Nonlinear interactions of fast and slow modes in rotating, stratified fluid flows

Submitted for the degree of Doctor of Philosophy

Paul David Williams
Balliol College

Trinity Term, 2003

Atmospheric, Oceanic and Planetary Physics
Department of Physics
University of Oxford

Nonlinear interactions of fast and slow modes in rotating, stratified fluid flows

Paul David Williams, Balliol College

Submitted for the degree of Doctor of Philosophy

Trinity Term, 2003

This thesis describes a combined model and laboratory investigation of the generation and mutual interactions of fluid waves whose characteristic scales differ by an order of magnitude or more. The principal aims are to study how waves on one scale can generate waves on another, much shorter scale, and to examine the subsequent nonlinear feedback of the short waves on the long waves. The underlying motive is to better understand such interactions in rotating, stratified, planetary fluids such as atmospheres and oceans.

The first part of the thesis describes a laboratory investigation using a rotating, two-layer annulus, forced by imposing a shear across the interface between the layers. A method is developed for making measurements of the two-dimensional interface height field which are very highly-resolved both in space and time. The system's linear normal modes fall into two distinct classes: "slow" waves which are relatively long in wavelength and intrinsic period, and "fast" waves which are much shorter and more quickly-evolving. Experiments are performed to categorize the flow at a wide range of points in the system's parameter space. At very small background rotation rates, the interface is completely devoid of waves of both types. At higher rates, fast modes only are generated, and are shown to be consistent with the Kelvin-Helmholtz instability mechanism based on a critical Richardson number. At rotation rates which are higher still, baroclinic instability gives rise to the onset of slow modes, with subsequent localized generation of fast modes superimposed in the troughs of the slow waves.

In order to examine the generation mechanism of these coexisting fast modes, and to assess the extent of their impact upon the evolution of the slow modes, a quasi-geostrophic numerical model of the laboratory annulus is developed in the second part of the thesis. Fast modes are filtered out of the model by construction, as the phase space trajectory is confined to the slow manifold, but the slow wave dynamics is accurately captured. Model velocity fields are used to diagnose a number of fast wave radiation indicators. In contrast to the case of isolated fast waves, the Richardson number is a poor indicator of the generation of the coexisting fast waves that are observed in the laboratory, and so it is inferred that these are not Kelvin-Helmholtz waves. The best indicator is one associated with the spontaneous emission of inertia-gravity waves, a generalization of geostrophic adjustment radiation.

A comparison is carried out between the equilibrated wavenumbers, phase speeds and amplitudes of slow waves in the laboratory (which coexist with fast modes), and slow waves in the model (which exist alone). There are significant differences between these wave properties, but it is shown that these discrepancies can be attributed to uncertainties in fluid properties, and to model approximations apart from the neglect of fast modes. The impact of the fast modes on the slow modes is therefore sufficiently small to evade illumination by this method of inquiry. As a stronger test of the interaction, a stochastic parameterization of the inertia-gravity waves is included in the model. Consistent with the laboratory/model intercomparison, the parameterized fast waves generally have only a small impact upon the slow waves. However, sufficiently close to a transition curve between two different slow modes in the system's parameter space, it is shown that the fast modes can exert a dominant influence. In particular, the fast modes can force spontaneous transitions from one slow mode to another, due to the phenomenon of stochastic resonance. This finding should be of interest to the meteorological and climate modelling communities, because of its potential to affect model reliability.

Acknowledgments

Many people have helped to make my Oxford D.Phil. experience unique and memorable.

On the academic side, thanks first and foremost to Peter Read and Tom Haine, my supervisors, whose guidance and expertise over the last four years have been invaluable. Thanks to all in the Geophysical and Planetary Fluid Dynamics group: Fiona, Henning, Richard, Edgar, Claire, Suzy, Steve and Hiro, for teaching me new things to do with a noisette at Christmas. Thanks to Ron Daniel for helpful lunchtime discussions about stochastic resonance, and to Jason Perry for helping to make mechanical improvements to the annulus lid.

On the administrative side, thanks to Sarah Harrington and Monika Porada for all their assistance; to Stuart Teasdale for much-needed computing support; and to Declan and Richard for keeping the Oven clean.

Thanks to the Natural Environment Research Council for funding this study; to the Royal Meteorological Society and the Institute of Physics for funding my attendance at overseas conferences; to the Oxford University Vice-Chancellors' Award committee for giving me a grant in my fourth year; to the Senior Common Room at Brasenose College for employing me as Senior Warden; and to my colleagues and friends — Leslie, Nathan and Caroline — for making the experience a pleasurable one.

Thanks to Paul Ingram and Mike Woodin, whose ideas about how the world should be have had a considerable influence on my own. What a difference six votes makes!

Thanks to my parents, Edward and grand-parents: if I have achieved success, it is because you have all made sacrifices for me, and that is something which I will never forget. Thanks to Paul for the Pink Floyd and the Jeffams; to Gosia for her Polish match-making service; to James for all those 80s nights at Fifth Avenue; to Fiona and Pippa for bringing part of the North to Oxford; to Tony for paying money to drink the Jeffams; and to Emma and Dorothee for everything.

Thank you all.

Contents

1 Introduction 5

1.1 Overview 5

1.2 Scale-separated interactions in non-fluid systems 10

1.3 The slow manifold 12

1.4 Application to weather forecasting 14

1.4.1 The importance of initialization 16

1.4.2 History of numerical weather prediction 18

1.4.3 When might inertia-gravity waves be resolved? 21

1.5 Dynamical similarity and laboratory experiments 22

1.6 Previous laboratory work 24

1.7 The current study 25

1.7.1 Aims of this investigation 26

1.7.2 Thesis outline 27

2 Description of the laboratory apparatus 28

2.1 The rotating, two-layer annulus 29

2.2 Fluid properties 32

2.2.1 Thermal expansivities 33

2.2.2 Refraction 35

2.2.3 Interfacial tension 36

2.2.4 Slow evolution of fluid properties with time 38

2.3 Flow visualization 39

2.4 The parabolic correction tank 43

2.5 Parallax effects 45

2.6 Chapter summary 46

3 Calibration of the laboratory experiment 47

3.1 Motivation for calibrating the experiment . . . 47

3.2 Theoretical approach . . . 48

3.3 Direct experimental approach . . . 49

3.4 Indirect experimental approach . . . 50

3.4.1 Equilibrium interface height field . . . 50

3.4.2 Previous approaches . . . 53

3.4.3 Torque balance calculation . . . 54

3.5 Implementation of the calibration scheme . . . 61

3.5.1 Choice of colour calibration variable . . . 63

3.5.2 Derivation of calibration curve . . . 64

3.6 Noise analysis . . . 68

3.7 Sample reconstructed height field . . . 71

3.8 Chapter summary . . . 73

4 Results of the laboratory experiments 74

4.1 Findings of Lovegrove . . . 75

4.2 Preliminary difficulties with current experiments . . . 77

4.3 Description of current experiments . . . 79

4.4 Experimental results: PAI series . . . 80

4.4.1 Analysis of small-scale waves in the *KH* regime . . . 87

4.4.2 Analysis of small-scale waves in the *MRW* and *MIW* regimes . . . 91

4.5 Experimental results: RAI series . . . 92

4.6 Experimental results: PAD series . . . 95

4.7 Experimental results: PEI series . . . 96

4.8 Calculation of wave amplitudes . . . 98

4.9 Discussion . . . 102

4.10 Chapter summary . . . 103

5 Description of the numerical model 104

5.1 Motivation for running numerical simulations . . . 105

5.2 Review of filtered models . . . 106

5.2.1 Candidate filtered models . . . 106

5.2.2 Candidate geometries . . . 107

5.3 Derivation of model equations 109
    5.3.1 Perturbation equations 114
    5.3.2 Normal mode decomposition of diagnostic equations 115
5.4 Perturbation streamfunction boundary conditions for the continuous equations 117
    5.4.1 Integral properties of the prognostic equations 119
    5.4.2 Integral properties of the diagnostic equations 120
5.5 Discretization of model equations 121
    5.5.1 Prognostic equations 123
    5.5.2 Diagnostic equations 125
5.6 Perturbation streamfunction boundary conditions for the discretized equations 127
5.7 Details of the numerical schemes 128
5.8 Initial conditions 130
5.9 Suitable values for numerical parameters 132
5.10 Chapter summary 134

6 Results of the numerical experiments 135

6.1 Model runs with zero interfacial tension 136
    6.1.1 Sample diagnostics 137
    6.1.2 Comparison between model and laboratory 144
6.2 Variation of model viscosity 148
6.3 Variation of model interfacial tension 151
6.4 Model runs with non-zero interfacial tension 154
6.5 Discussion of model/laboratory comparison 156
6.6 Radiation indicators 158
    6.6.1 Indicator definitions 158
    6.6.2 Indicator plots using model data 163
6.7 Chapter summary 167

7 A stochastic parameterization of the fast waves 169

7.1 Parameterizations of inertia-gravity waves 170
7.2 Model runs with the stochastic parameterization 172
    7.2.1 Reproducibility of the equilibrated state 172

7.2.2 Diagnosis of amplitude of stochastic small-scale features 175
7.2.3 Continuous variation of amplitude of stochastic small-scale features 178
7.3 Discussion 180
7.4 Chapter summary 181

8 Conclusions and future work 182
8.1 Summary of findings 182
8.2 Conclusions 185
8.3 Future work 190

A The *HSI* colour system 192

Chapter 1

Introduction

> *"We might say that the atmosphere is a musical instrument on which one can play many tunes. High notes are sound waves, low notes are long inertial waves, and nature is a musician more of the Beethoven than of the Chopin type."*
>
> Letter from Jule Charney to Phillip Thompson,
> 12 February 1947.

1.1 Overview

Like many physical systems, fluids often exhibit the coexistence of motions on a wide range of space and time scales. Correspondingly, the linear normal modes of the governing Navier-Stokes equations generally have spatio-temporal structures which fall naturally into distinct classes, when categorized according to the fundamental dynamical mechanisms which permit their existence. This property of the fluid equations was first identified by Margules (1893), who derived two species of solutions to Laplace's tidal equations. He named his solutions *"Wellen erster Art"* (waves of the first type) and *"Wellen zweiter Art"* (waves of the second type), which we now know as inertia-gravity and Rossby waves. Important studies of the characteristics of these modes have been presented by Hough (1898) and Longuet-Higgins (1968).

As a geophysical example of scale-separated fluid motions, the Earth's atmosphere and ocean support acoustic waves, which have relatively short wavelengths and are quickly

	λ/m	τ/s	c_{ph}/m s^{-1}
acoustic waves	1	10^{-2}	10^{2}
inertia-gravity waves	10^{5}	10^{4}	10
Rossby waves	10^{6}	10^{6}	1

Table 1.1: Characteristic horizontal wavelengths (λ), intrinsic periods (τ) and phase speeds (c_{ph}) of three different classes of waves observed in the Earth's atmosphere. These three quantities are related by $c_{\mathrm{ph}} = \lambda/\tau$. Values are given to the nearest typical order of magnitude; in reality there is a significant spread in characteristic scales about these mean values.

propagating; inertia-gravity waves, which are much longer and more slowly evolving; and Rossby waves, which are longer and slower still. Typical scales associated with these three linear eigen-modes in the atmosphere are shown in Table 1.1.

The presence of multiple, disparate scales in fluid motions presents a distinct difficulty for theoretical and numerical modelling, analysis and prediction. Simultaneously and accurately capturing the evolution of features whose characteristic scales differ by an order of magnitude or more, requires an exceptionally careful treatment and large computational effort. Moreover, computational resources aside, the primary focus of a study is usually restricted to just a subset of all the permissible motions. Including the entire span of scales would be distracting, and moreover, such a comprehensive analysis is not usually regarded as necessary for achieving reliable results. This is because interactions between motions on the scales of interest, and motions on dramatically different scales, are tacitly assumed to be negligible.

For these reasons it has become commonplace to filter from fluid dynamical models, motions on those scales which are not of primary interest. For example, a study of atmospheric Rossby and inertia-gravity waves is unlikely to be concerned with the acoustic waves with which they coexist. Since acoustic waves require fluid compressibility, they can be filtered out of the analysis by imposing an incompressibility approximation.

Inertia-gravity waves and Rossby waves are still permitted in this filtered model, but any nonlinear interaction which takes place between these modes and acoustic modes in the real system, cannot be captured. As another example, Haine & Williams (2002) have filtered out small-scale convective plumes from a surface ocean front model by imposing a hydrostatic approximation, and have investigated the effects of the filtering on the large-scale cross-frontal transfer.

The question which naturally arises is: are nonlinear wave interactions strong enough to significantly reduce the applicability of a filtered model, in which at least one of the wave types present in the real system is absent? There must be a matching of space and/or time scales in order for a significant interaction to occur, as we will quantify shortly. In the above example, the length, time and speed scales of the filtered (acoustic) and retained (inertia-gravity and Rossby) modes are so poorly matched, that for all practical purposes the answer is surely “no”. Though possible in principle — due to the *butterfly effect* (Drazin, 1992) — there is no evidence that the sound of a person’s voice can have any real impact upon tomorrow’s weather!

However, we face a potentially different situation when we consider the filtering out of inertia-gravity waves, through imposing a balance approximation such as quasi-geostrophy. The characteristic scale separation factors between inertia-gravity and Rossby waves are only around 10–100, as compared with 10^6 in the case of acoustic and Rossby modes. This makes it feasible that a small, but significant Rossby/inertia-gravity wave interaction could exist.

The intuitive notion that there must be a reasonable matching of scales for a strong interaction has been quantified by Benney (1977). He presented a simple, general theory for interactions between short and long waves, which is outlined here. For a resonant triad interaction to occur in a fluid system with dispersion relation $\omega(\boldsymbol{k})$, the wave vectors $\boldsymbol{k}_{\mathrm{a}}$, $\boldsymbol{k}_{\mathrm{b}}$ and $\boldsymbol{k}_{\mathrm{c}}$ are required (e.g. Gill, 1982) to satisfy

$$\boldsymbol{k}_{\mathrm{a}} - \boldsymbol{k}_{\mathrm{b}} = \boldsymbol{k}_{\mathrm{c}} \tag{1.1}$$

and

$$\omega(\boldsymbol{k}_{\mathrm{a}}) - \omega(\boldsymbol{k}_{\mathrm{b}}) = \omega(\boldsymbol{k}_{\mathrm{c}}) \ . \tag{1.2}$$

Equation (1.1) is satisfied by

$$\boldsymbol{k}_{\mathrm{a}} = \boldsymbol{k}_{\mathrm{s}} + \tfrac{1}{2}\boldsymbol{k}_{\mathrm{l}} \,, \quad \boldsymbol{k}_{\mathrm{b}} = \boldsymbol{k}_{\mathrm{s}} - \tfrac{1}{2}\boldsymbol{k}_{\mathrm{l}} \,, \quad \boldsymbol{k}_{\mathrm{c}} = \boldsymbol{k}_{\mathrm{l}} \,, \tag{1.3}$$

where the subscripts l and s refer to long and short waves, respectively. Equations (1.3) represent a triad consisting of two short waves (a and b) and one long wave (c). Using $\boldsymbol{k}_{\mathrm{l}} \ll \boldsymbol{k}_{\mathrm{s}}$ to employ a linear approximation to the dispersion relations, we find that equation (1.2) holds for this triad if

$$\boldsymbol{k}_{\mathrm{l}} \,.\, \boldsymbol{\nabla}_{\boldsymbol{k}}\, \omega(\boldsymbol{k}_{\mathrm{s}}) = \omega(\boldsymbol{k}_{\mathrm{l}}) \,, \tag{1.4}$$

where $\boldsymbol{\nabla}_{\boldsymbol{k}}$ is the gradient operator in wave vector space. A resonant triad interaction is therefore permitted between long and short waves if the phase velocity of the long wave is equal to the component of the group velocity of the short wave in the direction of travel of the long wave. In one dimension, this requirement simplifies to the phase speed of the long wave being equal to the group speed of the short wave. The physical interpretation is that the energy of the short modes, which travels at their group speed, must not drift relative to the phase of the long mode. This means that any energy transfer from short to long modes is focussed at particular locations fixed relative to the nodes of the long mode, rather than being input across all long mode phases, and this requirement evidently allows a resonant reinforcement of the energy transfer.

Figure 1.1 shows typical dispersion curves for the three classes of atmospheric waves previously mentioned. It is clear that there is no possibility of equality between the phase speed of a Rossby wave and the group speed of an acoustic wave, as acoustic waves propagate too quickly. This helps to justify the filtering of acoustic modes from a Rossby wave model, as previously discussed. But the possibility is open — in principle, at least — of equality between the phase speed of a long Rossby wave and the group speed of a short gravity wave, and therefore of a resonant interaction and energy exchange.

The above analysis throws a question mark over the conventional wisdom that the Rossby/inertia-gravity wave interaction is always negligible. Since Rossby waves are prototypes of the atmospheric disturbances which constitute our weather systems, and since inertia-gravity waves are actively filtered out of numerical weather prediction models by the initialization procedure (Section 1.4.1), a non-negligible interaction between

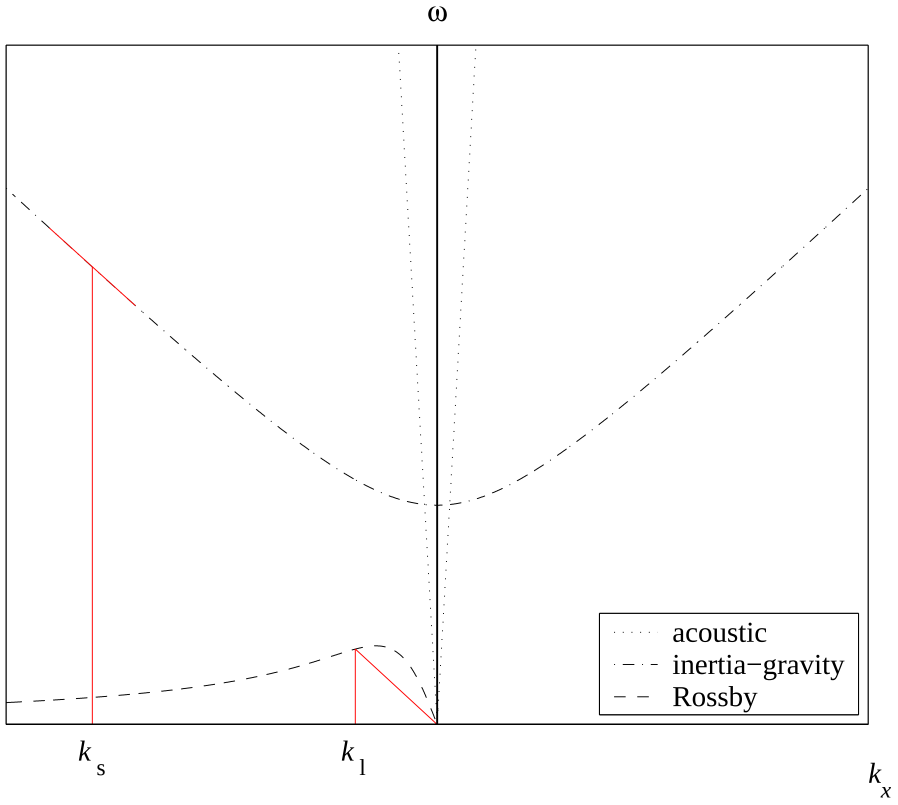


Figure 1.1: Schematic plot of the dispersion curves for three different zonally-propagating atmospheric wave modes, showing intrinsic angular frequency ω as a function of zonal (East-West) wave vector k_x. The chord slope at k_l is equal to the tangent slope at k_s, as shown in red.

the two modes would have possible implications for accurate weather forecasting.

This thesis is a study of the interactions between “fast” and “slow” modes, such as Rossby and inertia-gravity waves, in fluid flows. We are principally interested in investigating two key aspects of the interaction: firstly, the ability of slow motions to generate much faster motions; and secondly, the ability of the fast modes, once generated, to impact upon the evolution of the slow modes. In the present study, we investigate these issues using a laboratory experiment and accompanying numerical model. The underlying motive of the investigation is to understand the interaction properties on a much larger scale, in planetary atmospheres and oceans.

1.2 Scale-separated interactions in non-fluid systems

Systems which display interactions between multiple scale processes are ubiquitous across the entire spectrum of the natural sciences. Examples of systems with more than one timescale (*stiff* systems), from the fields of classical mechanics, chemistry and circuit theory, are briefly discussed in the following paragraphs. Though the governing equations for these systems may be very different from those of fluid systems, we might expect scale-separated interaction characteristics to be robust enough to withstand changes in the detailed phase space topologies. We may therefore be able to infer some of the qualitative features of scale-separated interactions in fluids, from those in the other systems.

The swinging spring

This system is also known as the *elastic pendulum*, since it is like the well-known simple pendulum except that the rigid rod is replaced with a spring. A point mass m is attached to a spring with force constant k in a gravitational field g, so that the equilibrium length of the spring is l. When constrained to move in a vertical plane, there are two modes of oscillation, distinguished by the physical mechanism that provides the restoring force. As with the simple pendulum, there is a rotational linear normal mode with timescale $\sqrt{l/g}$, but now there is also an elastic linear normal mode due to spring stretching and compression, with timescale $\sqrt{m/k}$. For the typical values $m = 1$ kg, $k = 10^3$ N m^{-1}, $g = 10$ m s^{-2} and $l = 1$ m, the timescales are 0.3 s and 0.03 s, showing a clear separation by a factor of ten.

Lynch (2000) has described an explicit analogy between the swinging spring system and atmospheric flow. He compares the high-frequency elastic modes to inertia-gravity waves, and the low-frequency rotational modes to Rossby waves. In the swinging spring, as in the atmosphere, there is nonlinear coupling between the fast and slow modes. The analogy proves to be useful, even though the swinging spring has a four-dimensional phase space whereas the atmosphere explores a manifold of dimension much larger than four in an infinite-dimensional phase space. In both cases, for general initial conditions both normal modes are excited and persist as the system evolves, but it is possible to determine initial conditions for which the fast modes remain absent. Lynch derives such initial conditions for the spring, and uses his analogy to illustrate the concept of atmospheric balance.

Lynch goes on to apply the Kolmogorov-Arnold-Moser (KAM) theorem to the spring system (Arnold, 1963). KAM theory provides a method for investigating how the phase space structures of a completely integrable system (e.g. the rigid pendulum) are modified when the system is slightly perturbed in such a way that integrability no longer exactly holds (e.g. by replacing the pendulum with a spring and allowing the fast and slow modes to interact). His conclusion is that if most of the energy is initially in the slow mode, then only an amount proportional to the ratio of fast to slow timescales can be transferred to the fast oscillations. This result is rigorously valid only for systems with four degrees of freedom, though Lynch speculates that the conclusion may be portable to the inertia-gravity/Rossby wave interaction in fluids with many more than four independent variables.

Chemical reactions

Chemical processes often exhibit the coexistence of chemical reactions and transport phenomena with a wide range of timescales. Vora & Daoutidis (2001) have developed a general method for removing the fast variables from analyses of chemical systems, something akin to making an assumption of balance in a fluid dynamics model. They test their method by comparing numerical integrations of some full-order models and associated reduced models, and find excellent agreement.

In particular, they study the esterification of a carboxylic acid with an alcohol in the presence of a catalyst. The system has 11 chemical species and 12 elementary reactions, of which 4 are slow and 8 are fast. Their method allows a reduction in the number of degrees of freedom from 11 to 7, significantly reducing the computational expense.

Chua's circuit

The chaotic behaviour of this nonlinear electronic circuit has been widely studied (Madan, 1993). It has two capacitors C_1 and C_2, an inductor L, and a diode with a nonlinear current-voltage response. These four components, together with a power supply, are all placed in parallel. There are two natural timescales in the equations: $\sqrt{LC_1}$ and $\sqrt{LC_2}$. In a common set-up, $L = 10$ mH, $C_1 = 0.08$ μF and $C_2 = 4.4$ nF, so that the fast and slow timescales are 7 μs and 28 μs.

Zhu et al. (2002) have shown in a laboratory experiment using Chua's circuit that the temporal regularity is enhanced by the introduction of a small amplitude noise source. This is a phenomenon known as *stochastic resonance*, which we investigate in fluids in Chapter 7 when we add a stochastic inertia-gravity wave parameterization to a quasi-geostrophic numerical model.

1.3 The slow manifold

The incompressible Navier-Stokes equations can be written in the general form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \ , \tag{1.5}$$

where the dot denotes differentiation with respect to time. The state vector $\boldsymbol{x}(t)$ contains the values of all the dynamical variables (velocity, pressure, etc.) at each spatial point at time t. The number of elements of $\boldsymbol{x}$ corresponds to the number of degrees of freedom of the fluid, which is formally infinite for the continuous system, but finite for a discretized or truncated normal mode approximation. The nonlinear vector operator $\boldsymbol{f}$ represents the dynamics embodied in the Navier-Stokes equations, including boundary conditions. The phase space trajectory traced out by $\boldsymbol{x}(t)$ describes the evolving state of the fluid.

Imposing a balance condition, to filter out inertia-gravity waves, corresponds to imposing a set of constraint equations which can be written in the general form

$$\boldsymbol{g}(\boldsymbol{x}) \; = \; \mathbf{0} \, . \tag{1.6}$$

Phase space trajectories $\boldsymbol{x}(t)$ are now constrained to move on the hyper-surface $\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$, which reduces the number of degrees of freedom of the system by the number of elements of the vector operator $\boldsymbol{g}$. The surface $\boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$ is called the fluid dynamical *slow manifold*, because trajectories which are constrained to evolve on it are completely devoid of fast inertia-gravity modes.[1] Slow manifolds can be defined in the same way for the other scale-separated systems discussed in Section 1.2.

It is more natural to picture the slow manifold in the phase space spanned by the linear normal mode variables, rather than by the physical variables. The linear normal modes of equations (1.5) are naturally partitioned into a fast class (inertia-gravity waves) and a slow class (Rossby waves). Since the spatio-temporal structures $\boldsymbol{x}_n^{\text{fast}}$ and $\boldsymbol{x}_n^{\text{slow}}$ of the linear normal modes form a complete set, solutions to equations (1.5) can be always written

$$\boldsymbol{x}(t) \; = \sum_n \; a_n^{\text{fast}}(t) \; \boldsymbol{x}_n^{\text{fast}} + \sum_n \; a_n^{\text{slow}}(t) \; \boldsymbol{x}_n^{\text{slow}} \; . \tag{1.7}$$

This equation represents a transformation between the elements of the state vector $\boldsymbol{x}(t)$ and the normal mode amplitudes $a_n(t)$. The slow manifold is now simply defined as $a_n^{\text{fast}} = 0 \quad \forall \; n$.

The concept of the slow manifold is intimately related to fast/slow interactions. If the interaction is negligible then we may constrain model trajectories to the slow manifold without loss of accuracy, justifying the use of filtered models.

---

[1] In turn, solutions of the full *incompressible* Navier-Stokes equations can also be said to exist on a slow manifold, namely that manifold of the *compressible* Navier-Stokes equations which is completely devoid of the “very fast” acoustic modes.

The existence of a strict *invariant* slow manifold within the unapproximated equations of fluid motion, upon which a real flow may evolve without ever exciting inertia-gravity wave modes, has been debated ever since it was defined by Leith (1980) and Lorenz (1980). There is no guarantee that a trajectory which starts out on the slow manifold but which evolves according to the full equations (1.5) will remain on the slow manifold for all time. This is a potential source of unreliability in any filtered model which makes the *a priori* assumption of perpetual slow manifold confinement.

Lorenz (1986) and Lorenz & Krishnamurthy (1987) have investigated the problem by constructing reduced-dimensional primitive equation models derived from the shallow water equations, containing three slow and two fast independent variables. They could not find initial conditions for which the fast motions remained absent during their numerical integrations, and concluded that this model did not possess an invariant slow manifold. Jacobs (1991) examined the same numerical model and discovered that a slow manifold did in fact exist. Lorenz (1992) argues that there is no mathematical inconsistency between the two conclusions, and that the apparent discrepancy is due to differing definitions of the slow manifold.

More recently, Yavneh & McWilliams (1994) report a distinct breakdown of the slow manifold at a critical Rossby number, in a numerical solution of the shallow water equations, though their breakdown is primarily associated with enhanced dissipation rather than an initiation of inertia-gravity wave propagation. Warn & Menard (1986) have argued that, when the slow manifold breaks down, it may simply be perturbed from a smooth subspace to a quasi-stochastic subspace, but that this "fuzzy manifold" may still possess many of the most useful properties of the original slow manifold. The questions of the existence of slow and fuzzy manifolds, and even of how to properly define them, clearly remain controversial.

1.4 Application to weather forecasting

Inertia-gravity waves exist ubiquitously in the Earth's atmosphere, and are of particularly large amplitude in the upper troposphere and lower stratosphere. They are observed in high resolution data from radiosondes, rockets and satellites (e.g. Sato (1994) and ref-

Figure 1.2: Gravity waves in noctilucent clouds photographed over Kiruna, Sweden at 23:08 on the night of 10th August 2000. On this occasion, the waves persisted for around 15 minutes before dissipating away. (Photograph by Dr. S. Kirkwood, Dr. P. Dalin and Dr. A. Moström, Swedish Institute of Space Physics.)

erences therein for radar observations). They are reported to have vertical wavelengths of around 1–5 km, horizontal wavelengths of around 200–1000 km and intrinsic time periods of around 10 hours.

Recently, extensive ground-based observations of atmospheric pure gravity waves have been made by a team at the Swedish Institute of Space Physics. They report that gravity waves of wavelengths in the range 5–50 km are visible from the ground in noctilucent clouds (NLCs) on around one night in three during the summer months, predominantly between latitudes 50–70°N and times 22:00–04:00 (Dr. P. Dalin, personal communication). A typical observation is shown in Figure 1.2.

Observations such as these have been mirrored in high-resolution numerical simulations of the atmosphere (e.g. O'Sullivan & Dunkerton, 1995; Sato et al., 1999) in which inertia-gravity waves have been generated by the geostrophic adjustment of a baroclinically-unstable tropospheric jetstream. An example of the inertia-gravity wave radiation produced during this process is shown in Figure 1.3.

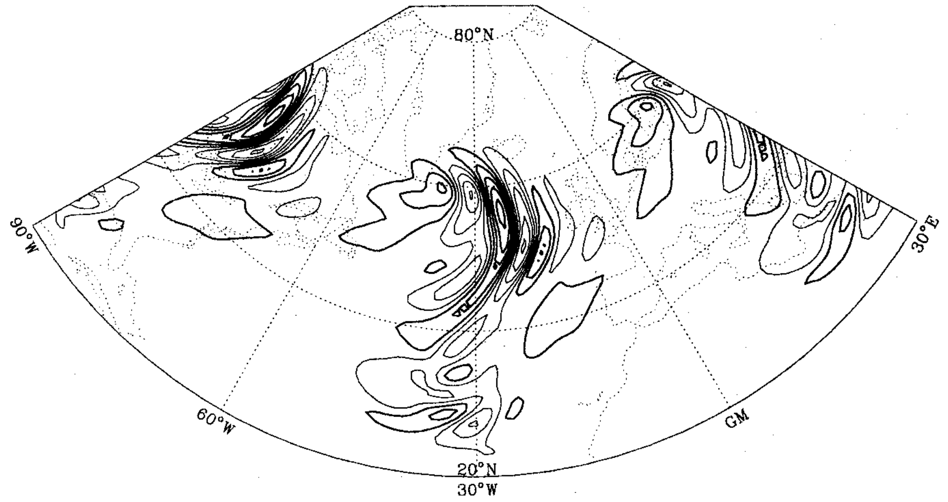


Figure 1.3: Contours of horizontal velocity divergence at 130 mb on day 11 of a high-resolution (T126, $\Delta z = 700$ m) numerical simulation by O'Sullivan & Dunkerton (1995). The model was initialized with a small-amplitude zonal wavenumber 6 mode, superimposed on a zonally-uniform baroclinically-unstable flow. By day 11, the perturbation has reached its maximum amplitude and begun to decay, generating the inertia-gravity waves shown.

The normal modes of a system will only appear if there is a mechanism present to excite them. There are two generation mechanisms for inertia-gravity waves in the free atmosphere (away from topography). Dalin's waves were generated due to a local wind shear, whereas O'Sullivan and Dunkerton's were radiated as excess energy, as a balanced mode lost and then re-established its balance through the geostrophic adjustment mechanism. It seems to be unclear which is the dominant of these two mechanisms in the atmosphere, but in Chapter 6 we will be in a position to determine the dominant mechanism in a laboratory experiment which is in many respects a scaled-down analogue of the atmosphere.

1.4.1 The importance of initialization

Despite the confirmed ubiquitous presence of inertia-gravity waves in the lower, meteorologically significant part of the atmosphere, a considerable portion of the effort that goes into producing weather forecasts is spent on the initialization process, in which inertia-gravity waves are filtered out of the observations before they are used as initial conditions. This is done for a number of reasons.

Firstly, atmospheric observations are often of such dubious quality that any apparent departures from balance are just as likely to be due to errors in the measurements, or in

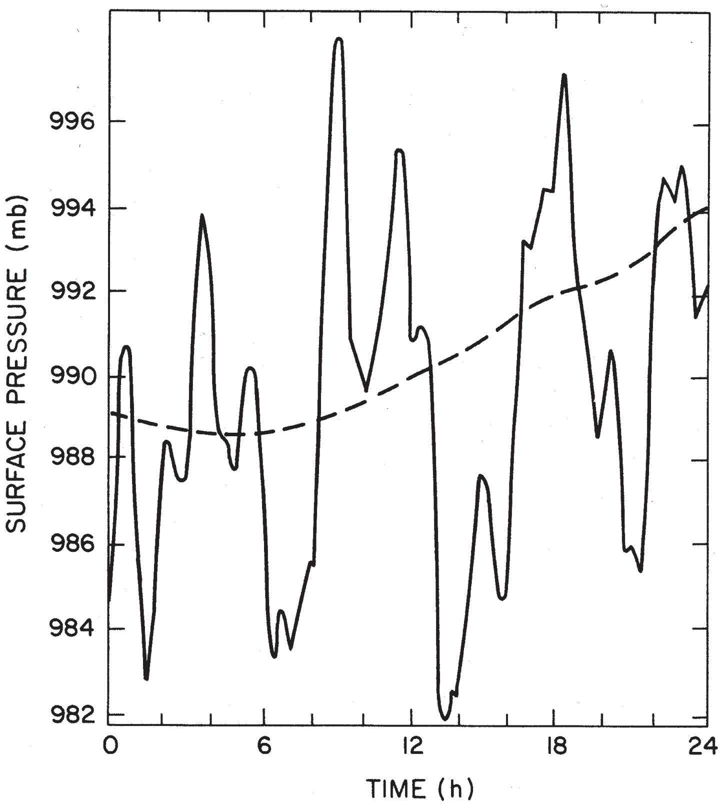


Figure 1.4: Time series of surface pressure at a fixed spatial point from a primitive equations model of the atmosphere, with uninitialized (solid) and initialized (dashed) starting fields. From Williamson & Temperton (1981).

the process which interpolates these measurements onto a regular grid, as to the presence of inertia-gravity waves. Secondly, even if inertia-gravity waves *are* the reason for departures from balance, typical observations are too sparse to fully resolve them, and the horizontal spatial resolution of numerical models is at present too low (11 km in the UK Meteorological Office mesoscale model) to accurately capture their evolution (but see Section 1.4.3). Furthermore, complete and permanent elimination of the fast modes permits the use of a significantly longer timestep whilst still satisfying the CFL criterion, which leads to more efficient integrations. Finally, the key aim of operational meteorologists is to model the synoptic-scale phenomena that constitute our weather systems, rather than trouble themselves with small-scale waves whose effects are secondary. For these reasons, it is desirable to have an initial atmospheric state which has no projection onto inertia-gravity modes, and so the observations are projected onto the slow manifold before being used as the initial conditions for the forecast.

An example of the likely consequences of running an uninitialized primitive equations model is shown in Figure 1.4. Using uninitialized starting conditions, based on raw observational data, the time series contains a large and spurious high-frequency inertia-gravity wave component which dominates the signal. Performing the same run but using initialized starting conditions, obtained by slightly adjusting the observations in a prescribed way to take them towards balance, completely removes the high-frequency component leaving only the slow trend of meteorological interest.

The issue of whether inertia-gravity waves in the atmosphere can affect the development of synoptic-scale weather systems is therefore an important and topical problem in meteorology, since if such an interaction exists it could place a fundamental limit on the accuracy of forecasting models which do not incorporate it. Operational meteorological centres include a parameterization of the effects of inertia-gravity waves in their forecasting models, in particular of their drag on the large-scale flow (Hines, 1997), but this is certainly an incomplete representation of the full interaction.

1.4.2 History of numerical weather prediction

Atmospheric inertia-gravity waves are not readily excited, since the length scale of the differential solar forcing between equator and poles is well-matched to typical Rossby mode wavelengths, but not to inertia-gravity wavelengths. Moreover, inertia-gravity modes are more efficiently dissipated since viscous effects are highly scale-selective. Therefore the vast majority of the energy of the atmosphere is in the vortical modes, and the atmospheric state is close to the slow manifold.

We can capitalize on this when constructing a model for numerical weather prediction. There would be a large redundancy in a model which captured the dynamics of solutions far from the slow manifold if it was known that the phase space trajectory would never visit there. To overcome this, we can either filter the fast modes out of the initial conditions, as described in Section 1.4.1, or alternatively filter the fast mode solutions out of the equations themselves by projecting the equations onto the slow manifold. We now briefly review the history of numerical weather prediction, in which both of these methods have been attempted, in order to illustrate the importance and difficulty of eliminating unbalanced modes from the forecast.

People have been fascinated by the weather since time immemorial, and have often made casual forecasts based on simple local observations such as the colour of the night sky. The first regular and systematic forecasts were those produced by Admiral Robert Fitzroy in 1860, which he published in the Times (Lynch, 2001). Forecasting methods employed then included maintaining a large catalogue of daily weather maps from the past, in order, and then finding a map which resembled the conditions for the present day and supposing that the atmosphere would do again what it had done before. Lorenz (1969) has shown that with this method, known as the *method of analogues*, it is difficult to find an acceptably close match even if the catalogue consists of 2000 maps.

Meteorology was finally given a firm scientific basis when when Bjerknes (1904) suggested the then revolutionary idea of solving partial differential equations to calculate the future weather. Two decades later this method was actually implemented by Richardson (1922), who performed the world's first numerical weather forecast. He integrated the discretized partial differential equations numerically on paper and, according to Lynch (1993), took two years to obtain his solution. He used the hydrostatic primitive equations (HPEs) — a slightly approximated form of the Navier-Stokes equations on a sphere, containing both fast and slow modes. He obtained completely unrealistic values (a surface pressure change of 145 mb in 6 hours) because the initial fields contained a significant but spurious inertia-gravity wave component, as in Figure 1.4, and not because of errors in his method.

The world's first computer forecast was performed by Charney, Fjörtoft and Von Neumann three decades later. In order to avoid Richardson's problem, they filtered the equations of motion to derive the quasi-geostrophic (Q-G) system (Charney, 1948) and their forecasts were reasonably successful.

The HPEs were used again a few years later, since they were shown to be more accurate than the Q-G equations. But in order to avoid Richardson's error, it was necessary to initialize the starting fields. Determining the most suitable initialization method formed a major area of research during the decades which followed. The following initialization methods have been proposed.

Hinkelmann (1951) suggested that the initial state should be modified so as to be geostrophically balanced, and then Charney (1955) suggested that it should instead satisfy the nonlinear balance equation, a diagnostic relation between the wind field and the geopotential. Phillips (1960) suggested that an even better initialization would result if the horizontal divergence of the initial wind were set equal to that implied by Q-G theory.

The technique of *dynamic initialization* was introduced by Miyakoda & Moyer (1968), and is discussed in the initialization review article by Lynch (1986). This involves integrating the raw observation data first forwards and then backwards in time, using the HPEs with enhanced dissipation. This forward/backward cycle is repeated several times to obtain fields in which the high frequency components have been damped out by the dissipation, and the fields so obtained are used as the initial conditions for the forecast.

Initialization techniques based on normal mode decompositions are discussed in the review article by Daley (1980). The technique of *linear normal mode initialization (LNMI)* was tested by Williamson (1976). The initial fields are separated into Rossby and inertia-gravity wave components (that is, projected onto the linear normal modes as in Section 1.3) and the amplitudes of the latter are set to zero. Unfortunately the inertia-gravity waves soon re-appear in the forecast: the primitive equations are nonlinear and evidently allow the Rossby components to interact in such a way as to generate new inertia-gravity waves.

The technique of *nonlinear normal mode initialization (NNMI)* was suggested independently by Machenhauer (1977) and Baer (1977) and Baer & Tribbia (1977). This involves not setting the initial inertia-gravity wave components to zero, but instead setting their initial rate-of-change to zero. NNMI takes into account the nonlinear nature of the equations. It works very well: the forecast is very smooth and the spurious inertia-gravity waves remain acceptably small throughout the integration. NNMI is the most popular method of initialization today, and it is used in many forecast centres.

Most recently, the technique of *digital filtering* has been suggested by Lynch (1991) as a simpler method than NNMI. It involves carrying out two short HPE model integrations starting with the raw data, one forwards in time and one backwards. This gives a time series (of typically 6 hours) centred on $t = 0$ for each model variable at each grid-point. A low-pass filter is then applied to these time-series, and the resulting values at $t = 0$ are used as the initial conditions. An optimal filter shape has been chosen and successfully implemented by Lynch (1996). Interestingly, Lynch (1999) has repeated Richardson's manual calculation on a computer, and reproduced the surface pressure tendency of 145 mb in 6 hours. In the same paper, he repeats the analysis after initialization of the starting fields with a digital filter, and obtains a realistic pressure tendency of 3 mb in 6 hours.

1.4.3 When might inertia-gravity waves be resolved?

The resolution of operational numerical weather prediction models has improved dramatically since the first ever computer forecasts were performed. The resolution is now sufficiently high that medium- to large-wavelength inertia-gravity waves can be partially resolved. Given a further reduction in the horizontal grid spacing by a factor of 10, with a corresponding improvement in observations, such waves could be fully resolved and potentially included explicitly in a forecast, rather than through an implicit parameterization of their effects.

If latitudinal and longitudinal grid spacings were each to drop by a factor of 10, we would also need to reduce the timestep by a factor of 10 to leave the Courant number unaffected. This would lead to an increase in computer time for a forecast by a factor of around 1000, or alternatively, we would need an increase in processing speeds by a factor of 1000 for the integration to take the same time. Since this factor corresponds to 10 doublings of clock speeds ($2^{10} \approx 10^3$) and since clock speeds double around every 18 months (Moore's law), this is expected to be achievable in around 15 years.

When this time comes, it would be useful for meteorologists to have a body of knowledge regarding the anticipated impacts of the inertia-gravity waves on the larger-scale flow. It is hoped that this thesis will form a small part of that collection of information, and that the present study will suggest directions for the research that will be needed between now and then to produce the rest.

1.5 Dynamical similarity and laboratory experiments

It is well-known (Douglas & Gasiorek, 2000) that, when written in non-dimensional form, the equations which govern the evolution of seemingly different fluid dynamical systems can be very similar. The aim of laboratory experiments in geophysical fluid dynamics is to exploit this *dynamical similarity*, to make inferences about atmospheric and oceanic phenomena from observations of the analogous laboratory flows.

As an example appropriate to this study, the shallow water equations (SWEs) applied to an approximated two-layer atmosphere or ocean, can very closely resemble the SWEs for a rotating, two-layer laboratory experiment. This statement holds despite the fact that typical length and time scales for corresponding geophysical and laboratory flows can differ by very many orders of magnitude. All that matters for dynamical similarity is equality of the relevant *non-dimensional* dynamical and geometrical parameters, such as the Rossby number and aspect ratio.

Dynamical similarity allows us to study geofluids in the laboratory, as suggested by Figure 1.5. Once we have solved a particular fluid flow problem by making observations in the laboratory, we have actually solved an infinite number of other fluid flow problems all of which are dynamically and geometrically similar, including on the planetary scale. For most of the remainder of this thesis, we study experiments in an isothermal, rotating, two-layer laboratory apparatus. In the final chapter we return the focus to geofluids to consider how portable our laboratory conclusions are to the atmosphere and ocean, mindful of the different non-dimensional parameters and boundary conditions we encounter in the laboratory.

Figure 1.6 shows a foretaste of the sort of flow we can observe using the present laboratory apparatus. In the image, baroclinic instability has led to the growth of a slow, large-scale mode with a dominant azimuthal wavenumber of two. Importantly for our purposes, two groups of fast, small-scale waves have developed and are superimposed onto the larger-amplitude baroclinic wave. Note the striking resemblance between small-scale waves in the laboratory experiment (Figure 1.6), and those in the atmosphere (Figure 1.2) and a numerical model (Figure 1.3). That waves on such dramatically different scales can appear so similar in form is testament to the power of the concept of dynamical similarity.

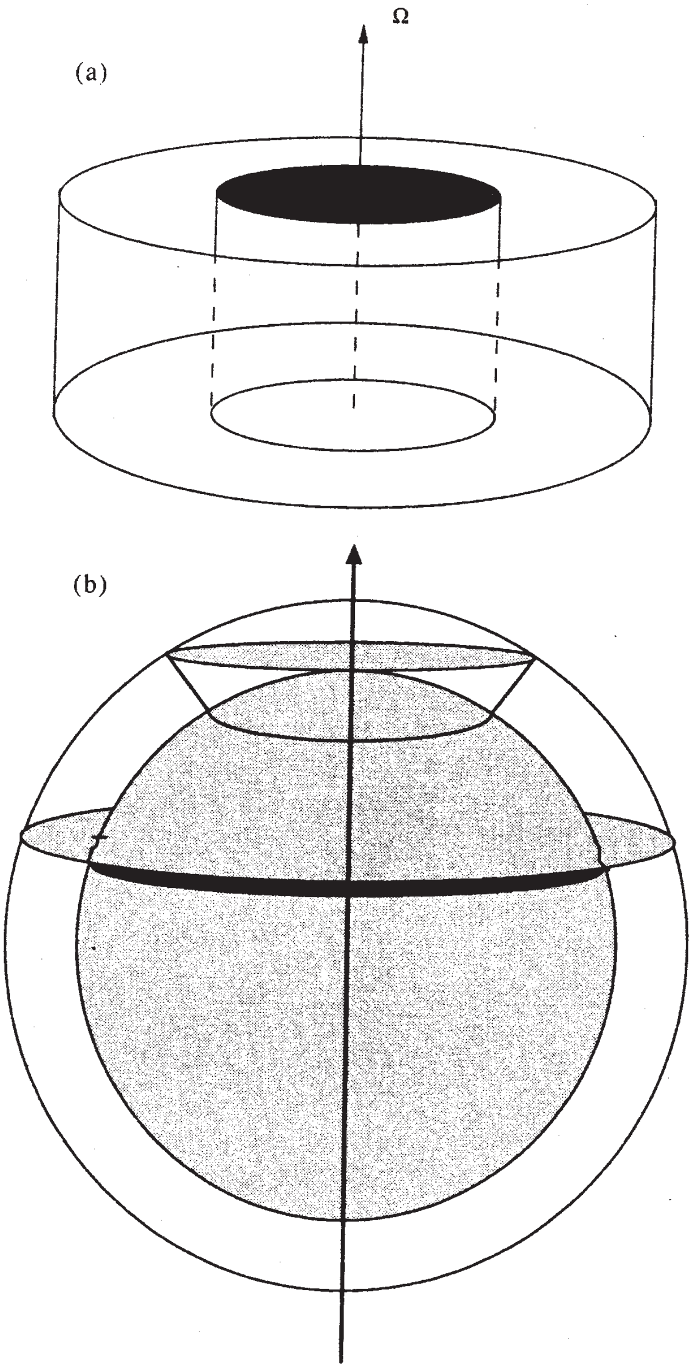


Figure 1.5: Diagram showing the analogy between (a) the fluid in a rotating annulus experiment in the laboratory, and (b) the fluid bounded by two latitude circles on a rotating planet. From Read et al. (1998).

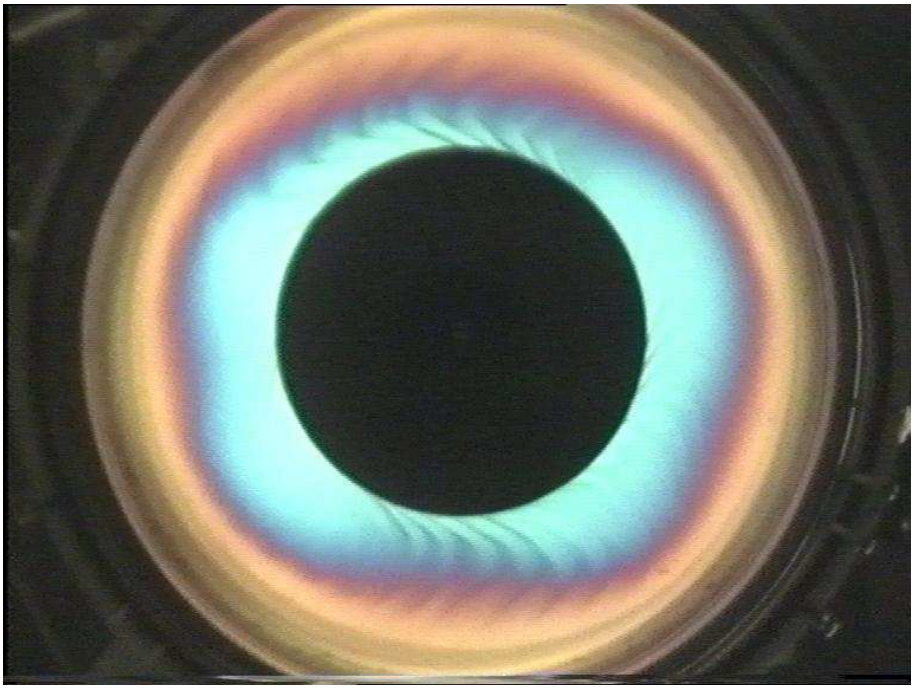


Figure 1.6: Sample laboratory experiment image. A large-scale baroclinic wave is shown to coexist with two groups of small-scale waves. There is a relationship, to be determined in Chapter 3, between the colour in the image and the height of the two-layer liquid interface.

1.6 Previous laboratory work

Laboratory investigations of non-rotating fluid flows began in the nineteenth century, and include the classic investigations of Reynolds (1883). At around the same time, Vettin (1884) became probably the first person to exploit dynamical similarity by carrying out rotating laboratory experiments as analogues of geophysical systems. He studied the surface flow in a rotating dishpan of fluid with a lump of ice near the centre, representing a polar ice cap, and (to the scorn of his contemporaries) he drew meteorological conclusions from his results.

The main benefits of studying geofluids indirectly in the laboratory are that the system is under the complete control of the experimenter, that global high-resolution measurements can be systematically taken, and that experiments can be repeated as many times as required. None of these statements hold when geofluids are studied directly rather than in the laboratory. A review of the role of laboratory experiments in geophysical fluid dynamics is given in Hide (1977).

As suggested by Vettin's experiment, for the closest resemblance between annulus and planet in Figure 1.5 we should apply heating and cooling at the outer and inner vertical sidewalls, respectively, to mimic the differential solar thermal forcing between equator and pole. The resulting *rotating thermal annulus* system with continuous fluid stratification has been extensively studied since the early 1950s, including the classic experiments of Hide et al. (1977).

It follows from the thermal (and gradient) wind balance equations for a rapidly-rotating annulus, that a radial temperature gradient will be accompanied by a vertical shear in the zonal velocity (such as that associated with the tropospheric jetstream in the atmosphere). Similar flows to those obtained in the thermal annulus can therefore be obtained in an isothermal annulus by imposing a velocity shear directly. For studying geoflows, an alternative to the thermal annulus is therefore the *rotating two-layer annulus*, with a shear provided across the fluid interface by differentially-rotating top and bottom horizontal boundaries.

The rotating, two layer annulus has also been studied extensively in the laboratory (e.g. Carrigan, 1978; King, 1979b; Appleby, 1982), and good agreement has been reached between the properties of balanced flows in the thermal and two-layer annuli, and of those in the corresponding theoretical and numerical models (Klein, 1990). During the two-layer annulus laboratory experiments of Lovegrove (1997), which were designed to investigate bifurcations between different large-scale modes, it was found under certain circumstances that fast, small-scale modes could develop. This finding, of secondary interest to Lovegrove's study and so not investigated in detail, has formed the starting point for the current work.

1.7 The current study

Lovegrove's findings showed that the interaction between balanced large-scale and unbalanced small-scale waves could be studied, for the first time, in the laboratory. Previous investigations had been exclusively based on highly-idealized and truncated numerical and theoretical models, whose representation of reality was far from complete.

This thesis is a study of scale-separated interactions in fluids in the laboratory, with the underlying motive being to better understand such interactions in geofluids such as the atmosphere and ocean. In the past, there have been laboratory investigations of large scale interactions (Section 1.6), and analytical investigations of scale-separated interactions and the slow manifold (Section 1.3). This study is believed to be the first laboratory investigation of the slow manifold.

1.7.1 Aims of this investigation

The key scientific questions we wish to answer are:

- **Under what circumstances do small-scale waves appear in the laboratory experiments?** We answer by undertaking a new series of experimental investigations.

- **Which mechanism causes the small-scale waves to appear in the laboratory experiments?** We answer by computing several radiation diagnostics associated with the candidate mechanisms. We do this using velocity data from a numerical model.

- **What are the effects of the laboratory small-scale waves on the large-scale, balanced flow?** We answer by comparing model (filtered) and laboratory (unfiltered) regime diagrams, and by performing model runs both with and without a stochastic inertia-gravity wave parameterization.

- **Having answered these questions for a laboratory experiment, what can we infer about answers to the analogous questions for geofluids?** We answer by comparing and contrasting the laboratory experiment with the atmosphere and ocean, including a comparison of nondimensional parameters and of boundary conditions.

1.7.2 Thesis outline

This thesis is split naturally into two parts. The first part describes the outcome of a laboratory investigation using the rotating, two-layer annulus. The apparatus is described in Chapter 2, and a method for calibrating interface heights is developed and applied in Chapter 3. In Chapter 4, the results of an extensive series of experiments are described and analyzed.

The second part of the thesis describes the outcome of a complementary method for investigating the rotating annulus, by running a purpose-built numerical model. In Chapter 5, we derive the model equations and obtain a reliable and fast integration scheme. In Chapter 6, we describe the results of the model runs, and compare them with the laboratory results. Then, in Chapter 7, we add a stochastic inertia-gravity wave parameterization to the model, and investigate its effects upon the large-scale flow.

Finally, in Chapter 8, we summarize the present work, and give our conclusions about the scale-separated wave-wave interaction in the laboratory by stating answers to the above four questions. We discuss the applicability of these conclusions to the analogous interaction in geofluids, and end by describing some possible avenues for future work.

Chapter 2

Description of the laboratory apparatus

"d-limonene ... can be harmful when vaporized and breathed."

US Environmental Protection Agency website

"...the primary ingredient of Citrus Burst ®, d-limonene, is plant derived. It is extremely safe..."

Florida Chemical Company, Inc. website

In this chapter, a description is given of the rotating, two-layer annulus apparatus which has been used for the laboratory component of this study. The apparatus was built at the U.K. Meteorological Office in the early 1970s, where it was used in the studies of King (1979b) and Appleby (1982). The apparatus was later moved to the University of Oxford, where is was used most recently by Lovegrove (1997).

In the following sections, we describe some modifications which have been made to the apparatus since the experiments of Lovegrove. These include an upgrade to a higher-quality video format for image transmission and storage, and the installation of a higher-resolution frame-grabber in the laboratory computer. We present some new results regarding the sensitivity of the working fluid properties to temperature fluctuations. Then we develop a simple method for reducing the mutual interfacial tension, and we discuss some previously unreported but important observed changes as the fluids age. Details of the employed flow visualization technique are given, and of the extent to which the visualization is distorted by refraction and parallax effects.

2.1 The rotating, two-layer annulus

Photos of the rotating, two-layer annulus apparatus are shown in Figures 2.1 and 2.2, and the annulus is shown schematically in Figure 2.3. The annulus consists of a cylindrical stainless steel tank of inner radius 125.00 mm and depth 250.00 mm, which has a fixed glass base and a removable glass lid. A solid steel cylinder, of radius 62.50 mm and depth 250.00 mm, is glued coaxially (to within an estimated 0.1 mm) to the base of the tank to form an annulus of gap width 62.50 mm. The annular region is filled to the brim with equal volumes of two immiscible liquids, to give a two-layer liquid with a well-defined interface and equal resting layer depths of 125 ± 1 mm. Details of the particular fluids used are given in Section 2.2.

The tank is mounted centrally (to within an estimated 0.1 mm) above a parabolic correction tank (discussed in Section 2.4) on a circular turntable 1 m above the laboratory floor, which can be made to rotate under computer control with angular velocity Ω. The annulus lid, which is in contact with the upper liquid, can be made to rotate under computer control with angular velocity $\Delta\Omega$ relative to the tank. This is possible because the lid is connected to the tank via a ball race, allowing low-friction relative motion powered by a servo motor and drive wheel. Both Ω and $\Delta\Omega$ can take either sign, and are stable to within 1% over a period of a few hours. The maximum achievable values are $|\Omega_{\mathrm{max}}| = 6.3$ rad s^{-1} and $|\Delta\Omega_{\mathrm{max}}| = 3.1$ rad s^{-1}.

There is a central circular hole in the turntable, of radius equal to the tank radius, so that white light from a bright 500 W tungsten-halogen source lamp on the laboratory floor may pass vertically into, through and out of the annular gap. The light is received by a colour charge-coupled device (CCD) video camera, which is on the rotation axis and co-rotates with the turntable 2 m above it. Communications between the laboratory frame and the rotating turntable frame (namely the camera power and output signal, and the servo motor power) are achieved through a commutator slip-ring, hidden from camera view by the inner cylinder.

Figure 2.1: The two-layer annulus apparatus and control equipment. From left to right: the rotating turntable, mounted with the annulus and with a tall metal frame to support the video camera; an amplifier which powers the turntable rotation and lid rotation; a television and video recorder to watch and record live images from the camera; and a computer with a frame-grabber installed to digitize and save selected images.

Figure 2.2: Close-up view into the annular tank from above.

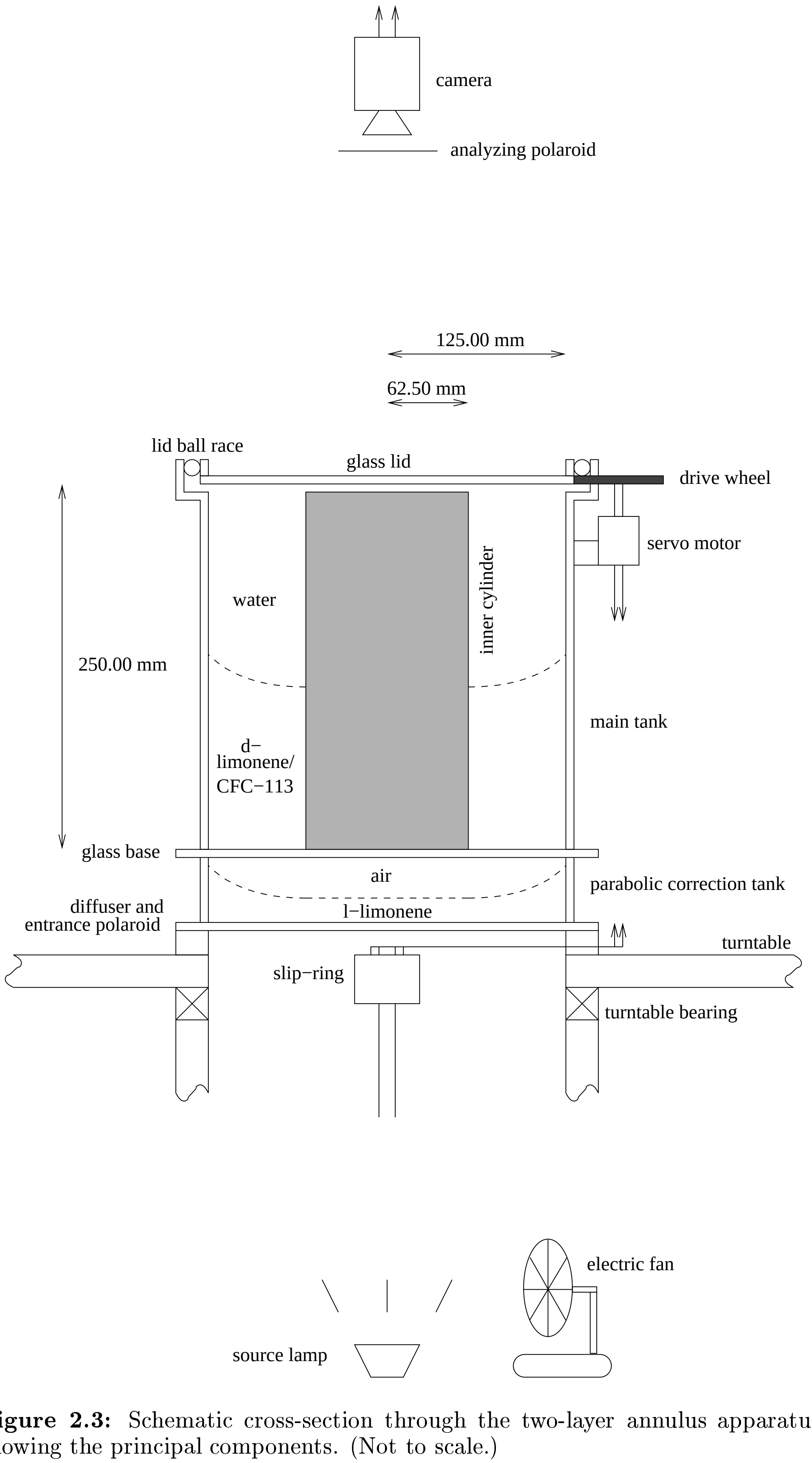


Figure 2.3: Schematic cross-section through the two-layer annulus apparatus, showing the principal components. (Not to scale.)

