gravity-waves caused by the transient change in the velocity field surrounding the wavepacket.

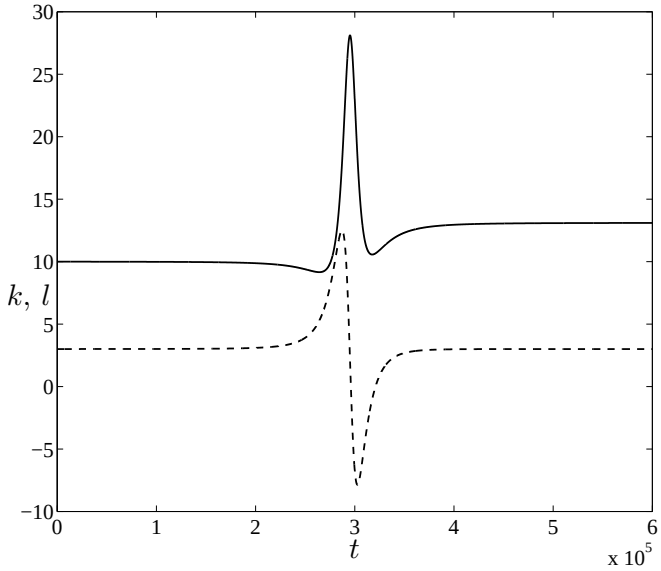

Figure 6.3: The evolution of the horizontal wavevectors, k (solid line) and l (dashed line), as the wavepacket sweeps past the dipole.

The strong dependence of the inertia-gravity wave amplitude on D can also clearly be seen here. This can be thought of as a proxy for the inverse Rossby number since $\epsilon = U/fD$. The dependence is demonstrated in figure 6.5 which shows the magnitude of the amplitude A_{g^+}, characterising the amplitude of the fast oscillations as $t \to \infty$, against D. The figure is plotted in linear-logarithmic coordinates so the fact that the points lie on a straight line shows the exponential dependence that the amplitude of the inertia-gravity waves has on the distance from the dipole axis. This exponential dependence is expected to be of the form $e^{-\alpha/\epsilon+\beta}$ and this is verified and plotted on the figure with $\alpha = 3.3$ and $\beta = 1$. It is worth noting that an exponential-asymptotic analysis similar to that in [65] could be carried out to obtain an explicit approximation for α and β.

6.7 Elliptical trajectories within a dipole

From figure 6.1 it is clear that in the case of the point vortex dipole closed elliptical stream lines exist in the near vicinity of each point vortex. By placing the wavepacket on one of these streamlines we can get a feel for how the wavepacket will evolve in an elliptical flow and get an insight into the possibility of growth in this situation.

The initial position of the wavepacket in the area of closed streamlines around one

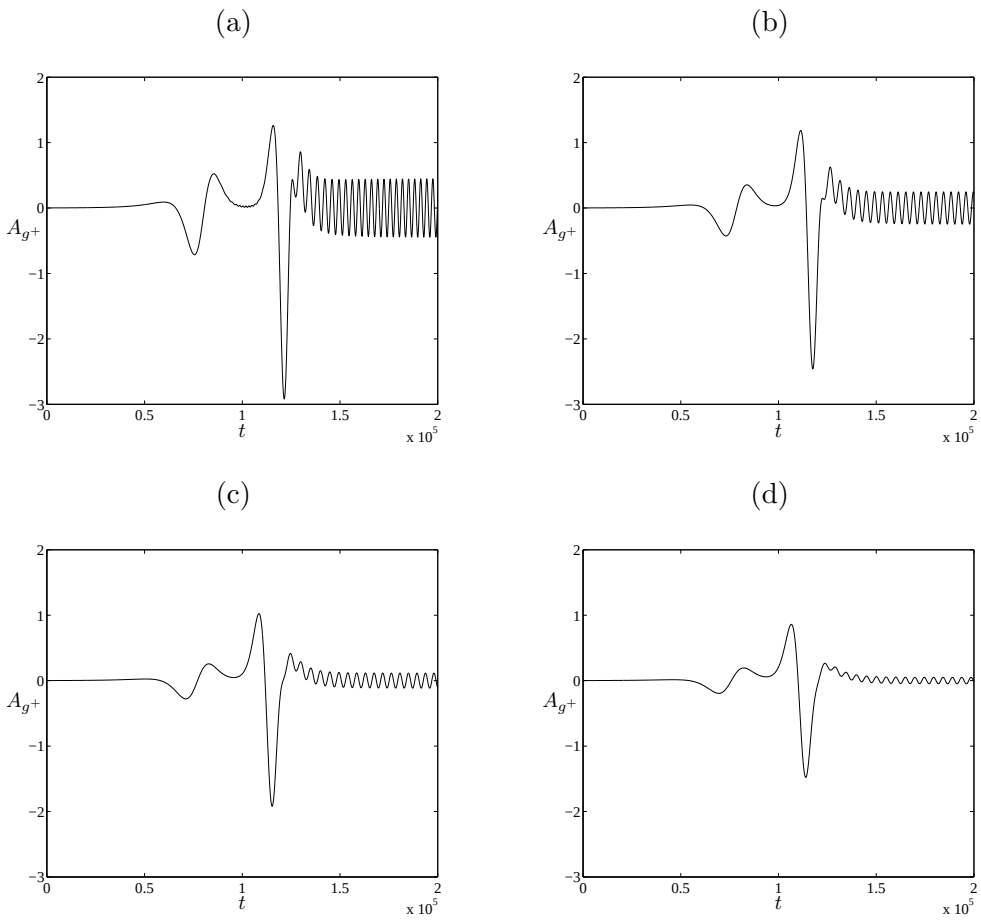

Figure 6.4: Inertia-gravity waves generated spontaneously as vortical-mode wavepackets sweep past a dipole. This shows the amplitude of A_{g^+} as a function of time for wavepackets located at distances, D (a) 50, (b) 75, (c) 100 and (d) 125 km from the dipole axis as $t \to \pm\infty$.

of the point vortices in the dipole affects the aspect ratio of the elliptical trajectory that the wavepacket will follow. The closer to the point vortex that the wavepacket starts, the more circular the trajectory will be and conversely the further away the wavepacket starts, the more elliptical the trajectory will be. It is also worth noting that along with the trajectory becoming more elliptical, the further away the wavepacket starts from the point vortex, the more deformed the ellipse of its trajectory will become. This is because the trajectory then passes closer to the dipole axis along which the streamlines are straight, so the elliptical trajectory gets slightly flattened along this side of the ellipse. This leaves us with a trade off between finding a trajectory that is not too circular, which is not desirable since it is known that circular flows are inherently stable, and finding a trajectory that is not too deformed which causes the periodic nature of

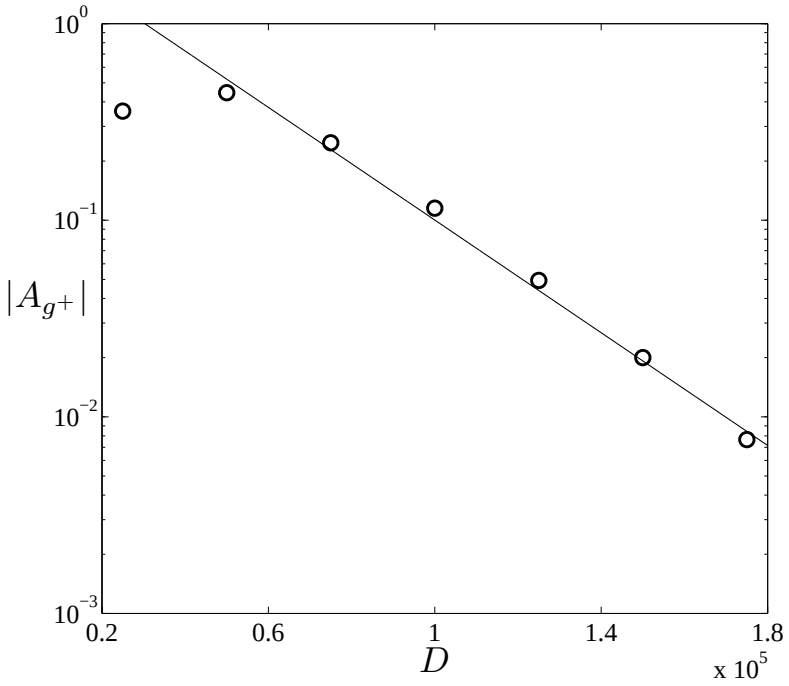

Figure 6.5: The amplitudes (circles) at $t \to \infty$ of the inertia-gravity wave mode generated spontaneously as a vortical-mode wavepacket is swept past a dipole as a function of the distance of the wavepacket to the dipole axis. A best fit line through the points shows the exponential dependence of these amplitudes on D.

the wavepacket to get complicated. Figure 6.6 shows a trajectory that appears to yield some growth. This trajectory was formed by initially placing a wavepacket at $(95, 120)$ km in a point vortex dipole, set up in the same way as the earlier part of this chapter, which has a propagation speed of 10 ms^{-1} and a Rossby number of $\epsilon = 2/3$. From (6.20) this means that we must set the point vortex separation to be $d = 210860$ m.

The motion of the wavepacket around the point vortex is periodic, so we may expect that the horizontal wavenumbers will behave in a periodic way. There is still no vertical motion, therefore we can expect the vertical wavenumber to remain constant. The evolution of the horizontal wavenumbers for the trajectory shown in figure 6.6 is shown in figure 6.7. As expected it is periodic but maybe not quite as sinusoidal as expected which is the result of the slight deformation of the elliptical trajectory.

The wavepacket's amplitudes are initially taken as $\delta(0) = 1$, $\zeta(0) = 0$ and since we are looking for instabilities caused by an unbalance we take $q(0) = 0$. Figure 6.8 shows the evolution of δ and ζ for the same trajectory as shown in figure 6.6. This figure clearly shows growth in these amplitudes and so there must be some sort of instability

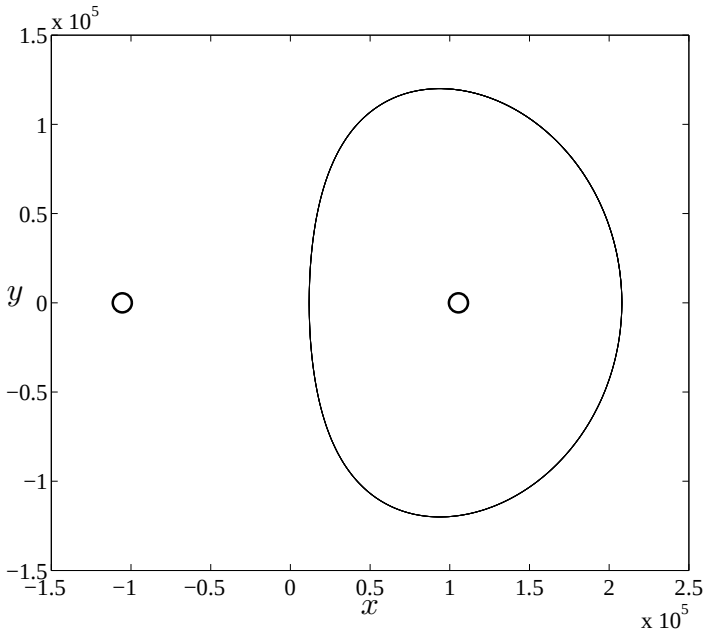

Figure 6.6: The elliptical trajectory of a wavepacket in the flow generated by a quasi-geostrophic dipole in three dimensions. The locations of the two point vortices are indicated by circles and are placed at $x = \pm 105430$ m. The wavepacket is in the plane $z = 0$ along with the dipole and is initially placed at $(95, 120)$ km.

in this elliptical flow.

This result gives motivation for the study of purely elliptical flows which will be carried out in the next chapter.

6.8 Complex time dependent flows

6.8.1 Introduction

A logical extension to the work that has already been carried out in the rest of this chapter is to use the point vortices, that have so far only been used to create a dipole, to create a complex time dependent flow. To do this we put multiple point vortices in the domain with the wavepacket and use the theory that we have already developed to evolve the system and track the wavepacket's properties. This is a flow that we discussed in less detail in [6]. Intuitively, in this system the wavepacket will be jostled about in a seemingly random manner by the changing strains and shears that the other point vortices assert on it. This gives the feel that the system behaves very much like a random-strain model and so we can draw on the results that are known for random-strain models to gain insight into what to expect in our simulations. As

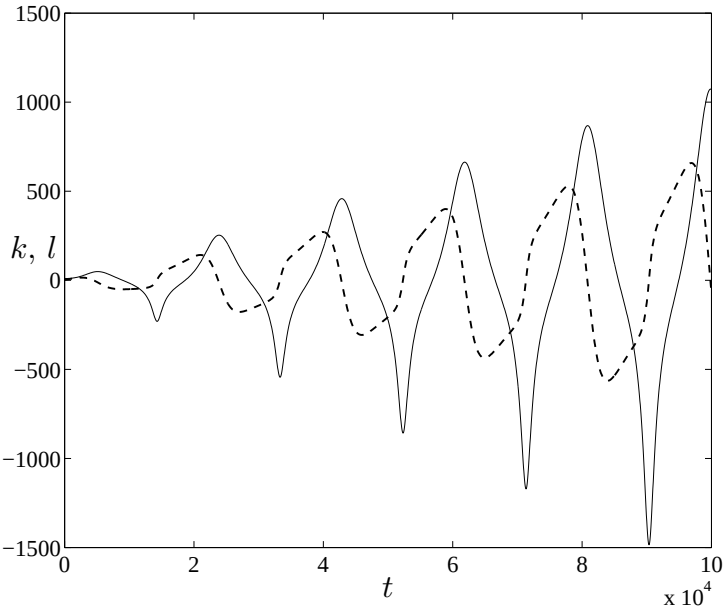

Figure 6.7: The evolution of the horizontal wavevectors, k (solid line) and l (dashed line), as the wavepacket follows the elliptical trajectory shown in figure 6.6 around the point vortex.

is well known from the study of particle advection, the trajectories of particles and hence wavepackets are typically chaotic when the velocity field is time dependent. The Lagrangian time dependence of the strain ∇U that appears on the righthand side of (4.26) is therefore very complicated; it is natural to model it by a stationary random process. This is the key idea of the random-strain models proposed in [29] in the context of passive-scalar advection (the scalar-concentration gradient obeys (4.26)).

Random-strain models have also been adapted to the layer-wise two-dimensional nature of geophysical flows with the conclusion that while typically $\kappa \to \infty$, the aspect ratio $m/\sqrt{k^2+l^2}$ reaches a stationary distribution, [24]. Since this ratio together with ∇U determines the coefficients of the wavepacket's amplitude equations (4.56)-(4.58) for random-strain models, these equations are essentially those of a linear oscillator with stationary random coefficients. This observation makes it possible to draw some conclusions about the behaviour of δ and ζ. First, these quantities typically grow exponentially, with a deterministic growth rate defined by $\lim_{t\to\infty} t^{-1}\log|\zeta|$, say, that can be recognised as the Lyapunov exponent of the system. Second, in the limit of small Rossby number, naturally defined using the correlation time of the random process determining the oscillator frequency, the growth rate can be expected to depend

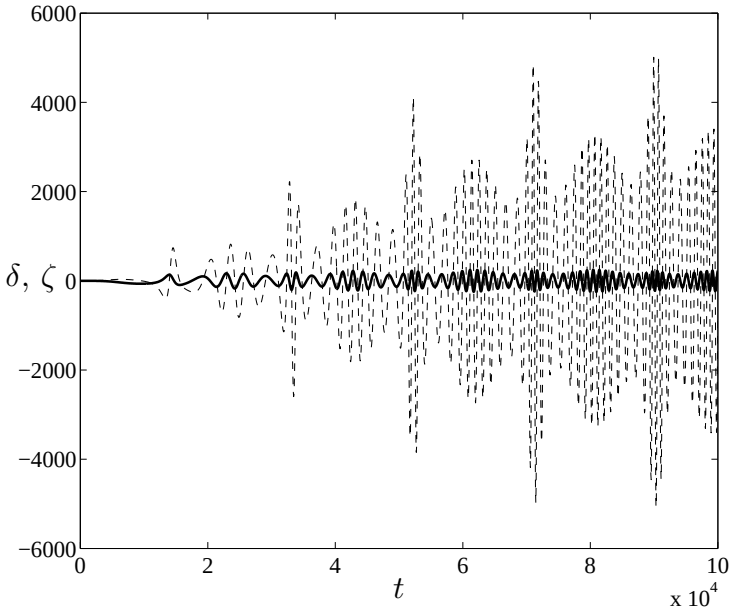

Figure 6.8: The evolution of δ (dashed line) and ζ (solid line), as the wavepacket follows the elliptical trajectory shown in figure 6.6 around the point vortex.

crucially on the smoothness of this process. Specifically, the explicit results available for closely related problems [4, 14], suggest that the growth rate is proportional to the power spectrum of the random process evaluated at twice the average oscillator frequency. This means that the growth of inertia-gravity waves will depend on how much energy the frequency of the oscillator has at the frequency of the gravity waves. When the Rossby number is small this average frequency lies at the tail end of the power spectrum of the oscillator frequency. Since the power spectrum's tail gets smaller as the smoothness of the process increases, there will be less power at the gravity wave frequency and so there will be less growth. Hence the growth of the wavepacket's amplitudes is entirely controlled by the smoothness of the process. In particular, if the process is real-analytic, the growth rate will be exponentially small in the Rossby number. Thus spontaneous inertia-gravity wave generation is predicted by random-strain models to have a similar Rossby number dependence in complex flows as in simple steady flows. The key assumption, which may not always be satisfied, is that the Lagrangian time series of ∇U is real-analytic. Note that a simple model for which analytic progress is possible would take ∇U to be a white noise so as to apply the techniques developed for the Kazantsev-Kraichnan models of a kinematic dynamo and passive-scalar advection, see [13, 44] and references therein. This would, however,

be appropriate only for flows with correlation times short compared to f^{-1} and to the inverse strain.

6.8.2 Initialisation

Throughout the study of this system we will use all the theory for point vortices that we derived previously in this chapter. In a slight contrast to the work done above though, here we will use a dimensional formulation, using in particular an unscaled variable, z, rather than the scaled version defined in (6.4). This introduces a factor of N^2/f^2 into the equation linking the potential vorticity to the stream function, (6.5), which has a knock-on effect of adding this factor into the equations of all the gradients of the velocity field. This has been done due to the fact that in this system we will be placing point vortices at different heights, and so to avoid confusion, and make the numbers more intuitive, it is easier if we do not scale the vertical coordinate.

Since we have dispensed with the vertical scaling, for our analysis of this system we can go back to using initial values of the wavenumbers that are in the more familiar ratio of $m/\sqrt{k^2+l^2} \approx N/f \approx 100$.

In contrast to the point vortex dipole where we looked for spontaneous growth and hence used $q \neq 0$, our aim of looking at this system is to find out if there are any unbalanced instabilities that lead to a growth of the wavepacket's amplitudes and so we need to set $q = 0$. Since this is the case we choose the wavepacket's initial conditions so that only one of its gravity wave modes is excited. It is worth noting here that this decomposition is an approximation since it is only strictly valid for uniform flows, therefore we can expect that the other gravity wave mode will be excited almost immediately.

We initially excite only one gravity wave mode by picking the initial values of δ, ζ and q so that (δ, ζ, q) is proportional to the right eigenvector corresponding to the gravity wave mode, as found in table 6.1. Without loss of generality we will use the gravity$^+$ mode throughout this study. From the table of eigenvectors it is clear that to achieve this proportionality we first need to set $q(0) = 0$, as already decided above, and then pick a value for $\zeta(0)$, which in this study we will take to be 10^{-5} s^{-1} which is a typical atmospheric value. To then make sure that just this mode is initially active, from the form of the right eigenvector, we need to pick

$$\delta(0) = \zeta(0)\frac{A + \sqrt{A^2 - 4\alpha B}}{2\alpha}, \tag{6.42}$$

where A, B and α are defined in (6.22).

The only parameters left to define now are the flow parameters, the initial positions and strengths of the point vortices and the wavepacket's initial position. The aim of this system is to make it as realistically accurate as possible to that found in the earth's atmosphere, to an order of magnitude. To that end we again take the now familiar values of $f = 10^{-4}$ s^{-1} and $N = 10^{-2}$ s^{-1}. The wavepacket was placed at a height of 10 km, which is around the tropopause, and for simplicity the horizontal coordinates were chosen so that the wavepacket's horizontal position was $(0, 0)$. In each simulation that was run the point vortices were all placed within a 150 km radius horizontally and within 1 km vertically of the wavepacket. The strength of the vortices were chosen so that the sum of all the strengths was zero and so that the speed of the wavepacket was always around 10 ms^{-1}. This was achieved by using a trial and error method and adjusting the strengths of the point vortices as necessary.

6.8.3 Results

Now that we have established the theory and set the system up, we are in the position to run some simulations of complex systems and see if we can gain any meaningful results by analysing their output. All the systems analysed below contain eleven point vortices and a wavepacket set up as described above.

The trajectories of the wavepacket and the point vortices for the first system to be analysed can be seen in panel (a) of figure 6.9. Panel (b) shows the wavepacket's wavenumbers, panel (c) shows the evolution of the gravity wave modes along with the wavepacket's amplitudes and panel (d) shows the 2-norm of the wavepacket's strain matrix over its frequency.

