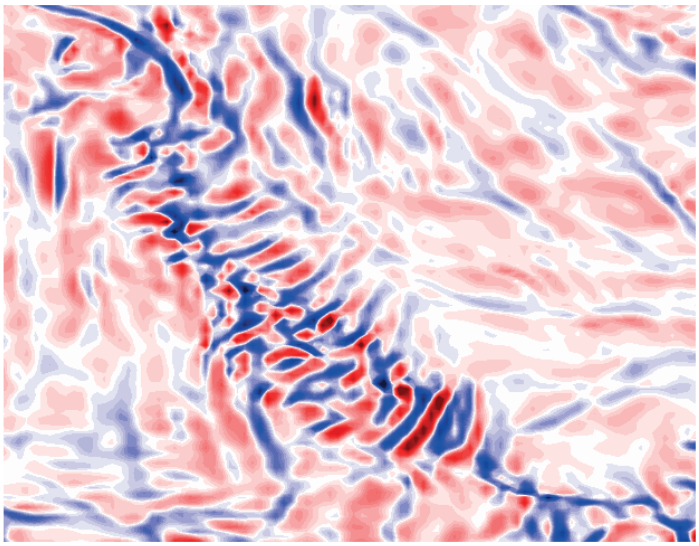
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Disentangling gravity waves from balanced flow



Speaketh but, O' mighty ocean!
In thy wavy world of fast and slow
How doth thee split up fast from slow?

Manita Chouksey

Hamburg 2018

Berichte zur Erdsystemforschung
Reports on Earth System Science

210
2018

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

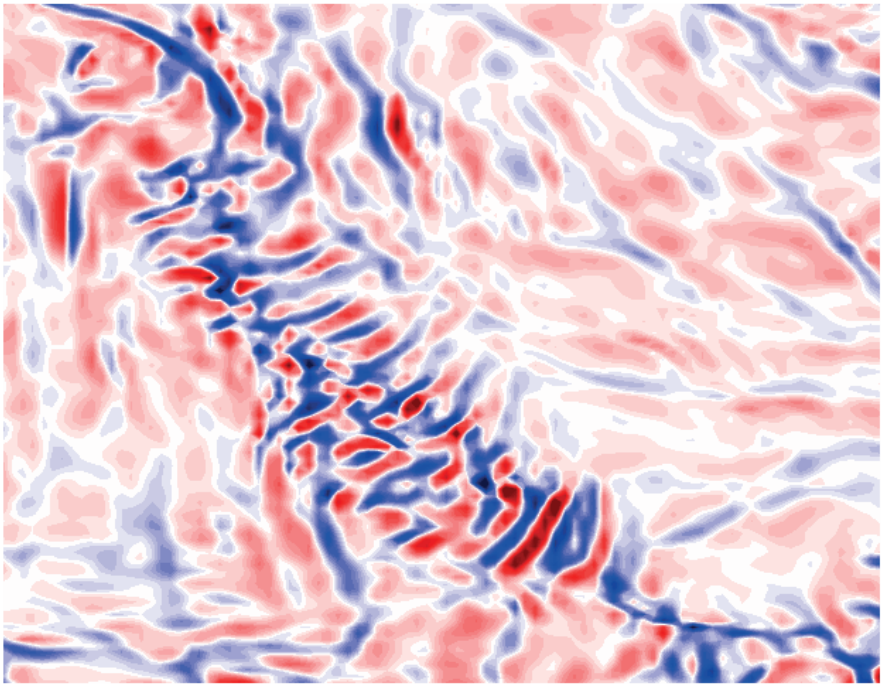
name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Disentangling gravity waves from balanced flow



Dissertation with the aim of achieving a doctoral degree at the
Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg

submitted by

Manita Chouksey

aus Jabalpur, Indien

Hamburg 2018

Manita Chouksey

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 16.01.2018

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Carsten Eden
Prof. Dr. Ulrich Achatz

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 210
Reports on Earth System Science / Max Planck Institute for Meteorology 2018

ISSN 1614-1199

O' mighty ocean!

Thou singeth tale of tangled waves

Waves thee maketh of all scales

Thou maketh wave of snail's pace

Thou maketh wave of rapid pace

Speaketh but, O' mighty ocean!

In thy wavy world of fast and slow

How doth thee split up fast from slow?

Zusammenfassung

Diese Arbeit zielt darauf ab, interne Schwerewellen von balancierten Strömungen zu trennen, um ihre Rolle für den Energietransfer zu kleineren Skalen zu untersuchen. Dafür werden zwei konzeptionell verschiedene diagnostische Methoden verwendet: eine modale Zerlegung nach Machenhauer (1977) sowie eine neuartige Methode, genannt quasi-geostrophischer Filter. Weiterhin untersucht die Arbeit mittels eines idealisierten numerischen Modells für baroklin-instabile Strömung in verschiedenen dynamischen Szenarien sowohl die Stärke der Erzeugung interner Schwerewellen als auch ihre Rolle für die Energiedissipation auf kleinen Skalen.

Interne Schwerewellen (ISW) werden zunächst mittels der von Machenhauer (1977) vorgeschlagenen nichtlinearen Initialisationstechnik diagnostiziert und in verschiedenen dynamischen Regimen, die von ageostrophisch bis quasi-geostrophisch reichen, untersucht. Aufgrund der komplexen Kopplung zwischen ISW und balancierten Strömungen, besonders im ageostrophischen Regime, eignen sich Frequenz-Wellenzahl-Spektren nicht für den präzisen Nachweis von ISW-Signalen. Um diese Komplikation zu umgehen, wird das gesamte Strömungsfeld mithilfe linearer sowie nichtlinearer modaler Zerlegung in seine balancierten und unbalancierten Komponenten aufgeteilt. Die lineare Zerlegung funktioniert relativ gut, bezieht aber the Nichtlinearität der Strömung nicht mit ein, was zu Ungenauigkeiten führt. Diese können jedoch in der nichtlinearen Zerlegung basierend auf Machenhauers Ansatz, die sich als eine effizientere und robustere diagnostische Methode zum Nachweis von ISW herausstellt, vermieden werden. Eine Untersuchung der diagnostizierten ISW in verschiedenen dynamischen Regimes zeigt eine deutlich gesteigerte Aktivität der Wellen im ageostrophischen Regime im Vergleich zum Quasi-geostrophischen. Außerdem dissipieren ISW vorzugsweise durch kleinskalige Dissipation und nehmen daher eine bedeutende Rolle im ozeanischen Energietransfer zu kleineren Skalen ein.

Eine neuartige Gleichgewichtsmethode, quasi-geostrophischer (QG) Filter genannt, wird zur Diagnose von ISW eingeführt. Im Gegensatz zu früheren auf Inversion der potentiellen Wirbelstärke beruhenden Methoden ermöglicht diese zusätzlich die Bestimmung ageostrophischer horizontaler Geschwindigkeiten. Der QG-Filter ist in einer auf dem diskreten Gitter konsistenten Version implementiert und kann sowohl als Gleichgewichtsprozedur als auch als diagnostisches Werkzeug verwendet werden. Die Ergebnisse zeigen, dass der QG-Filter eine effiziente Methode zum Erreichen des Gleichgewichts ist, allerdings schneidet eine umformulierte diskrete Version des Machenhauer'schen Ansatzes noch besser ab. Ferner werden die mittels QG-Filter identifizierten ISW genutzt, um die Erzeugung von

Schwerewellen im Ozean aus balancierten Strömungen durch verschiedene Mechanismen zu untersuchen: spontane Emission, konvektive Instabilität durch Frontogenese und laterale Grenzschicht-Instabilität. Dabei ergibt sich im Fall von konvektiver Instabilität eine höhere ISW-Aktivität als bei den anderen beiden Prozessen. Die Ergebnisse etablieren ISW als eine potentielle Energiesenke von balancierten Strömungen, was wesentliche Auswirkungen auf den Energiehaushalt des Ozeans hat.

Abstract

This thesis aims to disentangle internal gravity waves from balanced flow to investigate their role in the downscale transfer of energy in the ocean by using two conceptually different diagnostic procedures—a modal decomposition based on Machenhauer (1977) and a novel balance procedure called the quasi-geostrophic filter. Furthermore, using an idealized numerical model for baroclinically unstable flow in various dynamical scenarios, this thesis investigates both the strength of internal gravity wave generation and its role in the energy dissipation at small scales.

Internal gravity waves (IGWs) are first diagnosed using the non-linear initialization technique proposed by Machenhauer (1977) and are assessed in different dynamical regimes ranging from ageostrophic to quasi-geostrophic. The complex coupling between IGWs and balanced flow, in particular for an ageostrophic regime, renders frequency-wavenumber spectra inefficient for a precise detection of IGW signals. To tackle this complexity the full flow field is decomposed into its balanced and unbalanced components using linear and non-linear modal decompositions. The linear decomposition performs reasonably well but does not include the non-linearity of the flow leading to discrepancies. To account for this, a non-linear decomposition based on Machenhauer's technique is applied which turns out to be more efficient and a robust diagnostic tool to obtain IGWs. An assessment of the diagnosed IGWs in different dynamical regimes reveals that IGW activity is much higher in the ageostrophic than the quasi-geostrophic regime. Furthermore, IGWs dissipate predominantly through small-scale dissipation, and therefore play a significant role in the downscale transfer of energy in the ocean.

A new balance procedure called the quasi-geostrophic (QG) filter is introduced to diagnose IGWs, which additionally provides the ageostrophic horizontal velocities unavailable from previous potential vorticity inversion methods. The QG filter is implemented in a version consistent at the discrete grid level and can be used as a balance procedure or as a diagnostic tool. Results show that the QG filter is an efficient tool to obtain balance, however a reformulated discrete version of the Machenhauer's technique performs even better. Further, IGWs diagnosed from the QG filter are used to examine IGW emission from balanced flow in the ocean by different mechanisms: spontaneous emission, convective instability by frontogenesis, and lateral boundary instability, where convective instability shows more IGW activity than the other two processes. The results establish IGWs as a potential sink of balanced flow energy which has key implications for the ocean's energy budget.

Contents

Chapter 1

Introduction

"Dead water" is a phenomenon wherein the slow sailing vessels refuse to move forward owing to a mysterious resistance from the water below. Ancient mariners blamed the bewitched drowned or remora, the suckerfish, for holding their ships on the water and deceived by the water's calm demeanour at the surface, they called it 'dead water' (Nansen, 1897). With time, the expanding horizons of our cognizance and deeper dives into the mysteries of ocean dynamics have been revealing phenomena of boundless spectacles, therewith that cursed phenomenon has been given a lively new term: internal gravity waves.

Internal gravity waves (IGWs) are omnipresent in the Earth's atmosphere and ocean. These waves occur naturally and profoundly impact the atmospheric and oceanic flows as well as the Earth's climate. Contrary to IGWs which evolve at fast timescales, the motions which dictate the atmospheric and oceanic dynamics —and consequently weather and climatic patterns— manifest themselves at a lot longer timescales and evolve at a much slower pace, such as the mesoscale motions. Mesoscale motions and IGWs are the respective examples of balanced and unbalanced components of the geophysical flows. The definition of balanced and unbalanced is tricky as well as crucial for a precise depiction of the flow dynamics and is specified in Section 1.1. These balanced and unbalanced components interact and exchange momentum and hence influence the energy and momentum cycle.

But the tale has just begun, and hither comes the twist: the unbalanced motions are interlinked with the balanced motions in an intricate fashion such that their segregation becomes a challenging task. This intricacy is an obstacle in understanding the interaction mechanisms between these motions and their sources and sinks of energy, which have potential implications for the energy cycle of the atmospheric and oceanic flows and hence their overall dynamics. It is therefore of prime importance to disentangle the unbalanced

motions from the balanced ones to get a picture of the underlying dynamics and energetics of the geophysical flows.

The commonly used methods to separate balanced and unbalanced dynamics assign the extracted signals to unbalanced dynamics imprecisely leading to an ambiguous interpretation of the unbalanced dynamics. An inconsistent interpretation can have misleading consequences for the parameterization of unbalanced dynamics such as IGWs which are often parameterized in the models. This thesis aims at a more precise separation of the unbalanced and balanced motions by implementing two distinct procedures with dual capabilities: as a diagnostic tool to diagnose IGWs as well as a balance procedure to eliminate unbalanced IGW dynamics. An efficient extraction of IGWs is important in the understanding of the energy transfers between balanced and unbalanced dynamics, such as the downscale energy transfer from balanced motions, which is a key energy pathway of the ocean's energy cycle.

This thesis addresses the disentangling of the IGWs from balanced flows in idealized stratified and rotating flows and investigates the role of IGWs in the downscale transfer of energy. This separation is achieved by applying two different procedures: first, a nonlinear initialization technique that makes adjustments to initial conditions such that the generation of unbalanced motions is minimized; second, a novel balance procedure which constrains the flow dynamics under suitable approximations and as a result filters out unbalanced motions by construction. Both the procedures are designed to extract balanced motions from the flow, but they can also be employed as diagnostic tools to obtain the unbalanced IGW signals. The extraction of IGWs using these two diagnostic tools forms the theme of this thesis. Further, the results are evaluated in different dynamical scenarios which portray the interactions between unbalanced and balanced dynamics. For this, the IGW activity is assessed in different dynamical regimes ranging from ageostrophic to quasi-geostrophic (Chapter 2), as well as in different mechanisms of IGW emission in the ocean such as spontaneous emission, convective instability, and lateral boundary instability (Chapter 3).

This chapter unfolds with a brief characterization of the unbalanced and balanced dynamics. Then comes a description of the generation, dissipation, and the role of unbalanced and balanced motions in the ocean, such as IGWs and mesoscale eddies respectively. Next, specific research questions addressed in this thesis in relation to the downscale energy transfer in the ocean are laid out. The last section of this chapter contains an outline of the forthcoming chapters of the thesis.

1.1 Balanced and unbalanced dynamics

A stratified fluid, i.e. a fluid in which density varies with depth along the axis of gravity, when subject to rotation generates a variety of flow phenomena of geophysical interest. Geophysical flows, such as the Earth's ocean and atmosphere, are examples of such a stratified rotating system. In the ocean and atmosphere, their own stratification and the Earth's rotation shape the dynamics of the flow and generate motions of varied scales which primarily determine the weather and climate of the Earth. The characteristics of these motions can thus be described by the associated rotation and stratification which are measured respectively by the Coriolis frequency f and the Brunt-Väisälä (or stratification) frequency N:

$$f = 2\Omega \sin\varphi \quad , \quad N = \Big[-\frac{g}{\rho_0}\frac{d\rho}{dz} \Big]^{1/2} \tag{1.1}$$

where, Ω is the rotation rate of the Earth, φ is the geographic latitude, g is acceleration due to gravity, ρ_0 is a reference or mean value of density, and $d\rho/dz$ indicates density (ρ) variation with depth z. The time period associated with f is called inertial period and that with N is called stability period. The stability period $2\pi/N$ ranges between 20 minutes in the upper ocean to about a few hours in the ocean interior and the inertial period $2\pi/f$ is about 12 hours, changing with the geographic latitude as in Eq. (1.1) (e.g. Olbers et al., 2012).

The motions which are strongly influenced by rotation and stratification are approximately 'balanced' in the sense that they exhibit a balance of forces and are hence called balanced motions. The most common of these balances are the geostrophic balance and hydrostatic balance. A geostrophic balance results when the pressure gradient is 'nearly' balanced by the Coriolis acceleration in the horizontal:

$$f\underline{\boldsymbol{u}} = -\frac{1}{\rho_0}\nabla_h p \tag{1.2}$$

where $\underline{\boldsymbol{u}}$ denotes the anticlockwise rotation of $\boldsymbol{u}$ by 90°, and $\nabla_h p$ denotes the horizontal pressure gradient of the flow. Since the forces 'nearly' balance each other, such a nearly balance state is called quasi-geostrophic balance. In the vertical, the pressure gradient balances the buoyancy force resulting in a hydrostatic balance:

$$-b\rho_0 = g\rho = -\frac{\partial p}{\partial z} \tag{1.3}$$

where $b = -g(\rho/\rho_0)$ is the buoyancy, ρ and ρ_0 refer to density perturbation and a reference density as stated above. These balance approximations result in flows which are in a state of approximate balance and are thus called quasi-geostrophic flows, or simply balanced flows. Departures from this state of balance can occur, for instance in regions near boundaries where other factors such as turbulent friction become important and these balance of forces no longer hold. However, such departures can also occur in the interior of the fluid and generate unbalanced motions. Flows with such departures constitute the unbalanced flows. The balanced and unbalanced motions can be further characterized by dimensionless parameters that estimate the strength of rotation and stratification: Rossby number and Richardson number.

The Rossby number (Ro) can be defined as the ratio of frequencies (or time periods) of the given flow and the Coriolis (or inertial) frequency and thus it estimates the timescale separation between the balanced and unbalanced motions.

$$\mathrm{Ro} = \frac{1}{f\mathcal{T}} = \frac{\omega_{flow}}{f} \tag{1.4}$$

where $\mathcal{T} = \mathcal{L}/\mathcal{U}$ is the relevant timescale of the flow under consideration, with $\mathcal{U}$ and $\mathcal{L}$ as the relevant horizontal velocity and horizontal length scales of the flow. Here $\omega_{flow} = 1/\mathcal{T}$ refers to the frequency of the flow under consideration. So, a balanced flow with timescales much longer than the inertial period is a low frequency motion w.r.t. the Earth's rotation, and so represents a $\mathrm{Ro} \ll 1$ regime. On the other hand, an unbalanced flow with equal or shorter timescales than the inertial period is a high frequency motion w.r.t. the Earth's rotation, and is indicated by a relatively large Ro.

Accordingly, a smaller Ro implies a large separation between balanced and unbalanced timescales with the consequence that the balanced and unbalanced motions are weakly coupled; on the contrary, for a large Ro, the balanced and unbalanced motions interact more strongly and so the separation of their timescales is poorly defined (Zeitlin, 2008; Vanneste, 2013). For the typical scales of large-scale oceanic flows, with L$\sim 10^5$ m and U$\sim$ 0.1 m/s, the typical Rossby number is Ro = 0.01 or smaller (e.g. Vallis, 2006). An increasing Ro indicates an increasing degree of imbalance and hence more notable unbalanced motions. In the ocean, high Ro can occur in regions such as the mixed layer, weakly stratified regions, and near the equator. The notion of coupling between balanced and unbalanced motions is further elaborated in Section 1.1.1 and illustrated in Fig. 1.1.

The Richardson number (Ri) can be defined as the ratio of the vertical density gradient to the vertical gradient of horizontal velocity and thus measures the relative importance

of stratification and shear effects in a fluid column. For a given stratification and shear of velocity u, Ri can be estimated as:

$$\mathrm{Ri} = \frac{N^2}{(du/dz)^2} \tag{1.5}$$

For a large Ri, the stratification dominates the effects of shear and the flow is strongly stratified. Balanced flows are characterized by a large Ri, typically $\mathrm{Ri} \gg 1$, indicating the dominant effect of stratification. Unbalanced flows on the other hand show a typical Ri of order unity, $\mathrm{Ri} = \mathscr{O}(1)$, indicating that stratification effects are no longer dominant. For $\mathrm{Ri} \leq \mathscr{O}(1)$ other kinds of instabilities emerge, such as shear and symmetric instabilities.

Accordingly, balanced and unbalanced flows can be described by Ri and Ro values of the flow. Balanced flows are characterized by $\mathrm{Ro} \ll 1$ and $\mathrm{Ri} \gg 1$, in contrast the unbalanced flows are characterized by finite Ro and smaller Ri, but still $\mathrm{Ro} < 1$ and $\mathrm{Ri} > 1$. Further, the linear dynamics of a flow can be described by Rossby waves and gravity waves respectively. Rossby waves have a zonal phase propagation only in one direction (negative) i.e. to the west w.r.t. the mean flow, whereas gravity waves can propagate in all directions (positive and negative). This distinction can be readily seen from the dispersion relation of the waves, which has two solutions for gravity waves and only one for Rossby waves (cf. Olbers et al., 2012). On the other hand, non-linear dynamics of the flow results in motions like mesoscale eddies and unbalanced turbulence arising from the non-linear interactions of Rossby waves and gravity waves.

The characteristics, generation, and dissipation mechanisms of IGWs (unbalanced motions) in the ocean are elaborated in Sections 1.2 and that of mesoscale eddies (balanced motions) in Section 1.3.

1.1.1 The ‘slow’ manifold and the swinging pendulum

Balanced flows dominate geophysical flows and evolve at much slower timescales relative to the period of Earth's rotation. This slowly varying nature of balanced flows led to the hypothetical concept of the "slow manifold" (Leith, 1980; Lorenz, 1980, 1986) to describe geophysical flow systems. A slow manifold refers to a subspace of the phase space where unbalanced dynamics are absent. The slow manifold thus imposes a strictly "invariant" state on the flow system implying that a balanced flow on a slow manifold would continue to remain balanced.