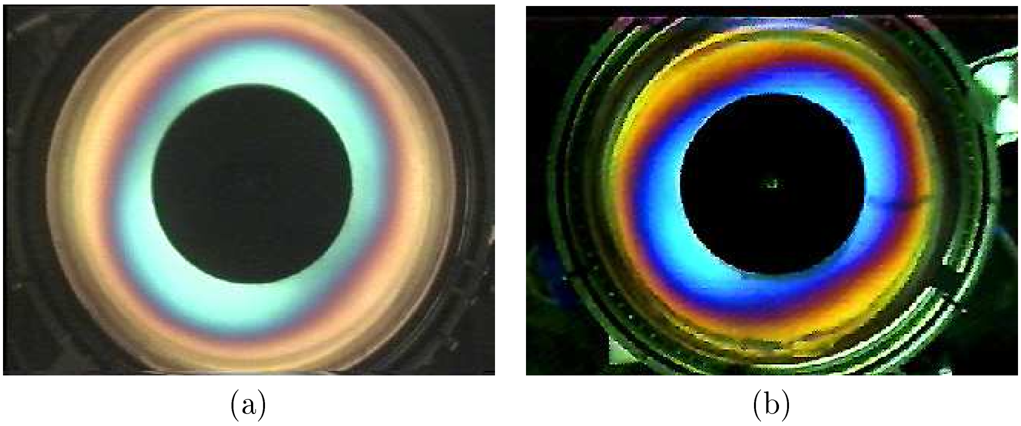

(a) (b)

Figure 2.4: Images captured from a video tape of the two-layer annulus experiment, photographed by the CCD camera. Digitization of the images was done using (a) the current frame-grabber and S-VHS equipment, and (b) the previous frame-grabber and regular VHS equipment.

The S-VHS signal[1] output by the camera can be recorded at 25 frames per second onto high-quality, S-VHS video tapes, which can be played back for subsequent analysis. The video signal can be input to a computer with a frame-grabber, to produce colour 24-bit digitized images measuring 768 pixels by 576 pixels. A sample image is shown in Figure 2.4(a), showing a wavenumber 2 mode which has grown due to baroclinic instability and which slowly drifts around the annulus with a period of around one minute. Different colours correspond to different depths of the lower liquid layer, for reasons to be discussed fully in Section 2.3. For reference, an image of a similar flow, captured using the older frame-grabber and ordinary VHS equipment used by Lovegrove et al. (2000), is shown in Figure 2.4(b). The image size is 320 pixels by 240 pixels, and so the use of the new frame-grabber has multiplied the resolution by a factor of 2.4 in each dimension.

2.2 Fluid properties

We choose to use water as the liquid for one of the layers. The water is first purified by being passed through a de-ionizing filter, as this slows down mould growth in the apparatus. The liquid for the other layer must then be hydrophobic, so that the two layers are mutually immiscible and give a well-defined interface. It should be transparent and colourless, to allow the passage of light without significant absorption. It should

---

[1] S-VHS, or Super-VHS, is a professional master grade version of regular VHS.

have an optical activity which is non-zero and which varies strongly with wavelength for visible light, in order for the flow visualization technique to be described in Section 2.3 to work. It would be advantageous if it had a density and viscosity close to that of water for comparison with theoretical two-layer models, many of which make assumptions of small stratification and equal viscosities (as discussed in the review article by Klein, 1990).

A liquid which satisfies most of these requirements is *limonene*[2], an organic oil which is distilled from orange peel. Unfortunately, its density (840 kg m^3) is significantly less than that of water. The limonene cannot be used as the upper layer, as it is highly volatile and rapidly evaporates to produce a harmful vapour (WHO, 1998), and the apparatus is not airtight. For this reason, the limonene is mixed with *CFC-113*[3], a heavier-than-water, colourless, hydrophobic, optically-inactive solvent, in such proportions that the composite liquid is slightly more dense than water. Water is then used as the upper layer liquid, thereby preventing harmful vapours from escaping into the laboratory. Some relevant physical properties of the liquids which make up the two layers are given in Table 2.1. The liquids are the same as those used by Lovegrove.

2.2.1 Thermal expansivities

It is seen from Table 2.1 that the two layers have quite different thermal expansivities. As the temperature rises, the lower layer becomes less dense more quickly than the upper layer. According to the values in the table, the ambient laboratory temperature needs to rise by only around 5 °C before the two layers have equal densities, leading to a possible Rayleigh-Taylor instability (Acheson, 1990) and layer inversion. In order to prevent this, the ambient laboratory temperature must be tightly controlled. An air conditioning system was used to achieve this, and was found to keep the temperature within 0.1 °C of 20.0 °C over a 24-hour period, as opposed to an observed fluctuation of 2.0 °C with the system switched off. The air conditioning system was permanently switched on during the experiments described in this thesis, to keep variations in the fractional density difference tolerably low, both during and between experiments.

---

[2] Limonene is also known by its synonyms *carvene* and *methylcyclohexene*.

[3] The full name is 1,1,2-trichlorotrifluoroethane.

	de-ionized water layer 1 (upper)	limonene/CFC-113 layer 2 (lower)
density, ρ (kg m^{-3})	997 ± 1	1003 ± 1
thermal expansivity, α (10^{-4} K^{-1})	2.07 ± 0.01	13.0 ± 0.1
mutual interfacial tension, S (10^{-2} N m^{-1})	2.85 ± 0.1	2.85 ± 0.1
kinematic viscosity, ν (10^{-6} m^2 s^{-1})	1.27 ± 0.02	1.08 ± 0.02
optical activity, ϕ, at 0.59 μm (° m^{-1})	0	770 ± 10
refractive index, n, at 0.59 μm	1.3328 ± 0.0001	1.4466 ± 0.0001

Table 2.1: Physical properties of the freshly-prepared working liquids at the ambient laboratory temperature (20.0 °C) and pressure. The densities and thermal expansivities were measured in the laboratory by the author using a density meter, which times the period of oscillation of an electromagnetically-excited sample tube in order to accurately calculate density. The mutual interfacial tension was measured in the laboratory by the author using a torsion balance, which measures the force required to pull a thin wire loop through the interface in order to determine the tension. The sources for the other property values are Lovegrove (1997), Hart & Kittelman (1986), Lide (1995) and Kaye & Laby (1995).

There is another problem associated with the large temperature sensitivity of the two-layer liquid, namely that the 500 W lamp directly beneath the annulus is not only an effective light source but also an effective heat source. During some of the initial experiments performed in this study, this heating was so strong that the liquids exhibited a spontaneous Rayleigh-Taylor inversion around two hours after the lamp was switched on. Because it was desired to run experiments which lasted for longer than two hours, an electric desktop fan was positioned on the laboratory floor to blow air across the lamp, as shown in Figure 2.3. The aim was to inhibit and destroy convection cells, in the hope that the majority of the heat transfer from lamp to liquids was via convection rather than radiation. This strategy worked, and experiments lasting three hours and longer could be carried out when the fan was switched on.

Lovegrove did not document taking these measures to control the liquid temperature, as he was probably not aware that the layer thermal expansivities were so different. There is therefore a possibility of reduced gravity drift in his experiments, which was not documented until now.

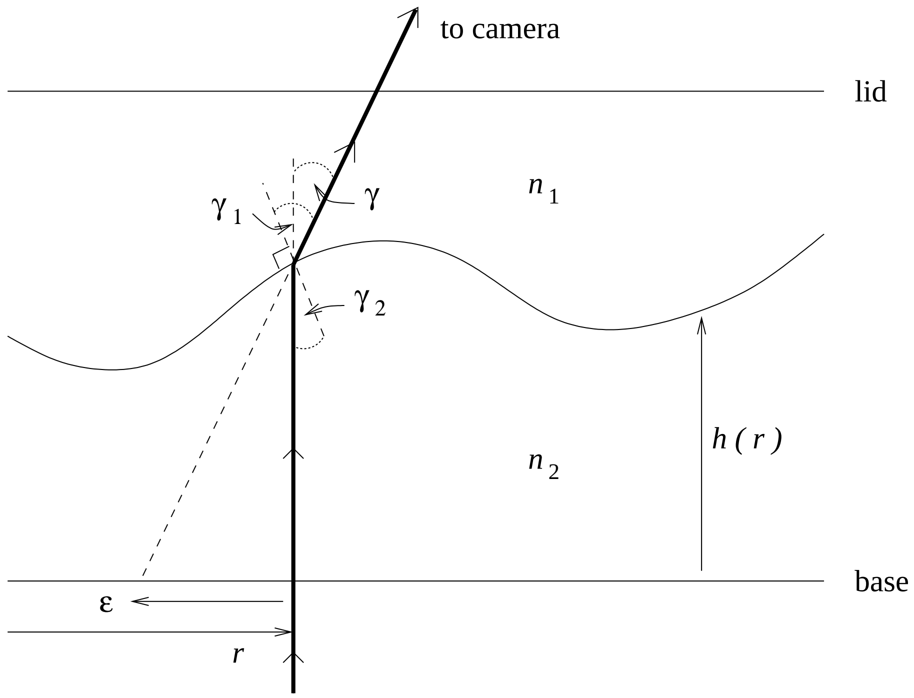


Figure 2.5: The path of a ray of light which enters the annulus vertically at radius r, where the interface height is $h(r)$ and the slope is $\mathrm{d}h/\mathrm{d}r$.

2.2.2 Refraction

The different refractive indices of the two layers (Table 2.1) will cause refraction to occur at the fluid interface. The images from the camera will therefore show a distorted version of the actual interface height shape, which means that the images output by the camera must be interpreted with caution.

Consider a ray of vertically-travelling light which enters the base of the tank at radius r, as shown in Figure 2.5. For this simplified analysis, the interface height $h(r)$ is taken to be a function of r only. The angle of incidence γ_2 is given by

$$\tan\gamma_2 = \frac{\mathrm{d}h}{\mathrm{d}r} \ , \tag{2.1}$$

and is related to the angle of refraction γ_1 by Snell's Law:

$$n_1 \sin\gamma_1 = n_2 \sin\gamma_2 \ . \tag{2.2}$$

Eliminating γ_2 between (2.1) and (2.2) gives

$$\sin\gamma_1 = \frac{n_2}{n_1}\frac{\mathrm{d}h/\mathrm{d}r}{\sqrt{1+(\mathrm{d}h/\mathrm{d}r)^2}} \ . \tag{2.3}$$

The angle through which the ray is deflected at the interface is then $\gamma \equiv \gamma_1 - \gamma_2$. Light which entered the base at radius r appears to the camera to have entered at apparent radius $r' \equiv r - \epsilon$, where

$$\epsilon = h \tan \gamma \ . \tag{2.4}$$

The camera therefore gives a distorted image of the annulus: if we observe a certain colour at a particular radius on a still from the camera, the radius at which the lower layer actually has the height corresponding to that colour is offset from the observation radius by an amount ϵ.

For a worst-case scenario of interfacial slope $\mathrm{d}h/\mathrm{d}r$=1, equations (2.1) and (2.3) give $\gamma_2 = 45.0°$ and $\gamma_1 = 50.1°$, so that the deflection angle is $\gamma = 5.1°$. At a point where $h = 12.5$ cm, equation (2.4) gives $\epsilon = 1.1$ cm, which is almost 20% of the annular gap width. It is important to use a certain amount of caution, then, when using images from the camera to infer distances in the annulus.

In practice there will also be azimuthal refractive distortion, neglected in this analysis, which will make no difference to wavelength determination but which will bias the kurtosis (peakiness) of the inferred wave shape. There will also be refraction at the glass lid, neglected here.

If it were necessary, we could compensate for radial refractive distortion by using equations (2.1), (2.3) and (2.4) to derive the transformation from apparent radius seen by the camera to actual radius. Such an analysis has not been carried out in this study, because the calculation to obtain the mapping becomes very involved; the maximum distortion is only a centimetre; and we are not concerned with locating exactly where the fast waves appear in the annulus.

2.2.3 Interfacial tension

The interfacial tension between the liquids (Table 2.1) is around four times as large as the values encountered in the two-layer experiments of King (1979b) and Appleby (1982), who used different working liquids from the present ones. The classical theoretical two-layer models first used by Phillips (1951) do not include the effects of interfacial

volume of washing-up liquid per unit surface area of water (ml m^{-2})	0	0.32	0.64	0.96
interfacial tension (10^{-3} N m^{-1})	28.5 ± 1.1	5.3 ± 0.4	1.4 ± 0.4	1.3 ± 0.4

Table 2.2: Dependence of interfacial tension on amount of added washing-up liquid, as measured in small samples of the working liquids using a torsion balance in the laboratory. A saturation limit is reached at around 0.6 ml m^{-2}. The quantity of added surfactant is given as a volume per unit surface area of the working liquid, rather than per unit volume, as it is assumed that all surfactant molecules are attracted to the surface rather than existing throughout the entire liquid volume, up to the saturation point.

tension. Moreover, interfacial tensions are non-existent in the atmosphere and are completely negligible in the ocean (except for motions on the very smallest of length scales). Because we would like to keep the laboratory experiment as close as possible to the simple two-layer models, and to atmospheric and oceanic flows, it is desirable to attempt to reduce it. We do this by adding a surfactant (*surface active agent*), a chemical whose molecules are made up of a water soluble and a water insoluble component, and which reduces the surface tension of water.

Various surfactants were tested in small (20 ml) samples of the working liquids.[4] Small added quantities of *Photo-Flo*, a surfactant used in the photo-developing industry, had the unexpected but reproducible effect of *increasing* the interfacial tension by around 30%. This is not entirely unexpected, as surfactants are defined with respect to a water-air interface, which is very different from the liquid-liquid interface in the annulus. Ordinary hand soap was tested but rejected, as it caused a significant cloudiness. The only surfactant to be successfully tested was ordinary washing-up liquid, and the results are shown in Table 2.2.

When washing-up liquid was added at 0.64 ml m^{-2}, the interfacial tension was reduced

---

[4] *In situ* measurements of interfacial tension in the annulus are not practicable using a torsion balance, and furthermore, we do not wish to contaminate the actual working liquids with the candidate surfactants.

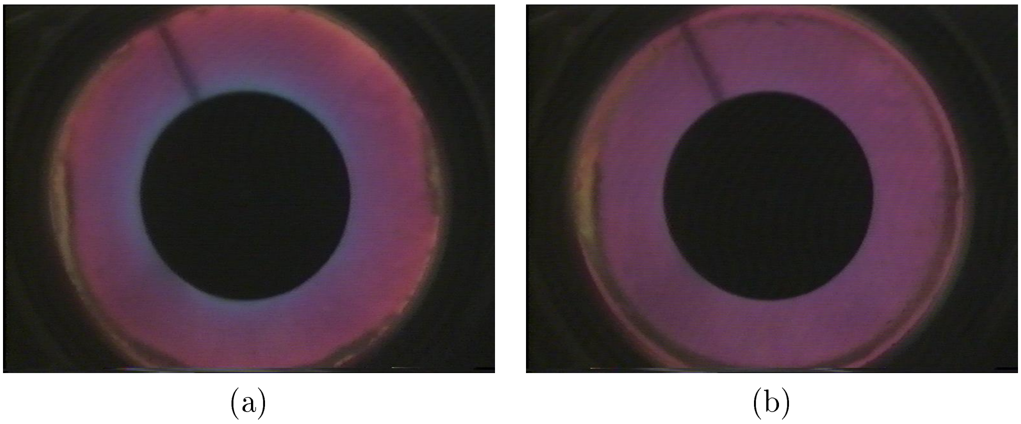

(a) (b)

Figure 2.6: Digitized experimental images, showing colours corresponding to the resting interface height shape (a) immediately before the addition of 0.12 ml of surfactant, and (b) a few seconds after. The dark radial spoke is a shadow due to the slip-ring connection wires shown in Figure 2.3. When the experiment is running and the interface height is not flat, this shadow is washed out by the dispersive effects of interfacial refraction (Section 2.2.2).

by more than an order of magnitude. However, this was found to facilitate the formation of bubbles of the lower-layer liquid in the upper layer near the interface, an effect which is undesirable. For all the experiments described in this thesis, washing-up liquid was added to the water in the annulus at 0.32 ml m^{-2} (i.e. a total of 0.12 ml), which we presume to give an interfacial tension of 5.3×10^{-3} N m^{-1}. A few seconds after the addition of the washing-up liquid, the resting interface shape recorded by the camera changed as shown in Figure 2.6. The interface became noticeably flatter, and the meniscus widths at the sidewalls were reduced, as expected. However, we will see in Section 2.2.4 that the interfacial tension will not necessarily remain at this reduced value as the fluids age.

2.2.4 Slow evolution of fluid properties with time

It has been casually observed in this study that the properties of the interface between the layers in the annulus seem to exhibit a long-term evolution as time passes. Direct observation into the tank through the lid revealed the slow formation of a skin at the interface, on timescales of weeks to months. Often this skin was so strong that it resisted piercing even by a sharpened pencil. This effect, previously unreported, suggests long-term chemical changes in the liquids, and possible accompanying variations in their physical properties. The values of the physical properties shown in Table 2.1 were all measured in freshly-prepared samples.

This long term interface evolution will be shown to have important consequences in Section 4.2. Various effects could explain the observed trends in interface properties:

- the pH of the water may be evolving as CO_2 and/or O_2 are exchanged with the laboratory air;
- the lower-layer liquid may be diffusing into the upper layer — both CFC-113 and limonene are actually sparingly soluble in water and are therefore expected to slowly diffuse out over time. CFC-113 has a solubility of 200 mg/litre in water at 20 °C (WHO, 2002), and limonene has a solubility of 13.8 mg/litre in water at 25 °C (Massaldi & King, 1973). Limonene has a strong, characteristic fruity smell which was observed to be taken on by the water as the fluids aged, consistent with the diffusion theory;
- there may be slow lower-layer interactions with the steel cylindrical container;
- the strong glue which fixed the inner cylinder to the base of the tank, and the black paint on the inner cylinder, were often found to have been corroded when the liquids were changed. This implies that the lower layer liquid is capable of slow dissolution of sealant and paint, which would alter its composition;
- there may be slow electro-chemistry with the walls (especially if the pH is changing);
- there may be chemical interactions with the added surfactant;
- any particulate matter in the laboratory air, such as dust, will tend to accumulate in the liquids over time as contaminants, and thereafter be possible candidates for dissolution.

Further work is needed to investigate which of these mechanisms is responsible for the observations.

2.3 Flow visualization

Previous flow measurement techniques in two-layer annulus experiments have included the capacitative method of Hart (1972), in which a thin vertical wire is inserted into the fluid across the interface. One of the liquids is chosen to be an electrical conductor, so that changes in interface height cause changes in the voltage drop along the wire, which can be measured and calibrated. A potential problem is that the wire can have a significant interaction with the flow (for example, as observed by Früh & Read, 1997). King (1979a) successfully applied a technique which exploited the absorption of an infra-red light beam by one of the layers, by an amount dependent upon the layer depth. This method is non-invasive, but like Hart's method, suffers from only returning measurements of interface height at one spatial location, rather than providing the global coverage which is desired. Shadowgraphy and Schlieren techniques (Goldstein, 1983) were the first to provide two-dimensional interface height fields, but were inaccurate and impractical.

The current visualization technique was first used by Hart & Kittelman (1986) in an open-cylinder experiment, and has since been used by Lovegrove in his annulus experiment. It provides non-invasive, global measurements of interface height which are highly-resolved in both space and time. The method relies on one of the liquid layers being optically active, in our case the lower layer due to the limonene. In order to understand how the technique works, we now track the passage of light through the apparatus from source lamp to video camera, via the two immiscible working liquids in the annular tank. This exercise is necessary to understand the relationship between the colour field recorded by the camera, and the interface height field.

As quasi-white light emitted by the source lamp travels vertically upwards through the apparatus shown in Figure 2.3, it first passes through a *diffuser*. This is a translucent plastic circular sheet of thickness 1 cm and radius equal to the tank radius, which is centrally mounted on the turntable. Its purpose is to diffuse the incoming light such that it illuminates the base of the tank uniformly. Without the diffuser, the video images would contain contrasting bright and dim regions, which would make interpretation and analysis more difficult.

The diffuse light next passes through an *entrance polaroid*, fixed to the upper side of the diffuser. This is a thin (1 mm) circular sheet of linearly-polarizing filter, of radius equal to the tank radius. The direction of its polarization vector determines a vertical plane of polarization for the emerging light. Importantly, the entrance polaroid is fixed to the rotating turntable. This means that the vertical polarization plane of the light entering the fluids will rotate in the laboratory frame, but is fixed in the camera frame.

The plane-polarized light next encounters the *parabolic correction tank*, whose purpose is discussed in Section 2.4.

Next, the light enters the *main tank* via its glass base, and travels through the optically-active lower layer liquid, whose effect is to rotate the plane of polarization of the light. This happens because limonene molecules are chiral, i.e. not superimposable on their mirror image (March, 1992), which is the cause of the optical activity. The amount of rotation depends on both the wavelength of the light and the depth of the liquid traversed (the latter dependency is one of proportionality, for depths greater than the molecular scale).

The rotation angle per unit depth for pure limonene has been determined experimentally by Hart & Kittelman (1986), for a range of wavelengths spanning the visible part of the spectrum. We can derive the rotation for the lower layer limonene/CFC-113 liquid by assuming that the rotation angle is reduced by a fraction equal to the volume-fraction of CFC-113 in the composite mix. This assumption is easily verified theoretically by taking the total rotation angle to be the same whether the constituent liquids are well-mixed or are separated into distinct layers, and has also been verified experimentally by Hart & Kittelman (1986). The resulting optical activity curve is shown in Figure 2.7.

Next, the light travels through the optically-inactive upper layer and leaves the tank via the glass lid, during which its plane of polarization is unchanged. The light then passes through an *analyzing polaroid*, which is a second thin sheet of linearly-polarizing filter fixed in front of the camera lens. This polaroid only allows the transmission of a certain fraction of the incident light intensity. This fraction varies from 1 if the analyzing polaroid axis and incident light polarization axis are parallel (or anti-parallel), to 0 if they are perpendicular (assuming perfect polaroids). For a given lower layer depth, therefore, only certain wavelengths will be rotated into close alignment with the analyzing polaroid and be transmitted to the camera. Other wavelengths will be extinguished

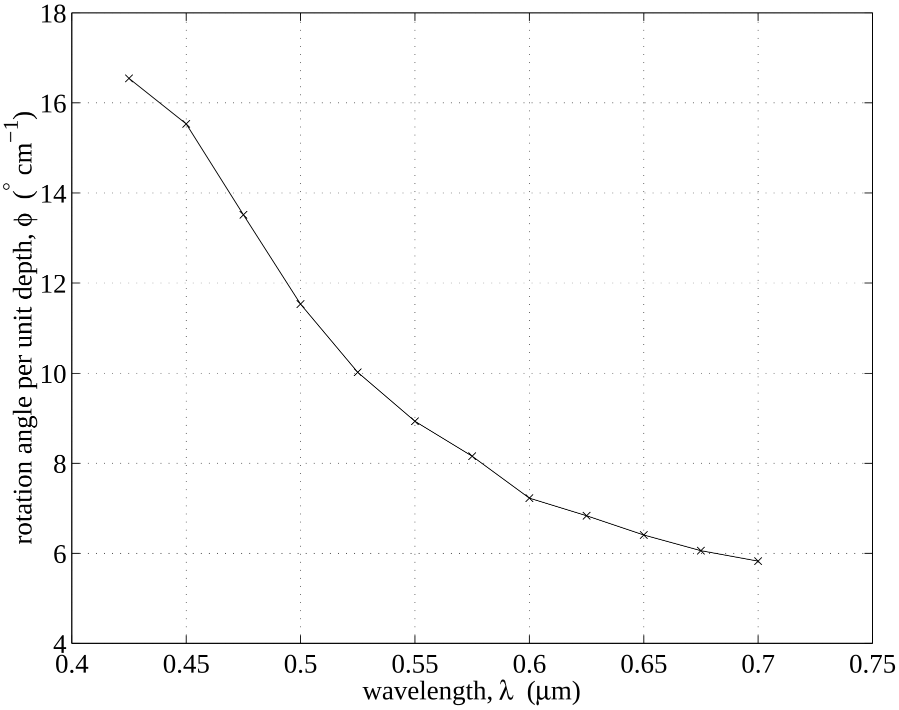
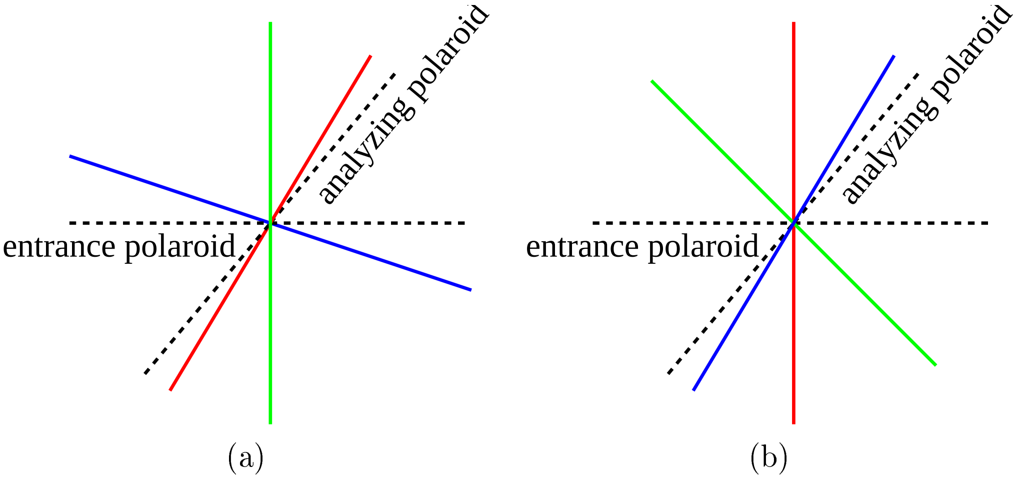


Figure 2.7: Optical rotary dispersion curve for the lower layer liquid of the present laboratory experiments, following Hart & Kittelman (1986).

by the polaroid. This is the origin of the causal relationship between lower layer depth, and colour recorded in the video images.

We now give an example which should help to elucidate the ideas of the previous paragraphs, by deducing qualitatively which colour will be the dominant one received by the camera for a given lower layer depth. Suppose that white light travels through a depth 10 cm of the lower layer (ignoring the parabolic correction tank). Then, from Figure 2.7, the red light component ($\lambda \sim 0.70\,\mu$m) will be rotated through an angle of about 60°, the green light ($\lambda \sim 0.55\,\mu$m) through 90° and the blue light ($\lambda \sim 0.44\,\mu$m) through 160°. These angles are shown in Figure 2.8(a), where the angle between the axes of the crossed entrance and analyzing polaroids is taken to be 50°. Most of the red light will be transmitted through the analyzing polaroid, plus some of the green light but hardly any of the blue light, and we therefore expect to see a red colour. Figure 2.8(b) shows the equivalent analysis for a lower layer depth of 15 cm, where we expect to see a predominantly blue colour. In Chapter 3, we quantify this analysis to derive the mathematical relationship between observed colour and height, allowing a calibration of the experiment.



(a) (b)

Figure 2.8: The polarization axes of red, green and blue light, after travelling through (a) 10 cm, and (b) 15 cm of the limonene/CFC-113 mixture, as viewed from above by the video camera. The entrance and analyzing polaroids are shown as dashed lines.

By rotating the analyzing polaroid attached to the camera lens, the angle between the axes of the crossed entrance and analyzing polaroids can be adjusted, which varies the colour observed for a given interface height. For some angles, the relationship between height and colour is more sensitive than for others, i.e. small changes in height produce relatively large changes in colour. Before any experiments were performed in this study, the differential angle was adjusted — by trial and error — to be such that the relationship was at its most sensitive, so that even very small changes in interface height produce a significant signal in the colour field. This is important if we are to be able to register the small-amplitude interfacial short waves.

2.4 The parabolic correction tank

Suppose that the apparatus so far described — without the parabolic correction tank — were set into rotation with turntable angular velocity $\Omega \neq 0$ and differential lid rotation $\Delta\Omega = 0$. After spin-up of both layers to angular velocity Ω, a balance is established between the radial pressure gradient force and the centripetal force in the fluid, and then the equilibrium height $h(r)$ of the fluid interface above the base of the tank is given (see also equations 3.6 and 5.22) by

$$h(r) = \frac{\Omega^2 r^2}{2g} + \text{constant} \ , \tag{2.5}$$

where r is the distance from the rotation axis and g is the acceleration due to gravity (*not* the reduced gravity, as the centripetal effect does not depend upon any density contrast between the layers). When the turntable is rotating at its maximum rate of $\Omega = 6.3$ rad s^{-1}, the associated interface height change across the annulus will be $h(12.5 \text{ cm}) - h(6.25 \text{ cm}) \approx 2$ cm. The interface is therefore not horizontal and colour gradients are seen by the camera. If a velocity shear is applied by rotating the lid, the interface height will be perturbed away from the parabolic shape given by equation (2.5), as described by equation (5.22). This manifests itself in the experiments of Hart & Kittelman (1986) as a reddening of their images at large radii.

Deciding that this effect was undesirable, Lovegrove (1997) devised a way of eliminating this background parabolic distortion from the view seen by the camera. His method relies on the fact that limonene exists in two different optical isomers: a dextrorotary isomer (*d*-limonene) which rotates plane-polarized light in the clockwise sense, and a laevorotary isomer (*l*-limonene) which rotates it by an equal angle in the anti-clockwise sense.

Lovegrove used *d*-limonene for the lower layer of the main tank, and introduced a second "parabolic correction" tank containing *l*-limonene and air, directly beneath it as shown in Figure 2.3. In equilibrium with $\Delta\Omega = 0$ and $\Omega \neq 0$, the surface of the *l*-limonene layer takes up exactly the same concave upwards parabolic shape as the interface in the main tank (equation 2.5), even though its domain is cylindrical rather than annular and it is covered with air rather than water. Clockwise light rotations in the main tank are then exactly cancelled out by anti-clockwise rotations in the parabolic correction tank, and so there are no colour gradients in the images output by the camera. When $\Delta\Omega \neq 0$, the camera then shows colours which correspond to the *deviation* of the interface height away from this basic parabolic shape, avoiding the background colour gradients of Hart & Kittelman (1986).

Since the *d*-limonene is diluted with CFC-113 to increase its density, the optical activities of the composite lower liquid in the main tank and the pure *l*-limonene in the parabolic correction tank will not be exact opposites, and the statement above will only be approximately true. It seems that the reason Lovegrove decided not to dilute the *l*-

limonene with CFC-113, which would have given liquids with exactly equal and opposite optical activities, was that the CFC is highly volatile and would have quickly evaporated without a protective covering of water.

2.5 Parallax effects

The diameter of the inner cylinder in an experimental image (e.g. Figure 2.6) is 310 pixels. We calibrate horizontal distances in the images by equating this to the physical diameter of 125 mm. This gives the side length of the (square) pixels to be 0.40 mm, which defines the horizontal resolution of the images.

Importantly, this length calibration takes place in the horizontal plane containing the annulus lid. Radii that we infer from an image, by converting distances from the annulus centre in pixels to distances in mm, will therefore correspond to radii at the lid. Because the camera is a finite distance away from the annulus (200 cm from the base), light paths from annulus to camera are not exactly vertical, and the radius at the lid r_{lid} will not be the same as the radius r at the fluid interface. Neglecting refraction, the relationship between them is given by geometry to be

$$\frac{r}{r_{\mathrm{lid}}} = \frac{200\ \mathrm{cm} - 12.5\ \mathrm{cm}}{200\ \mathrm{cm} - 25.0\ \mathrm{cm}} = 1.07\ . \tag{2.6}$$

Therefore, when we observe a colour at a particular point in an image, we must increase its apparent radius by 7% to obtain the real radius at which the interface height takes the value corresponding to the observed colour.

At the inner sidewall, r_{lid} takes the value 6.25 cm. From equation (2.6), r is then 6.70 cm, implying that when the interface is flat we will not be able to see the inner 0.45 cm because of a parallax effect. A similar calculation shows that the outer 0.78 cm of the annular gap will likewise not be visible to the camera.

2.6 Chapter summary

Descriptions have been given of the rotating, two-layer annulus laboratory apparatus, fluid properties and flow visualization technique.

Useful information from the laboratory experiment images, e.g. wavelengths and propagation speeds, can be extracted using the flow visualization technique. But the images contain much more information which has until now remained untapped, namely quantitative data about the interface height field, which is encoded in the image colours. In the next chapter we describe a new method for calibrating interface heights in the experiment, to allow this information to be extracted for the first time.

Chapter 3

Calibration of the laboratory experiment

> *"Mere colour, unspoiled by meaning, and unallied with definite form, can speak to the soul in a thousand different ways."*
>
> Oscar Wilde

In this chapter, we describe the calibration of the laboratory experiment, allowing the transformation from raw laboratory images to quantitative maps of interface height. Three candidate calibration schemes are described, and the one considered likely to be the most accurate is developed and implemented. The chosen scheme involves solving layerwise torque balance equations to determine the equilibrium interface height shape when the apparatus is in motion. This analysis is an extension of previous calculations, to include new and important physical effects.

The calibration attempt is successful, allowing the amplitudes of the interfacial fast and slow waves to be accurately measured for the first time with this apparatus.

3.1 Motivation for calibrating the experiment

The unprocessed colour video images from the laboratory experiment (e.g. Figure 1.6) are of limited use. Though they provide essential quantitative information about the horizontal structure of the flow (wavelengths and propagation speeds), the information regarding the vertical direction (in particular, interfacial wave amplitudes) is only qualitative. All we know so far is that changes in interface height correspond to changes in observed colour, as discussed in Section 2.3. The exact quantitative relationship between height and colour is not known, and so previous experimenters using this apparatus have had to *estimate* interfacial wave amplitudes by visual observation through the annulus lid. This method is fraught with difficulty, due to the rotating metal structure which supports the camera, inaccuracies due to refraction at the lid and, most of all, the inherent unreliability of guesswork. In this chapter, we develop a method for *measuring* interface height quantitatively.

The images obtained from the camera are recorded onto video tape, and can be subsequently digitized by the frame grabber to produce red $R(h)$, green $G(h)$ and blue $B(h)$ intensity components, which depend upon the lower layer depth $h(r, \theta, t)$. The task of this chapter is to determine the functions $R(h)$, $G(h)$ and $B(h)$, and hence their inverses so that we can compute h given R, G and B. In the following sections we consider three possible approaches to this problem.

3.2 Theoretical approach

Suppose that the angle between the axes of the entrance and analyzing polaroids is α, and that the rotation angle per unit lower layer depth shown in Figure 2.7 is $\phi(\lambda)$. Neglecting the liquid in the parabolic correction tank, if light of wavelength λ travels through a lower layer depth of $h(r, \theta, t)$, then a fraction $\cos^2[\alpha - h\phi(\lambda)]$ of the incident intensity will be transmitted through the analyzing polaroid to the camera. Assuming that the source lamp is perfectly emitting, the intensity spectrum input into the liquids is the Planck function $I_0(\lambda)$ at the temperature of the lamp (given as 3200 K by Trundle, 1987). Further assuming no absorption by the apparatus, the intensity spectrum $I(\lambda)$ received at the camera is given by

$$I(\lambda) = I_0(\lambda)\cos^2[\alpha - h\phi(\lambda)] \ . \tag{3.1}$$

This spectrum is sampled by the camera at three wavelengths (corresponding to red, green and blue light), to give pixel intensities of

$$\begin{aligned} R(h) &= A\, I_0(\lambda_{\mathrm{red}}) \cos^2[\alpha - h\phi(\lambda_{\mathrm{red}})] \ , && (3.2) \\ G(h) &= A\, I_0(\lambda_{\mathrm{green}}) \cos^2[\alpha - h\phi(\lambda_{\mathrm{green}})] \ , && (3.3) \\ B(h) &= A\, I_0(\lambda_{\mathrm{blue}}) \cos^2[\alpha - h\phi(\lambda_{\mathrm{blue}})] \ , && (3.4) \end{aligned}$$

where A is a constant of proportionality. Equations (3.2)–(3.4) could each be written as a weighted integral of (3.1) with respect to λ, where the weighting functions, or camera response functions, peak in the red, green and blue parts of the spectrum. The assumption is made here that the response functions are delta-functions, however. The camera then simply samples the incoming spectrum at three discrete wavelengths rather than over three narrow bands of finite width. We have succeeded in determining the functions $R(h)$, $G(h)$ and $B(h)$. Once we have determined R, G and B from an image, equations (3.2)–(3.4) represent three equations in three unknowns: A, α and h. It is tedious but straightforward to eliminate A and α to give an implicit, nonlinear equation for h in terms of R, G and B (Williams, 2000). However, this method is made unreliable by the assumptions which have been made (source lamp being a perfect black body at an assumed temperature, camera response functions being delta-functions, zero absorption in working liquids), none of which is particularly well-justified. We require a quantitatively accurate calibration scheme which is more reliable than this approximate theoretical method.

3.3 Direct experimental approach

For accuracy, it is desirable to take an experimental approach to the calibration problem. There is an obvious and direct experimental method. One can imagine filling up the initially-empty annular tank with the limonene/CFC mixture, in a series of discrete steps so that each time the depth rises by, say, 1 mm. A video recording could be made after each millimetre rise, allowing the colour in each video image to be measured and calibrated with interface height.

This method would require the limonene layer to be exposed to the laboratory air for a significant period of time. Because of the harmful vapour released by limonene when not covered by the water layer, this method would pose a health risk to the user. A Local Exhaust Ventilation (LEV) system could be used to reduce emissions into the laboratory, but the LEV is an effective extractor only when it is positioned directly above the annulus and therefore blocking the camera field of view, which would defeat the object. For these reasons, the direct experimental approach was rejected.

3.4 Indirect experimental approach

We have chosen for the present purposes to use an experimental calibration based on images taken when the experiment is in operation, that is, when both the turntable and lid are rotating at different rates. All that is needed is a method for obtaining the interface height field in just one special case. The method must be independent, in the sense that it does not rely on the colour information in the images, since that is what we wish to calibrate. Fortunately, it is possible to derive an analytical expression for the equilibrium interface height in the special case of no baroclinic instability. In this case, zonal wave modes are completely absent and the interface is axisymmetric, but the height can still vary strongly with radius.

We plan to take interface height as a function of radius from the analytical expression, and colour as a function of radius from a laboratory experiment, and to determine the relationship between interface height and colour from the two. We derive the required analytical expression over the following pages.

3.4.1 Equilibrium interface height field

We begin the calculation by deriving an expression for the equilibrium lower layer depth field $h(r)$, shown in Figure 3.1, in terms of the fluid interior solid-body rotation rates $\Delta\Omega_1$ and $\Delta\Omega_2$. The pressures in each layer are given by

$$p_i = \frac{1}{2}\rho_i \Omega_i^2 r^2 - \rho_i g z + \text{constant} \ , \tag{3.5}$$

where $i = 1$ refers to the upper layer and $i = 2$ to the lower layer. This equation represents hydrostatic balance in the z-direction, and a balance between the radial pressure gradient force and the acceleration experienced by a fluid parcel executing circular

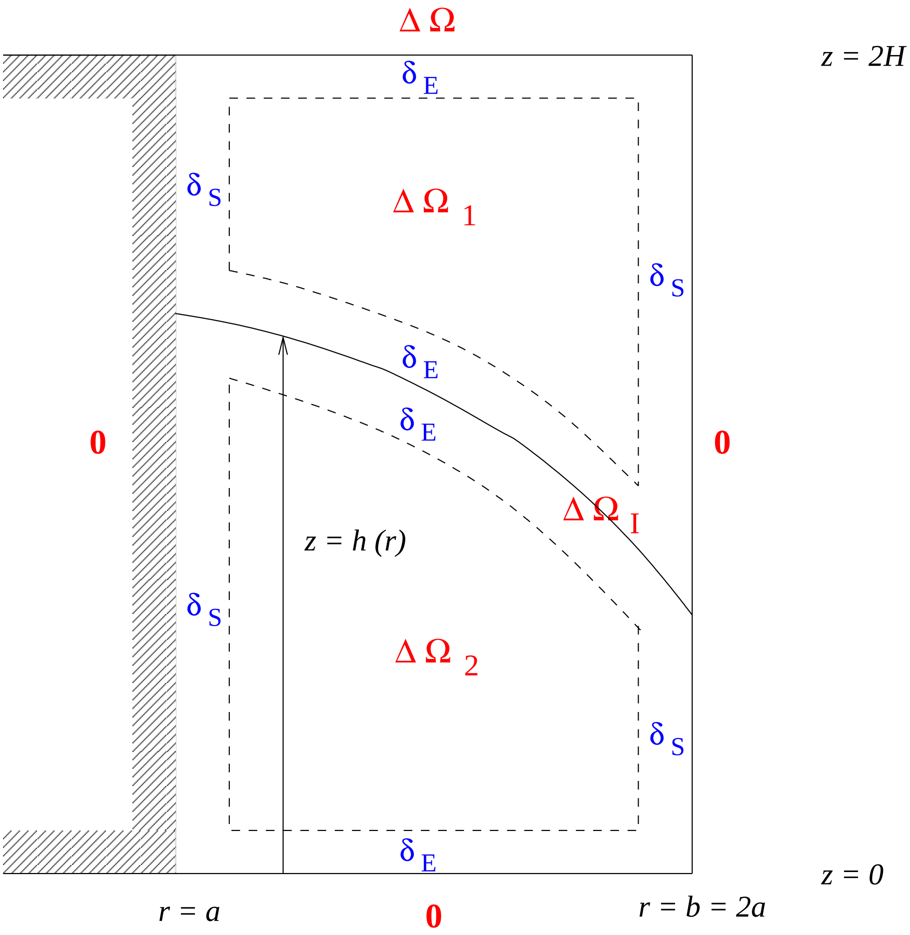


Figure 3.1: Variable definitions for the torque balance calculation. The boundary layer widths, labelled in blue, are shown greatly exaggerated. The angular velocities of the lid, base, sidewalls, interface and fluid interiors, about the rotation axis and relative to the base frame, are labelled in red.

motion with angular velocity $\Omega_i = \Omega + \Delta\Omega_i$ in the laboratory frame.

We may deduce the pressure difference between the annulus base and lid using the pressure field given by equation (3.5), and we can equate this pressure drop to the same quantity determined in an independent manner, namely by vertical integration of the hydrostatic equation. This gives

$$\frac{1}{2}(\rho_2\Omega_2^2 - \rho_1\Omega_1^2)r^2 + \text{constant} = \rho_2 g h + \rho_1 g(2H - h) - S\nabla^2 h \ , \tag{3.6}$$

where the term in $\nabla^2 h$ represents the drop in pressure across the interface due to interfacial tension S. From this equation we can evaluate $h(r)$. Using the fact that $\rho_1 \approx \rho_2$ and neglecting meniscus effects we obtain

$$h(r) = \frac{(\Omega_2^2 - \Omega_1^2)r^2}{2g'} + \text{constant} \ , \tag{3.7}$$

where $g' = g(\rho_2 - \rho_1)/[\frac{1}{2}(\rho_2 + \rho_1)]$. Finally, we determine the arbitrary constant by applying conservation of volume:

$$\int_{r=a}^{r=b} 2\pi r h(r)\,\mathrm{d}r = \pi(b^2 - a^2)H \ , \tag{3.8}$$

to obtain

$$h(r) = H + \frac{(\Omega_2^2 - \Omega_1^2)(r^2 - \frac{5}{2}a^2)}{2g'} \ . \tag{3.9}$$

We will need the following three formulae for the torque balance calculation. The values taken by the interface height at the sidewall boundaries are

$$h(r = a) = H + \frac{3}{2}a^2\left(\frac{\Omega_1^2 - \Omega_2^2}{2g'}\right) \equiv H_+ \tag{3.10}$$

and

$$h(r = b) = H - \frac{3}{2}a^2\left(\frac{\Omega_1^2 - \Omega_2^2}{2g'}\right) \equiv H_- \ , \tag{3.11}$$

and the interface slope at the outer sidewall is

$$\left.\frac{\mathrm{d}h}{\mathrm{d}r}\right|_{r=b} = -\left(\frac{b(\Omega_1^2 - \Omega_2^2)}{g'}\right) \ . \tag{3.12}$$

The substitution $b = 2a$ has been used in equations (3.9)–(3.11).

3.4.2 Previous approaches

All that remains to be done, to complete our derivation of the interface height field, is to determine the Ω_i in terms of Ω and $\Delta\Omega$. In much of Hart's work, including Hart (1972), Hart (1973) and Hart (1985), as well as in other studies including Bradford et al. (1981), these interior rotation rates are derived from the 2-layer quasi-geostrophic (Q-G) equations with linear, parameterized Ekman velocities, neglecting the influence of the sidewall boundaries and assuming a horizontal, flat interface. This calculation yields the simple result $\Delta\Omega_2 = \frac{1}{4}\Delta\Omega$ and $\Delta\Omega_1 = \frac{3}{4}\Delta\Omega$ for the case of exactly equal viscosities (see equations 5.20 and 5.21). These values can be substituted into equation (3.9) to obtain an explicit expression for $h(r)$.

The assumptions of geostrophy, non-interacting Stewartson layers and a horizontal interface mean that this method can only be considered a first approximation. Therefore, King (1979b) extended the Q-G analysis to include a non-horizontal interface, by including factors of cos(mean interface gradient) in the Ekman layer terms. This simple extension made the calculation significantly more difficult, as the formulae for $\Delta\Omega_i$ now include the mean interface slope. Therefore, on substituting into (3.9) an implicit, nonlinear equation for $h(r)$ is obtained, which must be solved numerically. This approach was taken in Williams (2001) to calibrate the present experiment, but it gave rotation rates which seemed too large due to the exclusion of Stewartson layer drag from this analysis.

Stewartson layers and ageostrophy cannot be captured by the simple Q-G approach, and so King (1979b) went on to take a different approach based on layer torque balance. He argued that, in equilibrium, the fluid interiors do not experience an angular acceleration, and so the net external torque on the interiors due to the boundary layers must be zero. Stewartson layers and ageostrophy are both included, but King resorted to using the horizontal interface assumption to make the calculation analytically tractable.

In the present problem, we specifically require a non-horizontal interface, as we want the calibration curve to span as wide a range of interface heights as possible. We therefore present, in the following section, an extension of King's torque balance calculation to include non-horizontal interface effects. A summary of the physical effects included in the previous and present calculations is shown in Table 3.1.

	Q-G method used by Hart and Bradford	Ekman flux calculation of King	torque balance calculation of King	torque balance calculation in this thesis
Ekman layers	✓	✓	✓	✓
Stewartson layers	×	×	✓	✓
sloping interface	×	✓	×	✓
ageostrophy	×	×	✓	✓

Table 3.1: A comparison of the physical effects taken into account in various calculations to determine the equilibrium interface height in a rotating, two-layer annulus. The methods in the first two columns are based on equating geostrophic Ekman pumping and suction velocities at the top and bottom of each layer. The methods in the last two columns are based on torque balance equations for each layer.

3.4.3 Torque balance calculation

For the torque balance calculation we model each fluid layer as an inviscid interior region, making up the vast majority of the volume of the layer, surrounded on all sides by thin viscous boundary layers which serve to change the fluid velocity from its interior value to its no-slip boundary value. In the two-layer annulus, the boundaries are the lid, base and fluid interface (at which the boundary layers are Ekman (1905) layers), and the inner and outer cylindrical sidewalls (at which the boundary layers are Stewartson (1957) layers). We assume that the interior flow in each layer is hydrostatic and columnar, and in solid-body rotation with the angular velocities (to be determined) shown in Figure 3.1.

We expect, when the imposed lid rotation $\Delta\Omega$ is positive, that

$$0 < \Delta\Omega_2 < \Delta\Omega_{\mathrm{I}} < \Delta\Omega_1 < \Delta\Omega \; , \tag{3.13}$$

where $\Delta\Omega_{\mathrm{I}}$ is the angular velocity of the interface. Qualitatively, the upper layer interior is being acted upon by a prograde (anti-clockwise) stress due to the Ekman layer at the lid, and by retrograde (clockwise) stresses due to the Ekman layer above the interface and both Stewartson layers. The boundary layer at the lid is tending to spin the layer up, and the remaining three boundary layers are tending to spin it down. In the lower layer, it is the interfacial boundary layer which gives a positive angular velocity tendency, and the remaining three which give a negative contribution.

The larger the velocity shear across a boundary layer, the larger the stress and the larger the torque exerted on the fluid interior by the boundary. If there is a non-zero net torque in either layer, there will be an angular acceleration. We expect the interior rotation rates to adjust themselves so that, in equilibrium, the net torque is zero and the rotation rates remain constant with time.

We now quantify the ideas of the previous paragraphs by writing down equations for the net torques in both layers in terms of the rotation rates, and then solving for the equilibrium rates by setting the torques equal to zero. To simplify the analysis we assume equal layer viscosities ν and densities ρ, both of which approximations are very good in the present context (see Table 2.1). The Ekman and Stewartson layer widths are derived in e.g. Read (1992b) to be, respectively,

$$\delta_{\mathrm{E}} = \left(\frac{\nu}{\Omega}\right)^{\frac{1}{2}} \tag{3.14}$$

and

$$\delta_{\mathrm{S}} = \left(\frac{a^2\nu}{\Omega}\right)^{\frac{1}{4}} \quad . \tag{3.15}$$

There are two distinct Stewartson layers at each sidewall in the rotating annulus. The one used here is that which serves to take the horizontal fluid velocity to zero at the boundary, and which has a nondimensional width of the Ekman number to the power of one-quarter. This is the appropriate layer for the present calculation, as it is the region in which the lateral velocity shear exists. The other Stewartson layer, of width equal to the Ekman number to the power of one-third, is responsible for returning vertical Ekman fluxes. This layer is not associated with a horizontal drag force at the sidewalls, and therefore does not make a contribution to the torque about the rotation axis. For a typical rotation rate of $\Omega = 1$ rad s^{-1}, we obtain $\delta_{\mathrm{E}} = 1$ mm and $\delta_{\mathrm{S}} = 8$ mm.

Rotation rate of interface

In equilibrium, the interface does not accelerate, and therefore must feel no net torque due to the thin Ekman layers above and below it. This means that the vertical shear in horizontal velocity across the upper interfacial Ekman layer must equal that across the lower one, giving

$$\Delta\Omega_{\mathrm{I}} = \frac{1}{2}(\Delta\Omega_1 + \Delta\Omega_2) \ . \tag{3.16}$$

Torque due to Stewartson layers

In general, shear stresses within fluids are given by the tensor $S_{i,\,j} = \rho\nu\partial u_i/\partial x_j,\ i \neq j$. So, for example, the stress *on* the upper layer interior *by* the vertical boundary $r = b$ is

$$-\rho\nu\frac{b\Delta\Omega_1}{\delta_{\mathrm{S}}} \ , \tag{3.17}$$

where the minus sign indicates that this stress represents a drag. A mean has been taken over the thin Stewartson layer, across which a velocity change of $b\Delta\Omega_1$ is achieved. The Stewartson layer has area $2\pi bH_+$ (equations 3.10 and 3.11) and is a distance b from the rotation axis, and so it exerts a torque on the fluid of

$$T^{\mathrm{Stewartson}}_{\mathrm{layer=1,}\ r=b} = -\frac{2\pi\rho\nu\Delta\Omega_1 H_+ b^3}{\delta_{\mathrm{S}}} \ . \tag{3.18}$$

Similar expressions are obtained for the torques $T^{\mathrm{Stewartson}}_{\mathrm{layer=1,}\ r=a}$, $T^{\mathrm{Stewartson}}_{\mathrm{layer=2,}\ r=b}$ and $T^{\mathrm{Stewartson}}_{\mathrm{layer=2,}\ r=a}$ due to the remaining three Stewartson layers.

Torque due to lid and base Ekman layers

The stress *on* the upper layer interior *by* the horizontal boundary $z = 2H$ at radius r is

$$\rho\nu\frac{r(\Delta\Omega - \Delta\Omega_1)}{\delta_{\mathrm{E}}} \ . \tag{3.19}$$

In this case the stress is dependent upon radius. An area element is $r\,\mathrm{d}r\,\mathrm{d}\theta$ and the distance from the axis is r, and so this Ekman layer exerts a torque on the fluid of

$$\begin{aligned} T^{\mathrm{Ekman}}_{\mathrm{layer=1,}\ z=2H} &= \int\limits_{\theta=0}^{2\pi}\int\limits_{r=a}^{b} \frac{\rho\nu(\Delta\Omega - \Delta\Omega_1)r^3}{\delta_{\mathrm{E}}}\,\mathrm{d}r\,\mathrm{d}\theta \qquad (3.20)\\ &= \frac{\pi\rho\nu(\Delta\Omega - \Delta\Omega_1)(b^4 - a^4)}{2\delta_{\mathrm{E}}} \ . \qquad (3.21) \end{aligned}$$

A similar expression is obtained for the torque $T^{\text{Ekman}}_{\text{layer}=2,\, z=0}$ on the lower layer interior due to the Ekman layer at the base.

Torque due to interfacial Ekman layers

The stress *on* the upper layer interior *by* the interface $z = h(r)$ at radius r is

$$-\rho\nu \frac{r(\Delta\Omega_1 - \Delta\Omega_I)}{\delta_{\text{E}}} \ . \tag{3.22}$$

In this case the area element is $\sqrt{1 + (\mathrm{d}h/\mathrm{d}r)^2}\, r\, \mathrm{d}r\, \mathrm{d}\theta$ and the distance from the axis is r, and so this Ekman layer exerts a torque on the fluid of

$$T^{\text{Ekman}}_{\text{layer}=1,\, z=h} = -\int\limits_{\theta=0}^{2\pi}\int\limits_{r=a}^{b} \frac{\rho\nu(\Delta\Omega_1 - \Delta\Omega_I)r^3}{\delta_{\text{E}}}\sqrt{1 + \left(\frac{\mathrm{d}h}{\mathrm{d}r}\right)^2}\, \mathrm{d}r\, \mathrm{d}\theta \tag{3.23}$$

$$\approx -\frac{\pi\rho\nu(\Delta\Omega_1 - \Delta\Omega_I)(b^4 - a^4)}{2\delta_{\text{E}}}\sqrt{1 + \left(\left.\frac{\mathrm{d}h}{\mathrm{d}r}\right|_{r=b}\right)^2} \ . \tag{3.24}$$

An approximation has been employed (without which further analytical progress becomes impossible) to replace the surd in the integrand of equation (3.23) with its value at $r = b$, since the r^3 factor heavily weights the integral towards larger r.

A similar expression is obtained for the torque $T^{\text{Ekman}}_{\text{layer}=2,\, z=h}$ on the lower layer interior due to the Ekman layer at the interface.

Torque balance equations

We now write down expressions for the net torque in each layer, and equate them to zero in equilibrium to give

$$T^{\text{Stewartson}}_{\text{layer}=1,\, r=a} + T^{\text{Stewartson}}_{\text{layer}=1,\, r=b} + T^{\text{Ekman}}_{\text{layer}=1,\, z=2H} + T^{\text{Ekman}}_{\text{layer}=1,\, z=h} = 0 \tag{3.25}$$

and

$$T^{\text{Stewartson}}_{\text{layer}=2,\, r=a} + T^{\text{Stewartson}}_{\text{layer}=2,\, r=b} + T^{\text{Ekman}}_{\text{layer}=2,\, z=0} + T^{\text{Ekman}}_{\text{layer}=2,\, z=h} = 0 \ . \tag{3.26}$$

Equations (3.25)–(3.26) are two nonlinear equations in the two unknowns $\Delta\Omega_1$ and $\Delta\Omega_2$. Rearranging, we may write the equations in matrix form:

$$\begin{bmatrix} 1 + \frac{1}{2}\sqrt{1 + \left(\frac{\mathrm{d}h}{\mathrm{d}r}\big|_b\right)^2} + \left(\frac{4}{15}\right)\left(\frac{\delta_{\text{E}}}{\delta_{\text{S}}}\right)\left(\frac{8H_+ + H_-}{a}\right) & -\frac{1}{2}\sqrt{1 + \left(\frac{\mathrm{d}h}{\mathrm{d}r}\big|_b\right)^2} \\ -\frac{1}{2}\sqrt{1 + \left(\frac{\mathrm{d}h}{\mathrm{d}r}\big|_b\right)^2} & 1 + \frac{1}{2}\sqrt{1 + \left(\frac{\mathrm{d}h}{\mathrm{d}r}\big|_b\right)^2} + \left(\frac{4}{15}\right)\left(\frac{\delta_{\text{E}}}{\delta_{\text{S}}}\right)\left(\frac{8H_- + H_+}{a}\right) \end{bmatrix}$$

$$\times \begin{bmatrix} \Delta\Omega_1 \\ \Delta\Omega_2 \end{bmatrix} = \begin{bmatrix} \Delta\Omega \\ 0 \end{bmatrix} . \tag{3.27}$$

The matrix equation is nonlinear because H_+, H_- and $\mathrm{d}h/\mathrm{d}r|_{r=b}$ all depend upon $\Delta\Omega_1$ and $\Delta\Omega_2$ through equations (3.10)–(3.12). If we now make the horizontal interface assumption, which is $H_+ = H_- = H$ and $\mathrm{d}h/\mathrm{d}r|_{r=b} = 0$, the equations linearize and we recover the results of King's torque balance calculation. Additionally neglecting the Stewartson layers by letting $\delta_S \to \infty$ reduces the matrix equation to

$$\begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 3/2 \end{bmatrix} \begin{bmatrix} \Delta\Omega_1 \\ \Delta\Omega_2 \end{bmatrix} = \begin{bmatrix} \Delta\Omega \\ 0 \end{bmatrix} , \tag{3.28}$$

for which the solution is

$$\begin{bmatrix} \Delta\Omega_1 \\ \Delta\Omega_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4}\Delta\Omega \\ \frac{1}{4}\Delta\Omega \end{bmatrix} , \tag{3.29}$$

which is the simple Q-G result, as expected.