One of the first things that stands out when looking at the figures produced by this simulation is that, as expected, although the simulation was started with just the gravity$^+$ mode activated, the gravity$^-$ mode also becomes active. Another initial observation is that there is a sharp change to the wavenumbers at around $t = 1.5 \times 10^5$ s. This sharp change is seen to also occur in the gravity wave modes and the wavepacket's amplitudes which shows, as expected from the equations, that these quantities are linked. It is worth noting however that this link appears to be quite powerful as the change in the wavepacket's amplitudes at this point is substantial. To try to get a better handle on why this sudden change occurs at this point, figure 6.10 shows the

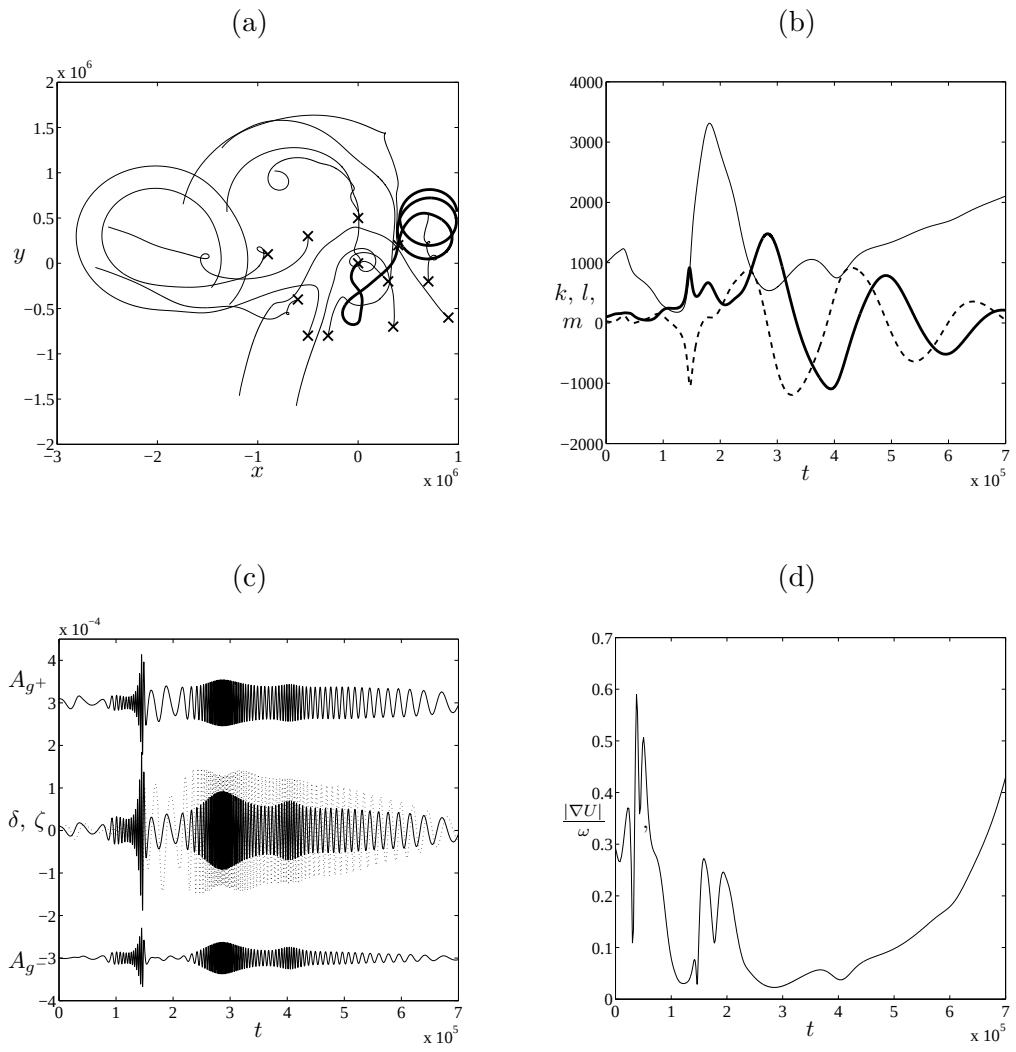

Figure 6.9: (a) The evolution of eleven point vortices, whose strengths sum to zero, placed within 150 km horizontally and 1 km vertically of a wavepacket initially placed at $(0, 0, 10)$ km. The initial positions of the point vortices are denoted by $\times$'s, the initial position of the wavepacket by a $+$ and the trajectory of the wavepacket by the heavy line. (b) The evolution of the vertical wavenumber m, shown as the thin line, and ten times the horizontal wavenumbers k, the thick line, and l, the dashed line. (c) The evolution of the two gravity wave modes and the amplitudes δ and ζ where the plus and minus gravity waves have been displaced by $\pm 3 \times 10^{-4}$ respectively. ζ is the solid line in the centre and δ, which has been multiplied by 5×10^{-4}, is the dotted line. (d) The norm of the wavepacket's strain matrix divided by its frequency.

evolution of the point vortices and wavepacket up to and including $t = 1.45 \times 10^5$ s. As can be seen in this figure at this time the wavepacket is very close to a point vortex just below and to the right of it. The wavepacket also has two point vortices, that are a bit further away, above and to the left of it. It is this close encounter with the

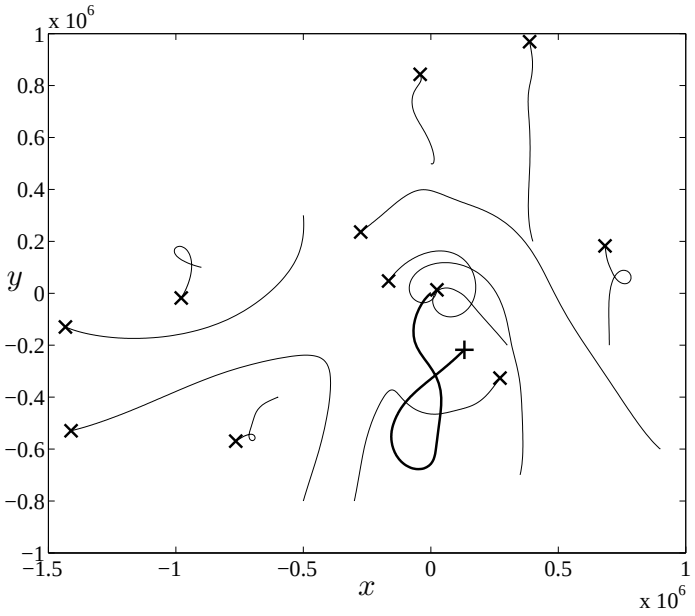

Figure 6.10: The evolution of the point vortices and the wavepacket, shown by the heavy line, up to $t = 1.45 \times 10^5$. The positions of the point vortices at this time are marked by $\times$'s and the wavepacket by a $+$.

point vortex that is responsible for the dramatic change in the wavenumber that we have observed. It is interesting to note that the point vortex that the wavepacket has a close encounter with is at a height of 9.6 km, i.e. 400 m below the wavepacket. This explains why the vertical wavenumber, m, of the wavepacket increases at this point; it is because it is being pushed from below, which condenses the wave crests giving rise to an increase in wavenumber. The direction of the change of the horizontal wavenumbers can also be anticipated from this figure. The strength of this closest point vortex is -6×10^{12} which means it is rotating in a clockwise fashion. From looking at the figure we can visualise that it will stretch out l and compress k giving rise to l falling and k increasing as is seen. While here we have just analysed the first big change in the wavenumbers and the amplitudes, it is clear from looking at the figures that all the big changes in the amplitudes coincide with big changes in the wavenumbers.

As well as looking at the wavenumbers for an indicator of large growth in the amplitudes, we can also look at a local Rossby number that is calculated as the norm

of the wavepacket's strain matrix, given by

$$\begin{pmatrix} \partial_x U & \partial_x V \\ \partial_y U & \partial_y V \end{pmatrix}, \tag{6.43}$$

over the wavepacket's frequency as derived in (6.25). This is a natural quantity to consider since we can think of the norm of the strain as a rate of change of the frequency, since the strain affects the rate of change of the wavenumbers which in turn affects the frequency. This means that this quantity is in essence a ratio of the rate of change of the frequency to the frequency. Hence, when this quantity is large we expect that the wavenumbers are rapidly changing and hence that there is rapid growth in the wavepacket's amplitudes.

By comparing panel (d) with panel (c) of figure 6.9 we can see that the changes to the gravity wave modes, and hence δ and ζ, occur when this local Rossby number is large. It is clear that this local Rossby number rapidly rises for the large dramatic changes in the amplitudes at the time discussed extensively above, before dropping and then slowly increasing again for the last third of the simulation which is where the amplitudes are dying away. Although this seems to be a good indicator for the majority of the flow there are questions that can be asked about its use at the beginning of the simulation. Here the largest values of the local Rossby number are found and while there is some small growth, from initially starting out very small to becoming steady waves, there is no large growth and so using the local Rossby number as a diagnostic for large growth seems to fail here. This could be down to the fact that the frequency of the wavepacket relies on the angle of the wavevector and not just on the values of the wavevectors alone. This means that even though there might be a large strain, and hence a large local Rossby number, the angle of the wavevector may not change so there may not be large growth. In this case using the local Rossby number as a diagnostic fails. Overall the local Rossby number is a good indicator of growth in the wavepacket's amplitudes, except when the angle of the wavenumber does not change despite the presence of a large strain. It is perhaps made more reliable when used alongside the wavenumbers.

The observations that we have made for this simulation are interesting but will only be useful if they also hold for other simulations. Figure 6.11 shows a similar simulation

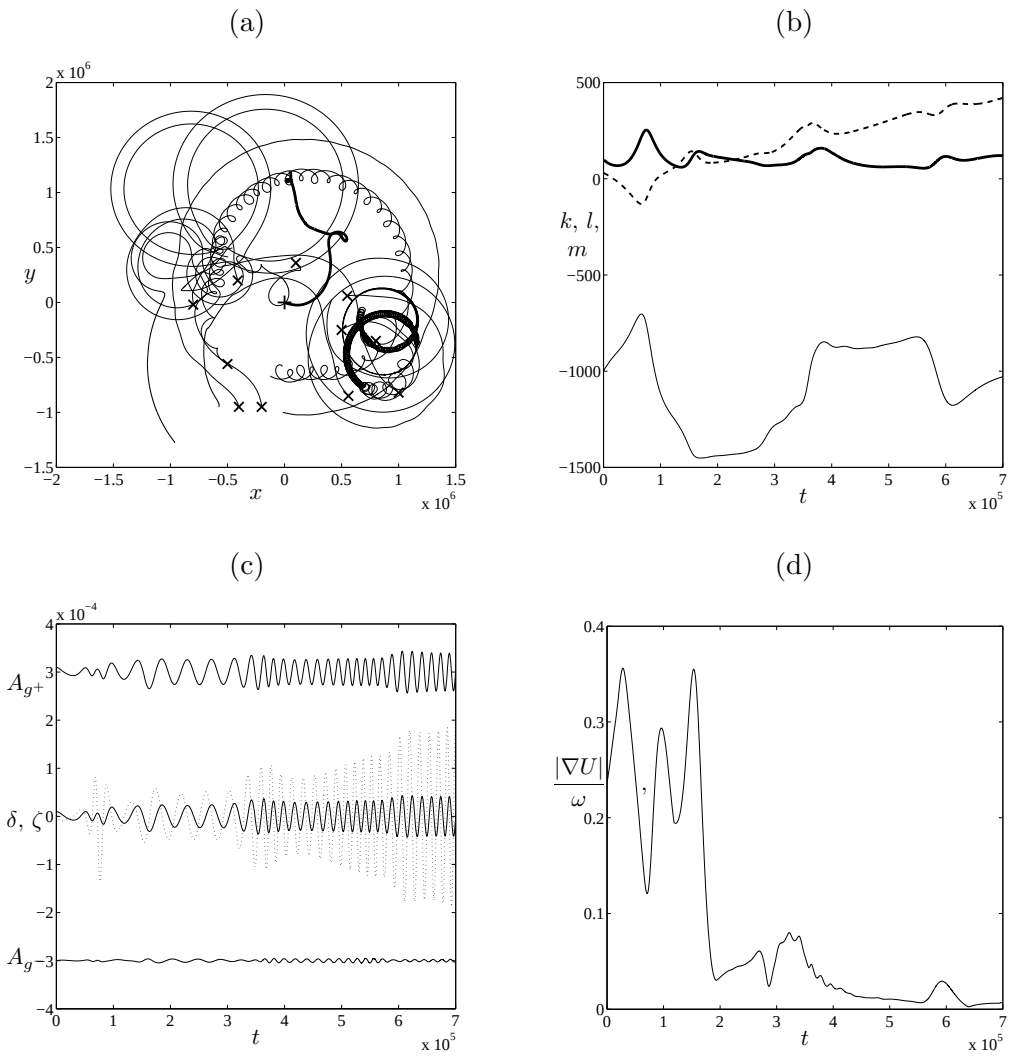

Figure 6.11: (a) The evolution of eleven point vortices, whose strengths sum to zero, placed within 150 km horizontally and 1 km vertically of a wavepacket initially placed at $(0, 0, 10)$ km. The initial positions of the point vortices are denoted by $\times$'s, the initial position of the wavepacket by a $+$ and the trajectory of the wavepacket by the heavy line. (b) The evolution of the vertical wavenumber m, shown as the thin line, and ten times the horizontal wavenumbers k, the thick line, and l, the dashed line. (c) The evolution of the two gravity wave modes and the amplitudes δ and ζ where the plus and minus gravity waves have been displaced by $\pm 3 \times 10^{-4}$ respectively. ζ, which has been multiplied by 1×10^{-3}, is the solid line in the centre and δ is the dotted line. (d) The norm of the wavepacket's strain matrix divided by its frequency.

that has been run with different initial conditions. It is worth noting here that these systems are very dependent on initial conditions with only a very slight change in the position of a vortex, for instance, making the system unrecognisable. As can be seen in this figure all the observations that were made about the previous set of results also hold true for this set. While appearing totally different to the last simulation we can

see that although the point vortex trajectories look a lot more complex, the wavepacket seems to make a relatively simple path through them which is reflected in the fact that the wavenumbers have a more uniform evolution. Even though this is the case it is still clear that the amplitudes of the wavepacket grow when the wavenumbers do. The link between the wavepacket's local Rossby number and the growth of the wavepacket's amplitudes also holds true for this simulation.

Although there are only results from two simulations shown here we have already been able to see some trends forming about the results. By looking at countless different examples of these systems it is clear that these trends continue.

6.9 Conclusion

We have shown that spontaneous generation of inertia-gravity waves can take place in a simple flow. The transient change in the velocity field caused by a point vortex dipole was sufficient to significantly change the wavevector and cause the gravity-wave modes to be excited. We also showed that the amplitude of this growth is exponentially dependent on the distance of the wavepacket from the dipole axis. Growth was then shown to exist in the almost elliptical trajectories that exist in close proximity to the point vortex dipole.

Following this, we used the theory that produced these results to study flows behaving like random-strain flows that are created by multiple point vortices. We have shown that growth exists in these flows and have linked the areas of growth to the areas where the wavenumbers experience rapid change. We also proposed a local Rossby number that along with the wavenumbers is a good indication of growth of the wavepacket.

The almost elliptical flows studied in this chapter will be the motivation for the next chapter. There we will attempt to quantify the growth in the wavepacket that was shown to exist in this chapter.

Chapter 7

Elliptical instability

7.1 Introduction

In this chapter we will study a flow of intermediate complexity, in the sense that its complexity lies between that of the dipole that was studied in the last chapter and the complexity of fully chaotic flows. This flow of intermediate complexity arises any time there are enclosed streamlines present in the flow which will cause any particle that is on one of these streamlines to have a periodic trajectory. The periodic nature of this trajectory means that any growth that is associated with transient events is repeated every period. The repetitive nature of these small transient growths can lead to instabilities arising. The simplest case of a flow that has enclosed streamlines, and hence periodic trajectories, is an elliptical flow. The simplest case of this flow is a circular flow but it is too symmetric for any interesting behaviour to arise and therefore we will discount this case from here on. Elliptical instabilities have been well studied (see [27] for a review of the main results up to 2002) but here we come at it from a slightly different viewpoint and see them as a mechanism for inertia-gravity wave generation.

The aim of this chapter is to interpret elliptical instabilities, that is the three-dimensional instability of two-dimensional flows with elliptic streamlines, of a rapidly rotating and strongly stratified fluid in the context of WKB. We derive an expression for the growth rate of these instabilities asymptotically. In the context of this chapter, we study the simplest instance of such flows, that of an unbounded elliptical vortex with uniform velocity. It is worth noting that this work can be extended to

apply to any flow that has elliptical streamlines at any position, or at any time in its evolution. One of these situations that helps motivate this study is the instability of two-dimensional vortices that are deformed elliptically by a large scale strain flow. This is especially important for the dynamics of the atmosphere and oceans since they are characterised by an abundance of vortices that are deformed through either mutual interactions or the effect of large scale flows. Another motivation is the elliptical streamlines that surround the point vortices in a point vortex dipole, as shown in the last chapter.

Rotation and stratification exert a strong influence on the stability of a wavepacket in an elliptical flow. This stems from the resonance between the periodic fluctuations associated with the elliptical flow and the free waves that the flow's dispersion relation permits. In the presence of both rotation and stratification, these waves are inertia-gravity waves whose frequency is bounded from below by the minimum of the Coriolis parameter, f, and the Brunt-Väisälä frequency, N. As a consequence, a vortex of fixed vorticity ceases to be unstable by the subharmonic instability responsible for the simplest form of elliptical instability when f and N exceed a certain threshold. As these parameters increase further, instabilities are limited to resonances of higher and higher order, leading to decreasing growth rates. This was clearly demonstrated by Miyazaki in [41] on the basis of numerical solutions of the Floquet problem that models elliptic instability.

As in the rest of this thesis, in this chapter we concentrate on the regime of rapid rotation and strong stratification with $N > f$, which is most relevant to the atmosphere and oceans. We are then motivated by the role that instabilities play in the generation of inertia-gravity wave-like motion and the resulting breakdown of the nearly geostrophic and hydrostatic balance that is typical of much of the atmosphere and ocean. This mirrors closely the motivation of the work that has been done by McWilliams and Yavneh, [40].

In this chapter the elliptical instabilities are studied by first formulating the problem using the equations derived in chapter 4. The instability problem is then formulated in terms of Floquet multipliers which are used to find an expression for the growth rate. This expression is calculated by directly relating the growth of the solutions to the occurrence of a Stokes phenomenon, which we capture using a combination of WKB expansion and matched asypmtotics in complex time. The analytical results are then confirmed by the numerical solution of the Floquet problem.

The work in this chapter is an expanded version of a paper that has been published in the journal, Physics of Fluids, [5].

7.2 Formulation

We consider the evolution of wavepackets in a horizontal elliptical flow in a three-dimensional stratified fluid, with constant Brunt–Väisälä frequency, N, rotating about the vertical axis at rate $f/2 > 0$. To achieve elliptic streamlines, the flows streamfunction is given by

$$\Psi = -\frac{1}{2}\left(bx^2 + ay^2\right), \tag{7.1}$$

which gives that its velocity field is

$$U = ay, \tag{7.2}$$

$$V = -bx, \tag{7.3}$$

$$W = 0, \tag{7.4}$$

and its vorticity is given by

$$\Omega = a + b. \tag{7.5}$$

In this formulation we always have $ab > 0$ which ensures that we always have ellipses rather than hyperbolas. From the above equations, there is no vertical motion in the system and so $R = -N^2 z$. We define

$$\varsigma = \operatorname{sgn} a = \operatorname{sgn} b, \tag{7.6}$$

and note that the flow is anticyclonic for $\varsigma = 1$ and cyclonic for $\varsigma = -1$. The flow can be characterised by three dimensionless parameters, namely

$$e = \sqrt{a/b}, \quad \epsilon = \sqrt{ab}/f, \quad \text{and} \quad f/N, \tag{7.7}$$

which are recognised respectively as the aspect ratio of the elliptical flow, a Rossby number and the Prandtl ratio. Without loss of generality we can assume that $e > 1$. This is because all the values of $e < 1$ can be thought of as an ellipse, with $e > 1$, rotated by $\pi/2$. Figure 7.1 shows the stream lines and velocity field for this flow where $a = 4$ and $b = 1$, which gives that $e = 2$.

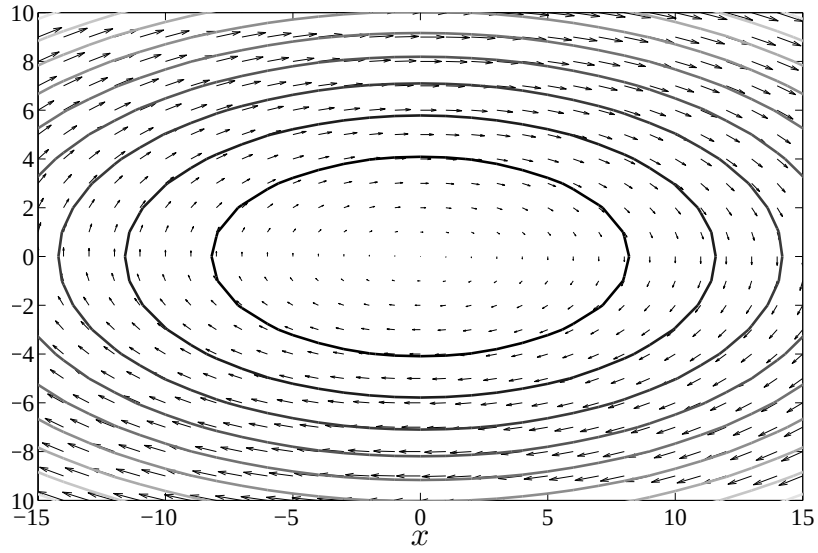

Figure 7.1: The stream lines and velocity field of an anticyclonic elliptical flow given by (7.1)-(7.4), where $a = 4$ and $b = 1$ which gives that $e = 2$.

To find the evolution of the wavepacket's wavenumbers we substitute (7.2)-(7.3) into (4.26) giving that

$$\dot{k} = bl, \tag{7.8}$$

$$\dot{l} = -ak, \tag{7.9}$$

$$\dot{m} = 0, \tag{7.10}$$

where the overdot denotes differentiation with respect to t. Solving this set of equations gives that

$$k = k_0 \cos(\sqrt{ab}t), \tag{7.11}$$

$$l = -\varsigma e k_0 \sin(\sqrt{ab}t), \tag{7.12}$$

$$m = m_0, \tag{7.13}$$

where $k_0 = k(0)$ and $m_0 = m(0)$ are constant. To simplify things, from this point on we will use a dimensionless time variable obtained by taking $(ab)^{-1/2}$ as a reference time, which means that the wavenumbers are now given by

$$k = k_0 \cos t, \tag{7.14}$$

$$l = -\varsigma e k_0 \sin t, \tag{7.15}$$

$$m = m_0. \tag{7.16}$$

The stability of the wavepackets that have been placed in the flow given by (7.2)-(7.4) depends on the behaviour of the amplitudes, δ, ζ and q, whose evolutions are governed by (4.56)-(4.58). These three equations can be combined into one second order equation, similar to that found in [40]. This is done by first differentiating (4.57) with respect to t and then substituting in (4.56). After assuming that the perturbations potential vorticity, q, vanishes and continuing to use the scaled version of t as above, we have that

$$\ddot{\zeta} + \frac{2\varsigma klm^2(e-e^{-1})}{\kappa^2(k^2+l^2)}\dot{\zeta} + \frac{1}{\epsilon^2}\left[\left(1-\varsigma\epsilon(e+e^{-1})\left(1-\frac{2\varsigma\epsilon e k_0^2}{k^2+l^2}\right)\frac{m^2}{\kappa^2} + \frac{N^2(k^2+l^2)}{f^2\kappa^2}\right]\zeta = 0. \tag{7.17}$$

There are four dimensionless parameters that appear in this equation: the three flow parameters that were given in (7.7) and the initial aspect ratio m_0/k_0 of the perturbation. Note that anticyclonic flows (with $\varsigma = 1$) are susceptible to centrifugal instability when the relative vorticity exceeds f, that is, for $\epsilon(e+e^{-1}) > 1$. Since we focus on a regime that is rapidly rotating, that is $\epsilon \ll 1$, we do not need to consider this instability mechanism any further.

The rest of this chapter focuses on a limiting case of (7.17) obtained by making the hydrostatic approximation, see section 2.4. This approximation uses the fact that the horizontal length scales are much larger than the vertical length scales to write the vertical gradient of the pressure as the density times the gravitational acceleration. This is done by letting $m_0 \gg k_0$ and $N \gg f$ in such a way that

$$\mu = \frac{f m_0}{N k_0} = O(1). \tag{7.18}$$

This is the regime most relevant to the dynamics of the atmosphere and oceans since there the horizontal length scales are clearly much larger than the vertical ones and also the condition that $N \gg f$ is verified. We will also demonstrate below that the largest growth rates of the elliptical instability correspond to $\mu = O(1)$. It might be useful to note here that the parameter μ can be recognised as the inverse square root of the Burger number, which is a measure of the scale of the motion of the fluid relative to its deformation radius. It can also be interpreted as the aspect ratio of the perturbation

scaled by f/N as is natural in rapidly rotating, strongly stratified fluids.