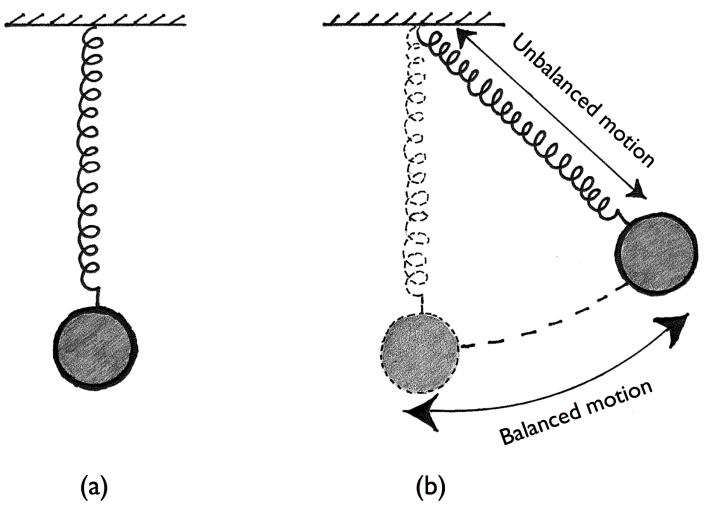


Figure 1.1: A sketch of the swinging pendulum system. (a) The spring-pendulum system at rest. (b) The system in motion. The oscillations in the spring have much higher frequency than the frequency of pendulum motion. In this way, the pendulum motion is analogous to the low frequency balanced motions that is accompanied by the oscillations in the spring which are analogous to the high frequency unbalanced motions (IGWs) in geophysical flows. The sketch is inspired from Lynch (2001) and Vanneste (2013).

Departures occur, however, from the balanced state leading to the generation of unbalanced motions like IGWs. Such departures contradict the slow manifold theory implying that a strictly invariant slow manifold does not exist. Several authors (e.g. Lorenz and Krishnamurthy, 1987; Lorenz, 1992; Ford et al., 2000; Vanneste and Yavneh, 2004) have discussed the existence –or rather the non-existence– of a strict slow manifold, leading to the notion that such a manifold is "fuzzy" rather than strict (e.g Warn, 1997; Ford et al., 2000). The 'fuzziness' of the slow manifold means that the balanced motion co-exists with unbalanced motion with different degrees of interaction determining the extent of 'fuzziness'. According to that, the flow dynamics can be thought of as a phase space with two interacting subspaces or manifolds: a balanced and an unbalanced fuzzy manifold, or in other words, a Rossby manifold and a gravity manifold (e.g. Leith, 1980; Theiss and Mohebalhojeh, 2009). Because it is fuzzy and not strict manifold, the balanced motion would inevitably generate unbalanced motions, for instance via spontaneous emission (Vanneste, 2013; also discussed in Chapters 2 and 3). The extent to which the unbalanced dynamics are generated depends upon certain flow parameters, such as the Rossby number described previously.

A useful piece of imagery to illustrate the concept of slow manifold with regard to balanced and unbalanced dynamics is by a simple mechanical system of a swinging pendulum

(e.g. Lynch, 2001; Vanneste, 2013). Consider a pendulum consisting of a heavy bob suspended from one end of a spring of lighter mass in which stretching is allowed but bending is not. The spring is fixed on the other end to complete the spring-pendulum system, as illustrated in Fig. 1.1(a). When the bob is displaced to a certain angle from its equilibrium position, as shown in Fig. 1.1(b), the bob is acted upon by restoring forces which set the pendulum into motion. The restoring for the bob comes from the force of gravity acting downward[1]. The oscillation of the pendulum is accompanied by another motion— that of the spring, subject to the restoring of the oscillating pendulum. The spring oscillates at a much higher frequency than the pendulum. The oscillations of the pendulum will induce spring oscillations related to the stretching and contracting of the spring. But the two types of motion are coupled by non-linear interactions due to the non-linear equations of motions that describes the dynamics of this system. As a consequence, the separation of these low and high frequency motions is not straightforward and becomes increasingly complex with a stronger coupling between these motions.

The ratio ε of frequencies of the fast and slow oscillations tells us the timescale separation between the two types of oscillations, just as the Rossby number estimates the timescale separation of the slow balanced and the fast unbalanced motions in geophysical flows (Eq. (1.4)). For a small ε, the spring oscillates at a much higher frequency than the pendulum and the timescales of these oscillations are well distinguished. The case of a larger ε implies that the two motions have comparable frequencies, and when ε becomes even larger the distinction between fast and slow oscillations is no longer feasible. In other words, this situation resembles a fuzzy manifold with a large extent of 'fuzziness'.

The scenario of the swinging pendulum system is analogous to geophysical flow systems where balanced motion coexists with the unbalanced IGWs. For a regime with $\mathrm{Ro} \ll 1$, the balanced and unbalanced timescales are well separated, which also means that the IGW generation in this regime is weak and consequently hard to detect. On the other hand, as the Rossby number approaches unity the distinction between the two timescales becomes less and less such that the balanced and unbalanced motions have comparable timescales and their separation becomes even more daunting.

This complication poses challenges to disentangle unbalanced motions from balanced motions and makes it difficult to obtain IGWs by conventional methods, such as the horizontal velocity divergence (e.g. Plougonven et al., 2005; Plougonven and Snyder, 2007) or small-scale vertical velocity obtained from spatial filtering (e.g. Sugimoto and Plougonven,

[1] An additional restoring force is provided by the tension force of the spring that acts to pull the bob upwards, thus introducing a non-linear coupling between the bob and the spring.

2016). To address this complication, two conceptually different methods are implemented to diagnose IGWs: the non-linear initialization technique by Machenhauer (1977) and a variant of the quasi-geostrophic balanced model. The methods are described in short in the next sections and are elaborated in detail in Chapters 2 and 3 respectively.

1.1.2 Modal decomposition

Modal decomposition is based on the non-linear normal mode initialization technique (NL-NMI) developed by Machenhauer (1977). The decomposition is a non-linear projection of the raw data onto the balanced manifold and by choosing appropriate initial conditions, the generation of high frequency motions can be minimized. The method proposed by Machenhauer (1977) is an iterative procedure which works by eliminating more and more of the high frequency motions with every iteration until an approximately balanced state is reached. Leith (1980) has shown that this method generates the quasi-geostrophic equations with the first iteration. Relating these initialization procedures to quasi-geostrophic balance, Leith (1980) derived decomposed modes for the hydrostatic Boussinesq equations, which was later generalised to the non-hydrostatic set of equations by Bartello (1995).

In this thesis, instead of initializing the model with this method, it is used to decompose the given full field into its balanced and unbalanced counterparts. Chapter 2 describes the details and implementation of the modal decomposition to diagnose unbalanced IGWs. The diagnosed IGWs are used to assess IGW activity in different dynamical regimes: from $\mathrm{Ro} \ll 1$ to $\mathrm{Ri} = \mathscr{O}(1)$ and to analyze the role of IGWs in the downscale transfer of energy.

1.1.3 Quasi-geostrophic filter

A balanced model is a model based on balance approximations such that the unbalanced dynamics are excluded. Such a model is helpful in understanding the balanced dynamics and a suite of balanced models have been developed (e.g. Gent and McWilliams, 1983). The most prominent of these is the quasi-geostrophic model (e.g. Vallis, 2006) which is a reasonable approximation to describe the flow dynamics of geophysical flows which are largely dominated by quasi-geostrophically balanced flows. The traditional methods based on quasi-geostrophic model diagnose IGWs by the inversion of potential vorticity (Vallis, 1996 and references therein). One such diagnostic tool is the quasi-geostrophic omega equation (e.g. Holton and Hakim, 2012) that extracts only the vertical velocity component to provide estimates of the unbalanced motions. But, the unbalanced motions are also

associated with the horizontal components of the velocity field which are not captured by the traditional methods. Consequently, the interpretation of the extracted IGW signals is ambiguous.

To tackle this issue, a variant of the quasi-geostrophic model that is better suited to diagnose the unbalanced IGW dynamics is introduced in Chapter 3. This quasi-geostrophic model variant is called the quasi-geostrophic filter (QG filter) and has the additional capacity to diagnose the unbalanced components associated with the horizontal velocities of the flow field. The method is explained in detail in Chapter 3. The IGWs diagnosed from the QG filter are explored in different scenarios in the ocean that generate IGWs: spontaneous emission, convective instability, and lateral boundary instability.

1.2 Energetics of internal gravity waves

IGWs are oscillations that occur in the interior of a density stratified fluid owing to the restoring force of the buoyancy. The interplay between buoyancy and gravity forces on a displaced fluid parcel sets the fluid parcel into an oscillatory wave motion. The wave motion is also influenced by the Coriolis force and because these oscillations occur in the interior of the fluid they are called 'internal' gravity waves, or simply internal waves.

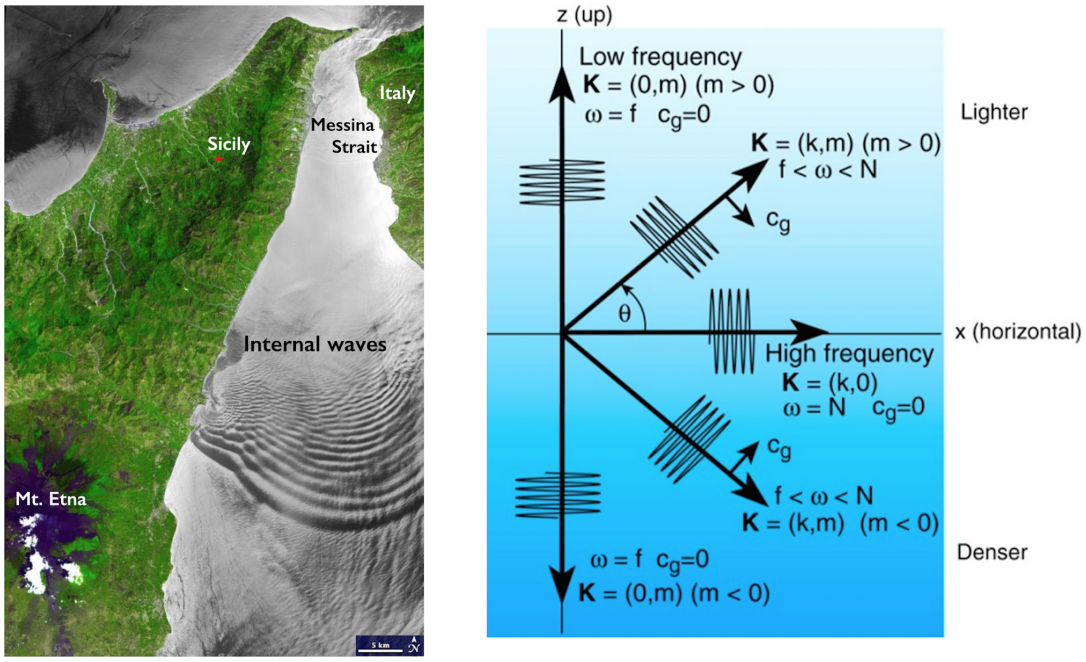
The characteristics of IGWs are influenced by both the stratification N and by the Earth's rotation (contained in f) at large scales. The frequency ω of IGWs is bounded by N and f which set the upper and lower bounds for ω. In the ocean, generally $N > f$ and so $f < \omega < N$. Accordingly, the time period of IGWs ranges between the stability period $2\pi/N$ and the inertial period $2\pi/f$. However, the frequency can be affected by the background flow, resulting in a Doppler shift. Thus, the frequency of encounter can be different from the intrinsic frequency.

The dispersion relation for IGWs relates the frequency ω and wavenumber $\boldsymbol{k}$, with k_h and m as the horizontal and vertical wavenumbers respectively and takes the following form:

$$\omega^2 = \frac{N^2 {k_h}^2 + f^2 m^2}{{k_h}^2 + m^2} \tag{1.6}$$

Another form of dispersion relation is in terms of the angle θ between the wave vector and the horizontal: $\omega^2 = N^2 \cos^2\theta + f^2 \sin^2\theta$. Here, $\cos^2\theta = {k_h}^2/\boldsymbol{k}^2$ and $\sin^2\theta = m^2/\boldsymbol{k}^2$ (for a

detailed discussion see Olbers et al., 2012). An example of the IGWs in the ocean and an illustration of IGW properties are shown in Fig. 1.2[2].



(a) (b)

Figure 1.2: (a) IGWs in the Strait of Messina captured from space. Strait of Messina is a narrow and shallow channel about 80 m deep, and lies between the island of Sicily and the Italian peninsula. Strong tidal currents occur in the Strait, which interact with the shallow sill and generate energetic internal wave trains. The dreaded conditions of the Strait, because of these energetic waves, are long known to have posed problems for sailors and has motivated the Greek myth of the sea monsters Scylla and Charybdis, who plagued Odysseus during his journey. IGWs show up as sea surface manifestations since they are associated with surface currents that modify the surface roughness which is seen by the satellite. Note that this is not related to sea surface elevations. The horizontal white line at the bottom indicates a scale of 5 km. Image courtesy of NASA. (b) Schematic illustrating IGW properties in a stably stratified fluid. IGW propagation for different values of the angle (θ) between the wave vector ($\boldsymbol{K}$) and the horizontal (x) is shown according to the dispersion relation in Eq. (1.6). Note that the direction of phase velocity (c_p), in the direction of wavevector the $\boldsymbol{K}$) is perpendicular to the direction of group velocity (c_g, shorter arrows). Schematic from Talley (2011).

Thus, the frequency ω of IGWs depends only on the angle θ of the wave vector i.e. the direction of wave propagation. The frequency, $\omega \approx N$ for an angle which is nearly horizontal ($\theta \approx 0$), whereas for a case with a nearly vertical angle ($\theta \approx \pm\pi/2$) the IGWs have frequency $\omega \approx f$, which are then called near-inertial waves. The direction of propagation

[2]The image in Fig. 1.2(a) is a courtesy of NASA/METI/AIST/Japan Space Systems and U.S./Japan ASTER Science Team and is taken from: https://asterweb.jpl.nasa.gov/gallery-detail.asp?name=messina-wave

of energy (group velocity) is perpendicular to the direction of propagation of wave crests and troughs (phase velocity). In a continuously stratified fluid, such as the ocean, the IGWs can propagate vertically as well as horizontally and exhibit large vertical amplitudes typically of the order of ten meters and wavelengths of up to a kilometer. IGW propagation seen from a tank experiment is shown in Fig. 1.3.

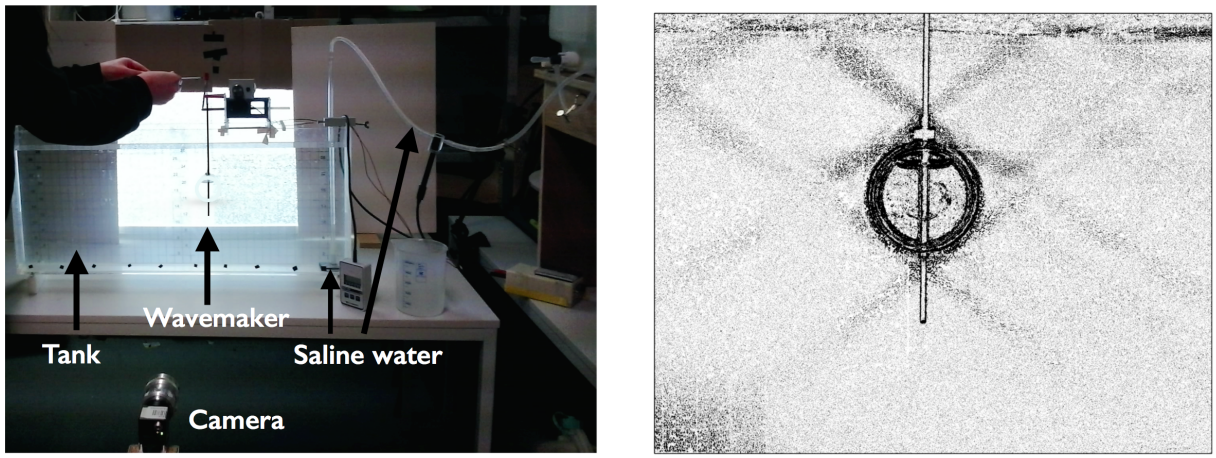


Figure 1.3: IGWs seen in a tank experiment using the synthetic Schlieren technique used to study the distribution of density gradients within a fluid as first demonstrated by Görtler (1943). The presence of density gradients leads to the deviation of light rays passing through these gradients. The interaction of the deflected light with an obstacle, such as the wavemaker in the tank, deflects the light rays to different degrees and leading to variations in the light distribution, which when illuminated makes these patterns visible. (a) The setup of the tank experiment. The water-filled tank is connected to a source of saline (denser) water which injects salty water at the bottom, and thus makes the water increasingly more denser and generates a linear density stratification. In the middle of the tank a wavemaker moves up and down with different frequencies to generate IGWs. Images of dots of different sizes and different dot densities are fixed behind the tank. The dots in the images appear to shift due to the deviation of light rays by density gradients, made visible by an illumination screen at the back and recorded by a camera. Removing the picture recorded at a state of rest from the subsequent frames with moving wavemaker makes the light refraction visible, similar to the classical Schlieren technique used e.g. Görtler (1943). The experiment was performed in the laboratory of Universität Hamburg by Prof. Carsten Eden, Lars Czeschel, Phlipp Jurgenowski, and Manita Chouksey and the setup was arranged by Ulrich Drübbisch. The hands in the picture belong to Philipp. (b) An image of IGW propagation obtained from the synthetic Schlieren technique. IGWs are clearly seen propagating in all directions away from the wavemaker (in the middle). The image was kindly provided by Prof. Carsten Eden.

1.2.1 Generation mechanisms

IGWs in the ocean are well known to be mainly forced by winds and tides. Wind stress fluctuations excite IGWs close to the inertial frequency f by inertial pumping (e.g. Gill, 1984) causing waves to radiate below the mixed layer (e.g. Alford et al (2016)). The tides generate IGWs at tidal frequencies, called internal tides[3], when the tidal flow oscillates over topographic obstacles at the bottom of the ocean, a process similar to the generation of lee waves (Bell, 1975; Nikurashin and Ferrari, 2011). Numerous studies related to these processes have contributed to global estimates of IGW conversion rates from winds (e.g., Alford 2001, Rimac et al. 2013) and tides (e.g. Arbic et al. (2004), Nycander (2005), Falahat et al. (2014)).

However, other sources of IGW generation, such as surface wave interactions and balanced flows, remain poorly understood. Surface wave interactions can generate IGWs through spontaneous creation (e.g. Olbers and Eden, 2016) in which two surface waves interact to produce an IGW by non-linear resonant interactions. Balanced oceanic flows, such as mesoscale eddies, generate IGWs by two broad mechanisms—(i) internal mechanisms of generation without any external interaction and (ii) external mechanisms of generation by interaction with external factors such as topography, wind, and waves.

Internal mechanisms of IGW generation by balanced flows include spontaneous emission, frontogenesis, and loss of balance. Spontaneous emission (e.g. review by Vanneste, 2013) refers to the emission of IGWs from a balanced flow without any external factor and is a well known mechanism of IGW generation from observations, numerical and laboratory experiments (e.g. review by Plougonven and Zhang, 2014). Spontaneous emission can also lead to IGW emission by frontogenesis in regions of strong strains (Shakespeare and Taylor, 2015, 2016). Loss of balance (e.g. Molemaker et al., 2005) is a related mechanism where the balance of the flow can break down and transfer energy into unbalanced motions. Geostrophic adjustment (e.g. Rossby, 1938, Blumen, 1972, Bartello, 1995) is linear mechanism of wave generation, where an arbitrary initial condition applied to the flow forces the flow away from its balanced state, and by generating IGWs the flow then 'adjusts' to its balanced state. Note that geostrophic adjustment is different from spontaneous generation, in which the already balanced flow generates IGWs.

External mechanisms of IGW generation by balanced flows involve processes such as stimulated loss of balance, topographic interaction, and direct extraction. Externally forced

[3]Internal tides can be barotropic or baroclinic. Barotropic tidal currents, i.e. periodic water motions accompanying the tidal changes in sea level resemble barotropic tides.

waves can 'stimulate' the emission of IGWs from a balanced flow by stimulated loss of balance (e.g. Gertz and Straub, 2009; Xie and Vanneste, 2015). Upon interaction with topography, balanced flow can emit unbalanced IGWs in the form of arrested Kelvin waves (e.g. Dewar and Hogg, 2010) or internal lee waves (Bell, 1975; Nikurashin and Ferrari, 2011) when balanced eddy encounters a topographic obstacle. The presence of winds can also cause the balanced motions to transfer their energy directly to IGWs by direct extraction (Barkan et al., 2017).

While an understanding of IGW generation by balanced flows and wave interactions has expanded, the global estimates have been developed only for a few processes, such as internal lee wave generation (Nikurashin and Ferrari, 2011) and IGWs generated by surface wave interactions (Olbers and Eden, 2016). A full fledged picture of IGW generation remains missing and this is partly associated with the complex interactions of IGWs with motions of different scales (e.g. mesoscales, waves, small-scale turbulence), that render it hard to specify the sources of IGWs.

1.2.2 Dissipation and energy transfers

The dissipation of IGWs can occur by wave drag and wave breaking. Wave drag involves an exchange of energy wherein IGWs transfer their energy to the mean flow by gravity wave drag (Eden and Olbers, 2017). Wave breaking involves transfer of energy to small-scale turbulence and thereby contribute to diapycnal[4] (across isopycnal) mixing. The main mechanisms for IGW breaking are convective and shear instability (e.g. Thorpe, 2007; Alford and Pinkel, 2000; Staquet and Sommeria, 2002).

Convective instability can occur when IGWs with large amplitudes relative to their vertical and horizontal wavelengths, carry denser water over lighter water as they propagate. The denser water being heavier pushes the isopycnals downwards generating an overturn of isopycnals. This overturning eventually causes the IGWs to break and in the breaking process IGWs mix the fluid patch in its vicinity. Shear instability occurs when shear effects dominate stratification effects, characterized by $\mathrm{Ri} < 1/4$, and can thus destabilize the stratification. The shear associated with IGWs can result in the release of kinetic energy (KE) which causes the waves to disintegrate and break. Another mechanism of wave breaking is overturning by self acceleration, which is similar to convective instability but it

[4]The isopycnals in the ocean are fairly horizontal and so across isopycnal mixing is considered analogous to vertical mixing.

is the wave induced mean flow that advects dense fluid over light resulting in overturning, rather than the wave itself (Sutherland, 2010).