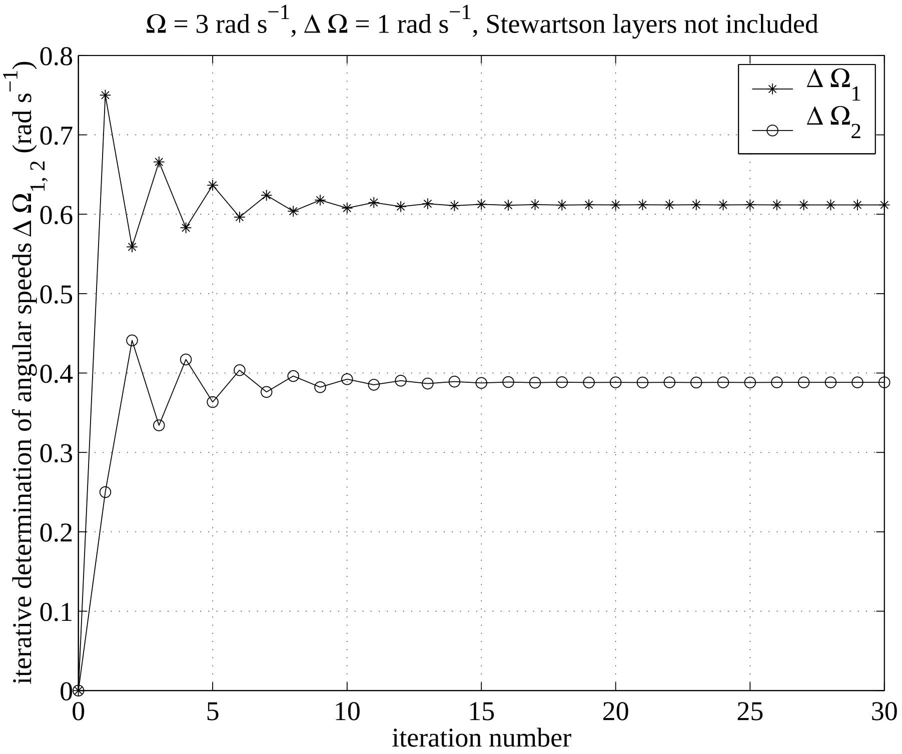
Iterative solutions

We use an iterative approach to solve the full, nonlinear matrix equation (3.27):

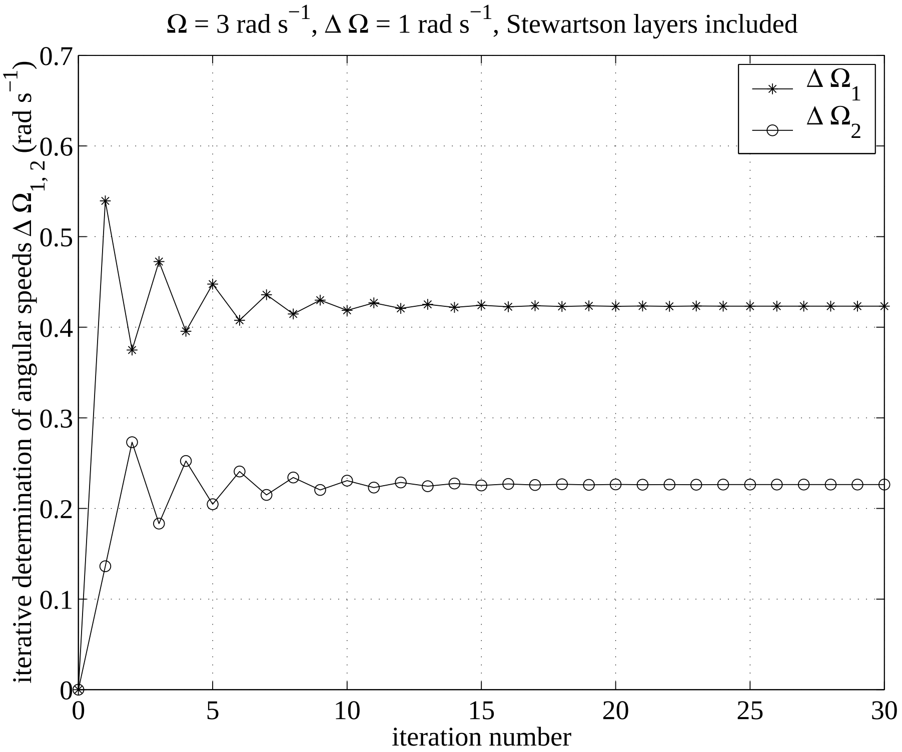
1. choose $\Delta\Omega_1 = \Delta\Omega_2 = 0$ as a first guess;
2. evaluate H_+, H_- and $\mathrm{d}h/\mathrm{d}r|_{r=b}$ for this $\Delta\Omega_1$, $\Delta\Omega_2$;
3. evaluate the four matrix elements for this H_+, H_-, $\mathrm{d}h/\mathrm{d}r|_{r=b}$;
4. invert the matrix equation to obtain an improved guess for $\Delta\Omega_1$ and $\Delta\Omega_2$;
5. if the original and improved solutions are not equal to within the required precision, return to step 2 for another iteration.

The iterations were found to converge in almost all cases. The exceptions occurred when both Ω and $\Delta\Omega$ were very large, when a feature with a period of two iterations persisted in the equilibrated iteration series. In these cases there is presumably no equilibrium solution to the torque balance equations.

Figure 3.2 shows the results of the iteration calculation (performed using Matlab), both with the Stewartson layers switched on and off, for $\Omega = 3$ rad s^{-1} and $\Delta\Omega = 1$ rad s^{-1}.



(a)



(b)

Figure 3.2: Results of the iterative numerical solution of the nonlinear matrix equation for $\Delta\Omega_{1,2}$, both (a) without and (b) with the Stewartson layer terms.

$\Delta\Omega_1$ (rad s^{-1})	horizontal interface	curved interface
Stewartson layers switched off	0.75	0.61
Stewartson layers switched on	0.54	0.42

(a)

$\Delta\Omega_2$ (rad s^{-1})	horizontal interface	curved interface
Stewartson layers switched off	0.25	0.39
Stewartson layers switched on	0.14	0.23

(b)

Table 3.2: Equilibrated values of (a) $\Delta\Omega_1$ and (b) $\Delta\Omega_2$ for the case $\Omega = 3$ rad s^{-1} and $\Delta\Omega = 1$ rad s^{-1}. The values are based on a torque balance analysis which always includes Ekman layers, and which can also additionally include Stewartson layers and/or the increased drag effects of a curved interface.

The fourth decimal place of the solutions is stable after around the 20th iteration. Note from the above iteration plan that the first improved guess is actually the solution of the horizontal interface problem, as when $\Delta\Omega_1 = \Delta\Omega_2 = 0$ we have $H_+ = H_- = H$ and $\mathrm{d}h/\mathrm{d}r|_{r=b} = 0$, so we have solved this problem for every combination of Stewartson layers present and absent, and interface horizontal and curved. The findings are summarized in Table 3.2.

In both layers, and both with and without a non-horizontal interface, the inclusion of Stewartson layer drag has significantly reduced the layer rotation rates. In the upper layer, allowing for a non-flat interface also reduces the rotation rate, because a curved interface has a greater surface area than a flat one, and hence gives a greater drag force. In the lower layer, the rotation rate is increased when the non-flat interface is included, as the Ekman layer at the interface provides the only positive torque in this layer.

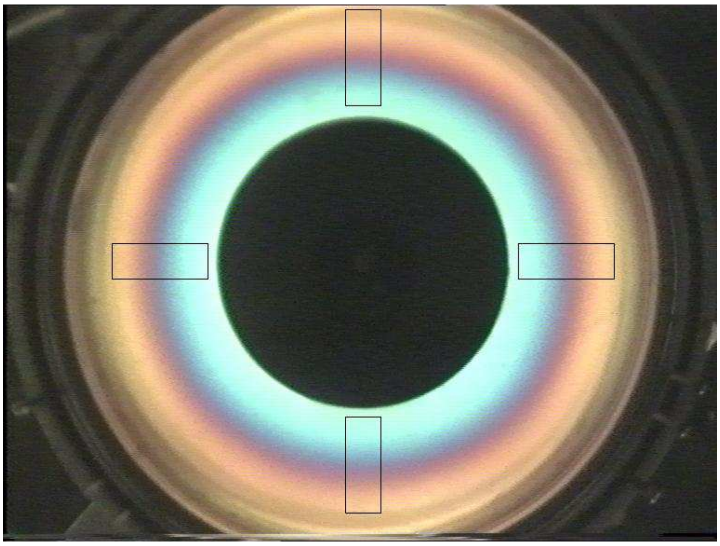


Figure 3.3: Laboratory experiment image used for the calibration, showing the equilibrated flow in the case $\Delta\Omega = 0.77$ rad s^{-1}, $\Omega = 1.87$ rad s^{-1}. The four boxes, each measuring 106 pixels by 40 pixels, indicate the areas from which colour information was extracted.

We conclude that both Stewartson layers and curved interface effects are important for determining quantitatively accurate layer rotation rates, which justifies the full analysis given above for the purposes of calibrating the laboratory experiment.

3.5 Implementation of the calibration scheme

We now describe the implementation of the calibration scheme. An image from a laboratory experiment, showing the equilibrated axisymmetric flow which is attained in the baroclinically-stable case $\Delta\Omega = 0.77$ rad s^{-1} and $\Omega = 1.87$ rad s^{-1}, is shown in Figure 3.3. In this case, the converged iterative solution to equation (3.27) with Stewartson layers and a sloping interface is found to be $\Delta\Omega_1 = 0.34$ rad s^{-1} and $\Delta\Omega_2 = 0.14$ rad s^{-1}. Figure 3.4 shows the interface height shape thereby obtained, calculated using equation (3.9). There are 106 pixels across the annular gap in Figure 3.3, and there will therefore be 106 points on the calibration curve we obtain.

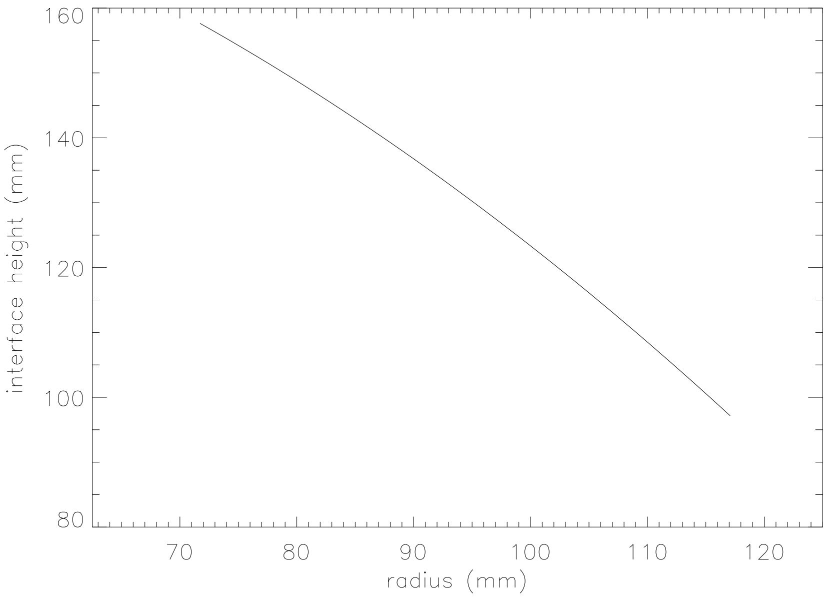


Figure 3.4: Lower layer depth $h(r)$ as a function of radius, calculated analytically for the flow shown in Figure 3.3.

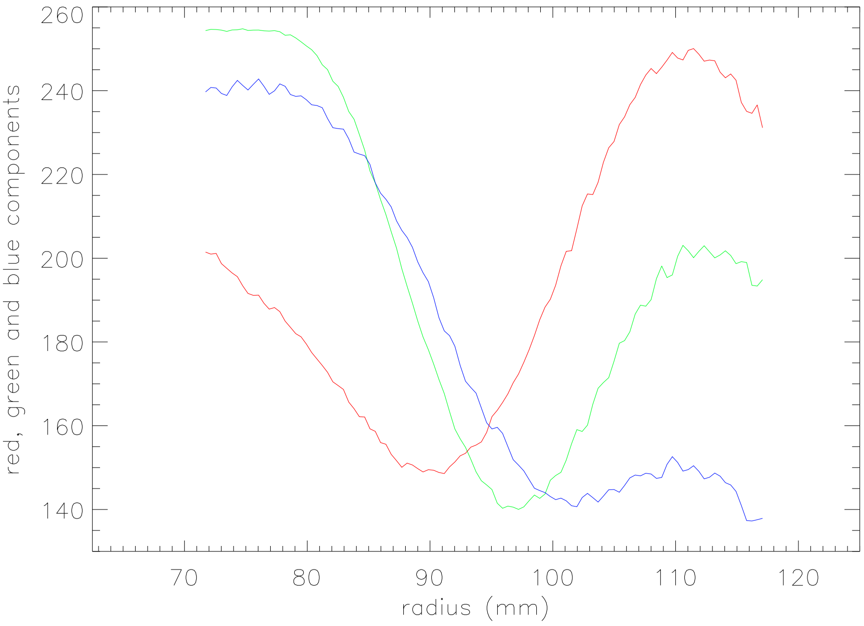


Figure 3.5: Azimuthally-averaged red, green and blue components as functions of radius, derived from Figure 3.3. The abscissa is drawn to span the entire annular gap, even though parallax effects discussed in Section 2.5 prevent the extraction of data near the sidewalls.

3.5.1 Choice of colour calibration variable

Now that we know the colour field and the interface height field, we are in a position to derive a calibration curve. The R, G and B values of the pixels in Figure 3.3 have been extracted by loading the jpeg image into IDL, from which they are each given as integers between 0 and $2^8 - 1 = 255$ inclusive (because the colour digitization is 8-bit for each of the three colour channels). The mean radial dependence of the R, G and B fields is shown in Figure 3.5, where the parallax effects of Section 2.5 have been included to give an unbiased measurement of the radius. As indicated in Figure 3.3, averages have been taken over the azimuthal angles corresponding to “3 o’clock”, “6 o’clock”, “9 o’clock” and “12 o’clock”, in case the flow is not perfectly axisymmetric. Also as indicated in the figure, averages have been taken over 40 azimuthally-neighbouring pixels at each of these four angles, to reduce contamination of the signal by noise.

We can eliminate radius between the curves in Figures 3.4 and 3.5 to obtain the red $R(h)$, green $G(h)$ and blue $B(h)$ components as functions of interface height h. Then in the three-dimensional space (R, G, B), the most complete calibration curve we can define is given parametrically by $(R(h), G(h), B(h))$. To find h for a given point (R^*, G^*, B^*), we would simply need to find the point on the 3-D calibration curve which is closest to the given point. We could do this by, for example, minimizing the cost function $[R(h) - R^*]^2 + [G(h) - G^*]^2 + [B(h) - B^*]^2$ with respect to h. This calculation is quite computationally-expensive, especially if interface heights are required at the majority of points on a 768 by 576 spatial grid, 25 times per second.

We would prefer a one-dimensional calibration curve, so that we can use simple linear interpolation to inexpensively return interface heights. For example, we could choose to use $R(h)$ as the calibration curve, abandoning $G(h)$ and $B(h)$, though it would seem wasteful to discard two-thirds of the available colour information. To avoid this redundancy, any function of $R(h)$, $G(h)$ and $B(h)$ could be used.

There are other colour systems apart from the (R, G, B) system, and there is no guarantee that R, G and B are in any way optimized as calibration variables. A commonly-used alternative is the (H, S, I) system (e.g. Foley & Van Dam, 1982), where H is the *hue*, S is the *saturation* and I is the *total intensity*. The transformation from (R, G, B) coor-

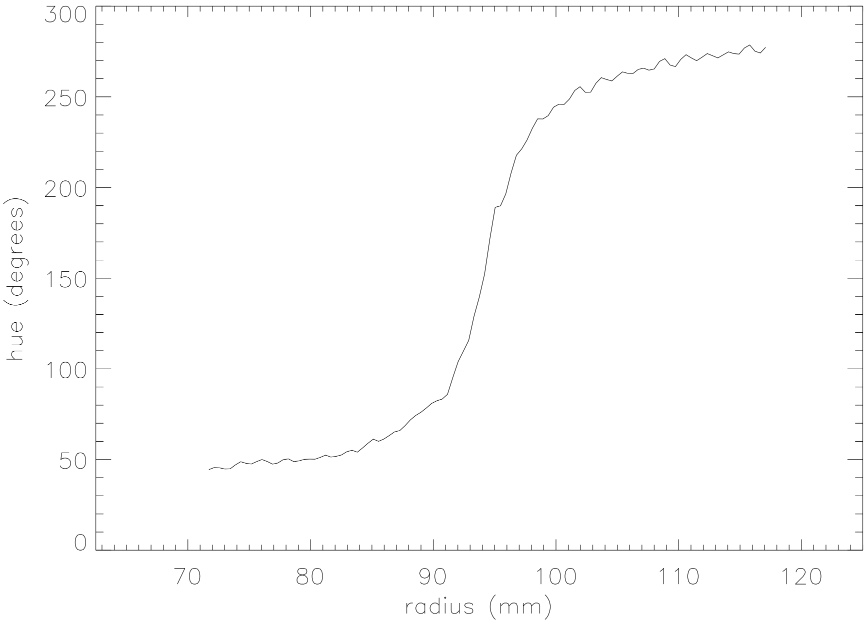


Figure 3.6: Azimuthally-averaged hue component as a function of radius.

dinates to (H, S, I) coordinates is outlined in Appendix A. The H, S and I fields in Figure 3.3, calculated from the R, G and B fields, are shown in Figures (3.6)–(3.8).

3.5.2 Derivation of calibration curve

In theory, any of the six variables R, G, B, H, S, I could be used to derive a calibration curve. However, the hue H is the most suitable, for two important reasons. Firstly, hue is the only colour variable of the six which is a one-to-one function of interface height over the height range being considered. The other five are many-to-one functions, and hence are not uniquely invertible. If any one of these five were to be used as the calibration variable, there would often be an ambiguity over which height had given rise to the observed colour component. This is clearly an undesirable feature of any calibration scheme.

Secondly, it follows from the definition of hue (equation A.9) that if R, G and B are all reduced in equal proportions then H will be unaffected. This is because hue is independent of the total intensity. The implication is that any local absorption in the fluids due to contamination by small particles, will be visible in all the colour fields except

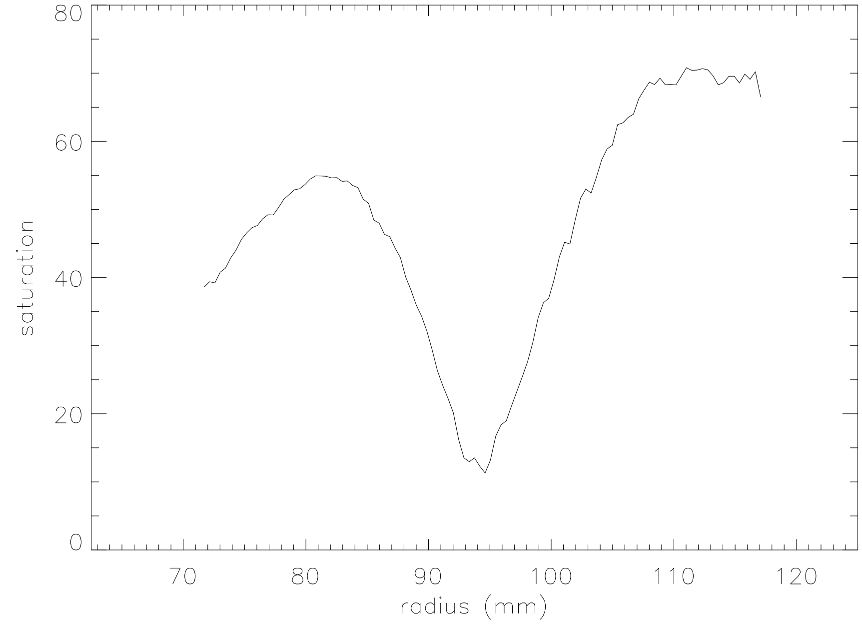


Figure 3.7: Azimuthally-averaged saturation component as a function of radius.

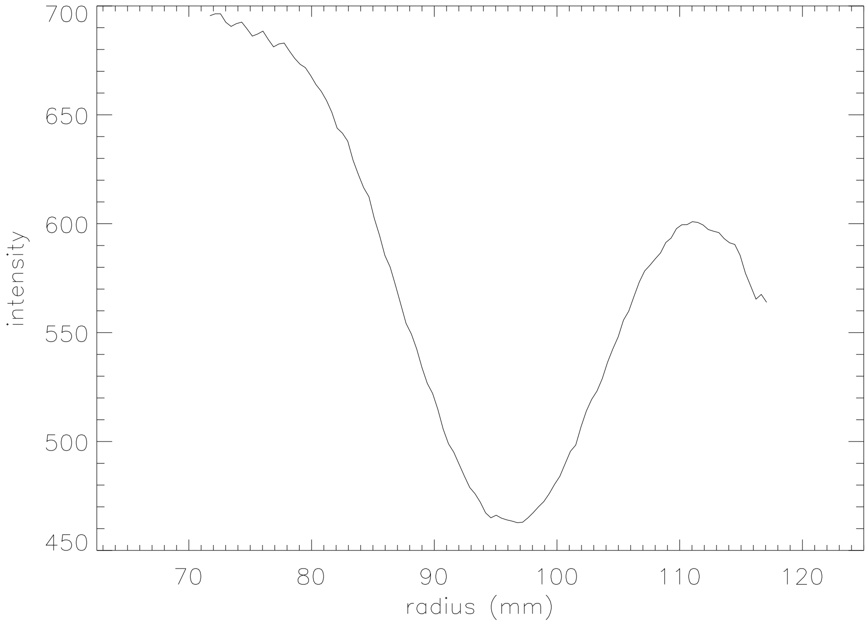


Figure 3.8: Azimuthally-averaged intensity component as a function of radius.

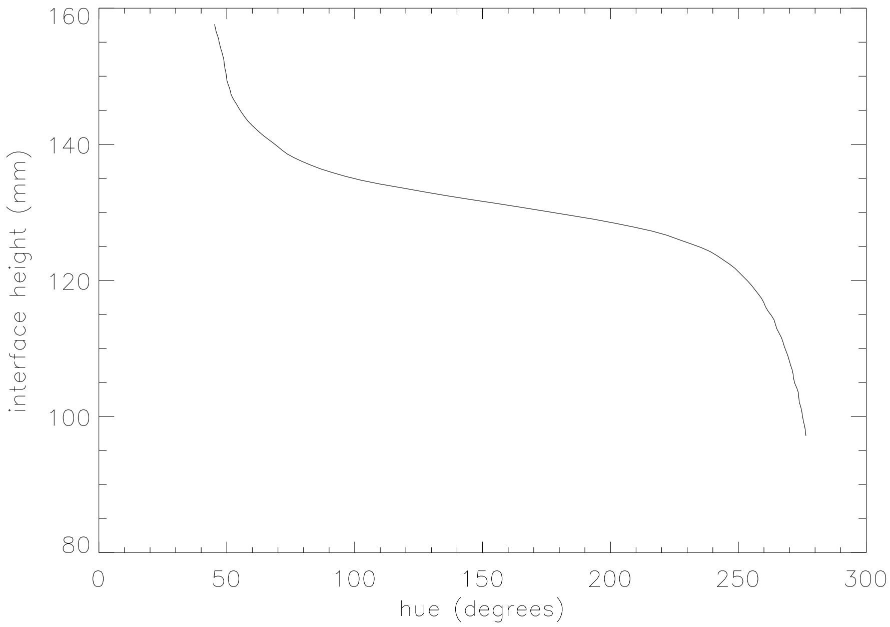


Figure 3.9: Calibration curve showing the relationship between the hue component and lower-layer depth. A slight smoothing has been applied to the curve, by taking a moving average, to remove the last remnants of noise.

hue (assuming equal absorption at all wavelengths). This explains why the hue field in Figure 3.6 is less noisy than the red, green and blue fields in Figure 3.5. A further source of light absorption in the fluids is a cloudiness that is often found to form in the liquids whenever the laboratory temperature cools slightly, possibly due to condensation of trace quantities of lower-layer liquid in the upper layer. Calibration curves using R, G, B, S, I would therefore be expected to return heights which varied from one day to the next, unlike a calibration curve using H.

Hue being a monotonic function of depth, coupled with its robustness to absorption effects, makes it the ideal calibration variable. The hue calibration curve, obtained by eliminating the radius from Figures 3.4 and 3.6, is shown in Figure 3.9. The curve is nonlinear, which means that interpretation of the raw experimental images must be done with caution. The regions of largest colour gradient do not necessarily correspond to the regions of steepest interface slope.

Strictly, we should calibrate hue against path length rather than interface height. The two are not exactly the same because of the parabolic correction tank, which is asso-

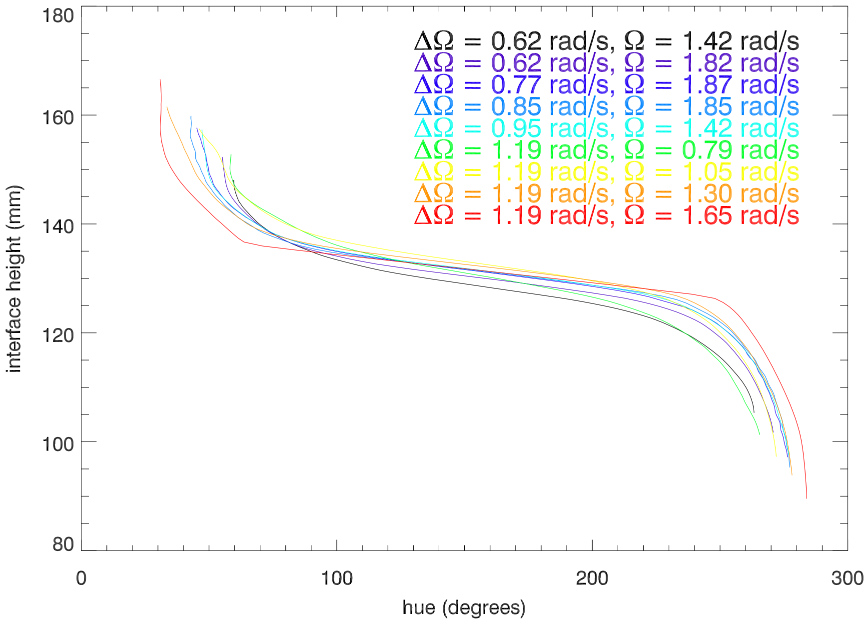


Figure 3.10: Family of calibration curves, each derived from different experimental conditions, as labelled. The curve shown in Figure 3.9 is included, together with eight others.

ciated with negative path lengths which vary spatially because of the parabolic shape taken up by the *l*-limonene. It follows from Section 2.4 that when $\Omega = 2$ rad s^{-1} the change in *l*-limonene depth across the gap is only 2 mm. This is sufficiently small that it is neglected here, i.e. we assume that the *l*-limonene surface is perfectly horizontal, and therefore has the same effect as simply altering the orientation of the entrance polaroid axis.

The method used to obtain the calibration curve in Figure 3.9 from the image in Figure 3.3 has been carried out eight further times, each time using a laboratory experiment image corresponding to a different combination of Ω and $\Delta\Omega$. The choices for these two variables were limited because only those combinations which give an axisymmetric, baroclinically-stable equilibrated flow will do. A calibration curve was derived in each case, and the nine curves are shown over-plotted in Figure 3.10.

Deriving such a family of calibration curves, for a range of values of Ω and $\Delta\Omega$, is an important test of the validity and accuracy of the calibration scheme. If the torque balance analysis had been an incomplete representation of the shear stresses acting on the annulus liquids, or if a mistake had crept into the calculation, we would expect a significant disagreement between different curves in the family. This is not the case, as the figure shows that all the curves have the same characteristic S-shape, and that there is good quantitative agreement between them. This implies that the error in the torque balance analysis is small, and that the analysis is a good representation of torques in the annulus. We conclude that the derived calibration curves are reasonably reliable and accurate. We may now proceed to use the curves to reconstruct interface height fields from given experimental images, including images in which the flow is not axisymmetric.

There is a spread in interface heights associated with any given hue, when inferred using the nine curves in Figure 3.10. This spread can be used to define an error in the inferred heights, which is less than ±3 mm across most of the range. It is important to note that errors in inferred wave amplitudes will be much less than this value, however. This is because inferred wave amplitudes are determined by the gradients of the calibration curves, rather than their absolute values, and there is excellent agreement across most of the range between gradients within the family of curves. In Section 3.6, we study another, much greater, source of error in inferred wave amplitudes.

3.6 Noise analysis

There is noise in the laboratory experiment images due to a phenomenon known as *pixel jitter*, which has contributions from the camera, the video recorder and the frame grabber. Pixel jitter causes the colour properties of each pixel to vary randomly in time about some mean value, even when the scene being shot by the camera remains exactly the same. It is the effects of this, and of the 8-bit colour quantization, which ultimately limit the vertical resolution of the inferred interface heights.

To estimate the size of the noise, colour information was extracted from a certain fixed pixel in 25 different frames, each showing the same resting interface in the annulus. Since there are no interface height changes between the frames, any variance in the colour properties is due entirely to pixel jitter. The standard deviations in the R, G, B, H, S and I data are shown in Table 3.3 for two cases: firstly, using images from a video recording, and secondly, using live images direct from the camera.

	video recording	direct from camera
red, R	3.6	2.5
green, G	4.2	2.9
blue, B	6.6	4.6
hue, H	9.6°	5.2°
saturation, S	3.4	1.9
intensity, I	10.3	9.0

Table 3.3: Noise in each of the six colour components, calculated as the standard deviation of the pixel jitter in 25 frames.

It is evident from the table that the noise can be reduced by about one third by using live images rather than video recordings, though video images were used in this thesis for reasons of convenience. In both cases the noise in the R, G and B signals is greater than one. This means that the accuracy of the inferred interface heights is ultimately limited by pixel jitter, rather than the discretization of R, G and B to integers.

A crucial issue is whether or not the signature of the small amplitude, small-scale waves will be visible through the noise in the hue field. We can investigate this by projecting the noise in the hue field onto the calibration curve to calculate the expected noise in the interface height field. For a calibration curve $h(H)$ giving interface height h in terms of hue H, we have

$$h(H + \epsilon_H) \approx h(H) + \epsilon_H \frac{\mathrm{d}h}{\mathrm{d}H} , \tag{3.30}$$

where ϵ_H is the noise in the hue field. The predicted pixel jitter noise ϵ_h in the inferred interface height fields is therefore given by

$$\epsilon_h \approx \epsilon_H \frac{\mathrm{d}h}{\mathrm{d}H} , \tag{3.31}$$

which is plotted against interface height for the calibration curve of Figure 3.9 in Figure 3.11, assuming images from a video recording with $\epsilon_H = 9.6°$ (see Table 3.3). We assume that ϵ_H is independent of interface height.

We can interpret the noise plotted in Figure 3.11 as the error in inferred heights, or alternatively as the smallest change in height that we can detect. A wave of amplitude

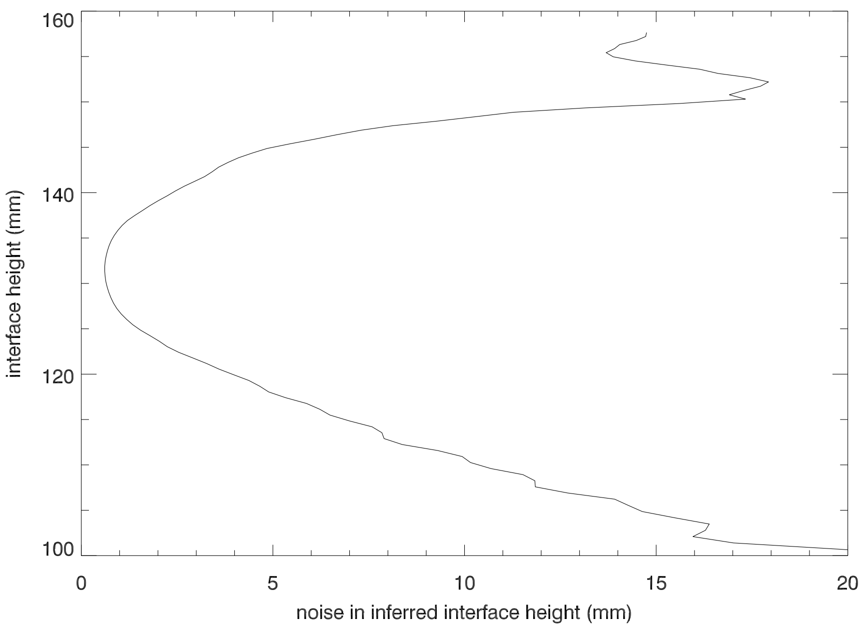


Figure 3.11: Resolution associated with the hue calibration curve, when used with images from a video recording.

1 mm could therefore be marginally resolved if it was superimposed on a background interface height of around 130 mm, but away from this height the resolution deteriorates rapidly.

If we average over a group of N neighbouring pixels in a frame, the noise decreases by a factor of $\sqrt{N}$. For example, if we averaged over a box measuring 20 pixels by 20 pixels centred on the pixel of interest, we would have $N = 400$ and noise reduction by a factor of 20. This would mean that we could resolve 1 mm amplitude waves at any observed interface height in the range of Figure 3.11, but the cost of this increase in vertical resolution is a decrease in horizontal resolution. 20 pixels corresponds to 8 mm, and so in this thesis the $N = 20 \times 20$ box averaging is performed only for inferring interface height features with large horizontal scales (the large-scale waves). Fortunately, the small-scale waves tend to occur near interface heights of 130 mm, and so for these we choose a box average of $N = 3 \times 3$ which gives adequate resolution in both the horizontal and vertical.

Another possible way to overcome the effects of pixel jitter would be to average over a number of frames which are sequential in time. The small-scale waves are so rapidly evolving that we could not average over more than around five frames (0.2 s) without losing resolution, and even then the signal-to-noise gain would be only $\sqrt{5} \approx 2.2$, so time-averaging was not attempted in this study.

3.7 Sample reconstructed height field

An example of an application of the calibration scheme is shown in Figure 3.12. The annulus rotation rates in this case were $\Omega = 0.46$ rad s^{-1} and $\Delta\Omega = 3.70$ rad s^{-1}, at which there is a rotationally-modified Kelvin-Helmholtz instability but no baroclinic instability. The equilibrated flow is a Kelvin-Helmholtz mode of azimuthal wavenumber 9. The wave has a particularly large wavelength and amplitude due to the very large lid rotation. The figure shows the reconstructed 2-D interface height, as well as an azimuthal cross-section at radius $r = 100$ mm, in which the angle increases in the anti-clockwise direction and has its zero at “3 o’clock”. Parallax effects described in Section 2.5 were taken into account to produce these images, and all other reconstructed interface height images in this thesis. We deduce that the amplitude of the interface height displacement is around 10 mm, at this radius.

Because there are still slight remnants of pixel jitter noise present, a filtered version of the azimuthal cross-section is shown over-plotted on the unfiltered curve. The filtering method used involves projecting the curve onto the Daubechies (1988) Discrete Wavelet Transform (DWT) functions, setting those coefficients associated with random noise equal to zero, and then reconstructing the curve from the remaining (non-zero) basis functions. This is implemented using the IDL routine `wtn`. The DWT functions are localized and compact (hence the term *wavelet*) and therefore optimized for noise removal, unlike filtering based on the Fast Fourier Transform (FFT) basis functions (global sine and cosine *waves*). The reconstructed height fields in this thesis are all filtered in this way.

Further interface height reconstructions are performed in Section 4.8, after a comprehensive series of laboratory experiments has been carried out.

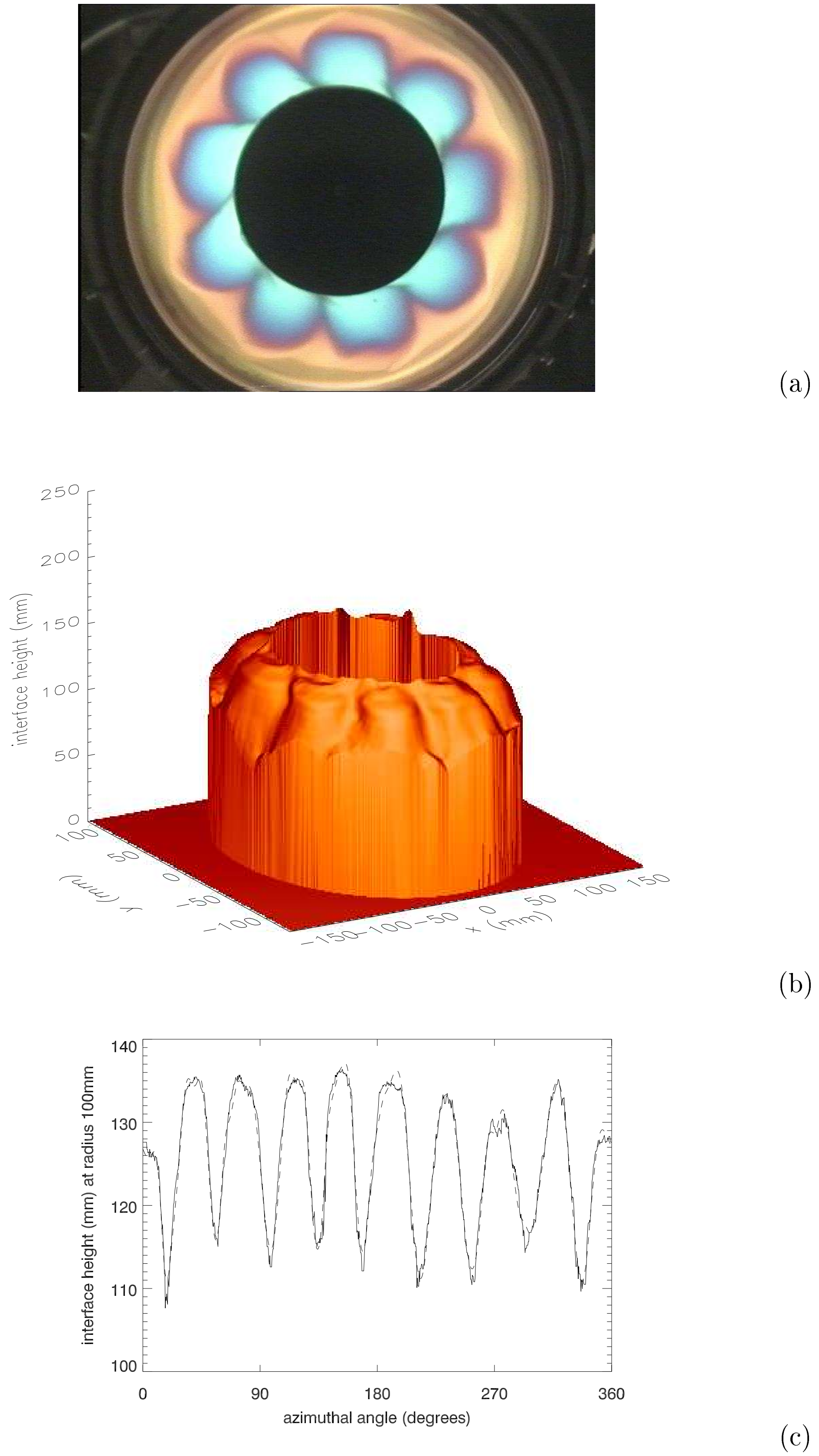


Figure 3.12: (a) Raw experimental image; (b) reconstructed 2D interface height; and (c) unfiltered (solid) and filtered (dashed) azimuthal interface height profile.

3.8 Chapter summary

We have successfully implemented a calibration scheme for interface height in the laboratory experiment. By iterative solution of the nonlinear layer torque balance equations, we have found that Ekman layers and Stewartson layers both exert a significant drag force on the layer interiors, and that allowing for sloping interface heights is also important for an accurate result. We have identified hue as an optimal colour calibration variable, and have used a multi-layer torque balance analysis to derive a calibration curve.

By projecting hue onto the calibration curve we have been able to make accurate reconstructions of interface height maps, and thereby derive wave amplitudes in the annulus for the first time. The vertical resolution of the inferred interface heights is limited by pixel jitter, but it can be better than 1 mm if live images are used and an average is taken over a number of pixels which are neighbouring in space or time.

In the next chapter, we describe the particular laboratory experiments which have been carried out in this study, and analyze the results.

Chapter 4

Results of the laboratory experiments

"Are you ready to kick it into action yet?"
"Oh, I'm ready to kick it alright!"

We begin this chapter by summarizing the experimental results obtained by Lovegrove (1997) and Lovegrove et al. (2000) using the rotating, two-layer annulus apparatus. We then describe three new series of experiments which have been carried out in the present study, mentioning some preliminary experimental difficulties which frustrated progress during the first year. The new series are designed to extend Lovegrove's results, both by exploring a more extreme range of parameters including the hitherto uninvestigated ageostrophic regime, and by investigating the effects of three particular system changes.

The new experiments prove to be fruitful, as large- and small-scale waves are found to coexist at a greater number of regions in parameter space than reported by Lovegrove. The locations of the major transition curves in the present experiments are shown to compare well with the predictions of simple theory.

With one particular set-up, we find that we can make video recordings of the waves with a horizontal spatial resolution of better than one fifth of a millimetre, corresponding to around 100 measurements per small-scale wavelength.

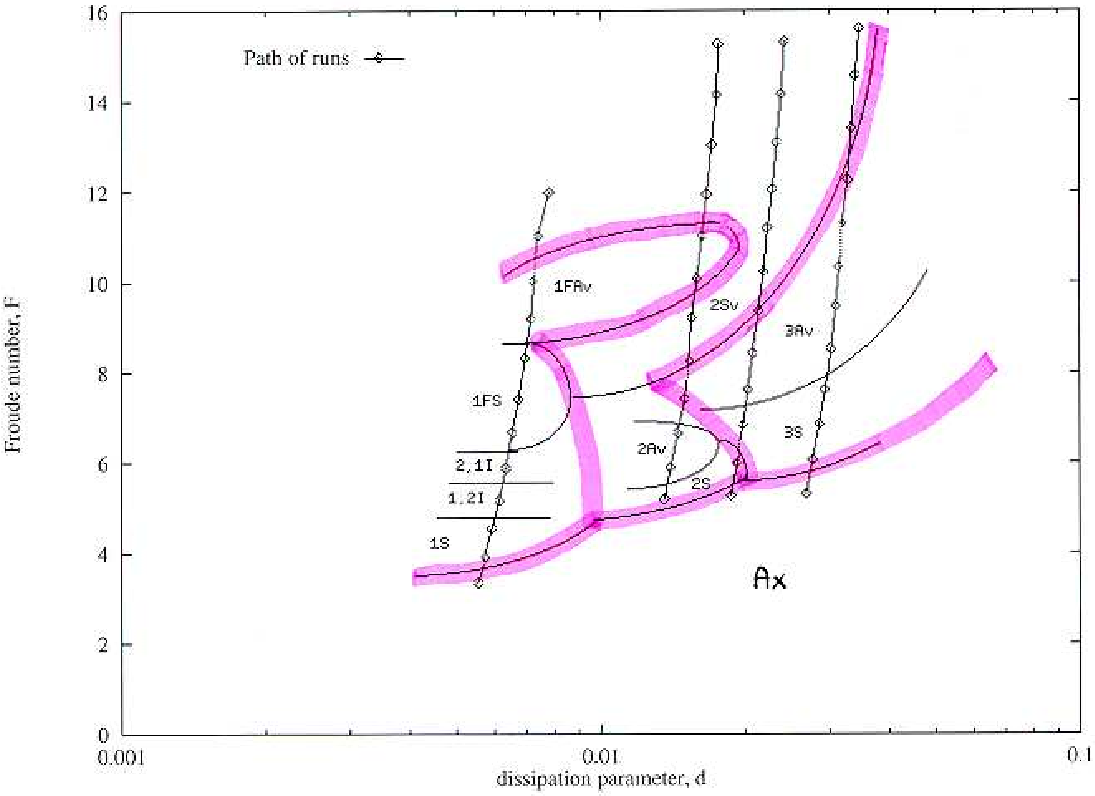


Figure 4.1: Regime diagram obtained by Lovegrove et al. (2000). The paths of the four scans are shown. In all cases the direction of the paths was towards higher Froude number. Categorization of the equilibrated large-scale flow was performed at each of the 48 points marked with a diamond, allowing the approximate locations of the transition curves to be inferred (thin curves). The diagram is also divided into regions according to the dominant azimuthal wavenumber (thick curves).

4.1 Findings of Lovegrove

Lovegrove performed a general survey of the different equilibrated large-scale flow types achieved in the baroclinically-unstable rotating, two-layer annulus. His results are given in terms of two dimensionless parameters: the *internal Froude number* (F) and *dissipation parameter* (d), which are defined by

$$F = \frac{f^2 L^2}{g' H} \tag{4.1}$$

and

$$d = \frac{\sqrt{\overline{\nu}\Omega}}{H \Delta\Omega}\ , \tag{4.2}$$

where $L = 6.25$ cm is the radial gap width, $H = 12.5$ cm is the layer depth, $f = 2\Omega$ is the Coriolis parameter, $g' = 2g(\rho_2 - \rho_1)/(\rho_1 + \rho_2) = 6$ cm s^{-2} is the reduced gravity and $\overline{\nu} = (\nu_1 + \nu_2)/2 = 1.18 \times 10^{-6}$ m^2 s^{-1} is the mean kinematic viscosity.

Lovegrove performed four scans in the (d, F) parameter space, in each one keeping $\Delta\Omega > 0$ constant, and increasing $\Omega > 0$ in twelve discrete steps over a six hour period. This meant that 30 minutes was spent at each of the 48 visited points of parameter space, sufficiently longer than the spin-up time of around ten minutes, and therefore allowing a reliable post spin-up flow categorization to be determined.

The results of Lovegrove's experiments are shown in Figure 4.1. The flow types are denoted by a number followed by one or more letters. The number refers to the dominant azimuthal wavenumber(s), and the letters refer to how the flow evolves with time: Ax denotes *axisymmetric* flow of azimuthal wavenumber zero; S denotes a *steady* wave whose amplitude is constant with time; Av denotes an *amplitude vacillation* wave whose amplitude periodically grows and decays with time. There are further, more complicated flow types present in the regime diagram which are not studied in this thesis.

The main focus of Lovegrove's work was transitions between these different large-scale flow types, and he found good agreement between the laboratory regime diagram, and one based on a bifurcation analysis of a theoretical model using spectral amplitude equations. However, he also reported the presence of small-scale waves during amplitude vacillation (Av) flows only. The small-scale waves would develop near the inner cylinder during decaying phases of the vacillation, and would be completely absent during growth phases.

This was the first time that a systematic small-scale wave presence had been reported in a rotating, two-layer annulus experiment. Read (1992a) had reported the presence of weak, high-frequency wave activity in the signal from a thermocouple probe embedded in a thermally-driven, continuously-stratified rotating annulus flow. The recorded frequency was close to the buoyancy frequency, suggesting that the origin of the signal was an inertia-gravity wave. However, the inertia-gravity wave could well have been generated by an interaction between the probe and the flow, and so this cannot be regarded as evidence of spontaneous emission.

4.2 Preliminary difficulties with current experiments

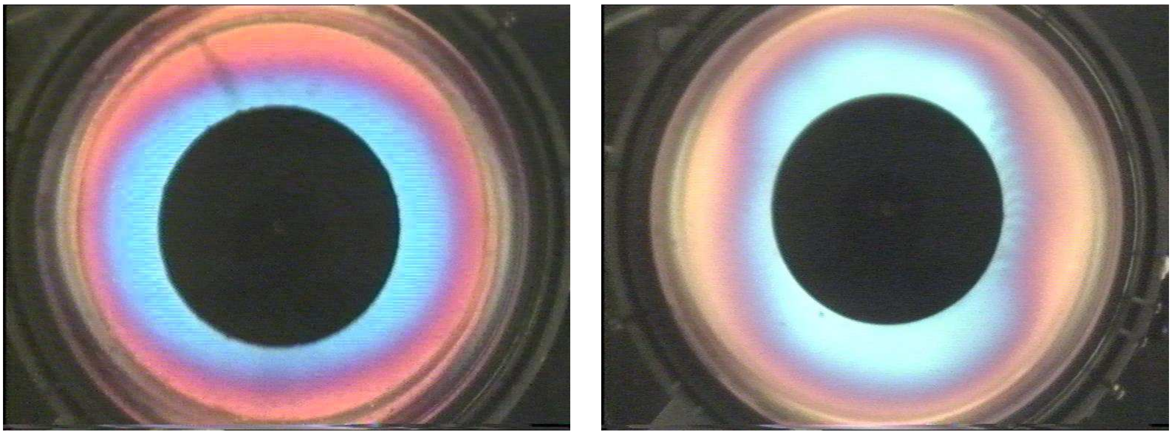
Difficulties were encountered in this study, when an attempt was made to reproduce the results of Lovegrove. In preliminary experiments, the fit between the observed regime diagram transition curves and those shown in Figure 4.1 was unsatisfactory. In particular, even though points across the entire regime diagram were sampled, equilibration was found to be nearly always to a wavenumber two flow. The fluid displayed a reluctance to undergo transitions to different states with wavenumbers other than two. Furthermore — in stark contrast with Lovegrove's findings — all flows were completely devoid of small-scale waves.[1] In an attempt to fix this problem, the working liquids were replaced with fresh preparations on a number of occasions, but the small-scale waves remained absent.

After almost a year of failed experiments, some time was spent working on a non-experimental part of the project. Surprisingly, when the experiments were attempted again after this break, the small-scale waves appeared in abundance. The only change in the apparatus between the unsuccessful and successful experiments, was that the fluids were six weeks older. We therefore infer a change in the liquids' physical and/or chemical properties over time, as previously suggested in Section 2.2.4. It is not surprising that such changes might affect small-scale waves more than large-scale ones, since viscous and interfacial tension effects are both scale-selective. Changes in the liquid properties could therefore make the difference between presence and absence of small-scale waves, whilst simultaneously exerting a significant but lesser impact upon the large-scale waves.

Figure 4.2 shows a comparison between the equilibrated flows obtained with fresh and aged fluids. The experimental details are otherwise identical, with $\Omega = 2.3$ rad s^{-1} and $\Delta\Omega = 0.62$ rad s^{-1}. A train of short waves is present in the experiment with old fluids, but with the fresh preparation no short waves are visible, even in the original video footage. In both cases, an experiment was performed in which $\Delta\Omega$ was held constant at the value just given, but Ω was slowly and continuously increased from its starting value at a rate of 4×10^{-4} rad s^{-2}, in anticipation of the experiments to be described in Section 4.3. In the system with short waves present, the flow underwent a transi-

---

[1] An undergraduate using the annulus had also previously failed to produce a flow with small-scale wave activity, during her Masters project.



(a) (b)

Figure 4.2: Raw laboratory images showing the post-transient flow obtained in experiments which used (a) fresh working liquids, and (b) aged working liquids, but which were otherwise dynamically identical. During the many weeks which passed between the two experiments, the source lamp broke and was replaced with a brighter one, the automatic gain control (AGC) feature of the camera was switched off, and the orientation of the analyzing polaroid was altered. These modifications explain the qualitative differences between the two images.

tion to a wavenumber 3 state when Ω reached 2.7 rad s^{-1}. The same transition was observed in the system without short waves, but it did not occur until Ω had reached 3.4 rad s^{-1}, suggesting that the short waves play a role in encouraging transitions. It was difficult to study this phenomenon systematically, because the fluid properties change in an uncontrolled and unknown way. Furthermore, the observation could be due to the direct impact of the changing fluid properties upon the large-scale, balanced dynamics, rather than to the impact of the short waves — whose generation is permitted by the changing fluid properties — on the balanced dynamics. Because of these difficulties, this phenomenon is not investigated further here, but is studied more carefully using a numerical model in Chapter 7.

There is a need for further work to determine exactly how the fluid properties change with age. Discrepancies between the results of this study and those of Lovegrove are discussed, in the context of evolving fluids, in Section 4.9. To explore the impact of the uncertain fluid properties, simulations using a numerical model with varying viscosity and interfacial tension are described in Sections 6.2 and 6.3, respectively.

The rest of this chapter describes the laboratory experiments which were carried out when the working liquids had been together in the annulus for between two and six

experiment number	$\Delta\Omega$ (rad s^{-1})
1	0.62
2	0.69
3	0.77
4	0.85
5	0.95
6	1.06
7	1.19
8	1.31
9	1.46
10	1.61
11	1.80
12	2.03
13	2.24
14	2.51
15	2.73
16	3.14

Table 4.1: The magnitude of the constant differential lid rotation $\Delta\Omega$ used in each of the present experiments, measured by timing the rotation period using a stopwatch. The error in the measurements is around 1%.

months. At this age, direct observation of the liquids and interface suggested that the fluid properties had become relatively stable.

4.3 Description of current experiments

In the current study, four series of laboratory experiments have been performed, each using a slightly different configuration of the two-layer annulus. Each series involves many scans across (d, F) parameter space, though the range of parameters covered is much greater here than in Lovegrove's experiments. Also, we are interested here not so much in determining the precise large-scale flow type, but primarily in locating the regions of existence and coexistence of large-scale and small-scale waves.

All experiments had $\Omega > 0$ so that, when viewed from above, the turntable rotation was anti-clockwise in the laboratory frame. In the first series, denoted *PAI (prograde, annu-*

lus, increasing), the lid rotation was prograde ($\Delta\Omega > 0$), the geometry was annular, and the runs were performed with increasing Ω (and therefore increasing Froude number). This series corresponds to the experiments of Lovegrove. In the second series, denoted *RAI (retrograde, annulus, increasing)*, the lid rotation was retrograde ($\Delta\Omega < 0$), the geometry was annular, and the runs were performed with increasing Froude number. This is an interesting case to study as there is an asymmetry between the large-scale dynamics of prograde and retrograde flow, as we will see in Section 4.5. In the third series, denoted *PAD (prograde, annulus, decreasing)*, the lid rotation was again prograde and the geometry annular, but the runs were performed with decreasing Froude number, which allows us to investigate the effects of hysteresis. In the fourth series, denoted *PEI (prograde, eccentric, increasing)*, the lid rotation was prograde and the Froude number was increasing, but the inner cylinder was displaced horizontally to give an eccentric annular geometry. The purpose of reversing the lid rotation and the direction of the parameter space scan, and of modifying the geometry, is to see whether any of these system changes affects the production of small-scale waves.

Each of the four series consisted of 16 experiments, in which the differential lid rotation $\Delta\Omega$ was held constant at the magnitude shown in Table 4.1. In each experiment, lasting three hours, the voltage supplied to the turntable motor was linearly increased or decreased with time under computer control. This caused the turntable rotation rate Ω to vary slowly as shown in Figure 4.3. The angular acceleration/deceleration $\mathrm{d}\Omega/\mathrm{d}t$ produced was closely constant at 4×10^{-4} rad s^{-2}. By holding $\Delta\Omega$ constant and gradually increasing Ω, a curve is traced out in the (d, F) parameter space. In contrast with Lovegrove's experiments, in which 48 discrete points in parameter space were sampled, in the present experiments we perform continuous scans.

We now look, in turn, at the results from each of the four series of laboratory experiments.

4.4 Experimental results: PAI series

During a typical experimental run in the *prograde, annulus, increasing* PAI series, the flow types observed fell into the following four distinct classes, when categorized according to the presence and absence of large- and small-scale waves. The flow types are

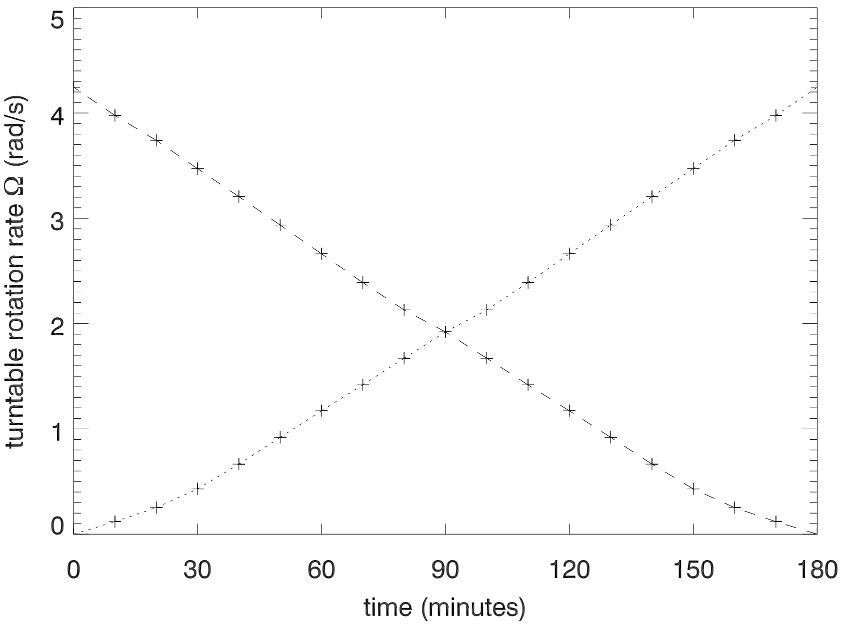
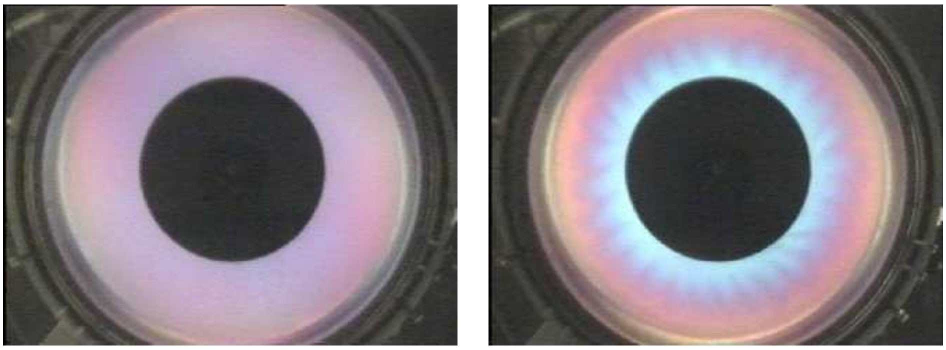
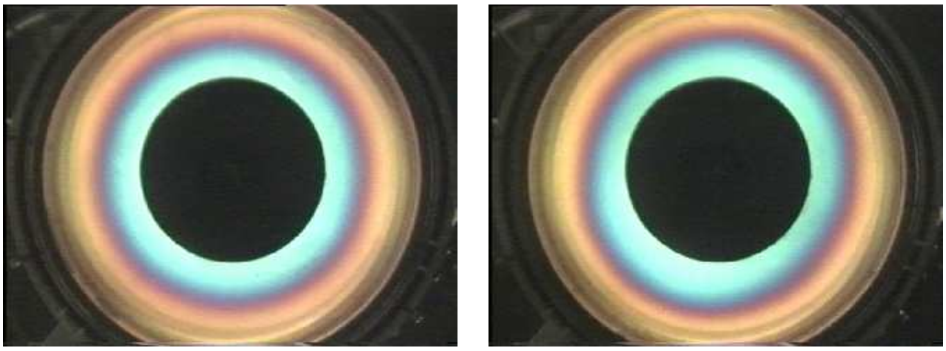


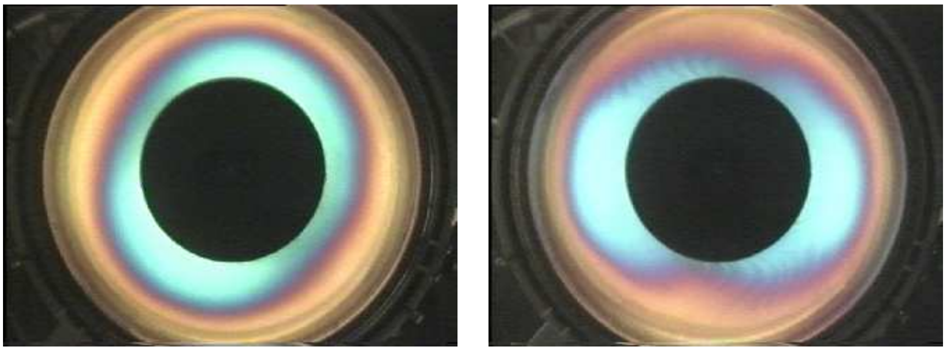
Figure 4.3: Measured temporal variation of the turntable rotation Ω used in the laboratory experiments. In the PAI, RAI and PEI experiments the computer-generated voltage, input to an amplifier which supplies power to the turntable motor, was linearly increased from zero to 3 V over three hours, to give the dotted curve. In the PAD experiments, the voltage was linearly decreased over three hours, to give the dashed curve. The curves were derived by making measurements of the rotation period every 10 minutes, by electronically timing successive breakings of a light beam by a tab fixed to the turntable at its circumference.



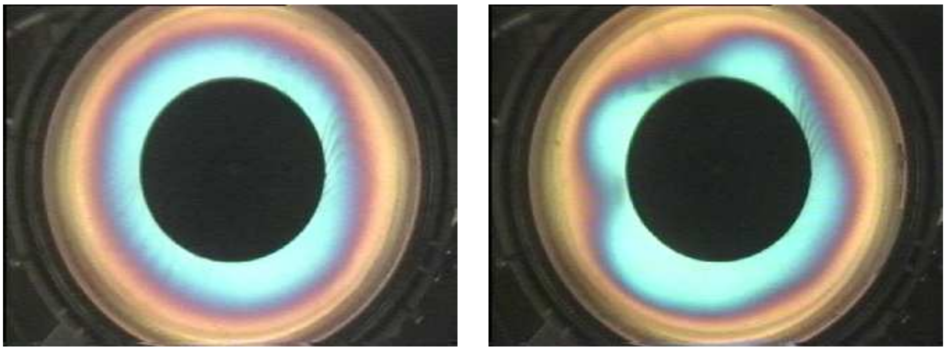
(a) 5 min (b) 35 min



(c) 70 min (d) 93 min 00 sec



(e) 93 min 30 sec (f) 94 min 15 sec



(g) 95 min 00 sec (h) 142 min

Figure 4.4: Images from experiment number PAI5, at the indicated times.

illustrated in Figure 4.4 by representative images from experiment number PAI5.