In the hydrostatic approximation, since $m_0 \gg k_0$, κ^2 is approximated by m^2, and (7.17) reduces to

$$\ddot{\zeta} + \frac{2\varsigma kl(e-e^{-1})}{k^2+l^2}\dot{\zeta} + \frac{1}{\epsilon^2}\left[\left(1-\varsigma\epsilon(e+e^{-1})\left(1-\frac{2\varsigma\epsilon e k_0^2}{k^2+l^2}\right)+\frac{N^2(k^2+l^2)}{f^2m^2}\right]\zeta = 0. \quad (7.19)$$

Using (7.14)-(7.16) and utilising the fact that $e > 1$ to define $\psi > 0$ by

$$e^2 = 1 + \psi^2, \quad (7.20)$$

we rewrite this equation as

$$\ddot{\zeta} - p(t)\dot{\zeta} + \frac{1}{\epsilon^2}\left[\omega^2(t) - \epsilon q(t) + \epsilon^2 r(t)\right]\zeta = 0. \quad (7.21)$$

Here

$$\omega^2 = 1 + \frac{N^2(k^2+l^2)}{f^2m^2} = 1 + \mu^{-2}(1+\psi^2\sin^2 t), \quad (7.22)$$

can be recognised as the square of the inertia-gravity wave frequency, as derived in section 3.3 after non-dimensionalising the time by scaling by $1/f$. The other functions in the equation are given by

$$p(t) = \frac{\psi^2\sin(2t)}{1+\psi^2\sin^2 t}, \quad (7.23)$$

$$q(t) = \varsigma\left(e + e^{-1} + \frac{2e}{1+\psi^2\sin^2 t}\right), \quad (7.24)$$

$$r(t) = \frac{2(e^2+1)}{1+\psi^2\sin^2 t}. \quad (7.25)$$

Equation (7.21) is a Hill equation, with coefficients that are π-periodic in t, whose stability can be determined using the Floquet theory for Hill equations found in [8]. Briefly, if $\boldsymbol{\zeta}(t) = (\zeta_1(t), \zeta_2(t))^{\mathrm{T}}$ is a column vector of independent solutions then the π-periodicity of the equation's coefficients means that we can write

$$\boldsymbol{\zeta}(t+\pi) = M\boldsymbol{\zeta}(t), \quad (7.26)$$

for some constant matrix M. The eigenvalues λ of M are then the Floquet multipliers,

and two fundamental solutions can be found for which

$$\zeta(t) = e^{\sigma t}\phi(t), \tag{7.27}$$

where

$$\sigma = \frac{1}{\pi}\log\lambda, \tag{7.28}$$

is the Floquet exponent, and $\phi(t)$ is π-periodic.

A useful and general result about the Floquet multipliers, λ_1 and λ_2, that will be used later, is that they satisfy $\lambda_1\lambda_2 = 1$. This can be derived by using the Wronskian of two solutions, ζ_1 and ζ_2, which is defined as

$$W = \zeta_1\dot{\zeta}_2 - \dot{\zeta}_1\zeta_2. \tag{7.29}$$

Introducing solutions of (7.21), of the form of (7.27), gives that

$$W = e^{(\sigma_1+\sigma_2)t}\chi, \tag{7.30}$$

where $\chi = \phi_1\phi_2(\sigma_2 - \sigma_1) + \phi_1\dot{\phi}_2 - \phi_2\dot{\phi}_1$. Next, the definition of the Wronskian, (7.29), can be differentiated with respect to t to give that

$$\dot{W} = \zeta_1\ddot{\zeta}_2 - \ddot{\zeta}_1\zeta_2, \tag{7.31}$$

which after substituting in (7.21), rearranged to be in terms of $\ddot{\zeta}$, gives that $\dot{W} = p(t)W$. We can now use the fact that $p(t)$ is π-periodic with zero mean to infer that W is also π-periodic. From this it is clear that $W(0) = W(\pi)$ which when used in (7.30) gives that

$$\chi(0) = \chi(\pi)e^{(\sigma_1+\sigma_2)\pi}. \tag{7.32}$$

Since ϕ was introduced as π-periodic we also have that $\chi(0) = \chi(\pi)$, so we are left with

$$e^{(\sigma_1+\sigma_2)\pi} = 1. \tag{7.33}$$

From the definition of λ given in (7.28) this can now be recognised as the product of the two eigenvalues involved and so we have proved that $\lambda_1\lambda_2 = 1$ as required.

The form of the solutions given in (7.27) guarantees that growth is going to occur

and so there are instabilities when $\mathrm{Re}\,\sigma > 0$. From (7.28) we see this to be equivalent to requiring that one of these eigenvalues is such that $|\lambda| > 1$. Hence to try and find these instabilities we can compute M and its eigenvalues and try to find values of the system's parameters that will satisfy this condition. The matrix M is computed by relating $\boldsymbol{\zeta}$ and $\dot{\boldsymbol{\zeta}}$ at two times t and $t+\pi$. Here we choose $t = -\pi/2$ and compute M as

$$M = [\boldsymbol{\zeta}(\pi/2), \dot{\boldsymbol{\zeta}}(\pi/2)][\boldsymbol{\zeta}(-\pi/2), \dot{\boldsymbol{\zeta}}(-\pi/2)]^{-1}. \tag{7.34}$$

Our aim is to determine the largest values of the growth rate $\mathrm{Re}\,\sigma$ for (7.21) analytically in the fast-rotation regime $\epsilon \ll 1$, with $N \gg f$, $\mu = O(1)$ and $\psi = O(1)$. In this regime (7.21) resembles the Hill equations with large parameters whose stability has been studied by Weinstein and Keller, [71], using a mapping to parabolic cylinder functions. However, there are difficulties in applying their results directly, related to the presence of a first derivative term that is singular for the complex values of t such that $k^2 + l^2 = 0$. It has therefore been simpler to develop a different approach, combining WKB analysis with complex-time matching. This approach, which has the advantage of demonstrating the link between the instability and the Stokes phenomenon, [1, 48], is described in the next few sections. The analytical results obtained there are confirmed and extended to finite N/f in section 7.9 by solving (7.21) numerically.

7.3 WKB analysis

We can now carry out WKB analysis on (7.21). It is worth noting that this analysis is carried out on the ordinary differential equations that were obtained from the original set of partial differential equations by a quite different WKB analysis in chapter 4. Here we use a WKB solution to (7.21) of the form

$$\zeta = A(t)e^{i\theta(t)/\epsilon}, \tag{7.35}$$

where the real functions $A(t)$ and $\theta(t)$ can be written as a power series in terms of ϵ like

$$A = A_0 + \epsilon A_1 + \ldots \quad \text{and} \quad \theta = \theta_0 + \epsilon\theta_1 + \ldots. \tag{7.36}$$

Substituting this solution into (7.21) gives that at leading order, $O(1/\epsilon^2)$, we find

$$\theta_0(t) = \int_{-\pi/2}^{t} \omega(t')\, dt'. \tag{7.37}$$

At the next order, we have

$$\frac{\dot{A}_0}{A_0} = -\frac{\dot{\omega}}{2\omega} + \frac{p}{2}, \tag{7.38}$$

$$\dot{\theta}_1 = -\frac{q}{2\omega}. \tag{7.39}$$

From this we can note that $A_0(t)$ is π-periodic since $p(t)$ has zero mean in $[-\pi/2, \pi/2]$. Pursuing the computation to higher orders suggests that all the $A_n(t)$, $n = 0, 1, \ldots$ are π-periodic which can be confirmed by calculating the Wronskian of the pair of solutions

$$\zeta = A(t)e^{i\theta(t)/\epsilon} \quad \text{and} \quad \bar{\zeta} = A(t)e^{-i\theta(t)/\epsilon}. \tag{7.40}$$

Using the definition of the Wronskian given in (7.29) we get that

$$W = 2i\epsilon^{-1}A^2\dot{\theta}, \tag{7.41}$$

which when differentiated with respect to t gives that

$$\dot{W} = W\left(\frac{\ddot{\theta}}{\dot{\theta}} + 2\frac{\dot{A}}{A}\right). \tag{7.42}$$

Using the leading order identities derived above for $\dot{\theta}$ and $\dot{A}/A$ in this expression yields that $\dot{W} = pW$, which is what was expected from previous results and which also again gives us that W is π-periodic. The recurrence relations for A_n and θ_n, of which (7.38) and (7.39) are the first terms, give $\dot{\theta}_n$ in terms of periodic functions and A_m for $m < n$. Using the Wronskian then gives that A_n is π-periodic which in turn shows that $\dot{\theta}_{n+1}$ is π-periodic and hence the periodicity of A_n, for all n, follows inductively.

The periodicity of the asymptotic series for $A(t)$ implies that there is no instability to any algebraic order in ϵ. To all algebraic orders, the fundamental solutions are given by (7.40) from which we can can calculate M from (7.34). After taking $\theta(-\pi/2) = 0$, which we can do without loss of generality, we find that

$$M = \begin{pmatrix} e^{i\theta/\epsilon} & 0 \\ 0 & e^{-i\theta/\epsilon} \end{pmatrix}, \tag{7.43}$$

where $\theta = \theta(\pi/2)$, which has eigenvalues $\lambda = \exp(\pm i\theta/\epsilon)$. These are the Floquet multipliers of the system and they clearly satisfy $|\lambda| = 1$. Since the instability condition is that $|\lambda| > 1$, this shows that to all orders in ϵ the system is inherently stable and so any instability is going to be due to exponentially small effects. Hence to capture these instabilities we will need to go beyond the standard WKB analysis done above and examine how exponentially small terms alter the form of the solutions in (7.40).

7.4 The Stokes phenomenon

The exponentially small terms that we need to find arise through a Stokes phenomenon associated with the presence of complex turning points, that is, complex values of t for which $\omega(t) = 0$. From these points emanate lines across which solutions defined by asymptotic series such as (7.36) acquire an additional exponentially small multiple of other linearly independent solutions. See [9, 23, 12] for analysis of the stokes phenomenon in a variety of problems.

Since the matrix M found in (7.34) depends only on the solution vector, $\boldsymbol{\zeta}$, at $t = \pm\pi/2$, the only Stokes lines relevant to our problem are those crossing the real t-axis between $-\pi/2$ and $\pi/2$. As mentioned above, the Stokes phenomenon occurs at complex values of t for which $\omega = 0$. From (7.22), these are located at

$$t_n = i \sinh^{-1} \frac{\sqrt{1+\mu^2}}{\psi} + n\pi, \quad n = 0, \pm 1, \pm 2, \cdots, \tag{7.44}$$

and $\bar{t}_n$. The Stokes phenomenon then occurs as $t \in \mathbb{R}$ crosses one of the Stokes lines joining t_n to $\bar{t}_n$. Our interest is in the Stokes lines that cross the interval $[-\pi/2, \pi/2]$ and it is easy to see that the segment of $\operatorname{Re} t = 0$ joining t_0 to $\bar{t}_0$ is such a Stokes line. From (7.44) it is also clear that for $n > 0$, t_n is going to be outside the interval we are looking at and so this line is the only one of interest to us. Across this line, the asymptotic solutions ζ and $\bar{\zeta}$ in (7.40) switch on an exponentially small multiple of $\bar{\zeta}$ and ζ, respectively. More specifically (7.40) defines a pair of independent solutions that are valid for $-\pi/2 \le t < -\delta$, for some $\epsilon^{1/2} \ll \delta \ll 1$. When crossing $\operatorname{Re} t = 0$ these

solutions gain additional terms and so for $\delta < t \leq \pi/2$, they become

$$\zeta = A(t)\left[e^{i\theta(t)/\epsilon} + Se^{-i\theta(t)/\epsilon}\right] \quad \text{and} \quad \bar{\zeta} = A(t)\left[e^{-i\theta(t)/\epsilon} + \bar{S}e^{i\theta(t)/\epsilon}\right], \tag{7.45}$$

where S is an exponentially small constant called the Stokes multiplier. Note that $A(t)$ and $\theta(t)$ appearing in (7.40) and (7.45) are defined by the asymptotic series (7.36), which diverge and must be truncated. The theory of the Stokes phenomenon, [9], shows that optimal truncation, in which the series are truncated near their smallest term, leads to truncation errors that are asymptotically smaller than the Stokes multiplier, S. Thus (7.40) and (7.45) can be interpreted consistently by regarding $A(t)$ and $\theta(t)$ as defined by their asymptotic series truncated optimally. We adopt this interpretation and carry out our analysis neglecting all $o(S)$ terms.

7.5 Calculating M

The Stokes multiplier, S, turns out to determine completely the leading-order form of the maximum growth rate of the instability. To show this we compute the matrix M that was defined in (7.34). To do this we start by working out the first matrix in the expression for M by taking $\boldsymbol{\zeta} = (\zeta, \bar{\zeta})^{\mathrm{T}}$ where ζ and $\bar{\zeta}$ were defined in (7.45) to get that

$$[\boldsymbol{\zeta}(\pi/2), \dot{\boldsymbol{\zeta}}(\pi/2)] = \begin{pmatrix} A(e^{i\theta/\epsilon} + Se^{-i\theta/\epsilon}) & (\dot{A} + i\epsilon^{-1}\dot{\theta}A)e^{i\theta/\epsilon} + S(\dot{A} - i\epsilon^{-1}\dot{\theta}A)e^{-i\theta/\epsilon} \\ A(e^{-i\theta/\epsilon} + \bar{S}e^{i\theta/\epsilon}) & (\dot{A} - i\epsilon^{-1}\dot{\theta}A)e^{-i\theta/\epsilon} + \bar{S}(\dot{A} + i\epsilon^{-1}\dot{\theta}A)e^{i\theta/\epsilon} \end{pmatrix}, \tag{7.46}$$

where A, $\dot{A}$, θ and $\dot{\theta}$ are all evaluated at $t = \pi/2$. The second matrix in the expression for M can also be calculated in a similar way where ζ and $\bar{\zeta}$ are now defined by (7.40). After noting that we earlier picked that $\theta(-\pi/2) = 0$ this matrix becomes

$$[\boldsymbol{\zeta}(-\pi/2), \dot{\boldsymbol{\zeta}}(-\pi/2)] = \begin{pmatrix} A & \dot{A} + i\epsilon^{-1}\dot{\theta}A \\ A & \dot{A} - i\epsilon^{-1}\dot{\theta}A \end{pmatrix}. \tag{7.47}$$

Here A, $\dot{A}$ and $\dot{\theta}$ are also evaluated at $t = \pi/2$, as above, since their periodicity gives that their values at $t = \pm\pi/2$ coincide. This matrix has determinant $-2i\epsilon^{-1}A^2\dot{\theta}$ and

so inverting it gives that

$$[\boldsymbol{\zeta}(-\pi/2), \dot{\boldsymbol{\zeta}}(-\pi/2)]^{-1} = \frac{i\epsilon}{2A^2\dot{\theta}} \begin{pmatrix} \dot{A} - i\epsilon^{-1}\dot{\theta}A & -\dot{A} - i\epsilon^{-1}\dot{\theta}A \\ -A & A \end{pmatrix}. \tag{7.48}$$

By using (7.46) and (7.48) we can now compute M which gives the surprisingly simple result that

$$M = \begin{pmatrix} e^{i\theta/\epsilon} & Se^{-i\theta/\epsilon} \\ \bar{S}e^{i\theta/\epsilon} & e^{-i\theta/\epsilon} \end{pmatrix}. \tag{7.49}$$

It is of interest to compare this expression for M with the one found in (7.43) and notice that exponentially small terms have appeared on the off-diagonal of the matrix. If this is done more generally, i.e. without arbitrarily choosing that $\theta(-\pi/2) = 0$, we get that $\theta = \theta(\pi/2) - \theta(-\pi/2)$ which can accommodate any convention for the arbitrary choice of $\theta(-\pi/2)$.

The Floquet multipliers, that is the eigenvalues, λ, of (7.49), are now solutions of the characteristic equation

$$(\lambda - e^{i\theta/\epsilon})(\lambda - e^{-i\theta/\epsilon}) = |S|^2. \tag{7.50}$$

Since $|S|^2$ is very small the solutions of this equation can be written as

$$\lambda_1 = e^{i\theta/\epsilon} + \delta x_1 + \delta^2 x_2 + \ldots, \tag{7.51}$$

$$\lambda_2 = e^{-i\theta/\epsilon} + \delta \bar{x}_1 + \delta^2 \bar{x}_2 + \ldots, \tag{7.52}$$

where, for simplicity, we have let $|S|^2 = \delta$ where $\delta \ll 1$. When the above expression for the first eigenvalue, λ_1, is introduced into (7.50) it gives that at $O(\delta)$,

$$x_1 = 1/(e^{i\theta/\epsilon} - e^{-i\theta/\epsilon}). \tag{7.53}$$

For the case where $e^{i\theta/\epsilon} \neq e^{-i\theta/\epsilon}$ we then get that $\lambda = e^{i\theta/\epsilon} + O(|S|^2)$ and so $|\lambda| = 1$, which from our earlier instability condition means that there will not be any instabilities. However, this breaks down when $e^{i\theta/\epsilon} = e^{-i\theta/\epsilon}$. In this case a different

expansion, of the form

$$\lambda_1 = e^{i\theta/\epsilon} + \delta^{1/2}x_1 + \delta x_2 + \ldots, \tag{7.54}$$

$$\lambda_2 = e^{-i\theta/\epsilon} + \delta^{1/2}\bar{x}_1 + \delta\bar{x}_2 + \ldots, \tag{7.55}$$

needs to be used, [25]. Following the same procedure with these new solutions gives that $\lambda = e^{i\theta/\epsilon} + O(|S|)$. Hence up to $o(S)$ errors, the Floquet multipliers, that is, the eigenvalues of (7.49), satisfy $|\lambda| = 1$ unless M is degenerate or nearly degenerate, in the sense that $\exp(i\theta) = \exp(-i\theta) + O(S)$. Thus the condition for instability, $|\lambda| > 1$, requires that $\exp(i\theta)$ is exponentially close to ± 1. This can be related to the fact that $\det M = 1$ which follows from the periodicity of the Wronskian, W, and implies that instability arises when the λ are purely real. However, they are complex conjugate to all orders in ϵ, a property which persists in the presence of a small perturbation in the nondegenerate cases $\lambda \neq \pm 1 + O(S)$. We therefore suppose that the parameters, e, μ and ϵ are such that

$$e^{i\theta/\epsilon} = \pm(1 + iT) + O(T^2), \tag{7.56}$$

$$e^{-i\theta/\epsilon} = \pm(1 - iT) + O(T^2). \tag{7.57}$$

for some $T \in \mathbb{R}$ of a similar order of magnitude as S. The Floquet multipliers of (7.49) are then given by

$$\lambda = \pm\left(1 + \sqrt{|S|^2 - T^2}\right) + o(S) \quad \text{and} \quad \lambda = \pm\left(1 - \sqrt{|S|^2 - T^2}\right) + o(S). \tag{7.58}$$

Since we need at least one of these to satisfy $|\lambda| > 1$ for the flow to be unstable, it is clear that we need $T^2 \leq |S|^2$, i.e. $-|S| \leq T \leq |S|$. This shows that since T and S are exponentially small the flow is unstable only in the exponentially narrow bands that satisfy this condition. From (7.28) we then have that the corresponding growth rate is $\sigma = \sqrt{|S|^2 - T^2}/\pi + o(S)$ which has a maximum value at the centre of these bands. This occurs at $T = 0$ and so the maximum growth rate is given by

$$\sigma_{\max} \sim \frac{|S|}{\pi}. \tag{7.59}$$

7.6 Using exponential asymptotics to calculate S

The task now is to compute the exponentially small Stokes multiplier, S, which quantifies the switching on of the branch of the WKB solution given in (7.45) as the Stokes line is crossed [1, 48]. We compute S by using matched asymptotics to examine how the solution to the left of the Stokes line, (7.40), connects to the solution to the right of the Stokes line, (7.45). To analyse the behaviour of the solution in the neighbourhood of t_0, we first expand ω^2 about t_0 which gives that

$$\omega \sim a^{1/2} e^{i\pi/4} (t - t_0)^{1/2}, \tag{7.60}$$

where

$$a = \frac{2}{\mu^2} \sqrt{(1 + \mu^2)(1 + \psi^2 + \mu^2)}. \tag{7.61}$$

We then rescale the evolution equation for ζ near t_0 by first realising that as $t \to t_0$, (7.21) becomes

$$\ddot{\zeta} + ia\epsilon^{-2}(t - t_0)\zeta \approx 0, \tag{7.62}$$

at leading order. Next we let τ be of the form $\tau = (t - t_0)/e^{\alpha}$ and pick α so that we get a dominant balance between the two terms of the above equation. This leads us to define the inner variables as

$$\tau = \epsilon^{-2/3} a^{1/3} (t - t_0) \quad \text{and} \quad Z(\tau) = \zeta(t), \tag{7.63}$$

which, at leading order, transforms (7.21) into

$$\frac{d^2 Z}{d\tau^2} + i\tau Z = 0. \tag{7.64}$$

Solutions to this equation can be written in terms of the Airy functions $\mathrm{Ai}(e^{-i\pi/6}\tau)$ and $\mathrm{Bi}(e^{-i\pi/6}\tau)$ and we claim that the solution of interest is

$$Z \sim C \left(\mathrm{Ai}(e^{-i\pi/6}\tau) + i\mathrm{Bi}(e^{-i\pi/6}\tau) \right). \tag{7.65}$$

One would usually expect there to be constants in front of both of the Airy functions in the above solution but since we know that we need Z to be a purely imaginary exponential we can combine these two constants to give us the form that is shown above. An expression for this constant C is found while verifying that this solution

matches $A(t)\exp(i\theta(t)/\epsilon)$ to the left of the Stokes line $\operatorname{Re} t = 0$. It is convenient to verify the matching along the line where the phase of τ is $-5\pi/6$, i.e. $\operatorname{ph}\tau = -5\pi/6$, which is an anti-Stokes line. This means that here the two independent solutions of (7.64) have similar orders of magnitude. Along this line, we can use the asymptotic formulae

$$\operatorname{Ai}(-z) \sim \frac{1}{\sqrt{\pi} z^{1/4}} \cos(2z^{3/2}/3 - \pi/4), \tag{7.66}$$

$$\operatorname{Bi}(-z) \sim -\frac{1}{\sqrt{\pi} z^{1/4}} \sin(2z^{3/2}/3 - \pi/4), \tag{7.67}$$

found in [2], with $z = -\exp(-i\pi/6)\tau = \exp(5i\pi/6)\tau \to +\infty$. Substituting these formulae back into (7.65) gives that along this line

$$Z \sim \frac{C e^{i\pi/24}}{\sqrt{\pi}\tau^{1/4}} e^{2ie^{i\pi/4}\tau^{3/2}/3}. \tag{7.68}$$

On the other hand, using (7.38)-(7.39) the solution $\zeta = A(t)\exp(i\theta(t))$, valid in the outer region away from t_0 and to the left of the Stokes line $\operatorname{Re} t = 0$, can be written as

$$\begin{aligned} \zeta &\sim \frac{1}{\omega^{1/2}} e^{\int_{-\pi/2}^{t} p(t')\,\mathrm{d}t'/2} e^{i\epsilon^{-1}\int_{-\pi/2}^{t}[\omega(t') - \epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'} \\ &\sim \frac{e^{-i\pi/8}}{(\epsilon a)^{1/6}\tau^{1/4}} e^{\int_{\Gamma_-} p(t')\,\mathrm{d}t'/2} e^{i\epsilon^{-1}\int_{\Gamma_-}[\omega(t') - \epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'} e^{2ie^{i\pi/4}\tau^{3/2}/3}, \end{aligned} \tag{7.69}$$

where we have used the asymptotic expression for ω that was given in (7.60). Here Γ_- denotes a contour joining $-\pi/2$ to t_0, while remaining to the left of the Stokes line $\operatorname{Re} t = 0$ as shown in figure 7.2. Comparing (7.69) with (7.68) shows that ζ correctly matches Z provided that

$$C = \frac{\sqrt{\pi} e^{-i\pi/6}}{(\epsilon a)^{1/6}} e^{\int_{\Gamma_-} p(t')\,\mathrm{d}t'/2} e^{i\epsilon^{-1}\int_{\Gamma_-}[\omega(t') - \epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'}. \tag{7.70}$$

We now match Z with the outer solution valid to the right of the Stokes line $\operatorname{Re} t = 0$. A connection formula for Airy functions, again found in [2], gives that

$$\operatorname{Ai}(ze^{\pm 2i\pi/3}) = \frac{1}{2} e^{\pm i\pi/3}(\operatorname{Ai}(z) \mp i\operatorname{Bi}(z)), \tag{7.71}$$

which we can use to rewrite (7.65) as

$$Z \sim 2Ce^{i\pi/3}\operatorname{Ai}(e^{-i5\pi/6}\tau). \tag{7.72}$$

Carrying out the matching on the anti-Stokes line $\mathrm{ph}\,\tau = -\pi/6$, we can use the asymptotic formula for $\mathrm{Ai}(-z)$ given in (7.66) to write that

$$Z \sim \frac{Ce^{i\pi/24}}{\sqrt{\pi}\tau^{1/4}}\left(e^{2ie^{i\pi/4}\tau^{3/2}/3} + ie^{-2ie^{i\pi/4}\tau^{3/2}/3}\right), \tag{7.73}$$

for $|\tau| \to \infty$. This should be matched with the form

$$\zeta(t) = A(t)[\exp(i\theta(t)/\epsilon) + S\exp(-i\theta(t)/\epsilon)], \tag{7.74}$$

of the solution to the right of the Stokes line. Using (7.38),(7.39) and (7.60) again gives the asymptotics

$$\begin{aligned}
\zeta &\sim \frac{1}{\omega^{1/2}} e^{\int_{-\pi/2}^{t} p(t')\,\mathrm{d}t'/2}\left(e^{i\epsilon^{-1}\int_{-\pi/2}^{t}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'}\right.\\
&\qquad\left. + Se^{-i\epsilon^{-1}\int_{-\pi/2}^{t}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'}\right)\\
&\sim \frac{e^{-i\pi/8}}{(\epsilon a)^{1/6}\tau^{1/4}} e^{\int_{\Gamma_+} p(t')\,\mathrm{d}t'/2}\left(e^{i\epsilon^{-1}\int_{\Gamma_+}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'} e^{2ie^{i\pi/4}\tau^{3/2}/3}\right.\\
&\qquad\left. + Se^{-i\epsilon^{-1}\int_{\Gamma_+}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'} e^{-2ie^{i\pi/4}\tau^{3/2}/3}\right),
\end{aligned} \tag{7.75}$$

where Γ_+ denotes a contour joining $-\pi/2$ to t_0 as also shown in figure 7.2. This contour crosses the Stokes line $\mathrm{Re}\,t = 0$ below the singular point t_p of $p(t)$ and $q(t)$, given by $t_\mathrm{p} = i\sinh^{-1}(1/\psi)$. Taking (7.70) into account, the matching with (7.73) gives the two equations

$$e^{\int_{\Gamma_-} p(t)\,\mathrm{d}t/2} e^{-i\int_{\Gamma_-} q(t)/\omega(t)\,\mathrm{d}t/2} = e^{\int_{\Gamma_+} p(t)\,\mathrm{d}t/2} e^{-i\int_{\Gamma_+} q(t)/\omega(t)\,\mathrm{d}t/2}, \tag{7.76}$$

$$\begin{aligned}
ie^{\int_{\Gamma_-} p(t')\,\mathrm{d}t'/2} e^{i\epsilon^{-1}\int_{\Gamma_-}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'} &=\\
+ Se^{\int_{\Gamma_+} p(t')\,\mathrm{d}t'/2} e^{-i\epsilon^{-1}\int_{\Gamma_+}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'}&.
\end{aligned} \tag{7.77}$$

We can now deform the integration contours. By doing so we get that the difference $\int_{\Gamma_+} - \int_{\Gamma_-}$ reduces to an integral along a closed contour encircling t_p. This is verified graphically in figure 7.2. To compute the integrals in the above equations we first need to find the residues of p and q/ω at $t = t_p$.