Wave-wave interactions as a consequence of resonant coupling can lead to energy exchange between waves as well as between waves and smaller scales. Two waves can interact to form a third wave and if the frequency of the new wave matches the natural frequency of the wave resulting from the combination of original waves, resonance occurs and the three waves are said to form a resonant triad (e.g. Gill, 1982; Pedlosky, 1987; Olbers et al., 2012). Many such combinations of interacting triads can occur, which can be largely explained by three mechanisms suggested by McComas and Bretherton (1977): induced diffusion, elastic scattering, parametric subharmonic instability (see also Gill, 1982; Olbers et al., 2012). Some of these triads can transfer energy to smaller scales, such as by induced diffusion (e.g. Müller et al., 1986 review) and by parametric subharmonic instability (e.g. Sutherland, 2010). However, energy transfer through wave-wave triadic interactions is a complex process and therefore much more complicated to identify and quantify.

1.2.3 Role of internal gravity waves in mixing, ocean circulation, and climate

IGWs influence processes and systems stretching far beyond their own scales. Their sphere of influence encompasses markedly diverse physical, chemical, and biological systems and a broad range of processes from small-scale turbulence to the long climatic scales.

Ocean mixing: biological, chemical, and physical impacts

Breaking IGWs play a leading role in driving diapycnal mixing in the ocean interior, as is well known from theory, models, observations, and laboratory experiments (e.g. review by MacKinnon et al., 2017). This diapycnal mixing is significantly important for the transfer and redistribution of heat and salt in the ocean interior. Furthermore, this mixing contributes to the vertical transport of nutrients, dissolved carbon-dioxide (CO_2), carbonates, and other components of the oceanic carbon system as well as other chemical species, that are important for the marine life, ocean circulation, as well as the global climate. In particular, IGW induced vertical mixing aids in the upward transport of nutrients into the upper ocean and thereby influences the biological productivity. IGWs have also been found to aid in coral health recovery by causing mixing which reduces the coral's thermal stress under increasing global temperatures (Schmidt et al., 2016). By mixing, IGWs lead

to dispersion of pollutants (e.g. Gregg, 1987) as well as other dissolved material in the ocean (e.g. James, 2002). In addition, breaking of IGWs can modify sediment transport rates (e.g. Heathershaw, 1985) over long timescales.

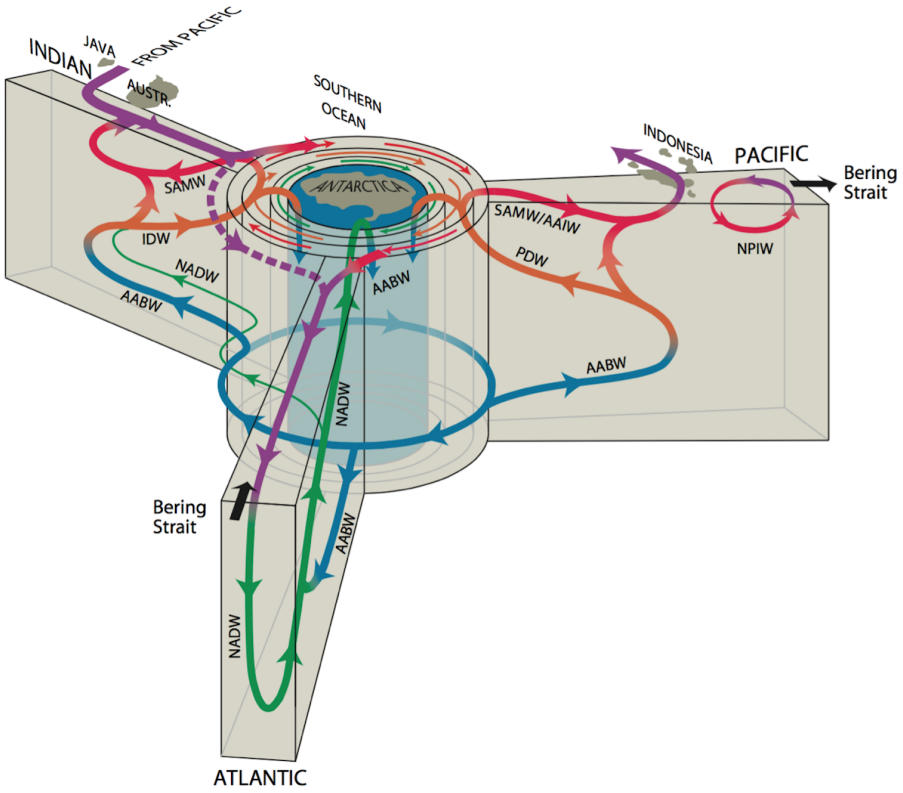


Figure 1.4: Schematic of the global meridional overturning circulation (MOC) from a Southern Ocean (SO) perspective (the cylinder in the middle). The SO connects all the major oceanic divisions: the Pacific, Atlantic, and Indian. The Atlantic meridional overturning circulation (AMOC) consists of two two overturning cells: a deep North Atlantic Deep Water (NADW) (green lines) and an abyssal Antarctic Bottom Water (AABW) (blue lines). The NADW, fresher and lighter than AABW, sinks in the North Atlantic and is upwelled by the action of the wind stress over the SO. A part of the upwelled NADW converts to AABW in the abyssal ocean close to Antarctica (blue cylinder). This AABW undergoes a mixing driven overturning spanning over the three basins of the global ocean as indicated in the figure. The shallower cell of NADW which upwells in the SO is largely controlled by mesoscale eddies. In the figure, thermocline waters are indicated by purple lines (cf. Talley, 2013 for details of other water masses). The schematic originally by Gordon (1986a,b) and Schmitz Jr (1996) has undergone several modifications. This schematic version is taken from Talley (2013).

The Meridional Overturning Circulation and the climate

IGW driven diapycnal mixing has been identified as one of the pivotal factors in the ocean's meridional overturning circulation (MOC) (Stommel, 1961; Munk and Wunsch, 1998; Staquet and Sommeria, 2002; Wunsch and Ferrari, 2004; Kuhlbrodt et al., 2007) (the other factor being wind-driven upwelling). The MOC refers to the circulation of oceanic water masses in the meridional-vertical plane as shown in Fig. 1.4.

A simplistic view of the mixing-driven branch of the MOC is as follows: wind-driven surface currents, such as the Gulf Stream, transport warmer and lighter surface waters polewards where the waters become denser due to cooling and sink into the ocean interior and the resulting deeper watermasses are subject to mixing during their return in the MOC loop; this mixing lightens the deep water masses and causes them to eventually rise at low latitudes to wind up the meridional loop. The other major branch of the MOC is the wind-driven upwelling in the Southern Ocean (e.g. Toggweiler and Samuels, 1993, 1995, 1998; Kuhlbrodt et al., 2007; Marshall and Speer, 2012). The strong westerly winds result in a substantial northward transport of surface waters by Ekman transport, and the associated horizontal divergence results in an upwelling in the vicinity of the Drake passage (e.g. Toggweiler and Samuels, 1995; Kuhlbrodt et al., 2007). Subsequently, the MOC predominantly dictates the heat and salt transport by the ocean, the ocean stratification, as well as the storage and recirculation of several chemical species such as carbon. These processes actively impact the weather and climatic patterns bringing the MOC to the forefront of the climate framework, underpinned by the actions of IGWs.

However, the understanding of this mixing-driven branch of the MOC presents difficulties since the mixing processes and estimates of mixing rates in the ocean remain poorly understood. Mixing is highly localized in space and time and the MOC might respond differently with the use of localized mixing rates than the uniform ones (e.g. Saenko, 2006). The effects of unresolved IGW dissipation in the ocean general circulation models are parameterized using prescribed values of mixing diffusivities which are sometimes arbitrarily chosen. In some models these diffusivities are related to resolved processes, for example stratification (Cummins et al., 1990) or energy input by tides (Simmons et al., 2004), but they represent only a subset of mixing processes.

Consequently, different models show different MOC strengths which hampers a robust interpretation of MOC projections, a key issue for this discrepancy being the representation of vertical mixing in the models (e.g. Randall et al., 2007). Clearly, for a consistent mixing parameterization a detailed understanding of the contributing processes to mixing— which

are mainly IGWs— is indispensable. Along this direction, a consistent IGW parametrization to represent IGW dissipation and mixing effects called IDEMIX (Internal gravity wave Dissipation, Energy and MIX-ing) has been recently developed (Olbers and Eden, 2013, 2017; Eden and Olbers, 2014, 2017) that includes energy exchange with the mean flow as well as non-linear wave-wave interactions. However, uncertainties related to the sources of IGW energy, such as balanced flows, still persist and a precise representation of these processes is missing. To account for this, an improved understanding of the energy pathways linking balanced flows and IGWs is clearly necessary.

Ocean's energy budget and energy sink for balanced motions

IGWs are an integral component of the ocean's energy budget and are a key player in the energetics of the ocean. The potential energy (PE) supplied to the ocean at large scales by the wind stress and the resulting Ekman pumping from the atmosphere is partly lost to the eddies by baroclinic or other instabilities in the form of KE. Large-scale PE is supplied by IGW mixing, IGWs in turn are driven by tides to a large extent (e.g. Eden, 2016). In ocean models where eddies are not resolved, the effects of eddies are often parameterized in an energy inconsistent way, and the eddy kinetic energy (EKE) and eddy potential energy (EPE) is simply lost. However, the dissipated EKE can potentially be used through IGWs for mixing and driving the deep overturning circulation. The eddies can lose their energy to IGWs through various mechanisms described in Section 1.2.1. In this way, IGWs can act as an energy sink for the eddies and may provide an important pathway linking the eddies down to the scales of molecular mixing (e.g. Bartello, 1995; Brüggemann and Eden, 2015 and Chapters 2 and 3 of this thesis).

An excursion back to the opening Section 1.1 tells us that besides the unbalanced IGW motions the geophysical flows also contain balanced motions. The picture of the global ocean remains largely incomplete without the ubiquitous mesoscale eddy field which contain most of the ocean's KE and are the most energetic component of the ocean's energy cycle. Mesoscale eddies interact and exchange energy with a variety of processes including IGWs. The balanced mesoscale field can generate IGWs by numerous mechanisms (Section 1.2.1) and the generated IGWs can in turn act as a potential energy sink to the balanced mesoscale field, as discussed above. This interlinking between balanced mesoscale eddies and unbalanced IGWs forms an important energy pathway and has crucial implications for the ocean's energy budget. To better understand the energy conversions between the mesoscale eddies and IGWs, an understanding of the dynamics and energy sources and

sinks of mesoscale eddies is crucial. The next section presents an overview of the dynamics of oceanic mesoscale eddies.

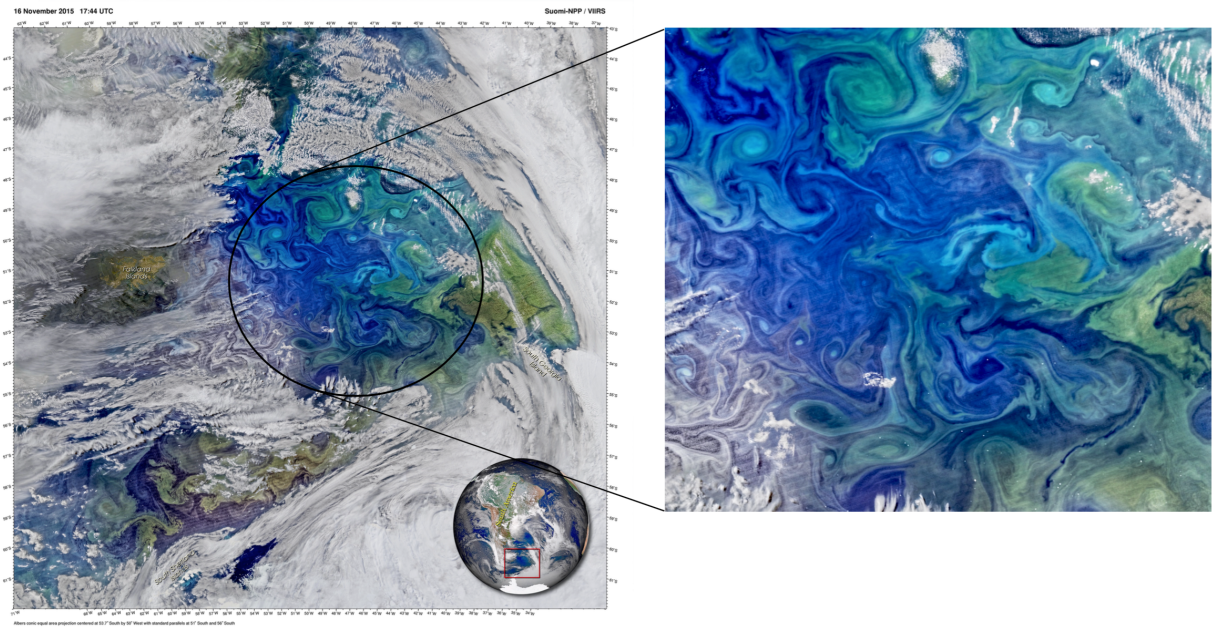


Figure 1.5: The South Atlantic in springtime near the Drake passage. (a) A satellite image of phytoplankton communities between the islands to the east of Argentina: the Falkland Islands to the west and South Georgia Island to the east, captured in November 2015 during spring in the Southern Hemisphere. This region also belongs to the Southern Ocean whose dynamics are largely controlled by eddies, as seen in the region enclosed by the circle. (b) An enlarged view of the eddy structures in the ocean. Clearly seen is a field of huge mesoscale eddies, about 100 km in size, and smaller sub-mesoscale eddies painted with rich hues by the phytoplankton communities imparting different shades to the region (owing to different chlorophyll concentrations). Eddies can vertically transport nutrient-rich water necessary to sustain phytoplanktons, as well as laterally transport with them several species of marine organisms which feed on phytoplanktons and other organisms, thus sustaining marine ecosystems. Eddies are aptly called 'the pastures of the ocean'. Image courtesy of NASA.

1.3 Dynamics of mesoscale eddies

The ocean is a sea of vortices. Vortices are whirling water masses, similar to little eddies in a stream. Mesoscale eddies are such vortices in the ocean with typical horizontal spatial scales of about $50 - 100$ km and evolve over timescales of weeks to several months. They are three dimensional structures and can occur in the upper, interior, as well as the deep

ocean. A satellite image of eddies in the South Atlantic is shown in Fig. 1.5[5]. A field of eddies coloured by various phytoplankton communities is clearly seen.

The vertical scales of the eddies, which may extend up to a few kilometers, are dwarfed by the substantially larger horizontal scales, and so the eddies can be considered as quasi two-dimensional features resembling a pancake-like structure. Mesoscale eddies belong to the class of balanced motions and are governed by the dynamics of Rossby waves (Section 1.1). Consequently, the eddy related fluctuations can be thought of as interaction between Rossby waves as described in the next section.

1.3.1 Generation mechanisms

Mesoscale eddies in the ocean derive their energy primarily through the barotropic and baroclinic instabilities of Rossby waves (e.g. Pedlosky, 1987; Vallis, 2006; Olbers et al., 2012).

Barotropic instability

Barotropic instability draws its energy from the horizontal shear of the currents which may be released by turbulent fluctuations to generate eddies. A barotropic fluid features isopycnals parallel to isobars which can also occur in a non-stratified fluid. In a fluid which is stratified, barotropic instability can co-exist with baroclinic instability. However, barotropic instability involves a KE exchange, unlike baroclinic instability which involves an exchange of PE.

Baroclinic instability

Baroclinic instability is relevant to rapidly rotating, strongly stratified flows with a meridional and vertical temperature gradient. Baroclinic instability feeds on the available potential energy[6] (APE) manifested in the horizontal density gradients. Since the isopycnals exhibit a tilting in the vertical for a baroclinic fluid, the associated instability generates eddies by the slumping of horizontal density gradients and converts the APE into the KE of the mesoscale eddies. The theory can be explained by the classical models of baroclinic instability: Eady model on the f-plane (Eady, 1949), and β-plane models of Charney and Phillips (Charney, 1947; Phillips, 1954). A variant of the Eady model, which is based on quasi-geostrophic equations, is developed by Stone (1966) for primitive equations. Such a

[5]The image is a courtesy of NASA/Ocean Biology Processing Group, NASA Goddard Space Flight Center, and is taken from https://oceancolor.gsfc.nasa.gov/feature/images/V2015320174400.SouthAtlantic.jpg

[6]Available potential energy is that part of the total potential energy which is 'available' for conversion into kinetic energy (cf. Lorenz, 1955; Vallis, 2006; Olbers et al., 2012).

model is implemented in this thesis to represent an idealized baroclinically unstable flow over a wide range of dynamical conditions from Ri $\gg 1$ to Ro $= \mathscr{O}(1)$.

Baroclinic instability is dominated by balanced flows and can exist for different dynamical regimes which can be distinguished by Ri of the flow: quasi-geostrophic baroclinic instability characterized by Ri $\gg 1$ (or Ro $\ll 1$) and ageostrophic baroclinic instability characterized by Ri $= \mathscr{O}(1)$ (or Ro < 1). However, several other kinds of instabilities pervade the ocean associated with smaller scales, such as sub-mesoscales. Sub-mesoscale motions or eddies are a characteristic feature of the upper-ocean and predominantly occur in the ocean mixed layer where departures from quasi-geostrophic balance give way to ageostrophic dynamics (although such departures might also occur in the ocean interior). For an extreme case with Ri < 1, other kinds of ageostrophic instabilities appear (e.g. Stone, 1966, 1971; Haine and Marshall, 1998; Fox-Kemper et al., 2008; Thomas et al., 2008) such as symmetric instability and Kelvin-Helmholtz instability for Ri < 0.25 and in these scenarios the unbalanced dynamics dominate.

1.3.2 Dissipation and energy transfers

Mesoscale eddies act as a vast reservoir of energy that mainly enters the ocean through solar heating, winds, tides, and geothermal heating. In particular, most of the energy input by winds is stored as PE that feeds eddies and other instabilities, while a minor portion is used for mixing (Wunsch and Ferrari, 2004). For this reason the dissipation of this vast reservoir of eddy energy is of tremendous concern, but the current understanding about the dissipation of eddy energy is much less.

The quasi two-dimensional structure of mesoscale eddies determines their mode of dissipation. In line with that, these eddies exhibit geostrophic turbulence (Charney, 1971), or two-dimensional turbulence, rather than the classical three-dimensional turbulence. Unlike three-dimensional turbulence in which KE is fluxed from larger to smaller scales dissipating finally at viscous molecular scales (Kolmogorov, 1941), the energy transfer in two-dimensional turbulence predominantly occurs from smaller scales towards large scales (Batchelor, 1969; Charney, 1971; Rhines, 1977). Thus, the former features a downscale energy transfer, whereas the latter an upscale energy transfer.

Mesoscale eddies are known to transfer their KE back to the large scales via upscale energy transfer. Evidences for this process in the ocean have been presented from satellite altimeter data (Scott and Wang, 2005) as well as models (Schlösser and Eden, 2007). This process falls short, however, to explain how the energy contained in this massive energy

reservoir is lost to the smallest molecular scales where the ocean's energy finally dissipates. For this to occur, these eddies must take a downscale route to dissipation.

Mesoscale eddies can directly transfer energy to small-scale turbulence upon interaction with rough ocean bottom through bottom friction in a turbulent bottom boundary layer (Arbic et al., 2009). Further potential sinks are given by the mechanisms of IGW generation described in Section 1.2.1. The IGWs eventually break and contribute to mixing wherein the energy is finally viscously dissipated. Yet another mechanism is the exchange of energy between balanced eddies and IGWs by wave-mean flow interaction resulting from the gravity wave drag (cf. Eden and Olbers, 2017). The conversion of the eddy kinetic energy (EKE) to the energy of IGWs is also crucial for the energy budget of IGWs, as suggested by the recent model-based estimates of small-scale dissipation rates which show digressions from observations without EKE effects taken into account (Pollmann et al. (2017)).

While global estimates of mesoscale eddy energy dissipation due to bottom friction and topographic interaction by lee wave generation have been developed (e.g. Arbic et al., 2009; Nikurashin and Ferrari, 2011), these processes appear still too small to act as a major eddy energy sink (Ferrari and Wunsch, 2009). On the other hand, eddy energy directly lost to IGWs has been identified as a potential candidate (e.g. Wunsch and Ferrari, 2004) but the global estimates of this energy sink remain unidentified. This thesis is an attempt in this direction, and chapters 2 and 3 discuss the role of IGWs in the dissipation of balanced eddy energy during the downscale transfer of energy.

1.3.3 Impact of mesoscale eddies on ocean circulation and climate

Mesoscale eddies are the "weather" of the ocean and are analogous to atmospheric highs and lows on a weather map. They drive the large-scale ocean circulation, sustain marine ecosystems, and profoundly impact the weather and climate of the Earth.

Ocean's transport system, marine ecosystems, and carbon-dioxide sink

Mesoscale eddies can act as a transport system of the ocean transporting heat and salt across huge horizontal distances owing to their huge horizontal scales. As they move along isopycnals, the eddies also transport water masses between the surface and the deeper layers of the ocean. By doing so, the mesoscale eddies help establish the ocean's circulation and promote large-scale mixing by distribution of different water masses. In this way, the eddies

determine the water mass properties in different regions of the global ocean impacting the ocean as well as the climate both locally and globally.