- **Axisymmetric flow regime, *AX*.** Absence of both large-scale and small-scale waves, e.g. Figure 4.4(a) and (c).
- **Kelvin-Helmholtz regime, *KH*.** Absence of large-scale waves with global presence of small-scale waves, e.g. Figure 4.4(b) [and also Figure 3.12(a)].
- **Mixed Regular Wave regime, *MRW*.** Presence of regular large-scale waves together with local presence of small-scale waves, e.g. Figures 4.4(d)–(g) which show a large-scale amplitude vacillation cycle.
- **Mixed Irregular Wave regime, *MIW*.** Presence of irregular large-scale waves together with local presence of small-scale waves, e.g. Figure 4.4(h).

In general, during a three hour run, the order in which the flow types were encountered was *AX*, *KH*, *AX*, *MRW*, *MIW*, so that there were four transitions in total. The KH regime is so-called in anticipation of evidence to be presented in Section 4.4.1 that the small-scale waves in that regime are, indeed, Kelvin-Helmholtz waves. For practical purposes, wavenumbers below five were counted as large-scale waves, and all others as small-scale waves, though typically the spectral gap was much larger than permitted by this definition.

In all flow regimes, the large- and small-scale waves described are superimposed onto a background parabolic interface shape given by equation (3.9), as radial pressure gradients are established to provide the required centripetal acceleration. The transition between the *AX* and *KH* regimes marks the onset of a rotationally-modified *Kelvin-Helmholtz instability* which gives rise to the observed global small-scale waves, as we will show shortly. The return to the *AX* regime marks an instability boundary, at which the turntable rotation is sufficiently large to re-stabilize the system against these waves. The start of the *MRW* regime corresponds to the onset of *baroclinic instability* which gives rise to the observed large-scale waves, with coexisting localized trains of small-scale waves. The transition to the *MIW* regime marks the point at which the forcing is so strong that the spatio-temporal regularity of the large-scale waves breaks down to leave a chaotic flow.

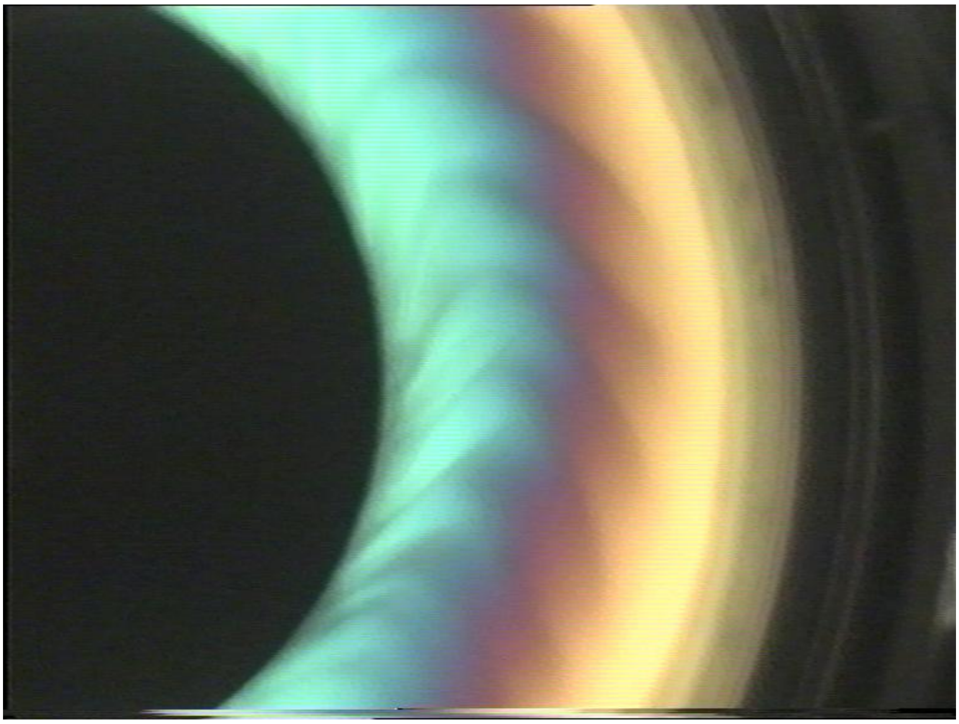


Figure 4.5: Close-up of the small-scale waves radiated during the decay phase of a 2Av large-scale flow in a laboratory experiment with $\Delta\Omega = 1.46$ rad s^{-1} and $\Omega = 1.94$ rad s^{-1}, corresponding to experiment PAI9 at a time of 90 minutes. The annular gap width (62.5 mm) measures 334 pixels in the image, and so the horizontal resolution in this close-up view is 0.19 mm. The small-scale wavelength is around 20 mm, so the waves are very well resolved.

Small-scale waves were observed during almost every baroclinically-unstable flow (the *MRW* and *MIR* regimes) in the present experiments, including flows in which the large-scale wave amplitude was constant in time. This is in contrast with Lovegrove's experiments, in which small-scale waves were reported to appear only during amplitude vacillation flows. Possible reasons for this apparent inconsistency are give in Section 4.9.

As a one-off, a special experiment was performed in which the camera zoom lens was used to zoom in to the annulus as far as possible. The camera was also shifted horizontally by around 10 cm, so that the field of view was centred not on the rotation axis but on the annular gap. This removes the parallax effect discussed in Section 2.5, which blocks from view a significant portion of the interface height adjacent to the inner cylinder, which is exactly where the small-scale waves tend to appear. A close-up of the radiated small-scale wave field so obtained is shown in Figure 4.5. Only an azimuthally-restricted part of the flow can be seen, but the advantage is that the horizontal structure of the wave-train can be seen at exceptionally high resolution (0.19 mm — see figure caption).

For each of the 16 experiments in the PAI series, the times of transitions between the four flow classes (in minutes since the start of the experiment) were recorded. $\Delta\Omega$ is known from Table 4.1 and Ω can be determined from the transition time using Figure 4.3, allowing the Froude number (F) and dissipation parameter (d) corresponding to each transition to be calculated using equations (4.1) and (4.2). The results thereby obtained are summarized in the regime diagrams of Figure 4.6. Contours of constant Rossby number $\mathrm{Ro} = \Delta\Omega/(2\Omega)$ are shown in Figure 4.6(b) for reference, showing that we have investigated flows in which this parameter varies by three orders of magnitude, encompassing the geostrophic $\mathrm{Ro} < 1$ and ageostrophic $\mathrm{Ro} > 1$ regimes. No attempt was made during this study to reproduce Lovegrove's detailed sub-classification within the *MRW* regime. The full Froude number span was not always achieved, as some of the higher-numbered experiments were terminated before the full three hours had elapsed. This was because the fluid interface had become so steep that it began to intersect the lid.

As regards the large-scale waves, there is good agreement between the location of the marginal baroclinic instability curve in the present and Lovegrove experiments. The Phillips model (e.g. Pedlosky, 1987) predicts a Froude number for this curve of $F_{\text{critical}} = \pi^2/2 \approx 4.9$, independent of the shear (i.e. the dissipation parameter). This is in reasonable agreement with the Froude numbers for the transitions to the *MRW* regime in Figure 4.6. There seems to be a weak dependence on shear in the laboratory which is not captured by the model, possibly because the model is for a channel rather than an annulus, and because of ageostrophic effects not captured by the model.

The small-scale modes of main interest have different characteristic properties when they appear in the *MRW* and *MIW* regimes rather than in the *KH* regime. This can be seen by comparing Figures 4.4(b)&(h), for example. In the *MRW* and *MIW* regimes, the short waves are generally smaller in amplitude and wavelength, and are radially and azimuthally confined unlike in the *KH* regime. This suggests that different generation mechanisms may be responsible in the two cases, a possibility which we investigate in Sections 4.4.1 and 4.4.2.

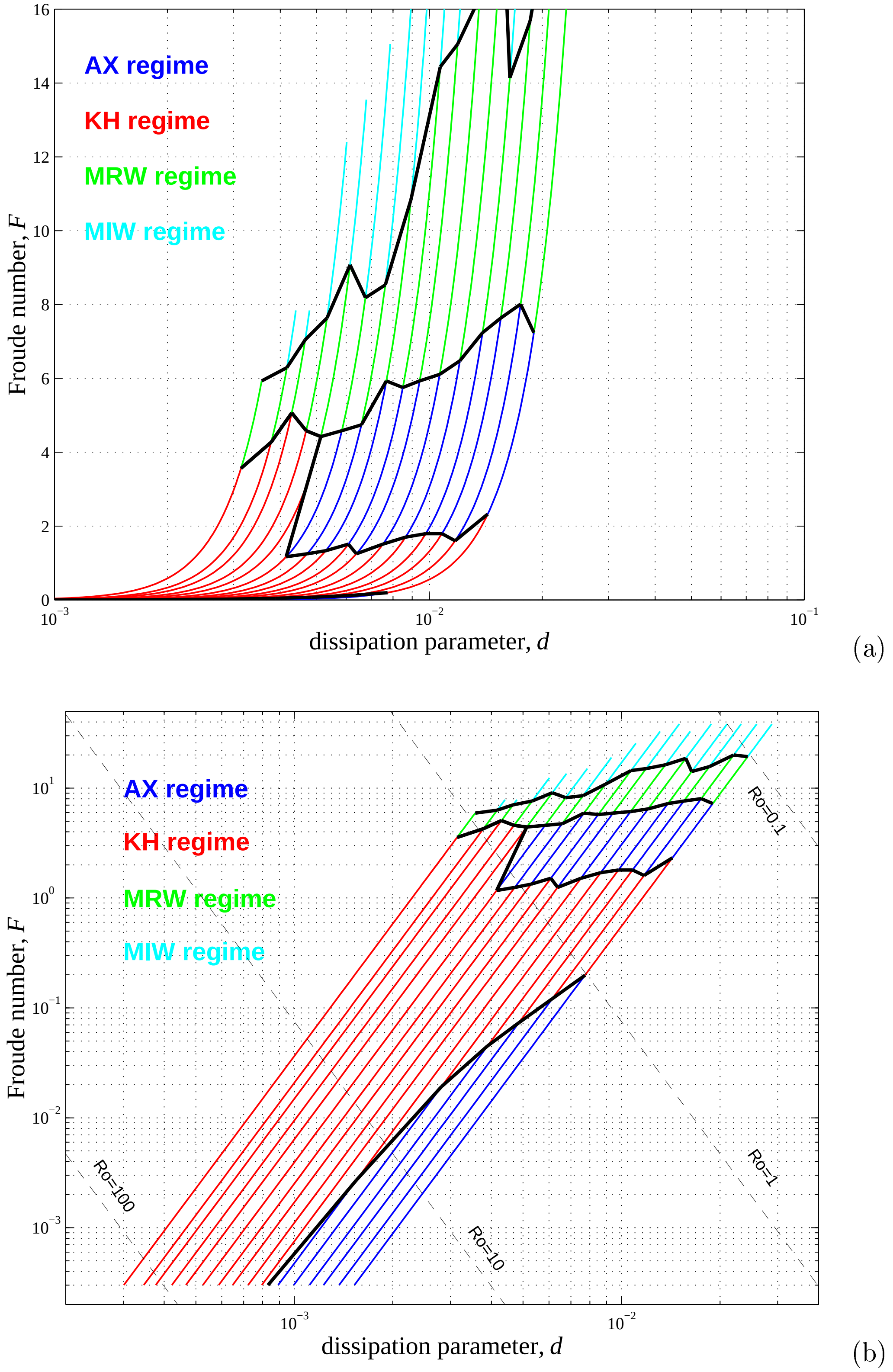


Figure 4.6: Experimental regime diagrams for the PAI series, plotted using two different sets of axis scales and limits. In (a) the axes are identical to those in Lovegrove's regime diagram, Figure 4.1, which focuses exclusively on the *MRW* regime. In (b) the full span of the present experimental paths is shown. Each multi-coloured line corresponds to one 3-hour experiment, with experiment number 1 furthest to the right and experiment number 16 furthest to the left.

4.4.1 Analysis of small-scale waves in the ***KH*** regime

We now show that the *KH* regime small-scale waves are Kelvin-Helmholtz modes generated by a shear instability, as anticipated by the label. The transfer of energy from the large-scale basic flow to the small-scale growing waves is achieved in this case through a wave-mean flow interaction, unlike the wave-wave triad interaction mechanism discussed in Section 1.1. Kelvin-Helmholtz modes are expected to be generated whenever the vertical shear in horizontal velocity exceeds a particular value. This criterion is expressed in non-dimensional terms as the Richardson number dropping below some critical value, usually taken to be $\frac{1}{4}$ or 1 depending upon the exact definition of the Richardson number (Acheson, 1990).

The gradient Richardson number for a continuously stratified flow is defined by

$$\mathrm{Ri} = \frac{-(g/\rho)(\partial\rho/\partial z)}{(\partial u/\partial z)^2} \; , \tag{4.3}$$

where $\rho(z)$ and $u(z)$ are the density and horizontal velocity profiles, respectively. For an axisymmetric two-layer annulus flow, with solid-body rotation rates in layers 1 and 2 of $\Delta\Omega_1$ and $\Delta\Omega_2$ relative to the turntable, the bulk Richardson number, vertically-averaged over the interfacial Ekman layers — both of width $\sqrt{\nu/\Omega}$ from equation (3.14) — is, at radius r,

$$\mathrm{Ri} = \frac{2g'\sqrt{\nu/\Omega}}{r^2(\Delta\Omega_1 - \Delta\Omega_2)^2} \; . \tag{4.4}$$

The Richardson number criterion stated above strictly only applies to non-rotating systems. The applicability of non-rotating Kelvin-Helmholtz instability theory to rotating systems has been investigated by James (1977). He derives an implicit fourth order polynomial dispersion relation for a rotating, two-layer channel, and numerically solves it to plot Kelvin-Helmholtz growth rate curves for both the rotating and non-rotating cases. A comparison of the curves leads him to conclude that

> "*the Kelvin-Helmholtz instability is but little affected by rotation ... broadly, this [rotating case] instability is adequately described by the non-rotating theory.*"

This statement holds because the laboratory small-scale waves evolve on timescales of much less than the rotation period. We conclude that we may proceed to apply the

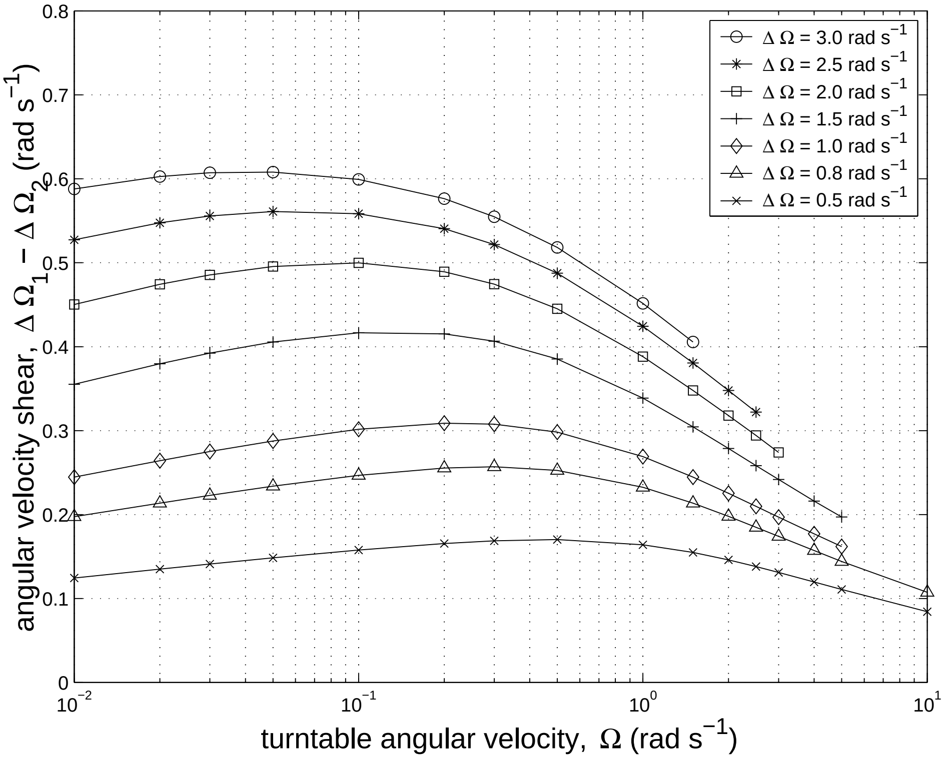


Figure 4.7: Variation of angular velocity difference across the fluid interface with turntable angular velocity Ω, for various fixed values of the differential lid rotation $\Delta\Omega$. The curves are obtained from solutions of the annulus torque balance equations (3.27) derived in Chapter 3. For combinations where both Ω and $\Delta\Omega$ were large, the iterative solution described in that chapter failed to converge, and so these points are missing from the figure.

non-rotating criterion to the rotating annulus system.

To a first approximation, $\Delta\Omega_1$ and $\Delta\Omega_2$ are both proportional to the differential lid rotation $\Delta\Omega$ and independent of the turntable rotation Ω. We developed, in Chapter 3, an iterative method for calculating $\Delta\Omega_1$ and $\Delta\Omega_2$ (and hence the shear $\Delta\Omega_1 - \Delta\Omega_2$) for given $\Delta\Omega$ and Ω. We can use the same torque balance analysis here, too, to improve on the first approximation for the shear. The results of this calculation for each combination of 7 different values of $\Delta\Omega$ and 16 different values of Ω, are shown in Figure 4.7. Both Stewartson layers and curved interface effects are included in the torque balance calculation, for accuracy. It can be seen that the shear actually shows a significant variation with turntable rotation for fixed differential lid rotation. This variation is unreported in the annulus literature, and is due to two effects. Firstly, as the turntable angular velocity is increased the fluid interface becomes more curved, resulting in a greater area over which interface drag forces act. Secondly, the Stewartson and Ekman

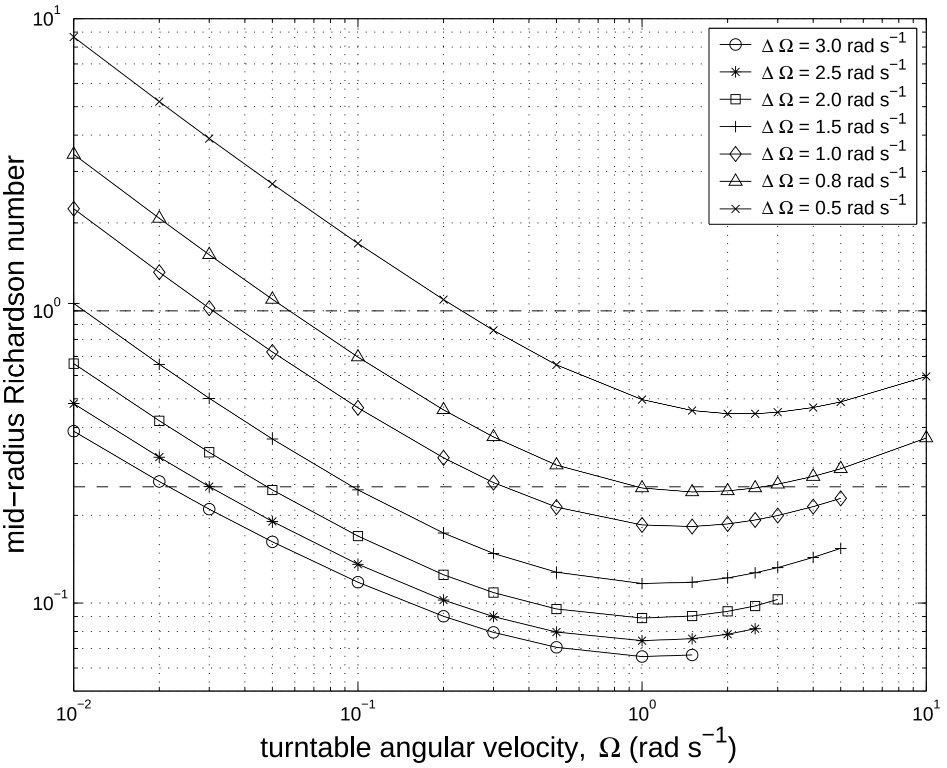


Figure 4.8: Variation of mid-radius Richardson number averaged over the interfacial Ekman layers, with turntable angular velocity, for various fixed values of the differential lid rotation. Dashed lines corresponding to critical Richardson numbers of 1/4 and 1 are also drawn.

layer thicknesses shrink according to different powers of Ω according to equations (3.14) and (3.15), resulting in a shift in the balance of drag forces due to these boundary layers. As anticipated by Table 3.2, the shear is around half that which would be obtained by the simpler calculation, with an assumed-horizontal interface and neglected Stewartson layers.

For each of the points plotted in Figure 4.7 we can calculate the Richardson number using equation (4.4). The results of this calculation at mid-radius $r = 9.4$ cm are shown in Figure 4.8. Because of the characteristic shapes of the velocity shear curves we can see that, as Ω is increased with $\Delta\Omega$ held constant, we encounter Richardson numbers first greater than the critical values required for Kelvin-Helmholtz instability, then less than, and then greater than again. This is exactly what we observed during the transitions to and from the AX and KH regimes in the experiments of Section 4.4, with the onset of Kelvin-Helmholtz instability being followed shortly after by a re-stabilization, as seen in Figure 4.6(b).

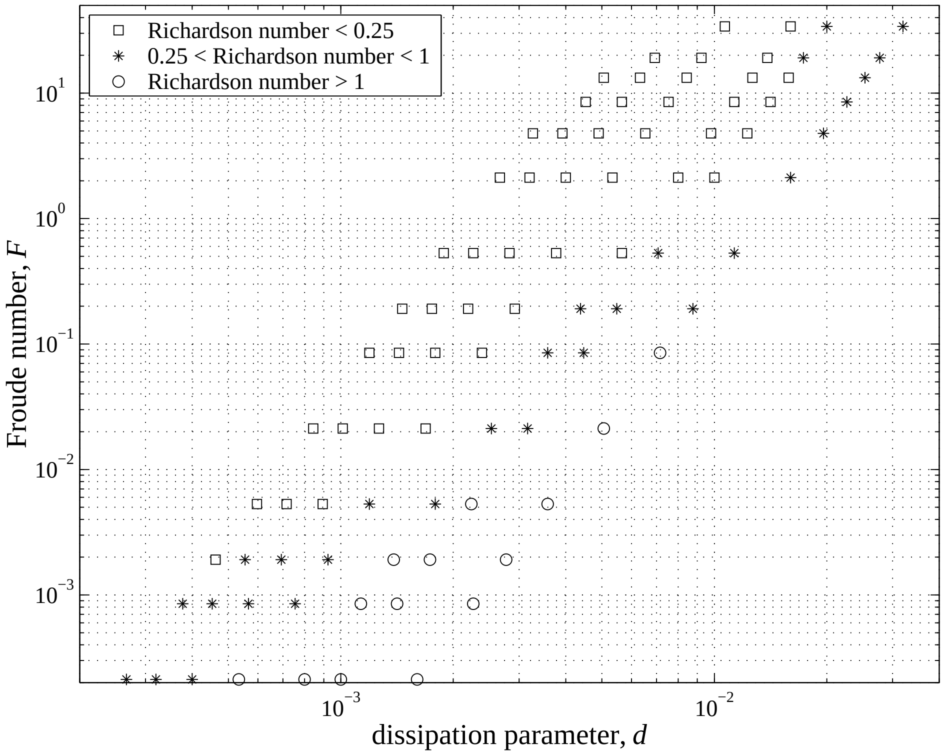


Figure 4.9: Regime diagram showing the predicted regions of Kelvin-Helmholtz stability and instability in the (d, F) plane. Instability is expected wherever the Richardson number drops below a critical value, which is shown in this plot as 0.25 and 1.

To further test the ability of this simple model to predict Kelvin-Helmholtz instability in the laboratory, we evaluate the dissipation parameter d and Froude number F at each of the points in Figure 4.8, and plot a point in (d, F) space according to whether the Richardson number is sub- or super-critical. The results are shown in Figure 4.9. There is good qualitative agreement between regions of Kelvin-Helmholtz instability in this figure and in the experimental regime diagram of Figure 4.6(b). The AX and KH flow types are both reproduced by this simple analysis for the two shown values of critical Richardson number. Their detailed shapes are not reproduced exactly, probably due to the neglect of interfacial tension in the analysis.

We are now in a position to explain the shape of the AX and KH regions in Figure 4.6(b). The initial AX regime is missing from experiments 7–16 because the shear is so large that the system is Kelvin-Helmholtz unstable even at the start of these experiments. Similarly, the larger Froude number AX regime is missing from experiments

12–16 because the shear is so large that the system does not re-stabilize with respect to the Kelvin-Helmholtz mechanism before baroclinic instability occurs.

4.4.2 Analysis of small-scale waves in the ***MRW*** and ***MIW*** regimes

We will conduct an investigation of the production of the *MRW* small-scale waves in Chapter 6, using velocity fields from a numerical model. However, there is a simple analysis that does not require velocity data, which can be done now. The analysis is based on the ship wave problem studied by Lighthill (1978), in which an object moving in a straight line at speed V in a fluid generates a wake of deep water surface gravity waves. It is shown in that analysis, using geometrical arguments based on the phase speed being twice the group speed, that waves at the edge of the wake will have a wavelength of $\lambda = (4\pi/3)(V^2/g)$, and that their crests will meet each other at an angle of $\alpha = 90^\circ + \sin^{-1}\left(\frac{1}{3}\right) \approx 109^\circ$.

This suggests a model for the *MRW* and *MIW* small-scale wave generation in the laboratory experiments, in which the entire large-scale wave is taken to be the extended moving object which causes inertia-gravity wave generation. The Lighthill theory is non-rotating, but as in Section 4.4.1 we assume that the laboratory small-scale waves are not strongly influenced by rotation. Since the speed at which the the large-scale wave travels around the annulus is proportional to $\Delta\Omega$, the model would suggest a small-scale wave crest intersection angle which does not vary between experiments, and a wavelength which increases with increasing experiment number.

A comparison between Figures 4.4(h) and 4.5, from mixed wave flows in experiments PAI5 and PAI9 respectively, reveals that these predictions are consistent with observations. The wave crest angle is around 90° in each case, reasonably close to the predicted angle. The drift period of the large-scale wave in Figure 4.5 was measured to be 33 s using a stop-watch. Converting to an angular phase speed and then to a velocity at mid-radius, $r = 9.4$ cm, gives $V = 1.78$ cm s^{-1}. With this velocity, and using the reduced gravity in place of g, the Lighthill theory predicts $\lambda = 23$ mm, in excellent agreement with the observed wavelength given Lighthill's assumptions of no rotation and a point generating object. The same analysis for the flow in Figure 4.4(h), for which the drift period is 45 s, gives $\lambda = 12$ mm, which is also in good agreement with observations.

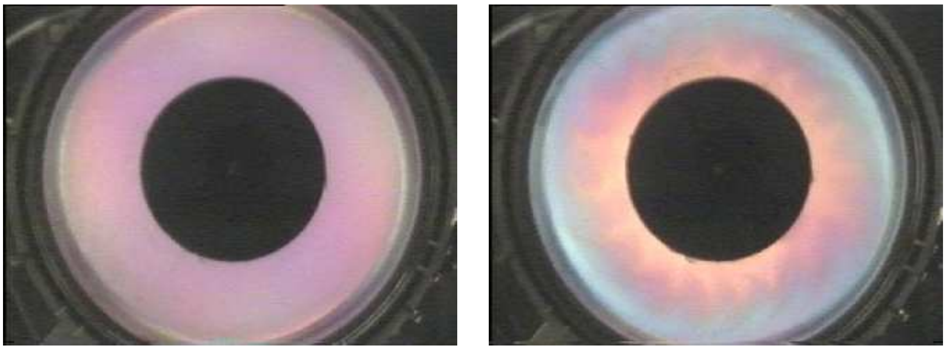
The small-scale waves in the Lighthill theory are not Kelvin-Helmholtz waves generated by a shear instability, but inertia-gravity waves generated by the motion of an object in the fluid. This gives our first indication that the observed laboratory small-scale waves in the *MRW* and *MIW* regimes may not be Kelvin-Helmholtz waves like those in the *KH* regime. This possibility is explored more fully in Sections 4.7 and 6.6.

4.5 Experimental results: RAI series

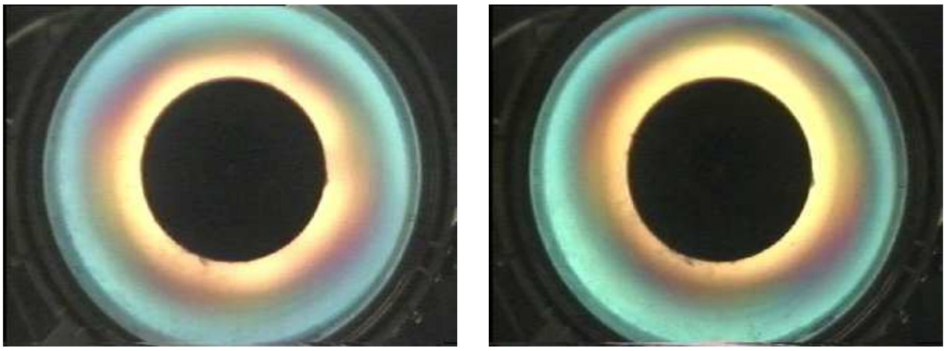
The same four regime types encountered in the PAI series were also observed in the *retrograde, annulus, increasing* RAI series, as illustrated in Figure 4.10 by representative images from experiment number RAI5. Note that the background radial colour gradients are different from those in Figure 4.4, because the equilibrium parabolic interface height $h(r)$ is now oppositely oriented in the z-direction. For prograde lid rotations in the PAI series we had $h \sim -r^2$, but for the present retrograde series we have $h \sim +r^2$ (see equation 5.22).

Figures 4.10(a), (b) and (c) show images from the *AX*, *KH* and *AX* regimes, respectively. Figures 4.10(d)–(g), each separated by 10 s, show a steady (non-vacillating) azimuthal wavenumber 1 flow (1S in Lovegrove's notation) corresponding to the *MRW* regime. The large-scale wave drifts around the annulus with a period of around 40 s. Small-scale waves are barely visible in these images, but can be seen in the original video recordings — they are easier to detect when a moving sequence of images is watched, rather than a single snapshot. They have amplitudes which are smaller than in the corresponding PAI experiments. Figure 4.10(h) shows an irregular large-scale wave corresponding to the *MIW* regime, with small-scale waves which are again barely detectable. The bottom part of the inner cylinder has been refracted into the field of view in this image, causing the dark feature at mid-radius and making the image difficult to interpret.

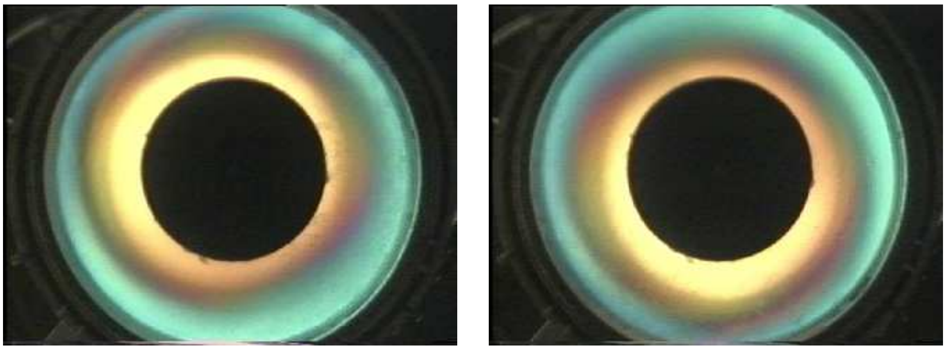
Figure 4.11 shows the regime diagram for the RAI series. It is broadly the same as the diagram for the PAI series shown in Figure 4.6(b), except that the shapes of the *AX*



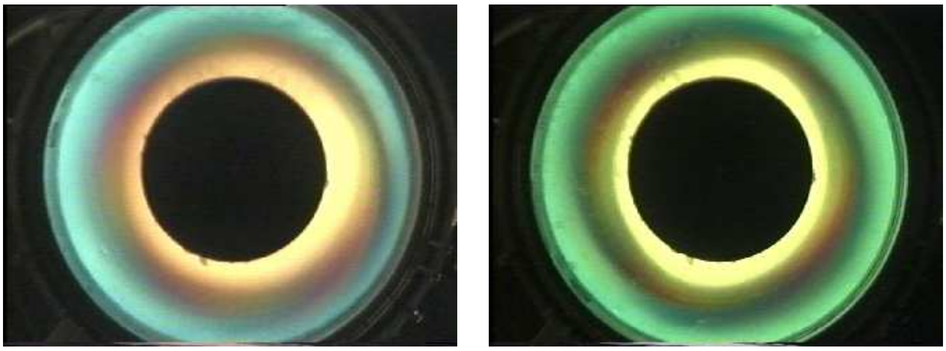
(a) 5 min (b) 40 min



(c) 70 min (d) 119 min 50 sec



(e) 120 min 00 sec (f) 120 min 10 sec



(g) 120 min 20 sec (h) 179 min

Figure 4.10: Images from experiment number RAI5, at the indicated times.

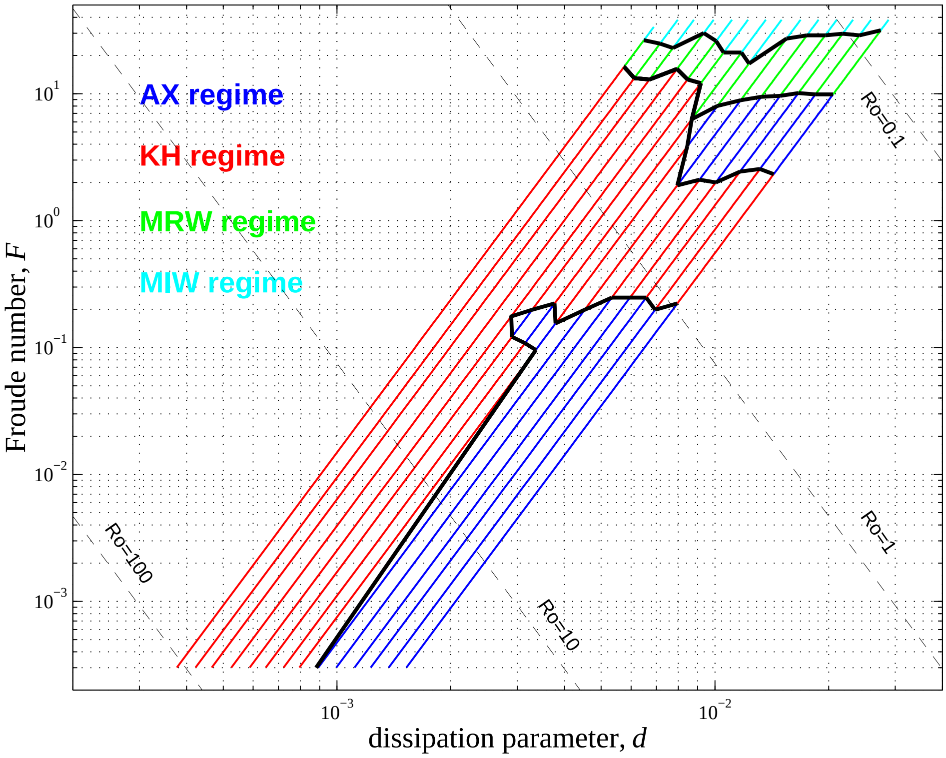


Figure 4.11: Experimental regime diagram for the RAI series. Each multi-coloured line corresponds to one 3-hour experiment, with experiment number 1 furthest to the right and experiment number 16 furthest to the left. The dissipation parameter, which is negative according to equation (4.2) since $\Delta\Omega < 0$, is plotted with its sign reversed for comparison with the other regime diagrams.

and *KH* regions have been modified slightly.

The most likely explanation for the reduced small-scale wave amplitudes in this series, compared with the PAI series, is that interfacial tension effects are larger in the present case. The quasi-geostrophic equilibrium parabolic interface height shape will shortly be shown to be given by equation (5.22). There are two contributions: the external centripetal effect $\Omega^2 r^2/(2g)$ is always positive, whereas the internal centripetal effect $-\Omega\Delta\Omega r^2/(2g')$ is positive for retrograde $\Omega\Delta\Omega < 0$ flow and negative for prograde $\Omega\Delta\Omega > 0$ flow. In the current series of experiments, then, the external and internal centripetal effects combine constructively to give an interface of larger curvature than in the PAI series, where there was partial cancellation between the two terms. This seems to have reinforced the effects of interfacial tension — which are proportional to curvature — to such an extent that the growth of small-scales waves has been suppressed in this case.

4.6 Experimental results: PAD series

Figure 4.12 shows the equivalent regime diagram for the *prograde, annulus, decreasing* PAD series. Individual frames from a typical experiment are not shown in this case, as they are almost identical to those from the PAI series except that their order is reversed in time. The small-scale wave amplitudes seemed to be generally the same as for the PAI series.

The regime diagram is almost exactly the same as the diagram for the PAI series shown in Figure 4.6(b), except that the *AX* regime at larger Froude numbers has vanished. The boundary between the *AX* and *KH* region is very well predicted by Figure 4.9 with a critical Richardson number of 1. The only dynamical difference between the large Froude number *KH* regions in the PAI and PAD series is the direction of approach in parameter space. Sensitivity to direction of approach is a manifestation of *intransitivity* and is a consequence of hysteresis in the system. Intransitivity has been observed before in experimental studies of the rotating annulus (e.g. Hide & Mason, 1975). The implication is that for a given (d, F) there are many possible equilibrated flows, and the particular one which is observed depends to an extent upon the system's memory of its

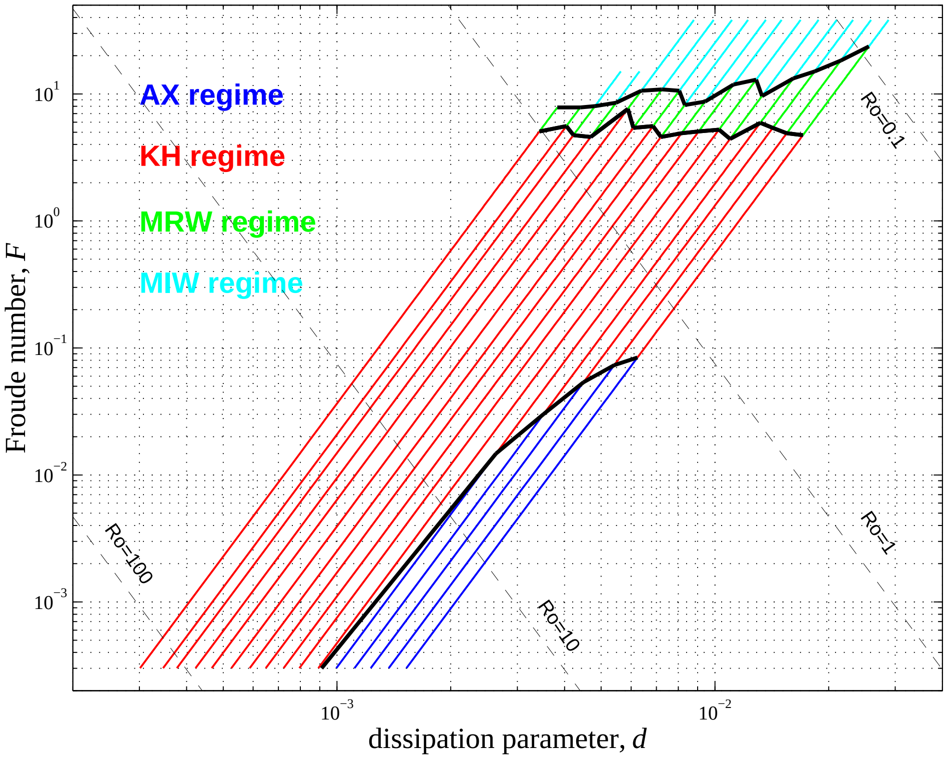


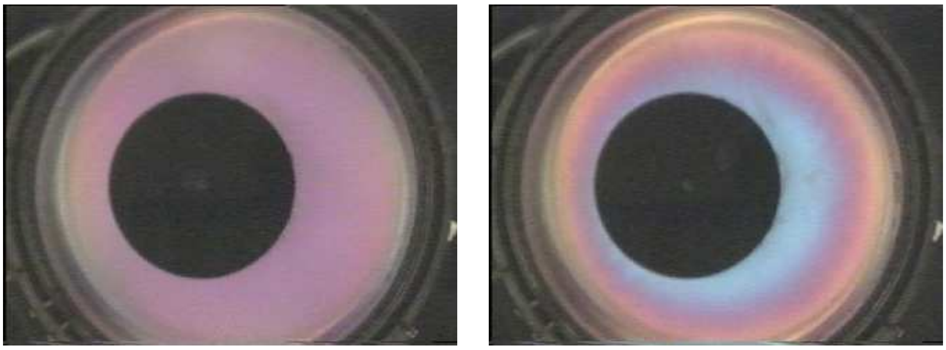
Figure 4.12: Experimental regime diagram for the PAD series. Each multi-coloured line corresponds to one 3-hour experiment, with experiment number 1 furthest to the right and experiment number 16 furthest to the left.

previous state.

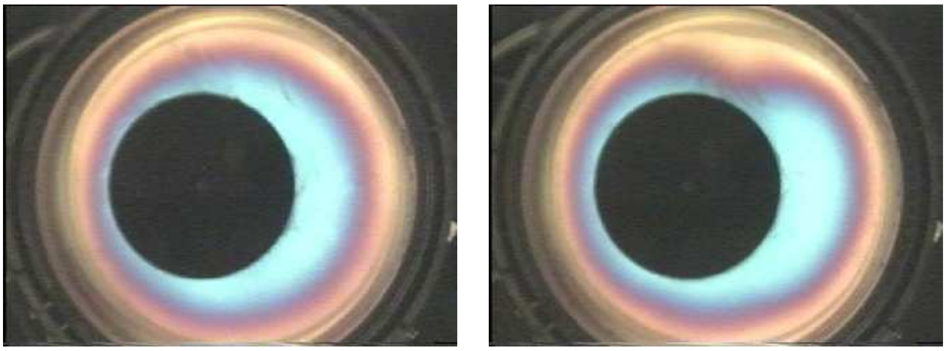
4.7 Experimental results: PEI series

For the *prograde, eccentric, increasing* PEI configuration, we displace the inner cylinder horizontally so that it is no longer aligned with the rotation axis. The concentric annulus thus becomes an eccentric annulus, with a distorted geometry in which the annular gap width and Froude number vary with azimuth. Streamlines become more tightly packed in the region of smallest gap width, enhancing the shear there and allowing an investigation into whether there are preferential azimuthal angles for short wave generation now that the azimuthal symmetry is broken.

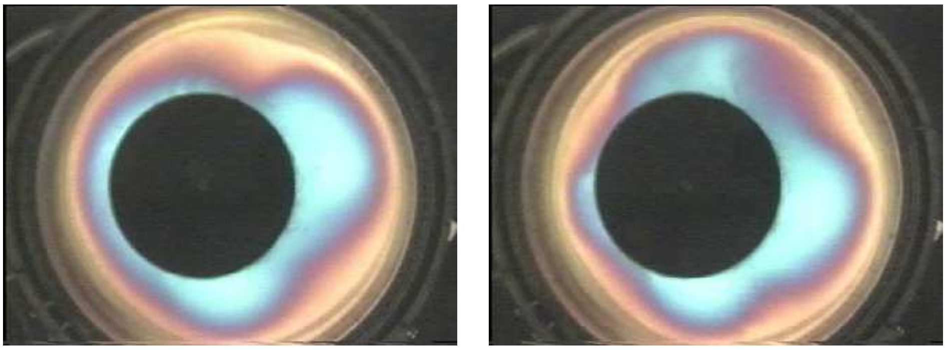
Because of time constraints, the full set of 16 PEI experiments could not be carried out, and so a regime diagram could not be drawn. Representative stills from experiment PEI3 are shown in Figure 4.13, however. The same four flow types previously described



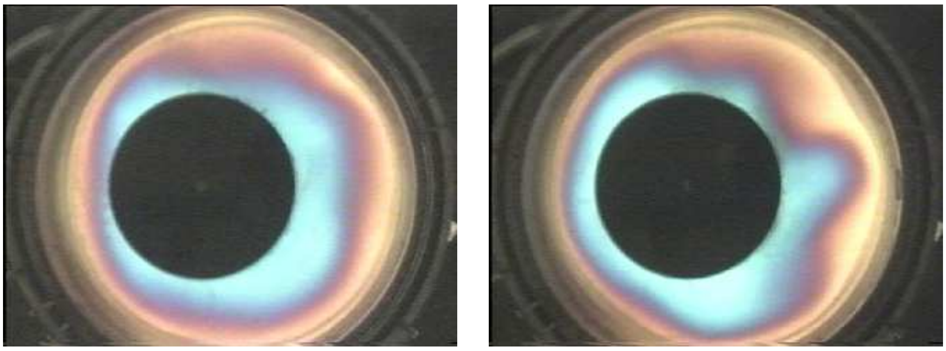
(a) 5 min (b) 40 min



(c) 70 min (d) 95 min



(e) 115 min (f) 131 min



(g) 132 min (h) 161 min

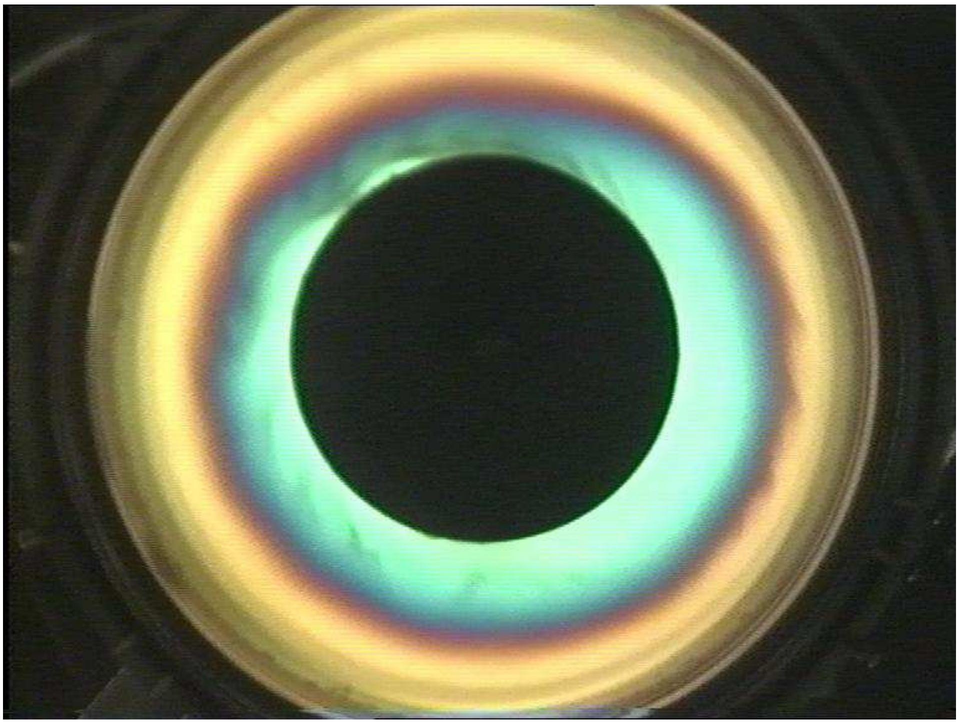
Figure 4.13: Images from experiment number PEI3, at the indicated times.

were observed. The AX flow type is shown in (a) and (c), and the KH regime in (b). The MRW regime is shown in (d)–(g), with (d) showing a 2S flow, (e) showing a 3S flow, and (f) and (g) showing times of maximum and minimum amplitude in the cycle of a 4Av flow. The MIW regime is shown in (h). Unlike in the KH regime of this preliminary experiment, in the mixed wave regimes there was no evidence of preferential short wave emission regions correlated with regions of large shear (i.e. regions near “9 o’clock”). This adds weight to the conclusion from Section 4.4.2 that the short waves in the MRW and MIW regimes are not generated by a shear instability.

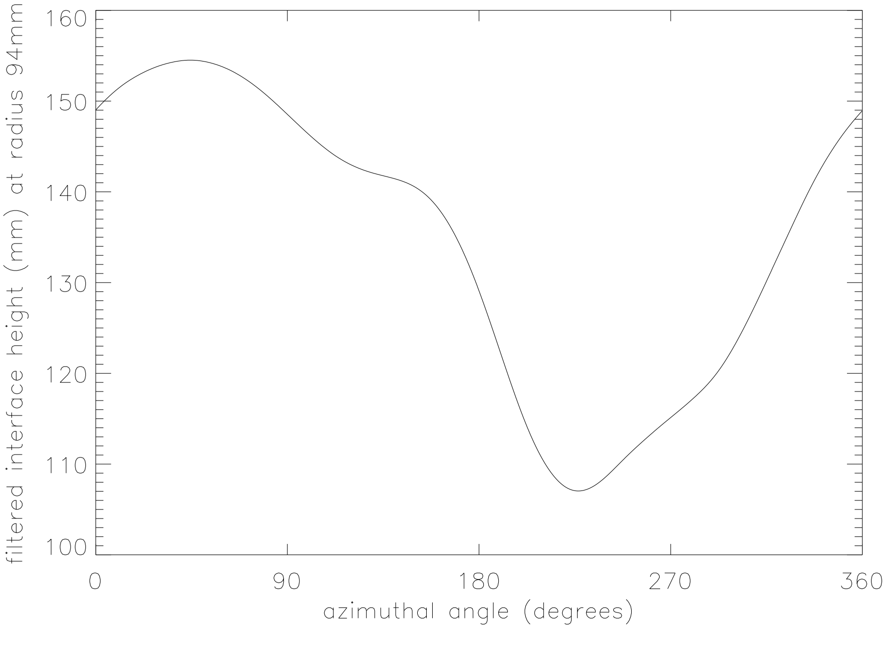
4.8 Calculation of wave amplitudes

We now use the calibration scheme of Chapter 3 to determine the amplitudes of waves in the MRW regime of the PAI experiments (Section 4.4), following the example of Section 3.7. Figures (4.14)–(4.16) show the amplitude calculation for large-scale, baroclinic waves of azimuthal wavenumbers 1, 2 and 3. The wave amplitudes in the middle of the annular gap (at $r = 94$ mm) are found to be around 25 mm, 8 mm and 7 mm, respectively. The azimuthal profiles show that the waves are not perfectly regular, and so the wave amplitudes read-off by eye and quoted here are approximate. This spatial irregularity can be seen in the raw images, and suggests the presence of azimuthal modes other than the dominant one. For example, the presence of a sub-dominant wavenumber 1 mode can be seen in the raw image of Figure 4.15(a). The amplitudes determined here will be compared with amplitudes from a numerical simulation in Section 6.1.2.

Figure 4.17 shows a similar analysis for the image shown in Figure 1.6, consisting of large-scale waves superimposed with two trains of small-scale waves. The amplitude of the small-scale waves near the inner cylindrical boundary (at $r = 70$ mm) is around 3 mm. This is around a third of the mid-radius baroclinic wave amplitude, implying that these small-scale waves are significantly larger, relative to the large-scale mode, than those reported by Read (1992a). The small-scale wave amplitude decreases quite rapidly with increasing radius, dropping below 1 mm at mid-radius (not shown).

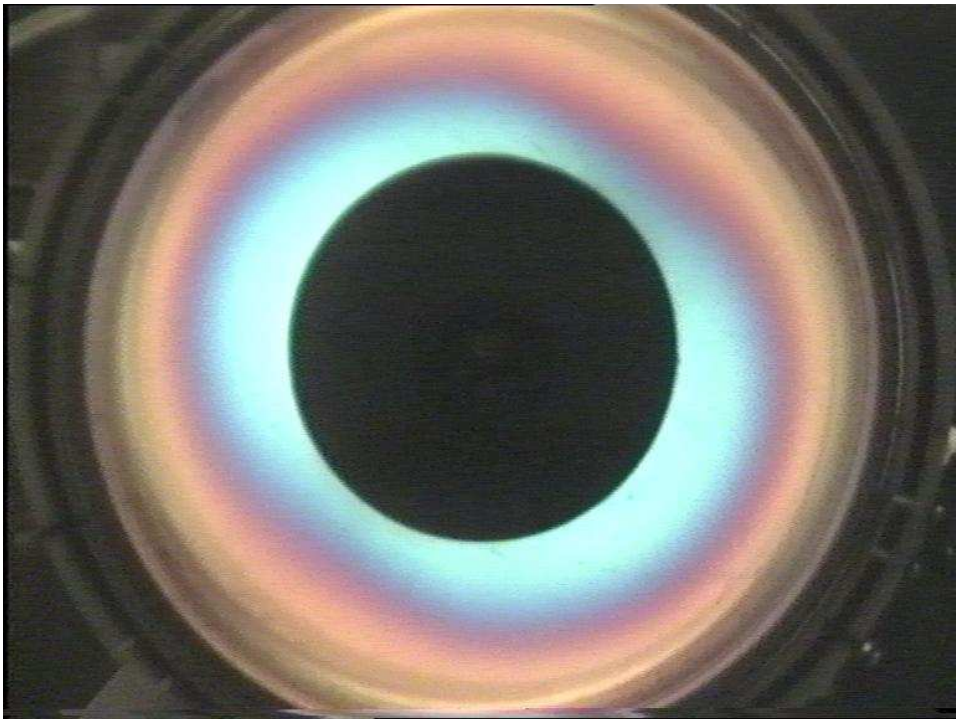


(a)

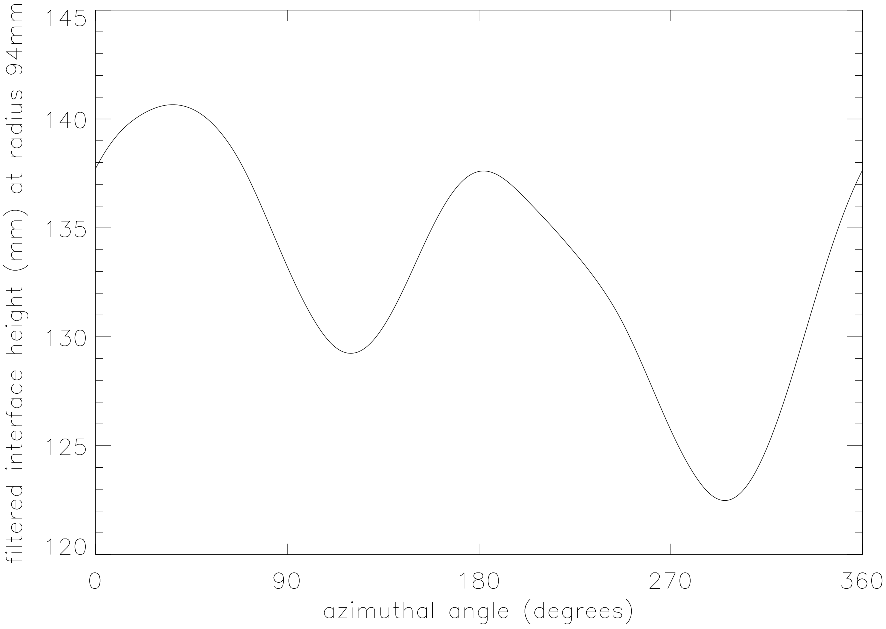


(b)

Figure 4.14: Large-scale baroclinic wave with azimuthal wavenumber 1. (a) Raw experimental image, and (b) reconstructed mid-radius azimuthal profile of interface height.

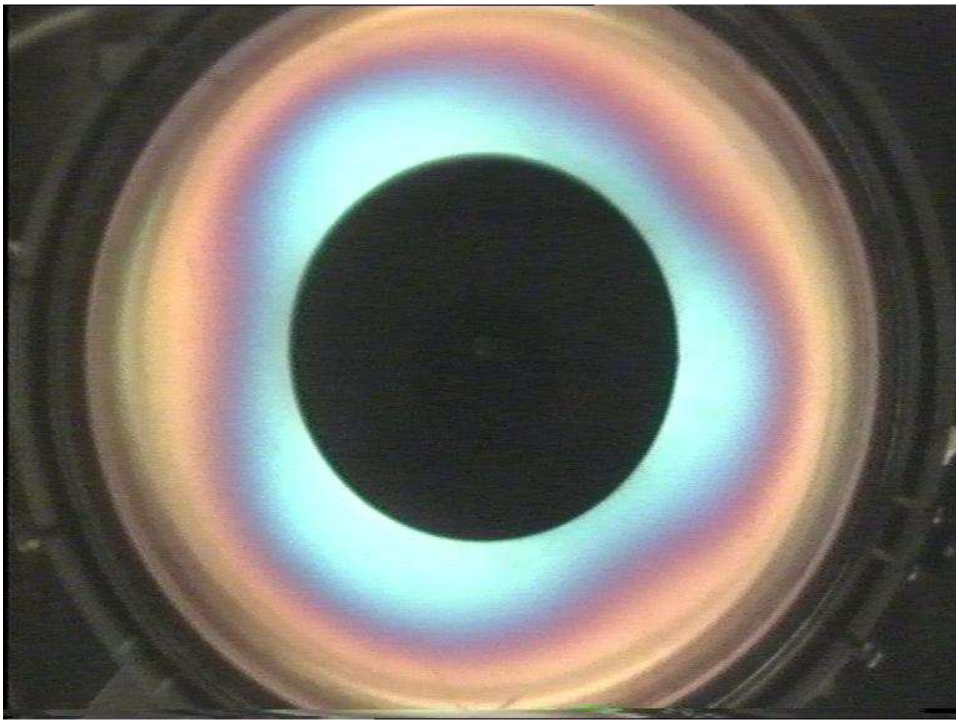


(a)

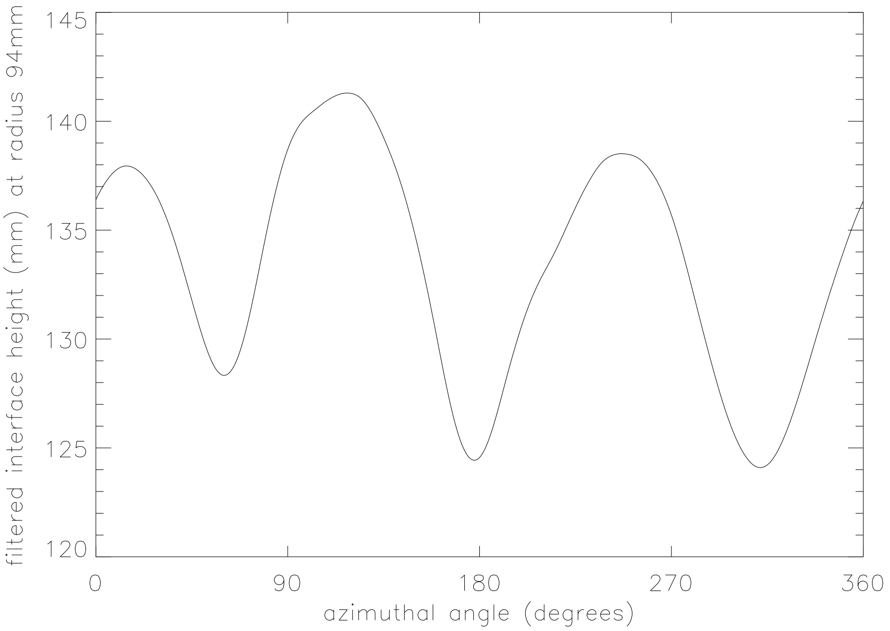


(b)

Figure 4.15: Large-scale baroclinic wave with azimuthal wavenumber 2. (a) Raw experimental image, and (b) reconstructed mid-radius azimuthal profile of interface height.



(a)



(b)

Figure 4.16: Large-scale baroclinic wave with azimuthal wavenumber 3. (a) Raw experimental image, and (b) reconstructed mid-radius azimuthal profile of interface height.