To find $\mathrm{Res}_{t_\mathrm{p}} p$ we use the definition of p found in (7.23) and then expand it as a Laurent series about $t = t_p$. We first expand $\sin(2t)$ and $\sin^2(t)$ about this point to get

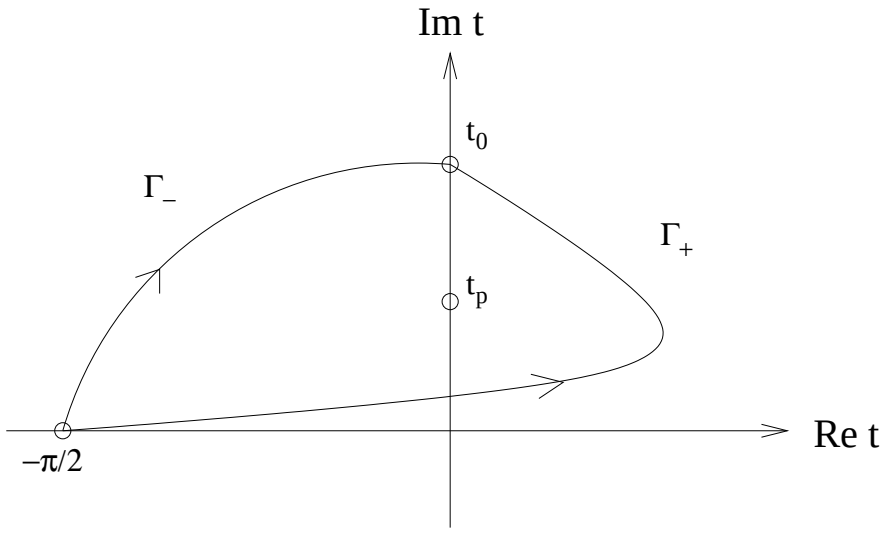

Figure 7.2: The paths of the integrals used in the calculation of the exponentially small Stokes multiplier S.

that

$$\sin(2t) = \frac{2i}{\psi}\sqrt{1+\frac{1}{\psi^2}}+\cdots, \tag{7.78}$$

$$\sin^2(t) = -\frac{1}{\psi^2}+\frac{2i}{\psi}\sqrt{1+\frac{1}{\psi^2}}(t-t_p)+\cdots, \tag{7.79}$$

where it is worth noticing that we have expanded $\sin^2(t)$ to the second order. Using these expansions in the definition of p gives that about t_p

$$\begin{aligned} p &= \frac{\psi^2\frac{2i}{\psi}\sqrt{1+\frac{1}{\psi^2}}}{1+\psi^2\left(-\frac{1}{\psi^2}+\frac{2i}{\psi}\sqrt{1+\frac{1}{\psi^2}}(t-t_p)\right)}+\cdots, \\ &= \frac{1}{t-t_p}+\cdots, \end{aligned} \tag{7.80}$$

and hence,

$$\mathrm{Res}_{t_p} p = 1. \tag{7.81}$$

The second residue that we require, $\mathrm{Res}_{t_p} q/\omega$, is calculated in the same way. Introducing the leading order behaviour of (7.79) in the expression for ω, (7.22), and the second order behaviour into the definition of q from (7.24) gives that

$$\begin{aligned}\frac{q}{\omega} &= \frac{\varsigma(e+e^{-1}+2e[2i\psi\sqrt{1+1/\psi^2}(t-t_p)]^{-1}}{1+\mu^{-2}(1+\psi^2[-1/\psi^2])}+\cdots,\\ &= \frac{-i\varsigma}{t-t_p}+\cdots,\end{aligned} \tag{7.82}$$

and so

$$\mathrm{Res}_{\,t_\mathrm{p}} q/\omega = -i\varsigma. \tag{7.83}$$

Using the values of the residues that we have just calculated gives that (7.76) is an identity. Next, by using $\mathrm{Res}_{\,t_\mathrm{p}} p = 1$ in (7.77) we get that the Stokes multiplier is given by

$$S = -ie^{2i\epsilon^{-1}\fint_{-\pi/2}^{t_0}[\omega(t')-\epsilon q(t')/(2\omega(t'))]\,\mathrm{d}t'}. \tag{7.84}$$

where $\fint$ is the Cauchy principal value. Since $q(t)$ has a pole at $t = t_\mathrm{p}$, to get the above expression we have used the fact that the integral can be written as the integral up to a distance of μ, where $\mu \to 0$, either side of the pole plus $i\pi$ times the residue at the pole. This technique can be seen in the following example where there is clearly a pole at $z = z_0$,

$$\int_{-\infty}^{\infty}\frac{g(z)}{z-z_0}\,\mathrm{d}z = \lim_{\mu\to 0}\left[\int_{-\infty}^{z_0-\mu}\frac{g(z)}{z-z_0}\,\mathrm{d}z+\int_{z_0+\mu}^{\infty}\frac{g(z)}{z-z_0}\,\mathrm{d}z\right]+i\pi\mathrm{Res}_{\,z_0} g(z) \tag{7.85}$$

$$= \fint_{-\infty}^{\infty}\frac{g(z)}{z-z_0}\,\mathrm{d}z+i\pi\mathrm{Res}_{\,z_0} g(z). \tag{7.86}$$

It follows that $|S|$, giving the instability growth rate, can be written as

$$|S| = e^{-\alpha/\epsilon+\varsigma\beta}, \tag{7.87}$$

where the two constants, α and β, given by

$$\alpha = -2i\int_0^{t_0}\omega(t)\,\mathrm{d}t \quad \text{and} \quad \beta = -i\fint_0^{t_0}\frac{\varsigma q(t)}{\omega(t)}\,\mathrm{d}t, \tag{7.88}$$

are real, positive and independent of ς. By using (7.22) and (7.23)-(7.25), these expressions can be given the more explicit forms

$$\alpha = \frac{2}{\mu}\int_0^{\sinh^{-1}(\sqrt{1+\mu^2}/\psi)}\sqrt{1-\psi^2\sinh^2 u+\mu^2}\,\mathrm{d}u, \tag{7.89}$$

and

$$\beta = \mu \varsigma \fint_0^{\sinh^{-1}(\sqrt{1+\mu^2}/\psi)} \left(e + e^{-1} + \frac{2e}{1 - \psi^2 \sinh^2 u} \right) \frac{\mathrm{d}u}{\sqrt{1 - \psi^2 \sinh^2 u + \mu^2}}. \quad (7.90)$$

These expressions can be transformed further to make them more convenient by letting $t = i \sin^{-1}(x)$ which gives that

$$\alpha = \frac{2}{\mu} \int_0^{\sqrt{1+\mu^2}/\psi} \sqrt{\frac{1 + \mu^2 - \psi^2 x^2}{1 + x^2}} \, \mathrm{d}x, \quad (7.91)$$

and

$$\beta = \mu \fint_0^{\sqrt{1+\mu^2}/\psi} \left(e + e^{-1} + \frac{2e}{1 - \psi^2 x^2} \right) \frac{\mathrm{d}x}{\sqrt{(1 + \mu^2 - \psi^2 x^2)(1 + x^2)}}. \quad (7.92)$$

Here $\fint$ denotes the Cauchy principal value of the integral, whose integrand is singular at $x = 1/\psi$.

By combing the expression for the growth rate σ, given in (7.59), along with the expression for $|S|$, given in (7.87), and its relevant constants (7.91) and (7.92), we now have a way to compute the maximum growth rate.

7.7 Analysis of the α and β integrals

Now that we have found the growth rate in terms of the two integral expressions for α and β, it will be insightful to draw some conclusions about their behaviour for large and small values of their defining parameters. Figure 7.3 shows the values of α and β as functions of e and μ which gives us an idea of how they behave. We can also gain more information about their behaviour by examining their explicit expressions (7.91) and (7.92).

7.7.1 The asymptotics of α for small and large values of μ

By examining figure 7.3 the first conclusion that we can draw is that $\alpha \to \infty$ in the limit of small μ. Specifically we can see that $\alpha = O(\mu^{-1})$ as $\mu \to 0$ since in this situation the integral for α reduces to

$$\alpha \sim \frac{2}{\mu} \int_0^{1/\psi} \sqrt{\frac{1 - \psi^2 x^2}{1 + x^2}} \, \mathrm{d}x. \quad (7.93)$$

The biggest contribution to this integral comes from the lower end of the integration range, i.e. when x is near zero which gives that $\alpha = O(\mu^{-1})$ as $\mu \to 0$.

The figure of the contours of α also suggests that $\alpha \to \infty$ as $\mu \to \infty$. From (7.91) the main contribution to this integral, when $\mu \to \infty$, comes from some intermediate part of the integration range and so to evaluate the integral we split it into two parts, [25]. The first part is between 0 and λ and the second part is between λ and $\sqrt{1+\mu^2}/\psi$, where we define λ such that $\mu \gg \lambda \gg 1$. The next step is to notice that in the first integral x^2 is always really small when compared to μ. If we scale x in the second integral by letting $x = X/\psi$ then we get that

$$\alpha \sim \frac{2}{\mu}\left(\sqrt{1+\mu^2}\int_0^{\lambda}\frac{\mathrm{d}x}{\sqrt{1+x^2}} + \int_{\lambda\psi}^{\sqrt{1+\mu^2}}\frac{\sqrt{1+\mu^2-X^2}}{X}\,\mathrm{d}X\right), \tag{7.94}$$

where $\sqrt{\psi^2+X^2}$ has been replaced with X in the second integral since $X \gg \psi$. Evaluating this expression gives that

$$\begin{aligned}\alpha &\sim \frac{2\sqrt{1+\mu^2}}{\mu}\left[\sinh^{-1}(\lambda) - \log\left(2\sqrt{1+\mu^2}\right) - \sqrt{\frac{1+\mu^2-\lambda^2\psi^2}{1+\mu^2}}\right.\\ &\qquad\left. + \log\left(2\frac{1+\mu^2+\sqrt{1+\mu^2}\sqrt{1+\mu^2-\lambda^2\psi^2}}{\lambda\psi}\right)\right] \qquad (7.95)\\ &\sim -\frac{2\sqrt{1+\mu^2}}{\mu}\left[\log\left(1+\mu^2\right) + \sqrt{\frac{1+\mu^2-\lambda^2\psi^2}{1+\mu^2}} + \log(\psi)\right.\\ &\qquad\left. - \log\left(1+\mu^2+\sqrt{1+\mu^2}\sqrt{1+\mu^2-\lambda^2\psi^2}\right)\right]. \qquad (7.96)\end{aligned}$$

Taking the limit of this expression as $\mu \to \infty$ gives that

$$\alpha \sim 2\log(\mu) + 2\log(2) - 2 - 2\log(\psi), \tag{7.97}$$

as $\mu \to \infty$.

This suggests, and is confirmed by Figure 7.3, that the largest growth rates are attained for $\mu = O(1)$. Thus from (7.18) the perturbations that grow as a result of the elliptical instability of vortices should be expected to have an aspect ratio of $m_0/k_0 = O(N/f)$ which is the Prandtl ratio.

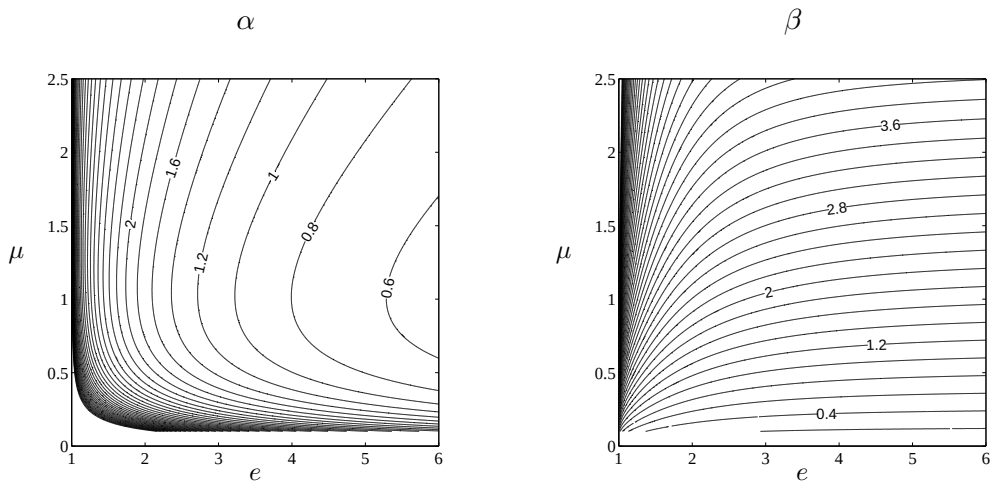

Figure 7.3: Contours of the parameters α and β governing the maximum growth rates according to (7.59)-(7.87) as functions of e and μ.

7.7.2 The asymptotics of α for small and large values of ψ

It is also of interest to look at how α behaves for small and large values of the eccentricity. The analysis of the integral definition of α, given in (7.91), for $\psi \to 0$ follows the same lines as the analysis for $\mu \to \infty$ as far as the expression found in (7.96), the only difference being that this time we have defined λ so that $1/\psi \gg \lambda \gg 1$. Taking the limit as $\psi \to 0$ now gives that

$$\alpha \sim -\frac{2\sqrt{1+\mu^2}}{\mu}\left(\log\psi + 1 - 2\log 2 - \frac{1}{2}\log(1+\mu^2)\right). \tag{7.98}$$

Secondly, we are interested in the behaviour of α as $\psi \to \infty$. As $\psi \to \infty$ it is clear that the upper limit of the integral for α tends to zero which suggests using the scaling $\psi x = X$ which gives that

$$\alpha \sim \frac{2}{\mu\psi}\int_0^{\sqrt{1+\mu^2}} \sqrt{1+\mu^2-X^2}\ \mathrm{d}X, \tag{7.99}$$

where we have taken $\sqrt{1+X^2/\psi^2} \approx 1$. To take this further we now use the scaling $X = \sqrt{1+\mu^2}y$ which, after simplifying, gives that

$$\alpha \sim \frac{2}{\mu\psi}(1+\mu^2)\int_0^1 \sqrt{1-y^2}\ \mathrm{d}y, \tag{7.100}$$

and hence

$$\alpha \sim \frac{(1+\mu^2)\pi}{2\mu\psi} \text{ as } \psi \to \infty. \tag{7.101}$$

This expression for large ψ can actually be used to estimate α for values of ψ as small as 1, which makes it very useful. If we take $\mu = 1$, for instance, the error in (7.101) is 15%, 10% and 5% for $\psi = 1$, 1.5 and 2 respectively, which shows how accurate this approximation is. This expression also shows that the largest growth rates are attained precisely for $\mu \sim 1$ and when ψ is large.

7.7.3 The effect of β

From (7.92) it is clear that $\beta > 0$, which when combined with the expression for $|S|$ shows that anticyclonic flows ($\varsigma = 1$) are more unstable than cyclonic flows ($\varsigma = -1$). According to (7.87), the growth rate in an anticyclonic flow is a factor $\exp(2\beta)$ larger than the growth rate of the corresponding cyclonic flow. Formally, this is an $O(1)$ factor, but the typical values of β are such that it is numerically very small which means that the instability of cyclones is exceedingly weak and probably negligible in most circumstances. Note that because β is a decreasing function of e, which is clear from figure 7.3, the asymmetry between cyclones and anticyclones is the largest for small eccentricity.

7.8 Position and thickness of the instability bands

The formulae (7.59), (7.87), (7.91) and (7.92) for σ_{max}, $|S|$, α and β respectively, give complete and explicit expressions for the maximum growth rates of the elliptical instabilities in terms of the three parameters ϵ, μ and e (recall that $\psi = \sqrt{e^2-1}$). From earlier we know that these growth rates are achieved when the three parameters are related in such a way that $\exp(i\theta/\epsilon) = \pm 1$, that is,

$$\theta = n\pi\epsilon, \quad n = 1, 2, \cdots . \tag{7.102}$$

This condition can be recognised as a resonance condition between the phase of the inertia-gravity oscillations and the period of rotation around the elliptical vortex which is 2π in the dimensionless time used here.

The growth rates can be written more directly in terms of ϵ, μ and e by solving

(7.102) perturbatively, with $\theta = \theta_0 + \epsilon\theta_1 + \cdots$, where θ_0 and θ_1 are obtained from (7.37) and (7.39). This gives the approximate position of the instability bands as well as their width. To leading order, the instability bands are centred at values of e and μ satisfying

$$\begin{aligned} \theta_0 &= \frac{1}{\mu}\int_{-\pi/2}^{\pi/2}\sqrt{1+\mu^2+\psi^2\sin^2 t}\,\mathrm{d}t \\ &= \frac{2}{\mu}\int_0^1 \sqrt{\frac{1+\mu^2+\psi^2x^2}{1-x^2}}\,\mathrm{d}x = n\pi\epsilon, \quad n=1,2,\cdots . \end{aligned} \tag{7.103}$$

The computation of the second order correction, $\epsilon\theta_1$, is more involved but it is worth noting that it is, in principle, needed to obtain a leading-order approximation to the growth rate $\mathrm{Re}\,\sigma$ as a function of e and μ. This is because the error in α needs to be $o(\epsilon)$, which requires us to approximate the resonance values of e and μ with an $o(\epsilon)$ error too. We do not pursue these detailed computations here.

Since the values of e and μ satisfying the resonance condition, (7.102), are ϵ-close together, the expresion for σ_{max}, (7.59), provides a useful approximation to the growth rates of the instability without the need to locate the resonances accurately. This is demonstrated in the next section where we compare the prediction (7.59) with numerical solutions of the Floquet problem associated with (7.19).

Note that the band width can be deduced directly from the expression (7.103) for θ_0. For fixed ϵ and e, for instance, T in (7.56) can be written as $T = \epsilon^{-1}\Delta\mu\,\partial_\mu\theta$, where $\Delta\mu$ is the distance between μ and the resonant values, and the derivative is evaluated at resonance. According to (7.58), the instability-band width is therefore $\Delta\mu = 2\epsilon|S|/\partial_\mu\theta$ where θ can be approximated by θ_0.

7.9 Comparison with numerical results

The Floquet problem associated with equation (7.21) for the amplitude ζ was solved numerically using Matlab's standard Runge-Kutta solver. The growth rates $\mathrm{Re}\,\sigma$ obtained in this manner are compared with the asymptotic estimate for $\sigma_{\max}$ that is found in (7.59). To emphasise the exponential dependence on the inverse Rossby number $1/\epsilon$, it is convenient to display $\mathrm{Re}\,\sigma$ as a function of $1/\epsilon$ for fixed values of μ and of e. Figure 7.4, panels (a) and (b), summarise the results obtained for several anticyclonic vortices (with $\varsigma = 1$) and cyclonic vortices ($\varsigma = -1$) respectively. These figures display in linear-logarithmic coordinates the local maxima of the growth rates

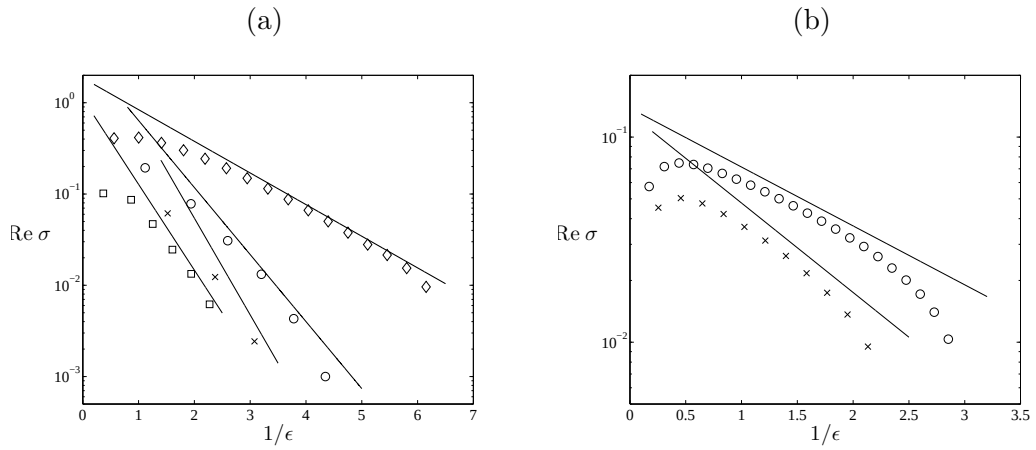

Figure 7.4: Numerical estimates of the local maxima of the growth rates $\mathrm{Re}\,\sigma$ as functions of the inverse Rossby number $1/\epsilon$. (a) Anticyclonic flows with $e = 1.5$, $\mu = 1$ ($\times$); $e = 2$, $\mu = 1$ ($\circ$); $e = 2$, $\mu = 0.5$ ($\square$) and $e = 4$, $\mu = 1$ ($\Diamond$). (b) Cyclonic flows with $e = 4$, $\mu = 0.5$ ($\times$) and $e = 6$, $\mu = 0.5$ ($\circ$). The straight lines in these linear-logarithmic plots indicate the exponential dependence predicted by the asymptotic estimate (7.59).

obtained numerically as ϵ is varied as well as showing the corresponding asymptotic predictions, (7.59), which are represented as line segments. The maximum growth rates displayed correspond to the peaks of the instability bands which occur for resonant values of $1/\epsilon$ for fixed e and μ. Figures 7.5 and 7.6 illustrate the complete structure of the instability bands, in linear coordinates, that surround the resonant values by showing all the nonzero growth rates obtained numerically. The parameters used in these figures correspond to the parameters that were used in figure 7.4, panels (a) and (b).