Moreover, eddies act as a huge sink for the anthropogenic CO_2, which has marked implications for the increasingly warming global climate. The eddies are especially related to the CO_2 sink in the Southern Ocean (e.g. Sallée et al., 2012) which stores most of the anthropogenic CO_2 taken up by the global ocean. Modeling results also show that eddies affect the CO_2 storage and induces changes in the overturning circulation (e.g. Munday et al., 2014). In addition, mesoscale eddies transport nutrients as well as marine organisms and are thus of central importance in the growth and sustenance of the ocean's primary production, linking the base of the oceanic food chain to more complex marine ecosystems (e.g. Condie and Condie, 2016; Shulzitski et al., 2016).

Southern Ocean, the Meridional Overturning Circulation, and the climate

The Southern Ocean (SO) is a salient component of the global overturning circulation of the ocean and one of its prominent features is the inter-basin connection, connecting the principal oceanic divisions: the Pacific, Atlantic, and Indian Oceans (cf. Fig. 1.4). In this way, the SO circulation acts like a "steering wheel" of the global MOC by circulating heat, salt, CO_2, biological, and chemical elements among the major oceanic branches. The SO circulation is a result of several dynamical features at play, but the most prominent ones are the Antarctic Circumpolar Current (ACC), mesoscale eddies, and the Ekman transport (e.g. Olbers et al., 2012).

The ACC, the largest oceanic current resulting due to the intense westerly winds at the "roaring forties and furious fifties" latitudes, i.e. 40°S and 50°S, contributes to significant transports of heat, salt, and tracers eastward (e.g. Vallis, 2006; Olbers et al., 2012). But the lack of zonal continental boundaries in the SO does not allow for meridional transports by the ACC, which is fundamental to sustain the global MOC. It is mainly the eddies which drive a poleward volume transport opposing the equatorward wind driven Ekman transports. The sum of Ekman and eddy-driven transport establish the MOC in the SO. In addition, the eddies also aid the vertical transport of momentum into the deep ocean.

Mesoscale eddies play a similar (but less prominent) role also in the circulation of other oceanic basins, and hence in the overall dynamics of the global MOC. The meridional transport of heat by mesoscale eddies impacts the overlying atmosphere which in turn governs wind and precipitation patterns influencing weather on a global scale. Clearly, the changes in the eddy transport would significantly impact the MOC and hence the climate.

Although eddies play a momentous role in the ocean's energetics, they are not resolved or resolved partially in OGCMs or climate models because of which the representation of eddy effects in the models relies on parameterizations. Current mesoscale eddy parameterizations account for certain aspects of the effects of eddies, such as isopycnal mixing; but other major aspects which modify eddy energy such as the interaction and exchange of eddy momentum with the mean flow and the interlinking of mesoscale eddies with IGWs, remain unclear and pose a challenge for a precise parameterization of these processes and consequently affect the interpretations of future climate projections in the models.

1.4 Specific questions addressed in this thesis

The question of the downscale energy transfer from balanced flows to the smallest oceanic scales is a difficult one. It gets complicated further because the preferred energy transfer by the balanced flows is in the opposite direction, i.e. they feature an upscale energy transfer (Charney, 1971, also Section 1.3.2). This general picture of the balanced flows, however, changes with the flow characteristics estimated by the dimensionless parameters Ri and Ro described in Section 1.1. For a range of dynamical regimes of baroclinic instability ranging from ageostrophic (Ri $= \mathcal{O}(1)$ or Ro < 1) to quasi-geostrophic (Ri $\gg 1$ or Ro $\ll 1$), Brüggemann and Eden (2015) discuss the spectral energy fluxes and find that the regimes with smaller and smaller Ri show stronger and stronger energy flux towards large wavenumbers, i.e. a more and more downscale transfer of energy. In view of this, Brüggemann and Eden (2015) have emphasized that a direct ageostrophic route can potentially facilitate the dissipation of balanced flows towards small scales. But specific processes involved in the downscale energy transfer via an ageostrophic route were not identified.

IGWs put forward a strong candidature for this task, and this thesis addresses the question if IGWs diagnosed from the two diagnostic tools (Sections 1.1.2 and 1.1.3) are indeed related to the ageostrophic route suggested by Brüggemann and Eden (2015) and if so, to what extent IGWs contribute the downscale energy transfer. But the detection of IGW signals is a fundamental issue because of their coupling with the balanced flows. This directs to the need of separating balanced flows and IGWs which is the central point addressed in this thesis by using a model setup similar to Brüggemann and Eden (2015). Further, the questions of the efficiency of the two diagnostic tools in diagnosing IGWs and

the strength of IGW emission for a range of baroclinic instability regimes as well as in different emission scenarios are considered. The thesis ultimately directs to the question of the role of IGWs in the downscale energy transfer in the ocean.

1.5 Thesis Overview

As outlined in this chapter, this thesis addresses the disentangling of IGWs generated from balanced flows and assesses the diagnosed IGW activity in different dynamical scenarios.

The Chapters 2 and 3 of this thesis comprise of two research papers written in the style of journal publications. As a consequence, they contain their own abstract, introduction and conclusions.

Chapter 2 consists of a research paper titled “Fast gravity waves and how to find them" submitted to the *Journal of Physical Oceanography* and is currently under review. It addresses the spontaneous emission of IGWs from baroclinic instability in an idealized numerical model in different dynamical regimes, from ageostrophic ($\mathrm{Ri} = \mathcal{O}(1)$) to quasi-geostrophic ($\mathrm{Ri} \gg 1$) and the role of IGWs in downscale energy transfer in the ocean, using modal decomposition based on the method proposed by Machenhauer (1977).

Chapter 3 consists of a research paper titled “Gravity wave emission by different mechanisms" in preparation to be submitted to the *Journal of Physical Oceanography*. It introduces a new diagnostic tool called the QG filter based on a potential vorticity inversion method to diagnose IGWs generated by different mechanisms: spontaneous emission, convective instability, lateral boundary instability.

Chapter 4 presents the key results and conclusions based on the research undertaken in the research papers. An outlook further discusses the results with respect to open questions and opens doors for new questions.

O' mighty ocean!

Thy tryst with winds and tides

Maketh waves far and wide

Speaketh but, O' mighty ocean!

Sans winds, sans tide

How doth thee maketh waves inside?

Chapter 2

Fast gravity waves and how to find them

This chapter is under review in the Journal of Physical Oceanography.

Citation: Chouksey, Manita, Carsten Eden, and Nils Brüggemann, 2018: Internal gravity wave emission in different dynamical regimes. J. Phys. Oceanogr.

Note: This is the version submitted to the journal and it might undergo some changes after revision. The final form after publication might appear different than the present one.

Abstract

We aim to diagnose internal gravity waves emitted from balanced flow and investigate their role in the downscale transfer of energy. We use an idealized numerical model to simulate a range of baroclinically unstable flows to mimic dynamical regimes ranging from ageostrophic to quasi-geostrophic flows. Wave-like signals present in the simulated flows, seen for instance in the vertical velocity, can be related to gravity wave activity identified by frequency and frequency-wavenumber spectra. To explicitly assign the energy contributions to the slow balanced and fast unbalanced gravity modes, we perform linear and non-linear modal decomposition to decompose the full state variable into its slow balanced and fast unbalanced gravity wave counterparts. The linear decomposition shows a reasonable separation of the slow and fast modes, but is limited in its application to a non-linear system. To account for the non-linearity in our system, we apply a decomposition using the normal mode initialization technique proposed by Machenhauer in 1977. Further, we assess the strength of the gravity wave activity and dissipation related to the decomposed modes for different dynamical regimes. The results show that there is much more gravity wave activity for an ageostrophic regime than for a quasi-geostrophic regime. Furthermore, internal gravity waves dissipate predominantly through small-scale dissipation. Thus, internal gravity waves could possibly be catalysed by ageostrophic baroclinic instability and might therefore contribute to a downscale energy transfer in the ocean.

2.1 Introduction

Internal gravity waves (IGWs) occur naturally in the atmospheric and oceanic flows and influence the atmosphere mainly by vertical momentum transport and the ocean by density mixing. Despite their ubiquity and importance in the geophysical flows and numerous observational and numerical studies, the emission and dissipation of IGWs are not well understood. Due to this inadequacy, the parameterization of IGWs remains a strenuous task. This gap is in part related to the short spatial and temporal scales of IGWs which render them hard to observe and difficult to resolve, in part to the specification of their sources, and to a certain extent to the difficulty of separating IGW from other motions. The latter is the theme of this paper, in which we diagnose IGWs emitted from balanced flows for different dynamical regimes and investigate their role in the downscale transfer of energy.

IGWs are forced mainly by orography, convection, and jet/front systems in the atmosphere, and winds and tides in the ocean. Of these, the emission of IGWs from geostrophically balanced flows, referred to as spontaneous generation (see Vanneste (2013), and references therein), is of marked interest as it provides an avenue to the internal mechanisms in the flow that lead to IGW generation. IGWs generated spontaneously have been discussed extensively in both oceanic and atmospheric literature: in observations (e.g. Section 2 of Plougonven and Zhang, 2014), laboratory experiments (e.g. Williams et al., 2008), and several numerical simulations (discussed below; also see reviews by Vanneste (2013) and Plougonven and Zhang (2014), and references therein). Spontaneous generation (or emission) is also of special interest for it allows for understanding the fundamental nature of balanced flows, which dominate much of the atmosphere and ocean.

The slow geostrophically balanced motions evolve over long timescales, whereas the IGW correspond to the fast unbalanced motions. The slowly varying nature of the balanced motion led to the concept of "slow manifold" (Leith, 1980; Lorenz, 1980, 1986) that is defined as a subspace of the phase space which is strictly "invariant" and completely devoid of any IGW activity. However, the routinely observed emission of unbalanced fast motion (IGW) in geophysical flows questions the validity of such a manifold. The existence of the slow manifold –or rather its non-existence– has been discussed by several authors (e.g. Lorenz and Krishnamurthy, 1987; Lorenz, 1992; Ford et al., 2000; Vanneste and Yavneh, 2004) which makes clear that an exactly invariant slow manifold does not exist but rather manifolds with different degrees of "invariance", where the slow balanced and fast unbalanced motions coexist but their degree of interaction differs. Accordingly, such a manifold more suitably came to be called as a fuzzy manifold (Warn, 1997) or a slow quasimanifold (Ford et al., 2000). The non-existence of an exactly invariant slow manifold implies that immaterial of the initially balanced conditions the slowly evolving balanced motion will always co-occur with the fast motion and hence "spontaneous generation of IGWs for geophysical flows is inevitable" (Vanneste, 2013). An insight into this mechanism also facilitates our understanding of the emission of IGWs from balanced mesoscale flows (mesoscale eddies) in the ocean, that has strong implications for the energy budget of the ocean.

Mesoscale eddies, a consequence of baroclinic instability, are ubiquitous in the ocean and are one of the most energetic components of the ocean energy budget. The eddies act as a reservoir of energy which enters the ocean at large scales, but energy in the ocean is finally dissipated at the viscous molecular scales via a downscale energy transfer. Balanced mesoscale flows are known to lose their energy to large oceanic scales through an upscale

energy transfer (e.g. Charney, 1971). How the balanced mesoscale eddies lose their energy to dissipative molecular scales is however unclear; this is where the unbalanced motions such as IGWs come into the picture. Studies suggest that balanced flows could lose their energy to unbalanced motions, like IGWs, through processes such as geostrophic adjustment, spontaneous generation, loss of balance, stimulated loss of balance, topographic interaction, direct extraction, gravity wave drag. Geostrophic adjustment (e.g. Rossby, 1938; Blumen, 1972; Bartello, 1995), differs from previously discussed spontaneous generation in that for the former process the flow is forced away from its balanced state by an arbitrary initial condition and the flow then adjusts to its balanced state (geostrophy) while emitting unbalanced IGWs. Another mechanism is the loss of balance (e.g. Molemaker et al., 2005) where the balance of the flow can break down and transfer energy into unbalanced motions. Stimulated loss of balance (e.g. Gertz and Straub, 2009; Xie and Vanneste, 2015) is different from loss of balance and it refers to the process by which externally forced waves can further 'stimulate' the emission of waves and this wave-mean interaction extracts energy from the balanced flow. Topographic interaction (e.g. Dewar and Hogg, 2010) requires that the balanced flow interact with topography and in the process emit unbalanced motions, whereas the energy transfer from mesoscales directly to IGWs has been referred to as direct extraction and was discussed in the context of a wind-driven channel flow by Barkan et al. (2017). Yet another mechanism is the exchange of energy between balanced flow and gravity waves by wave-mean flow interaction resulting from the gravity wave drag (see Eden and Olbers, 2017). In these ways the balanced motions could find an energy pathway via unbalanced motions *en route* to viscous dissipation, resulting in the downscale transfer of energy. The idea is further motivated by numerous atmospheric and oceanic studies and there seems to be a general consensus on this notion; some of these studies are briefly mentioned in the following passage.

The loss of balance in a baroclinically unstable flow results in a transfer of energy from balanced mean flow to unbalanced motions and eventually to dissipation by means of a downscale energy transfer as shown by Molemaker et al. (2005). A direct interior route to dissipation by means of unbalanced motions in a Boussinesq flow that can initiate a downscale energy transfer has been discussed in an idealized flow configuration by Molemaker et al. (2010). For a range of dynamical regimes, downscale energy transfer from baroclinically unstable flows in an idealized setting has been discussed by Brüggemann and Eden (2015), and they emphasize that an ageostrophic direct route to dissipation might be of importance in the energy budget of the ocean. For a realistic flow configuration, Capet et al. (2008)c find an increase in the downscale energy flux related to the ageostrophic dis-

sipation route with an increased horizontal resolution that can better resolve ageostrophic dynamics. Another unbalanced route is the stimulated loss of balance that can propel the downscale energy transfer as shown by Gertz and Straub (2009) for an unstratified thin-aspect-ratio fluid. Spontaneous generation of IGWs from idealized dipoles, e.g. Sugimoto and Plougonven (2016), also suggest that the balanced flows could dissipate via unbalanced IGWs during a downscale energy transfer. An alternative perspective that links IGWs and balanced mesoscale eddies are the recent model-based estimates of small-scale dissipation rates in the ocean, which fail to reproduce the observations without eddy forcing taken into account in a parameterization of IGW, implying that eddy dissipation is necessary for the IGW energy budget (Pollmann et al. (2017)).

Despite numerous observational and numerical studies on IGW emission from balanced motions, the exact mechanism behind this process remains poorly understood and a puzzling part is the identification of IGWs. The complication related to the identification of the IGW signals is in part associated with the coupling of the balanced motions and IGWs. The extent of this coupling can be estimated by Rossby number (Ro) (or an equivalent Richardson number (Ri)) which is a measure of the timescale separation between the slow balanced and fast unbalanced IGW motions. This coupling is weak for a regime with Ro $\ll 1$, equivalent of a Ri $\gg 1$, such that the fast and slow motions are "well separated"; on the contrary, for a large Ro, equivalent to Ri $= \mathcal{O}(1)$ or smaller, the fast and slow motions interact more strongly and the separation of these processes is not well defined (Vanneste, 2013; Zeitlin, 2008). The coupling between these motions adds to the intricacy of separation and detection of IGW signals from the balanced flow field. The emphasis of this paper is on the diagnosis of gravity wave signals.

The traditional approach to identify IGWs, or more generally unbalanced motions, is to use the fast fields such as the horizontal velocity divergence, the vertical component of the vorticity to determine the IGWs (e.g. Plougonven et al., 2005; Plougonven and Snyder, 2007), or a spatial filtering to obtain small-scale vertical velocity as the signature of the IGWs (e.g. Sugimoto and Plougonven, 2016). As another way, the full field of interest could be separated into horizontally non-divergent and vertically irrotational components (e.g. Molemaker et al., 2005) or simply put into rotational and divergent parts (Molemaker et al., 2010; Brüggemann and Eden, 2015) which give the balanced and unbalanced contributions respectively. Another technique to obtain IGW signals is from the quasi-geostrophic omega equation which gives the ageostrophic vertical velocity whose difference with the full vertical velocity yields the unbalanced IGW contribution (e.g. Danioux et al., 2012; Nagai et al., 2015). Although these methods work well in identifying IGWs owing to the waves' distinct

spatial characteristics, the methods have the restraint that the unbalanced part interpreted as the IGW could still contain a notable amount of balanced part which hampers a concise interpretation of the signals.

A conceptually different approach is the linear modal decomposition, similar to projection onto the balanced manifold, which separates the slow balanced and fast IGW modes by decomposing the full field. Such a decomposition has been implemented for a linear system, for example by Molemaker et al. (2010) and Borchert et al. (2014). However, the application of a linear decomposition to a nonlinear system is hindered by its limitation to account for the non-linear evolution, and thus an extension of this method to a non-linear framework is desirable. This issue is addressed by a non-linear normal mode initialization technique (NLNMI) developed independently by Machenhauer (1977) and Baer and Tribbia (1977), that allows for adjustments to the initial conditions in order to minimize the tendency of the system to generate fast motions. Relating these initialization procedures to quasi-geostrophic balance, Leith (1980) derived decomposed modes for the hydrostatic Boussinesq equations, which was later generalised to the non-hydrostatic set of equations by Bartello (1995). More recently, the non-linear initialisation scheme of Baer and Tribbia (1977), was applied by Kafiabad and Bartello (2016) for balanced rotating dynamics to identify the energy cascades for differently initialized balanced regimes, and by Kafiabad and Bartello (2017) to identify mechanism and scales of spontaneous imbalance in a rotating stratified turbulence system. In this paper, we apply the non-linear initialization procedure of Machenhauer (1977) with an aim to diagnose IGWs by decomposing the full field into its balanced and unbalanced counterparts, for a range of dynamical regimes from ageostrophic to quasi-geostrophic.

The present work is also motivated by the previous work of Brüggemann and Eden (2015) who discussed spectral energy fluxes for a range of dynamical regimes ranging from ageostrophic (small Ri) to quasi-geostrophic (larger Ri). The regimes with smaller and smaller Ri show stronger and stronger energy flux towards large wavenumbers, i.e. a larger and larger downscale transfer of energy (although there is also still an inverse energy transfer for the smaller wavenumber range). We use a very similar model setup as Brüggemann and Eden (2015) and by diagnosing gravity waves we also aim to answer the question whether this ageostrophic route towards dissipation is generated by gravity wave emission during ageostrophic baroclinic instability.

To weave together the numerous threads sketched up above, we use a simple model of baroclinic instability to simulate flows representing low to high Ri regimes from ageostrophic to quasi-geostrophic (Section 2.2). This allows us to study the evolution of the

flows individually and evaluate the characteristics of the unbalanced motions without interference from other processes as is the case in a more complex flow configuration. To characterize the waves we first explore them in the Fourier space (Section 2.3). We then apply Machenhauer's non-linear initialization technique to the model data to obtain balanced and unbalanced modes and analyze gravity wave activity for different flow regimes (Section 2.4). Further we estimate the energy dissipation related to balanced and unbalanced modes for different dynamical regimes under study (Section 2.5). Finally, we summarize the results and the key conclusions (Section 2.6).

2.2 Baroclinic instability in different dynamical regimes

2.2.1 Numerical model

Baroclinic instability in our setup is represented in an idealized channel configuration resembling the classical Eady model (Eady, 1949) of baroclinic instability: flow on a f-plane with a prescribed stratification and a vertically sheared background zonal flow, under Boussinesq and hydrostatic approximations. Our model configuration differs from the Eady model by using a free surface and a meridional buoyancy gradient which is expressed as a sine function with an amplitude of M_0^2. The latter allows to apply double periodic boundaries to exclude lateral boundary instabilities. The presence of boundaries itself can lead to IGW generation (e.g. Borchert et al., 2014, laboratory experiments by Williams et al., 2008) and we suppress it with double periodic boundaries since we focus on studying IGW emission from balanced flows. In its initial state, the model has a background flow in thermal wind balance with a constant stratification N_0^2 and a meridional buoyancy, this combination makes the setup baroclinically unstable. An example of the initial state temperature of the setup is shown in Fig. 2.1(a). Temperature is the only active tracer in our setup and hence temperature and buoyancy are equivalent. The model is forced with a restoring of the zonal mean flow and zonal mean buoyancy towards the initial state, there is no additional surface forcing and no bottom friction.

We simulate baroclinic instability for a range of dynamical regimes characterized by the Richardson number (Ri) which is defined here as the ratio of the vertical density stratification and vertical shear of the horizontal velocity. The parameters Ri, N_0 and M_0 are related as:

$$\mathrm{Ri} = \frac{{N_0}^2 f^2}{{M_0}^4} \tag{2.1}$$

where f the Coriolis parameter. Ri sets the initial buoyancy restoring in the model, and hence the flow dynamics ranging from ageostrophic ($\mathrm{Ri} = \mathscr{O}(1)$) to quasi-geostrophic ($\mathrm{Ri} \gg 1$) as described in Brüggemann and Eden (2015). The potential energy (PE) deposited into the model by the restoring is lost to the kinetic energy (KE) predominantly by baroclinic instability, as in the Eady model. This energy is dissipated in the model by lateral biharmonic (A_h) and vertical harmonic friction (A_v) at small scales and by a linear drag acting on the zonal mean zonal flow at the large scales (as in Brüggemann and Eden, 2015). Dissipation by biharmonic and harmonic friction is controlled by a grid Ekman number (Ek) which is set to $Ek = 0.1,\ 0.06$ for $\mathrm{Ri} = 3, 13$ and $Ek = 0.01$ for $\mathrm{Ri} = 377, 915$. The time scale of the linear drag λ_u and the relaxation time scale λ_T are set proportional to the time scale of the fastest growing modes (see Table 2.1). The numerical code for the model is identical to the one in Eden (2016).