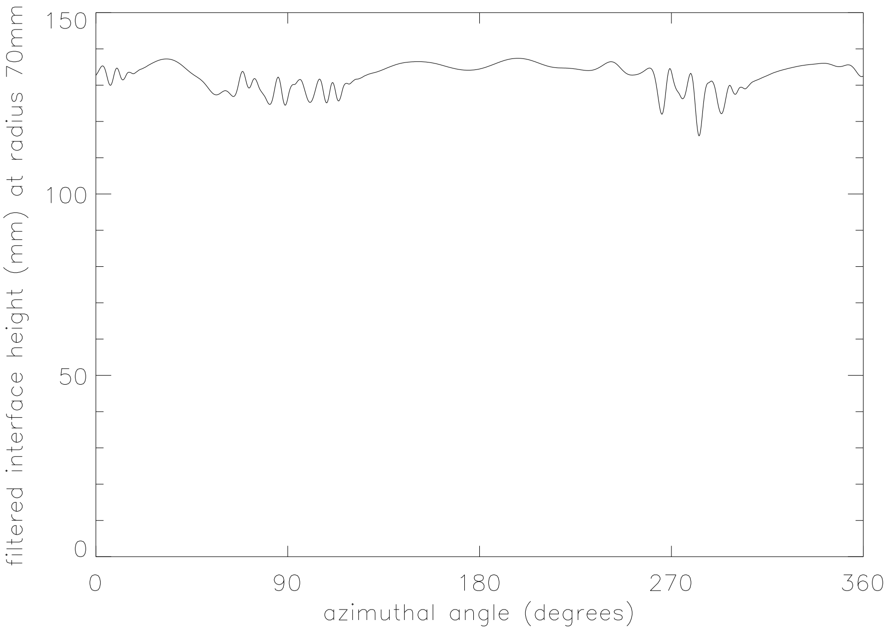


Figure 4.17: Reconstructed azimuthal profile of interface height near the inner cylindrical boundary of Figure 1.6.

4.9 Discussion

The small-scale waves in the current experiments are more ubiquitous than those in Lovegrove's experiments. In particular, whereas Lovegrove reported small-scale wave generation only during amplitude vacillating large-scale modes, we have observed them in almost all non-vacillating large-scale flows, though their amplitudes are generally smaller when the large-scale amplitude is constant. The large-scale flow types during which small-scale waves were observed to be generated in Lovegrove's experiments, are therefore a subset of those in the present experiments.

This apparent discrepancy can be explained in either (or both) of two ways. We could assume that small-scale waves were actually just as ubiquitous in Lovegrove's experiments as in the present ones, but that his flow visualization had sufficient resolution to capture only those with the largest amplitude. This could be due to his lower-grade video signal or frame-grabber (Figure 2.4), or to a non-optimized crossed polaroid angle (Section 2.3). Alternatively it may have been that the differences in fluid properties, which evolve in time as inferred in Sections 2.2.4 and 4.2, have had a significant impact upon the small-scale wave production mechanism. It is not clear which of these two explanations is responsible, though it seems more likely to be the latter.

4.10 Chapter summary

The wave modes occurring in the rotating two-layer annulus equations fall into two distinct classes, both of which we have observed in the present series of laboratory experiments. We have identified those regions of the principal 2-D parameter space in which the large-scale and small-scale modes exist and coexist, and labelled them appropriately. Kelvin-Helmholtz shear instability theory, based on a critical Richardson number, appears to explain the locations of the $AX \rightarrow KH$ and $KH \rightarrow AX$ transition curves, and baroclinic instability theory based on a critical Froude number successfully accounts for the locations of the $AX \rightarrow MRW$ and $KH \rightarrow MRW$ transition curves.

The mechanism by which the *MRW* and *MIW* small-scale waves are generated remains to be explained, though we have shown that the waves are robust to various system changes, appearing in four different experimental configurations, and are therefore not just a feature peculiar to the particular configuration used by Lovegrove. The generation mechanism responsible will be investigated in detail using a numerical model in Chapter 6.

This chapter marks the end of the first part of the thesis. Though we have been able to derive wavelengths and amplitudes in the experiment, there is no practicable way to measure velocity fields. These are needed to investigate the production of short, fast waves in baroclinically-unstable flow regimes. Motivated by this, in the following chapters we develop and run a numerical model of the laboratory experiment, which will allow us to derive high-resolution velocity data for this purpose. It will also give interface heights and azimuthal wavenumbers for comparison with those in the laboratory annulus.

Chapter 5

Description of the numerical model

> *"quagmire* (noun) 1. *an area of soft wet ground which you sink into if you try and walk on it.* 2. *a difficult and dangerous situation."*
>
> Cambridge Advanced Learner's Dictionary, 2003.

In this chapter, we describe the design and construction of a new numerical model for simulating fluid flows in rotating annulus laboratory experiments. Having discussed a variety of candidate model types, each with different dynamical assumptions, we decide to use a multi-layer quasi-geostrophic model. A model with a full representation of the annular geometry is preferable to a Cartesian channel model, for a number of important reasons which are discussed.

The two-layer continuous quasi-geostrophic equations in cylindrical coordinates are derived, and are decomposed into vertical and azimuthal normal mode form to simplify their solution. Suitable sidewall boundary conditions are derived by considering integral properties of the governing equations. Then the equations are carefully discretized in such a way as to preserve discrete analogues of the integral properties. Suitable numerical parameter values and initial conditions are given, and the model code units are tested to ensure that they are free from errors.

The model has become known as QUAGMIRE, the *QUAsi-Geostrophic annulus Model for Investigating Rotating fluids Experiments.*

5.1 Motivation for running numerical simulations

In this, the second part of the thesis, we embark upon an investigation of the rotating, two-layer annulus using a numerical model. This computational approach is complementary to the laboratory investigation undertaken in the previous chapters, and is intended to enhance and extend our understanding of the dynamical mechanisms at play. The key aims are to determine the fluid velocity fields (unavailable from the laboratory experiments) in order to investigate sources of the observed small-scale wave emission in the *MRW* regime, and to run simulations both with and without a representation of the fast waves in order to investigate their impacts on the large-scale flow.

One possible numerical approach would be to carry out a direct numerical simulation (DNS) of the Navier-Stokes equations for the system, and to examine the model's ability to simulate the production of short waves as observed in the laboratory. DNS codes have been developed for the rotating continuously-stratified thermal annulus (e.g. White, 1986; Hignett et al., 1985) but these would require significant modification in order to be applicable to the discrete-layer isothermal system. Furthermore, DNS codes are computationally expensive, and could be used to examine not more than a few case studies at the resolution required to simulate the fast, small-scale waves.

As an alternative to a DNS for the numerical simulations, it was decided to use a balanced model, in which small-scale waves are filtered out by construction (Section 1.1). Because of the filtering of unbalanced modes, balanced models have fewer dynamical degrees of freedom and therefore run much more quickly than DNS models, allowing large numbers of simulations to be performed.

A key additional benefit is that a comparison of the laboratory and numerical results allows us to assess the ability of a filtered model to simulate a system in which motions occur on a wider spectrum of scales than that permitted by the filtering. This is equivalent to an assessment of the impact of the small-scale waves upon the large-scale balanced flow. If there are found to be discrepancies between model and laboratory system behaviour, and if the only significant difference between model and laboratory is the presence of small-scale waves in the laboratory, then we can infer that those discrepancies are likely to be due to the presence of the small-scale waves.

5.2 Review of filtered models

The relative merits of three candidate filtered models and two candidate geometries are now discussed.

5.2.1 Candidate filtered models

Three commonly-used filtered models for simulating rapidly-rotating, two-layer flows are those based on the *quasi-geostrophic equations*, the *balance equations* and the *slow equations*. These three equation sets can each be derived from the *shallow water equations*, which in turn are derived from the *Navier-Stokes equations* under the assumptions of hydrostatic balance and columnar flow. Discussions of these and other filtered models are given by McWilliams & Gent (1980) and by McIntyre & Norton (2000).

The main assumptions made in the derivation of the quasi-geostrophic equations, first used by Charney et al. (1950), are that the potential vorticity is advected only by the geostrophic component of the flow, and that the amplitudes of perturbations to the fluid surfaces are much smaller than the mean fluid depths. A list of the complete set of approximations is given in Section 5.3.

The balance equations (Charney, 1955) are derived by performing a horizontal velocity decomposition into rotational and divergent components, and then truncating with respect to the divergent component. The balance that they describe is more complicated, but also more accurate, than geostrophic balance, and efficient procedures have been developed to integrate them (Daley, 1982). However, it has been pointed out by Moura (1976) that, in their most general form, the balance equations have spurious non-physical wave solutions with phase speeds much larger than those of inertia-gravity waves.

The slow equations (Lynch, 1989) are derived in a similar way to the balance equations, except that the velocity truncation is performed in a more systematic manner (based on normal mode initialization, discussed in Section 1.4), which results in the vanishing of the spurious solutions. Numerical integrations of the slow equations show excellent agreement with initialized numerical integrations of the shallow water equations.

Of these three candidate models, the quasi-geostrophic (Q-G) model was selected to simulate flows in the annulus. This is because only one scalar function of horizontal position is needed per layer to uniquely define the state of the system using a Q-G model (streamfunction), whereas three are needed per layer using a balance or slow equations model (streamfunction, velocity potential and geopotential). With three times fewer independent variables, the computational advantages gained from using a Q-G model were felt to outweigh the disadvantages of its slightly lower formal accuracy.

5.2.2 Candidate geometries

A number of numerical Q-G models have been developed for systems consisting of superposed immiscible fluid layers in a rectangular channel (e.g. Brugge et al., 1987). Before constructing a new numerical model, we first considered whether any of these Cartesian models could meet our requirements. For the following reasons, it was decided that they could not.

Firstly, the channel equations with periodic boundary conditions are a good approximation to the annulus equations only if the ratio of the width of the annular gap to its mean radius is much smaller than unity (King, 1979b). With this geometry, the curvature becomes negligible, and we would be justified in using a channel model to simulate the flow in the annulus. For the present laboratory apparatus, though, the ratio is 6.25 cm/9.375 cm $\sim$ 0.7, which is only slightly smaller than 1.

Secondly, channel models have additional, shift-reflect symmetries (Cattaneo & Hart, 1990) not present in annulus models. This is the case because, though the annulus and periodic channel are topologically similar, the geometry of their boundaries is fundamentally different. For example, there is a reflect symmetry in the channel in the plane which is equidistant from the sidewall boundaries, but there is no analogous symmetry in the annulus. Kwon & Mak (1988) show that the existence of such additional symmetries in the periodic channel leads to certain large-scale wave-wave interaction coefficients being identically zero. Importantly, an annular model would allow the complete set of large-scale wave-wave interactions that take place in the laboratory experiments, to be included in the model, which is important for quantitative agreement.

Furthermore, a model in cylindrical coordinates would be more general, and potentially applicable to laboratory experiments other than the present one. For example, it would keep open the possibility of running simulations in an open cylinder with no inner sidewall, as well as in an annulus, though it would then be necessary to include an inner sidewall of small nominal radius in the model, to avoid the singularity at $r = 0$. We would need to assume that the flow is insensitive to the inclusion of this additional boundary.

There are background potential vorticity (PV) gradients present in both the channel and the annulus, due to the sloping of equilibrium geopotential height surfaces in the presence of a vertical shear in horizontal velocity. In the channel, these geopotential heights and PV gradients are linear in the across-channel direction (giving an effective β-effect), whereas in the annulus they are quadratic because of the parabolic equilibrium interface height shape (Section 2.4). This gives a *quadratic β-effect*, with the possibility of qualitatively different dynamics than in the presence of the usual linear β-effect. Furthermore, the quadratic β-effect can be quite large (Section 5.3).

As a final point, not connected with geometry, few of the existing Q-G layer channel models include the effects of interfacial tension, which are not necessarily always negligible in the laboratory (Section 2.2.3).

Since we desire quantitative agreement with the laboratory experiments, we conclude for these reasons that we need to construct a new multi-layer Q-G model which takes into account the cylindrical geometry and interfacial tension. Such a model is described in the remaining sections of this chapter.

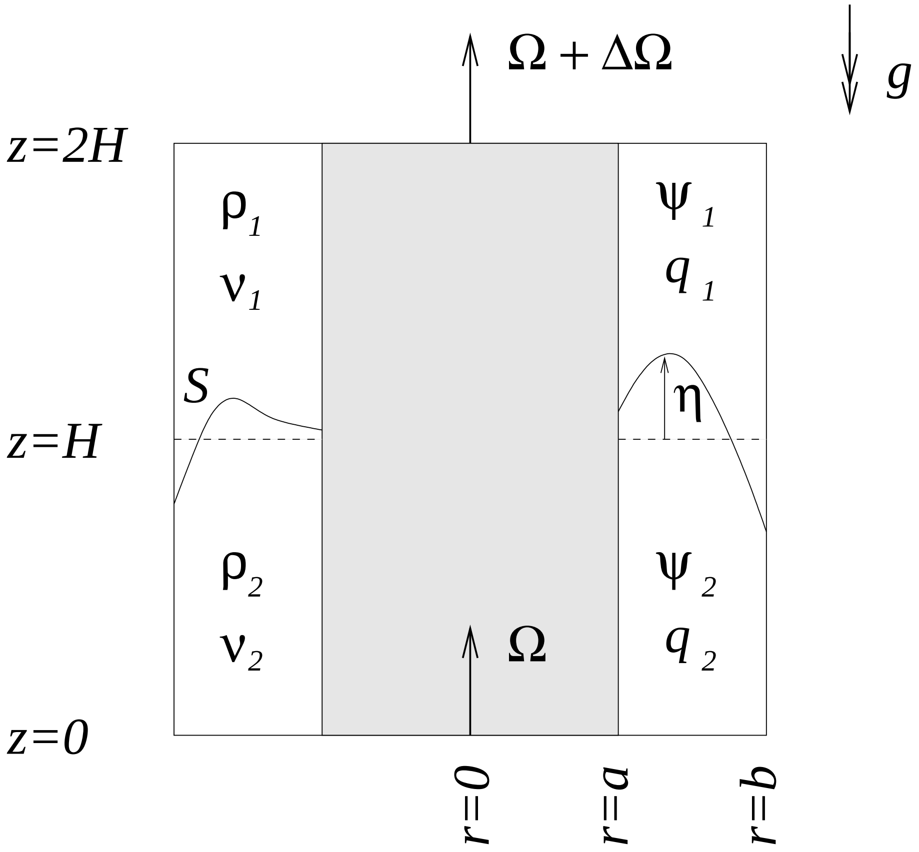


Figure 5.1: Schematic diagram showing a vertical cross-section through the two-layer annulus system being modelled. The dashed line shows the resting interface height. See text for definitions.

5.3 Derivation of model equations

The system to be modelled is shown schematically in Figure 5.1. Such a model is often informally referred to as two-and-a-half dimensional, as the representation of the vertical is achieved through only two discrete layers. Cylindrical polar coordinates $\boldsymbol{r} = (r, \theta, z)$ are used, the z-axis being coincident with the vertical rotation axis. The fluid is bounded by a flat base at $z = 0$, a flat lid at $z = 2H > 0$ and cylindrical walls at $r = a$ and $r = b > a$. The two immiscible layers have densities ρ_i, kinematic viscosities ν_i and mutual interfacial tension S. The undisturbed layer depth is H and the disturbed lower layer depth is $H + \eta$. The acceleration due to gravity is g. The annulus base and walls rotate about the axis of symmetry with angular velocity Ω, and the lid with angular velocity $\Omega + \Delta\Omega$.

Working in the frame of the base, the four fundamental equations for the pressure $p_i(\boldsymbol{r}, t)$

and the velocity $\boldsymbol{u}_i(\boldsymbol{r},t)$ in layer $i=1,2$ are the Navier-Stokes equations:

$$\frac{\partial \boldsymbol{u}_i}{\partial t} + (\boldsymbol{u}_i \cdot \boldsymbol{\nabla})\boldsymbol{u}_i + 2\boldsymbol{\Omega} \times \boldsymbol{u}_i + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}) = -\frac{1}{\rho_i}\boldsymbol{\nabla} p_i + \nu_i \nabla^2 \boldsymbol{u}_i + \boldsymbol{g} \tag{5.1}$$

and the equation of volume conservation for the incompressible liquid:

$$\boldsymbol{\nabla} \cdot \boldsymbol{u}_i = 0 \; . \tag{5.2}$$

We take the curl of equation (5.1) and use vector identities to obtain an equation for the layer vorticities $\boldsymbol{\omega}_i = \boldsymbol{\nabla} \times \boldsymbol{u}_i$:

$$\frac{\partial \boldsymbol{\omega}_i}{\partial t} + (\boldsymbol{u}_i \cdot \boldsymbol{\nabla})\boldsymbol{\omega}_i = [(2\boldsymbol{\Omega} + \boldsymbol{\omega}_i) \cdot \boldsymbol{\nabla}]\boldsymbol{u}_i + \nu_i \nabla^2 \boldsymbol{\omega}_i \; , \tag{5.3}$$

the z-component of which, in the layer interiors where the flow is assumed to be vertically-columnar and inviscid, is

$$\frac{\partial \xi_i}{\partial t} + (\boldsymbol{u}_i \cdot \boldsymbol{\nabla})\xi_i = (f + \xi_i)\frac{\partial u_{i,\,z}}{\partial z} \; , \tag{5.4}$$

where ξ_i is the z-component of $\boldsymbol{\omega}_i$, $f = 2\Omega$ is the Coriolis parameter and $u_{i,\,z}$ is the vertical velocity.

We next vertically integrate equation (5.4) over the fluid interiors, parameterizing vertical Ekman pumping/suction velocities at the lid, base and interface (Gill, 1982). Assuming that the Ekman layer depths are much smaller than the total layer depths, and making the quasi-geostrophic assumptions $\eta \ll H$ and $\xi_i \ll f$, we obtain, after rearrangement:

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_1 \cdot \boldsymbol{\nabla}\right) q_1 = -\frac{\sqrt{\Omega\nu_1}}{H}\left[\xi_1 + \chi_2(\xi_1 - \xi_2)\right] + 2\Delta\Omega\frac{\sqrt{\Omega\nu_1}}{H} \; , \tag{5.5}$$

$$\left(\frac{\partial}{\partial t} + \boldsymbol{u}_2 \cdot \boldsymbol{\nabla}\right) q_2 = -\frac{\sqrt{\Omega\nu_2}}{H}\left[\xi_2 + \chi_1(\xi_2 - \xi_1)\right] \; , \tag{5.6}$$

where $\chi_i = \sqrt{\nu_i}/(\sqrt{\nu_1} + \sqrt{\nu_2})$, and $q_i(r,\theta,t)/H$ are the perturbation potential vorticities (PPVs), given by

$$q_1(r,\theta,t) = \xi_1 + \frac{f\eta}{H} \tag{5.7}$$

and

$$q_2(r,\theta,t) = \xi_2 - \frac{f\eta}{H} \; . \tag{5.8}$$

To complete the derivation, we write all of the independent variables ($\boldsymbol{u}_i$, ξ_i and η) in equations (5.5)–(5.8) in terms of the layer streamfunctions $\psi_i(r, \theta, t)$ defined by

$$u_{i,\,\theta} = \frac{\partial \psi_i}{\partial r} \tag{5.9}$$

and

$$u_{i,\,r} = -\frac{1}{r}\frac{\partial \psi_i}{\partial \theta} \ . \tag{5.10}$$

The streamfunctions ψ_1 and ψ_2 are defined only to within arbitrary additive constants, which will be discussed in Section 5.4.2. The vorticities are given by

$$\xi_i = \nabla^2 \psi_i \ . \tag{5.11}$$

Assuming hydrostatic balance and nearly equal layer densities, the interface height perturbation is given in terms of the streamfunctions (to within an additive constant) by

$$\eta - \delta_m^2 \nabla^2 \eta = \frac{f}{g'}(\psi_2 - \psi_1) + \frac{r^2\Omega^2}{2g} \ , \tag{5.12}$$

where $g' = 2g(\rho_2 - \rho_1)/(\rho_2 + \rho_1)$ is the reduced gravity. The term in $\delta_m = \sqrt{S/[g(\rho_2 - \rho_1)]}$ represents the effects of interfacial tension for an interface of small curvature. δ_m is the characteristic static meniscus width, as can be seen by considering solutions to equation (5.12) when the tank is at rest ($\Omega = 0$) and the fluid velocities are zero ($\psi_i =$ constant). The equation is a forced Helmholtz equation for η given ψ_i, where the boundary conditions are the slopes $\partial\eta/\partial r$ at the annulus walls, which are related to the interface contact angle. We require an explicit formula for η, and so we seek a first order solution to the Helmholtz equation for weak interfacial tension, by estimating the $\nabla^2\eta$ term using the solution for η when $\delta_m = 0$. This gives

$$\eta = \frac{f}{g'}(1 + \delta_m^2 \nabla^2)(\psi_2 - \psi_1) + \frac{r^2\Omega^2}{2g} \ , \tag{5.13}$$

where 1 and $\delta_m^2 \nabla^2$ are the first two terms in a power series solution. On simple grounds, the series would be expected to converge rapidly if $\delta_m^2 \nabla^2 \eta \ll \eta$, which is the case if $\delta_m^2 \ll \lambda^2$ for waves of wavelength λ. We expect waves to form on the scale of the internal Rossby radius $\sqrt{g'H}/|f|$, so the convergence criterion becomes $\delta_m^2 f^2/g'H \ll 1$. This is equivalent to $FI \ll 1$ where F is the Froude number, given by equation (4.1), and the non-dimensional parameter $I = \delta_m^2/(b-a)^2$ is the interfacial tension number (Appleby, 1982).

We finally substitute equations (5.9), (5.10), (5.11) and (5.13) into (5.5) and (5.6) to obtain the two coupled partial differential equations governing the evolution of quasi-geostrophic motions in the two-layer annulus:

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_1 q_1 = -\frac{\sqrt{\Omega\nu_1}}{H}\left[\nabla^2\psi_1 + \chi_2\nabla^2(\psi_1-\psi_2)\right] + \frac{2\Delta\Omega\sqrt{\Omega\nu_1}}{H} \tag{5.14}$$

and

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_2 q_2 = -\frac{\sqrt{\Omega\nu_2}}{H}\left[\nabla^2\psi_2 + \chi_1\nabla^2(\psi_2-\psi_1)\right] \ . \tag{5.15}$$

The total derivative operators are given by

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_i = \frac{\partial}{\partial t} - \frac{1}{r}\frac{\partial\psi_i}{\partial\theta}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial\psi_i}{\partial r}\frac{\partial}{\partial\theta} \tag{5.16}$$

and the horizontal Laplacian operator is given by

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2} \ . \tag{5.17}$$

From equations (5.7) and (5.8), the quantities q_1 and q_2 are given in terms of ψ_1 and ψ_2 by

$$q_1 = \nabla^2\psi_1 + \frac{f^2}{g'H}(1+\delta_{\mathrm{m}}^2\nabla^2)(\psi_2-\psi_1) + \frac{f}{H}\frac{r^2\Omega^2}{2g} \tag{5.18}$$

and

$$q_2 = \nabla^2\psi_2 - \frac{f^2}{g'H}(1+\delta_{\mathrm{m}}^2\nabla^2)(\psi_2-\psi_1) - \frac{f}{H}\frac{r^2\Omega^2}{2g} \ . \tag{5.19}$$

On the right side of equation (5.14), the first term represents spin-down by the frictional Ekman layers at the lid ($\nabla^2\psi_1$) and interface ($\nabla^2(\psi_1-\psi_2)$). The second term is the (constant) forcing term, and represents generation of PV by the rotating lid, communicated to the fluid interior by the Ekman layer. The terms on the right side of (5.15) have a similar interpretation, except that there is no forcing term in this case.

Equations (5.18) and (5.19) are similar to the PV-streamfunction relationships in the channel model of Brugge et al. (1987), except that the present equations include an interfacial tension modification, and Brugge's βy term has been replaced with our $\beta^* r^2$ term. This is the quadratic β-effect discussed in Section 5.2.2. It is equal and opposite in the upper and lower layers, corresponding to the fact that depth increases in one layer are accompanied by equal decreases in the other layer. The radial interface height change across the annulus, associated with these quadratic β-effect terms, can be up to

20 mm (Section 2.4). Since mid-radius large-scale wave amplitudes reach only 25 mm (Section 4.8), interface perturbations due to the quadratic β-effect are not small compared to those due to large-scale waves, and can therefore not be neglected.

Upon non-dimensionalization of equations (5.14), (5.15), (5.18) and (5.19), using a time scale $(\Delta\Omega)^{-1}$ and horizontal length scale $(b-a)$, the definitions of Froude number and dissipation parameter given in Section 4.1 appear naturally. We choose to code the model using dimensional units, however, and therefore do not carry out the non-dimensionalization here.

We now summarize the assumptions which were required to derive equations (5.14)–(5.19). It is important to bear these approximations in mind, since they limit the applicability of the model:

- incompressible fluids
- vertically-columnar fluid interiors
- inviscid fluid interiors (Reynolds number $\mathrm{Re} \gg 1$)
- linear Ekman pumping/suction
- Ekman layer depths $\delta_{\mathrm{E}} \ll H \pm \eta$
- $\eta \ll H$
- $\xi_i \ll f$ (Rossby number $\mathrm{Ro} \ll 1$)
- hydrostatic balance $\mathrm{D}w/\mathrm{D}t \ll g$
- $g' \ll g$
- $|\boldsymbol{\nabla}\eta \ll 1|$
- $FI \ll 1$
- passive Stewartson layers which do not exchange fluid with the interiors
- Stewartson layer widths $\delta_{\mathrm{S}} \ll b-a$

The final two assumptions are discussed in Section 5.4, but are included here for completeness.

5.3.1 Perturbation equations

There is an equilibrium solution to equations (5.14)–(5.19) of the form $u_{i,r} = 0$, $u_{i,\theta} = r\Delta\Omega_i$. Substituting allows us to determine the interior solid-body rotation rates:

$$\frac{\Delta\Omega_1}{\Delta\Omega} = \frac{2+\chi}{2(1+\chi)} \tag{5.20}$$

and

$$\frac{\Delta\Omega_2}{\Delta\Omega} = \frac{1}{2(1+\chi)} \ , \tag{5.21}$$

where $\chi = \sqrt{\nu_2/\nu_1}$. For $\chi = 1$ this is the same result as the solution obtained from the torque balance analysis of Chapter 3 in the absence of Stewartson layers and with a horizontal interface (see Section 3.4.2 and Table 3.2). The corresponding interface height (to within an additive constant) is given by equation (5.13) to be

$$\overline{\eta} = \frac{\Omega^2 r^2}{2g}\left(1 - \frac{\Delta\Omega/\Omega}{g'/g}\right) \ . \tag{5.22}$$

Equations (5.20)–(5.22) describe the basic, equilibrium state upon which baroclinically-unstable perturbations may grow. We refer to this as the *mean flow* and label the corresponding streamfunctions and PPVs as $\overline{\psi_i}(r)$ and $\overline{q_i}(r)$, respectively.

Governing equations for perturbations to the streamfunction $\psi_i'(r,\theta,t)$ and PPV $q_i'(r,\theta,t)$ are obtained by substituting $\psi_i = \overline{\psi_i}(r) + \psi_i'(r,\theta,t)$ and $q_i = \overline{q_i}(r) + q_i'(r,\theta,t)$ into equations (5.14)–(5.19) to obtain

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_{1'} q_1' = -\frac{\sqrt{\Omega\nu_1}}{H}\left[\nabla^2\psi_1' + \chi_2\nabla^2(\psi_1' - \psi_2')\right] - \Delta\Omega_1\frac{\partial q_1'}{\partial\theta} + \frac{f^2}{2H}\left(\frac{\Omega}{g} - \frac{\Delta\Omega}{g'}\right)\frac{\partial\psi_1'}{\partial\theta} \tag{5.23}$$

and

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_{2'} q_2' = -\frac{\sqrt{\Omega\nu_2}}{H}\left[\nabla^2\psi_2' + \chi_1\nabla^2(\psi_2' - \psi_1')\right] - \Delta\Omega_2\frac{\partial q_2'}{\partial\theta} - \frac{f^2}{2H}\left(\frac{\Omega}{g} - \frac{\Delta\Omega}{g'}\right)\frac{\partial\psi_2'}{\partial\theta} \ , \tag{5.24}$$

where

$$q_1' = \nabla^2\psi_1' + \frac{f^2}{g'H}(1 + \delta_{\mathrm{m}}^2\nabla^2)(\psi_2' - \psi_1') \tag{5.25}$$

and

$$q_2' = \nabla^2\psi_2' - \frac{f^2}{g'H}(1 + \delta_{\mathrm{m}}^2\nabla^2)(\psi_2' - \psi_1') \ . \tag{5.26}$$

The total derivatives now advect according to the perturbation streamfunctions, i.e.

$$\left(\frac{\mathrm{D}}{\mathrm{D}t}\right)_{i'} = \frac{\partial}{\partial t} - \frac{1}{r}\frac{\partial\psi_i'}{\partial\theta}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial\psi_i'}{\partial r}\frac{\partial}{\partial\theta} \ . \tag{5.27}$$

Equations (5.23)–(5.26) are the fully nonlinear model equations which we solve. The constant forcing term in equation (5.14), which represents forcing of the full flow by the lid rotation, has been replaced in equations (5.23) and (5.24) with more complicated terms which represent forcing of the perturbation flow by the equilibrium state. An analytical assessment of the stability of small perturbations could begin by linearizing equations (5.23)–(5.26), but for the model we retain all of the nonlinear terms.

The perturbation velocity fields are given in terms of the perturbation streamfunctions by

$$u'_{i,\,\theta} = \frac{\partial \psi'_i}{\partial r} \tag{5.28}$$

and

$$u'_{i,\,r} = -\frac{1}{r}\frac{\partial \psi'_i}{\partial \theta} \ , \tag{5.29}$$

which are the perturbation forms of equations (5.9) and (5.10). The perturbation interface height field is given (to within an additive constant) by

$$\eta' = \frac{f}{g'}(1 + \delta_m^2 \nabla^2)(\psi'_2 - \psi'_1) \ , \tag{5.30}$$

which is the perturbation form of equation (5.13).

5.3.2 Normal mode decomposition of diagnostic equations

Given the fields ψ'_i and q'_i at any time, we can evaluate $\partial q'_i/\partial t$ at that time using the prognostic equations (5.23) and (5.24), and thereby determine q'_i at a short time in the future. We may then use this to invert the diagnostic Helmholtz equations (5.25) and (5.26) to obtain ψ'_i at that time, and then begin the loop again using the updated fields.

The Helmholtz equations are coupled, and the inversion is made easier by first writing them in vertical normal mode form to remove the coupling. We take the sum and difference of the equations to obtain, respectively,

$$\nabla^2(\psi'_1 + \psi'_2) = q'_1 + q'_2 \tag{5.31}$$

and

$$\nabla^2(\psi'_2 - \psi'_1) - C_{\text{itcc}}\frac{2f^2}{g'H}(\psi'_2 - \psi'_1) = C_{\text{itcc}}(q'_2 - q'_1) \ , \tag{5.32}$$

where C_{itcc} is an interfacial tension correction coefficient given by

$$C_{\text{itcc}} = \frac{1}{1-(2f^2\delta_m^2)/(g'H)} \; . \tag{5.33}$$

We know that $f^2\delta_m^2/g'H \ll 1$ (Section 5.3), and so C_{itcc} is slightly larger than unity, and is exactly equal to unity if the interfacial tension is zero.

Defining the barotropic (bt) and baroclinic (bc) vertical normal mode variables to be

$$\Psi'_{\text{bt}} = \psi'_1 + \psi'_2 \; , \tag{5.34}$$

$$\Psi'_{\text{bc}} = \psi'_2 - \psi'_1 \; , \tag{5.35}$$

$$Q'_{\text{bt}} = q'_1 + q'_2 \; , \tag{5.36}$$

$$Q'_{\text{bc}} = C_{\text{itcc}}(q'_2 - q'_1) \; , \tag{5.37}$$

equations (5.31) and (5.32) both become uncoupled Helmholtz equations of the form

$$\nabla^2\Psi'_m - \lambda_m\Psi'_m = Q'_m \tag{5.38}$$

for $m = \text{bt}, \text{bc}$. The eigenvalues are $\lambda_{\text{bt}} = 0$ and $\lambda_{\text{bc}} = 2C_{\text{itcc}}f^2/g'H$.

We now perform a second normal mode decomposition, this time into azimuthal modes, to further simplify the solution of the Helmholtz equations. At each timestep, we expand

$$\Psi'_m(r,\theta) = \sum_{n=-\infty}^{\infty} \hat{\Psi}'^n_m(r)e^{\sqrt{-1}n\theta} \; , \tag{5.39}$$

$$Q'_m(r,\theta) = \sum_{n=-\infty}^{\infty} \hat{Q}'^n_m(r)e^{\sqrt{-1}n\theta} \; . \tag{5.40}$$

The complex functions $\hat{\Psi}'^n_m$ and $\hat{Q}'^n_m$ satisfy $\hat{\Psi}'^n_m = \hat{\Psi}'^{-n^*}_m$ and $\hat{Q}'^n_m = \hat{Q}'^{-n^*}_m$, where the asterisk represents complex conjugation, because $\Psi'_m(r,\theta)$ and $Q'_m(r,\theta)$ are real. The $n = 0$ term is called the *mean flow correction* (a correction to the zonal flow that is generated by nonlinear self interactions of the waves), and is equal to the zonal average of the perturbation quantities as can be seen from the zonal integration of equations (5.39) and (5.40). The $n \neq 0$ terms represent eddy (wave) components. Substituting equations (5.39) and (5.40) into (5.38) gives the radial structure equation:

$$\frac{\mathrm{d}^2\hat{\Psi}'^n_m}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}\hat{\Psi}'^n_m}{\mathrm{d}r} - \left(\lambda_m + \frac{n^2}{r^2}\right)\hat{\Psi}'^n_m = \hat{Q}'^n_m(r) \; . \tag{5.41}$$

This complex ordinary differential equation must be solved for each combination of vertical modes $m \in \{\text{bt}, \text{bc}\}$ and azimuthal modes $n \in \{0, \pm1, \pm2, \ldots\}$ to determine

$\hat{\Psi}'^{n}_{m}(r)$ given $\hat{Q}'^{n}_{m}(r)$. The inversion process required to obtain $\psi'_i(r,\theta)$ from $q'_i(r,\theta)$, which are linked by equations (5.25) and (5.26), is summarized as:

$$\boxed{\quad q'_i \xrightarrow{(5.36)\,\&\,(5.37)} Q'_{\mathrm{m}} \xrightarrow{(5.40)} \hat{Q}'^{n}_{m} \xrightarrow{(5.41)} \hat{\Psi}'^{n}_{m} \xrightarrow{(5.39)} \Psi'_{\mathrm{m}} \xrightarrow{(5.34)\,\&\,(5.35)} \psi'_i \quad}$$

We could now perform a third normal mode decomposition, this time in the radial coordinate, by projecting $\hat{\Psi}'^{n}_{m}(r)$ and $\hat{Q}'^{n}_{m}(r)$ onto the eigenfunctions of the linear operator on the left side of equation (5.41). The baroclinic eigenfunctions are *modified Bessel functions* of order n in the scaled radial coordinate $\tilde{r} = \sqrt{\lambda_{\mathrm{bc}}}r$ (Boas, 1983), and the barotropic eigenfunctions are of the form $r^{\pm n}$. However, this approach would force the streamfunction and PPV to satisfy the same boundary conditions, for which there is no justification. In the present model, we therefore solve the discretized radial structure equation directly rather than projecting onto radial modes.

5.4 Perturbation streamfunction boundary conditions for the continuous equations

We must now choose boundary conditions to apply to the perturbation streamfunction when integrating equation (5.41). The equation was derived under the assumption of inviscid flow. It therefore cannot describe the viscous Stewartson layers of width δ_S, and so applies only to the fluid interior $a+\delta_S < r < b-\delta_S$. We assume $\delta_S \ll a, b$ so that we may still write the integration range as $a < r < b$, but when we refer to $r = a$ or $r = b$ we now mean the boundary between the fluid interior and Stewartson layer, rather than the physical lateral boundary itself.[1]

There are a number of candidate boundary conditions. To impose passive Stewartson layers which do not anywhere exchange fluid with the interior, we would apply the impermeability condition on the radial perturbation velocity $u'_{i,\,r}|_{r=a,\,b} = 0 \quad \forall\ \theta, i$, which

---

[1] An alternative method for keeping the Stewartson layers out of the analysis would be to imagine that our laboratory apparatus is equivalent to a *gedanken* experiment in which, at all times in each layer, the lateral boundaries rotate at the same rate as the fluid interiors, so that the Stewartson layers vanish.

in the normal mode variables corresponds to Dirichlet boundary conditions

$$\hat{\Psi}'^{n}_{m}|_{r=a,\,b} = 0 \quad \forall\, n \neq 0, m\ . \tag{5.42}$$

The mean flow correction $n = 0$ velocity is purely zonal, and so this component automatically satisfies impermeability. Impermeability alone is therefore not a sufficient condition to uniquely specify a solution. No-slip boundary conditions for the zonal perturbation velocity $u'_{i,\theta}|_{r=a,\,b} = 0 \quad \forall\, \theta, i$ correspond to the Neumann conditions

$$\left.\frac{\mathrm{d}\hat{\Psi}'^{n}_{m}}{\mathrm{d}r}\right|_{r=a,\,b} = 0 \quad \forall\, n, m\ . \tag{5.43}$$

The equilibrium solid-body rotation flow about which we perturb satisfies impermeability, but is not no-slip.

Since we are solving a second order differential equation, only two independent boundary conditions are required. We cannot therefore impose both impermeable and no-slip flow at both boundaries, as that would require four independent conditions. This over-constrained nature of the PPV inversion in Q-G models is discussed in Williams (1979). A comprehensive study of the comparative effects of using no-slip boundary conditions rather than the more traditional free-slip conditions is described by Mundt et al. (1995).

We are therefore forced to use a reduced set of boundary conditions, but we must choose carefully and consistently which conditions to retain and which to abandon, to avoid any possibility of non-physical behaviour. We are, of course, free to employ different boundary conditions for the different normal mode components specified by m and n.

The debate over suitable lateral Q-G boundary conditions has had a long and contentious history in the literature. In the classic periodic channel models of Phillips (1954) and Phillips (1956), boundary conditions corresponding to equation (5.42) are used for the wave $n \neq 0$ terms, and equation (5.43) is used for the mean flow correction $n = 0$ component only. The latter condition was not imposed (but the former was retained) in the studies of Phillips (1963) and Pedlosky (1964), but McIntyre (1967) showed that relaxing this mean flow correction boundary condition leads to a spurious, unspecified energy flux through the sidewalls. The condition was included again in Pedlosky (1970), but replaced in Pedlosky (1971) and Pedlosky (1972) with an *ad-hoc* condition chosen for mathematical convenience. Smith (1974) points out that the resulting non-physical energy source might well invalidate Pedlosky's results, and repeats Pedlosky's calculations with the proper boundary condition retained (Smith & Pedlosky, 1975; Smith, 1977). More recent studies (Appleby, 1982; Yoshida & Hart, 1986; Lewis, 1992; Stephen, 1998) have avoided the spurious energy and associated unreliable conclusions by applying both conditions in full, as in Phillips' original paper.

A useful interpretation of Phillips' mean flow correction boundary condition has been given by Davey (1978). For non-zero zonal perturbation velocities $u'_{i,\,\theta}|_{r=a,\,b}$ at the boundary between the interior and a Stewartson layer, there will be a corresponding return volume flux between the Ekman layers and the Stewartson layer due to the asymmetry of the Ekman spiral (Pedlosky, 1987), which will have a non-zero radial component proportional to $u'_{i,\,\theta}|_{r=a,\,b}$. We can therefore ensure that there is no net build-up of mass in the Stewartson layers by setting

$$\int_0^{2\pi} u'_{i,\,\theta}|_{r=a,\,b}\,\mathrm{d}\theta = 0 \quad \forall\, i\ . \tag{5.44}$$

This condition is automatically satisfied for the wave $n \neq 0$ components, and is equivalent to equation (5.43) with $n = 0$, which is the condition used by Phillips. With this condition, there is no *net* exchange of fluid due to the perturbation flow between each Ekman layer and the Stewartson layers, though *local* exchange is allowed.

Next, we attempt to derive a consistent and plausible set of boundary conditions for the annulus, which do not lead to non-physical behaviour, by considering integral properties of both the prognostic and diagnostic model equations.

5.4.1 Integral properties of the prognostic equations

Consider the area-integral of the perturbation PPV tendencies over the annular domain:

$$\int_{\theta=0}^{2\pi}\int_{r=a}^{b} \frac{\partial q'_i}{\partial t}\, r\,\mathrm{d}r\,\mathrm{d}\theta\ , \tag{5.45}$$

as given by the prognostic equations (5.23) and (5.24). The linear $\partial/\partial\theta$ forcing terms integrate to give zero unconditionally. The advection terms in the total derivatives integrate to give zero (Salmon & Talley, 1989) if

$$\left.\frac{\partial \psi'_i}{\partial \theta}\right|_{r=a,\,b} = 0\ , \tag{5.46}$$

and the Laplacian (∇^2) terms integrate to give zero if

$$\int_0^{2\pi} \left.\frac{\partial \psi_i'}{\partial r}\right|_{r=a,\,b} \mathrm{d}\theta = 0 \ . \tag{5.47}$$

The two conditions (5.46) and (5.47) are equivalent to impermeability for the waves and no-slip for the mean flow correction, as originally used by Phillips. With these conditions, the mean layer PPVs are conserved by the continuous equations and there is no spurious energy flux. We choose to apply these conditions to the present model, except that the second condition leads to an ill-posed PPV inversion for the special case $n = 0$, $m = \mathrm{bt}$, as we will see in Section 5.4.2.

5.4.2 Integral properties of the diagnostic equations

Equation (5.41) for the barotropic mean flow correction is

$$\frac{\mathrm{d}^2 \hat{\Psi}'^0_{bt}}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}\hat{\Psi}'^0_{bt}}{\mathrm{d}r} = \hat{Q}'^0_{bt} \ . \tag{5.48}$$

Since $\lambda_{\mathrm{bt}} = 0$ and $n = 0$ for this case, one of the terms in the radial structure equation has vanished, making the left side an exact differential. Equation (5.48) can therefore be integrated analytically between $r = a$ and $r = b$ to give

$$b\left.\frac{\mathrm{d}\hat{\Psi}'^0_{bt}}{\mathrm{d}r}\right|_{r=b} - a\left.\frac{\mathrm{d}\hat{\Psi}'^0_{bt}}{\mathrm{d}r}\right|_{r=a} = \int_a^b \hat{Q}'^0_{bt}\, r\, \mathrm{d}r \ . \tag{5.49}$$

We choose initial conditions for which the right side of this equation is zero, i.e. the barotropic PPV averaged over the 2-D annular domain is zero, and it is then guaranteed to remain so for all time, as shown in Section 5.4.1. This means that we need only explicitly set

$$\left.\frac{\mathrm{d}\hat{\Psi}'^0_{bt}}{\mathrm{d}r}\right|_{r=a} = 0 \tag{5.50}$$

and we will automatically have

$$\left.\frac{\mathrm{d}\hat{\Psi}'^0_{bt}}{\mathrm{d}r}\right|_{r=b} = 0 \tag{5.51}$$

from equation (5.49). If we explicitly set both (5.50) and (5.51) when solving (5.48), we have an underconstrained problem. We need to find an additional constraint, therefore, to close the solution.

We have defined two streamfunctions in the model — one per layer or, equivalently, one per vertical normal mode — and each of these has an integration constant associated with it (Section 5.3). Just because these two arbitrary constants have no physical meaning does not mean that they do not need to be defined in the numerical model. Now that we know that equations (5.50) and (5.51) are not independent boundary conditions, and therefore that to explicitly impose both would lead to an underconstrained PPV inversion, we choose to explicitly impose only equation (5.50). We then take the opportunity to use the remaining degree of freedom associated with the solution of equation (5.48) to define one of the streamfunction integration constants, by arbitrarily setting

$$\hat{\Psi}'^{0}_{bt}|_{r=b} = 0 \ , \tag{5.52}$$

which completes the set of two boundary conditions for the $m = \mathrm{bt}$, $n = 0$ case, and gives a well-posed problem.

Incidentally, the second streamfunction integration constant is defined by requiring the mean interface perturbation to be zero using equation (5.13), which follows from volume conservation for either layer. This requirement is imposed by adding a suitably-chosen constant to one of the streamfunction fields after the PPV inversion, and not as a boundary condition during the inversion.

A summary of the boundary conditions which we must explicitly set when integrating equation (5.41) is given in Table 5.1. With these conditions, the sidewall boundaries are impermeable to each component of the full flow — the solid-body rotation equilibrium flow, the mean flow correction and the eddy components. The boundaries are slippery to the solid-body rotation flow and the eddies, but no-slip to the mean flow correction.

5.5 Discretization of model equations

We have derived a set of model partial differential equations and boundary conditions which are both sensible and well-posed. We now discretize the equations so that they are suitable for numerical solution on a computer. We must take great care to ensure that the discretized equations and boundary conditions retain the important properties possessed by the continuous equations. In particular, it is important that they satisfy discretized analogues of the integral properties discussed in Section 5.4.

	$n = 0$	$n \neq 0$
$m = \mathrm{bt}$	$\left.\frac{\mathrm{d}\hat{\Psi}'^n_m}{\mathrm{d}r}\right\|_{r=a} = 0$ $\hat{\Psi}'^n_m\|_{r=b} = 0$	$\hat{\Psi}'^n_m\|_{r=a} = 0$ $\hat{\Psi}'^n_m\|_{r=b} = 0$
$m = \mathrm{bc}$	$\left.\frac{\mathrm{d}\hat{\Psi}'^n_m}{\mathrm{d}r}\right\|_{r=a} = 0$ $\left.\frac{\mathrm{d}\hat{\Psi}'^n_m}{\mathrm{d}r}\right\|_{r=b} = 0$	$\hat{\Psi}'^n_m\|_{r=a} = 0$ $\hat{\Psi}'^n_m\|_{r=b} = 0$

Table 5.1: Summary of suitable boundary conditions to apply to the streamfunction when integrating the continuous equations. Because the diagnostic Helmholtz equation relating ψ and q is second order, two conditions (one at each boundary) are required for each combination of vertical and azimuthal normal modes, denoted by m and n respectively.

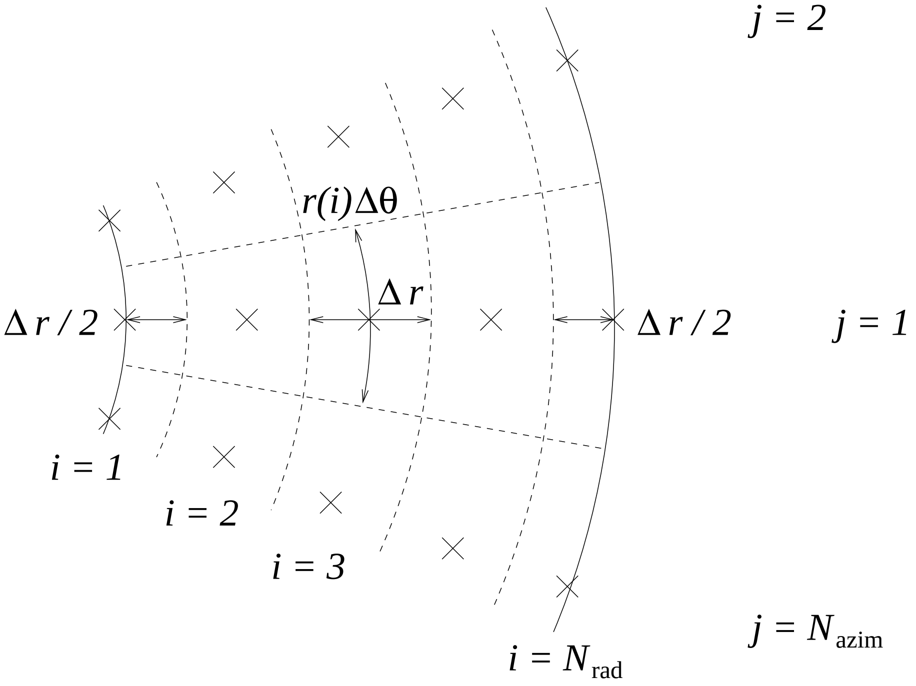


Figure 5.2: Definition of the model grid. Grid-points are marked with a "×" sign, grid-boxes with dashed lines, and the two cylindrical boundaries with solid lines. The dimensions of typical grid-boxes, both in the interior and at the boundary, are shown.

The grid on which we discretize the equations is shown in Figure 5.2. The grid consists of N_{rad} points in the radial dimension (including one point on each boundary $r = a$ and $r = b$), and N_{azim} points in the azimuthal dimension. We define

$$\Delta r = \frac{b - a}{N_{\text{rad}} - 1} \tag{5.53}$$

and

$$\Delta\theta = \frac{2\pi}{N_{\text{azim}}} \ , \tag{5.54}$$

and then we have

$$r(i) = a + (i - 1)\Delta r \ , \qquad i = 1, 2, \ldots, N_{\text{rad}} \tag{5.55}$$

and

$$\theta(j) = j\theta \ , \qquad j = 1, 2, \ldots, N_{\text{azim}} \ . \tag{5.56}$$

The point $(i, N_{\text{azim}} + 1)$ is equivalent to the point $(i, 1)$. We define the perturbation streamfunction $\psi'(i, j, k)$ and PPV $q'(i, j, k)$ at each of these points in each layer $k = 1, 2$, so that ψ' and q' are co-located on the grid. The area of the gridbox with coordinates (i, j) is approximately $[1 - \frac{1}{2}\delta_{i,\,1} - \frac{1}{2}\delta_{i,\,N_{\text{rad}}}]r(i)\Delta r\Delta\theta$, where δ is the Kronecker delta function.

5.5.1 Prognostic equations

In the continuous case, we chose perturbation streamfunction boundary conditions such that each of the three contributions to the area-integrated perturbation PPV tendency was zero. We would now like to choose discretizations of these contributions, together with discretizations of the boundary conditions, for which this statement still holds *exactly*. If our discretization only conserves mean PPV approximately, then there is the possibility of a non-physical and explosive increase in the PPV, even if the error is small, due to the compound effects of very many timesteps. Following Section 5.4.1, we therefore next examine the discretizations and boundary conditions necessary to ensure that

$$\sum_{i=1}^{N_{\text{rad}}} \sum_{j=1}^{N_{\text{azim}}} [1 - \frac{1}{2}\delta_{i,1} - \frac{1}{2}\delta_{i,N_{\text{rad}}}]\, f(i, j, k) r(i) \Delta r \Delta\theta = 0 \tag{5.57}$$

for $k = 1, 2$, where $f(i, j, k)$ is, in turn, the discretized azimuthal derivative, Jacobian and Laplacian.

Azimuthal derivative

The centred, second order discretization of the azimuthal derivative:

$$f(i,j,k) = \frac{\psi'(i,j+1,k) - \psi'(i,j-1,k)}{2\Delta\theta} \tag{5.58}$$

satisfies equation (5.57) unconditionally, as in the continuous case.

Jacobian

The second order Arakawa (1966) discretization of the Jacobian satisfies equation (5.57) if

$$\frac{\psi'(i,j+1,k) - \psi'(i,j,k)}{\Delta\theta} = 0 \quad \forall\, j,k,i = 1, N_{\text{rad}}\ , \tag{5.59}$$

which is a discretized version of the condition (5.46) for the continuous case.

Laplacian

It is tedious but straightforward to show that the five-point discretization of the Laplacian (whose continuous definition is given in equation (5.17) for reference):

$$\begin{aligned} f(i,j,k) &= \frac{\psi'(i+1,j,k) - 2\psi'(i,j,k) + \psi'(i-1,j,k)}{(\Delta r)^2} \\ &+ \frac{\psi'(i+1,j,k) - \psi'(i-1,j,k)}{2r(i)\Delta r} \\ &+ \frac{\psi'(i,j+1,k) - 2\psi'(i,j,k) + \psi'(i,j-1,k)}{[r(i)\Delta\theta]^2}\ , \end{aligned} \tag{5.60}$$

with *ghost point* values $\psi'(0,j,k)$ and $\psi'(N_{\text{rad}}+1,j,k)$ given by linear extrapolation:

$$\psi'(2,j,k) - \psi'(1,j,k) = \psi'(1,j,k) - \psi'(0,j,k) \tag{5.61}$$

$$\psi'(N_{\text{rad}}+1,j,k) - \psi'(N_{\text{rad}},j,k) = \psi'(N_{\text{rad}},j,k) - \psi'(N_{\text{rad}}-1,j,k)\ , \tag{5.62}$$

satisfies equation (5.57) if

$$\sum_{j=1}^{N_{\text{azim}}} \frac{\psi'(2,j,k) - \psi'(1,j,k)}{\Delta r} = 0 \quad \forall\, k \tag{5.63}$$

and

$$\sum_{j=1}^{N_{\text{azim}}} \frac{\psi'(N_{\text{rad}},j,k) - \psi'(N_{\text{rad}}-1,j,k)}{\Delta r} = 0 \quad \forall\, k\ , \tag{5.64}$$

which are discretized versions of the condition (5.47) for the continuous case. There will be a small error in the value of the discretized Laplacian at the boundaries due to the assumption of linearly-extrapolated ghost points, but there seems to be no other simple way to discretize the Laplacian in such a way that analogues of its integral properties are fully preserved.

5.5.2 Diagnostic equations

The discretized versions of equations (5.39) and (5.40) are

$$\begin{aligned}\Psi'_m(i,j) &= \sum_{n=0}^{N_{\text{azim}}-1} \hat{\Psi}'^n_m(i) e^{2\pi\sqrt{-1}nj/N_{\text{azim}}} \ , \quad (5.65)\\ Q'_m(i,j) &= \sum_{n=0}^{N_{\text{azim}}-1} \hat{Q}'^n_m(i) e^{2\pi\sqrt{-1}nj/N_{\text{azim}}} \ . \quad (5.66)\end{aligned}$$

The summations have been truncated, compared to equations (5.39) and (5.40), because there are only N_{azim} independent Fourier components associated with the discrete Fourier transform of a series of N_{azim} numbers.

Because $\Psi'_m(i,j)$ is real, we have

$$\hat{\Psi}'^{N_{\text{azim}}-n}_m(i) = \left[\hat{\Psi}'^n_m(i)\right]^* \ , \qquad n = 1, 2, \ldots, N_{\text{azim}} - 1 \ . \quad (5.67)$$

We choose N_{azim} to be even, and then we need only explicitly solve equation (5.41) for $n = 0, 1, 2, \ldots, N_{\text{azim}}/2$. Solutions for $n = N_{\text{azim}}/2 + 1, \ldots, N_{\text{azim}} - 1$ are given in terms of solutions for $n = N_{\text{azim}}/2 - 1, \ldots, 1$ by equation (5.67), halving the processing time required for the PPV inversions. The maximum resolvable wavenumber is the Nyquist wavenumber, $N_{\text{azim}}/2$.

In terms of the normal mode variables, the discretized boundary conditions (5.59), (5.63) and (5.64) reduce, on substitution into equations (5.65) and (5.66), to

$$\left.\begin{aligned}\hat{\Psi}'^n_m(1) &= 0\\ \hat{\Psi}'^n_m(N_{\text{rad}}) &= 0\end{aligned}\right\} \quad \forall\ m, n \neq 0 \quad (5.68)$$

and

$$\left.\begin{aligned}\hat{\Psi}'^0_m(1) &= \hat{\Psi}'^0_m(2)\\ \hat{\Psi}'^0_m(N_{\text{rad}}) &= \hat{\Psi}'^0_m(N_{\text{rad}} - 1)\end{aligned}\right\} \quad \forall\ m \ . \quad (5.69)$$

We now consider the discretization of the radial structure equation (5.41). Using centred three-point finite differences at the interior points $i = 2, 3, \ldots, N_{\text{rad}} - 1$, we obtain

$$\begin{aligned} \frac{\hat{\Psi}'^n_m(i-1) - 2\hat{\Psi}'^n_m(i) + \hat{\Psi}'^n_m(i+1)}{(\Delta r)^2} & \\ + \frac{\hat{\Psi}'^n_m(i+1) - \hat{\Psi}'^n_m(i-1)}{2r(i)\Delta r} & \\ - \left[\lambda_m + \frac{n^2}{[r(i)]^2}\right] \hat{\Psi}'^n_m(i) &= \hat{Q}'^n_m(i) \; . \end{aligned} \tag{5.70}$$

Re-grouping terms according to grid-points gives

$$\alpha^-(i)\hat{\Psi}'^n_m(i-1) + \gamma(i)\hat{\Psi}'^n_m(i) + \alpha^+(i)\hat{\Psi}'^n_m(i+1) = \hat{Q}'^n_m(i)(\Delta r)^2 \; , \tag{5.71}$$

where the dimensionless quantities $\alpha^\pm$ and γ are given by

$$\alpha^\pm(i) = 1 \pm \frac{\Delta r}{2r(i)} \tag{5.72}$$

and

$$\gamma(i) = -2 - \left[\lambda_m + \frac{n^2}{[r(i)]^2}\right] (\Delta r)^2 \; . \tag{5.73}$$

In Cartesian geometry we would have $\alpha^\pm(i) = 1$.

The $N_{\text{rad}} - 2$ equations (5.71), together with 2 boundary conditions, complete the set of N_{rad} equations in the N_{rad} unknowns $\hat{\Psi}'^n_m(i), \quad i = 1, 2, \ldots, N_{\text{rad}}$. These linear equations can be written in matrix form:

$$\begin{pmatrix} \text{bdy} & \text{bdy} & & & & \cdots \\ \alpha^-(2) & \gamma(2) & \alpha^+(2) & & & \cdots \\ & \alpha^-(3) & \gamma(3) & \alpha^+(3) & & \cdots \\ & & \alpha^-(4) & \gamma(4) & \alpha^+(4) & \cdots \\ & & & \alpha^-(5) & \gamma(5) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \hat{\Psi}'^n_m(1) \\ \hat{\Psi}'^n_m(2) \\ \hat{\Psi}'^n_m(3) \\ \hat{\Psi}'^n_m(4) \\ \hat{\Psi}'^n_m(5) \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{Q}'^n_m(2)(\Delta r)^2 \\ \hat{Q}'^n_m(3)(\Delta r)^2 \\ \hat{Q}'^n_m(4)(\Delta r)^2 \\ \hat{Q}'^n_m(5)(\Delta r)^2 \\ \vdots \end{pmatrix} \tag{5.74}$$

where the zero elements in the tridiagonal N_{rad} by N_{rad} matrix have been left blank. The two elements labelled “bdy” are boundary condition elements, dependent upon m and n, and there are two more such elements in the final two columns of the bottom row.

5.6 Perturbation streamfunction boundary conditions for the discretized equations

In the continuous case, we found that the boundary conditions for the barotropic mean flow correction component ($m = \mathrm{bt}$, $n = 0$) were ill-posed as originally stated, and remained so until we replaced a redundant boundary condition with an equation to define an integration constant (Section 5.4). This happens in the discretized case, too: the square matrix in equation (5.74) is singular for the barotropic mean flow correction, when the boundary condition elements "bdy" are $(-1, 1)$ in the top row and $(1, -1)$ in the bottom row. The analytical proof of this, which involves showing that a certain linear combination of rows is zero, is tedious but straightforward. By analogy with the continuous case, we replace the two boundary condition elements in the bottom row with $(0, 1)$ to define the integration constant by setting the streamfunction for this component to zero on the outer boundary, and then the matrix is no longer singular (typical condition numbers are given in Section 5.9).

In the continuous system, we set the $n = 0$, $m = \mathrm{bt}$ normal streamfunction derivative to zero at one boundary and found that, if the mean barotropic PPV was zero, the streamfunction derivative would automatically be zero at the other boundary (Section 5.4.2). Importantly, in contrast with the continuous system, this statement does not hold exactly for the discretized system. This is because $\hat{Q}'^{n}_{m}(1)$ and $\hat{Q}'^{n}_{m}(N_{\mathrm{rad}})$ do not appear in equation (5.74); we do not apply the discretized differential equation at the boundaries, as we need to use these two degrees of freedom to set the boundary conditions.

The error corresponding to this PPV leak is small ($\sim (\Delta r)^2$), but even small errors can grow to dominate the solution after a large number of timesteps. To fix this problem with the barotropic mean flow correction, we discard the outer boundary streamfunction $\hat{\Psi}'^{0}_{\mathrm{bt}}(N_{\mathrm{rad}})$ obtained through inversion of equation (5.74) and define a new value for it by setting $\hat{\Psi}'^{0}_{\mathrm{bt}}(N_{\mathrm{rad}}) = \hat{\Psi}'^{0}_{\mathrm{bt}}(N_{\mathrm{rad}} - 1)$. This ensures that the boundary conditions (5.69) required for conservation of mean PPV are satisfied, but the consequence is that the discretized differential equation (5.70) is not exactly satisfied at the point $N_{\mathrm{rad}} - 1$. The imposed boundary conditions are summarized in Table 5.2.