The figures confirm the validity of our asymptotic estimate. They also suggest that this estimate remains useful for moderately small values of ϵ, say $\epsilon \lesssim 1/2$. It is worth noting that the dimensional growth rates are obtained by multiplying σ by$\sqrt{ab}$ which is related to the relative vorticity, $\Omega = a+b$, of the flow by $\sqrt{ab} = \Omega/(e+e^{-1})$. As expected from our asymptotics, the growth rates in the case of cyclonic flows are exceedingly small for $\epsilon \ll 1$, even for the large eccentricities used in Figure 7.6. Nonetheless, our results clarify the fact that all elliptical flows are unstable, regardless of the sense of the rotation, the strength of the flow and the strength of the stratification. Note that the match between asymptotic and numerical results for cyclonic flows appears to degrade for small ϵ, that is large $1/\epsilon$. This is because the smallness of both the growth rate and the instability bands' width makes the maximum growth rate delicate to estimate

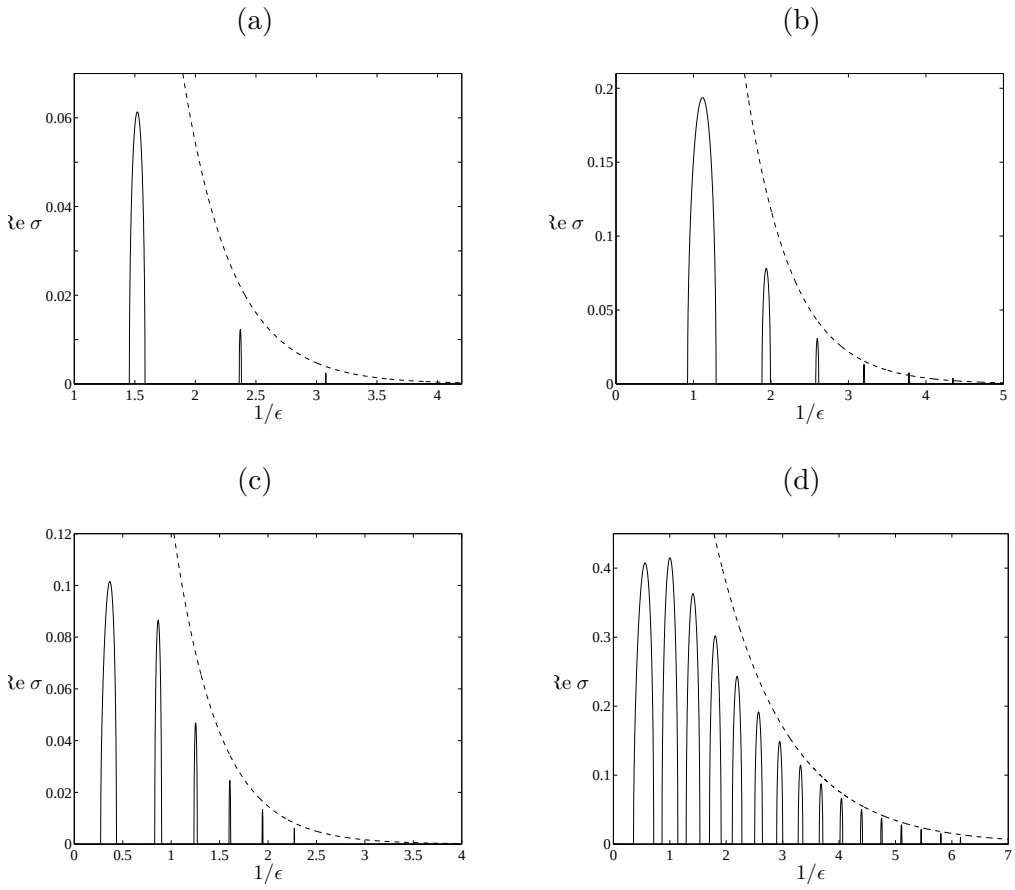

Figure 7.5: Growth rates $\mathrm{Re}\,\sigma$ in anticyclonic flows as functions of the inverse Rossby number, $1/\epsilon$, for (a) $e = 1.5$, $\mu = 1$; (b) $e = 2$, $\mu = 1$; (c) $e = 2$, $\mu = 0.5$ and (d) $e = 4$, $\mu = 1$. The growth rates computed numerically (solid lines) are compared with the asymptotic estimate of the maximum growth rates $\sigma_{\max}$ (dashed lines).

numerically. In particular, the agreement between the asymptotic and numerical results could be improved by using a finer resolution in $1/\epsilon$ but this would need a large increase in computing power.

The separation between instability bands can be estimated from the asymptotic formula (7.103). In terms of the varying $1/\epsilon$ used in the figures, it is given by

$$\gamma = \frac{\pi\mu}{2\displaystyle\int_0^1 \sqrt{\frac{1+\mu^2+\psi^2x^2}{1-x^2}}\,\mathrm{d}x}. \tag{7.104}$$

Evaluating this quantity for the parameters chosen for the figures gives $\gamma = 0.62$, 0.54, 0.31 and 0.34 for the figure 7.5, panels (a)-(d), and $\gamma = 0.18$ and 0.12 for figure 7.6, panels (a) and (b). By examining the relevant graphs it can be seen that these results are in good agreement with the numerical results.

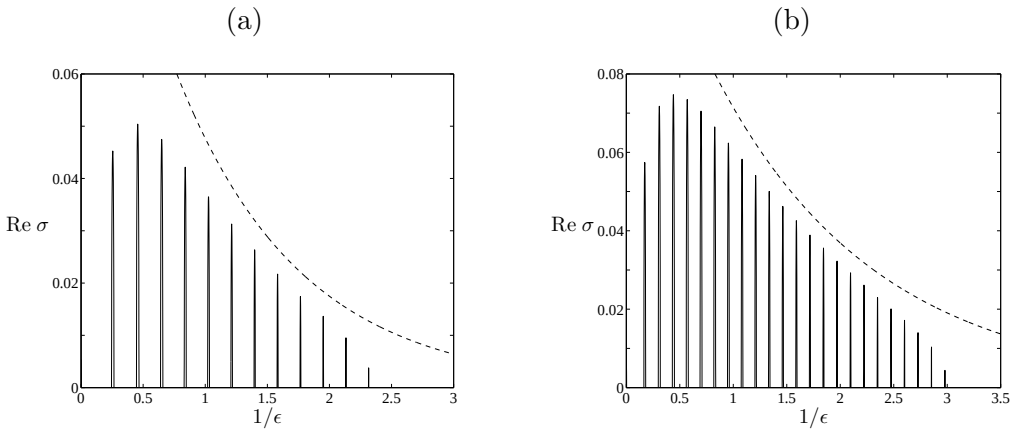

Figure 7.6: Growth rates $\mathrm{Re}\,\sigma$ in cyclonic flows as functions of the inverse Rossby number, $1/\epsilon$, for (a) $e = 4$, $\mu = 0.5$ and (b) $e = 6$, $\mu = 0.5$. The growth rates computed numerically (solid lines) are compared with the asymptotic estimate of the maximum growth rates $\sigma_{\max}$ (dashed lines).

7.10 Justifying the hydrostatic approximation

The derivation of the asymptotic expression for the growth rate made use of the hydrostatic approximation, which assumes that $N \gg f$, $m_0 \gg k_0$ and $\mu = fm_0/(Nk_0) = O(1)$. This assumption, which could be relaxed, is made because it corresponds to the regime most relevant to atmospheric and oceanic flows. It is also consistent with the results in the sense that the growth rates obtained are maximised for $\mu = O(1)$ and decay rapidly for $\mu \gg 1$ or $\mu \ll 1$. To test the sensitivity of the results to the hydrostatic approximation, we can solve the Floquet problem numerically for the full, unapproximated equation, (7.17), for moderately large values of N/f and m_0/k_0. The results obtained for $\mu = 1$ and $e = 4$ are displayed in figure 7.7. This compares the growth rates obtained in the hydrostatic approximation with those obtained for $N/f = m_0/k_0 = 3$ and 6. Except for $\epsilon \gtrsim 1$, there is relatively little difference between the results, that is to say, the maximum growth rates fall on the same curve that is well described by the hydrostatic asymptotics. Of course, the location of the instability bands change depending on N/f, but this is not significant, since they would also change if m_0/k_0 was varied independently of N/f as is physically relevant.

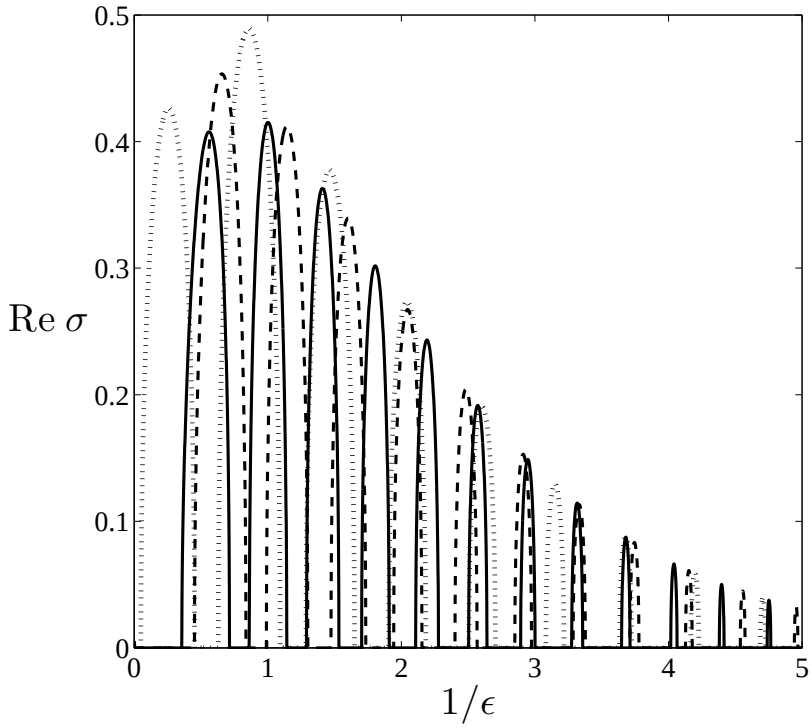

Figure 7.7: Effect of the hydrostatic approximation: the growth rate $\mathrm{Re}\,\sigma$ is plotted as a function of the inverse Rossby number, $1/\epsilon$, for an anticyclonic flow with $e = 4$ and $\mu = 1$, in the hydrostatic limit $N/f \to \infty$ (solid lines), for $N/f = 6$ (dashed lines) and for $N/f = 3$ (dotted lines).

7.11 Conclusion

In this chapter we have shown that all elliptical flows are inherently unstable regardless of the flow's sense of rotation, its strength, or the strength of its stratification. This was done by deriving an asymptotic formula that predicts the maximum growth rates of these elliptical flows in terms of the flow's eccentricity and the aspect ratio of its perturbation. We then showed that these results matched the numerical results that were found by solving the problem numerically.

The asymptotic formula that we have derived also makes it possible to quantify the difference in the strengths of the instabilities that occur between anticyclonic and cyclonic flows. It has already been recognised that cyclones are less unstable than anticyclones, to the extent that McWilliams and Yavneh, [40], considered only the instability of the latter. We have shown that cyclones are in fact linearly unstable, with growth rates that have the same exponential dependence on the Rossby number as the corresponding anticyclones but differ by a factor which, although formally of

order one, turns out to be numerically very small.

An important point to note, as mentioned in the introduction, is that these results apply to a far wider range of problems than just the purely elliptical system. These results can be applied to any flow that has, or develops, elliptical streamlines at any point in its evolution. Hence there are many areas in the field of geophysical fluid dynamics that can use the results that have been attained in this chapter.

The next step is to take the results and intuition we have gained by studying fairly simple and well defined flows and try to apply them to the study of a complex flow more like that found in the earth's atmosphere and oceans.

Chapter 8

Baroclinic lifecycle

8.1 Introduction

The flows we have studied so far, while increasing in complexity, have all been idealised to a large degree. The aim of this chapter is to study a more realistic flow, namely an idealised baroclinic life cycle, that is, the formation, evolution and dissipation of a baroclinic instability which is a process that generates midlatitude large-scale perturbations. This flow is realistic enough to allow comparisons with atmospheric observations and yet simple enough that the initial conditions can be controlled and moist processes can be excluded. Previous numerical studies of inertia-gravity wave generation in baroclinic life cycles, [47, 73], have agreed with the atmospheric observations, [21, 54], that inertia-gravity wave generation occurs in exit regions of jet streaks, that is, fast flowing, relatively narrow air currents found in the atmosphere. In this chapter, data that has previously been obtained from the Weather Research and Forecast Model (WRF) is used to provide a jet streak that occurs as the result of baroclinic instability. The simulations were carried out by Plougonven and Snyder for a study of inertia-gravity wave generation by baroclinically unstable flows, [51, 52]. By using this flow we are able to compare the direct results they found, namely, finding the areas of inertia-gravity wave generation and the directions of the wavevectors that produce the largest growth, with predictions that our WKB approach can provide. We also link the growth of the inertia-gravity waves to the Lagrangian transience of the flow.

To achieve this, we first change the coordinate system that the data, provided by

[51, 52], from the WRF model is in from a system in terms of the geopotential and temperature to the Boussinesq system set out in section 2.3.3. We then smooth the data to remove any inertia-gravity waves that are in the flow in the form of small scale, large amplitude variations and interpolate the data to the position of the wavepacket. We can then solve the equations for the evolution of a wavepacket, as found in section 4.2.9, in an idealised baroclinic lifecycle, and analyse the results.

8.2 Baroclinic instability

The flow that we study in this chapter, an idealised baroclinic lifecycle, is created by baroclinic instabilities, [61]. These instabilities are a form of hydrodynamic instability, which occur in stably stratified, rapidly rotating fluids, of which the earth's atmosphere and oceans are prime examples. Baroclinic instabilities are a type of hydrodynamic instability that affect the weather the most since they are the main creation mechanism for cyclones and anticyclones, which dominate the weather patterns of the earth. These instabilities also occur in the oceans and on other planets.

To understand the processes involved in creating baroclinic instabilities, we consider the scenario that occurs as a result of the temperature difference between the equator and the poles. This results in domes of cold air forming over the poles while warm air gathers at the equator. This creates a temperature gradient between the poles and the equator and, since cold air is denser than warm air, it also creates a density gradient. As a consequence of this gradient the density increases as the distance to the poles decreases, along with the density decreasing with altitude as usual. As a result, the contours of constant density angle towards the poles, as shown in figure 8.1, rather than lying horizontally, as they would if this temperature gradient did not exist.

In the scenario set out in figure 8.1, we may intuitively expect that the domes of cold air at the poles will settle and spread out radially to lower altitudes under the force of gravity. This intuition holds true in the absence of rotation, but since the earth is rotating, the Coriolis force, see section 2.2.1, deflects the flow of settling fluid to the right in the northern hemisphere and to the left in the the southern hemisphere. Hence the Coriolis force acts as an opposing force to the settling fluid, creating a vertically sheared flow that circulates the dome of cold air. This flow, called a thermal wind, supports the dome of cold air so that it does not settle. An expression for the thermal wind can be found by combining the hydrostatic and geostrophic balance relations, see

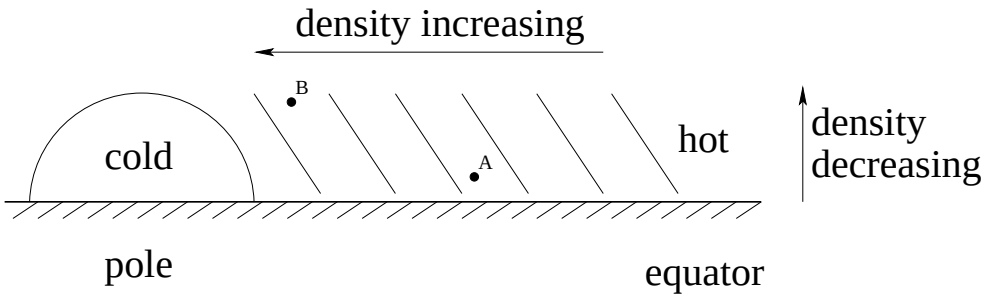

Figure 8.1: The temperature gradient between the poles and the equator can cause domes of cold air to form over the poles. This creates a density gradient which leads to the contours of constant density, represented by the solid lines, tilting towards the poles rather than lying horizontally.

section 2.4, to get

$$\frac{\partial \mathbf{u}}{\partial z} = \frac{g}{\rho_b f}\left(\frac{\partial \rho}{\partial y}, -\frac{\partial \rho}{\partial x}\right). \tag{8.1}$$

The higher the density of the cold air, the higher the speed the thermal wind must be to support it and stop it from settling. When the mass of air gets significantly dense, the thermal winds can become baroclinically unstable, allowing small perturbations to grow with time.

A simple model of this happening can be created by placing two fluid particles, call them A and B with B heavier than A, as shown in figure 8.1. If the position of these particles is now inverted through a perturbation, particle A becomes surrounded by denser fluid and so is buoyant, and particle B becomes surrounded by lighter particles and so is negatively buoyant. The net effect is that the centre of gravity of the fluid has been lowered, therefore its overall potential energy has also been lowered. The potential energy that the fluid has lost is gained as kinetic energy by the perturbation. This energy gain will amplify the perturbation causing it to grow. This is baroclinic instability.

Spiralling eddies are produced by these opposing forces. In the northern hemisphere, these occur as the cold air that is trying to spread outwards is deflected clockwise, while the warm air that is trying to converge towards the pole, which is the rotation axis, is deflected anticlockwise. Part of the gravitational potential energy that is stored in the dome of cold air is converted into kinetic energy in the eddies as they are formed, and is removed from the system when they break away. The energy that these eddies take with them is eventually totally dissipated and the system settles into a lower energy state. The kinetic energy of these eddies is felt in the form of strong winds that are

often present during storms [43].

In this chapter we study the potential of baroclinic lifecycles for inertia-gravity wave generation. In this case the energy of the system is transferred from the large scale baroclinic modes to the large scale barotropic modes. These are modes in which the pressure is a function of the density only. As the energy goes between these two modes it first passes through lower scale baroclinic modes and then by baroclinic instabilities, to barotropic modes of a similar scale. Finally the energy is transferred to larger barotropic scales. This whole process of baroclinic instabilities and energy transfer can be thought of as an inverse cascade in which energy passes from large wavenumbers to smaller wavenumbers.

8.3 Model setup

The data used for this study was provided by R. Plougonven and is the same data that he used along with C. Snyder, [51, 52], for their study of inertia-gravity wave generation. Here we give an outline of how this data was originally generated.

The data for the simulated baroclinic lifecycle that we use is generated by the Weather Research and Forecast Model (WRF), [57]. This model solves the compressible, nonhydrostatic fluid equations to simulate the evolution of the baroclinic instability of a jet on the f-plane. This jet is comparable to the thermal wind that we described above. The flow is confined to a channel that is periodic and 4000 km long in the x direction, with 10 000 km between the walls in the y direction and a depth of 20 km.

Since the aim of the initial study was to study inertia-gravity wave generation, the flow was designed to be as free from them as possible. This was achieved by initialising the flow so that it consists of the superposition of a geostrophically balanced zonal jet and its most unstable normal mode, with a small amplitude and a zonal wavelength of 4000 km. The jet is created in a way similar to the method employed by Rotunno, [55], which involves inverting a two-dimensional potential vorticity distribution in the y-z plane. The potential vorticity is distributed so that a sharp tropopause separates the troposphere, with a uniform potential vorticity of 0.4 PV units, and the stratosphere, with a uniform potential vorticity of 4 PV units. The most unstable mode of the jet was calculated by evolving a small perturbation to the jet over 4 days, then reducing its amplitude and repeating four times so that the total integration time was 16 days.

After this time the normal mode was extracted and rescaled before being superimposed on top of the jet.

8.4 Modifying the data

The WRF model solves the compressible fluid equations and so the data it produces is in a coordinate system based on the geopotential and the temperature, the system that was mentioned at the end of section 2.3.3. Hence to be able to use the data with our model, the Boussinesq model which uses density and pressure, we first need to modify it. The details of this change of coordinates can be found in Appendix A. As laid out in that discussion, we need to use $-\theta$ as our ρ, where θ is decomposed as

$$\theta_{\text{total}} = \theta_b + \bar{\theta}(z) + \theta'(x, y, z, t). \tag{8.2}$$

By comparing this to our decomposition of ρ, (2.26), it is clear that we need to take

$$\rho_b = \theta_b \qquad \text{and} \qquad R = -\frac{g}{\rho_b}(\bar{\theta} + \theta'). \tag{8.3}$$

The horizontal gradients of R are easily derived, since $\bar{\theta}$ only depends on z, but this dependence means that the vertical gradient is given by

$$\frac{\partial R}{\partial z} = -\frac{g}{\rho_b}\left(\frac{\partial \bar{\theta}}{\partial z} + \frac{\partial \theta}{\partial z}\right). \tag{8.4}$$

The data from the WRF model includes the second term in this expression, $\partial_z\theta$, but it does not include the first one, $\partial_z\bar{\theta}$ and so we construct this term from the data for $\bar{\theta}$ and z using a finite difference method. Another quantity that needs to be carefully considered is the potential vorticity. In the WRF simulation the potential vorticity is defined as

$$q_{\text{wrf}} = \frac{1}{\rho}(\nabla\theta \cdot \omega_a), \tag{8.5}$$

where ω_a is the absolute vorticity. If we compare this to the Boussinesq definition of potential vorticity that we have used,

$$q_{\text{wkb}} = \frac{g}{\rho_b}(\nabla\rho_{\text{total}} \cdot \omega_a), \tag{8.6}$$

it is clear that the two are related by

$$q_{\text{wkb}} = \frac{g\rho_b}{\theta_b} q_{\text{wrf}}. \tag{8.7}$$

It is also important to keep in mind when making calculations, that the domain the baroclinic lifecycle is taking place in is periodic in the x direction and so the flow's data fields are also periodic in the x direction. Before we can use the data from the WRF model with the equations that we have derived, we need to set up an interpolation system that can determine all the flow's fields at the position of the wavepacket for any position and time.

8.5 Interpolation

To solve the ordinary differential equations that govern the wavepacket's evolution, we require the values of the flow's data fields wherever the wavepacket may be. The WRF model produces a data file every 6 hours that includes all the flow's data that we need. This data is laid out on a grid which has a horizontal resolution of 50 km and a vertical resolution of 250 m. The first step in interpolating the data to any time and position is to run a linear time interpolation to determine the data at the time needed:

$$U = \frac{(t_1 - t)U_0 + (t - t_0)U_1}{t_1 - t_0}, \tag{8.8}$$

where U_0 and U_1 are the fields being interpolated to create U, the field in question at the time, t. In this expression U_0 and U_1 are the field's values at t_0 and t_1, the times of the closest data files before and after t respectively. After interpolating all the fields to the required time we need to spatially interpolate them to the wavepacket's present location.

Two methods of spatial interpolation are used; the first, while simple and quicker, is rather crude and so a second, more refined one is used.

The first method of spatial interpolation initially determines the closest point in the data grid to the wavepacket's position in three dimensional space and then reads the value of the data at that point from the time interpolated data file. This method is very crude and we can improve on it by using a second method that makes use of cubic-spline interpolations. The first step in this procedure is to find the position of each data point, on the horizontal level below it, in the four by four grid surrounding it,

as shown in figure 8.2. A cubic spline interpolation is used to interpolate these points in the x direction to produce four more points, which we can then interpolate in the y direction to find the point in question. This same procedure is then carried out on the level below this one and the two levels above the point, to produce two interpolated points exactly above and two exactly below the point in question. These four points are then used in a final cubic spline interpolation to find the value of the field at the point in question. When finding the grid of points around the point in question, it is important to remember the domain is periodic and so the points required might be from the other end of the domain.