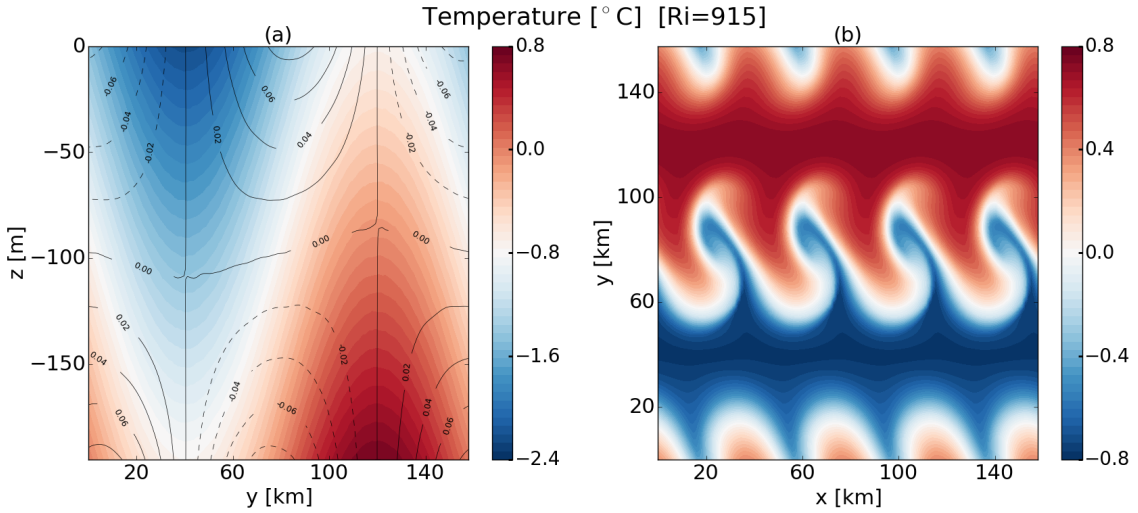


Figure 2.1: Temperature field shown at different instances of the model run for Ri = 915 from the standard setup. (a) Temperature at the initial state (time=0) of the model with zonal velocity contours on top (10 days later) shown in black. (b) Temperature at the surface after 32 days, notice the four wavelengths of the fastest growing mode.. Note that the vertical axis in (a) is depth and in (b) the meridional extent.

A small random initial perturbation provided in temperature grows exponentially with time as the baroclinic instability sets in. For each stratification and shear, there exists a particular perturbation of a certain spatial scale which grows faster than perturbation of other scales. This particular spatial scale is of the order of the deformation radius (Eady, 1949; Stone, 1966) and is referred to as the fastest growing mode, which becomes dominant

over all the other perturbations and is therefore assumed to be the mode at which most of the PE is converted to KE. The linear growth rate σ_{max} and the corresponding wavenumber k_{max} of the fastest growing mode can be expressed for the primitive equations as derived by Stone (1966):

$$k_{max} \approx \sqrt{\frac{\mathrm{Ri}}{1+\mathrm{Ri}}}\sqrt{\frac{5}{2}}L_r^{-1}, \;\; \sigma_{max} \approx \sqrt{\frac{\mathrm{Ri}}{1+\mathrm{Ri}}}\sqrt{\frac{5}{54}}\frac{f}{\sqrt{\mathrm{Ri}}} \tag{2.2}$$

For quasi-geostrophic approximation (large Ri), the length and time scales of this fastest growing mode turn into expressions as derived by Eady (1949):

$$k_{max} \approx 1.61 \; L_r^{-1}, \;\; \sigma_{max} \approx 0.31 \; \frac{f}{\sqrt{\mathrm{Ri}}} \tag{2.3}$$

where $L_r = N_0 H/f$ is the Rossby radius of deformation. Since our model is based on primitive equations, in our simulations we use Stone's formulation and the model domain allows for four wavelengths of the fastest growing mode ($4 \times 2\pi/k_{max}$) (for example, see Fig. 2.1(b)). The domain width is chosen to be equal in zonal and meridional directions and so $L_x = L_y$. In our standard model setup the number of grid points in both horizontal directions is $nx(= ny)$ =120. Note, however that the actual horizontal resolution, which determines the smallest resolved scales, depends on k_{max} and varies as we vary Ri for different simulations. On the contrary, a fixed vertical depth of $H = 200$ m with $nz = 40$ layers for the standard setup provides a constant vertical resolution of 5m for all simulations. In addition, we also run simulations at a higher resolution with 240 and 80 grid points in the horizontal and vertical respectively, the vertical resolution in this case is 2.5m. We use our standard model setup for the bulk of our analysis, and we mention the resolution used for the respective diagnosis in the text or the figure caption. The model time step depends on the Courant-Friedrichs-Lewy (CFL) condition, mean flow, and the horizontal resolution. CFL is set to 0.001, 0.003, 0.005, and 0.005 for Ri = 3, 13, 377, and 915 respectively. An overview of the model parameters for the standard setup is presented in Table 2.1. Next we discuss the numerical simulations used in this work.

Symbol	*Meaning*	*Value*	*Unit*
nx, ny, nz	Number of grid points in (x, y, z) directions	120, 120, 40	-
H	Depth of the basin	200	m
α	Aspect ratio	0.02	-
f	Coriolis frequency	7×10^{-5}	s^{-1}
N_0	Brunt-Väisälä frequency	f/α	s^{-1}
M_0	Meridional stratification	$\sqrt{fU_0/H}$	s^{-1}
U_0	Mean flow	$N_0 H/\sqrt{\mathrm{Ri}}$	m s^{-1}
L_r	Rossby radius of deformation	$N_0 H/f$	m
k_{max}	Wavelength of the fastest growing mode	$\sqrt{5/2}\sqrt{\mathrm{Ri}/(1+\mathrm{Ri})}\ L_r^{-1}$	m^{-1}
σ_{max}	Growth rate of the fastest growing mode	$\sqrt{5/54}\sqrt{\mathrm{Ri}/(1+\mathrm{Ri})}\ f/\sqrt{\mathrm{Ri}}$	s^{-1}
L_x	Length of the domain	$4 \times 2\pi/k_{max}$	m
dx(=dy)	Horizontal resolution	L_x/nx	m
dz	Vertical resolution	H/nz	m
dt	Time step	dx CFL$/U_0$	s
A_h	Biharmonic horizontal friction	Ek $f dx^4$	m^2 s^{-1}
A_v	Harmonic vertical friction	Ek $f dz^2$	m^2 s^{-1}
λ_u	Linear drag coefficient	0.75 σ_{max}	s^{-1}
λ_T	Restoring time scale	2 σ_{max}	s^{-1}

Table 2.1: An overview of the model parameters for our standard model setup.

2.2.2 Numerical simulations

The numerical simulations can be used to investigate different dynamical regimes depending upon the choice of the parameter Ri: ageostrophic (Ri = $\mathcal{O}(1)$) to quasi-geostrophic (Ri $\gg$ 1). After about 45 days, all model simulations are in a quasi-steady equilibrium between the buoyancy forcing and the large and small-scale dissipation. We disregard the spinup period here and consider the quasi-steady integrations only. Snapshots of KE and buoyancy for our setup's two extreme Ri (3 and 915) from the quasi-steady state are shown in Fig. 2.2. The differences between the two extreme regimes are evident from the spatial scales of the associated features both in buoyancy and KE, where the ageostrophic regime with Ri = 3 exhibits small scale features with filament-like structures and has much higher KE than

the quasi-geostrophic regime with Ri = 915 which exhibits mesoscale eddy-like features with large spatial scales.

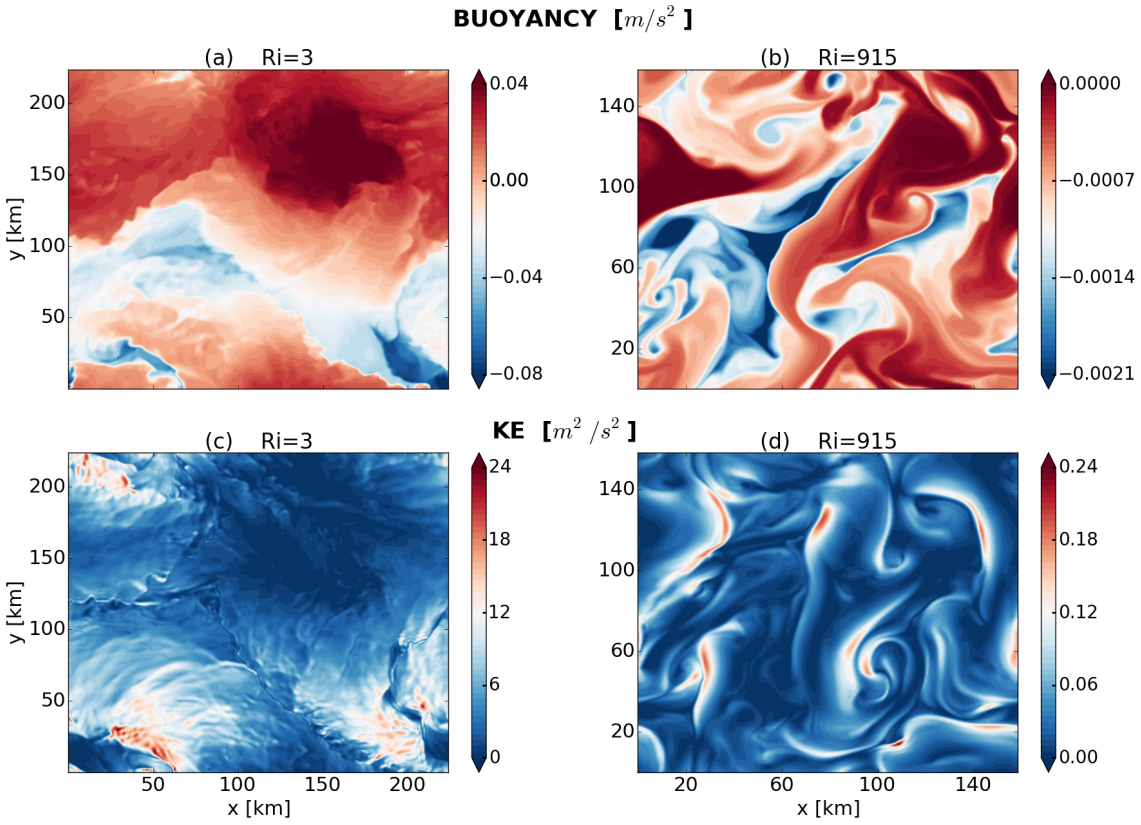


Figure 2.2: Snapshots of buoyancy (a, b) and KE (c, d) at a depth of 1.25 m from the surface in the quasi-steady state of the high resolution model simulations. The differences in the structure and magnitude are clearly evident between the dynamical regimes with Ri = 3 (a, c) and Ri = 915 (b, d), equivalent of an ageostrophic and quasi-geostrophic regime respectively.

Snapshots of the vertical velocity corresponding to the Ri in Fig. 2.2 are shown in Fig. 2.3. These suggest the existence of wave-like features for both ageostrophic and quasi-geostrophic regimes, similar to e.g. what is described in Plougonven and Snyder (2007). These wave-like features manifest themselves as wave trains, as can be seen for instance in the vertical velocity for Ri = 3 in Fig. 2.3. Since the crests and troughs seen in the figure above are akin to gravity wave activity, those signals are accordingly interpreted by e.g. Plougonven and Snyder (2007) as gravity waves generated by the baroclinic instability process in the simulations. It is the aim of this study to investigate if those signatures are indeed gravity waves in a more qualitative manner, or vertical velocities associated with the balanced mode. We proceed further to explore these wave signals in the Fourier space.

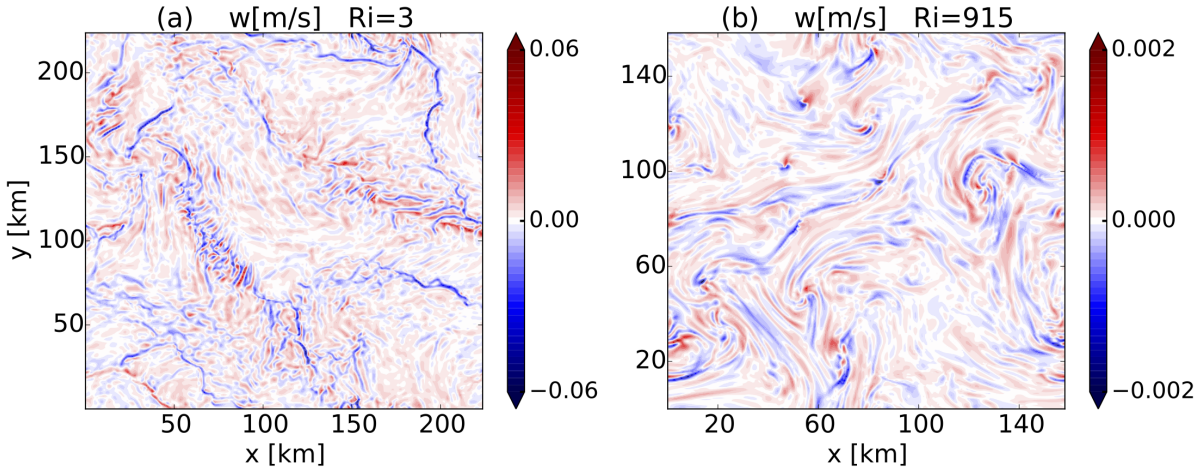


Figure 2.3: Snapshots of vertical velocity at about 20 m depth, for (a) Ri = 3 and (b) 915 from the quasi-steady state of the high resolution model runs. Note that the colorscales are different.

2.3 Analysis in Fourier space

2.3.1 Frequency spectrum

The presence of wave-like features in the physical space motivates us to delve further into Fourier space to identify the characteristic properties of these wave signals. A frequency (ω) spectrum of KE at 50 m depth for different Ri is shown in Fig. 2.4. The spectra shown are averages of 11 chunks, each of 45 day period, of the quasi-steady integrations of the standard model setup and averages over the model domain. We show results from only one depth but frequency spectra calculated at other depths give similar results. Most of the energy is concentrated at the smallest frequencies, i.e. the spectrum is red, but a certain amount of energy is also contained in super-inertial frequencies with $\omega > f$ indicative of gravity waves, in particular for small Ri. The percentage of the KE content above the inertial frequency f is also indicated in the figure. It shows that the relative energy level contained in the super-inertial frequencies is much higher for an ageostrophic regime, than it is for other dynamical regimes. This energy in the super-inertial frequencies could be associated with gravity waves, which have frequencies higher than f. On the other hand, the gravity waves can be Doppler shifted by the mean flow such that the frequency analysis alone does not provide a clear separation of the balanced and gravity mode, for which a frequency and wavenumber spectrum is better suited.

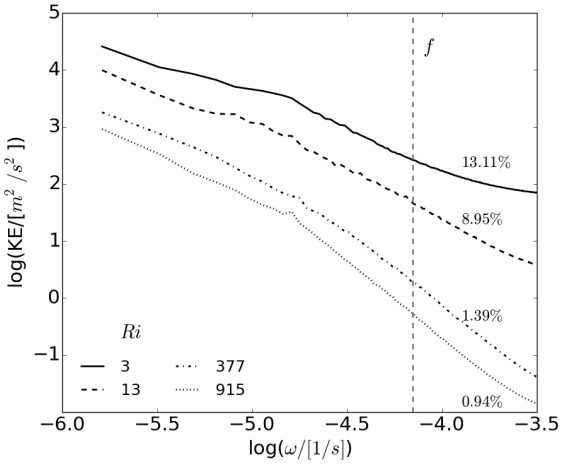


Figure 2.4: Frequency spectrum for KE for different Ri as indicated in the Figure, from the standard model setup. The spectrum shown is an 11-chunk average, and is shown at 50m depth. The vertical dashed grey line indicates the Coriolis frequency f, and the numbers to its right indicate the percentage of energy contained in the frequencies larger than f for each Ri. The values are similar also at other depths (not shown).

2.3.2 Energy in vertical modes

We begin with the vertical wavenumber and consider the energy distribution in vertical modes. All model variables are projected on vertical modes by transformations in the vertical, i.e. discrete sine transformation of the vertical velocity (w) and buoyancy (b), and a discrete cosine transformation of the horizontal velocity (u, v) and pressure (p). The vertical eigenvalues m, for $N = const$ are given by $m = n\pi/H$ for the vertical mode number $n = 0, 1, 2, 3, ...$ After decomposition into the vertical modes we calculate KE, available potential energy (APE), and total energy (TE) contained in each mode. APE is defined here as $P = b'^2/(2N_m^2)$, where $b' = b - N_m^2\ z$ gives the difference between the local buoyancy b and the reference buoyancy $N_m^2 z$ of the time and global mean of N_m^2, which is the stratification of the equilibrated flow.

The distribution of TE and KE as a fraction of TE in the barotropic and first four baroclinic modes is shown in Table 2.2 for different Ri, again using 11 chunks of quasi-steady integrations of the standard setup. The values shown in Table 2.2 are averaged in time and horizontally. The breakdown of energy into vertical modes shows that both KE and APE (the remaining fraction of TE in Table 2.2) rapidly decrease in general with higher vertical modes, although APE appears to share a larger portion of TE for higher

modes. For the first baroclinic mode KE dominates APE for small Ri while the reverse is true for large Ri, whereas for higher baroclinic modes APE dominates KE for all Ri.

Ri →	3		13		377		915	
modes ↓	TE	KE/TE	TE	KE/TE	TE	KE/TE	TE	KE/TE
0	5236.00	1.00	1553.00	1.00	341.00	1.00	163.00	1.00
1	2481.48	0.66	268.01	0.51	36.27	0.19	30.66	0.13
2	861.81	0.33	124.11	0.22	27.92	0.02	25.69	0.01
3	329.12	0.40	49.94	0.26	9.53	0.03	8.27	0.01
4	267.72	0.26	40.65	0.16	9.96	0.01	9.20	0.00

Table 2.2: Total energy (TE) (in m^2s^{-2}) and KE as a fraction of TE in the first five vertical modes for different regimes indicated by Ri. TE and KE are averaged in time and horizontally after they are are computed from vertically decomposed buoyancy and horizontal velocities.

Further, frequency and wavenumber spectra of KE in different vertical modes (not shown) shows distinct differences between odd and even modes, the trend being more pronounced for regimes with higher Ri. This disparity between odd and even modes might be related to the behaviour of the fastest growing mode in the simulations. To test this we project the fastest growing mode ϕ on the vertical eigenfunctions Φ_n (see Appendix A for a detailed derivation). The projection can be written as $\phi(z) = \sum_{n=1}^{\infty} A_n \Phi_n(z)$ and upon solving the coefficient A_n takes the simple form for baroclinic modes[1]:

$$A_n = \frac{2d}{H(1 + n^2\pi^2 d^2/H^2)} \left[(-1)^n \ \chi + 1\right] \tag{2.4}$$

Here $d = f/Nk_h^2$, k_h is the horizontal wavenumber, and χ is a function of d (for the full expression of χ refer to Appendix A). Since the energy in vertical modes depends on the coefficient A_n and A_n in turn is proportional to $\left[(-1)^n \ \chi + 1\right]$, this disparity between modes might be inferred from in Eq. 2.4 (f, N and H being constant). The decisive factors in this expression are $(-1)^n$ and the '+1' in the square brackets, the factor $(-1)^n$ acts as a switch and gives rise to the odd-even non-uniformity in the modes. Note, however, that this projection is performed for the Eady modes which might differ from the ones in the model since the model is based on primitive equations (Section 2.2.1). Although the model background state is not the same as the Eady state, we find a similar behaviour in

[1]For barotropic mode, the factor 2 vanishes in the RHS expression of A_n. See Appendix A.

our simulations. That being so, the projection explains that the energy distribution not only tends to decrease with increasing vertical modes, it also shows a distinct distribution between odd and even modes. Further, we extend the analysis to the frequency and horizontal wavenumber space, to better identify gravity waves in the model simulations.

2.3.3 Frequency-wavenumber analysis

KE and APE in frequency-wavenumber ($\omega - k_h$) space are obtained from a three-dimensional Fourier Transform of the horizontal velocity in time and space (zonal and meridional direction), for different vertical modes. The zonal (k) and meridional (l) wavenumbers are collapsed together to give the horizontal wavenumber (k_h). Variance preserving ω-k_h spectra of KE (shown as $\log_{10}(\omega * k_h * \text{KE})$) for the first and second baroclinic modes ($n = 1, 2$) for Ri = 3 and Ri = 915 are shown in Fig. 2.5. The figure also shows the gravity wave dispersion relation, which can be expressed as $\omega^2 = f^2 + gHk_h^2$ (barotropic mode) and $\omega^2 = f^2 + c_n^2 k_h^2$ (baroclinic modes), where g is acceleration due to gravity and $c_n = N/m$. As the wave's frequency can be influenced by the Doppler-shift, we show $\omega \pm U_0 k_h$ in the figure, where U_0 is the mean flow (see Table 2.1). We assume that the possible region for gravity wave lies approximately within the envelope of the Doppler-shifted extrema. We call this guideline region, henceforth, as the gravity-wave branch.

As is evident from Fig. 2.5, there is a substantial amount of energy in the gravity-wave branch for Ri = 3 for both modes, while it is much smaller for Ri = 915 and outside of the gravity-wave branch. Instead, most of the energy for Ri = 915 is located at the wavelength of the fastest growing mode confirming that there is not much energy related to gravity waves for a quasi-geostrophic regime. For an ageostrophic regime, on the contrary, the energy in the gravity wave branch suggests that ageostrophic dynamics resulting from baroclinic instability at small Ri could generate a significant amount of gravity wave energy. However, a $\omega - k_h$ spectrum is not enough to confirm this statement because the energy of the balanced could also be within the gravity-wave branch. Especially for Ri = 3 in Fig. 2.5, the balanced mode lies mostly within the gravity-wave branch. The co-existence of these processes makes it difficult to isolate the energy contributions from gravity waves or unbalanced modes and the balanced mode. To treat this difficulty and to clearly ascribe this energy to the gravity waves, we employ a modal decomposition method to decompose the full flow vector into these two modes, elaborated in the next section.