	$n = 0$	$n \neq 0$
$m = \mathrm{bt}$	$\hat{\Psi}'^n_m(2) - \hat{\Psi}'^n_m(1) = 0$ $\hat{\Psi}'^n_m(N_{\mathrm{rad}}) = 0^\dagger$	$\hat{\Psi}'^n_m(1) = 0$ $\hat{\Psi}'^n_m(N_{\mathrm{rad}}) = 0$
$m = \mathrm{bc}$	$\hat{\Psi}'^n_m(2) - \hat{\Psi}'^n_m(1) = 0$ $\hat{\Psi}'^n_m(N_{\mathrm{rad}}) - \hat{\Psi}'^n_m(N_{\mathrm{rad}} - 1) = 0$	$\hat{\Psi}'^n_m(1) = 0$ $\hat{\Psi}'^n_m(N_{\mathrm{rad}}) = 0$

Table 5.2: Summary of the boundary conditions applied to the streamfunction when integrating the discretized equations. The analogous conditions for the for the continuous case are given in Table 5.1. †After the inversion, $\hat{\Psi}'^0_{\mathrm{bt}}(N_{\mathrm{rad}})$ is redefined by $\hat{\Psi}'^0_{\mathrm{bt}}(N_{\mathrm{rad}}) - \hat{\Psi}'^0_{\mathrm{bt}}(N_{\mathrm{rad}} - 1) = 0$, as discussed in the text.

5.7 Details of the numerical schemes

Time stepping

For the time-stepping we use a leapfrog scheme with a Robert (1966) 3-level time filter applied at each timestep, to suppress the computational mode splitting between even and odd numbered steps (Mesinger & Arakawa, 1976). At each step, of size Δt, q^{t+1} is determined at each grid point using the leapfrog scheme:

$$q^{t+1} = q^{t-1} + 2\Delta t \ q^t_{\mathrm{tendency}} \ , \tag{5.75}$$

and then the value of q^t is adjusted in such a way as to move it closer to the mean of q^{t-1} and q^{t+1}:

$$q^t \to q^t + R\left(\frac{q^{t-1} + q^{t+1}}{2} - q^t\right) \ . \tag{5.76}$$

The old value of q^t is abandoned and the new, filtered value is used in its place. The *Robert filter parameter* $R > 0$ is chosen to be as small as possible whilst still suppressing the leapfrog decoupling.

Time-lagged diffusion

Numerical solutions of the simple diffusion equation, using the leapfrog scheme for the time-discretization and a time-centred three-point finite difference for the space-discretization, are unconditionally unstable due to a computational mode (Haltiner & Williams, 1980). To avoid this in the present model, we time-lag the diffusion terms by one timestep when evaluating the right sides of the discretized analogues of equations (5.23) and (5.24). This means that, when evaluating the PPV tendency at timestep t, we calculate the forcing ($\partial/\partial\theta$) and advection terms using the fields at timestep t, but calculate the diffusion (∇^2) terms using the fields at timestep $t-1$.

Hyperdiffusion

To represent sub-gridscale effects we add a hyperdiffusion term to the right sides of the prognostic equations (5.23) and (5.24), as is usual in numerical models (e.g. Lewis, 1992).

At first, a fourth-order streamfunction hyperdiffusion term $\nu_{\text{hyper}}\nabla^4\psi_i'$ was tried, but significant gridscale features were always found to form at the lateral boundaries whenever the model was run. This is because during the PPV inversion, any gridscale features in the PPV field will give rise to corresponding grid-scale features in the perturbation streamfunction field, and then the $\nu_{\text{hyper}}\nabla^4\psi'$ contribution to the PPV tendency will tend to damp out these features in the PPV field. Unfortunately this does not happen at the boundaries in the discretized system, because boundary values of the PPV are not used when performing the inversion. As already discussed, $\hat{Q}'^n_m(1)$ and $\hat{Q}'^n_m(N_{\text{rad}})$ are missing from equation (5.74). Values of PPV therefore are able to feed back into the PPV tendency field only at interior points, and there is nothing to suppress grid-scale features in the PPV field at the boundaries.

To avoid this, we instead use second-order hyperdiffusion applied to the PPV, by adding a term $\nu_{\text{hyper}}\nabla^2 q_i'$ to the prognostic equations. This term is also time-lagged by one timestep, as discussed above. The hyperdiffusion term does not exactly satisfy equation (5.57), though the error is small. In order to keep the model solutions as close as possible to the continuous equations solutions, we reset the mean PPV to zero after each timestep, by adding a very small constant whose value is chosen to fit this requirement.

Summary of numerical integration scheme

A flow chart summarizing the details of the numerical integration scheme is shown in Figure 5.3. Given the PPV fields at times $t-1$ and t, we invert to obtain the streamfunction fields at those times, which then allows us to calculate all the contributions to the PPV tendency. We perform a leapfrog time integration to obtain the PPV field at time $t+1$, and then modify the PPV field at time t by applying a Robert filter. Once we have obtained $q'(t)$ and $q'(t+1)$ from $q'(t-1)$ and $q'(t)$, we discard $q'(t-1)$ and $\psi'(t-1)$, we dump $q'(t)$ and $\psi'(t)$ to disk, then we re-label $t \rightarrow t-1$ and begin the loop again.

The system state is completely determined by ψ'. Note that it is also completely determined by q' together with the boundary conditions, because equations (5.25) and (5.26) are uniquely invertible. It is not necessary to dump both ψ' and q' to disk in order to have a complete description of the system, therefore. Nevertheless, we choose to save both fields, in order to reduce the need for further calculations when plotting model diagnostics.

5.8 Initial conditions

A feature of the leapfrog timestepping scheme is that initial condition fields are required at two separate times, in order to begin the integration. As shown in Figure 5.3, we choose to specify the PPV fields as initial conditions. We use small amplitude random noise for these fields, seeding the system to permit the growth of unstable perturbations of any azimuthal and radial wavenumber. The intrinsic Fortran function `RANDOM_NUMBER` is used to generate random numbers with a uniform distribution which are shifted to a chosen interval centred on zero. We then subtract the mean PPV in each layer at both timesteps, which makes the fields satisfy the zero mean barotropic PPV condition of Sections 5.4.2 and 5.6.

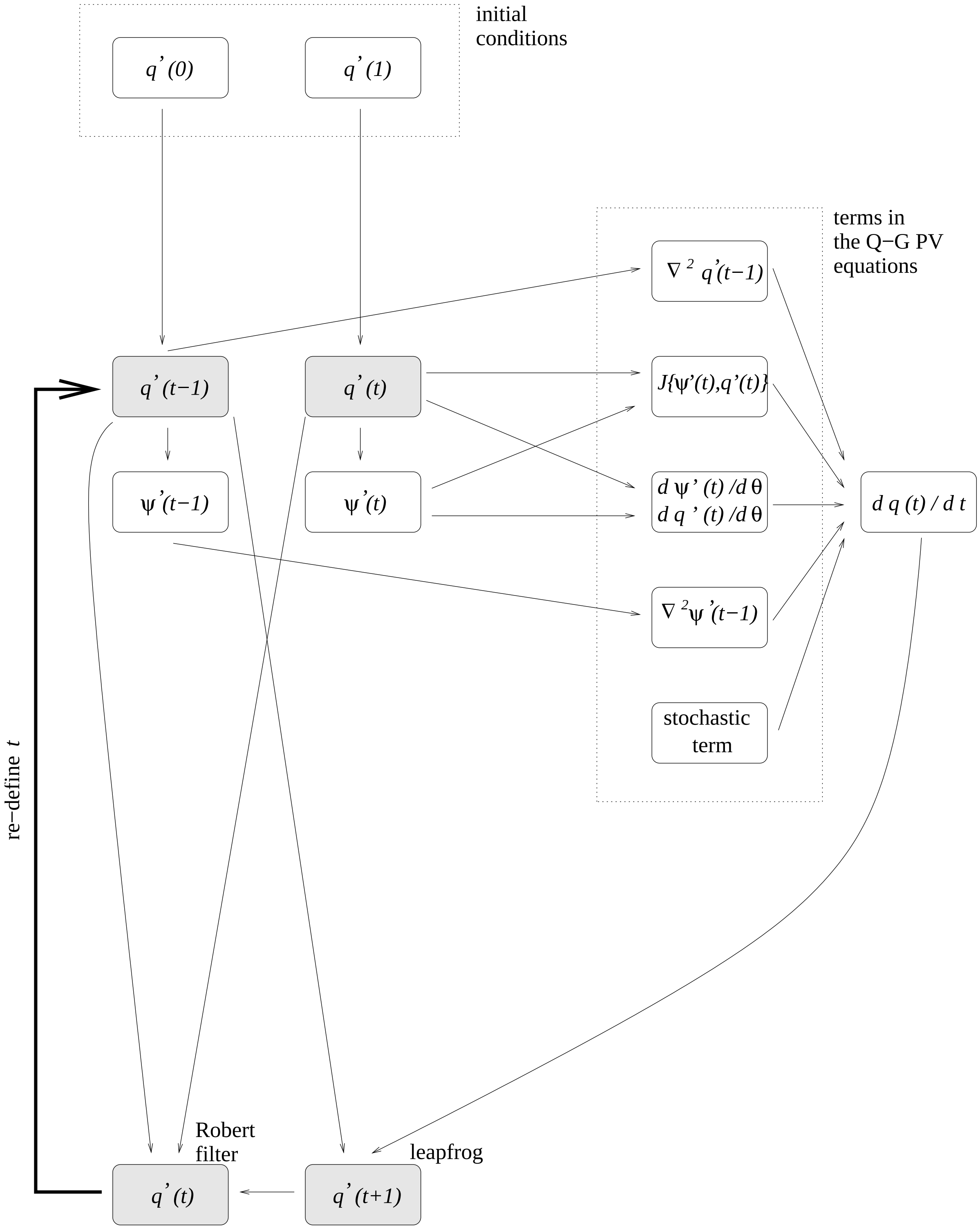


Figure 5.3: Organigram showing how the the model integrations progress, starting with initial conditions $q'(0)$ and $q'(1)$. Each timestep has inputs $q'(t-1)$ and $q'(t)$ and outputs $q'(t)$ and $q'(t+1)$, shown shaded. $J(\psi', q') = [(\partial\psi'/\partial r)(\partial q'/\partial\theta) - (\partial\psi'/\partial\theta)(\partial q'/\partial r)]/r$ is the Jacobian. The stochastic term, designed to represent the effects of small-scale waves, is introduced in Chapter 7 and is left switched off until then.

	m=bt	m=bc
$n = 0$	389	59
$n = 1$	112	35
$n = 2$	99	33
$n = 3$	82	31
$n = 4$	67	29
$n = 5$	54	26
$n = 6$	44	24
$n = 7$	36	21
$n = 8$	31	19
$n = 9$	26	17

Table 5.3: Estimates of the condition numbers (in the infinity-norm) of the tridiagonal matrices in equation (5.74), corresponding to the first 10 azimuthal modes for both of the vertical modes. Values given are rounded to the nearest integer.

5.9 Suitable values for numerical parameters

Code to carry out the numerical integrations described in this chapter has been written in Fortran 95 by the author and his supervisors, and compiled using the Numerical Algorithms Group (NAG) f95 compiler for Linux. Routines from the NAG library were employed: `nag_fft` for the transformations between real and spectral space described by equations (5.65) and (5.66), and `nag_gen_bnd_lin_sys` for solving the complex band matrix equation (5.74) a large number ($\sim N_{\text{azim}}$) of times each timestep.

All model runs described in this thesis were performed using double numerical precision (retaining 16 significant figures) for the calculations, and single numerical precision (retaining 8 significant figures) for the dumps to disk. The factor by which relative errors in the perturbation streamfunction are greater than relative errors in the PPV is known as the *condition number* of the corresponding matrix. Some typical condition numbers for the matrices in equations (5.74) are shown in Table 5.3. The largest condition number in the system has a value of a few hundred, implying that only the last two significant figures of the inferred perturbation streamfunctions will be uncertain, and that errors due to rounding are therefore small.

The azimuthal derivative, Laplacian and Jacobian routines were each tested using input

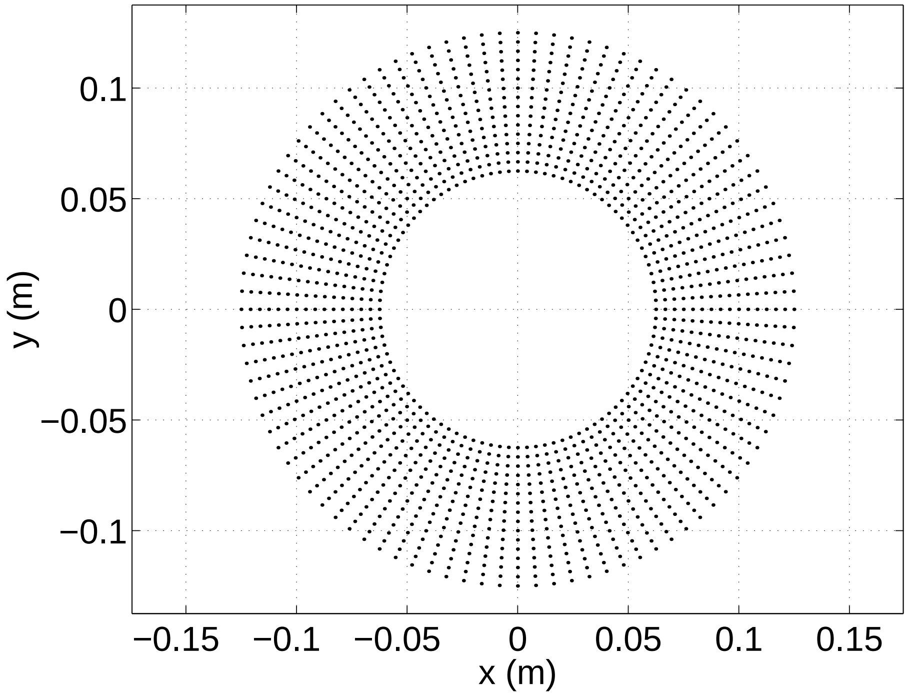


Figure 5.4: Gridpoint positions used for the model runs in this thesis. There are 96 points in the azimuthal direction and 16 in the radial direction, giving gridboxes which are approximately square near the inner boundary.

fields consisting of random numbers satisfying the boundary conditions. The mean PPV tendency due to each contribution was found to be zero to within numerical precision, implying that the code for these routines is free from errors. The Helmholtz solver was tested by first using the forward formulae (5.25) and (5.26) with our discretized Laplacian (5.60)–(5.62) to calculate the PPV fields corresponding to given random perturbation streamfunction fields, and then using the Helmholtz solver routine to reconstruct the streamfunction fields from the calculated PPVs. The root-mean-square difference between the original and reconstructed streamfunction fields was around 0.1%, implying that the solver code was also free from errors. The reason that the agreement is not exact, to within numerical precision, is that we assume linearly-extrapolated ghost points to evaluate the Laplacian in the forward formulae — an assumption which is not made during the inversion.

For all the model runs described in this thesis (unless stated otherwise), the annulus dimensions a, b and H have the values indicated in Figure 2.3, the fluid properties ρ_i, ν_i and S are as indicated in Table 2.1, and the acceleration due to gravity is $g = 9.81$ m s^{-2}. The FFT is much faster if the only prime factors of N_{azim} are 2, 3 and 5, and so we use a grid defined by $N_{\text{azim}} = 2^5 \times 3 = 96$ and $N_{\text{rad}} = 16$, as shown in Figure 5.4. The Robert filter parameter is $R = 0.01$. For given Ω and $\Delta\Omega$, we take the amplitude of the random initial PPV perturbation to be $\Delta\Omega/100$ so that we are assessing the growth of very small perturbations, we choose the timestep Δt to be such that the azimuthal Courant number $\frac{1}{2}\Delta\Omega\Delta t/\Delta\theta$ is 0.01, and we choose the hyperdiffusion coefficient ν_{hyper} to be such that the e-folding time $1/(\nu_{\text{hyper}} k^2_{\text{Nyquist}})$ for damping of mid-radius gridscale waves with the Nyquist wave vector $k_{\text{Nyquist}} = N_{\text{azim}}/(a + b)$ is equal to one lid rotation period.

In order to demonstrate insensitivity to the numerical parameters, comparative runs were done with (separately) the hyperdiffusion coefficient decreased by a factor of 10, the Robert filter parameter decreased by a factor of 10 and the gridspacing doubled in both directions, but all other parameters unmodified. The equilibrated wave number was the same in each case, and the mid-radius wave amplitude and phase speed differed by at most 0.3%. We have therefore demonstrated that both rounding errors and discretization errors are small, and that the equilibrated state is insensitive to the values of the numerical parameters, implying that the model output gives an accurate representation of the true solutions of the continuous model equations.

The code is very efficient: on a Linux workstation with a 1.4 GHz *AMD Athlon* processor and 100% of the CPU usage, and with $N_{\text{azim}} = 96$ and $N_{\text{rad}} = 16$, a model integration speed of 120 timesteps per second is attained. Since timesteps of up to around 0.1 s can be used stably, the model can run ten times faster than the laboratory annulus. The run-time memory requirement is 3.1 MB.

5.10 Chapter summary

We have constructed a multi-layer cylindrical quasi-geostrophic numerical model of the rotating annulus laboratory experiment, and named it QUAGMIRE. Great care has been taken to choose discretizations and boundary conditions which are both physically sensible and computationally stable, and as a result the model gives reliable solutions of the continuous equations. Large series of model runs have been carried out for comparison with the laboratory results, and are described in the next chapter.

Chapter 6

Results of the numerical experiments

"The purpose of models is not to fit the data but to sharpen the questions."

Samuel Karlin,
11th R. A. Fisher Memorial Lecture,
The Royal Society, 20 April 1983.

A large number of annulus flows have been simulated using QUAGMIRE, and the results are described in this chapter. A Matlab diagnostics package has been written by the author, to read in the raw data dumped to disk and plot it and other derived quantities. Comparisons are made between flow properties in the model and the laboratory, both as a check that the model works properly and reliably, and to investigate whether we can attribute any differences to short waves present in the laboratory but not in the model.

A major advantage of the numerical model is that velocity fields — unavailable in the laboratory — can easily be derived from the streamfunctions. The model velocity fields are used in this chapter to compute various diagnostics, each of which is expected to have some ability in predicting regions of generation of small-scale waves, either due to a shear instability or to spontaneous emission. By identifying the indicator with the best predictive skill, we draw conclusions about the mechanism which is most likely to be responsible for the observed mixed-wave short emissions in the laboratory.

6.1 Model runs with zero interfacial tension

The main series of model simulations carried out during this study consisted of 210 separate runs, one for each combination of 10 values of Ω and 21 values of $\Delta\Omega$ given by

$$\Omega/\text{rad s}^{-1} \in \{1.00, 1.50, 1.75, 2.00, 2.25, 2.50, 2.75, 3.00, 3.25, 3.50\}$$

and

$$\begin{aligned} \Delta\Omega/\text{rad s}^{-1} \in\ & \{0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.08, 0.10, 0.12, 0.15, \\ & 0.20, 0.23, 0.30, 0.40, 0.50, 0.60, 0.70, 0.85, 1.06, 1.31, 1.61\}. \end{aligned}$$

These values were chosen to give a roughly uniform density of sampled points in the $(\log[d], F)$ parameter space. The interfacial tension S was set to zero for these runs, a condition which will be relaxed for the runs to be described in Section 6.4. Starting from noisy initial conditions (Section 5.8), each run was continued for an integration time equal to 60 lid rotation periods, which was usually found to be sufficient for waves arising from any baroclinic instability to have equilibrated at finite amplitude. In a few cases full equilibration was not achieved within this time and so the integration was continued for a further 60 lid rotation periods. The 210 runs required around six days of computer time to complete, and took up 3 GB of disk space dumping both the streamfunction and PPV fields once every 500 timesteps.

A difference between the QUAGMIRE runs and the laboratory scans of Chapter 4 is that for the model simulations the state was reset to the appropriate initial condition (small-amplitude noise superimposed onto the background equilibrium state) before each new run. In contrast, in the continuous laboratory experiment scans, the previously attained flow served as the effective initial condition. We do this so that for each parameter combination, we are examining the stability of small perturbations to the axisymmetric equilibrium state, rather than the stability of an equilibrated finite amplitude large-scale mode corresponding to a neighbouring point of parameter space. Though this approach maintains a close association between the model and the theoretical studies of baroclinic instability, it does mean that the model and laboratory experiments do not correspond to exactly the same problem.

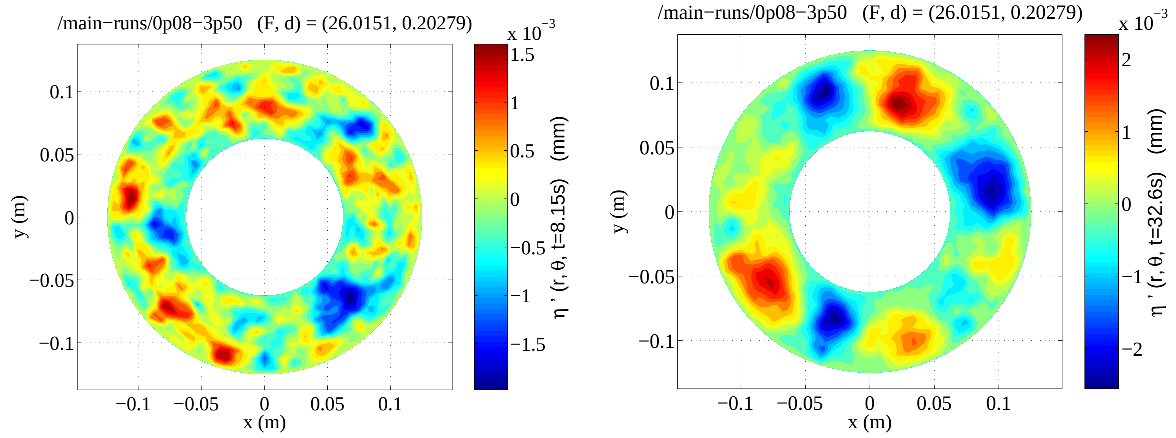
6.1.1 Sample diagnostics

We now show some sample model diagnostics for the simulation with $\Delta\Omega = 0.08$ rad s^{-1} and $\Omega = 3.50$ rad s^{-1}, demonstrating how we can diagnose azimuthal wavenumbers, interfacial wave amplitudes and wave phase speeds from the raw model data.

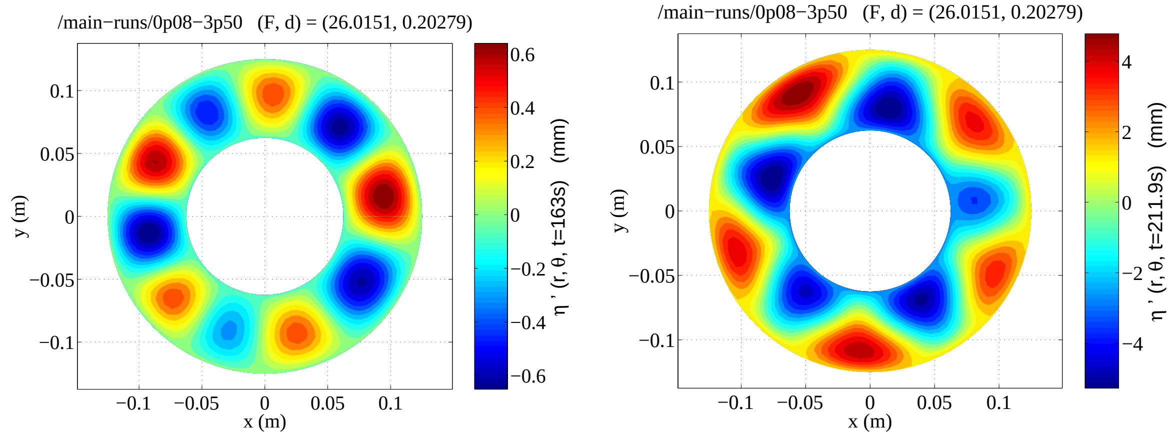
Figure 6.1 shows the perturbation interface height field (calculated from equation 5.30) at six different timesteps, showing how the system evolves from the noisy initial state in (a), via an azimuthal wavenumber 5 mode with initial largest growth rate in (b)–(e), to the equilibrated wavenumber 3 mode in (f). Note that the selected mode is not that with the initial largest growth rate, which was quite a common occurrence in the model runs. This is due to the development of a radial mode between timesteps 10,000 and 14,000 as seen in (c), (d) and (e), which seems to alter the relative stability of the azimuthal wavenumber 3 and 5 modes. Appleby (1988) has discussed the difficulty of predicting, in a two-layer flow simultaneously baroclinically-unstable to two discrete wavelengths, which of the two modes will eventually predominate. Note the close resemblance between the wave shapes in Figures 6.1(f) and 4.16(a), giving our first evidence that the simulated flows are reasonable. A quantitative model/laboratory comparison is carried out in Section 6.1.2.

Azimuthal wavenumbers can easily be read off by eye from the interface height plots, but it is convenient to automate this procedure when it needs to be done for many hundreds of model runs. To this end, the azimuthal Fourier component amplitudes corresponding to the field in Figure 6.1(f) are plotted in Figure 6.2. The wavenumber 3 component is the largest, as expected, but the azimuthal profiles are not perfectly sinusoidal. There are significant wavenumber 6 and 9 harmonics, with amplitudes around a factor of 10 smaller than the dominant component. There is only slight evidence of an energy build-up at the gridscale (wavenumbers of around $N_{\mathrm{azim}}/2 = 48$), showing that the numerical hyperdiffusion term (Section 5.7) is successfully suppressing any spurious growth.

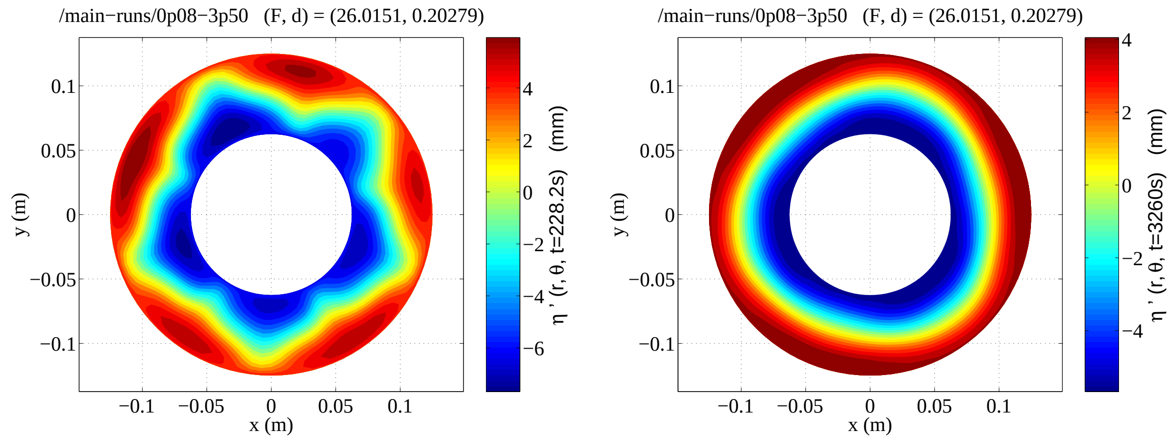
In order to determine wave amplitudes, Figure 6.3 shows azimuthal profiles of full (i.e. mean plus perturbation) interface height corresponding to Figure 6.1(f), at each of the 16 model radii. The height shows no variation with azimuth at the two sidewall boundaries, a consequence of both layer streamfunctions being constant there. Wave



(a) 500 steps (b) 2000 steps



(c) 10,000 steps (d) 13,000 steps



(e) 14,000 steps (f) 200,000 steps

Figure 6.1: Evolution of the perturbation interface height field η', from random small-amplitude initial conditions to an equilibrated large-scale mode with azimuthal wavenumber 3. Note that the colourbar scales vary between the plots.

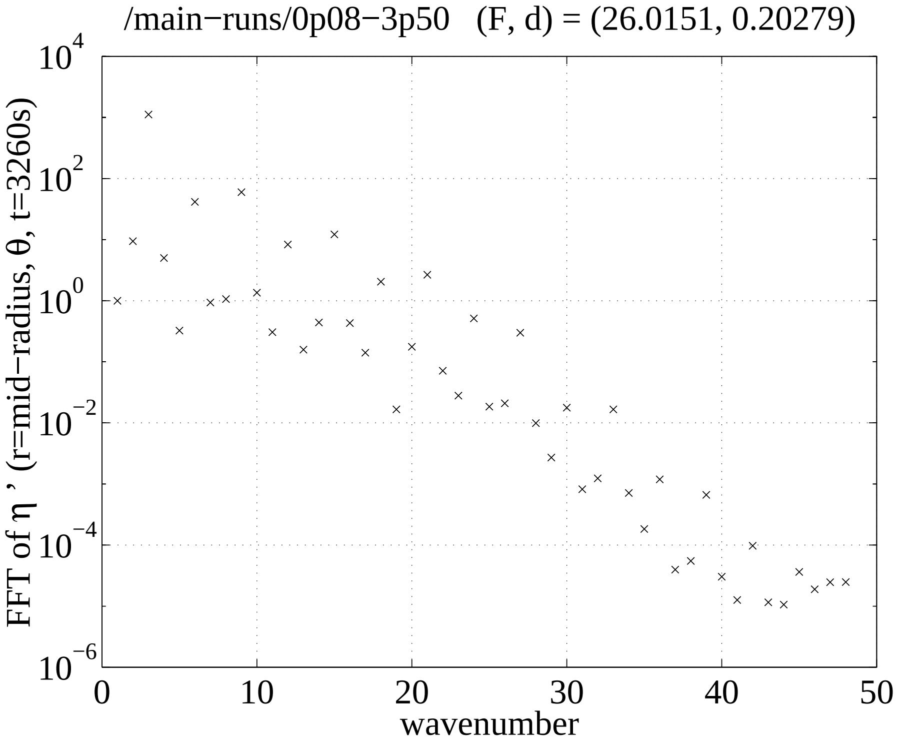


Figure 6.2: Amplitudes of azimuthal Fourier components of the mid-radius perturbation interface height η' at timestep 200,000, normalized by the wavenumber 1 amplitude.

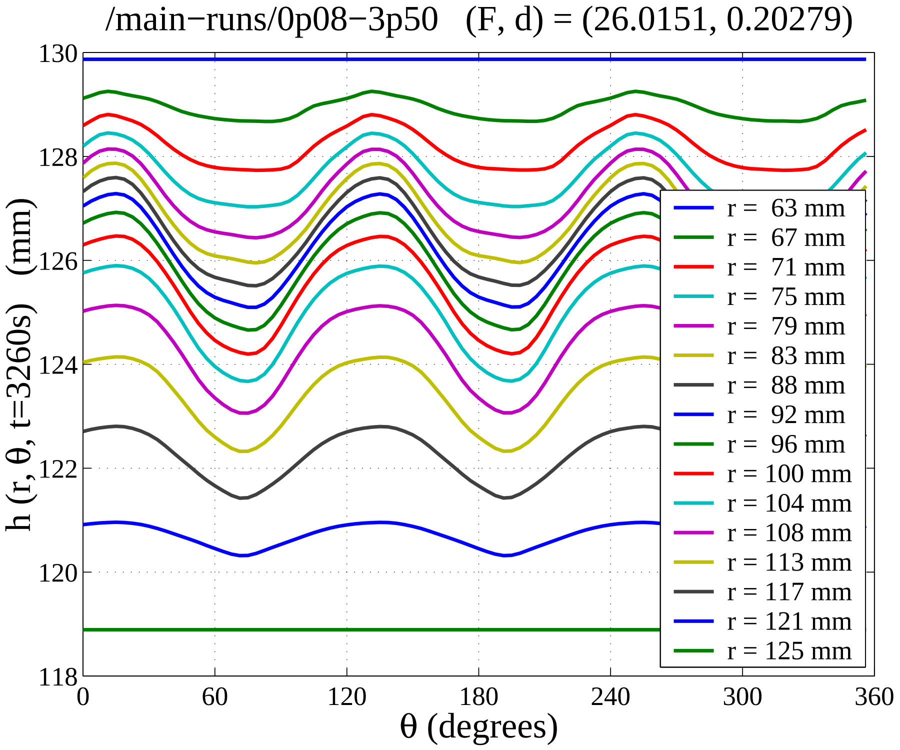


Figure 6.3: Azimuthal profiles of the full interface height field $h = H + \overline{\eta} + \eta'$ at timestep 200,000, for various radii shown rounded to the nearest millimetre.

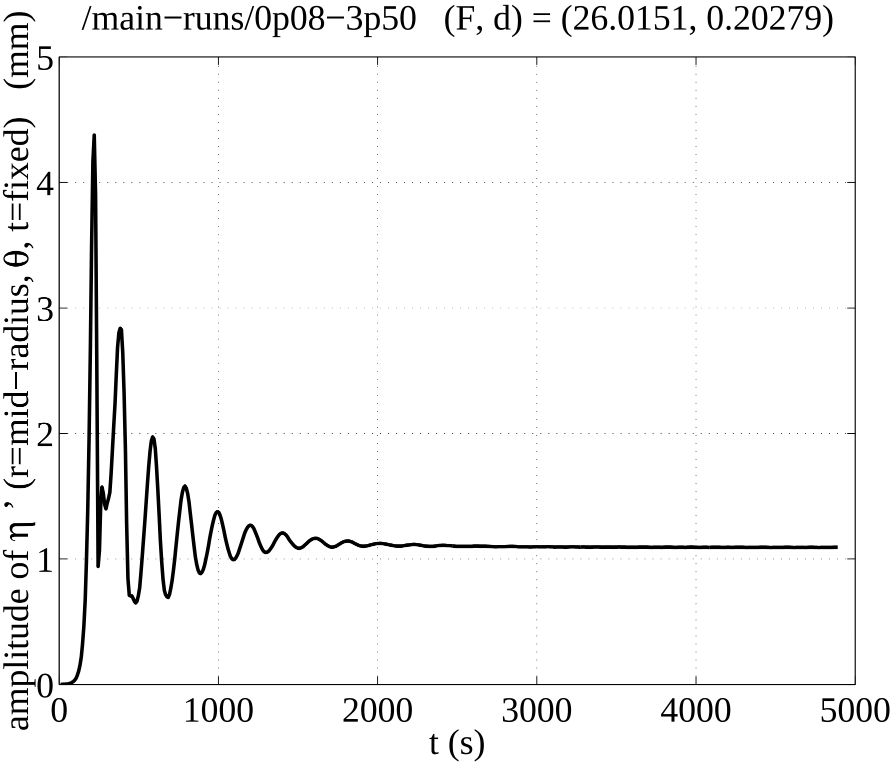


Figure 6.4: Time series of amplitude of mid-radius perturbation interface height η' up to timestep 300,000, starting at time zero with random small-amplitude initial conditions.

amplitudes at each radius can be computed from diagrams such as this, by halving the difference between the maximum and minimum displacements. In this case the amplitude is largest near the middle of the annular gap, where it reaches around 1 mm.

Figure 6.4 shows a timeseries of mid-radius wave amplitude for the entire model run. The wavenumber 5 mode is seen to grow very rapidly at the start of the run, but then gives way to the selected wavenumber 3 mode after around 300 s. The mode 3 amplitude temporarily vacillates as it decays towards its equilibrated value of 1.1 mm, after which it remains constant.

By watching the equilibrated wave as it drifts around the annulus, we can derive its phase speed. Figure 6.5 shows a post-equilibrium timeseries of perturbation interface height displacement at a fixed mid-radius point, from which we can verify our previous finding that the wave amplitude is 1.1 mm. By taking a temporal Fourier transform we find the predominant period to be 51.7 s, which we must multiply by the azimuthal wavenumber to obtain the drift period of 155 s. This is closely equal to double the lid

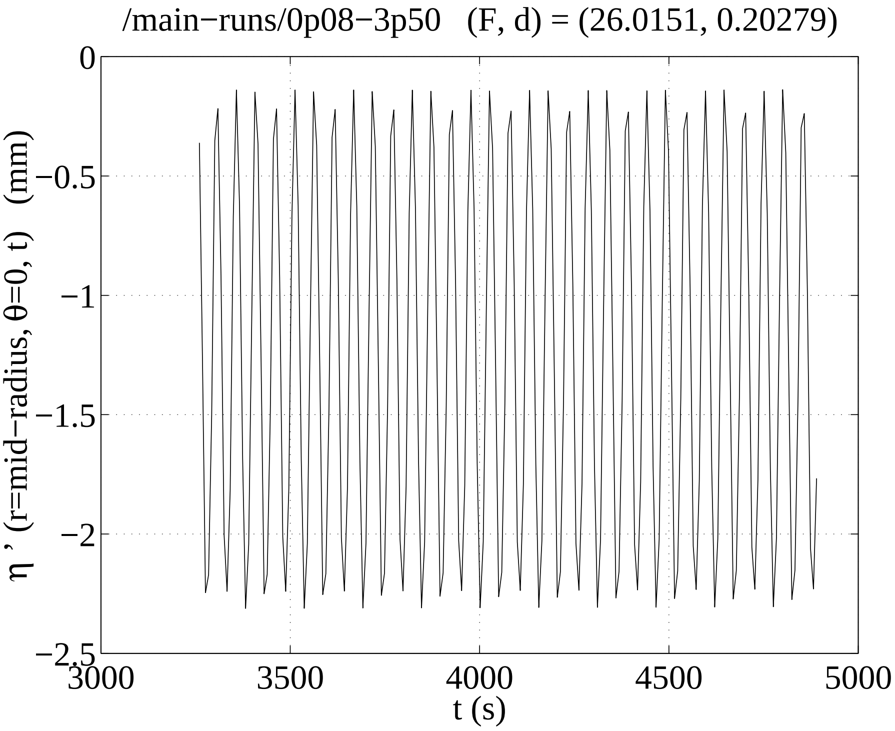


Figure 6.5: Time series of mid-radius perturbation interface height displacement η' at a fixed azimuthal point, between timesteps 200,000 and 300,000. The three lobes of the large-scale wave, which pass the measurement point in turn, have slightly different amplitudes suggesting that perfect equilibration has not quite been reached.

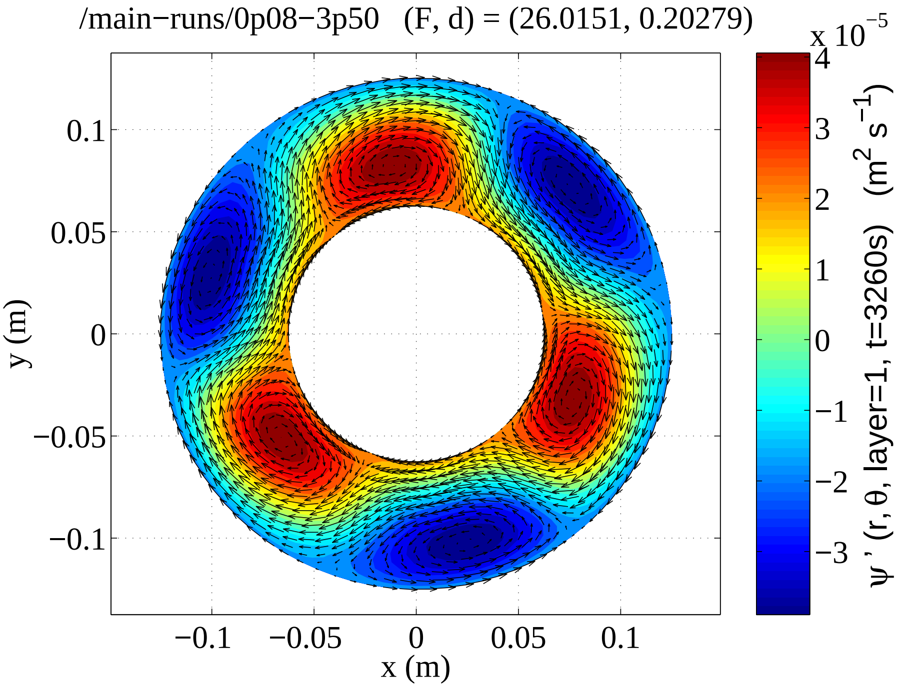


Figure 6.6: Contour plot of perturbation streamfunction ψ_1' in the upper layer at timestep 200,000. Perturbation velocity vectors $\boldsymbol{u}_1'$, derived from the streamfunction field, are over-plotted. The largest perturbation velocity vector in the plot has a magnitude of 2.1 mm s^{-1}.

rotation period $2 \times 2\pi/\Delta\Omega = 157$ s, so that the wave angular phase speed is half the lid rotation speed in this case.

We can derive the perturbation velocity components from the perturbation streamfunctions, using equations (5.28) and (5.29). These two fields in the upper layer, after equilibration, are shown over-plotted in Figure 6.6. The streamfunction is constant along the boundaries, meaning that the radial velocity component is zero there. Both constants have adjusted to take different values, however, meaning that there is a non-zero radially-averaged azimuthal velocity component, which corresponds to the mean flow correction (Section 5.3.2). Six eddies have developed: three with rotation in the clockwise sense and three in the anti-clockwise sense. The net azimuthal fluid transport due to the perturbation velocities of these eddies is in the retrograde (clockwise) direction, though this transport is weaker than that of the mean flow and so the overall transport is still in the prograde direction. In the lower layer, the net perturbation transport is in the prograde (anti-clockwise) direction (not shown). The wave arising from the baroclinic instability has therefore tended to reduce the mean velocity shear across the interface which caused the instability.

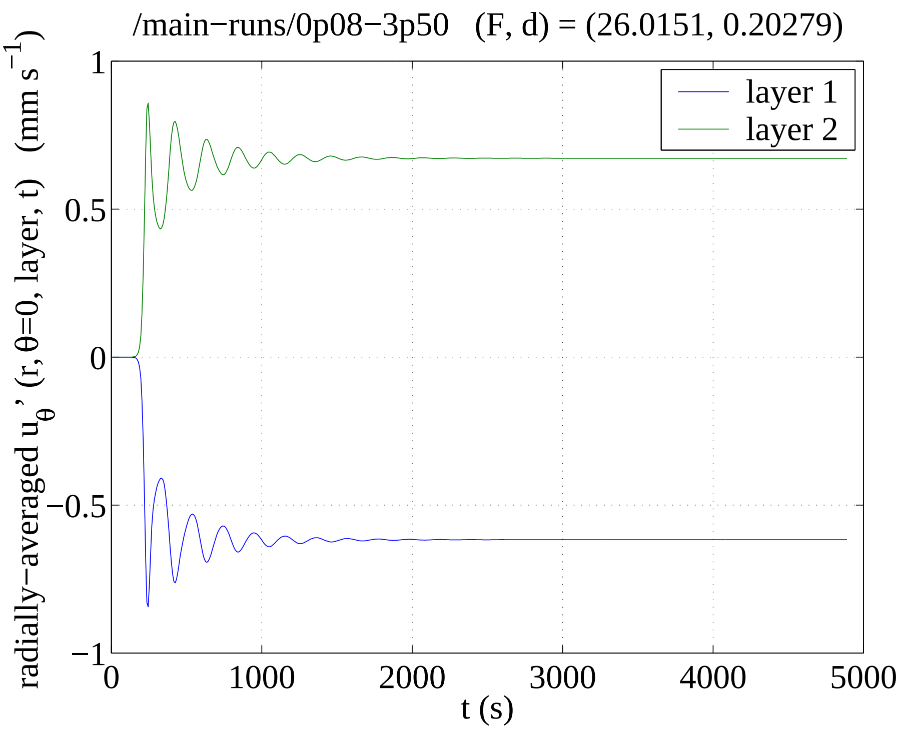


Figure 6.7: Time series of radially-averaged zonal perturbation velocity $\overline{u_\theta{}'}$ in both layers, up to timestep 300,000.

We can evaluate the radially-averaged azimuthal velocity components responsible for reducing the shear, by taking the difference between the streamfunctions at $r = a$ and $r = b$ and dividing by $b - a$. The results of this for both layers are shown in Figure 6.7 as functions of time. The series are almost exactly equal and opposite, implying that the mean azimuthal velocity is almost purely baroclinic, with a much smaller barotropic component. This finding is consistent with the highly-truncated two-layer model of Lovegrove (1997), in which certain interaction coefficients are shown to be zero because of an additional symmetry introduced due to the layer depths being equal. This leads to the equilibrated barotropic mean flow correction component being zero (see Lovegrove's equation (1.17) and following comments).

As an example of an amplitude vacillation simulated by QUAGMIRE, Figure 6.8 shows a timeseries of post-transient amplitude from the run with $\Delta\Omega = 0.70$ rad s^{-1} and $\Omega = 3.25$ rad s^{-1}. In this case, the equilibrated state is a 1AV, with an amplitude envelope that is not sinusoidal. The vacillation period is 47.5 s and the wave drift period in this case is 17.3 s (not shown), so that the wave drifts completely around the tank about

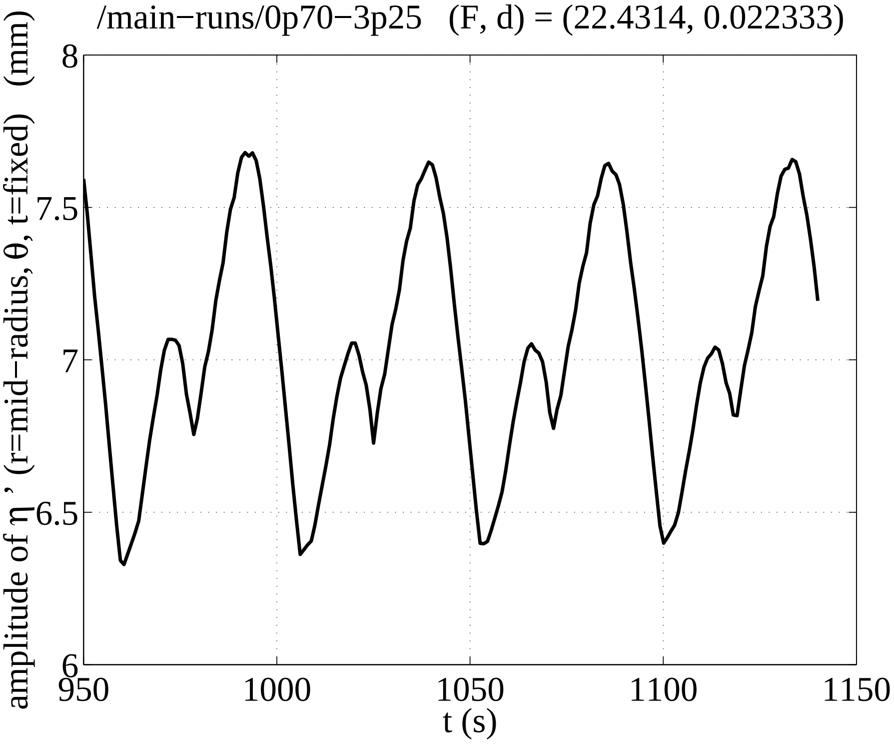


Figure 6.8: Time series of amplitude of mid-radius perturbation interface height η', after the initial transients have decayed away, showing an amplitude vacillation envelope.

three times between successive peaks of the vacillation cycle. This is in good agreement with typical vacillation periods observed in the laboratory experiment.

6.1.2 Comparison between model and laboratory

We have shown how wavenumbers, amplitudes and phase speeds can be derived from the raw data generated by the QUAGMIRE numerical experiments. These three quantities are also readily available from the calibrated laboratory experiments, and we now undertake a comparison between the two as an important test of QUAGMIRE's reliability.

Wavenumber comparison

The dominant azimuthal wavenumber after equilibration has been determined for each of the 210 model runs described in Section 6.1, and the resulting numerical regime diagram is shown in Figure 6.9. There are well-defined regimes everywhere, apart from at low d and high F where the wavenumber 1 and 2 regions become entangled and confused.

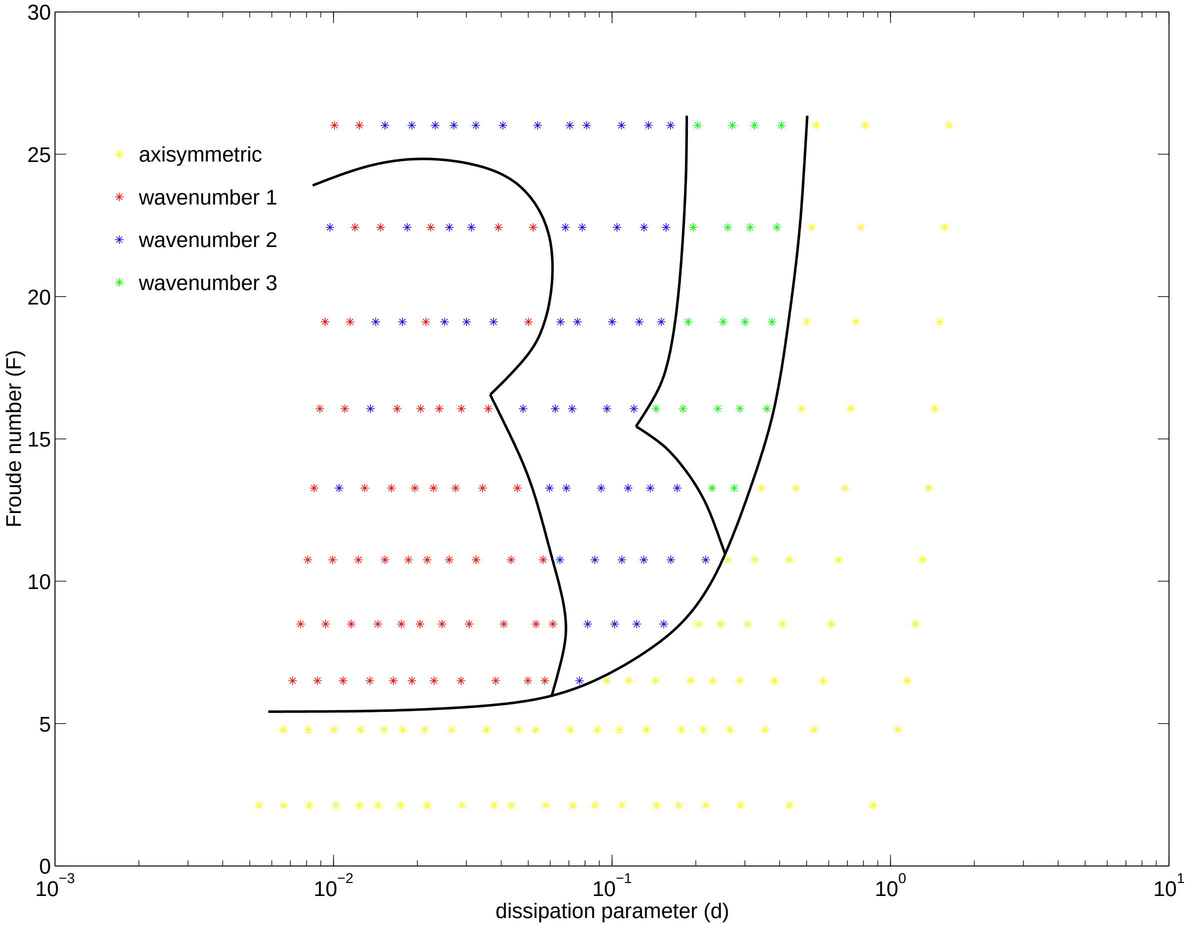


Figure 6.9: Model regime diagram, showing equilibrated azimuthal wavenumber $m \in \{0, 1, 2, 3\}$ at each of the 210 points investigated. Wavenumber transition curves have been inferred and over-plotted.

A small number of 10-member ensemble runs was carried out, in order to test sensitivity to initial conditions. The runs within each ensemble had the same run parameters, but different random numbers for the noise in the initial state. All ten members gave equilibrated flows which were identical in terms of wavenumber, wave speed and amplitude, except in the low-d and high-F corner where there was a probability partition between wavenumber 1 and 2. In this region there is high sensitivity to initial conditions, which help to determine the final state, whereas elsewhere in the regime diagram there is insensitivity to the precise details of the noise in the initial conditions. We return to examine the response of the model to small-scale random noise in Chapter 7, when we use a noisy forcing term to represent the laboratory small-scale waves.

The laboratory regime diagram corresponding to Figure 6.9 is shown in Figure 4.1. There is excellent qualitative agreement between the shapes of the model and laboratory wavenumber transition curves. Quantitative agreement is limited by a shift in the regime features in the (d, F) plane between the two diagrams. For example, consideration of the coordinates of the $m = 0, 1, 2$ and $m = 0, 2, 3$ transition curve triple points shows that the model overestimates F by a factor of 1–2 and d by a factor of 5–10. The error in F is small, and can be attributed to the many approximations made when deriving the model Q-G equations. The error in d is significantly larger. In the non-dimensionalized governing equations, d is the coefficient of the Ekman layer terms. The mismatch between QUAGMIRE and laboratory regime diagrams therefore suggests that the model assumption of linear, parameterized Ekman layers is inadequate. This is perhaps not surprising, as the Ekman velocity formulae used in the model equations are derived under the assumption of geostrophy, but the Rossby numbers reached in the laboratory *MRW* regime can be as large as 1 (see Figure 4.6(b)).

Wave speed comparison

The post-transient angular phase speed of the waves has been determined for each of the baroclinically-unstable model runs, and is shown in Figure 6.10. The wave speed shows no variation with turntable speed Ω, as in the laboratory experiments, and in each case it is close to half the lid rotation speed $\Delta\Omega$. These two statements hold even

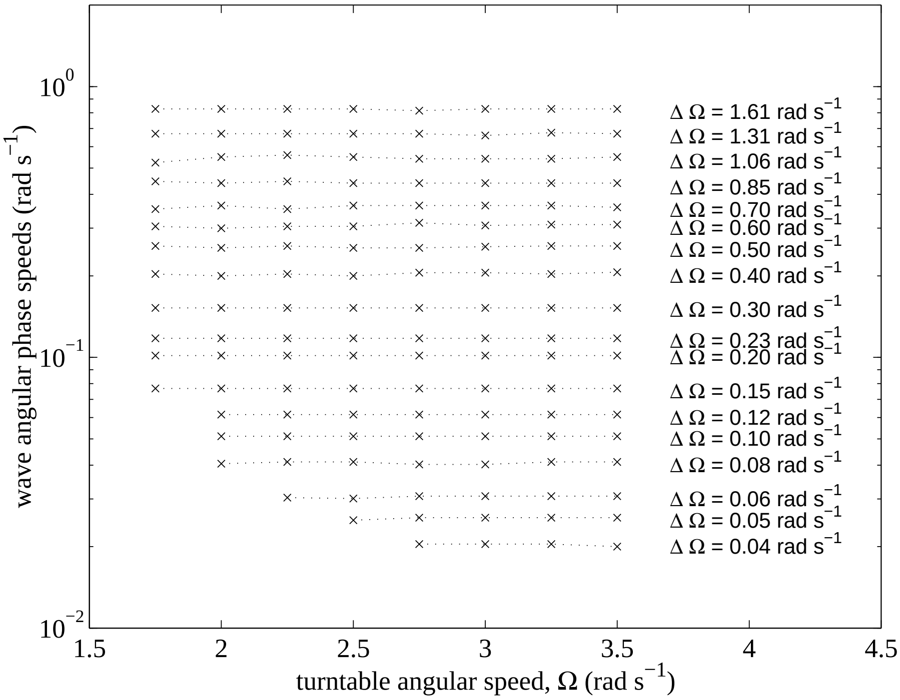


Figure 6.10: Angular phase speed of the model waves after equilibration, plotted against Ω for various $\Delta\Omega$, for each model run which exhibited baroclinic instability.

as wavenumber transition curves are crossed. Since the equilibrium Q-G rotation rates are closely one-quarter and three-quarters of the lid rotation speed (Section 5.3.1), we conclude that the model waves travel at the mean of the equilibrium layer rotation rates.

The variation of model wave speed with lid speed is shown in Figure 6.11 for the case $\Omega = 2.0$ rad s^{-1}. Over-plotted on the same figure is the equivalent laboratory data, obtained from experiments PAI1–10 after 90 minutes by timing drift periods with a stopwatch. The model overestimates the wave speeds by a factor of four, presumably due (at least in part) to the importance of Stewartson layer drag discussed in Section 3.4.3, which is present in the laboratory but absent in the model, and also to the uncertainty in the fluid properties, to be investigated in Sections 6.2 and 6.3.

Wave amplitude comparison

A systematic model/laboratory amplitude comparison is difficult because a given Ω and $\Delta\Omega$ will correspond to different wavenumber regimes in the laboratory and model, due to the shift of features in the (d, F) parameter space, discussed above. Whilst wave

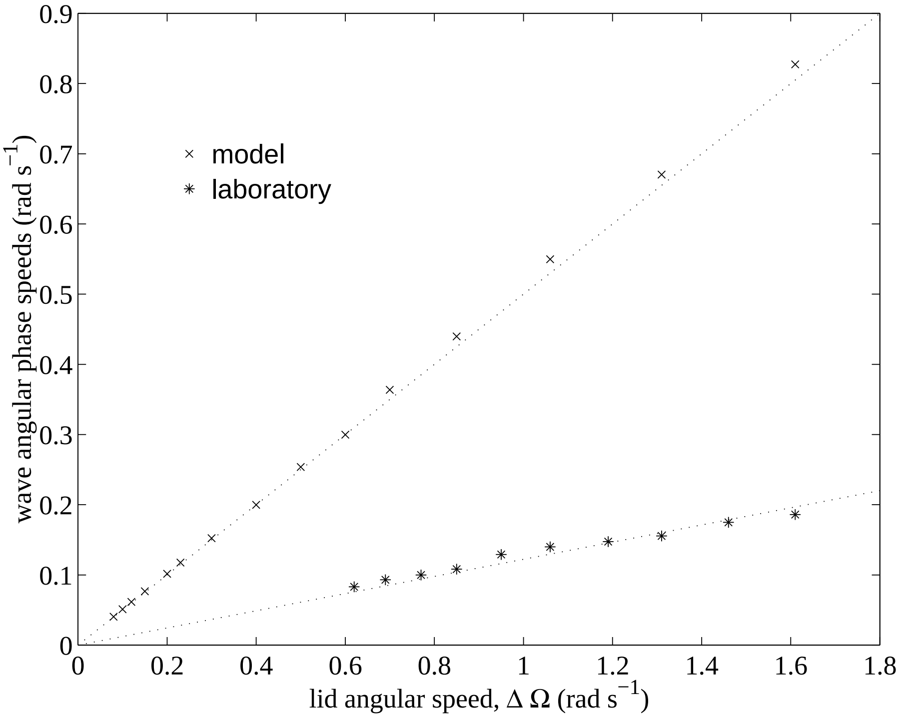


Figure 6.11: Angular phase speed of the wave after equilibration, plotted against $\Delta\Omega$ for $\Omega = 2.0$ rad s^{-1}. Data from both the model and laboratory experiments are shown, together with straight lines (dotted) of gradient 0.50 and 0.12 passing through the origin.

speeds are independent of wavenumber, as shown above, we might expect that wave amplitudes are not. Instead, the mid-radius wave amplitudes after equilibration were determined for each of the baroclinically-unstable model runs, and typical values were used for the comparison, because amplitude variations within wavenumber regimes were small. Typical amplitudes of wavenumber 1, 2 and 3 flows in the model were found to be around 5 mm, 2 mm and 1 mm respectively, a factor of a few smaller than the laboratory amplitudes reported in Section 4.8. We investigate this mismatch, together with the phase speed mismatch, in Sections 6.2 and 6.3.

6.2 Variation of model viscosity

Sufficient evidence has been accumulated (Sections 2.2.4 and 4.2) to suggest that there may be slow changes in the physical properties of the working liquids. In this section and the next, we vary the model viscosities and interfacial tension, respectively, to investigate whether these changes can help to explain the observed discrepancies between the

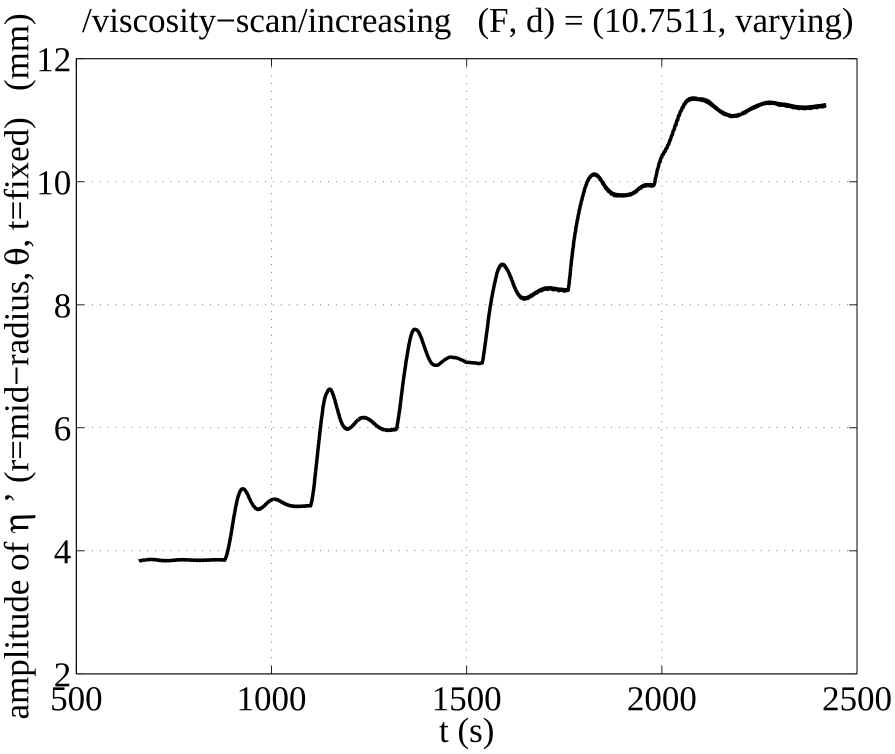


Figure 6.12: Timeseries of mid-radius wave amplitude, showing how the amplitude responds to six discontinuous increases in lower layer viscosity.

amplitudes and phase speeds of the large-scale waves reported above. Preliminary laboratory evidence for changes in the large-scale waves with time, and therefore with fluid properties, has been presented in Section 4.2. On these grounds we expect the model waves to display a variation with fluid properties also.