8.5.1 Smoothing the data fields

Inertia-gravity waves are present in the flow in the form of large amplitude, small scale variations that produce strong gradients in the flow's data fields. Since we are interested in inertia-gravity wave generation we want to eliminate the inertia-gravity waves that are already there while still keeping the overall behaviour of the flow by smoothing the data field. This is done by convolving the data field with a smoothing kernel, that is, a data field that has an impulse like spike. The strength of the smoothing is controlled by a single parameter, γ, that alters the width of the smoothing kernel. When the width is narrow, which corresponds to $\gamma = 2$ ($\gamma = 1$ gives no smoothing), the data field remains relatively unchanged and when the width is large, which corresponds to γ increasing, the data field is more aggressively smoothed. Figure 8.3 shows the data field corresponding to dU/dx for four different values of the smoothing function. It is clear from the figure that as the smoothing increases, the strong gradients associated with the large amplitude, small scale variations start to disappear but the data's main characteristics remain intact.

8.6 Initialisation

Now that we are able to transform the data into the coordinate system that the amplitude equations are based in, and that we are able to interpolate this data to any position in the flow's domain at any time, we can start to study the evolution of a wavepacket in this flow.

This involves a number of steps. Firstly, the initial conditions of the wavepacket must be set, then the correct data files need to be determined from the initial time

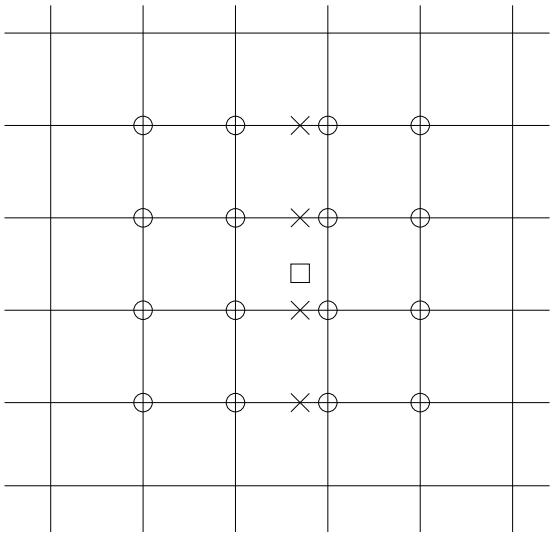

Figure 8.2: A diagram showing how the data is interpolated from a grid to a desired point, □, where the data is known at the grid's intersection points. The first step is to determine the surrounding data points, ∘, of the point in question and then interpolate them in the x direction to obtain the points shown by ×. The data values at these four points can now be interpolated in the y direction to give the data at the desired point, □.

point. Following this, the data from the files is modified to make it usable in our equations and smoothed to remove any small scale, large-amplitude variations. The next step involves interpolating the data in time and space to the wavepacket's current position and solving the equations using this data. This step is then repeated, whilst constantly checking and updating where necessary the data files that are loaded, until the end time of the simulation is reached.

The system is initialised with the Coriolis parameter set as $f = 10^{-4}$ s^{-1}, which is an appropriate value for the atmosphere, and the gravitational constant set as $g = 9.81$ ms^{-2}. The part of the idealised baroclinic lifecycle that is of most interest to us is between day 7 and day 13, the time just after the jet streak has run its course, and so these times are taken as the start and end of the simulation. The wavepacket's altitude needs to be chosen so that it corresponds with an altitude where the vertical velocity is small. This is because the derivation of our equations governing the evolution of a wavepacket relied on the fact that the vertical velocity was zero. From the maximum and minimum vertical velocity profile of the flow, shown in figure 8.4, it is clear that the vertical velocity drops significantly after its peak at $z = 5$ km and is very low, and decreasing further, from $z = 10$ km upwards. Ideally we would then concentrate on trajectories at the top of the domain, but there is very little activity there and so

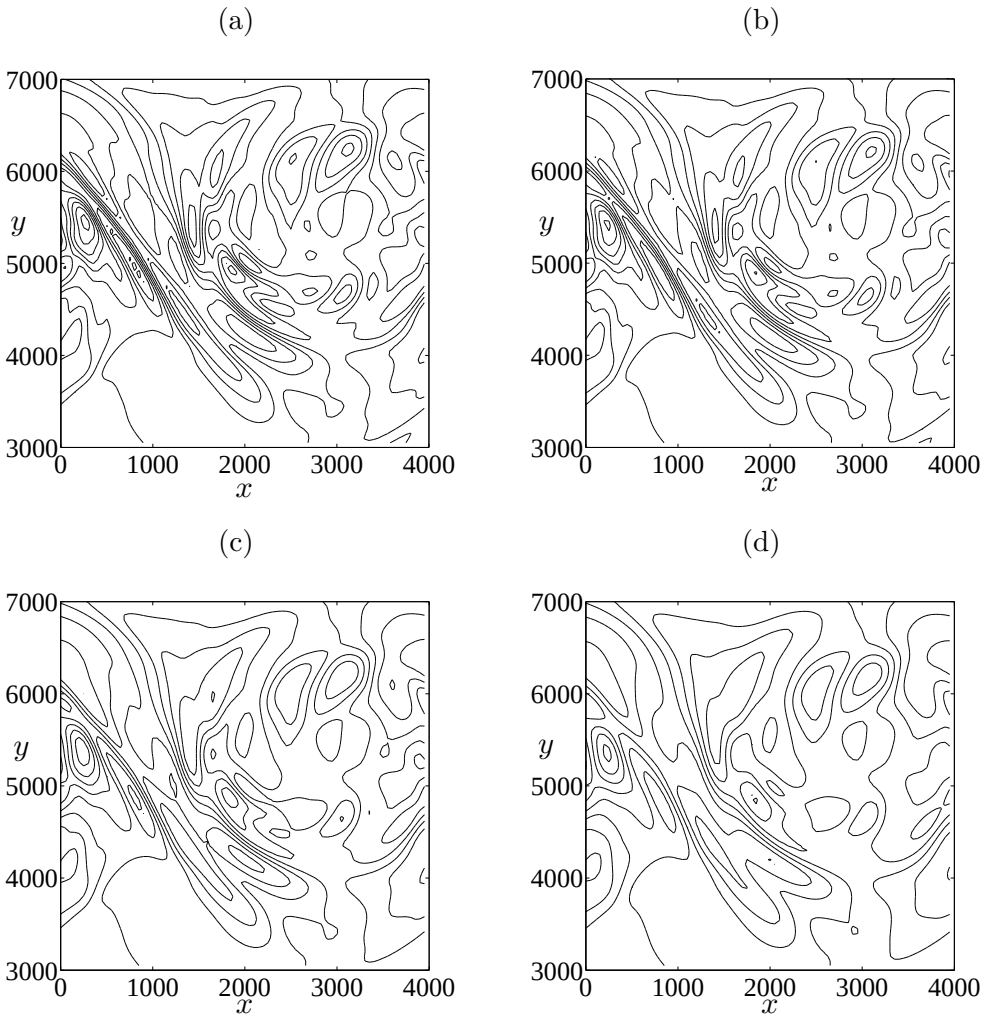

Figure 8.3: Contours of dU/dx with no smoothing, $\gamma = 1$ in panel (a) and the smoothing parameter $\gamma = 4$, (b), $\gamma = 6$, (c) and $\gamma = 8$, (d).

we compromise by concentrating our study on two different altitudes, $z = 10$km and $z = 12$km, which are in the vicinity of the tropopause. The main areas of the flow that are of interest to us are those where the velocity field is constantly changing and where the divergence field of the flow is also varying. These are the areas that were identified as being areas of high inertia-gravity wave activity in [51, 52]. We will consider in some detail one such trajectory that goes through these areas we are interested in, starting at $(3000, 3750)$ km on the horizontal plane. This trajectory can be seen in figures 8.5 and 8.6 for $z = 10$ km and in figures 8.7 and 8.8 for $z = 12$ km. In the case where $z = 10$ km, the remnants of the jet streak can still be seen in the velocity and divergence fields in the first few panels. The flow then proceeds to wrap up cyclonically, creating a high level of activity, that is, rapid changes, in the divergence field in front of it. The case

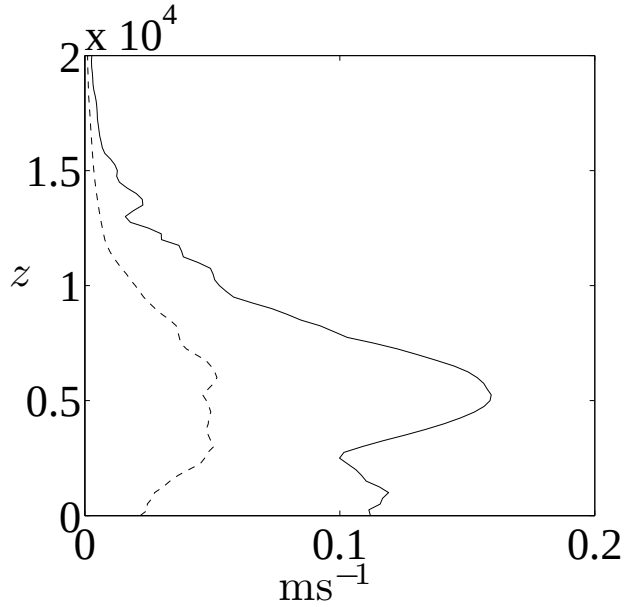

Figure 8.4: The maximum (solid line) and minimum (dashed line) of the vertical velocity against height in metres, for the time period of the simulated baroclinic lifecycle that we are interested in, i.e. from day 7 till the 18th hour of day 13.

where $z = 12$ km is very similar although the activity is slightly weaker. This trajectory is also contrasted with one that starts at $(500, 3000, 10)$ km which stays away from the areas of high activity.

The initial wavenumbers are chosen so that they satisfy the relation

$$\frac{m}{\sqrt{k^2 + l^2}} \approx 4\frac{N}{f}, \tag{8.9}$$

where $N = 10^{-2}$. This relation, when the 4 is not present, is motivated by the fact that the large-scale motion is of this order and also because previous studies, [65, 66], and chapter 7 of this thesis, suggest that this is the right scaling for the fastest growing inertia-gravity waves. The 4 has been added to this relation to ensure that the case where $\alpha = 0$, where α is defined in (4.59), is avoided as this causes problems in the numerics.

To satisfy 8.9 we take $m = \pm 4500$ with $\sqrt{k^2 + l^2} = 10$, where four different configurations of the horizontal wavenumbers are used: $(10, 0)$, $(7, 7)$, $(0, 10)$ and $(-7, 7)$. The vertical wavenumber is initially taken as positive or negative to avoid, if at all possible, it passing through zero. It is worth noting here that we could take any value of m and the corresponding values of k and l to satisfy (8.9) . These

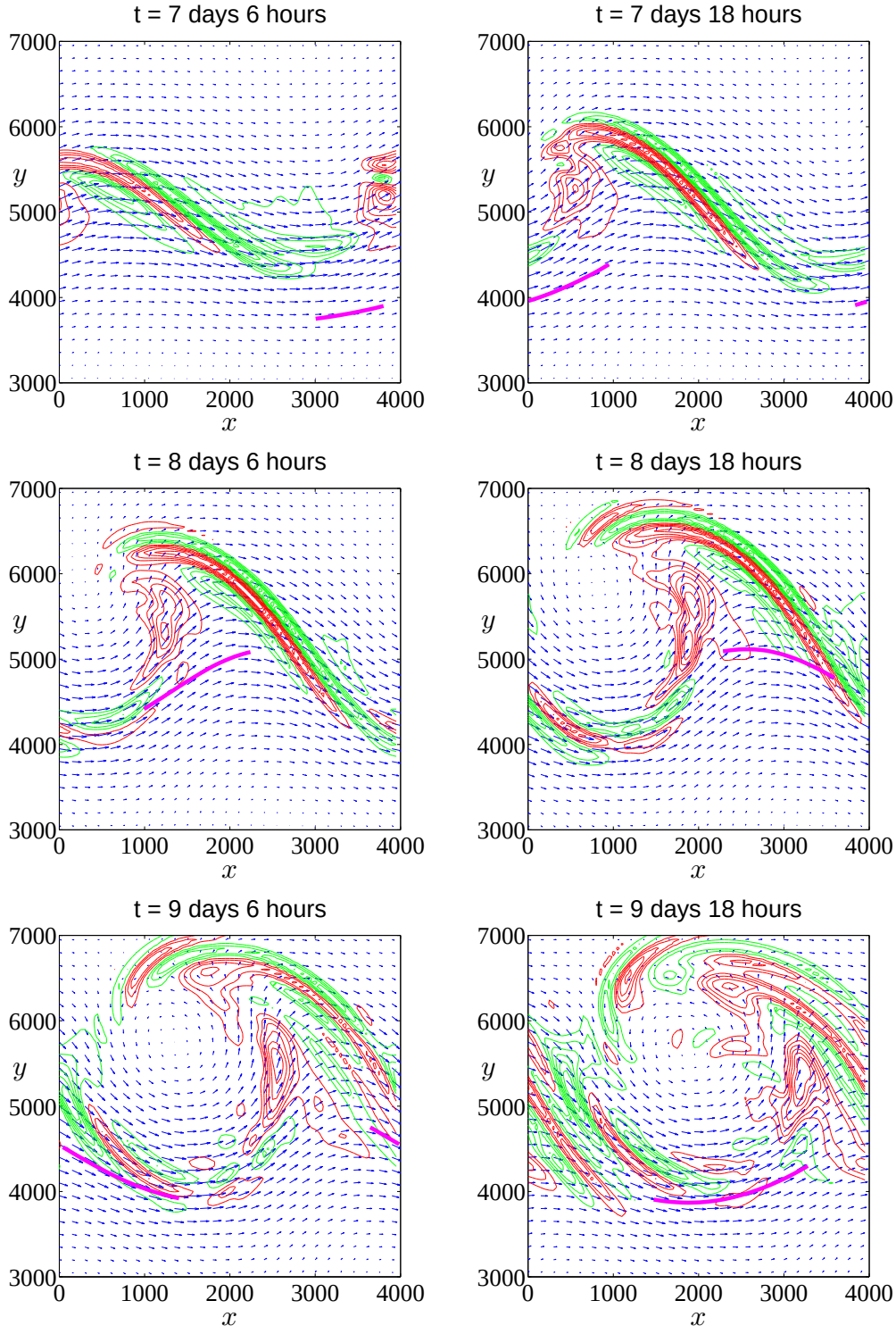

Figure 8.5: The first part of a 3 day, 18 hour simulation of a wavepacket that was initially placed, at the start of day 7, at $(3000, 3750, 10)$ km where x and y are measured in km. Each panel shows how the trajectory, the pink line, of the wavepacket evolves over a 12 hour period. The velocity field, the blue arrows, and the divergence field, red contours for positive divergence and green for negative, are shown for the time that is displayed above each figure, which is the middle of the 12 hour period.

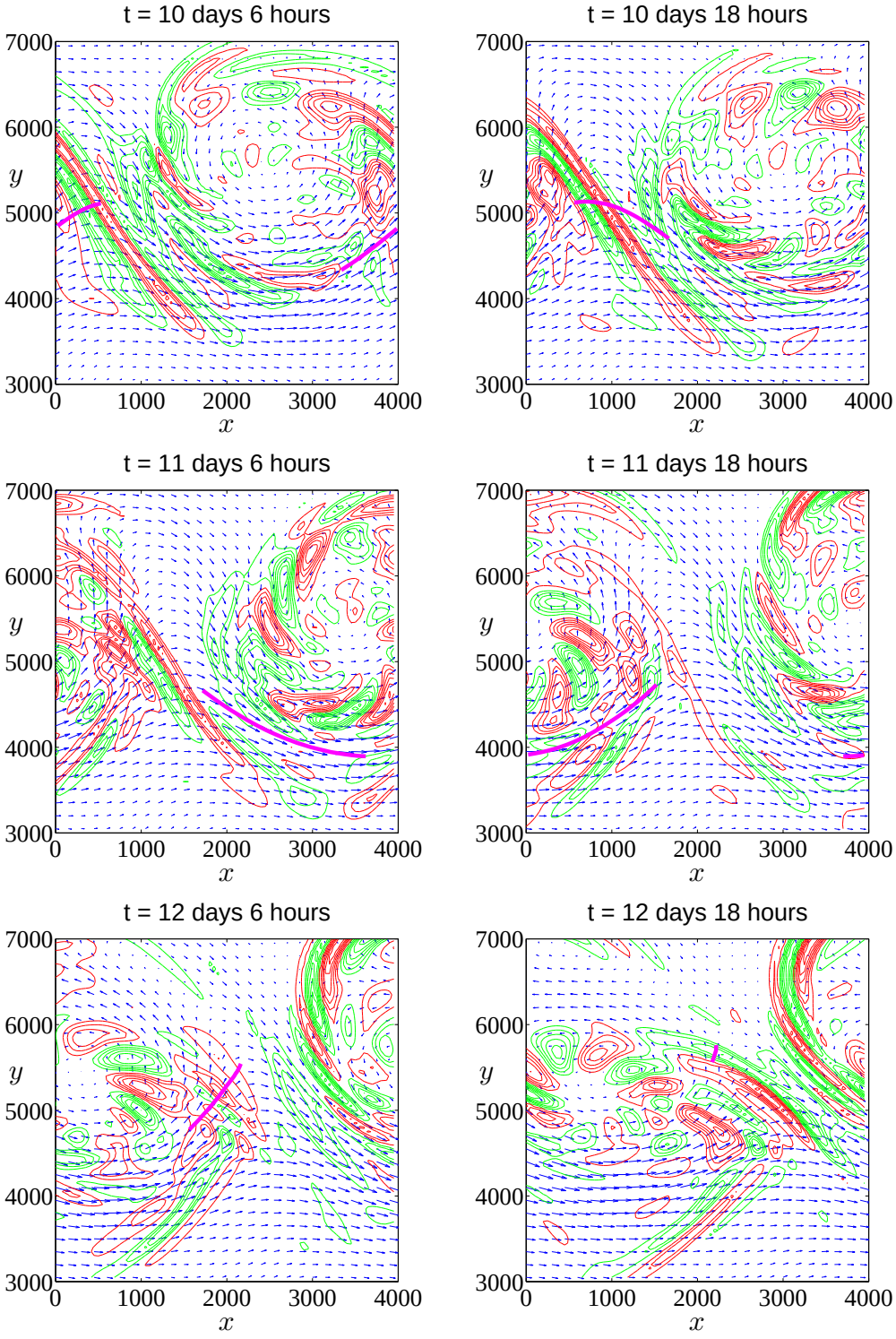

Figure 8.6: The second part of a 3 day, 18 hour simulation of a wavepacket that was initially placed, at the start of day 7, at $(3000, 3750, 10)$ km where x and y are measured in km. Each panel shows how the trajectory, the pink line, of the wavepacket evolves over a 12 hour period. The velocity field, the blue arrows, and the divergence field, red contours for positive divergence and green for negative, are shown for the time that is displayed above each figure, which is the middle of the 12 hour period.

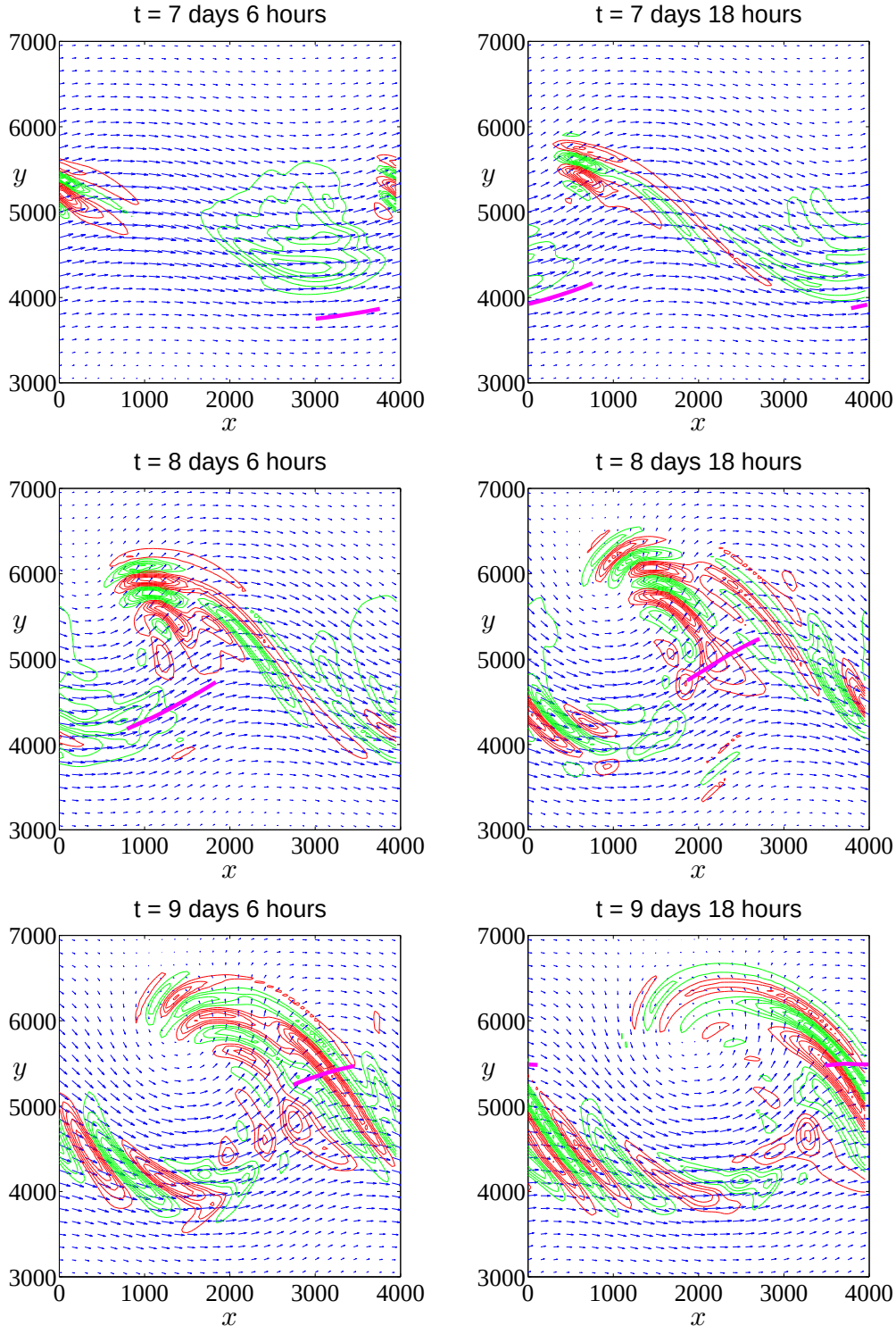

Figure 8.7: The first part of a 3 day, 18 hour simulation of a wavepacket that was initially placed, at the start of day 7, at $(3000, 3750, 12)$ km where x and y are measured in km. Each panel shows how the trajectory, the pink line, of the wavepacket evolves over a 12 hour period. The velocity field, the blue arrows, and the divergence field, red contours for positive divergence and green for negative, are shown for the time that is displayed above each figure, which is the middle of the 12 hour period.