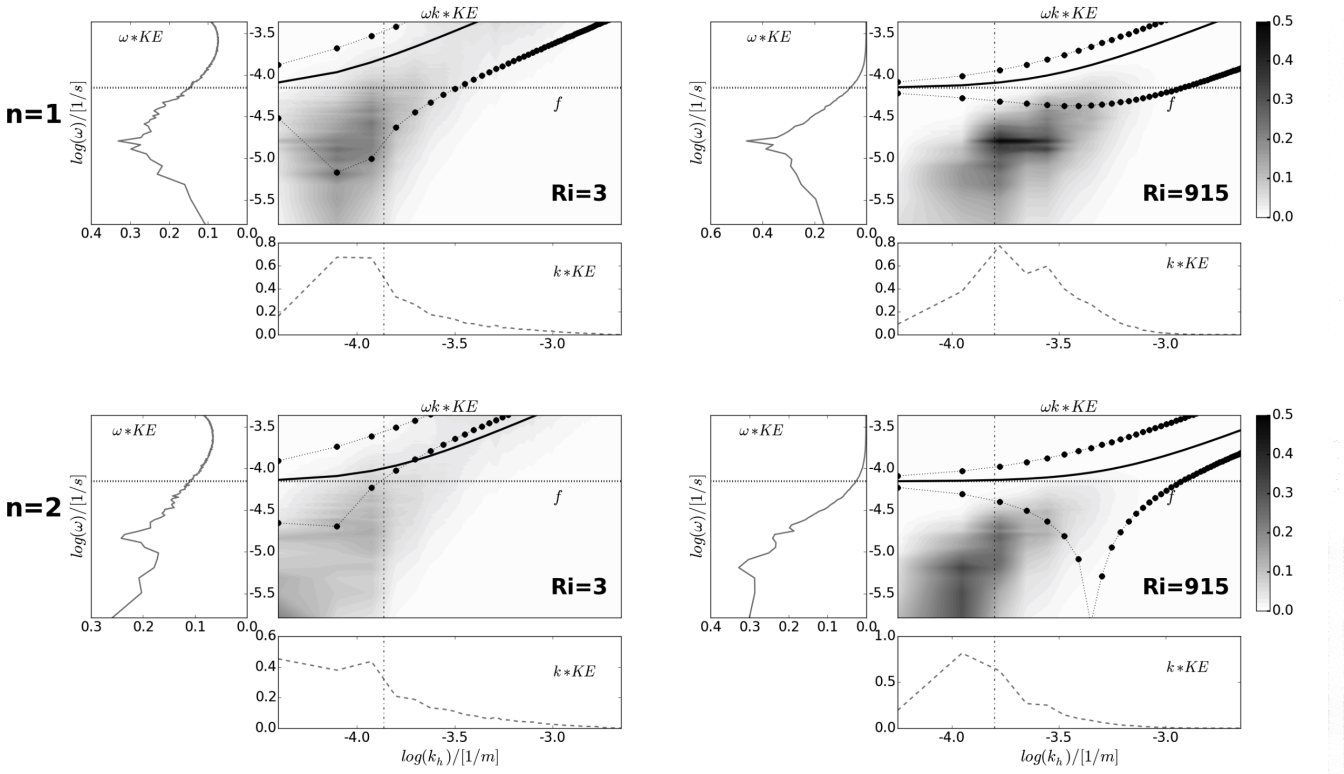


Figure 2.5: Variance preserving frequency-wavenumber spectrum (log-log scale) for KE for the first (top) and second (bottom) baroclinic modes, compared for Ri = 3 (left) and Ri = 915 (right). In each subplot:- Left: solid grey curve shows ωKE as a function of frequency (ω) averaged in wavenumber space; Bottom: the dashed grey curve shows k_hKE as a function of horizontal wavenumber (k_h) averaged in frequency space, and the dashed-dotted grey line indicates k_{max}, the wavenumber of the fastest growing mode; Middle: the shading shows ωk_hKE as a function of ω and k_h, the solid black curve indicates the gravity wave dispersion relation, lines with black circles indicate the Doppler shifted dispersion relation. The gravity wave branch is enclosed by the Doppler-shifted extrema. The dotted black line (left and middle subplot) represents the Coriolis frequency (f). The spectra are calculated and averaged from 11 chunks of 45 day length each from the standard setup. KE shown is normalized with the total KE, and hence the figure shows at which scales most of the KE is concentrated.

2.4 Modal decomposition: balanced and unbalanced modes

We use here a linear modal decomposition to diagnose the gravity wave oscillations in our simulations and extend this decomposition by a non-linear normal mode initialization technique (NLNMI) by Machenhauer (1977), used in numerical weather prediction to generate an appropriate balanced initial state. This section presents an overview of the decomposition methods, details appear in Appendix B.

The hydrostatic and Boussinesq system of equations,

$$\partial_t u + \boldsymbol{u}\cdot\boldsymbol{\nabla} u = -\partial_x p + fv \quad , \quad \partial_t v + \boldsymbol{u}\cdot\boldsymbol{\nabla} v = -\partial_y p - fu \quad , \quad \partial_t b + \boldsymbol{u}\cdot\boldsymbol{\nabla} b = -wN^2 \tag{2.5}$$

complemented by the diagnostic relations $\partial_z p = b$ and $\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$, can be written for the state vector $\boldsymbol{x}$ containing the relevant state variables as:

$$\partial_t \boldsymbol{x} = i\mathscr{L}\cdot\boldsymbol{x} + \mathscr{N}(\boldsymbol{x}) \tag{2.6}$$

where $\mathscr{L}\cdot\boldsymbol{x}$ contains all the linear terms and the vector $\mathscr{N}$ contains the non-linear and forcing terms.

After decomposition into vertical modes and then a Fourier transformation in x and y, the system in matrix notation yields,

$$\tilde{\boldsymbol{x}} = \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{p} \end{pmatrix} \quad , \quad \mathscr{L} = \begin{pmatrix} 0 & -if & -k \\ if & 0 & -l \\ -kc_n^2 & -lc_n^2 & 0 \end{pmatrix} \tag{2.7}$$

where a Fourier transformed quantity is represented by a tilde ($\tilde{}$). The spectrum of the matrix $\mathscr{L}$, which is the set of eigenvalues of $\mathscr{L}$, describe the characteristic frequencies of the system and are given by $\omega = \omega^0 = 0$ and $\omega = \omega^\pm = \pm\sqrt{f^2 + c_n^2 k_h^2}$. Notice that $\omega^\pm$ resembles the dispersion relation for gravity waves with $c_{n>0} = N/m = NH/(n\pi)$ and $c_0 = \sqrt{gH}$.

The corresponding right ($\boldsymbol{q}^{0,\pm}$) and left ($\boldsymbol{p}^{0,\pm}$) eigenvectors[2] to the matrix $\mathscr{L}$ then give the state vector $\tilde{\boldsymbol{x}}$ in Fourier space, which can be expressed as:

$$\tilde{\boldsymbol{x}} = b\,\boldsymbol{q}^0 + g^+\boldsymbol{q}^+ + g^-\boldsymbol{q}^- \quad \text{with} \quad b = \boldsymbol{p}^0\cdot\tilde{\boldsymbol{x}} \ , \ g^+ = \boldsymbol{p}^+\cdot\tilde{\boldsymbol{x}} \ , \ g^- = \boldsymbol{p}^-\cdot\tilde{\boldsymbol{x}} \tag{2.8}$$

[2]For a given matrix, a right eigenvector is a column vector while a left eigenvector is a row vector. In the context of matrices, the commonly used "eigenvector" is the right eigenvector. Here we use the two eigenvectors separately. See Appendix B for details.

This yields a projection of the state vector $\tilde{\boldsymbol{x}}$ on three eigenmodes or normal modes: one balanced slow mode and two unbalanced fast gravity wave modes corresponding to the characteristic frequencies ω^0 and $\omega^{\pm}$ respectively. Henceforth, we call these modes as balanced mode and unbalanced gravity modes (or simply unbalanced modes). The left and right eigenvectors are given explicitly in Appendix B.

We define vectors $\tilde{\boldsymbol{x}}_B$ and $\tilde{\boldsymbol{x}}_G$ associated with the balanced manifold and unbalanced manifold respectively:

$$\tilde{\boldsymbol{x}}_B = b\,\boldsymbol{q}^0 = \mathscr{B}\cdot\tilde{\boldsymbol{x}} \tag{2.9}$$

$$\tilde{\boldsymbol{x}}_G = g^+\boldsymbol{q}^+ + g^-\boldsymbol{q}^- = \mathscr{G}^+\cdot\tilde{\boldsymbol{x}} + \mathscr{G}^-\cdot\tilde{\boldsymbol{x}} \tag{2.10}$$

Since the matrix $\mathscr{L}$ is Hermitian, the eigenvectors of $\mathscr{L}$ are mutually orthogonal, and so are the balanced and unbalanced modes and the associated linear manifolds. As the manifolds $\mathscr{B}$ and $\mathscr{G}$ are mutually orthogonal, they span the whole phase space and therefore the vector $\tilde{\boldsymbol{x}}$ can be written as a linear combination of the two modes implying that $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}_B + \tilde{\boldsymbol{x}}_G$ holds true (cf. Leith, 1980; Theiss and Mohebalhojeh, 2009). After reverse transformation from $\tilde{\boldsymbol{x}}$ to $\boldsymbol{x}$ (i.e. $\tilde{\boldsymbol{x}}_B$ to $\boldsymbol{x}_B$ and $\tilde{\boldsymbol{x}}_G$ to $\boldsymbol{x}_G$) the energy contained in the balanced and unbalanced gravity modes can be obtained.

2.4.1 Linear modal decomposition

In the linear case ($\mathscr{N} = 0$ in Eq. (2.6)) any $\tilde{\boldsymbol{x}}$ can be projected on the slow linear manifold. This part of $\tilde{\boldsymbol{x}}$ i.e. $\tilde{\boldsymbol{x}_B}$ becomes stationary,

$$\partial_t\tilde{\boldsymbol{x}}_B = i\mathscr{L}\cdot\tilde{\boldsymbol{x}}_B = ib\;\omega^0\boldsymbol{q}^0 = 0 \tag{2.11}$$

and only the fast modes will evolve in time according to

$$\partial_t\tilde{\boldsymbol{x}}_G = i\mathscr{L}\cdot\tilde{\boldsymbol{x}}_G = \sum_{d=\pm} ig^d\omega^d\boldsymbol{q}^d \tag{2.12}$$

The balanced and the gravity modes are then contained in $\tilde{\boldsymbol{x}}_B$ and $\tilde{\boldsymbol{x}}_G$ respectively. We call the modes resulting from the linear normal mode decomposition of baroclinically unstable fully non-linear model state as the linear balanced mode (BAL_L) and linear unbalanced modes (UNB_L).

2.4.2 Non-linear modal decomposition

The linear modes refer to the linear system and using them for the realistic non-linear case does not provide a consistent decomposition. To handle this discrepancy we include in the decomposition the non-linearity in the system. For $\mathscr{N} \neq 0$, the time evolution of the slow mode is non-zero and its behaviour becomes

$$\partial_t \tilde{\boldsymbol{x}}_B = \tilde{\mathscr{N}}(\boldsymbol{x}_B) \neq 0 \tag{2.13}$$

where $\boldsymbol{x}_B$ is the inverse Fourier transform of $\tilde{\boldsymbol{x}}_B$ and $\tilde{\mathscr{N}}$ the Fourier transform of $\mathscr{N}$. One approach to separate the modes in the non-linear case is to choose a state that eliminates the fast modes such that only the slow mode remains, whose difference with the full vector then gives the isolated fast modes. According to Machenhauer (1977), the time changes in the fast modes, i.e. $\mathscr{G} = \sum_{d=\pm} \mathscr{G}^d$, can be eliminated by a non-linear normal mode initialization technique requiring that the time derivative of the fast modes is zero, that is

$$\mathscr{G} \cdot \partial_t \tilde{\boldsymbol{x}} = 0 \tag{2.14}$$

which for Eq. (2.6)

becomes

$$\begin{aligned} \mathscr{G} \cdot \partial_t \tilde{\boldsymbol{x}} = i\mathscr{G} \cdot (\mathscr{L} \cdot \tilde{\boldsymbol{x}}) + \mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}) = 0 \;\rightarrow\; (\mathscr{L} \cdot \mathscr{G}) \cdot (\mathscr{G} \cdot \tilde{\boldsymbol{x}}) = i\mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}) \\ \rightarrow\; \mathscr{G} \cdot \tilde{\boldsymbol{x}} = i(\mathscr{L} \cdot \mathscr{G})^{-1} \cdot \mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}) \end{aligned} \tag{2.15}$$

Here the linear operator $(\mathscr{L} \cdot \mathscr{G})^{-1} \cdot \mathscr{G} = \sum_{d=\pm} (\omega^d)^{-1} \mathscr{G}^s$ operates in the gravity mode space only where the eigenvalues are non-zero to avoid problems by singularities for the inversion $(\mathscr{L} \cdot \mathscr{G})^{-1}$ as claimed by Leith (1980).

Eq. (2.15) is a non-linear condition on $\boldsymbol{x}$ which is proposed by Machenhauer (1977) to be solved iteratively until convergence is reached, which is usually the case after a few steps. Starting with a linear slow mode $\boldsymbol{x}_B$ the iteration for the initialisation technique is given by:

$$\tilde{\boldsymbol{x}}_1 = \tilde{\boldsymbol{x}}_B + i(\mathscr{L} \cdot \mathscr{G})^{-1} \cdot \mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}_B) \tag{2.16}$$

$$\tilde{\boldsymbol{x}}_2 = \tilde{\boldsymbol{x}}_B + i(\mathscr{L} \cdot \mathscr{G})^{-1} \cdot \mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}_1) \tag{2.17}$$

$$\ldots \tag{2.18}$$

$$\tilde{\boldsymbol{x}}_k = \tilde{\boldsymbol{x}}_B + i(\mathscr{L} \cdot \mathscr{G})^{-1} \cdot \mathscr{G} \cdot \tilde{\mathscr{N}}(\boldsymbol{x}_{k-1}) \tag{2.19}$$

It was shown by Leith (1980) that the first iteration step corresponds to the quasi-geostrophic approximation. Hence, we use only the first step in our analysis, further iterations do not change the results much. The result of the iteration can now be used to calculate the non-linear balanced mode, the difference of this balanced mode to the actual state vector can be interpreted as the non-linear gravity mode. We call the modes resulting from the non-linear decomposition as the non-linear balanced mode (BAL_NL) and non-linear unbalanced modes (UNB_NL).

2.4.3 Decomposition results

As stated above, a non-linear decomposition is more suitable for a fully non-linear model state of a baroclinically unstable flow, such as our setup, so we first present an example of the non-linear decomposition in physical space. Next we consider both the decompositions by means of $\omega - k_h$ spectra and elaborate on the differences. Fig. 2.6 shows snapshots of zonal velocity for Ri = 915 for the full velocity (FULL) and velocities from the non-linear balanced mode (BAL_NL) and non-linear unbalanced modes (UNB_NL) as an example of the non-linear decomposition. For the balanced mode more large-scale features are present as in the full component, whereas the unbalanced mode shows indeed more small-scale features akin to gravity wave activity. This is a first indication that the separation of the slow and fast modes using the modal decomposition is effective. To demonstrate this more quantitatively we apply next a frequency-wavenumber analysis to the decomposed fields.

The $\omega - k_h$ spectra (see Fig. 2.7, Fig. 2.8, Fig. 2.9) are computed similar to the method described in Section 2.3.3 and then averaged for 11 chunks of 45 day length each. The spectra shown are at 50 m depth and are normalized with the total KE (or APE) under the volume. Recall from Section 2.1, that the coupling between balanced and unbalanced modes tends to be weaker for Ri $\gg$ 1 regime such that the temporal scales of the slow balanced and the fast unbalanced motions are well separated. The opposite is true for small Ri where this coupling is much stronger and the separation of the modes gets more difficult. We tackle the less complicated case first (for Ri $\gg$ 1) before expounding on a case with small Ri.

A $\omega - k_h$ spectra of KE for Ri = 915 is shown in Fig. 2.7 for modes obtained from the linear and non-linear decomposition. As the balanced modes correspond to the motions with large temporal scales, in $\omega - k_h$ space the energy associated with the balanced modes is expected to lie towards low frequencies, away from the high frequencies. The energy associated with the unbalanced motions, on the other hand, is expected to lie in the region

confined by the gravity-wave branch, which is the super-inertial frequency range that allows for gravity waves enveloped by the corresponding Doppler-shifted dispersion relation. The expectation is fulfilled in the case of balanced modes (BAL_L, BAL_NL), as seen from Fig. 2.7(a, b), where most of the KE lies outside the gravity-wave branch and towards small frequencies and wavenumbers.

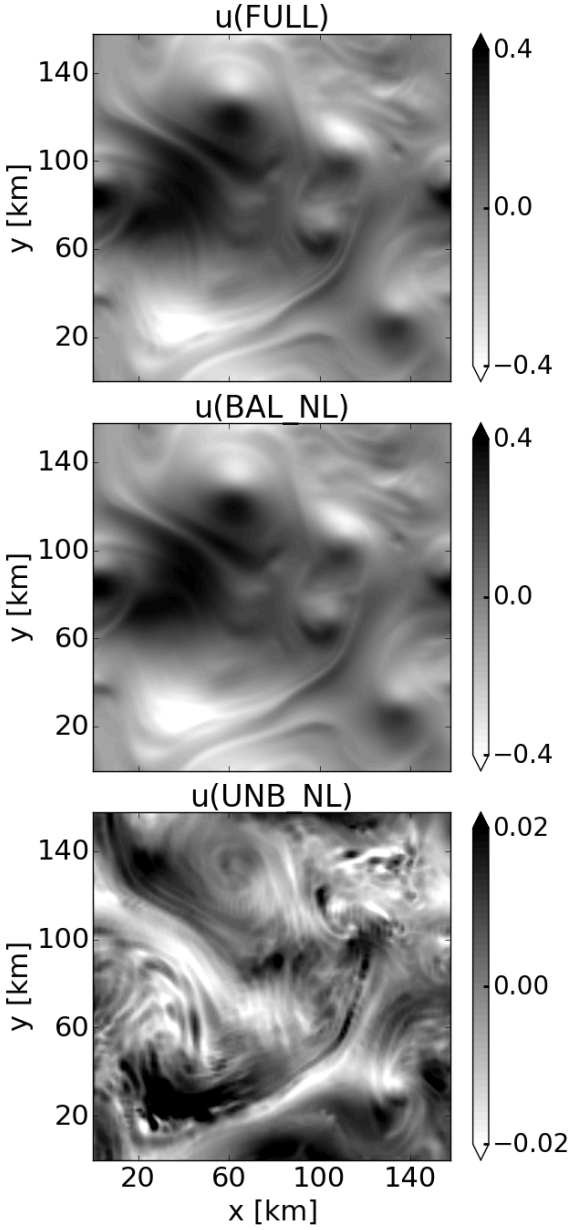


Figure 2.6: Snapshots of zonal velocity for Ri = 915 with its full component, FULL (top) and modally decomposed components: balanced, BAL_NL (middle) and unbalanced, UNB_NL (bottom) modes using non-linear decomposition. Note the difference in the colorscale in the bottom panel.