The experiment with $\Delta\Omega = 0.60$ rad s^{-1} and $\Omega = 2.25$ rad s^{-1} was repeated with the lower layer viscosity ν_2 varying throughout the run. Starting with a wave which had equilibrated with the measured laboratory viscosity (Table 2.1), ν_2 was increased discontinuously to a new value and the system was allowed to re-equilibrate, and then ν_2 was increased again, etc. The increase was by a factor of around two in each case, and the flow remained 1S throughout. The resulting amplitude trace is shown in Figure 6.12. There is a ringing effect as the system adjusts to each of the six changes. Though the viscosities reached by the end of the run were unrealistically large, just a single doubling of ν_2 from its assumed value produces a significant increase in wave amplitude, of around 25%.

The viscosity scan experiment described above was repeated once more, this time with

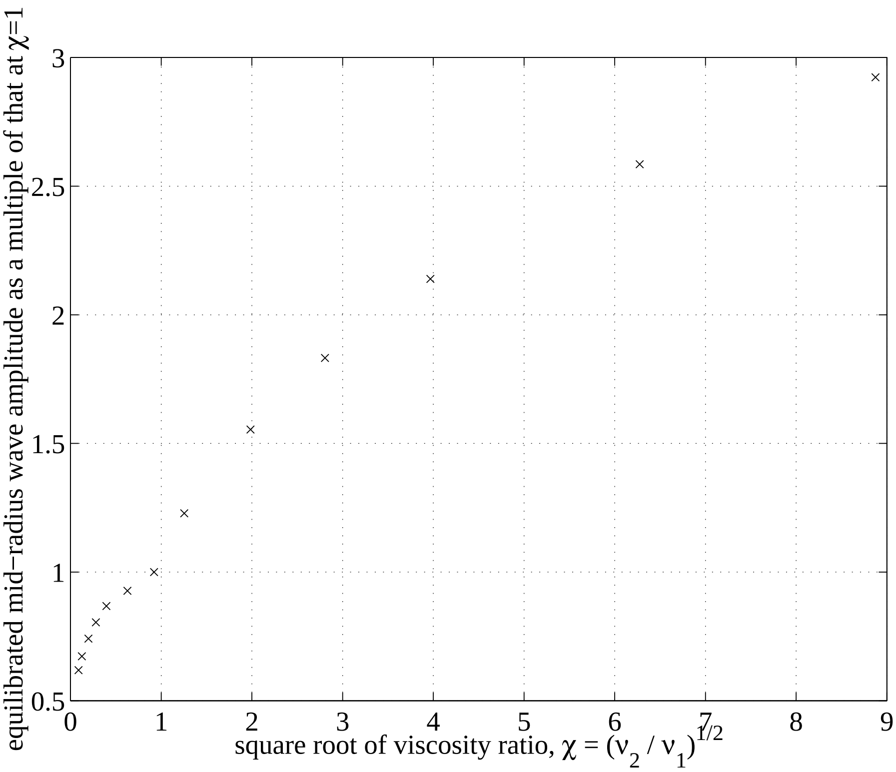


Figure 6.13: Variation of equilibrated mid-radius wave amplitude with viscosity ratio, given as a multiple of its value when the viscosities are equal.

six discontinuous decreases in lower layer viscosity. The equilibrated wave amplitudes obtained from each of the 13 viscosities investigated are shown in Figure 6.13, as a function of $\chi = \sqrt{\nu_2/\nu_1}$. There are two distinct branches of the curve corresponding to increasing and decreasing χ. If ν_2 were twice as large as its assumed value, and ν_1 were half as large, then χ would be twice its assumed value and the wave amplitude would be over 50% larger.

The reason that wave amplitudes vary with viscosity seems to be that viscosity affects the energetics of the system. From equations (5.20) and (5.21), as $\chi \to 0$, $\Delta\Omega_1 \to \Delta\Omega$ and $\Delta\Omega_2 \to \frac{1}{2}\Delta\Omega$, whereas as $\chi \to \infty$, $\Delta\Omega_1 \to \frac{1}{2}\Delta\Omega$ and $\Delta\Omega_2 \to 0$. Therefore, as χ increases, the kinetic energy of the fluids decreases. There seems to be a corresponding increase in the gravitational potential energy, which manifests itself as an increase in the interfacial wave amplitude.

The wave angular phase speed was also determined for each of the 13 viscosities investigated, and is plotted as a function of χ in Figure 6.14. There is a strong variation with

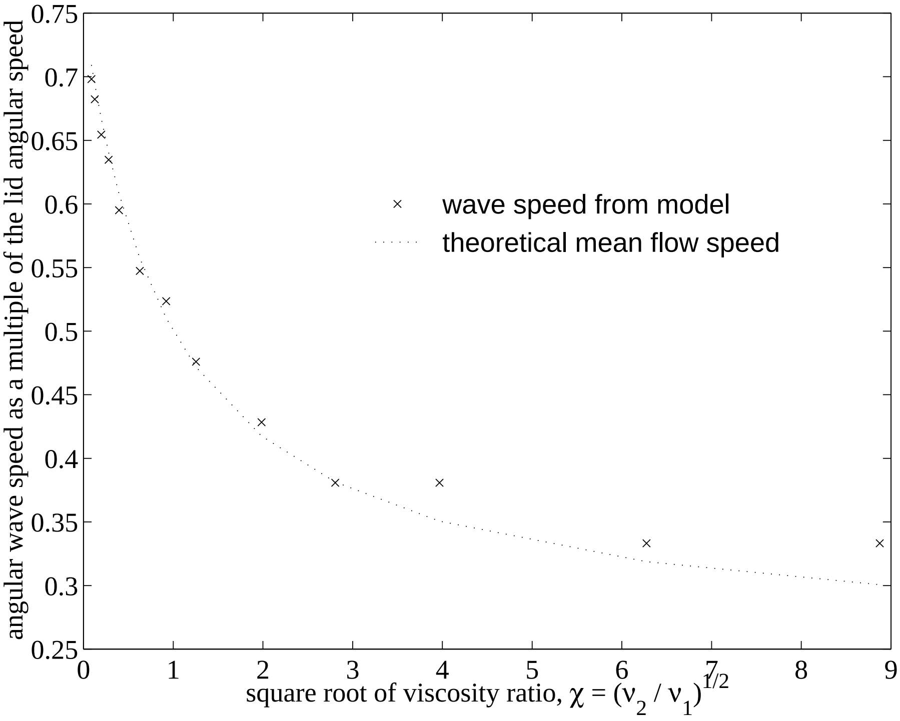


Figure 6.14: Variation of equilibrated wave angular phase speed with viscosity ratio, given as a fraction of the angular lid speed. The dotted line represents the mean of the equilibrium layer rotation rates (see text).

χ. Over-plotted is a curve corresponding to the mean of the equilibrium layer rotation rates, $\frac{1}{2}(\Delta\Omega_1 + \Delta\Omega_2)$, determined theoretically as a function of χ from equations (5.20) and (5.21). There is a good fit between the curve and the 13 points, showing that QUAGMIRE waves travel at the mean layer speed for all viscosities in this range. An increase in χ from 1 to 2 would decrease the model wave speeds by around 20%.

The increase in model wave amplitude and decrease in model wave speed as χ is doubled do not fully account for the laboratory/model disagreement reported in Section 6.1.2. Since uncertainties in the viscosity can only partially account for the discrepancy, in the following section we investigate the effects of uncertainties in the interfacial tension.

6.3 Variation of model interfacial tension

The experiment with $\Delta\Omega = 0.60$ rad s^{-1} and $\Omega = 2.25$ rad s^{-1} was repeated with the interfacial tension S varying throughout the run. Starting with a wave which had equilibrated with zero tension, S was increased discontinuously to a new value and the system

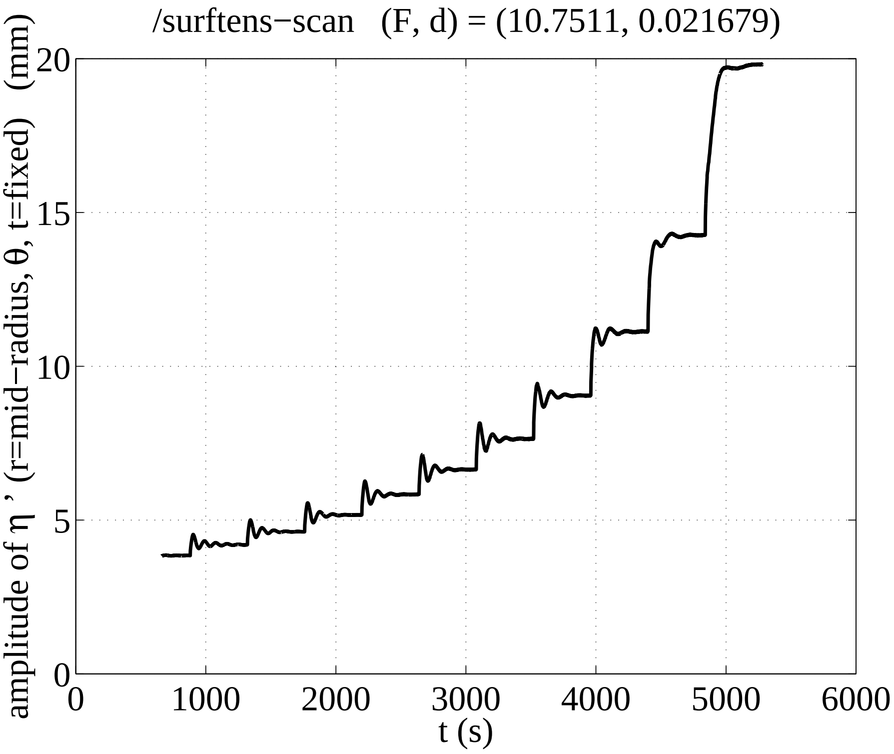


Figure 6.15: Timeseries of mid-radius wave amplitude, showing how the amplitude responds to ten discontinuous jumps in interfacial tension.

was allowed to re-equilibrate, and then S was increased again, etc. The increase was by 10^{-3} N m^{-1} in each case, and the flow remained 1S throughout. The resulting amplitude trace is shown in Figure 6.15. As with the viscosity scans, there is a ringing effect as the system adjusts to each change.

The equilibrated wave amplitudes obtained from each of the 11 interfacial tensions investigated are shown in Figure 6.16, as a function of S. The interfacial tension reached at the end of the run was 10^{-2} N m^{-1}, only around a third of the assumed laboratory value in the absence of a surfactant (Table 2.1), but the product FI had reached 0.47 by this stage and QUAGMIRE requires $FI \ll 1$ (Section 5.3). We therefore do not expect the rapid amplitude growth with S shown in the figure to be continued as FI approaches unity, but we can infer that the assumed interfacial tension can increase QUAGMIRE amplitudes by at least a factor of five relative to the $S = 0$ case.

The wave angular phase speed was also determined for each of the 11 interfacial tensions investigated, and is plotted as a function of S in Figure 6.17. There is no variation with S, the speed taking on the mean layer speed $0.5\Delta\Omega$ for all investigated interfacial

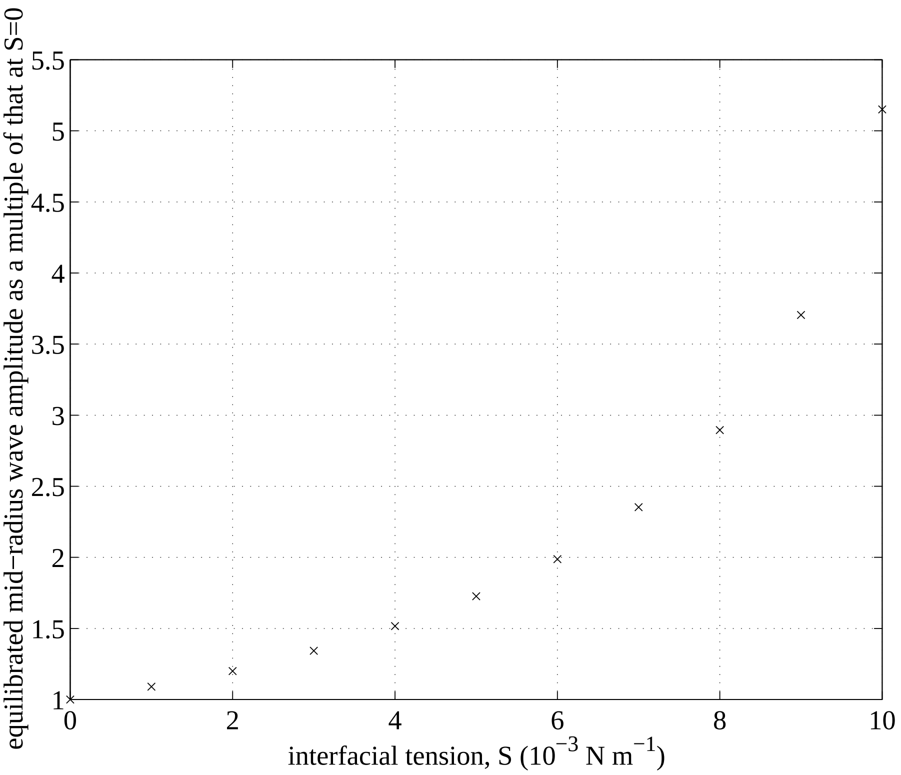


Figure 6.16: Variation of equilibrated mid-radius wave amplitude with interfacial tension, given as a multiple of its value when the tension is zero.

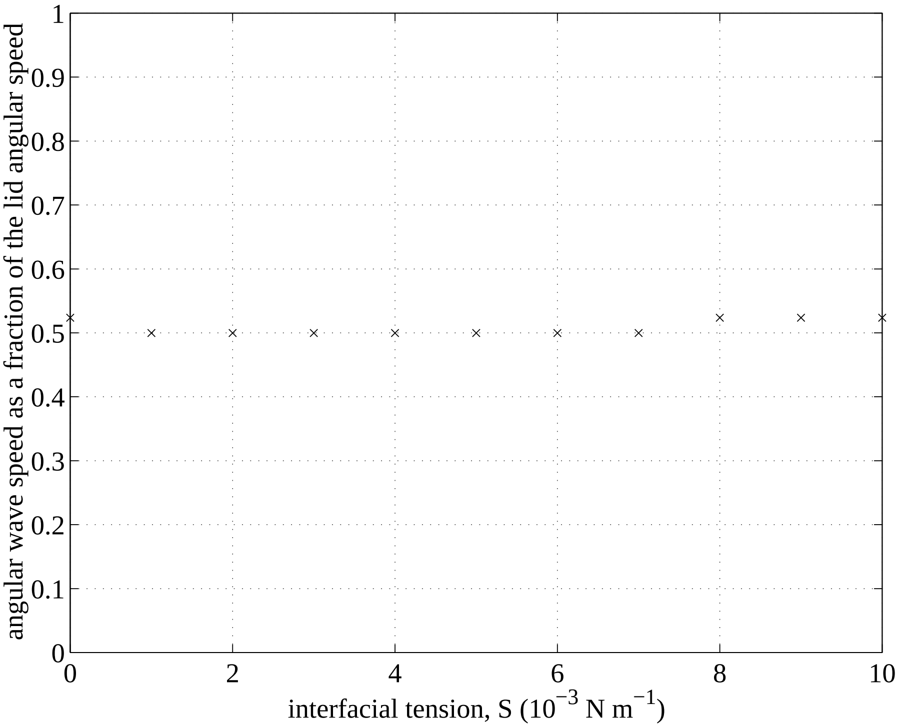


Figure 6.17: Variation of equilibrated wave angular phase speed with interfacial tension, given as a fraction of the angular lid speed.

tensions.

We have been unable to find detailed studies of the impact of interfacial tension on baroclinic instability in the geophysical literature, and so it is difficult to corroborate the above findings. The effects of interfacial tension on Kelvin-Helmholtz instability have been more widely studied, however. From equation (9.3) of Acheson (1990), the amplitude growth rate of a Kelvin-Helmholtz mode increases with interfacial tension, and the wave phase speed is independent of it. In this section, we have come to similar conclusions about waves due to baroclinic instability — equilibrated amplitudes grow with increased tension, but speeds are unaffected — and so the comparison with Kelvin-Helmholtz instability makes our findings plausible.

The model/laboratory intercomparison has been improved by increasing the model interfacial tension. The comparison is discussed in more detail in Section 6.5.

6.4 Model runs with non-zero interfacial tension

A second series of 210 model runs has been carried out, identical to the first series (Section 6.1) except that the interfacial tension is now set to be $S = 5.0 \times 10^{-3}$ N m^{-1}, close to the assumed laboratory value in the presence of a surfactant (Section 2.2.3). The resulting wavenumber regime diagram is shown in Figure 6.18. The 21 runs with $\Omega = 3.50$ rad s^{-1} all crashed due to an arithmetic exception before the 500th timestep, as the product FI was then 0.57 and the interfacial tension correction coefficient $C_{\mathrm{itcc}} = 1/(1-2FI)$ was negative, and so there are no points corresponding to these runs in the diagram. When Ω is 3.25 rad s^{-1}, 3.00 rad s^{-1} and 2.75 rad s^{-1}, the product FI is 0.49, 0.42 and 0.35 respectively. None of these is much smaller than unity, and so wavenumbers in the top three rows of points in the regime diagram are likely to be unreliable.

Neglecting these top three rows, there is good agreement between QUAGMIRE wavenumber transition curves both without and with small interfacial tension (Figures 6.9 and 6.18, respectively). We conclude, based on the present regime diagram and the analysis of Section 6.3, that small model interfacial tension has no impact upon equilibrated wavenumbers or wave speeds, but significantly increases wave amplitudes.

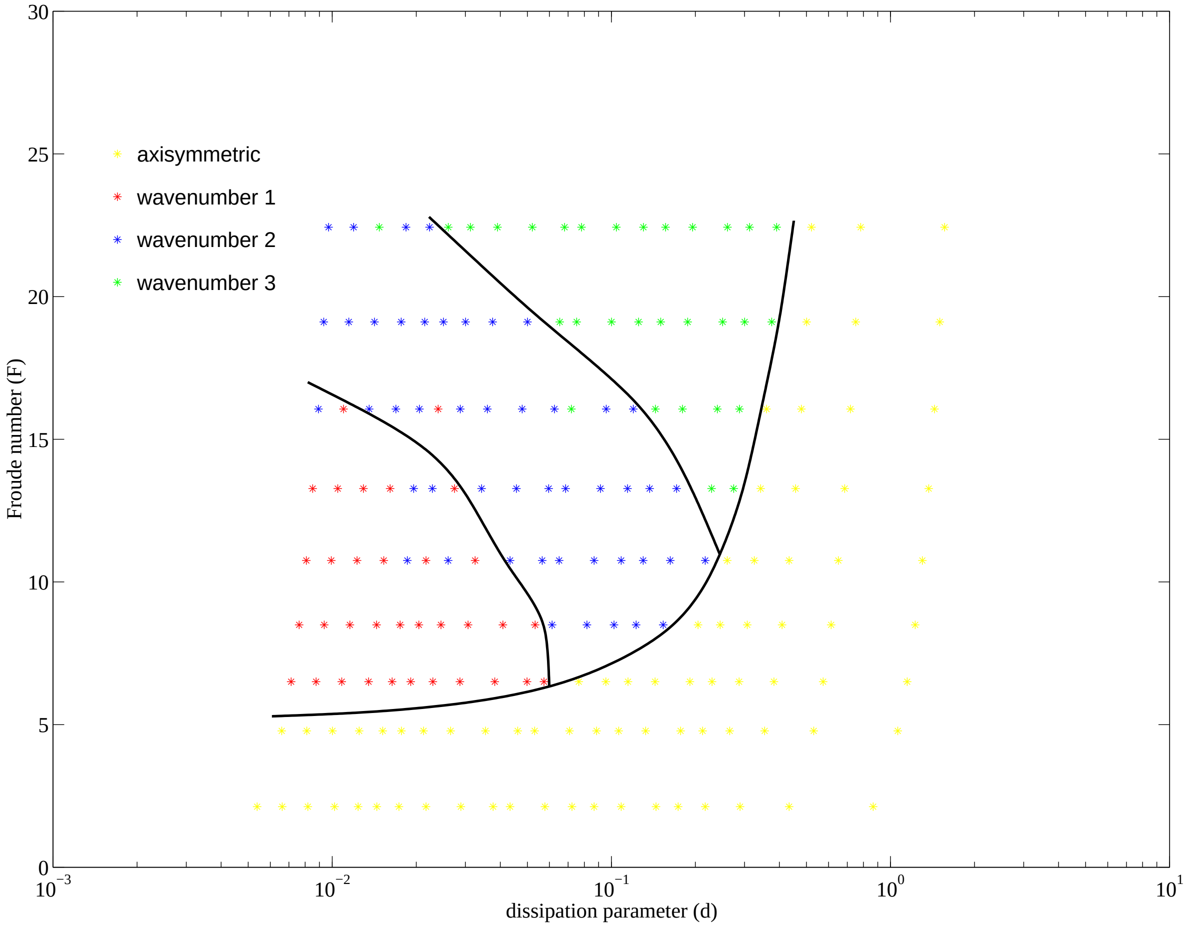


Figure 6.18: Model regime diagram with non-zero interfacial tension, showing equilibrated azimuthal wavenumber $m \in \{0, 1, 2, 3\}$ at each of the points investigated. Wavenumber transition curves have been inferred and over-plotted.

	laboratory	numerical model		
		$S=0$ $\chi=1$	$S=0$ $\chi=2$	$S=5$ mN m^{-1} $\chi=1$
***(d, F)* coordinates of triple points**				
m = 0, 1, 2	(0.01, 4.8)	(0.06, 6)	?	(+0%, +0%)
m = 0, 2, 3	(0.02, 5.5)	(0.25, 11)	?	(+0%, +0%)
mid-radius wave amplitude				
m = 1	25 mm	5 mm		
m = 2	8 mm	2 mm	+55%	+75%
m = 3	7 mm	1 mm		
angular wave speed	0.12 $\Delta\Omega$	0.50 $\Delta\Omega$	−20%	+0%

Table 6.1: Comparison of three important wave properties between the laboratory experiments and the QUAGMIRE runs with model parameters $S = 0$ and $\chi = 1$. The changes in the properties when χ is increased to 2 and S is increased to 5.0×10^{-3} N m^{-1} are also shown. The change in the triple point coordinates as χ is doubled has not been investigated, and accordingly these entries are labelled with "?".

6.5 Discussion of model/laboratory comparison

A summary of the findings of Sections 6.1.2, 6.2, 6.3 and 6.4, regarding the quantitative comparison between large-scale waves in the laboratory and the model, is given in Table 6.1. In the present section, we consider in turn each of the comparisons in the table, with the aim of investigating whether or not it is possible to confidently assign reasons for the discrepancies. In particular, laboratory/model differences could be due to (a) uncertainties in the fluid parameters; (b) unreasonable model assumptions *apart from* the neglect of fast modes; and (c) the neglect of fast modes in the model. Before we can attribute the discrepancies to reason (c), and thereby claim that we have found evidence of an observable fast wave impact upon the slow wave dynamics, we must first be able to reject the hypotheses that the discrepancies are due to reasons (a) and (b).

Model and laboratory Froude numbers agree to within a factor of up to 2, which seems reasonable given the many model approximations, including the small Rossby number assumption. The significant disagreement between dissipation parameters, by a factor of up to around 10, can probably be put down to uncertainties in layer viscosities (which appear in the dissipation parameter formula), and to the assumption of geostrophic parameterized Ekman velocities and other model approximations, including boundary conditions. Lovegrove (1997) constructed a spectral, channel model of the annulus, and came to very similar conclusions about the laboratory transition curve comparison. He notes that:

> "*... while the Froude numbers of experimental runs are of the same magnitude as those present in the theoretical regime diagram, the experimental values of the dissipation parameter are actually about an order of magnitude smaller than the predicted theoretical values.*"

He accounts for the discrepancy as being due to system differences, and without recourse to short waves.

Model wave amplitudes with zero interfacial tension are a factor of up to 7 times smaller than those measured in the laboratory. Realistic errors in the viscosity could increase the model amplitudes by one-half. Interfacial tensions smaller than the assumed laboratory value in the absence of a surfactant could almost double model wave amplitudes. Though we have been unable to run the model with a realistically large tension, it seems believable from Figure 6.16 that the measured tension could give the required amplitude amplification to account for this discrepancy.

Finally, the model waves travel around four times faster than those in the laboratory. This could partially be explained by uncertainties in the viscosities, to which model wave speeds are moderately sensitive: a realistic viscosity error could reduce model speeds by 20%. The dominant mechanism, though, is the neglect of Stewartson boundary layers in the model. We know that such boundary layers exist in the laboratory experiment, and we have seen in Section 3.4.3 that they exert a significant drag force on the layers, slowing down waves by up to an additional 40% (Table 3.2). Further allowing for the approximation in equation (3.24), and the model assumptions, is probably enough to explain this discrepancy.

Despite the model/laboratory discrepancies reported in the above paragraphs, many aspects of the comparison are positive. The model gives a realistic variety of selected wavenumbers, waves of reasonable shape and form, and vacillations with reasonable periods. We conclude that it seems likely that the discrepancies can be attributed to mechanisms (a) and (b). It follows that, since we are unable to reject these mechanisms, the discrepancies are not proof of an observable fast wave impact upon the balanced modes.

6.6 Radiation indicators

In the first part of this chapter, we have found reasonable agreement between the numerical and laboratory experiments, in terms of wave speeds, amplitudes and wavenumbers. It is therefore reasonable to assume that QUAGMIRE is also adequately simulating velocity fields. In the remainder of this chapter, we use the model velocity fields to compute five diagnostics of small-scale wave generation in the *MRW* regime. Some of the diagnostics are predictors of small-scale waves due to a shear instability mechanism, and others due to a nonlinear spontaneous emission mechanism. By investigating which of the five indicators best predicts the spatial locations of the laboratory small-scale waves, we will be able to infer which of the two generation mechanisms is responsible.

6.6.1 Indicator definitions

We now review, in turn, each of the five radiation indicators to be diagnosed using the model velocity fields.

Horizontal divergence, δ

To first order in the Rossby number, the velocity fields in any quasi-geostrophic model are horizontally non-divergent, permitting the introduction of a streamfunction. At higher order, though, there must be a small non-zero horizontal divergence $\delta = \boldsymbol{\nabla}_h . \boldsymbol{u}_h$ in order to allow the interface height to slowly evolve. In a velocity decomposition, vertical components correspond to balanced modes and divergent components to unbalanced modes, which suggests that δ may be a good indicator of small-scale wave generation.

From mass conservation $\boldsymbol{\nabla} . \boldsymbol{u} = 0$ for an incompressible fluid, we have

$$\delta = -\frac{\partial w}{\partial z} \; . \tag{6.1}$$

Integrating over the lower layer (including Ekman layers), at the top and bottom of which the vertical velocities are $(\mathrm{D}/\mathrm{D}t)_2 h$ and 0 respectively, gives

$$\delta_2 = -\frac{1}{h}\left(\frac{\partial}{\partial t} - \frac{1}{r}\frac{\partial \psi_2}{\partial \theta}\frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial \psi_2}{\partial r}\frac{\partial}{\partial \theta}\right) h \; , \tag{6.2}$$

with a similar expression for the upper layer. The horizontal divergence is a general indicator of short wave emission, i.e. it is not specific to either the shear or spontaneous emission mechanisms.

Local Richardson number, Ri

The Richardson number discussed in Section 4.4.1 is expected to be a good indicator of small-scale wave generation by a shear instability mechanism. In that section we derived an expression for the Richardson number in terms of the velocity fields. In the baroclinically-stable regime these velocity fields were simply those associated with solid-body rotation, for which we could write down analytical expressions. By doing this we were able to show that the Richardson number was a good indicator of the production of small-scale waves in the *KH* regime. We were unable to repeat the analysis for the small-scale waves in the *MRW* regime as we could not determine the velocity fields. However, these fields are now known from the model, enabling us to complete this avenue of inquiry.

The general definition of Richardson number for the annulus is

$$\mathrm{Ri} = \frac{2g'\sqrt{\nu/\Omega}}{(\Delta \boldsymbol{u})^2} \; , \tag{6.3}$$

where $(\Delta \boldsymbol{u})^2 = (\Delta u_r)^2 + (\Delta u_\theta)^2$ is the velocity shear across the interface, which can be diagnosed from the streamfunction. Equation (6.3) is a generalization of equation (4.4), which applies only to the axisymmetric equilibrium flow.

Brown indicator, ϕ, and turbulent energy dissipation rate, ϵ

A number of shear instability indicators have been developed as complementary alternatives to the Richardson number. Two of these are the Brown indicator ϕ and turbulent energy dissipation rate ϵ, first studied by Roach (1970) as indicators of Clear Air Turbulence (CAT) in the atmosphere. CAT occurs in cloudless conditions at altitudes of around 10 km, and is due to small-scale Kelvin-Helmholtz billows. It is occasionally severe enough to lift aeroplane passengers from their seats and cause injury or death (Roach & Bysouth, 2002), and so there are important practical reasons for developing a reliable indicator.

Roach begins his analysis by noting that, on the one hand, there are dynamical processes which tend to increase the vertical shear in horizontal velocity, e.g. thermal wind balance giving a tropospheric jetstream in the atmosphere, or the imposed differential lid rotation in the annulus. On the other hand, viscous energy dissipation due to small-scale waves tends to reduce the shear. Roach makes an assumption of approximate balance between these two competing effects on short timescales, leaving the shear (and Richardson number) constant.

Roach proceeds by imagining a thought experiment in the atmosphere in which the dissipation effect is switched off, destroying the balance and allowing an increase in shear. He argues that the rate at which the small-scale features were dissipating energy just before the switch-off must equal the rate of energy increase of the system just after, which is analytically derivable from the dynamical equations by setting the viscosity to zero. Using this approach, he calculates an energy dissipation rate of

$$\epsilon = \begin{cases} \dfrac{(\Delta \boldsymbol{u})^2}{24}\phi & : \quad \phi > 0 \\ 0 & : \quad \phi < 0 \end{cases} \tag{6.4}$$

where

$$\phi = -\frac{1}{\mathrm{Ri}}\frac{\mathrm{DRi}}{\mathrm{D}t} \; . \tag{6.5}$$

$\Delta \boldsymbol{u}$ is the vertial shear in horizontal velocity associated with the tropospheric jetstream, taken to be the shear across the interface in the annulus, and Ri is the Richardson number for the large-scale flow, defined for the annulus by equation (6.3). D/Dt is the total derivative operator calculated from equation (5.16). The indicator ϕ takes its name from Brown (1973), who derived an approximate form which was more practical for operational diagnosis of CAT, though in the present study we use the direct definition (6.5) involving Ri. In Brown's paper, both ϕ and ϵ are shown to be better indicators of CAT than Ri, and so we might expect the same to be true in the annulus.

Lighthill radiation term, LRT

Lighthill (1952) has presented a theory for the generation of sound waves by large-scale motions in a 3-D compressible adiabatic gas. The governing equations for Lighthill's system are isomorphic to the non-rotating shallow water equations, with a correspondence between acoustic and gravity modes, and so the problem of generation of pure gravity waves had also unintentionally been solved by Lighthill. Ford (1994) extended the theory to include rotation, and thereby derived an inertia-gravity wave radiation term. The generation mechanism in this case is an evolving vortical motion rather than a velocity shear, making this indicator fundamentally different from the previous three.

Ford's derivation begins by taking the f-plane inviscid barotropic shallow water equations, in flux form. Two equations, obtained by taking the curl and the divergence of the momentum equation, are combined to produce a single equation:

$$\left(\frac{\partial^2}{\partial t^2} + f^2 - gH\nabla^2\right)\frac{\partial h}{\partial t} = \frac{\partial}{\partial t}\boldsymbol{\nabla}.\boldsymbol{F} + f\boldsymbol{k}.\boldsymbol{\nabla}\times\boldsymbol{F} + \frac{g}{2}\frac{\partial}{\partial t}\nabla^2 h^2 \ , \tag{6.6}$$

where

$$\boldsymbol{F} = \boldsymbol{u}\boldsymbol{\nabla}.(h\boldsymbol{u}) + (h\boldsymbol{u}.\boldsymbol{\nabla})\boldsymbol{u} \tag{6.7}$$

and $\boldsymbol{k}$ is the unit vertical vector. The left side of equation (6.6) is the linear shallow-water inertia-gravity wave operator acting on $\partial h/\partial t$, which turns out to be a more convenient variable than h. The right side contains all of the nonlinear terms, which we refer to collectively as the Lighthill Radiation Term (LRT). The linear normal modes of equation (6.6) are shallow-water inertia-gravity waves, for which the intrinsic angular frequency ω and total wave vector K are related by dispersion equation:

$$\omega^2 = f^2 + gHK^2 \ . \tag{6.8}$$

The right side of Ford's original equation is written explicitly in Cartesian coordinates, but we retain the vector form here as we would like to diagnose LRT in the cylindrical geometry of the annulus. Ford goes on to derive an approximate form for LRT based on small Froude number, though Froude numbers are larger than unity for flows of interest in the present system, and so we use the unapproximated form given in equations (6.6) and (6.7).

Ford argues that inertia-gravity waves will be generated in any region for which LRT is non-zero, so that all vortical flows will emit freely-propagating inertia-gravity waves, disproving the existence of a strict slow manifold. The radiation mechanism is termed *spontaneous-adjustment emission radiation (SER)* by Ford et al. (2000). SER is a generalization of geostrophic adjustment radiation (GAR), as it includes GAR as a sub-class but does not necessarily take the flow towards a state of geostrophic balance.

It is important to note that Ford's theory is based on the shallow water equations, and so the expression for LRT given by equation (6.6) is an indicator of *shallow-water* inertia-gravity wave emission by an evolving *shallow-water* large-scale mode. Though the large-scale modes in the present system can reasonably be classified as shallow, it appears that the small-scale modes cannot. Their typical wavelengths (around 20 mm, from Section 4.4.2) are significantly smaller than the layer depth (125 mm), suggesting that the observed short waves are in the deep-water regime with a different dispersion relation from that above. Nevertheless, Lovegrove (1997) was able to demonstrate good agreement between short wave periods measured in the laboratory, and those predicted by the shallow water dispersion relation (6.8), suggesting that the boundary between the shallow and deep limits is determined by more than just the wave aspect ratio. We therefore speculate that Ford's theory is appropriate to the current system, even though the system seems to be in a regime which is formally outside the limits of the theory's applicability.

An approximation needs to be made before we can apply Ford's one-layer theory to the two-layer annulus. We approximate $\nabla p_1 = 3\nabla p_2$, which enables us to write the

indicator name	definition	mechanism	radiation condition
horizontal divergence	$\delta = \boldsymbol{\nabla}_h . \boldsymbol{u}_h$	not specific	$\delta \neq 0$
Richardson number	$\mathrm{Ri} \propto \frac{1}{(\Delta \boldsymbol{u})^2}$	KH	$\mathrm{Ri} < 1$
Brown indicator	$\phi = -\frac{\mathrm{D}}{\mathrm{D}t} \ln \mathrm{Ri}$	KH	ϕ large and positive
energy dissipation rate	$\epsilon \propto \phi (\Delta \boldsymbol{u})^2$	KH	$\epsilon > 0$
Lighthill radiation term	non-linear IGW terms	SER	$\lvert\mathrm{LRT}\rvert > 0$

Table 6.2: Summary of the five radiation indicators to be diagnosed using data from QUAGMIRE. The particular generation mechanism associated with each indicator is listed (Kelvin-Helmholtz shear *KH*, or Spontaneous Emission Radiation *SER*) together with the condition which needs to be satisfied in order for radiation to be expected.

horizontal pressure gradient force in the lower layer as $(1/\rho_2)\nabla p_2 = -(g'/2)\nabla h$. This assumption applies because equilibrium zonal velocities in the upper layer are three times those in the lower layer (for equal viscosities, from Section 5.3.1) and so three times the radial pressure gradient is required to support them. The implication is that equation (6.6) holds for the lower layer in the rotating annulus, so long as we replace g with $-g'/2$. All of the terms in the expression for LRT can be calculated from the QUAGMIRE output.

Table 6.2 gives a summary of the properties of the five radiation indicators discussed in this section, including the conditions under which short wave emission is expected.

6.6.2 Indicator plots using model data

Figure 6.19 shows plots of the five radiation indicators as calculated from the main series QUAGMIRE experiment ($S = 0$) with $\Delta\Omega = 0.15$ rad s^{-1} and $\Omega = 3.00$ rad s^{-1}, for

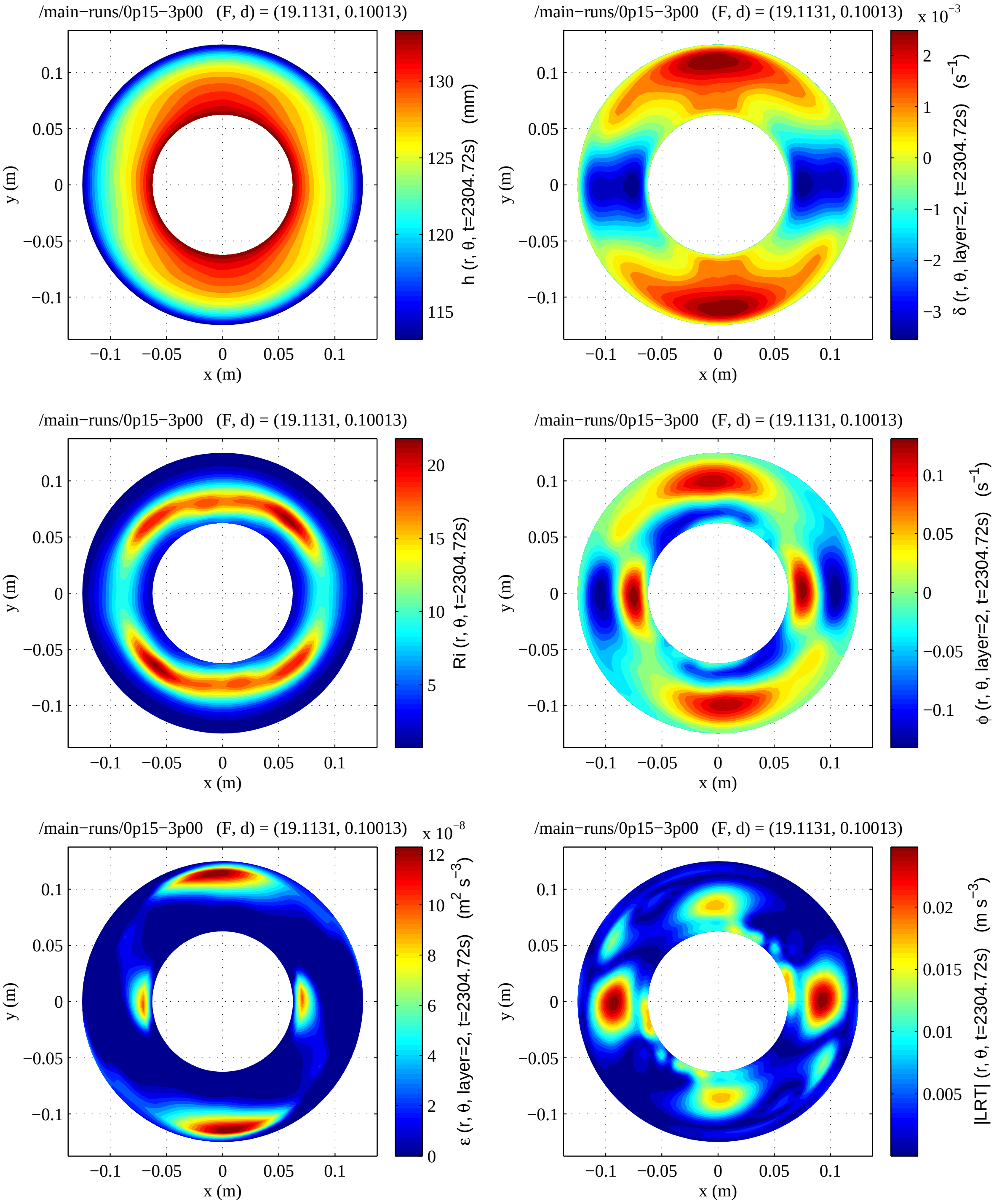


Figure 6.19: Plots of interface height (top left), lower layer horizontal velocity divergence (top right), Richardson number (middle left), lower layer Brown indicator (middle right), lower layer energy dissipation rate (bottom left) and lower layer Lighthill radiation term (bottom right) from a 2S QUAGMIRE simulation. The corresponding plots for the upper layer are similar.

which the equilibrated flow is 2S, denoting a steady (non-vacillating) wavenumber two flow in the notation of Section 4.1. The interface height field is also shown in the figure. If we observed this interface height field in the laboratory experiment, we would expect small-scale wave generation at low- to mid-radii near $\theta = 0°$ ("3 o'clock") and $\theta = 180°$ ("9 o'clock"), as can be seen by comparison with Figure 1.6 rotated through 90°. We are therefore interested in which (if any) of the five indicators would predict radiation at these (and only these) angular positions, based on the radiation criteria in Table 6.2.

The horizontal divergence indicator shows four large amplitude regions, but this includes two at which small-scale waves are not observed in the laboratory. The Richardson number shows local minima with respect to azimuth at the two expected regions, but is at its smallest close to the outer sidewall which is not a laboratory generation region. The Brown indicator has large positive maxima exactly where the short laboratory waves appear, but there are two equally large maxima elsewhere in the annulus. Similarly, the energy dissipation rate has two maxima too many to be a reliable indicator, and both of the unwanted maxima are larger than the maxima in the expected locations. The Lighthill radiation term has large global maxima at the two expected regions, and two weaker local maxima at other locations in the annulus.

Figure 6.20 shows a similar analysis for an equilibrated 1S (steady wavenumber one) flow with $\Delta\Omega = 0.50$ rad s^{-1} and $\Omega = 2.75$ rad s^{-1}. In this case, based on the laboratory experiments, we would expect small-scale radiation at $\theta = 90°$ ("12 o'clock").

The Lighthill radiation term formulae (6.6) and (6.7) contain terms with up to four derivatives in them, which amplify small-scale features relative to large-scale features. This heavy differentiation gives rise to the high level of noise present in the LRT plots.

The three Kelvin-Helmholtz instability indicators have each over-predicted regions of small-scale wave generation, in both the 2S and 1S cases. The Lighthill diagnostic, an indicator of spontaneous emission radiation, gives the best fit with the laboratory observations. It has large values exactly where the short waves appear in the laboratory. There are smaller subsidiary local maxima in other regions, but the values taken there are presumably not large enough for the laboratory short waves to overcome the effects

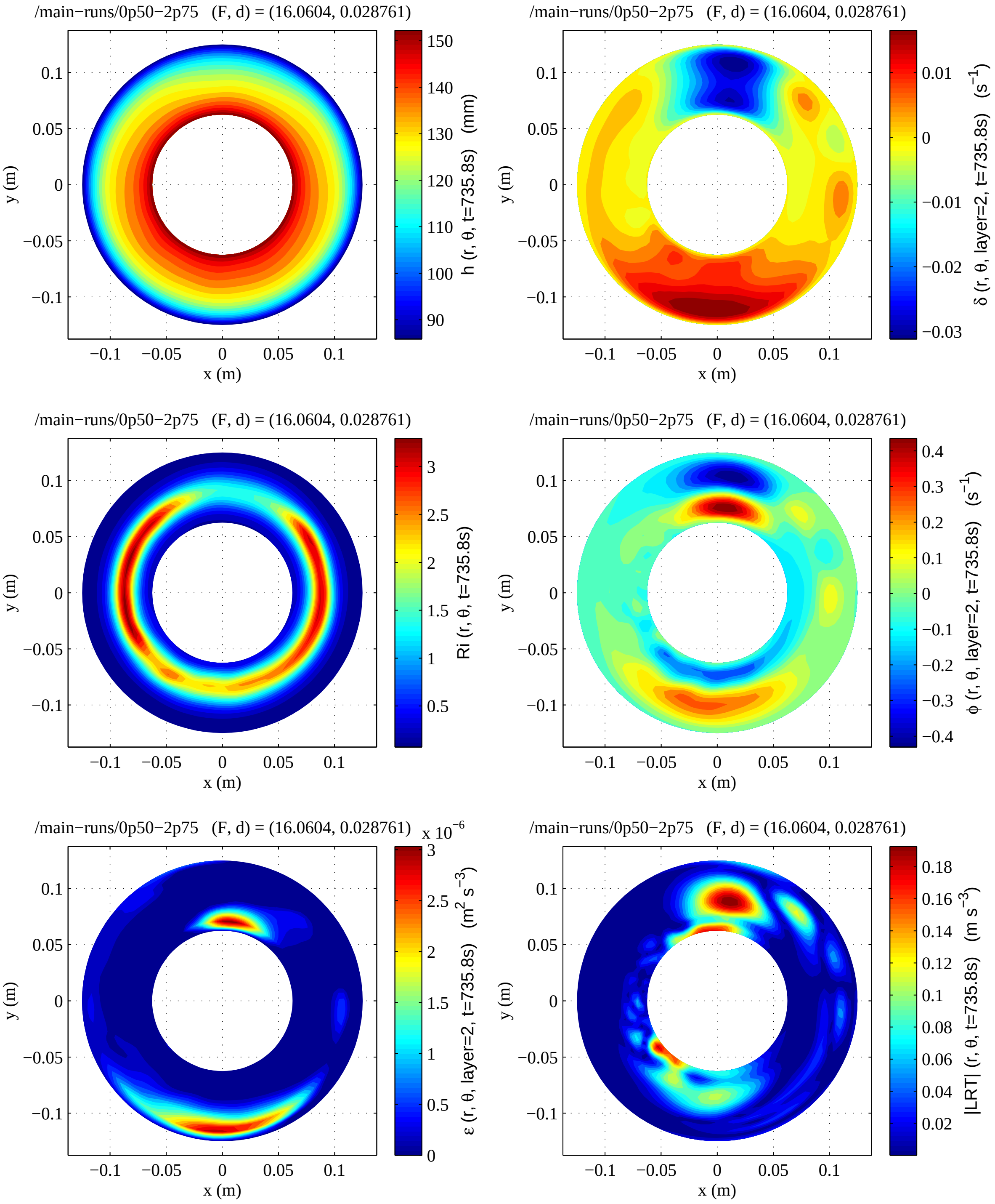


Figure 6.20: Plots of interface height (top left), lower layer horizontal velocity divergence (top right), Richardson number (middle left), lower layer Brown indicator (middle right), lower layer energy dissipation rate (bottom left) and lower layer Lighthill radiation term (bottom right) from a 1S QUAGMIRE simulation. The corresponding plots for the upper layer are similar.

of viscous dissipation — not included in Ford's theory — and grow to an amplitude which would make them visible. We conclude that the SER mechanism is likely to be responsible for the *MRW* regime small-scale wave generation in the laboratory.

As regards the baroclinically-stable flow regime, δ, ϕ and ϵ are each identically zero since the flow is then steady and axisymmetric. Most of the contributions to LRT are also zero, and those which aren't are very small due to the heavy differentiation, giving an LRT which is around 10^6 times smaller than in Figures 6.19 and 6.20 (not shown). Only Ri is non-negligible, reinforcing the conclusions of Section 4.4.1 that the laboratory short waves in the baroclinically-stable regime are generated by a shear instability.

6.7 Chapter summary

In this chapter we have investigated the results of simulations using the rotating, two layer annulus model described in Chapter 5. The basic model behaviour is the same as that seen in the laboratory, confirming that the model is reliable and that the code is free from errors. For example, the model displays baroclinic instability with a variety of realistic equilibrated wavenumbers for super-critical Froude numbers, and stability with relaxation back to an axisymmetric state otherwise.

We have shown how wave amplitudes, phase speeds and wavenumbers can be derived from the raw data produced by the model. These quantities are in reasonable agreement with measurements from the laboratory annulus. Specifically, we have found that it seems likely that all observed laboratory/model discrepancies can be attributed to fluid property errors and model approximations other than the neglect of fast modes. This means that we are able to state, based on the comparisons that have been carried out in this chapter, that we have found no evidence of an observable small-scale wave impact upon the large-scale balanced flow in the laboratory.

By diagnosing five candidate radiation indicators using the model velocity fields, we have been able to conclude that the observed short laboratory waves in the *MRW* regime (and presumably also the *MIW* regime) are best explained by the spontaneous emission radiation mechanism. This is because shear mechanisms predict short wave generation at regions other than those observed.

There were many differences between model and laboratory — apart from the neglect of the fast modes in the model — in the comparison described in this chapter. In the following chapter, we incorporate a stochastic inertia-gravity wave parameterization into the model. This allows us to run comparative simulations in which the *only* difference is the presence and absence of inertia-gravity waves, allowing a stronger test of their impacts than has been achieved here.

Chapter 7

A stochastic parameterization of the fast waves

"Démons et merveilles
Vents et marées
Au loin déjà la mer s'est retirée
Mais dans tes yeux entrouverts
Deux petites vagues sont restées
Démons et merveilles
Vents et marées
Deux petites vagues pour me noyer."

Sables Mouvants, Jacques Prévert.

In this chapter, we design and implement a simple parameterization of inertia-gravity waves in QUAGMIRE. Having first justified the need for a parameterization, we review previous deterministic and stochastic parameterization schemes, and give details of the chosen present scheme. We then describe numerical runs designed to measure the dependence of the equilibrated wavenumber, wave amplitude and phase speed of the large-scale waves on the amplitude of the parameterized inertia-gravity waves. We are particularly interested in investigating differences between runs with the parameterization switched on and off. The short wave parameterization proves to be fruitful, as we are able to identify regions of parameter space in which the parameterized waves exert a large influence on the balanced flow, in particular by forcing spontaneous transitions between regimes of different azimuthal wavenumber.

7.1 Parameterizations of inertia-gravity waves

Based on the comparison between laboratory experiments and numerical model runs in Chapter 6, we were able to draw the conclusion that no evidence had been found of an observable small-scale wave impact upon the large-scale flow. However, since the model differs from the laboratory experiments in more ways than just through the absence of small-scale waves, this conclusion was necessarily weak. Any discrepancy between laboratory and model could be put down to these differences, including uncertain fluid properties and model approximations, rather than to the filtering out of small scales.

We would like to design a stronger test of the scale-separated interaction. Ideally, we would like to run two laboratory experiments, one with and one without small-scale waves, but identical in all other ways. Unfortunately this is impossible, as the laboratory small-scale waves cannot easily be switched off at will, and so the next best thing is to include them in the numerical model. QUAGMIRE is a quasi-geostrophic model, and so by construction cannot explicitly capture the evolution of the short ageostrophic waves. It can, however, represent them implicitly by including a parameterization of their effects on the balanced flow. Such a parameterization for the two-layer annulus is described in this chapter, and model runs are compared both with the parameterization switched on and off.

Most conventional parameterization schemes are *deterministic*, i.e. they describe the effects of sub-gridscale processes by deterministic bulk formulae which depend upon local resolved scale variables and a number of adjustable parameters (Palmer, 2001). For example, a well-known deterministic parameterization is that for the momentum deposition due a continuous spectrum of gravity waves developed by Hines (1997). Recently, Piani & Norton (2003) have shown that the deterministic Hines parameterization, which has one adjustable parameter a, significantly underestimates the variability of the quasi-biennial oscillation in simulations using the UK Meteorological Office Unified Model. They have shown that a *stochastic* parameterization, in which a is allowed to vary randomly according to some chosen probability distribution, gives an increased variability and better agreement with observations.

Following Piani, we include a stochastic parameterization of small-scale waves in QUAG-

MIRE, by adding a random noise term to the right side of the prognostic model equations (5.23) and (5.24) for each layer. To do this, we must assume that the precise details and structure of the laboratory small-scale waves are irrelevant, and that they have the same impact on the balanced flow as would random noise.

As previously noted, QUAGMIRE cannot capture the evolution of the small-scale waves. However, we can reasonably expect it to capture the response of the balanced modes to PV anomalies induced by the small-scale modes. Inertia-gravity waves have zero PV anomaly only in the linear limit, and so any finite amplitude inertia-gravity waves will carry a non-zero PPV. It is this quantity which we parameterize in the model equations, as a stochastic perturbation to the PPV tendency fields. It was pointed out in Section 5.7 that the system state is completely specified by the PPV field. By perturbing the PPV tendency field with noise, therefore, we are effectively perturbing *all* of the dynamical fields, including the horizontal divergence field which we expect the laboratory inertia-gravity waves to perturb directly.

We choose the simplest possible form for the stochastic noise terms. At each gridpoint and at each timestep, a random number is drawn from the uniform distribution on the interval $[0, 1]$, and then shifted to the interval $[-\mathrm{amp}, \mathrm{amp}]$ before being used as an additive contribution to the PPV tendency as shown in Figure 5.3. The constant *amp* is a given amplitude with units s^{-2}, and is related to interface height wave amplitudes in a way to be determined in Section 7.2.2. The noise fields are chosen to be purely baroclinic, i.e. equal and opposite in both layers, as any increase in the depth of one layer due to an interfacial small-scale wave is matched by a corresponding reduction in the depth of the other layer. The discretized noise fields so defined contain no correlations in either time or horizontal position. An important difference between laboratory and model is that the parameterized short model waves are present throughout the entire annular domain, whereas the laboratory short waves are localized in space and time, appearing only where the Lighthill radiation term is large (Section 6.6.2). This strengthens the analogy between the model and the atmosphere, where inertia-gravity waves are more ubiquitous than in the laboratory annulus.

7.2 Model runs with the stochastic parameterization

In the following sections we show the results of some model runs with the stochastic terms switched on, for comparison with the runs described in Chapter 6.

7.2.1 Reproducibility of the equilibrated state

In this section, we investigate the possible role that small-scale features play in large-scale wavenumber selection. We use an experiment with $\Delta\Omega = 0.23$ rad s^{-1} and $\Omega = 2.25$ rad s^{-1}, which is quite close to the wavenumber $m = 1, 2$ transition curve. A 30-member ensemble was carried out for each of 21 values of the noise tendency amplitude parameter, ranging from 0 to 2.0 s^{-2} in steps of 0.1 s^{-2}. Within each ensemble, the only difference between the 30 members was the random numbers in the stochastic forcing fields. In each case, the equilibrated azimuthal wavenumber m was noted, and found to be either 1 or 2.

Typical post-transient model fields are shown in Figure 7.1, for a noise amplitude of 0.5 s^{-2}, giving an indication of the relative amplitudes of the large-scale wavenumber 2 mode and the small-scale stochastic noise. Since the model gridspacing is approximately equal to the wavelengths of the laboratory short waves (see Figures 1.6 and 5.4), there is a reasonable matching of lengthscales between laboratory and stochastic model short waves. At first sight, the plots of PPV in Figures 7.1(a) and (b) appear unrealistically noisy compared to the laboratory, but this is simply because the Laplacian operator — which amplifies small scales relative to large scales — is required to obtain the PPV from the streamfunction. For this reason, the plot of interface height in Figure 7.1(c) is much less noisy. We will show in Section 7.2.2 that there is a good matching of the amplitudes of interface perturbations between the laboratory and model short waves.

For each ensemble of constant noise amplitude, the probability of equilibration to wavenumber 2 was calculated and is plotted in Figure 7.2. There is a clear and strong dependence of probability partition on noise amplitude. The results are consistent with a linear drop-off in the probability of $m = 2$ as the noise increases to around 1.0 s^{-2}, followed by a

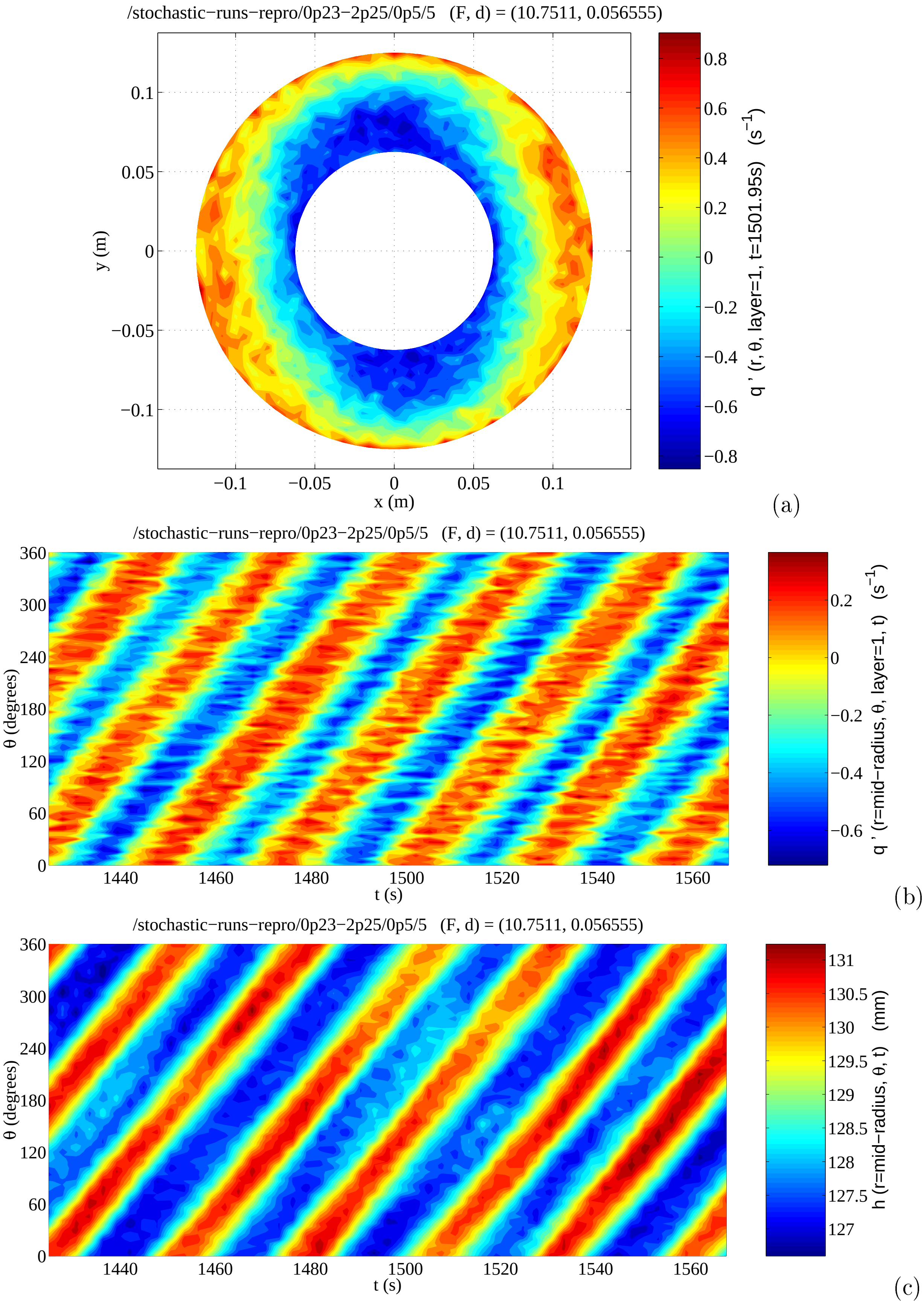


Figure 7.1: Model output with a stochastic PPV tendency term of amplitude 0.5 s^{-2}, showing (a) PPV as a function of radius and azimuth, (b) mid-radius PPV as a function of azimuth and time, and (c) mid-radius interface height as a function of azimuth and time.