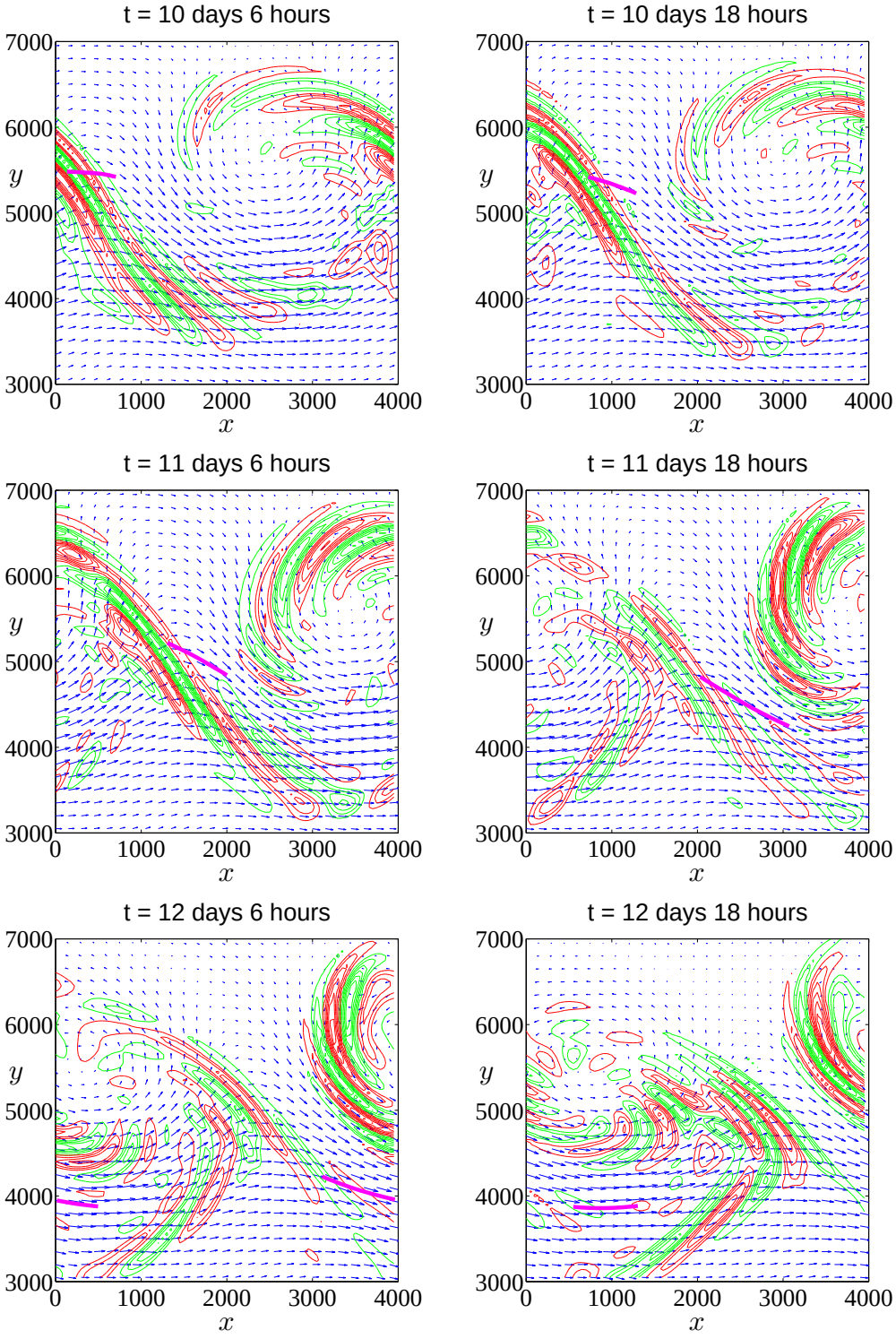

Figure 8.8: The second part of a 3 day, 18 hour simulation of a wavepacket that was initially placed, at the start of day 7, at $(3000, 3750, 10)$ km where x and y are measured in km. Each panel shows how the trajectory, the pink line, of the wavepacket evolves over a 12 hour period. The velocity field, the blue arrows, and the divergence field, red contours for positive divergence and green for negative, are shown for the time that is displayed above each figure, which is the middle of the 12 hour period.

Wavepacket label	Initial position (km)	Initial wavevector	Trajectory shown in figure(s)
A	$(3000, 3750, 10)$	$(10, 0, -4500)$	8.5 & 8.6
B	$(3000, 3750, 10)$	$(7, 7, -4500)$	8.5 & 8.6
C	$(3000, 3750, 10)$	$(0, 10, 4500)$	8.5 & 8.6
D	$(3000, 3750, 10)$	$(-7, 7, 4500)$	8.5 & 8.6
E	$(3000, 3750, 12)$	$(-7, 7, 4500)$	8.7 & 8.8
F	$(500, 8000, 10)$	$(7, 7, -4500)$	8.12

Table 8.1: The initial conditions of the wavepackets whose evolutions are discussed in this chapter.

four are all that are needed to cover the main eight points of a compass, since the negative of each case produces the same evolution of the wavepacket's amplitudes as they only appear as squared terms in the equations. It is worth pointing out that the combination of wavenumbers chosen here is not the only possible combination and that any combination that satisfies (8.9) is perfectly acceptable. The initial conditions for all the trajectories discussed in this chapter can be seen in table 8.1.

Two different states of the wavepacket are used to initialise its amplitudes. As discussed in chapter 6, by picking the initial values of the wavepacket's amplitudes correctly we can start with a vortical mode wavepacket or with just one of its inertia-gravity wave modes excited. Although this decomposition only makes sense in a uniform flow and so we expect the other modes of the wavepacket to be excited almost immediately, it is interesting to compare the difference between the two configurations. The first case is achieved by choosing the initial values of the wavepacket's amplitudes to be proportional to the right eigenvector of the vortical mode in table 6.1 and the second case is achieved by picking them to be proportional to either the gravity$^+$ or gravity$^-$ mode, also shown in table 6.1.

8.7 Results

Now that we have established how the data needs to be modified and have set up a system that we can use it in to solve the equations governing the evolution of a wavepacket, we are in a position to run simulations and study their results. A second order Runge-Kutta algorithm was used to solve the wavepacket's governing equations.

It is worth noting first that all the simulations of wavepackets that we consider below start with the wavepacket being initially balanced, that is, just its vortical mode excited. This is because, after a comparison, the evolution of the wavepacket when it is initially balanced was found to be almost indistinguishable from its evolution when it is not, due to the fact our notion of balance only holds in a uniform flow. When the flow is not uniform, the gravity wave modes of a wavepacket become excited very quickly and so this, along with the fact that the differences between the initial amplitudes in each case is small compared to the values they achieve with even a small amount of growth, means that it is just as if the wavepacket was unbalanced from the start.

By smoothing the data fields of the flow we want to remove the strong gradients associated with inertia-gravity waves that are present while making sure the results are not overly dependent on the strength of the smoothing that we choose. Ideally we would like the smoothing to remove any small-scale transient behaviour that exists and leave behind the dominant balanced behaviour. As a result of this we expect that the amplitudes of the wavepacket may slightly decrease but not by a large amount. Figure 8.9 shows the wavepacket's amplitudes, δ and ζ, for wavepacket C, as defined in table 8.1. This figure shows that while the amplitudes have decreased slightly the overall evolution of the wavepacket is robust to changes in the smoothing. Throughout the rest of this chapter we will set $\gamma = 6$.

Figure 8.10 and the left-hand column of figure 8.11 show the evolution of the wavenumbers, amplitudes and local Rossby number for wavepackets A, B and D, as defined in table 8.1. The first thing to notice is that the wavenumbers appear quite lively which is what would be expected by studying the flow's velocity field. The velocity field is changing frequently and quite rapidly along the wavepacket's trajectory and so it follows that the wavevector must also be constantly changing. If we compare the figures for the evolution of these three wavepackets' wavevectors, we find that although their initial values are different, their evolutions, albeit with slightly different magnitudes, follow a very similar pattern. Note that the wavenumbers' evolution is inverted for the

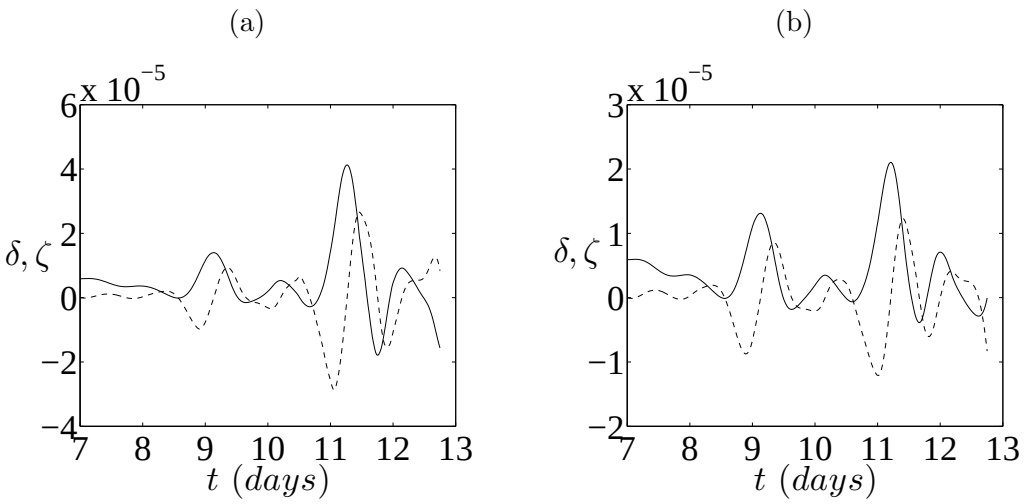

Figure 8.9: The evolution of amplitudes, δ, solid line, and ζ, dashed line, of wavepacket C, as defined in table 8.1. In panel (a) $\gamma = 4$ and in panel (b) $\gamma = 10$.

third case where m is initially positive rather than negative. However, the similarity in the shape of the wavenumbers' evolution is not passed onto the evolution of the wavepacket's amplitudes. In the first case presented, the amplitudes have two areas where they grow more substantially than anywhere else. The first area appears just after day 9 and the second appears just after day 11. By comparing the evolution of the wavevector to the evolution of the amplitudes, it is clear that the periods of growth in the amplitudes coincide with the periods of greatest change of the wavevector, which in turn are caused by Lagrangian transience of the flow. This phenomenon of the Lagrangian transience of the flow causing growth in the amplitudes is also seen in the evolution of wavepackets B and D. It is interesting to note that the amplitudes grow to a whole order of magnitude less in wavepacket B compared to A and D.

These three wavepackets' evolutions can be compared to wavepacket F, where there is relatively little change to the wavepacket's trajectory, figure 8.12. As seen in the figure, the converse of what is discussed above happens. The wavepacket's trajectory is fairly simple, its wavenumbers remain fairly constant and its amplitudes show no sign of growth. This serves to support the discussion laid out above.

The sensitivity to the initial wavenumber that was demonstrated above can be further confirmed by considering the evolution of the amplitudes of wavepacket C shown in figure 8.9. In this case the horizontal wavevector is directly perpendicular to the flow. This figure shows amplitudes that are substantially smaller than those created in

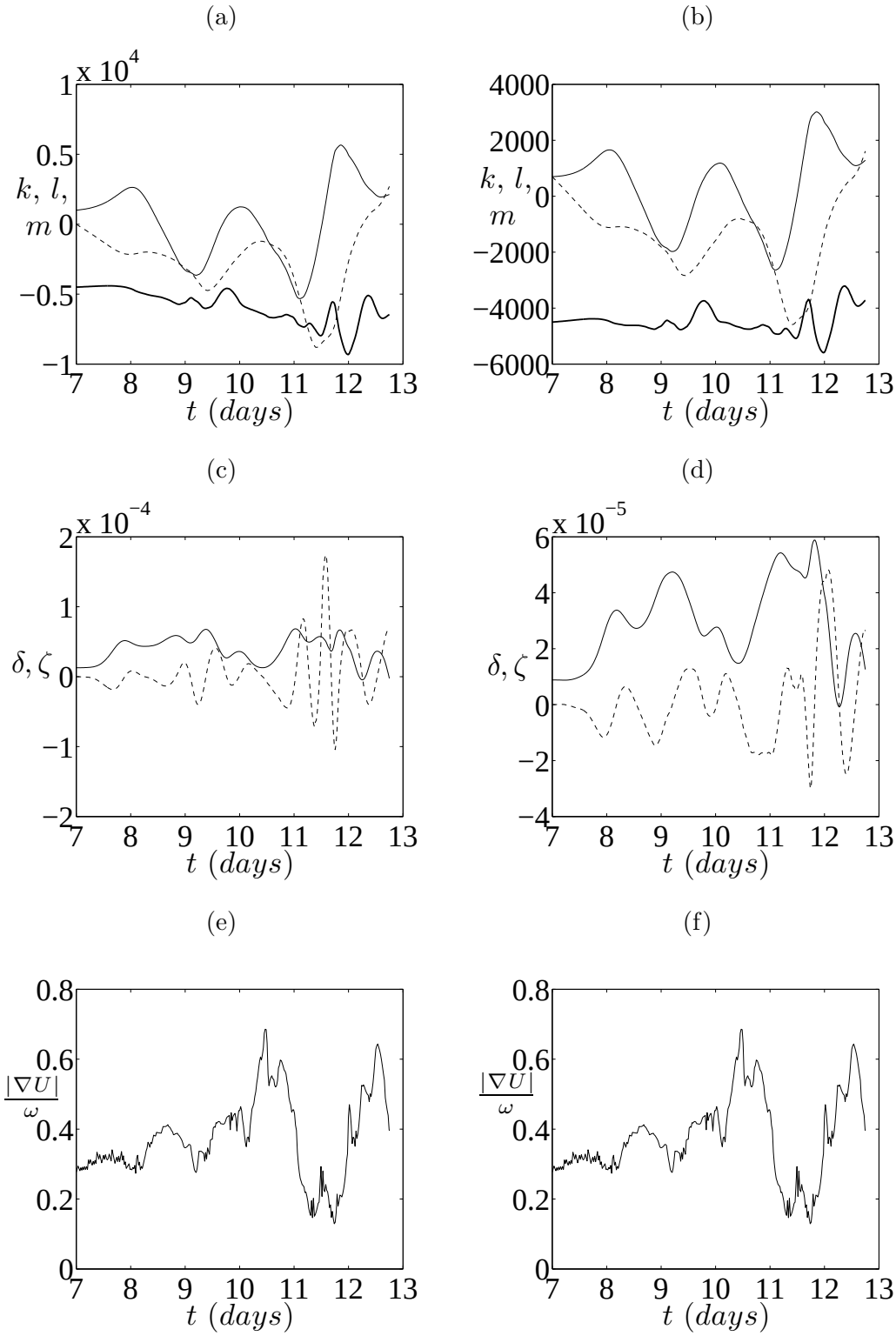

Figure 8.10: The evolution of wavepacket A, left-hand column, and B, right-hand column, see table 8.1, where panels (a) and (b) are the wavenumbers, $100 \times k$, solid line, $100 \times l$ dashed line and m, thick line, (c) and (d) are the amplitudes, δ, dashed line, and ζ, solid line, and (e) and (f) are the local Rossby number.

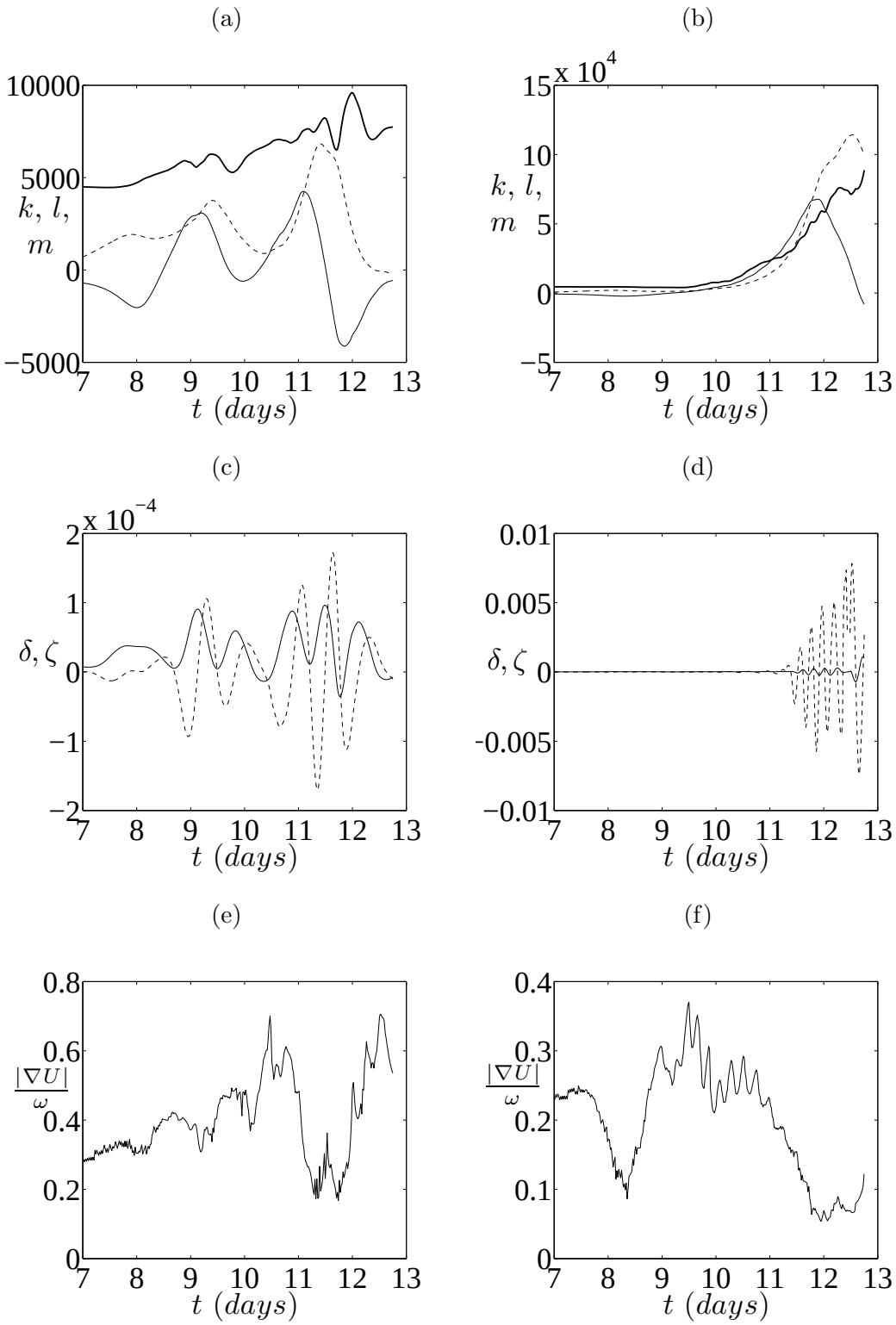

Figure 8.11: The evolution of wavepacket D, left-hand column, and E, right-hand column see table 8.1, where panels (a) and (b) are the wavenumbers, $100 \times k$, solid line, $100 \times l$, dashed line and m, thick line, (c) and (d) are the amplitudes, δ, dashed line, and ζ, solid line and (e) and (f) are the local Rossby number.

wavepackets A and D but are a similar order to those of wavepacket B.

A different situation emerges when we consider a trajectory that takes place at a height of 12 km. The right-hand column of figure 8.11 shows the evolution of the wavevector, amplitudes and the local Rossby number of wavepacket E. This wavepacket exhibits clear signs of being stretched in all directions. This is seen through the fact that all the wavepacket's wavenumbers grow exponentially, meaning that the wavepacket's wavelength decreases exponentially. As a consequence of this the amplitude of the wavepacket also grows exponentially, passing all wave breaking thresholds, until either the wavepacket breaks or another dissipation mechanism occurs.

In section 6.8.3 we introduced and discussed the idea of a local Rossby number that was defined as the norm of the wavepacket's strain matrix over its frequency. By defining it here with the inertia-gravity wave frequency, (3.13), taken as the wavepacket's frequency, we can assess its viability as an indicator for growth of the wavepacket. The local Rossby number for the wavepackets A, B, D and E are shown in figures 8.10 and 8.11. By studying the figures we can see that there is a relationship between the spikes in the local Rossby number and the times when the growth of the wavepackets' amplitudes occur. It is clear that the amplitude of the spike in the local Rossby number does not correspond to the size of the growth observed. For instance, there is a small spike in the low value of the local Rossby number that is present between day 11 and day 12 for wavepacket A, which corresponds to the area of largest growth. The low value at this area and the small amplitude of the spike are caused by the high frequency of the wavepacket, which does not affect the wavenumbers, at this time.

This does not hold for wavepacket E whose evolution takes place at 12 km, but the values of the local Rossby number for this case are a lot lower. We can compare the local Rossby number for wavepackets A, B and D with the case where there is little change to the wavepacket's trajectory, wavepacket F, as we did with the wavenumbers and amplitudes. In panel (d) of figure 8.12 the local Rossby number for this trajectory is shown. From this figure it is clear that there is very little change in a very small local Rossby number which goes along with the fact that there is not much happening in the wavepacket's wavenumbers and amplitudes. These results show that although the local Rossby number is not a completely robust and accurate tool for determining areas of growth, it can be used as a good first indicator.

The areas of the flow that correspond to the points where we see the growth of the wavepacket, correspond to the areas that were found in the same flow by different

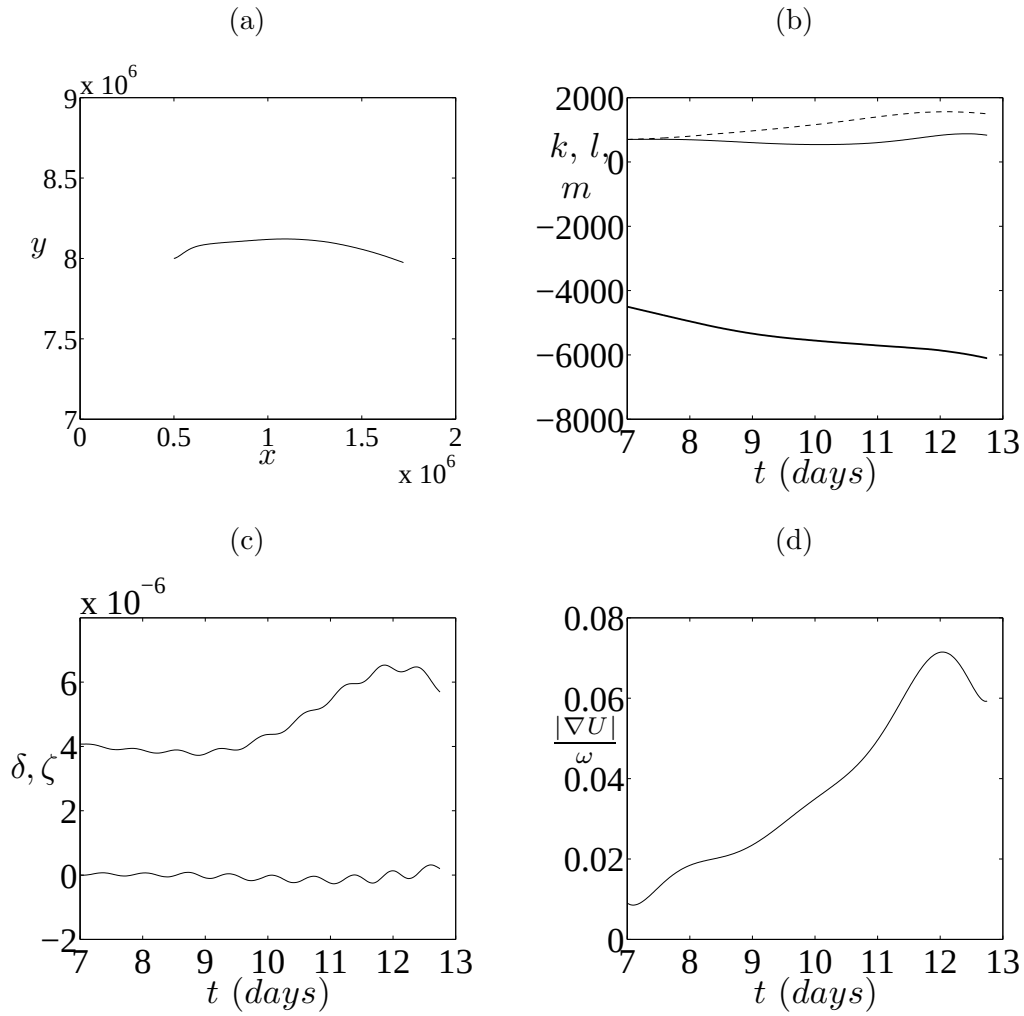

Figure 8.12: The evolution of wavepacket F, see table 8.1, where panel (a) is the trajectory, (b) the wavenumbers, $100 \times k$, solid line, $100 \times l$ dashed line and m, thick line, (c) the amplitudes, δ, dashed line, and ζ, solid line, and (d) the local Rossby number.

methods, in [52]. There, the inertia-gravity waves were detected by studying the divergence field of the flow. From the divergence fields in figures 8.5 and 8.6 we can see that inertia-gravity waves are present at the wavepacket's location at around day 9 and 11. This is also where the areas of large growth are found in our simulation.