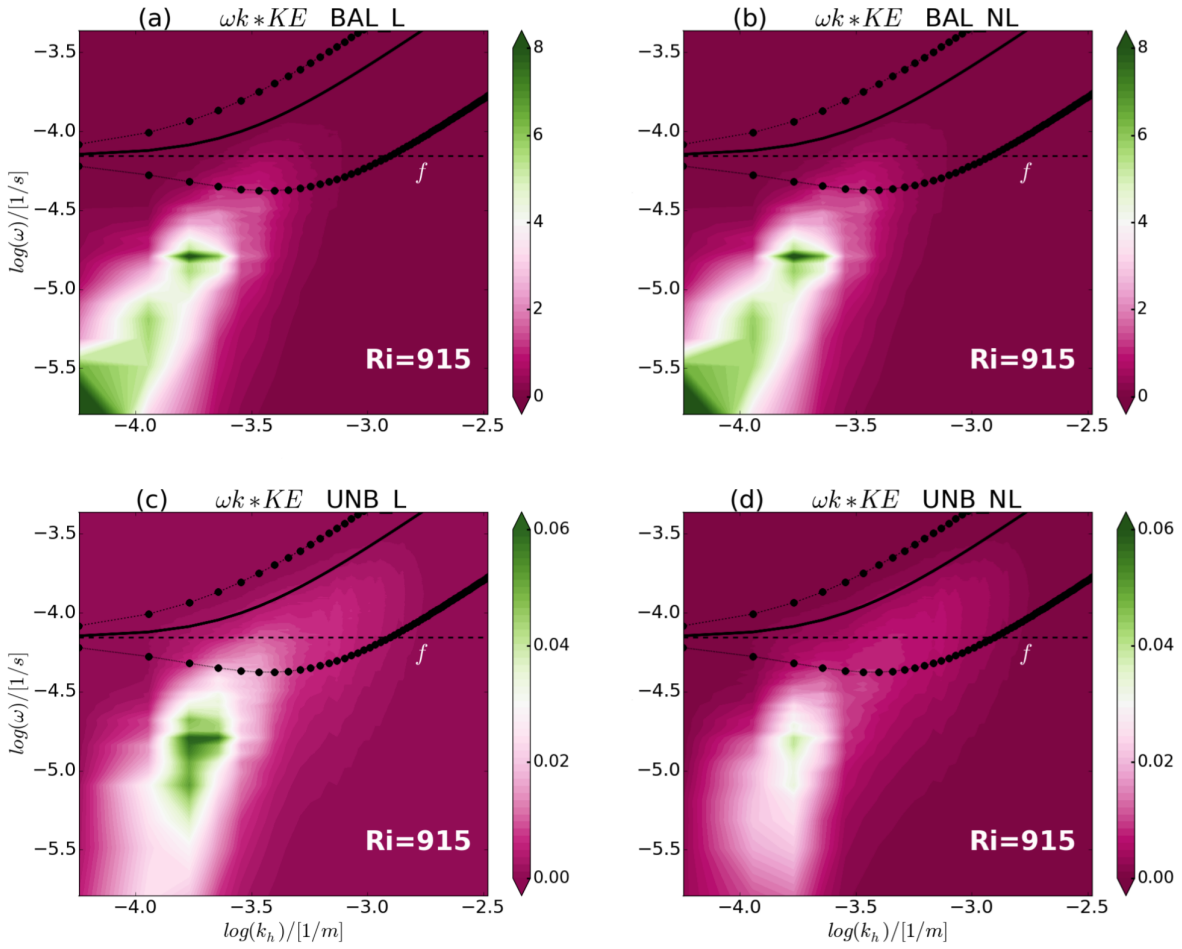


Figure 2.7: Variance preserving frequency-wavenumber spectra of KE of the balanced (a,b) and unbalanced (c,d) modes for Ri = 915 at 50 m depth. The spectra are calculated and averaged from 11 chunks of 45 day length each from the standard setup and the KE shown is normalized with the total KE under the volume. The spectra in (a) and (c) are obtained from the linear decomposition (BAL_L, UNB_L), and the ones in (b) and (d) from the non-linear decomposition (BAL_NL, UNB_NL). In each subplot: the color shading shows the respective KE as a function of ω and k_h, the dashed black line indicates the Coriolis frequency (f), the solid black curve indicates the gravity wave dispersion relation and lines with black circles show the Doppler shifted dispersion relation. Note that the gravity-wave branch is enclosed by the Doppler-shifted extrema. Notice a magnitude difference of two orders between the balanced and unbalanced modes.

However, also for the unbalanced modes (UNB_L, UNB_NL) (Fig. 2.7(c, d)) most of the KE lies outside the gravity-wave branch against the expectation. This KE is much higher for the linear unbalanced modes than for the non-linear unbalanced modes (compare Fig. 2.7(c) and (d)). This suggests that this discrepancy could be a consequence of

applying a linear decomposition to a non-linear flow system. The non-linear decomposition in Fig. 2.7(d) indeed shows a reduction of KE that sits outside the gravity-wave branch, as compared to the linear decomposition in Fig. 2.7(c). Thus, the non-linear decomposition appears to be an improvement over the linear one and a better suited tool than a linear decomposition to decompose the modes. We find similar improvement for the other model experiments (not shown).

However, there is still energy outside the gravity wave branch in Fig. 2.7(d). This could be related to a misinterpretation of the gravity wave branch, since, as mentioned in Section 2.3.3, we use the mean flow prescribed initially U_0 (see Table 2.1) to calculate the Doppler shift. This mean flow might not be well suited for an "effective" Doppler shift of the waves, which means that the gravity wave branch would also change. On the other hand, the residual unbalanced energy outside the gravity wave branch could also be related to an (unknown) artifact of the non-linear decomposition. A possibility for such an artifact is that we use the eigenvectors $\boldsymbol{q}^{0,\pm}$ and $\boldsymbol{p}^{0,\pm}$ of the analytical instead of the discrete system for the method. We will explore this issue further in later studies and in the following we consider the results only from the non-linear decomposition since they show improvement with respect to our expectation.

Further, notice the difference in magnitude of two orders between the balanced and unbalanced modes in Fig. 2.7, stating that a significant amount of KE is contained in the balanced mode for large Ri. The negligible amount of KE in the gravity-wave branch even for the unbalanced modes signifies that the gravity wave emission is weak in a $\mathrm{Ri} \gg 1$ regime. We now consider examples from other Ri.

Balanced (BAL_NL) and unbalanced (UNB_NL) modes from the non-linear decomposition for the extreme Ri (3) and an intermediate Ri (13) in our simulations are shown in Fig. 2.8. For the balanced modes (see Fig. 2.8 (a, c)), for $\mathrm{Ri} = 3$ and 13, most of the KE lies outside the gravity-wave branch, but some KE is also within this branch. This is associated with the strong coupling between the balanced and unbalanced motions for small Ri, for which the timescale separation of these modes is not well-defined. In contrast, most of the KE in the balanced mode for the aforementioned $\mathrm{Ri} = 915$ clearly lies outside this branch. Also, the balanced mode has much higher KE for $\mathrm{Ri} = 915$ than for smaller Ri.

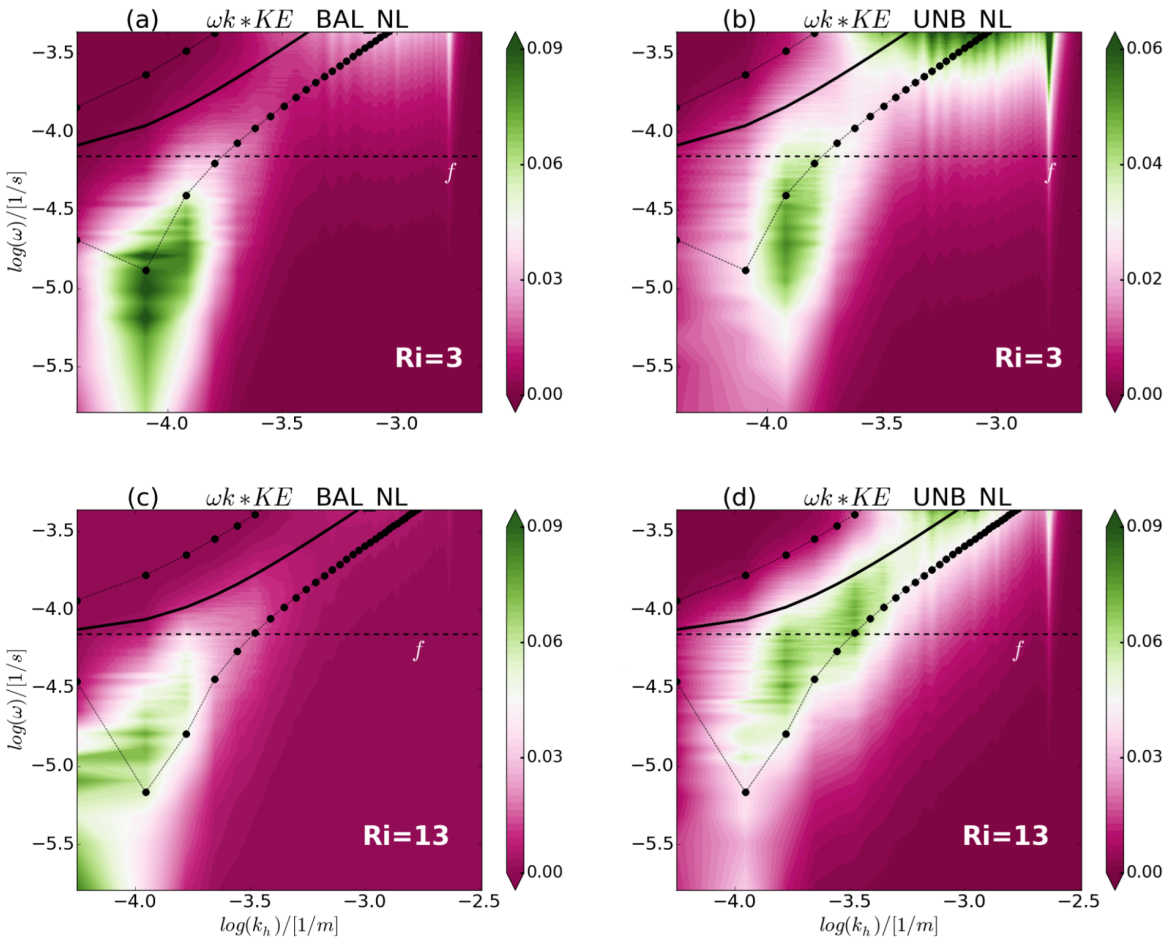


Figure 2.8: Variance preserving frequency-wavenumber spectra of KE similar to Fig. 2.7 but for different Ri: 3 (top) and 13 (bottom), from the non-linear decomposition. The spectra are for the non-linear balanced modes (BAL_NL) are shown in (a) and (c) and the ones for the non-linear unbalanced modes (UNB_NL) are shown in (b) and (d). The gravity-wave branch is enclosed by the Doppler-shifted extrema. Note the difference in the colorscales.

This high KE for Ri = 915 is not surprising, since as one moves from ageostrophy towards quasi-geostrophy, one progresses towards a more 'balanced' state, and not unexpectedly would one find balanced modes dominating the unbalanced modes. Put in other words, it implies that a regime in a quasi-geostrophic balance will emit weak unbalanced gravity waves. In the context of modal decomposition, this suggests that for a regime in a quasi-geostrophic balance the energy in unbalanced modes would be far less than in an ageostrophic regime. This conjecture is supported by Fig. 2.8 (b, d) and Fig. 2.7 (d) for the non-linear unbalanced modes. The figure clearly illustrates the negligibly small KE in

the gravity-wave branch for the unbalanced modes of Ri = 915 in contrast to Ri = 3 and 13, where a significant amount of KE is concentrated within the gravity-wave branch. For the intermediate Ri = 13, KE in the unbalanced mode is aligned along the gravity-wave dispersion relation. For Ri = 3, the reason for high KE at higher frequencies is not known, and might also be related to numerical issues which we have not tested further.

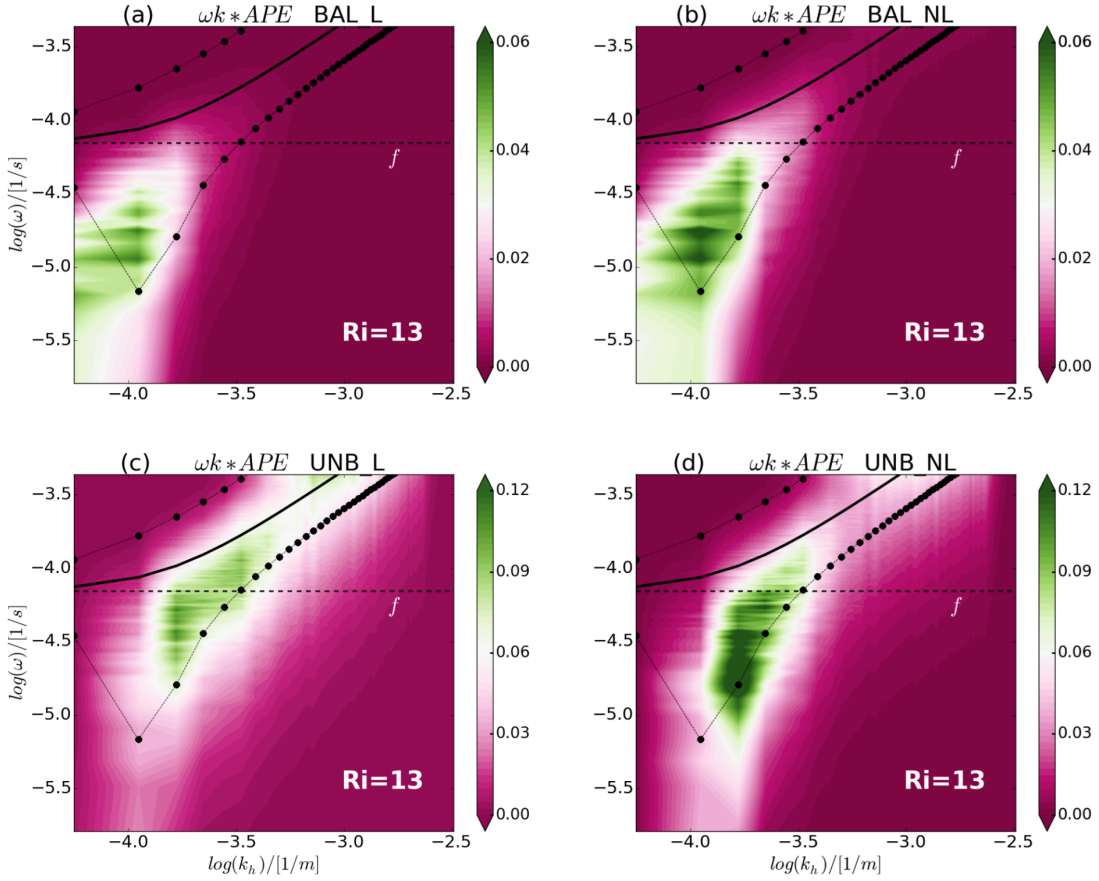


Figure 2.9: Variance preserving frequency-wavenumber spectra of APE similar to Fig. 2.7 but for Ri = 13. The balanced modes are shown in (a,b) and unbalanced modes in (c,d). The spectra in (a) and (c) are obtained from the linear decomposition (BAL_L, UNB_L), and the ones in (b) and (d) from the non-linear decomposition (BAL_NL, UNB_NL). The gravity-wave branch is enclosed by the Doppler-shifted extrema. Notice that the colorscales differ for unbalanced and balanced modes.

Further we discuss an $\omega - k_h$ spectra for APE for the linear and non-linear decompositions, and as an example we show the APE spectra for Ri = 13 in Fig. 2.9. The APE spectra exhibit a distribution between balanced and unbalanced modes similar to what is described before for KE (Fig. 2.8 (b, c)). In the balanced modes (BAL_L, BAL_NL)

(Fig. 2.9 (a, b)), most of the APE is present at lower frequencies, whereas for the unbalanced modes (UNB_L, UNB_NL) APE tends to be present at higher frequencies. However, there appears to be more APE in the non-linear modes (BAL_NL, UNB_NL) than in the linear modes (BAL_L, UNB_L), and unbalanced modes show more APE than the balanced modes.

2.5 Energy dissipation under different dynamical regimes

The energy in the model setup is dissipated at small scales by horizontal biharmonic friction and harmonic vertical friction, and at large scales by the zonal mean drag. We use the modal decomposition results to assess the differences in small and large-scale dissipation associated with the velocities of the full model state, balanced and unbalanced modes.

The equation for KE obtained from the horizontal momentum equation is written as

$$\partial_t K = -\boldsymbol{\nabla} \cdot \boldsymbol{u} K - \boldsymbol{\nabla} \cdot \boldsymbol{u} p + w b' + \boldsymbol{u}_h \cdot \lambda_u(\overline{\boldsymbol{u}_h}) + \boldsymbol{u}_h \cdot \boldsymbol{F}_u \tag{2.20}$$

where $K = \boldsymbol{u}_h^2/2$ denotes the kinetic energy, $\boldsymbol{u} = (u, v, w)$ is the full, $\boldsymbol{u}_h$ is the horizontal, and $\overline{\boldsymbol{u}_h}$ is the zonal mean velocity. The dissipation terms, which are the last two terms in the RHS of Eq. (2.20), extract KE from the flow. Large-scale dissipation, which acts on the large scales, is denoted by the term $D_L = \boldsymbol{u}_h \cdot \lambda_u(\overline{\boldsymbol{u}_h})$, where λ_u is the linear drag coefficient which extracts energy from the mean flow and is related to the maximum growth rate as $\lambda_u = 0.75\,\sigma_{max}$. Small-scale dissipation on the other hand damps the smallest scales and is denoted by $D_S = \boldsymbol{u_h} \cdot \boldsymbol{F}_u$, where $\boldsymbol{F}_u = A_h \nabla^4 \boldsymbol{u}_h + A_v \partial^2 \boldsymbol{u}_h / \partial z^2$ indicates the dissipation due to biharmonic and vertical friction respectively (for A_h and A_v see Table 2.1).

The global mean values of KE dissipation are illustrated in Fig. 2.10 for all Ri for their full velocity component (FULL) and modally decomposed components: non-linear balanced mode (BAL_NL), and non-linear unbalanced modes (UNB_NL). The figure illustrates the large-scale (D_L) and small-scale (D_S) dissipation values, shown as a fraction of the total dissipation ($D_S + D_L$) for the respective mode; note that $D_S + D_L = 1$. In addition, the dissipation values for the modes from linear decomposition and contributions to D_S by biharmonic friction (D_b) and vertical friction (D_v) are tabulated in Table C of Appendix C.

For small Ri, D_S of the full component is larger than the corresponding D_L, while D_S becomes smaller and negligible for higher Ri, as seen from the figure. The opposite is true

for D_L which dominates D_S at larger Ri and becomes smaller, but not negligible, for small Ri. The impact of vertical friction (D_v) is in general much smaller than the biharmonic friction (D_b) in all cases (Table C, Appendix C). Further, we weigh the variations in the dissipation due to the decomposed modes for different Ri, but note that the unscaled dissipation rates related to the balanced and unbalanced modes will not add up to the dissipation of the full velocities (i.e. D_S(UNB_NL) + D_S(BAL_NL) $\neq$ D_S(FULL); also for D_L) because we disregard covariances between the decomposed velocities here. Only for the total energy of the linear decomposition, the covariances of the state vector components $\boldsymbol{x}$ cancel.

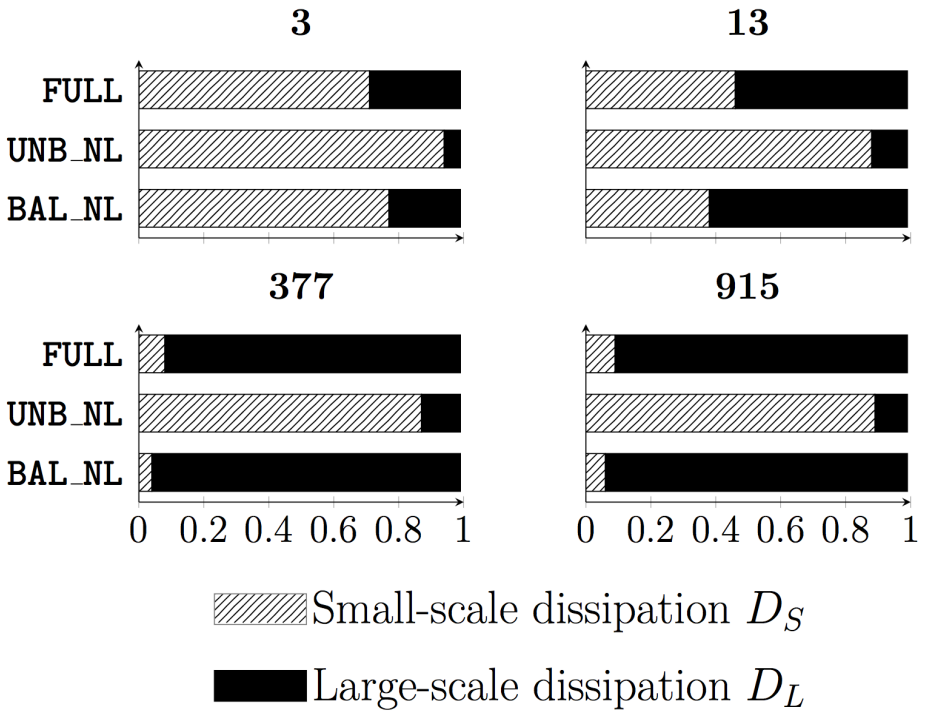


Figure 2.10: Globally integrated values of KE dissipation for the full and the non-linear balanced and unbalanced components (BAL_NL, UNB_NL) for all Ri. Dissipation contribution to the total KE dissipation from small-scales (D_S) is shown by the hatched region and from large-scales (D_L) by the black shaded region. Note that $D_S + D_L = 1$ for the respective mode.

The dissipation related to the non-linear balanced modes, BAL_NL, has higher values for D_L than D_S for all Ri except for Ri = 3. For Ri = 377 and Ri = 915, most of the dissipation in BAL_NL occurs at large-scales while it is almost negligible for small-scales.

For Ri = 3, however, a significant amount of dissipation still occurs at the small-scales. For the non-linear unbalanced modes, UNB_NL, D_S is much larger than D_L for all Ri, with a maximum of D_S for Ri = 3, while D_L for UNB_NL tends to become almost negligible for small Ri. This directs to the inference that in the gravity wave mode dissipation occurs predominantly at small-scales for all Ri.

The dissipation results for the decomposed modes elucidate that regimes with larger Ri dissipate energy preferably at large-scales contrary to the small Ri regimes which dissipate energy predominantly at small-scales. This result is consistent with previous results, e.g. Brüggemann and Eden (2015), which predict that in wavenumber space regimes with large and small Ri show a dominant KE flux towards large scales and small scales, respectively. In their study, the velocity field was decomposed into its rotational and divergent components, analogous to the balanced and unbalanced gravity modes in our study. Brüggemann and Eden (2015) showed that a downscale energy transfer is associated with a divergent flow field. However, a specific connection to a process, such as IGWs, was not made. We suggest based on our results that gravity waves could be a potential participant in the downscale energy transfer via the ageostrophic route.