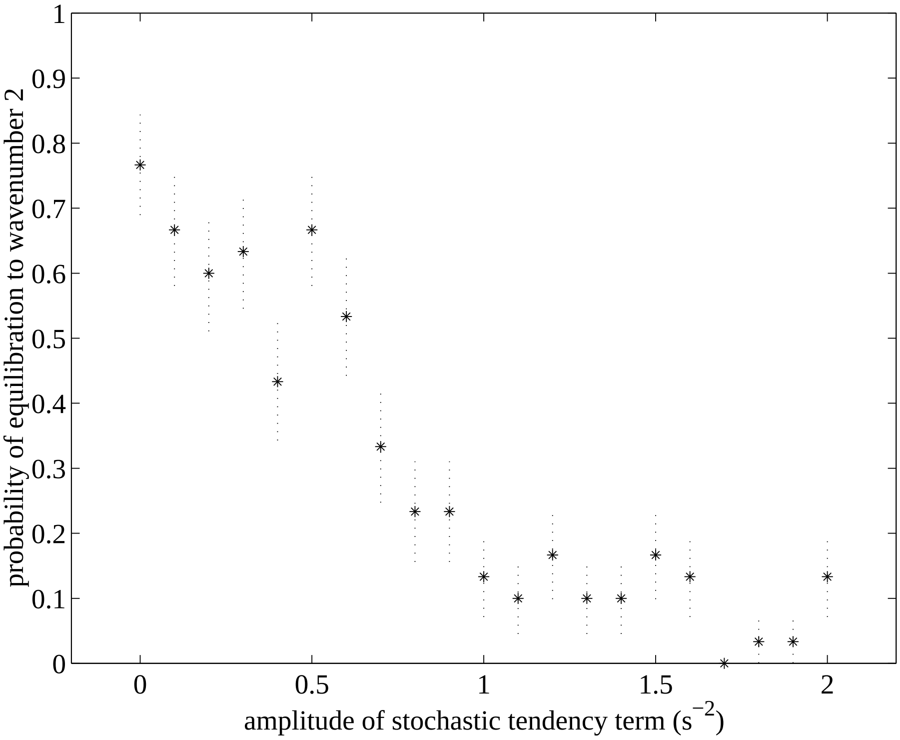


Figure 7.2: Probability of equilibration to azimuthal wavenumber two as a function of stochastic noise amplitude. The error bars correspond to plus and minus the standard deviation of the appropriate binomial distribution.

saturation at a probability of around 10% up to a noise of 2.0 s^{-2}.

We will see in Section 7.2.2 that, as suggested by the interface height plot in Figure 7.1(c), the interfacial wave amplitude for the noise is much smaller than for the large-scale mode. The addition of small-amplitude noise has therefore had a very significant impact upon the system's predictability, at this point in parameter space. An extensive investigation of the effects of the stochastic parameterization at other points in parameter space has not been performed, due to constraints on time and computational resources. 10-member ensembles at the centres of the $m = 1, 2, 3$ regions demonstrated 100% equilibration to the given wavenumber, irrespective of noise amplitude up to 2.0 s^{-2}. This suggests that the regions of parameter space in which inertia-gravity waves can exert a strong influence on the large-scale modes, are confined to finite width strips adjacent to transition curves.

The model wave phase speeds were measured and found to be completely unaffected by the introduction of noise up to 2.0 s^{-2} (not shown).

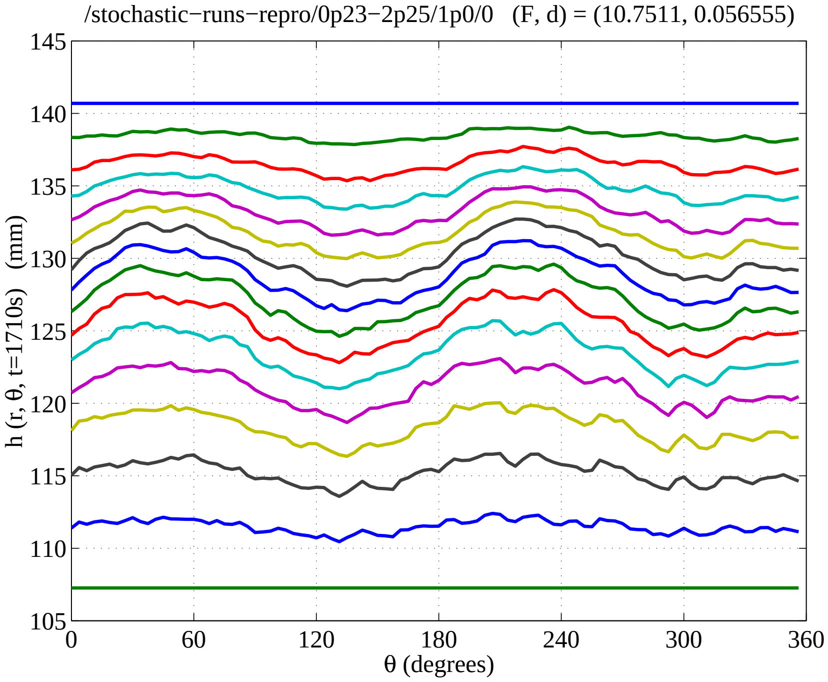
7.2.2 Diagnosis of amplitude of stochastic small-scale features

Figure 7.3(a) shows azimuthal interface height profiles after equilibration, from one of the ensemble runs described in the previous section with a noise amplitude of 1.0 s^{-2}. Figure 7.3(b) shows the same profiles after filtering by taking a running mean in the azimuthal direction with a window size of 8 gridpoints. This is sufficiently large to remove features at the scale of the stochastic forcing, and is sufficiently small to leave intact the large-scale features.

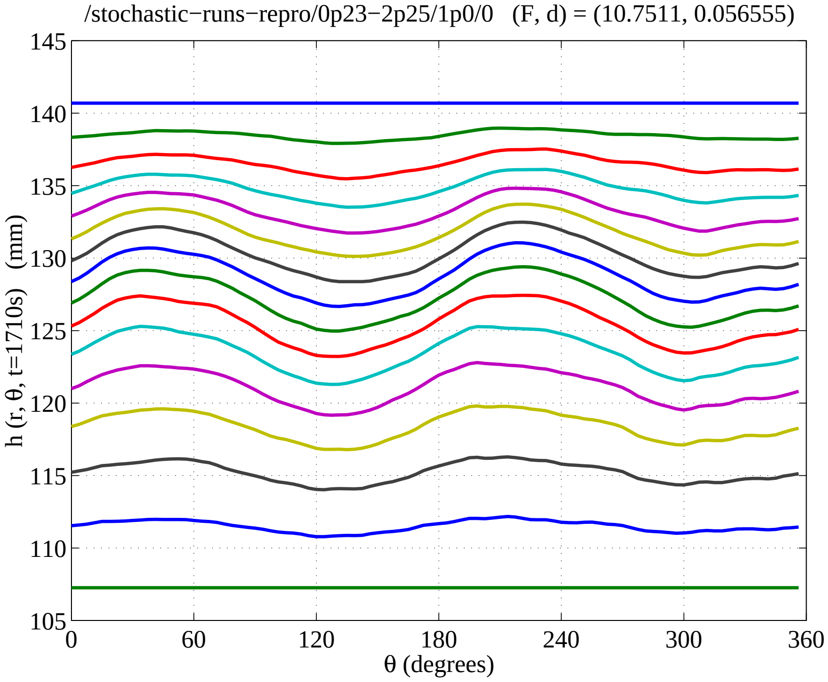
We can derive interfacial wave amplitudes (defined as half the difference between maximum and minimum displacements) from Figure 7.3. The amplitude so calculated from (b) is interpreted as the amplitude of the underlying large-scale wave, and that from (a) as the sum of the large-scale and small-scale wave amplitudes. By taking the difference, we can infer the amplitude of the stochastic small-scale waves. The results of this analysis at mid-radius, for each of the 21 noise amplitudes used and for both wavenumbers 1 and 2, are shown in Figure 7.4.

For a stochastic noise tendency amplitude of zero, the unfiltered and filtered amplitudes are almost identical, implying that the filtering has not modified the structure of the long modes. The amplitudes of the large-scale waves ("filtered" curves in the figure) increase significantly with stochastic noise amplitude. This is because the parameterized inertia-gravity waves have added energy to the system, which is expected because an interface height field containing short ripples has more gravitational potential energy than the same field with the ripples smoothed out. Importantly, the amplitudes of the small-scale stochastic features ("unfiltered–filtered" curves) are consistent between the large-scale wavenumber 1 and 2 cases, validating the analysis and giving us a direct linear correspondence between the stochastic noise tendency parameter in the model (in s^{-2}) and the corresponding interfacial amplitude of the stochastic small-scale features (in mm).

It is clear from Figure 7.4 that the stochastic small-scale features which caused the re-partitioning of the probability distribution in Figure 7.2 were many times smaller in amplitude than the large-scale wave with which they coexisted. Their typical amplitudes are similar to those observed in the laboratory.



(a)



(b)

Figure 7.3: (a) Interface height profiles after equilibration, showing an azimuthal wavenumber 2 mode with superimposed small-scale noise representing inertia-gravity waves. (b) Same profiles but with inertia-gravity waves filtered out. The legend relating colour to radius (not shown here to allow the full profiles to be seen) is identical to that in Figure 6.3.

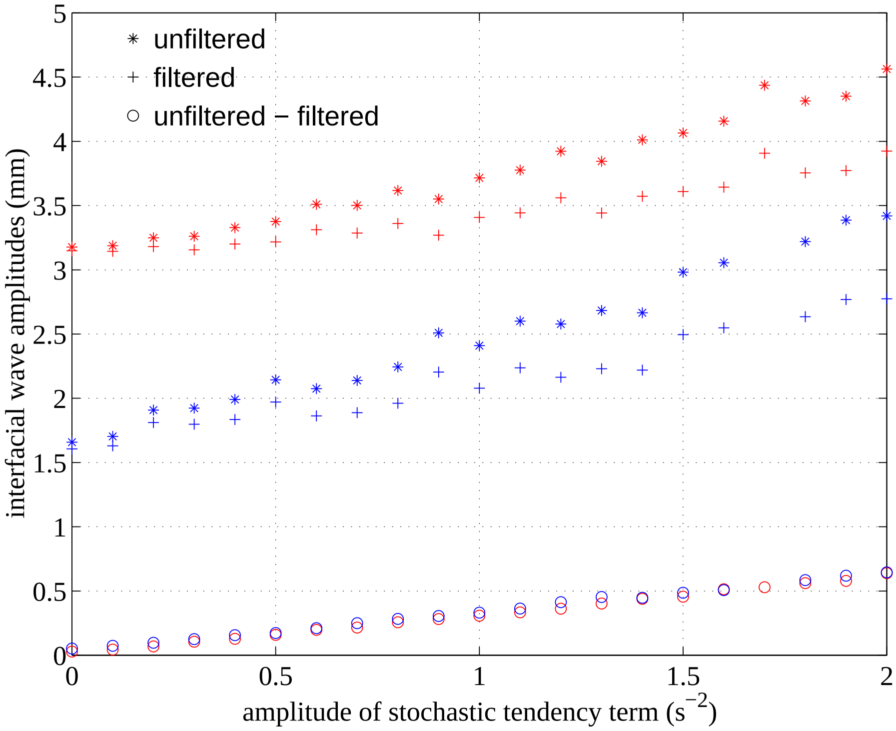


Figure 7.4: Unfiltered, filtered and unfiltered–filtered interfacial wave amplitudes as a function of stochastic noise tendency amplitude. Red points correspond to wavenumber 1, and blue to wavenumber 2. Each of the 30 ensemble members for a noise tendency amplitude of 1.7 s^{-2} equilibrated to wavenumber 1, and so there is no data corresponding to wavenumber 2 at this noise amplitude.

7.2.3 Continuous variation of amplitude of stochastic small-scale features

In Section 7.2.1 we investigated the stability of an equilibrium axisymmetric annulus flow continuously seeded with stochastic noise, which is in many ways a simple model of an axisymmetric atmospheric jetstream in the presence of inertia-gravity waves. A more likely scenario in the atmosphere is for a particular large-scale azimuthal mode to have already equilibrated, and so there are good geophysical reasons to be more interested in the stability of an equilibrated large-scale wave in the presence of inertia-gravity waves, rather than the stability of an axisymmetric jetstream upon which a large-scale wave is soon to grow in the presence of inertia-gravity waves.

In order to investigate this, we now take a wavenumber 2 flow with $\Delta\Omega = 0.23$ rad s^{-1} and $\Omega = 2.25$ rad s^{-1}, which has equilibrated at finite amplitude in the absence of stochastic forcing. As with the investigation of Section 7.2.1, which also used these parameters, the system is quite close to the wavenumber $m = 1, 2$ transition curve. In the present investigation, we continue the model integrations but increase the stochastic noise amplitude from 0 to 2.0 s^{-2}, by 10^{-6} s^{-2} each timestep so that the increase is quasi-continuous.

When this numerical experiment is performed, the wavenumber 2 mode persists until the noise reaches a certain threshold level, at which point a spontaneous transition is observed to a wavenumber 1 mode. A Hovmüller diagram showing the transition, which takes place over around 100 s, or the time taken for the large-scale wave to travel around the annulus twice, is shown in Figure 7.5. This kind of transition was never observed without the inertia-gravity wave parameterization switched on, and so we can conclude that the transition was caused by the parameterization. At the time of the transition, the stochastic noise parameter had reached a value of 1.1 s^{-2}, corresponding from Figure 7.4 to an interface perturbation of amplitude of 0.3 mm.

After the transition to wavenumber 1, the stochastic noise amplitude was decreased back to zero by 10^{-6} s^{-2} each timestep, but the reverse transition back to wavenumber 2 did

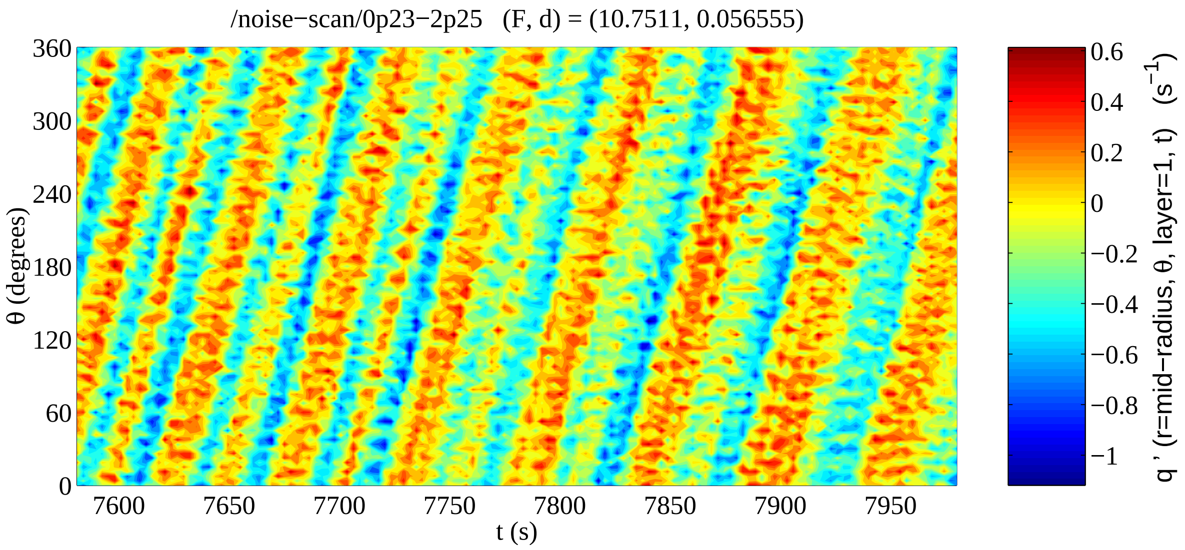


Figure 7.5: Hovmüller diagram, showing a mid-radius azimuth-time contour plot of PPV in the upper layer around the time of a spontaneous wavenumber transition.

not occur. At the end of the integration, when the noise had reached zero, the wavenumber 1 mode was still dominant, indicating the presence of hysteresis in the system. If these conclusions are portable to the atmosphere (Chapter 8) then the implication is that a short but sufficiently intense burst of small-amplitude inertia-gravity waves could force a large-scale regime change which could persist long after the inertia-gravity waves have been dissipated away.

This result suggests a simple schematic model for explaining the spontaneous transitions, in which the stable equilibrium states $m = 1$ and $m = 2$ are represented by minima of the potential well shown in Figure 7.6. With the system in the $m = 2$ state, a short burst of sufficiently large amplitude stochastic forcing permits the system to overcome the transition barrier and thereby undergo an irreversible transition to the $m = 1$ state, in which the system will remain after the end of the burst.

As in Section 7.2.1, when the above experiment was repeated with parameters corresponding to the centre of a wavenumber regime in parameter space, spontaneous transitions were not observed.

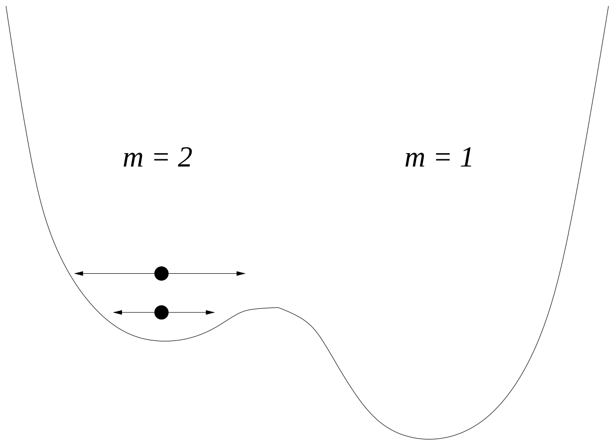


Figure 7.6: Schematic double-well potential for a bistable system, which can explain the observed model regime transitions close to the $m = 1, 2$ transition curve in (d, F) parameter space.

7.3 Discussion

We have seen that the addition of small-amplitude noise has had a very significant impact on the system. This phenomenon is a form of *stochastic resonance* (Pikovsky et al., 2001). This is a nonlinear resonance which is not dependent upon any matching of timescales, as is the familiar criterion required for linear resonance. The phenomenon allows a small (stochastic) forcing to produce a large (resonant) response. If stochastic resonance is exhibited by a nonlinear system, then the introduction of very small amplitude noise can dramatically affect the system state.

Stochastic resonance has been observed before in fluid systems. De Swart & Grasman (1987) have studied the effects of adding a stochastic forcing term to a low-order atmospheric spectral model based on the barotropic potential vorticity equation, and found that the noise forces the system to alternately visit different regimes due to a stochastic resonance. The phenomenon is widely observed across the entire spectrum of the natural sciences. For example, the human eye can detect signals otherwise too faint to be seen if random noise is added to the field of vision (Hogan, 2003). And, as discussed in Section 1.2, Chua's electronic circuit displays an altered temporal regularity upon the introduction of small amplitude noise.

The addition of noise terms to the governing model equations has led to a better agreement between one of the the laboratory/model comparisons. We found in Section 6.1.2

that model wave amplitudes were significantly lower their laboratory counterparts. However, we see from Figure 7.4 that the inclusion in the model of parameterized short waves of less than 1 mm in amplitude increases the long wave amplitudes by up to 60%, giving a better fit with the laboratory results. This amplitude increase is due to energy from the short modes filtering upscale into the long modes.

Our findings regarding stochastic resonance mirror an observation we made in the laboratory. In the current chapter, we found that the presence of small-amplitude, fast waves could increase the likelihood of a model state transition. Correspondingly, we found in the preliminary laboratory experiments of Section 4.2 that the annulus with a complete absence of fast waves exhibited a strong reluctance to undergo a transition away from the wavenumber 2 mode. We are not in a position at the moment to be able to fully attribute this reluctance to the absence of fast modes, as there were also unknown changes in the fluid properties which could have been responsible. However, we can state with certainty that the short waves do appear to have an influence on wavenumber transitions, in both the laboratory and the model.

7.4 Chapter summary

By implementing a simple stochastic inertia-gravity wave parameterization in the numerical model, we have shown that short modes can play a crucial role in large mode wavenumber selection. This finding seems to apply only to regions of parameter space which are some finite distance away from a wavenumber transition curve. These short modes also significantly increase the long wave amplitudes, but leave their propagation speeds unaltered.

Via the phenomenon of stochastic resonance, the stochastic inertia-gravity wave parameterization has the ability to induce spontaneous azimuthal wavenumber transitions which would not occur if the inertia-gravity waves were absent.

Chapter 8

Conclusions and future work

> *"Habe nun, ach! Philosophie,*
> *Juristerei und Medizin,*
> *Und leider auch Theologie*
> *Durchaus studiert, mit heißem Bemühn.*
> *Da steh' ich nun, ich armer Tor,*
> *Und bin so klug als wie zuvor."*
>
> *Faust*, J. W. von Goethe.

8.1 Summary of findings

The first part of this thesis focused on the results of laboratory experiments using a rotating, two-layer annulus, in which relative motion was forced between the two isothermal and immiscible layers by a differentially-rotating lid in contact with the upper layer. The natural interfacial tension between the liquids was reduced by the addition of a surfactant. Based on direct observations of the fluid interface, and on an initial inability to reproduce the results of previous experimental studies, we speculated that the physical properties of the liquids were exhibiting slow changes with time.

Because the lower-layer liquid was optically active and the apparatus was seen through crossed polaroids, a video camera viewing the fluids from above registered colours which were related to the depth of the lower layer. The relationship between hue and interface height was quantified by deriving a calibration curve, based on a torque balance calculation which gives an analytical expression for the equilibrium interface height field in the absence of any instability. An upgrade of the flow visualization hardware, together with the implemented calibration scheme, allowed two-dimensional maps of interface height to be inferred with a horizontal resolution of up to 0.2 mm, a vertical resolution of up to 1 mm and a temporal resolution of 1/25 s. The vertical resolution could be further improved as long as a reduction in the horizontal or temporal resolution could be tolerated.

Four series of new laboratory experiments were performed: one with prograde differential rotation in an annulus with increasing Froude number (PAI); one with retrograde differential rotation in an annulus with increasing Froude number (RAI); one with prograde differential rotation in an annulus with decreasing Froude number (PAD); and one with prograde differential rotation in an eccentric annulus with increasing Froude number (PEI). As predicted by standard theory, the motions observed in the fluids fell into two distinct categories and were robust to the changes in experimental configuration. The short, fast waves had wavelengths of around 20 mm and interfacial amplitudes of around 3 mm, and the long, slow waves had wavelengths of around 200 mm and interfacial amplitudes of up to 25 mm.

Each flow observed contained either no waves at all (the axisymmetric flow regime, AX), short waves only (the Kelvin-Helmholtz regime, *KH*), or both long and short waves coexisting (the mixed regular and irregular wave regimes, *MRW* and *MIR*). Flows containing long waves only, with a complete absence of short waves, were observed only in preliminary experiments with fresh preparations of the working liquids, and were not investigated in detail in this thesis. These experimental results are in concordance with the assertion by Ford et al. (2000) that every single evolving vortical flow emits inertia-gravity waves, which in our case are large enough to be visible if the fluid properties are permitting.

The mechanism which gives rise to the long waves is well understood from previous studies to be baroclinic instability. The mechanism which gives rise to the short waves *when they develop in the absence of long waves* was shown in this study to be consistent with a Kelvin-Helmholtz instability based on a critical Richardson number. Such a simple analysis was not possible for the case when the short waves develop in the presence of long waves, as the long waves perturb the layer velocity fields to an extent which cannot be predicted by simple linear theory.

In order to assess the mechanism by which the laboratory short waves are generated in the presence of a large-scale mode, and to examine the feedback impact of the short waves on the long waves, a quasi-geostrophic numerical model of the laboratory experiment was developed in the second part of the thesis and named QUAGMIRE. Short waves are permitted in the laboratory experiment but not the model. There was found to be excellent agreement, regarding the shapes of azimuthal wavenumber regimes in the system parameter space, between numerical and laboratory experiments. Quantitative agreement was not perfect but, due to model approximations and suspected uncertainties in assumed laboratory fluid properties, this was not thought to be due to the presence of short waves in the laboratory.

The model velocity fields were used to address the question of the generation mechanism of the short waves in the presence of long waves. Kelvin-Helmholtz instability theory, which successfully predicted the generation of laboratory short waves in the absence of long waves, could not explain the coexisting short waves, and neither could three other indicators of shear instability. The best predictor was found to be the Lighthill radiation term, which is an indicator of spontaneous emission radiation which could take place even in a purely barotropic fluid with no vertical shear.

Finally, we incorporated a simple stochastic parameterization of the short waves into the numerical model. In general, the effect of the parameterized short waves on the long waves was limited to an increase in the long wave amplitude. Sufficiently close to a wavenumber transition curve, however, a stochastic resonance effect allowed the short waves to exert a dominant influence over long mode wavenumber selection. In particular, spontaneous transitions were observed between different azimuthal modes, which were directly attributable to the presence of the stochastic short waves. This finding supported a similar observation we made in the laboratory, in which a flow devoid of short waves displayed a reluctance to undergo state transitions which occurred if the short waves were present.

8.2 Conclusions

We are now in a position to return to the four questions posed in Section 1.7.1, and to give answers based on the investigations of this thesis.

Under what circumstances do small-scale waves appear in the laboratory experiments?

In the laboratory investigations of Chapter 4, we found that if the fluids were baroclinically-stable, short waves were globally generated whenever the Richardson number dropped below a critical value. Short waves were locally generated in every single observed baroclinically-unstable flow with aged fluids, superimposed in the nodal regions of the long baroclinic mode. The short wave amplitudes were larger if the long mode was undergoing an amplitude vacillation, but were still generally present with reduced amplitudes (sometimes barely visible in stills but clearly present in the video footage) even when the long mode amplitude remained constant.

Which mechanism causes the small-scale waves to appear in the laboratory experiments?

Two different generation mechanisms are both responsible for short wave emissions in the laboratory experiments, though the circumstances under which they are responsible differ between the mechanisms. As shown in the laboratory investigations of Chapter 4, a Kelvin-Helmholtz shear instability is responsible for small-scale wave generation in the absence of long waves. When the short waves appear locally in the nodes of long waves, the generation mechanism was shown using the numerical model in Chapter 6 to be spontaneous emission by the evolving large-scale flow. This conclusion is further supported by the ship wake analysis of Chapter 4, which is an alternative way of analyzing emission of short waves by a moving "object" long wave. Further independent corroboration for this conclusion comes from the eccentric annulus laboratory experiments of Chapter 4, in which an azimuthally-varying velocity shear was not associated with azimuthally-varying short wave emission.

What are the effects of the laboratory small-scale waves on the large-scale, balanced flow?

A comparison of the laboratory regime diagram of Chapter 4, derived from experiments which included short waves, and the numerical regime diagram of Chapter 6, derived from experiments which did not include short waves, leads us to conclude that the laboratory short waves do not have a dominant impact upon the large-scale flow, in general. Though we found significant discrepancies between the numerical and laboratory results, in terms of equilibrated wavenumbers, amplitudes and phase speeds, it was felt that these differences could be explained by other factors such as model approximations and uncertain fluid properties.

For example, laboratory wave speeds are around a factor of four smaller than model wave speeds. Though it might be tempting to partially attribute this to inertia-gravity wave drag on the balanced flow, the discrepancy is adequately explained by Stewartson layer drag and viscosity uncertainty. This is not to say that we have found *evidence of absence* of a fast wave impact upon the slow modes, but rather that this particular test has given an *absence of evidence.*

However, the more explicit tests described in Chapter 7, based on model runs with a stochastic inertia-gravity wave parameterization switched on, did find evidence of an impact. The results showed that, sufficiently close to a regime transition curve, short waves play a key role in wavenumber selection, and can force spontaneous long wave transitions which would otherwise not occur. The preliminary laboratory experiments of Chapter 4, based on a comparison between flows with and without small-scale waves, give further independent evidence to corroborate this conclusion.

Having answered these questions for a laboratory experiment, what can we infer about answers to the analogous questions for geofluids?

If the laboratory annulus system and the atmosphere on a rotating planet were exactly dynamically and geometrically similar, then the fluid flows would also be mathematically similar (Section 1.5) and our conclusions about the laboratory system would be portable to the atmosphere.

Geometrical similarity is limited by differing horizontal/vertical aspect ratios; by the presence of annular sidewall boundaries and a rigid lid in the annulus which have no counterpart in the atmosphere; by the discretization of the fluid in the annulus to two discrete homogeneous layers, rather than a single continuously-stratified layer; and by the absence of bottom topography in the annulus. Additionally, the atmosphere is forced differently and is coupled to other components of the climate system which are not represented in the annulus, which causes the flow in the atmosphere to be generally more irregular and chaotic.

Mindful that geometrical similarity does not hold exactly, which will limit any comparison, we can determine the extent to which dynamical similarity holds by evaluating the non-dimensional system parameters for an approximated two-layer atmosphere. Referring back to Figure 1.5, we take the annular gap width L of the "atmospheric annulus" to be a quarter of the circumference of the Earth, and H to be the scale height of around 10 km. We take the reduced gravity g' to be 2 m s^{-2} in value[1], and the kinematic viscosity ν to be the turbulent eddy value of 5 m^2 s^{-1}. We take the background rotation rate Ω to be the local component of the Earth's rotation vector at latitude 45°, and the differential rotation rate $\Delta\Omega$ to be the zonal tropospheric jetstream speed — typically 40 m s^{-1}, from Figure 1 of O'Sullivan & Dunkerton (1995) — divided by the radius of the 45° latitude circle.

A comparison between non-dimensional parameters in the laboratory experiment and in the two-layer annulus approximation to the atmosphere is shown in Table 8.1. The atmosphere is seen to explore a significantly different regime from that explored in the laboratory experiments. The atmospheric Froude number and dissipation parameter are both larger, corresponding to the fact that the atmosphere typically exhibits higher wavenumber states, and more irregularity, than the laboratory annulus. The Rossby numbers are similar. Viscous effects are much more significant in the laboratory than in the the atmosphere, though they are still relatively weak in both systems compared

---

[1] We could naïvely compute the reduced gravity using $g' = g\Delta\rho/\overline{\rho}$, where the densities at the surface and the scale height are $\rho_2 = 1$ kg m^{-3} and $\rho_1 = \mathrm{e}^{-1}$ kg m^{-3}, respectively. Most of this density difference is due to static compressibility rather than static stability, however, giving an over-estimate of g'. It is more appropriate to use $g' = g\Delta\theta/\overline{\theta}$, where θ is potential temperature. We use $\overline{\theta} = 300$ K and $\Delta\theta = 60$ K, from Figure 1 of O'Sullivan & Dunkerton (1995), to obtain the quoted value for the reduced gravity.

	laboratory	atmosphere
$F = \dfrac{4\Omega^2 L^2}{g' H}$	9	50
$d = \dfrac{\sqrt{\nu\Omega}}{H\Delta\Omega}$	0.02	0.2
$\mathrm{Ro} = \dfrac{\Delta\Omega}{2\Omega}$	0.1	0.09
$\mathrm{Re} = \dfrac{L^2\Delta\Omega}{\nu}$	2×10^3	2×10^8
$\mathrm{Ek} = \dfrac{\nu}{2\Omega L^2}$	6×10^{-5}	5×10^{-10}
$I = \dfrac{S}{g(\rho_2 - \rho_1)L^2}$	0.1	0

Table 8.1: A comparison between typical values of the Froude number, dissipation parameter, Rossby number, Reynolds number, Ekman number and interfacial tension number in the laboratory experiment and a two-layer annulus approximation to the atmosphere.

to other dynamical effects, as demonstrated by the large Reynolds and small Ekman numbers. Interfacial tension effects, completely absent in the atmosphere, are also reasonably significant in the laboratory, as indicated by the size of the interfacial tension number.

The results of this thesis are not expected to depend upon whether the observed zonal wavenumbers are 1, 2 or 3 such as in the laboratory, or around 6 as in the atmosphere. Furthermore, though the non-dimensionalized viscosity and interfacial tension are quite different for the two systems, these effects remain small compared to other effects in both cases. Therefore we expect these other dynamical effects — which are similar for both systems — to be the ones which determine the system characteristics.

An explicit comparison between inertia-gravity wave properties in the laboratory and the atmosphere is also possible. In both cases, the amplitudes are generally around an order of magnitude smaller than that of the main, large-scale mode. Atmospheric pure gravity waves are sufficiently short in wavelength to be in the deep regime, like the laboratory short waves, though larger wavelength inertia-gravity and pure inertial waves in the atmosphere are in the shallow regime. In terms of the comparison of short wave impacts, this difference is unlikely to alter our conclusions. If anything, the impact of a short wave in a shallow fluid would be expected to be greater than the impact of a short wave in a deep fluid, because in the latter case the region of dynamical influence is vertically-confined. This suggests that the laboratory short waves, which are in a deep fluid, are not able to exert as great an influence on the balanced flow as are short waves in a shallow atmosphere. This means that, if anything, our laboratory investigation may have underestimated the strength of the atmospheric interaction.

Based on the above comparisons, there is every reason to suspect that the conclusions of this study regarding rotating laboratory experiments, will have counterparts regarding flows in the atmosphere. For example, the stochastic resonance phenomenon discussed here would also be expected to be observed in an atmospheric general circulation model. The implication is that, in a region of the atmosphere which is simultaneously unstable to two different modes with approximately equal growth rates, a local burst of inertia-gravity wave activity could determine which mode grows to equilibration, or could alternatively force a spontaneous transition from one mode to the other. The mechanism which permits such behaviour is not captured by any weather forecasting model which does not include inertia-gravity modes, either explicitly or through a stochastic parameterization.

8.3 Future work

We conclude the thesis by giving some possible avenues for future work suggested by the results.

Experimental work

In addition to the PAI, RAI, PAD and PEI experimental runs described in Chapter 4, there is a further configuration which warrants investigation. The inner cylinder could be removed, so that the fluid occupies a cylindrical domain rather than an annular one. This change would make the apparatus very similar to that used by Hart (1972). Losing the inner sidewall boundary has the advantage that the geometrical similarity between laboratory and atmosphere is stronger. Also, in practical terms, it has the benefit of removing the parallax effect which blocks from view the short wave generation region. Experimental runs in an open cylinder would allow an investigation of the role of the inner sidewall boundary in locally enhancing the vertical shear across the interface. QUAGMIRE runs could also be done for this configuration, as long as we insert a model inner sidewall of very small radius to avoid the singularity in the model equations on the rotation axis.

Numerical modelling work

A possible extension of the numerical modelling work would be to localize the inertia-gravity wave parameterization, so that the stochastic terms are only active in those regions where the magnitude of the Lighthill radiation term is large. This would help us to investigate whether the stochastic resonance phenomenon still occurs when the noise is localized. Spatio-temporal correlations could also be included in the noise terms, with realistic auto-correlations based on the observed short wave frequencies and wavelengths.

A second avenue for future research might be to focus on improving the laboratory/model agreement, for example by running a more accurate dynamical model which explicitly permits the short, fast waves. An extensive set of simulations, such as that presented in this thesis, could not be performed due to the computational expense. However, a small set of case studies should still provide enough material for a fruitful analysis.

A further important investigation would be to determine how close to a transition curve the system needs to be in order for stochastic resonance to take place, for a given noise amplitude. We would expect that as distance from a transition curve increases, the threshold noise amplitude for resonance would increase as the potential barrier in Figure 7.6 becomes taller. It would be useful to quantify this by performing numerical experiments to determine the threshold amplitude as a function of position in the parameter space. Such an investigation would allow us to make a balanced assessment of how frequently short waves in the atmosphere are expected to resonantly interact with long waves, an issue which should be of significant interest to the meteorological community due to the potential for forecast error that this phenomenon could inflict.

Appendix A

The *HSI* colour system

There are three different types of cones in the retina of the human eye, each containing different photosensitive pigments. Therefore every colour which is perceivable by humans is defined by only three independent quantities, and is representable by a single point in a three-dimensional space.[1] This is known as Young's trichromatic theory of colour vision (Longhurst, 1973). One of the most common colour systems uses (R, G, B) coordinates to define this space. Respectively, these are the red, green and blue components which, if combined, would give a colour which was indistinguishable to the normal human eye from the colour being represented.

Another common colour system uses (H, S, I) coordinates. The *intensity (I)* gives an indication of the total brightness of the colour, the *hue (H)* gives an indication of the dominant wavelength, and the *saturation (S)* gives an indication of the strength of the dominance. We now derive the transformation from the (R, G, B) to the (H, S, I) colour system, following Foley & Van Dam (1982).

Figure A.1 shows a general colour represented by the coordinates $\boldsymbol{C} = (R, G, B)$. The *achromatic axis* (or *grey axis*) is defined by the unit vector $\hat{\boldsymbol{a}} = \frac{1}{\sqrt{3}}(1, 1, 1)$. To determine (H, S, I) we first decompose $\boldsymbol{C}$ into a component along the achromatic axis and a component perpendicular to it:

$$\boldsymbol{C} \;=\; [(\boldsymbol{C}.\hat{\boldsymbol{a}})\hat{\boldsymbol{a}}] \;+\; [\boldsymbol{C} - (\boldsymbol{C}.\hat{\boldsymbol{a}})\hat{\boldsymbol{a}}] \;. \tag{A.1}$$

The greater the projection onto the achromatic axis, the brighter the colour. The greater

---

[1] Colour-blind people have only two different types of cones, and every colour they can perceive can be represented in a two-dimensional space.

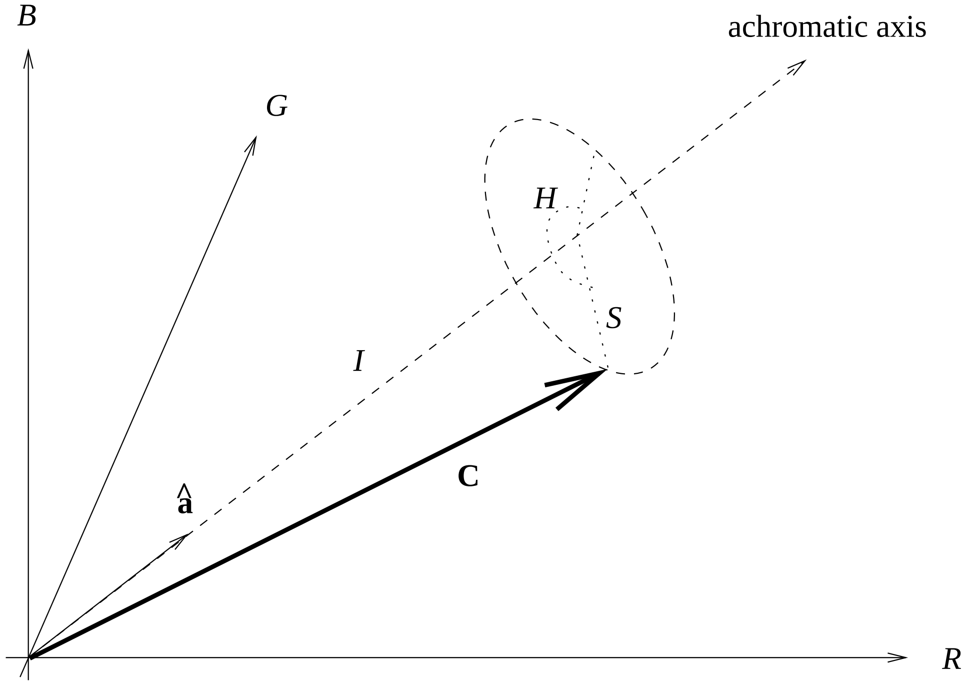


Figure A.1: The relationship between the Cartesian (R, G, B) and cylindrical (H, S, I) colour systems, shown geometrically. Hue describes the colour in terms of its angular position on a "colour wheel".

the distance from the achromatic axis, the more saturated and pure the colour (colours exactly on the axis are grey). In the plane perpendicular to the achromatic axis, the angle measured anti-clockwise from the vector pointing in the G-direction $(-1, 2, -1)$ determines the dominant wavelength, as shown in Table A.1. In this definition, green is arbitrarily assigned a hue of zero. This is the most useful definition for our purposes, as green hues are rarely (if ever) observed in the laboratory experiment images, and so there is no need to worry about the connection between $H = 0°$ and $H = 360°$ in the calibration curve.

Correspondingly, we define

$$I = \sqrt{3}\,\boldsymbol{C.\hat{a}} \quad , \tag{A.2}$$

$$S = |\ \boldsymbol{C}-(\boldsymbol{C.\hat{a}})\boldsymbol{\hat{a}}\ |\ , \tag{A.3}$$

and

$$H = \cos^{-1}\left[\frac{(-1,2,-1)}{|(-1,2,-1)|}.\frac{\boldsymbol{C}-(\boldsymbol{C.\hat{a}})\boldsymbol{\hat{a}}}{|\ \boldsymbol{C}-(\boldsymbol{C.\hat{a}})\boldsymbol{\hat{a}}|}\right]\ , \tag{A.4}$$

where, for uniqueness, we require

$$0° < H < 180° \text{ if } R < B\ , \tag{A.5}$$

$$180° < H < 360° \text{ if } R > B\ . \tag{A.6}$$

	hue, H (degrees)
green	0
cyan	60
blue	120
magenta	180
red	240
yellow	300
green	360

Table A.1: Hue, given as an angular position on a colour wheel. The zero of hue is here arbitrarily assigned to green, though it is more common to assign it to red so that the colours of the rainbow are cycled through in order as hue increases from 0° to 360°.

Evaluating the expressions in (A.2)–(A.4) leads to

$$I = R + G + B \ , \tag{A.7}$$

$$S = \sqrt{\frac{1}{3}\left[(R-G)^2 + (R-B)^2 + (G-B)^2\right]} \ , \tag{A.8}$$

and

$$H = \cos^{-1}\left[\frac{2G - R - B}{\sqrt{2[(R-G)^2 + (R-B)^2 + (G-B)^2]}}\right] \ . \tag{A.9}$$

The definitions (A.7)–(A.9) are used in the calibration analysis of Chapter 3.

Bibliography

Acheson, D. J. (1990). *Elementary Fluid Dynamics.* Oxford University Press.

Appleby, J. C. (1982). *Comparative theoretical and experimental studies of baroclinic waves in a two-layer system.* Ph.D. thesis, University of Leeds.

Appleby, J. C. (1988). Selection of baroclinic waves. *Q. J. R. Meteorol. Soc.*, **114**, 1173–1179.

Arakawa, A. (1966). Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. *J. Comp. Phys.*, **1**, 119–143.

Arnold, V. I. (1963). Small denominators and the problem of the stability of motion in classical and celestial mechanics. *Russ. Math. Surv.*, **18**, 85–191.

Baer, F. (1977). Adjustments of initial conditions required to suppress gravity oscillations in non-linear flows. *Beiträge zur Physik der Atmosphäre*, **50**, 350–366.

Baer, F. & Tribbia, J. (1977). On complete filtering of gravity modes through non-linear initialization. *Monthly Weather Rev.*, **105**, 1536–1539.

Benney, D. J. (1977). A general theory for interactions between short and long waves. *Stud. Appl. Math.*, **56**, 81–94.

Bjerknes, V. (1904). Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. *Meteor. Zeit.*, **21**, 1–7.

Boas, M. L. (1983). *Mathematical methods in the physical sciences.* Wiley, second edition.

Bradford, J., Berman, A. S., & Lundgren, T. S. (1981). Nongeostrophic baroclinic instability in a two-layer rotating system. *J. Atmos. Sci.*, **38**, 1376–1389.

Brown, R. (1973). New indices to locate clear-air turbulence. *Meteorological Magazine*, **102**, 347–361.

Brugge, R., Nurser, A. J. G., & Marshall, J. C. (1987). A quasi-geostrophic ocean model: some introductory notes. Technical report, Blackett Laboratory, Imperial College, London.

Carrigan, C. R. (1978). Instability of a two-layer baroclinic flow in a channel. *Tellus*, **30**, 468–471.

Cattaneo, F. & Hart, J. E. (1990). Multiple states for quasi-geostrophic channel flows. *Geophys. Astrophys. Fluid Dyn.*, **54**, 1–33.

Charney, J. G. (1948). On the scale of atmospheric motions. *Geofysiske Publikasjoner (Geophysica Norvegica)*, **17**(2), 17pp.

Charney, J. G. (1955). The use of the primitive equations of motion in numerical prediction. *Tellus*, **7**, 22–26.

Charney, J. G., Fjørtoft, R., & von Neumann, J. (1950). Numerical integration of the barotropic vorticity equation. *Tellus*, **2**(4), 237–254.

Daley, R. (1980). Normal mode initialization. In *ECMWF seminar proceedings: Data Assimilation Methods, 15–19 September 1980*, pages 107–158.

Daley, R. (1982). A non-iterative procedure for the time integration of the balance equations. *Monthly Weather Rev.*, **110**, 1821–1830.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Commun. Pur Appl. Math.*, **41**, 909–996.

Davey, M. K. (1978). Recycling flow over bottom topography in a rotating annulus. *J. Fluid Mech.*, **87**(3), 497–520.

De Swart, H. E. & Grasman, J. (1987). Effect of stochastic perturbations on a low-order spectral model of the atmospheric circulation. *Tellus*, **39A**, 10–24.

Douglas, J. F. & Gasiorek, J. M. K. (2000). *Fluid Mechanics*. Pearson Higher Education.

Drazin, P. G. (1992). *Nonlinear systems*. Cambridge University Press.

Ekman, V. W. (1905). On the influence of the earth's rotation on ocean currents. *Ark. Math. Astr. Fys.*, **2**, 1–52.

Foley, J. D. & Van Dam, A. (1982). *Fundamentals of interactive computer graphics.* Addison-Wesley Publishing Co., Massachusetts.

Ford, R. (1994). Gravity wave radiation from vortex trains in rotating shallow water. *J. Fluid Mech.*, **281**, 81–118.

Ford, R., McIntyre, M. E., & Norton, W. A. (2000). Balance and the slow quasimanifold: some explicit results. *J. Atmos. Sci.*, **57**, 1236–1254.

Früh, W.-G. & Read, P. L. (1997). Wave interactions and the transition to chaos of baroclinic waves in a thermally driven rotating annulus. *Phil. Trans. R. Soc. London*, **355**(1722), 101–153.

Gill, A. E. (1982). *Atmosphere-Ocean Dynamics.* Academic Press.

Goldstein, R. J. (1983). *Fluid dynamics measurements.* Springer-Verlag.

Haine, T. W. N. & Williams, P. D. (2002). The role of nonhydrostatic dynamics in controlling development of a surface ocean front. *Ocean Modelling*, **4**, 121–135.

Haltiner, G. J. & Williams, R. T. (1980). *Numerical prediction and dynamic meteorology.* Wiley, 2nd edition.

Hart, J. E. (1972). A laboratory study of baroclinic instability. *Geophys. Fluid Dyn.*, **3**, 181–209.

Hart, J. E. (1973). On the behaviour of large-amplitude baroclinic waves. *J. Atmos. Sci.*, **30**, 1017–1034.

Hart, J. E. (1985). A laboratory study of baroclinic chaos on the f-plane. *Tellus*, **37A**, 286–296.

Hart, J. E. & Kittelman, S. (1986). A method for measuring interfacial wave fields in the laboratory. *Geophys. Astrophys. Fluid Dyn.*, **36**, 179–185.

Hide, R. (1977). Experiments with rotating fluids. *Q. J. R. Meteorol. Soc.*, **103**(435), 1–28.

Hide, R. & Mason, P. J. (1975). Sloping convection in a rotating fluid. *Adv. in Phys.*, **24**(1), 47–100.

Hide, R., Mason, P. J., & Plumb, R. A. (1977). Thermal convection in a rotating fluid subject to a horizontal temperature gradient: spatial and temporal characteristics of fully developed baroclinic waves. *J. Atmos. Sci.*, **34**, 930–950.

Hignett, P., White, A. A., Carter, R. D., Jackson, W. D. N., & Small, R. M. (1985). A comparison of laboratory measurements and numerical simulations of baroclinic wave flows in a rotating cylindrical annulus. *Q. J. R. Meteorol. Soc.*, **111**, 131–154.

Hines, C. O. (1997). Doppler spread parameterization of gravity wave momentum deposition in the middle atmosphere. Part 1: Basic formulation; Part 2: Broad and quasi monochromatic spectra and implementation. *J. Atmos. Solar Terr. Phys.*, **59**(4), 371–386 and 387–400.

Hinkelmann, K. (1951). Der Mechanismus des meteorologischen Lärmes. *Tellus*, **3**, 285–296.

Hogan, J. (2003). Eye can see better when it's noisy. *New Scientist*, (7th June 2003), 20–21.

Hough, S. S. (1898). On the application of harmonic analysis to the dynamical theory of the tides. II: On the general integration of Laplace's dynamical equations. *Phil. Trans. R. Soc. London*, **191**, 139–85.

Jacobs, S. J. (1991). On the existence of a slow manifold in a model system of equations. *J. Atmos. Sci.*, **48**, 893–901.

James, I. N. (1977). Stability of flow in a slowly rotating two-layer system. Technical report, Geophysical Fluid Dynamics Laboratory, UK Meteorological Office. Occasional Note Met O 21/77/2.

Kaye, G. W. C. & Laby, T. H. (1995). *Tables of Physical and Chemical Constants*. Longman, London, 16th edition.

King, J. C. (1979a). An experimental study of baroclinic wave interactions in a two-layer system. *Geophys. Astrophys. Fluid Dyn.*, **13**, 153–167.

King, J. C. (1979b). *Instabilities and nonlinear wave interactions in a two-layer rotating fluid*. Ph.D. thesis, University of Leeds.

Klein, P. (1990). Transition to chaos in unstable baroclinic systems: a review. *Fluid Dyn. Res.*, **5**, 235–254.

Kwon, H. J. & Mak, M. (1988). On the equilibration in nonlinear barotropic instability. *J. Atmos. Sci.*, **45**(2), 294–308.

Leith, C. E. (1980). Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.*, **37**, 958–968.

Lewis, S. R. (1992). A quasi-geostrophic numerical model of a rotating internally heated fluid. *Geophys. Astrophys. Fluid Dyn.*, **65**, 31–55.

Lide, D. R., editor (1995). *Handbook of Chemistry and Physics*. Chemcal Rubber Publishing Company, 76th edition.

Lighthill, M. J. (1952). On sound generated aerodynamically. I: General theory. *Proc. R. Soc. Lond.*, **211**, 564–587.

Lighthill, M. J. (1978). *Waves in fluids*. Cambridge University Press.

Longhurst, R. S. (1973). *Geometrical and Physical Optics*. Longman, 3rd edition.

Longuet-Higgins, M. (1968). The eigenfunctions of Laplace's tidal equations over a sphere. *Phil. Trans. R. Soc. London*, **262**(1132), 511–607.

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646.

Lorenz, E. N. (1980). Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.*, **37**, 1685–1699.

Lorenz, E. N. (1986). On the existence of a slow manifold. *J. Atmos. Sci.*, **43**(15), 1547–1557.

Lorenz, E. N. (1992). The slow manifold — what is it? *J. Atmos. Sci.*, **49**(24), 2449–2451.

Lorenz, E. N. & Krishnamurthy, V. (1987). On the nonexistence of a slow manifold. *J. Atmos. Sci.*, **44**, 2940–2950.

Lovegrove, A. F. (1997). *Bifurcations and instabilities in rotating two-layer fluids.* Ph.D. thesis, Oxford University.

Lovegrove, A. F., Read, P. L., & Richards, C. J. (2000). Generation of inertia-gravity waves in a baroclinically unstable fluid. *Q. J. R. Meteorol. Soc.*, **126**, 3233–3254.

Lynch, P. (1986). Techniques of initialization. *Weather*, **42**(3), 66–70.

Lynch, P. (1989). The slow equations. *Q. J. R. Meteorol. Soc.*, **115**, 201–219.

Lynch, P. (1991). Filtered equations and filtering integration schemes. In *ECMWF seminar proceedings: Numerical methods in atmospheric models, 9–13 September 1991, Volume 1*, pages 119–159.

Lynch, P. (1993). Richardson's forecast factory: the $64,000 question. *The Meteorological Magazine*, **122**, 69–70.

Lynch, P. (1996). The Dolph-Chebyshev window: a simple optimal filter. *Monthly Weather Rev.*, **125**, 655–660.

Lynch, P. (1999). *The life cycles of extratropical cyclones (editors: Shapiro, M. A. and Grønås, S.)*, chapter *Richardson's marvellous forecast*, pages 61–73. American Meteorological Society, Boston.

Lynch, P. (2000). The swinging spring: a simple model of atmospheric balance. In *Proceedings of the Symposium on the Mathematics of Atmosphere-Ocean Dynamics, Isaac Newton Institute, June–December 1996.* Cambridge University Press.

Lynch, P. (2001). *Meteorology at the Millennium (editor: Pearce, R.)*, chapter *Weather forecasting: from woolly art to solid science.* Academic Press.

Machenhauer, B. (1977). On the dynamics of gravity oscillations in a shallow water model, with applications to non-linear normal mode initialization. *Beiträge zur Physik der Atmosphäre*, **50**, 253–271.

Madan, R. (1993). *Chua's Circuit: A Paradigm for Chaos.* World Scientific Publishing Company, Singapore.

March, J. (1992). *Advanced Organic Chemistry.* Wiley, 4th edition.

Margules, M. (1893). Luftbewegungen in einer rotierenden Sphäroidschale. *Sitzungsberichte der Kaiserliche Akad. Wiss. Wien, Teil IIA*, **102**, 11–56.

Massaldi, H. A. & King, C. J. (1973). Simple technique to determine solubilities of sparingly soluble organics: solubility and activity coefficients of d-limonene, n-butylbenzene, and n-hexyl acetate in water and sucrose solutions. *J. Chem. Eng. Data*, **18**(4), 393–397.

McIntyre, M. E. (1967). *Convection and baroclinic instability in rotating fluids.* Ph.D. thesis, Cambridge University.

McIntyre, M. E. & Norton, W. A. (2000). Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, **57**, 1214–1235.

McWilliams, J. C. & Gent, P. R. (1980). Intermediate models of planetary circulations in the atmosphere and ocean. *J. Atmos. Sci.*, **37**(8), 1657–1678.

Mesinger, F. & Arakawa, A. (1976). Numerical methods used in atmospheric models. *Global Atmospheric Research Programme Publications Series No. 17.*

Miyakoda, K. & Moyer, R. (1968). A method of initialization for dynamical weather forecasting. *Tellus*, **20**, 115–128.

Moura, A. D. (1976). The eigensolutions of the linearized balance equations over a sphere. *J. Atmos. Sci.*, **33**, 877–907.

Mundt, M. D., Brummell, N. H., & Hart, J. E. (1995). Linear and nonlinear baroclinic instability with rigid sidewalls. *J. Fluid Mech.*, **291**, 109–138.

O'Sullivan, D. & Dunkerton, T. J. (1995). Generation of inertia-gravity waves in a simulated life cycle of baroclinic instability. *J. Atmos. Sci.*, **52**(21), 3695–3716.

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Q. J. R. Meteorol. Soc.*, **127**, 279–304.

Pedlosky, J. (1964). The stability of currents in the atmosphere and the ocean: Part I. *J. Atmos. Sci.*, **21**, 201–219.

Pedlosky, J. (1970). Finite-amplitude baroclinic waves. *J. Atmos. Sci.*, **27**, 15–30.

Pedlosky, J. (1971). Finite-amplitude baroclinic waves with small dissipation. *J. Atmos. Sci.*, **28**, 587–597.

Pedlosky, J. (1972). Limit cycles and unstable baroclinic waves. *J. Atmos. Sci.*, **29**, 53–63.

Pedlosky, J. (1987). *Geophysical Fluid Dynamics*. Springer-Verlag.

Phillips, N. A. (1951). A simple three-dimensional model for the study of large-scale extratropical flow patterns. *J. Meteor.*, **8**, 381–394.

Phillips, N. A. (1954). Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus*, **3**, 273–286.

Phillips, N. A. (1956). The general circulation of the atmosphere: a numerical experiment. *Q. J. R. Meteorol. Soc.*, **82**(352), 123–164.

Phillips, N. A. (1960). On the problem on initial data for the primitive equations. *Tellus*, **12**, 121–126.

Phillips, N. A. (1963). Geostrophic motion. *Rev. Geophys.*, **1**(2), 123–176.

Piani, C. & Norton, W. A. (2003). The equatorial stratospheric response to variations in deterministic and stochastic gravity-wave parameterizations. *J. Geophys. Res.* submitted.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: a universal concept in nonlinear sciences*. Cambridge University Press.

Read, P. L. (1992a). Applications of singular systems analysis to 'baroclinic chaos'. *Physica D*, **58**, 455–468.

Read, P. L. (1992b). *Rotating fluids in geophysical and industrial applications (editor: Hopfinger, E. J.)*, chapter III.1 *Dynamics and instabilities of Ekman and Stewartson layers*, pages 49–84. Springer-Verlag.

Read, P. L., Collins, M., Früh, W.-G., Lewis, S. L., & Lovegrove, A. F. (1998). Wave interactions and baroclinic chaos: a paradigm for long timescale variability in planetary atmospheres. *Chaos, Solitons and Fractals*, **9**(1/2), 231–249.

Reynolds, O. (1883). An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Phil. Trans. R. Soc. London*, **174**, 935–982.

Richardson, L. F. (1922). *Weather prediction by numerical process.* Cambridge University Press. (Reprinted by Dover Publications, New York, 1965).

Roach, W. T. (1970). On the influence of synoptic development on the production of high level turbulence. *Q. J. R. Meteorol. Soc.*, **96**, 413–429.

Roach, W. T. & Bysouth, C. E. (2002). How often does severe clear air turbulence occur over tropical oceans. *Weather*, **57**, 8–19.

Robert, A. J. (1966). The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan*, **44**(5), 237–245.

Salmon, R. & Talley, L. D. (1989). Generalizations of Arakawa's Jacobian. *J. Comp. Phys.*, **83**, 247–259.

Sato, K. (1994). A statistical study of the structure, saturation and sources of inertio-gravity waves in the lower stratosphere observed with the MU radar. *J. Atmos. Terr. Phys.*, **56**(6), 755–774.

Sato, K., Kumakura, T., & Takahashi, M. (1999). Gravity waves appearing in a high-resolution GCM simulation. *J. Atmos. Sci.*, **56**(8), 1005–1018.

Smith, R. K. (1974). On limit cycles and vacillating baroclinic waves. *J. Atmos. Sci.*, **31**, 2008–2011.

Smith, R. K. (1977). On a theory of amplitude vacillation in baroclinic waves. *J. Fluid Mech.*, **79**(2), 289–306.

Smith, R. K. & Pedlosky, J. (1975). A note on a theory of vacillating baroclinic waves *and* Reply. *J. Atmos. Sci.*, **32**, 2027.

Stephen, A. V. (1998). *POD methods in baroclinic flows.* Ph.D. thesis, University of Oxford.

Stewartson, K. (1957). On almost rigid rotations. *J. Fluid Mech.*, **3**, 17–26.

Trundle, E. (1987). *Television and video engineer's pocket book (page 75).* Heinemann Newnes.

Vettin, F. (1884). Experimentelle Darstellung von Luftbewegungen unter dem Einfluss von Temperatur-Unterschieden und Rotations-Impulsen. *Z. Meteor.*, **1**, 227–230.

Vora, N. & Daoutidis, P. (2001). Nonlinear model reduction of chemical reaction systems. *Am. Inst. Chem. Eng. Journal*, **47**(10).

Warn, T. & Menard, R. (1986). Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model. *Tellus*, **38A**, 285–294.

White, A. A. (1986). Documentation of the finite difference schemes used by the Met O 21 2D Navier-Stokes model. Technical report, Geophysical Fluid Dynamics Laboratory, UK Meteorological Office. Internal Report Met O 21 IR86/3.

WHO (1998). *Concise International Chemical Assessment Document no. 5: Limonene.* International Programme on Chemical Safety, World Health Organization, Geneva.

WHO (2002). *International Chemical Safety Card 0050: 1,1,2-trichloro-1,2,2-trifluoroethane.* International Programme on Chemical Safety, World Health Organization, Geneva.

Williams, G. P. (1979). Planetary circulations: 2. The Jovian quasi-geostrophic regime. *J. Atmos. Sci.*, **36**, 932–968.

Williams, P. D. (2000). Interactions of 'fast' and 'slow' modes in rotating, stratified flows. First Year Report, Department of Physics, Oxford University.

Williams, P. D. (2001). Interactions of 'fast' and 'slow' modes in rotating, stratified flows. Second Year Report, Department of Physics, Oxford University.

Williamson, D. (1976). Normal mode initialization procedure applied to forecasts with the global shallow water equations. *Monthly Weather Rev.*, **104**, 195–206.

Williamson, D. & Temperton, C. (1981). Normal mode initialization for a multilevel grid-point model. Part II: Non-linear aspects. *Monthly Weather Rev.*, **109**, 745–757.

Yavneh, I. & McWilliams, J. C. (1994). Breakdown of the slow manifold in the shallow-water equations. *Geophys. Astrophys. Fluid Dyn.*, **75**, 131–161.

Yoshida, A. & Hart, J. E. (1986). A numerical study of baroclinic chaos. *Geophys. Astrophys. Fluid Dyn.*, **37**, 1–56.

Zhu, L., Lai, Y.-C., Liu, Z., & Raghu, A. (2002). Can noise make nonbursting chaotic systems more regular? *Physical Review E*, **66**(015204(R)).