As stated earlier, the initial wavenumbers were chosen so that the case where $\alpha = 0$ is never encountered. On first inspection one would expect α to be dominated by f but this breaks down when m is small making it possible that α can pass through zero. When α approaches zero we expect that the terms containing α^{-1} in the equation for $D_t\delta$, (4.56), will cancel each other out but this only happens when the flow is an exact solution of the Boussinesq equations, section 2.3.3. The apparent singularity

arises due to the fact that we have assumed that the flow's potential vorticity and the wavepacket's potential vorticity are both independently conserved. When running numerical simulations the balances between terms that are needed for this to hold are rarely met and so problems arise causing α to pass through zero. Adding a factor of four into 8.9 ensures that this singularity is not encountered in our simulations.

8.8 Conclusion

By using the data from a simulated idealised baroclinic lifecycle we have shown that the system of equations that we have derived can be used both as a means to study inertia-gravity wave generation and as a diagnostic tool to predict areas where this generation may occur.

We have linked the generation of inertia-gravity waves to areas of Lagrangian transience in the flow, that is, areas where the trajectory of the wavepacket changes direction. This transience causes sharp changes in the wavepacket's wavenumbers, which in turn lead to growth of the wavepacket's amplitudes. However, this growth has a limited dependence on the initial direction of the wavevector. We have also shown that there is a link between the local Rossby number of the wavepacket and the growth that is seen in its amplitudes although this link is not as strong as that found for the wavenumbers.

The results also show that our methods could be used to diagnose areas where inertia-gravity wave generation is likely to occur. The growth seen in the trajectories that we have run has coincided with areas that were highlighted in other studies of this flow, [51, 52].

Chapter 9

Conclusion

In this thesis, we have derived a new set of equations that can be used in the study of inertia-gravity wave generation. These equations, which describe the evolution of a small-scale, small-amplitude wavepacket in a rapidly rotating flow, are derived from the Boussinesq set of fluid equations through the use of a WKB technique based on small scales. We then used these equations to analyse a few simple flows showing that although there are some interesting features of the flow, such as wave breaking, there was no growth of the wavepacket in the long time limit.

Next we progressed to consider flows that were created by point vortices with particular attention given to the case of a point-vortex dipole. By exploiting the fact that the flow is uniform at large distances from the dipole we were able to decompose the wavepacket's amplitude equations into their three modes. This enabled us to show that the transient change of the velocity field upon passing the dipole excites a vortical mode wavepacket's gravity wave modes, its amplitude being exponentially dependent on the distance of the wavepacket from the dipole axis. This method could be used in any model where the flow is uniform before and after an event to study spontaneous generation. We then made some observations and proposed some indicators of growth for time dependent models created by point vortices that behave like random strain flows. This model could be taken further in future research through a more detailed and systematic study. By running a large number of simulations, statistics could be used to determine the general trend of wavepackets in random strains. This could be done in conjunction with using random processes to determine the initial conditions of the point vortices. These statistics could then be compared with statistics gained

through using random processes to create the strain instead of point vortices.

Next we showed that all elliptical flows are inherently unstable regardless of the flow's sense of the rotation, its strength, or the strength of its stratification. This was done by deriving an asymptotic formula that predicts the maximum growth rates of these elliptical flows in terms of the flow's eccentricity and the aspect ratio of its perturbation. These results were then verified by solving the problem numerically. These results apply to a far wider range of problems than just the purely elliptical system. They can be applied to any flow that has, or develops, any axisymmetric closed streamlines at any point in its evolution. Hence there are many areas in the field of geophysical fluid dynamics that these results can be applied to.

Finally, by considering a simulated idealised baroclinic lifecycle, we linked the generation of inertia-gravity waves to areas of Lagrangian transience in a flow. This transience causes sharp changes in the wavepacket's wavenumbers, which in turn leads to growth of the wavepacket's amplitudes. We also showed that this growth is dependent on the initial wavevector. The idea of a local Rossby number was discussed, highlighting links between its growth and the growth of the wavepacket. We suggested that the system of equations that we have derived may be used as a diagnostic tool to predict areas where amplification may occur. One application of this may be predicting areas of clear-air turbulence, that is, the erratic movement of air masses in the absence of any visual clues such as clouds. These areas, which can be hazardous to the comfort, and even safety, of air travel, are often found when the troposphere meets the tropopause and in the vicinity of jet streaks. It is possible that areas of high amplification, such as those found in this chapter, are also areas of high clear-air turbulence [28, 53]. The technique derived to accomplish this study of a baroclinic lifecycle could in future be used to study inertia-gravity wave generation in any complex flow.

To further this research it would be interesting to add the vertical component of velocity to our derivation of the wavepacket's governing equations. This would enable the study of new sets of flows with far greater complexity. If this was also coupled with adding in the group velocity, the scope of the equations would be even further increased. At present the wavepacket is always captured by the flow, in the sense that it gets frozen into the flow and its wavecrests behave like passive tracers. If the group velocity was added to the derivation, it would be possible to simulate scenarios where the wavepacket was not frozen into the flow and cases where the spontaneous generation of inertia-gravity waves could result in the waves travelling away in different

directions with different speeds. Another avenue of further research would be to study the singularity that occurs at $\alpha = 0$ that was found in the last chapter. This could involve finding a subtle numerical way of avoiding it or deriving the equations in a slightly different way so that it is never created.

To conclude, through our derivation of equations that govern a wavepacket's properties, we have added to the understanding of inertia-gravity wave generation through spontaneous generation and through unbalanced instabilities. The main contributions are the demonstration that Lagrangian transience of the flow leads to inertia-gravity wave generation, with a local Rossby number being a good indicator of this growth, and that elliptical flows are always unstable.

Appendix A

Change of coordinates

The following work comes largely from private correspondence with Riwal Plougonven.

The primitive equations under the Boussinesq approximation, that are defined in (2.27)-(2.31), can be changed into a set of primitive equations for the atmosphere in terms of the geopotential and temperature which are given by

$$\begin{aligned}
\frac{Du}{Dt} - fv &= -\frac{\partial \phi}{\partial x}, && \text{(A.1)}\\
\frac{Dv}{Dt} + fu &= -\frac{\partial \phi}{\partial y}, && \text{(A.2)}\\
\frac{\partial \phi}{\partial z} - g\frac{\theta}{\theta_r} &= 0, && \text{(A.3)}\\
\frac{D\theta}{Dt} &= 0, && \text{(A.4)}\\
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} &= 0, && \text{(A.5)}
\end{aligned}$$

where $\phi = gz$ is the geopotential defined as the potential of the earth's gravitational field. The other new quantity, θ, is the potential temperature, which is the temperature that a fluid parcel would gain if moved to a reference pressure, p_0, and is given by

$$\theta = T\left(\frac{p_0}{p}\right)^{\frac{R}{C_p}}, \qquad \text{(A.6)}$$

where T is the temperature in degrees Kelvin, R is the gas constant and c_p is the specific heat capacity at a constant pressure. Since we are dealing with an ideal gas we

can link some of these new quantities through the relation of state that is given by

$$P = \rho RT. \tag{A.7}$$

We have now completely defined this system in terms of seven variables, namely $(u, v, w, P, T, \phi, \theta)$ and so the task is to now find a way of changing our original equations, (2.27)-(2.31), that are defined using the variables $(u, v, w, p, \rho, \bar{\rho}, N)$, into this new set of equations.

We can start this transformation by realising that a key aspect of (A.1)-(A.5) is the simple relation between geopotential and the potential temperature when the hydrostatic approximation is made. This relation takes the form $\partial_z \phi \propto \theta$ with θ a conserved quantity. In order to keep this property we have to search for a new vertical coordinate, ξ, such that

$$\frac{\partial \phi}{\partial \xi} = g \frac{\partial z}{\partial \xi} = \theta. \tag{A.8}$$

By introducing the definition of θ we find that

$$\frac{\partial \xi}{\partial z} = \frac{g}{\theta} = \frac{g}{T} \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}}. \tag{A.9}$$

We can now use the hydrostatic approximation to rewrite $\partial_z \xi$ as

$$\frac{\partial \xi}{\partial p} \frac{\partial p}{\partial \xi} = \frac{\partial \xi}{\partial p}(-\rho g) = \frac{g}{T} \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}}, \tag{A.10}$$

which can be simplified using the equation of state, (A.7), to give

$$\frac{\partial \xi}{\partial p} = -\frac{R}{p} \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}}. \tag{A.11}$$

From this we obtain that ξ should have the form

$$\xi = C - c_p \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}}, \tag{A.12}$$

where C is a constant. Following the choice of [26] we pick C such that

$$\xi = \left(1 - \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}} \right) \frac{c_p}{R} H_s, \tag{A.13}$$

where $H_s = p_0/(\rho_0 g)$.

Now that we are in a system that has a vertical coordinate that is based on pressure, the $\rho^{-1}\nabla p$ terms in the horizontal momentum equations change to $\nabla\phi$ which is what we required.

The final equation to look at is the continuity equation. The hydrostatic equilibrium can be used to gain a continuity equation in isobaric coordinates of the form

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} = 0, \tag{A.14}$$

where $\omega = Dp/Dt$. If we now note that

$$\omega = -\frac{p}{H_s}\left(\frac{p}{p_0}\right)^{\frac{R}{c_p}}\omega^* = \frac{p_0^{\frac{R}{c_p}}}{H_s}p^{\frac{c_v}{c_p}}\omega^*, \tag{A.15}$$

and that

$$\frac{\partial \xi}{\partial p} = -\frac{H_s}{p_0^{\frac{R}{c_p}}}p^{-\frac{c_v}{c_p}}, \tag{A.16}$$

we find that

$$\frac{\partial \omega}{\partial p} = p^{-\frac{c_v}{c_p}}\frac{\partial}{\partial \xi}\left(p^{\frac{c_v}{c_p}}\omega^*\right). \tag{A.17}$$

Using this result we can conveniently write the continuity equation as

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{1}{\rho^*}\frac{\partial \rho^*\omega^*}{\partial \xi} = 0, \tag{A.18}$$

where $\rho^* \propto p^{c_v/c_p}$ is a pseudo-density. Hence the final form of the continuity equation is

$$\nabla \cdot (r\mathbf{u}) = 0, \tag{A.19}$$

where

$$r(\xi) = \rho_0\left(\frac{p}{p_0}\right)^{\frac{c_v}{c_p}} = \rho_0\left(1 - \frac{zR}{c_p H_s}\right)^{-\frac{R}{c_p}}, \tag{A.20}$$

is a pseudo-density introduced by [26].

The final step in the change of coordinates is to take the Boussinesq approximation. This simplifies the equations and leaves us with (A.1)-(A.5) as required.

Bibliography

[1] M. J. Ablowitz and A. S. Fokas. *Complex variables: introduction and applications.* Cambridge University Press, 1997.

[2] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions.* Dover, 1965.

[3] D. G. Andrews, J. R. Holton, and C. B. Leovy. *Middle atmosphere dynamics.* Academic Press, 1987.

[4] L. Arnold, G. Papanicolaou, and V. Wihstutz. Asymptotic analysis of the Lyapunov exponent and rotation number of the random oscillator and applications. *SIAM Journal on Applied Mathematics*, 46:427–450, 1986.

[5] J. Aspden and J. Vanneste. Elliptical instability of a rapidly rotating, strongly stratified fluid. *Physics of Fluids*, 21, 2009.

[6] J. Aspden and J. Vanneste. Inertia-gravity-wave generation: a geometric-optics approach. *INI/IUTAM proceedings*, 2009.

[7] S. I. Badulin and V. I. Shrira. On the irreversibility of internal-wave dynamics due to wave trapping by mean flow inhomogeneities. Part 1. Local analysis. *Journal of Fluid Mechanics*, 251:21–53, 1993.

[8] C. M. Bender and S. A. Orszag. *Advanced mathematical methods for scientists and engineers: Asymptotic methods and perturbation theory.* Springer, 1999.

[9] M. V. Berry. Uniform asymptotic smoothing of Stokes's discontinuities. *Proceedings of the Royal Society London A*, 422:7–21, 1989.

[10] O. Buhler and M. E. McIntyre. Wave capture and wave vortex duality. *Journal of Fluid Mechanics*, 534:67–95, 2005.

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

[12] A. B. Olde Daalhuis, S. J. Chapman, J.R. King, J. R. Ockendon, and R. H. Tew. Stokes phenomenon and matched asymptotic expansions. *SIAM Journal of Applied Mathematics.*, 55:1469–1483, 1995.

[13] G. Falkovich, K. Gawędzki, and M. Vergassola. Particles and fields in fluid turbulence. *Reviews of Modern Physics*, 73(4):913–975, 2001.

[14] A. Figotin and L. Pastur. *Spectra of random and almost-periodic operators*. Springer, 1992.

[15] R. Ford, M. E. McIntyre, and W. A. Norton. Balance and the slow quasimanifold: some explicit results. *Journal of the Atmospheric Sciences*, 57(9):1236–1254, 2000.

[16] S. J. Friedlander and A. Lipton-Lifschitz. Localised instabilities in fluids. In S.J. Friedlander and D. Serre, editors, *Handbook of Mathematical Fluid Dynamics, vol. II*, pages 289–353. Elsevier Science, 2003.

[17] D. C. Fritts and M. J. Alexander. Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, 41, 2003.

[18] A. E. Gill. *Atmosphere-Ocean Dynamics*. Academic Press, 1982.

[19] M. Griffths and M. J. Reeder. Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis. I: Model solutions. *Royal Meteorological Society*, 122:1153–1174, 1996.

[20] M. Griffths and M. J. Reeder. Stratospheric inertia-gravity waves generated in a numerical model of frontogenesis. II: Wave sources, generation mechanisms and momentum fluxes. *Royal Meteorological Society*, 122:1175–1195, 1996.

[21] F. Guest, M. Reeder, C. Marks, and D. Karoly. Inertia-gravity waves observed in the lower stratosphere over Macquarie island. *Journal of the Atmospheric Sciences*, 57:737–752, 2000.

[22] D. Guimbard and S. Leblanc. Local stability of the Abrashkin-Yakubovich family of vortices. *Journal of Fluid Mechanics*, 567:91–110, 2006.

[23] V. Hakim. Computation of transcendental effects in growth problems: linear solvability conditions and nonlinear methods - the example of the geometric model. In H. Segur, S. Tanveer, and H. Levine, editors, *Asymptotics beyond all orders*, pages 15–28. Plenum Press, 1991.

[24] P. Haynes and J. Anglade. The vertical-scale cascade in atmospheric tracers due to large-scale differential advection. *Journal of the Atmospheric Sciences*, 54(9):1121–1136, 1997.

[25] E. J. Hinch. *Perturbation methods*. Cambridge University Press, 1991.

[26] B. J. Hoskins and F. P. Bretherton. Atmospheric frontogenesis models: mathematical formulation and solutions. *Journal of the Atmospheric Sciences*, 29:11–37, 1972.

[27] R. R. Kerswell. Elliptical instability. *Annual Review of Fluid Mechanics*, 34:83–113, 2002.

[28] J. Knox, D. McCann, and P. Williams. Application of the lighthillford theory of spontaneous imbalance to clear-air turbulence forecasting. *Journal of the Atmospheric Sciences*, 65:3292–3304, 2008.

[29] R. H. Kraichnan. Convection of a passive scalar by a quasi-uniform random straining field. *Journal of Fluid Mechanics*, 64:737–762, 1974.

[30] C. E. Leith. Nonlinear normal mode initialisation and quasi-geostrophic theory. *Journal of the Atmospheric Sciences*, 37:958–968, 1980.

[31] E. N. Lorenz. Attractor sets and quasi-geostrophic equilibrium. *Journal of the Atmospheric Sciences*, 37:1685–1699, 1980.

[32] E. N. Lorenz. On the existence of a slow manifold. *Journal of the Atmospheric Sciences*, 43:1547–1557, 1986.

[33] R. S. MacKay. Slow manifolds. In T.Dauxois, A.Litvak-Hinenzon, R.S.MacKay, and A.Spanoudaki, editors, *Energy localisation and transfer*, pages 149–192. World Scientific Publishing, 2004.

[34] M. E. McIntyre. Balanced flow. *Encyclopedia of Atmospheric Sciences*, 2003.

[35] M. E. McIntyre. Potential vorticity. *Encyclopedia of Atmospheric Sciences*, 2003.

[36] M. E. McIntyre. Spontaneous imbalance and hybrid vortex gravity structures. *Journal of the Atmospheric Sciences*, 66(5):1315–1326, 2009.

[37] J. C. McWilliams. *Fundamentals of Geophysical Fluid Dynamics*. Cambridge university press, 2006.

[38] J. C. McWilliams and P. R. Gent. Intermediate models of planetary circulations in the atmosphere and ocean. *Journal of the Atmospheric Sciences*, 37(8):1657–1678, 1980.

[39] J. C. McWilliams, M. J. Molemaker, and I. Yavneh. Ageostrophic, anticyclonic instability of a geostrophic, barotropic boundary current. *Physics of Fluids*, 16:3720–3725, 2004.

[40] J. C. McWilliams and I. Yavneh. Fluctuation growth and instability associated with a singularity of the balance equations. *Physics of Fluids*, 10(10):2587–2596, 1998.

[41] T. Miyazaki. Elliptical instability in a stably stratified rotating fluid. *Physics of Fluids A*, 5:2702–2709, 1993.

[42] M. J. Molemaker, J. C. McWilliams, and I. Yavneh. Baroclinic instability and loss of balance. *Journal of Physical Oceanography*, 35:1505–1517, 2005.

[43] B. T. Nadiga and J. M. Aurnou. A tabletop demonstration of atmospheric dynamics: baroclinic instability. *Oceanography*, 21(4):196–201, 2008.

[44] S. Nazarenko, R. J. West, and O. Zaboronski. Fourier space intermittency of the small-scale turbulent dynamo. *Physical Review E*, 68:026311, 2003.

[45] K. Ngan, D. N. Straub, and P. Bartello. Three-dimensionalisation of freely-decaying two-dimensional turbulence. *Physics of Fluids*, 16:2918–2932, 2004.

[46] E. I. Ólafsdóttir, A. B. Olde Daalhuis, and J. Vanneste. Inertia-gravity-wave radiation by a sheared vortex. *Journal of Fluid Mechanics*, 596:169–189, 2007.

[47] D. O'Sullivan and T. J. Dunkerton. Generation of inertia-gravity waves in a simulated life cycle of baroclinic instability. *Journal of the Atmosphere Sciences*, 52:3695–3716, 1995.

[48] R. B. Paris and A. D. Wood. Stokes phenomenon demystified. *Bulletin of the Institute of Mathematics and its Applications*, 31:21–28, 1995.

[49] J. Pedlosky. *Geophysical Fluid Dynamics*. Springer, 1986.

[50] R. Plougonven, D. J. Muraki, and C. Snyder. A baroclinic instability that couples balanced motions and gravity waves. *Journal of the Atmospheric Sciences*, 62:1545–1559, 2005.

[51] R. Plougonven and C. Snyder. Gravity waves excited by jets: propagation versus generation. *Geophysical Research Letters*, 32, 2005.

[52] R. Plougonven and C. Snyder. Inertia-gravity waves spontaneously generated by jets and fronts. Part I: different baroclinic life cycles. *Journal of the Atmospheric Sciences*, 64(7):2502–2520, 2007.

[53] R. Plougonven, C. Snyder, and F. Zhang. Comments on application of the Lighthill-Ford theory of spontaneous imbalance to clear-air turbulence forecasting. *Journal of the Atmospheric Sciences*, 66:2506–2510, 2009.

[54] R. Plougonven and H. Teitelbaum. Comparison of a large-scale inertia-gravity wave as seen in the ECMWF and from radiosondes. *Geophysical research letters*, 30(18), 2003.

[55] R. Rutunno, C. Snyder, and W. C. Skamarock. An analysis of frontogenesis in numerical simulations of baroclinic waves. *Journal of the Atmospheric Sciences*, 51:3373–3398, 1994.

[56] R. Salmon. *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 1998.

[57] W. C. Skamarock, J. N. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers. A description of the advanced research WRF version 2. *NCAR technical note NCAR/TN468+STR*, 2005.

[58] C. Snyder, D. J. Muraki, R. Plougonven, and F. Zhang. Inertia-gravity waves generated within a dipole vortex. *Journal of the Atmospheric Sciences*, 64(12):4417–4431, 2007.

[59] C. Snyder, R. Plougonven, and D. J. Muraki. Mechanisms for spontaneous gravity wave generation within a dipole vortex. *Journal of the Atmospheric Sciences*, 66(11):3464–3478, 2009.

[60] D. N. Straub. Instability of 2D flows to hydrostatic 3D perturbations. *Journal of the Atmospheric Sciences*, 60:79–87, 2003.

[61] C. D. Thorncroft, B. J. Hoskins, and M. E. McIntyre. Two paradigms of baroclinic-wave life-cycle behaviour. *Quarterly journal of the Royal Meteorological Society*, 119:17–55, 1993.

[62] G. K. Vallis. *Atmospheric and Oceanic Fluid Dynamics, Fundamentals and Large-Scale Circulation.* Cambridge University Press, 2006.

[63] J. Vanneste. Inertia-gravity-wave generation by balanced motion: revisiting the Lorentz-Krishnamurthy model. *Journal of the Atmospheric Sciences*, 61:224–234, 2004.

[64] J. Vanneste. Exponential smallness of inertia-gravity-wave generation at small Rossby number. *Journal of the Atmospheric Sciences*, 65:1622–1637, 2008.

[65] J. Vanneste and I. Yavneh. Exponentially small inertia-gravity waves and the breakdown of quasi-geostrophic balance. *Journal of the Atmospheric Sciences*, 61:211–223, 2004.

[66] J. Vanneste and I. Yavneh. Unbalanced instabilities of rapidly rotating stratified shear flows. *Journal of Fluid Mechanics*, 584:373–396, 2007.

[67] A. Viúdez. Spiral patterns of inertia-gravity waves in geophysical flows. *Journal of Fluid Mechanics*, 562:73–82, 2006.

[68] A. Viúdez. The origin of the stationary frontal wave packet spontaneously generated in rotating stratified vortex dipoles. *Journal of Fluid Mechanics*, 593:359–383, 2007.

[69] S. Wang, F. Zhang, and C. Snyder. Generation and propagation of inertia-gravity waves from vortex dipoles and jets. *Journal of the Atmospheric Sciences*, 66(5):1294–1314, 2009.

[70] T. Warn, O. Bokhove, T. G. Shepherd, and G. K. Vallis. Rossby number expansions, slaving principles,and balance dynamics. *Quarterly journal of the Royal Meteorological Society*, 121:723–739, 1995.

[71] M. I. Weinstein and J. B. Keller. Asymptotic behaviours of stability regions for Hill's equation. *Journal of Applied Mathematics*, 47(5):941–958, 1987.

[72] I. Yavneh, J. C. McWilliams, and M. J. Molemaker. Non-axisymmetric instability of centrifugally-stable stratified Taylor-Couette flow. *Journal of Fluid Mechanics*, 448:1–21, 2001.

[73] F. Zhang. Generation of mesoscale gravity waves in upper-tropospheric jet-front systems. *Journal of the Atmospheric Sciences*, 61:440–457, 2004.