2.6 Summary and Conclusions

In this study, we diagnose internal gravity waves emitted from balanced flow using the non-linear initialization technique of Machenhauer (1977). We use an idealized numerical setup which is baroclinically unstable and the choice of the Richardson number (Ri) allows us to emulate different dynamical regimes ranging from ageostrophic (Ri = $\mathcal{O}(1)$) to quasi-geostrophic (Ri$\gg$1) flows. We first diagnose internal gravity waves in frequency and wavenumber space, and then using linear and non-linear modal decomposition. The modal decomposition yields balanced and unbalanced gravity modes, which we discuss in frequency-wavenumber space, followed by an assessment of the dissipation associated with the decomposed modes. The key results are as follows:

1. An ageostrophic regime shows much more gravity wave activity than a quasi-geostrophic regime. In frequency-wavenumber space, most of the energy is concentrated at small scales for the former and at large scales for the latter.

2. The non-linear initialization technique used is able to segregate balanced and unbalanced gravity modes, although this segregation becomes difficult for Ri = $\mathcal{O}(1)$ or

less. The non-linear decomposition is promising and an improvement over the linear decomposition.

3. Gravity waves dissipate predominantly through small-scale dissipation. Hence, gravity waves could be responsible for the downscale energy transfer.

It should be noted that the non-linear decomposition might be less efficient for Ri = $\mathcal{O}(1)$ since the interactions between the balanced and unbalanced motions are much stronger and the time scale separation between these modes is minimal, which renders it hard to separate one mode from the other. Nonetheless, for regimes which allow the separation of fast and slow motions, the procedure seems promising and can be applied to isolate gravity wave modes from balanced modes for future studies.

We have attributed the IGW signals in the simulations as emissions from balanced flows. However, the unbalanced motions generated from balanced motions could form triads and can further generate unbalanced motions. This process might also contribute to IGW activity. Triad interactions between slow balanced and fast unbalanced motions have been studied, for e.g. by Bartello (1995), which suggest that in a slow-fast-fast triad the slow mode can catalyze the flow of energy from one fast mode to another. This slow-fast-fast interaction also sweeps the balanced energy from the slow mode downscale to the scales of dissipation (Bartello, 1995). This supports the notion of a downscale energy transfer via an ageostrophic route suggested by e.g. Brüggemann and Eden (2015). Our results suggest that gravity waves could be a potential candidate in the downscale energy transfer via the ageostrophic route and hence could be catalyzed by ageostrophic baroclinic instability.

2.7 Appendix A: Eady mode projection on vertical modes

The wave solution for the Eady case can be expressed as:

$$\phi(z) = A\cosh(z/d) + B\sinh(z/d) \tag{2.21}$$

where $d = f/Nk_h^2$ (where $k_h^2 = k^2 + l^2$) and $A = BH(U_0 - c)/(U_0 d)$ and $B = 1$ are constants evaluated from the initial condition (for details of the Eady solution, see e.g. Olbers et al., 2012.) The phase speed, c, is given as:

$$c = \frac{U_0}{2} \pm \frac{U_0 d}{H}\sqrt{1 + \frac{H^2}{4d^2} - \frac{H}{d}\coth(\frac{H}{d})} \tag{2.22}$$

Note that c becomes imaginary when the term under the square root becomes negative. The set of vertical eigenfunction Φ_n is expressed as:

$$\Phi_n(z) = \cos(\frac{n\pi}{H})z = \cos mz \tag{2.23}$$

$$m = \frac{n\pi}{H}, n = 0, 1, 2, 3, ... \tag{2.24}$$

The projection of Eady mode on the vertical modes then can be written as:

$$\phi(z) = \sum_{n=1}^{\infty} A_n \Phi_n(z) \tag{2.25}$$

and the coefficient A_n can be estimated from:

$$\int_{-H}^{0} A_n \Phi_n^2(z)\,\mathrm{d}z = \int_{-H}^{0} \phi(z)\Phi_n(z)\,\mathrm{d}z \tag{2.26}$$

$$\int_{-H}^{0} A_n \cos^2(mz)\,\mathrm{d}z = \int_{-H}^{0} \{\Re(A)\cosh(z/d) + \sinh(z/d)\}\cos(mz)\,\mathrm{d}z \tag{2.27}$$

$$= \frac{d}{1+m^2d^2}\big[\Re(A)\big\{\sinh\big(\frac{H}{d}\big)(-1)^n\big\} - \cosh\big(\frac{H}{d}\big)(-1)^n + 1\big] \tag{2.28}$$

for $n = 0$ (barotropic mode)

$$A_0 = \frac{d}{H(1+m^2d^2)}\big[\Re(A)\sinh(H/d) - \cosh(H/d) + 1\big] \tag{2.29}$$

and for $n = 1, 2, 3, ...$ (baroclinic modes)

$$A_n = \frac{2d}{H(1+m^2d^2)}\big[(-1)^n\big\{\Re(A)\sinh(H/d) - \cosh(H/d)\big\} + 1\big] \tag{2.30}$$

2.8 Appendix B: Modal decomposition

2.8.1 Eigenvectors

The three right (column) eigenvectors $\boldsymbol{q}^{\pm,0}$ to the matrix $\mathscr{L}$ with $\mathscr{L} \cdot \boldsymbol{q}^{\pm,0} = \omega^{\pm,0}\boldsymbol{q}^{\pm,0}$ are given by

$$\boldsymbol{q}^{\pm} = \begin{pmatrix} (ifl - k\omega)/(\omega^2 - f^2) \\ (-ifk - l\omega)/(\omega^2 - f^2) \\ 1 \end{pmatrix} \ , \ \boldsymbol{q}^{0} = \begin{pmatrix} -il/f \\ ik/f \\ 1 \end{pmatrix} \tag{2.31}$$

and the three left (row) eigenvector $\boldsymbol{p}^{\pm,0}$ to the matrix $\mathscr{L}$ with $\boldsymbol{p}^{\pm,0} \cdot \mathscr{L} = \omega^{\pm,0}\boldsymbol{p}^{\pm,0}$ are given by

$$\boldsymbol{p}^{\pm} = \frac{\omega^2 - f^2}{2\omega^2} \left(\frac{(-ifl - k\omega)c_m^2}{\omega^2 - f^2} \ , \ \frac{(ikf - l\omega)c_m^2}{\omega^2 - f^2} \ , \ 1 \right) \ , \ \boldsymbol{p}^{0} = \frac{f^2}{(\omega^{\pm})^2} \left(ilc_m^2/f \ , \ -ikc_m^2/f \ , \ 1 \right) \tag{2.32}$$

Note that it holds that

$$\boldsymbol{p}^{\pm} \cdot \boldsymbol{q}^{\pm} = \boldsymbol{p}^{0} \cdot \boldsymbol{q}^{0} = 1 \ , \ \boldsymbol{p}^{\pm} \cdot \boldsymbol{q}^{\mp} = \boldsymbol{p}^{\mp} \cdot \boldsymbol{q}^{\pm} = \boldsymbol{p}^{0} \cdot \boldsymbol{q}^{\pm} = \boldsymbol{p}^{\pm} \cdot \boldsymbol{q}^{0} = 0 \tag{2.33}$$

(For $\boldsymbol{q}^{\pm}$ and $\boldsymbol{p}^{\pm}$, read ω as $\omega^{\pm}$.)

2.8.2 Projection matrices

Mathematically, $\mathscr{B}$ and $\mathscr{G}$ are the projection matrices $\mathscr{B} = \boldsymbol{q}^{0} \cdot \boldsymbol{p}^{0}$ and $\mathscr{G}^{\pm} = \boldsymbol{q}^{\pm} \cdot \boldsymbol{p}^{\pm}$,

$$\mathscr{B} = \boldsymbol{q}^{0} \cdot \boldsymbol{p}^{0} = \frac{c_m^2}{(\omega^{\pm})^2} \begin{pmatrix} l^2 & -kl & -ilf/c_m^2 \\ -kl & k^2 & ikf/c_m^2 \\ ilf & -ikf & f^2/c_m^2 \end{pmatrix} \tag{2.34}$$

$$\mathscr{G}^{\pm} = \boldsymbol{q}^{\pm} \cdot \boldsymbol{p}^{\pm} = \frac{1}{2(\omega^{\pm})^2} \begin{pmatrix} (f^2l^2 + k^2(\omega^{\pm})^2)/k_h^2 & klc_m^2 - if\omega^{\pm} & ifl - k\omega^{\pm} \\ klc_m^2 + if\omega^{\pm} & (f^2k^2 + l^2(\omega^{\pm})^2)/k_h^2 & -ifk - l\omega^{\pm} \\ (-ifl - k\omega^{\pm})c_m^2 & (ifk - l\omega^{\pm})c_m^2 & k_h^2c_m^2 \end{pmatrix} \tag{2.35}$$

with the properties

$$\mathscr{L} = \omega^0 \mathscr{B} + \sum_{d=\pm} \omega^d \mathscr{G}^d = \begin{pmatrix} 0 & -if & -k \\ if & 0 & -l \\ -kc_m^2 & -lc_m^2 & 0 \end{pmatrix} \tag{2.36}$$

and

$$\mathscr{G}^\pm \cdot \mathscr{G}^\pm = \mathscr{G}^\pm \ , \ \mathscr{G}^\pm \cdot \mathscr{G}^\mp = 0 \ \rightarrow \ \mathscr{L}^n = \sum_{s=\pm} (\omega^s)^n \mathscr{G}^s \ \rightarrow \ f(\mathscr{L}) = \sum_{s=\pm} f(\omega^s) \mathscr{G}^s \tag{2.37}$$

for a general function f, but note that $\mathscr{L}$ is singular since $\omega^0 = 0$.

2.8.3 Operator $i(\mathscr{G} \cdot \mathscr{L})^{-1} \cdot \mathscr{G}$

The matrix $i(\mathscr{G} \cdot \mathscr{L})^{-1} \cdot \mathscr{G}$ introduced in Section 2.4.2 is given by

$$i(\mathscr{G} \cdot \mathscr{L})^{-1} \cdot \mathscr{G} = \frac{1}{(\omega^\pm)^2} \begin{pmatrix} 0 & f & -ik \\ -f & 0 & -il \\ -ikc_m^2 & -ilc_m^2 & 0 \end{pmatrix} \tag{2.38}$$

2.9 Appendix C: Dissipation

Ri →	3				13				377				915			
modes ↓	D_b	D_v	D_S	D_L	D_b	D_v	D_S	D_L	D_b	D_v	D_S	D_L	D_b	D_v	D_S	D_L
FULL	0.69	0.01	0.71	0.28	0.44	0.01	0.46	0.53	0.08	0.00	0.08	0.91	0.09	0.00	0.09	0.90
UNB_L	0.69	0.05	0.75	0.24	0.77	0.09	0.87	0.12	0.71	0.12	0.83	0.16	0.75	0.13	0.88	0.11
UNB_NL	0.86	0.08	0.94	0.05	0.80	0.08	0.88	0.11	0.76	0.11	0.87	0.12	0.78	0.11	0.89	0.10
BAL_L	0.35	0.00	0.36	0.63	0.21	0.00	0.22	0.77	0.04	0.00	0.04	0.95	0.06	0.001	0.06	0.93
BAL_NL	0.72	0.05	0.77	0.22	0.36	0.02	0.38	0.61	0.04	0.00	0.04	0.95	0.06	0.00	0.06	0.93

Table 2.3: Globally integrated values for KE dissipation shown as a fraction of the total dissipation. The total dissipation is a sum of large-scale ($D_L = \boldsymbol{u}_h \cdot \lambda_u \overline{\boldsymbol{u}_h}$) and small-scale ($D_S = \boldsymbol{u}_h \cdot \boldsymbol{F}_u$) dissipation, and D_S comprises of the contributions from biharmonic friction (D_b) and vertical friction (D_v) (i.e. $D_b + D_v = D_S$). The dissipation values are shown for different regimes indicated by Ri for their full velocity component (FULL) and modally decomposed components: linear and non-linear balanced modes (BAL_L, BAL_NL), and linear and non-linear unbalanced modes (UNB_L, UNB_NL).

O' mighty ocean!

Sans winds, sans tide

Yet thou maketh waves inside

Speaketh but, O' mighty ocean!

What doth thee maketh of those waves

When walls in thy path stays?

Chapter 3

Gravity wave emission by different mechanisms

This chapter is in preparation for submission.

Abstract

We introduce a new diagnostic tool to diagnose internal gravity waves emitted from balanced flow and examine the wave emission in three different scenarios: spontaneous emission, convective instability, and lateral boundary instability. The new diagnostic tool, quasi-geostrophic (QG) filter, is based on a potential vorticity inversion method to diagnose internal gravity waves. The QG filter has the additional capability to diagnose ageostrophic horizontal velocity components from the flow, thus providing a robust estimate of the unbalanced component. The QG filter is implemented in a version consistent on the discrete grid level and compared with a reformulated discrete implementation of the non-linear initialization technique proposed by Machenhauer (1977). The QG filter comes forth as an efficient diagnostic tool, however the discrete version of Machenhauer's method extended to three iterations performs even better by minimizing wave activity in the initial conditions. Internal gravity waves diagnosed from the QG filter are examined in three scenarios: spontaneous emission, generation at fronts, and emission from boundary instabilities. The emission in the case of frontal instability is higher as compared to spontaneous emission and boundary instabilities. Further, the waves emitted from boundary instabilities are similar to Kelvin waves arrested by the balanced flow. The QG filter provides a robust extraction of wave signals revealing considerable wave activity in the three scenarios presented. The results bring forth the prominent role of internal gravity waves as an energy sink of balanced flow.

3.1 Introduction

Quasi-geostrophically balanced flows, which span a range of motions from mesoscale eddies to the large-scale circulation, dominate atmospheric and oceanic flows. Since balanced flows dominate the geophysical flows, balanced models such as based on quasi-geostrophic equations are a reasonable approximation to describe geophysical flows. But besides the low frequency balanced motions, high frequency unbalanced motions, such as internal gravity waves (IGWs), forced internally or externally occur naturally in geophysical flows and play a vital role in shaping the flow dynamics. Balanced models exclude by construction the high frequency dynamics in order to facilitate the theoretical understanding of the geophysical flows which are largely balanced. This exclusion sets a restriction to explore the interaction between balanced and unbalanced dynamics which is germane to conjure up a picture of the energy transfers among different dynamical processes. Whilst the

unbalanced motions are crucial to understand the energetics of the flow, the rapidity of their temporal scales renders them hard to observe or resolve in models as well as makes it complex and tenuous to identify and disentangle them from other motions. In this paper, we introduce a potential vorticity inversion method called a quasi-geostrophic filter (QG filter) embedded in a quasi-geostrophic balanced model to filter and diagnose the high frequency unbalanced dynamics in different scenarios.

In case of the ocean, the primitive equations with the Boussinesq and hydrostatic approximation comprise of the evolution equations of horizontal velocities and buoyancy. A balanced quasi-geostrophic model under the same approximation reduces to a single evolution equation of potential vorticity. This evolution equation corresponds to the low frequency mode of the flow. On the contrary, the primitive equations for such a system, as well as for a system without hydrostatic approximation, consist of one low frequency mode and two high frequency modes (or fast modes) associated with IGWs. In this way, a balanced model excludes the high frequency dynamics. But the high and low frequency motions do interact and exchange energy. This interaction between the low and high frequency motions is crucial to understand the energy transfers between these motions which influence the ocean dynamics as well as the ocean energy cycle. For this, a brief review of the energy sources and sinks in the ocean is laid out.

In the ocean, the energy primarily from penetrative solar heating, winds, tides, and geothermal heating drives the large-scale ocean circulation and feeds energy into the balanced motions. This energy is lost ultimately to the smallest molecular oceanic scales. So, the energy enters and leaves the ocean at scales of motions furthest apart in the oceanic spectra. But the tricky riddle that the ocean has in store is: how do the balanced motions lose their energy to the smallest dissipative scales? This question is further complicated by the preferred energy transfer of the balanced flow towards large scales. Therefore, this energy pathway that links the balanced flow to the unbalanced ones is the key to get a complete picture of the energy budget of the ocean. IGWs are a potential candidate in this energy pathway– a notion that resonates with previous studies (e.g. Bartello, 1995; Brüggemann and Eden, 2015; also Chapter 2)

Numerous mechanisms have been proposed for the possible energy pathway(s) through which energy from the balanced flow can be lost to IGWs on the way to molecular dissipation. Internal mechanisms that can generate IGWs from the balanced flow include geostrophic adjustment (e.g. Rossby, 1938; Blumen, 1972; Bartello, 1995), spontaneous generation (e.g. review by Vanneste, 2013), and loss of balance (e.g. Molemaker et al.,

2005). External ingredients can also stimulate the excitation of IGWs from the balanced flow, mainly by winds, tides, and topography. In this paper we discuss gravity wave emission in three different scenarios: spontaneous emission, convective instability at fronts, and lateral boundary instability. Spontaneous emission is the generation of IGWs by the balanced flow spontaneously without any external stimulus (see reviews by Vanneste, 2013 and Plougonven and Zhang, 2014). Spontaneous emission is a ubiquitous phenomenon and is inevitably tied to the emission of IGWs (e.g. Vanneste, 2013). Convective instability at fronts generated by balanced flow arises due to an unstable density column which can lead to isopycnal overturning followed by emission of IGWs. Such processes during frontogenesis have been found to generate IGWs via frontogenesis (e.g. Shakespeare and Taylor, 2015, 2016). Lateral boundary instability can result in emission of IGWs when a balanced eddy collides with a lateral wall. This process is similar to the interaction of a balanced vortex with topography which can result in emission of unbalanced wave motions (e.g. Dewar and Hogg, 2010).

However, the understanding of these processes is hindered by the fast timescales of these waves which not only makes them hard to observe but also difficult to resolve in models. The other aspect that complicates the issue further is the coupling between the high frequency unbalanced and low frequency balanced component of the flow. In different dynamical regimes, the timescales of the two different motions are coupled to different extents. In particular, for a regime with Rossby number $\mathrm{Ro} = \mathscr{O}(1)$, the timescale separation between the high and low frequency modes becomes much smaller (Zeitlin, 2008; Vanneste, 2013) and the identification and separation of the wave motions from the balanced flow becomes a challenging task. Several procedures have been developed to address this issue as discussed below, and in this paper we introduce a new procedure.

A family of quasi-geostrophic balanced models (e.g. Vallis, 1996 and references therein) have evolved with the aim to achieve a balanced state. These balanced procedures can be modified and applied to diagnose IGWs. One such diagnostic tool is the quasi-geostrophic omega equation (e.g. Houghton et al., 1971; Hoskins et al., 1978, 2003), based on the quasi-geostrophic potential vorticity equation, that has been applied to diagnose IGWs from the flow field (e.g. Pinot et al., 1996; Danioux et al., 2012; Nagai et al., 2015). However, in a primitive equation model, Pinot et al., 1996 evaluated the diagnostic performance of the omega equation in diagnosing vertical velocities at ocean fronts and have pointed out the importance of including the ageostrophic advection to accurately diagnose vertical velocities. In agreement, the new diagnostic tool introduced in this paper has the capability

to diagnose ageostrophic horizontal velocities aiming to provide improved detection of unbalanced wave signals.

Another balance procedure is the Optimal Potential Vorticity (OPV) method introduced by Viúdez and Dritschel (2004) which begins with an IGW-permitting dynamics to achieve a balanced solution with minimal IGW activity. This method works by iterations, where IGWs are removed in every iteration and potential vorticity (PV) is restored, which grows slowly over a long enough time period towards a prescribed PV field which is that of the state of balance. However, at the end of the iteration weak residual IGWs are still present due to spontaneous generation (cf. Viúdez and Dritschel, 2004 and McKiver and Dritschel, 2008). An alternative balance procedure is the nonlinear quasi-geostrophic (NQG) balance put forward by McKiver and Dritschel (2008). This method works by removing the ageostrophic component of the horizontal vorticity from the equations to obtain a pair of diagnostic balance relations which filters inertia gravity waves.

A different iterative procedure, known as the non-linear normal mode initialization procedure has been introduced by Machenhauer (1977), Baer (1977), and Baer and Tribbia (1977) that allows for adjustments to the initial conditions in order to minimize the tendency of the system to generate IGWs. This has been further extended to the hydrostatic system by Leith (1980) and to the non-hydrostatic set of equations by Bartello (1995). Recently, the non-linear initialisation scheme of Baer and Tribbia (1977), has been applied by Kafiabad and Bartello (2016, 2017) for balanced rotating dynamics to identify the energy cascades for differently initialized balanced regimes. In Chapter 2, the non-linear initialization procedure of Machenhauer (1977) is implemented in a primitive equation ocean model by to diagnose IGWs for a range of dynamical regimes from ageostrophic to quasi-geostrophic.

In this work, we introduce a new diagnostic procedure named the quasi-geostrophic (QG) filter implemented in a consistent way on the discrete level to filter and hence diagnose IGWs. The QG filter is based on quasi-geostrophic equations and uses a potential vorticity inversion method with the additional capacity to diagnose the horizontal ageostrophic components, which is absent in the traditional QG-omega equation based methods. This new feature allows for extraction of the ageostrophic balanced part from the flow to obtain unbalanced motions with minimal balanced part. A QG model can be used to filter the high frequency unbalanced dynamics and it can also be used to initialize the model with QG balance conditions such that the emission of fast motions is reduced. The gravity waves diagnosed from the QG filter are then used to understand gravity wave emis-

sion in different scenarios such as spontaneous emission, lateral boundary instability and convective instability.

The paper unfolds as follows: the theoretical background and a demonstration of the QG filter is described in Section 3.2. A comparison with the normal mode initialization technique of Machenhauer is also presented and discussed. Gravity waves diagnosed with the QG filter are used to study gravity wave emission in different scenarios *viz.* spontaneous emission, convective instability, and lateral boundary instability are presented in Section 3.3. A summary and the main conclusions are stated in Section 3.4.

3.2 Quasi-geostrophic filter

3.2.1 Theoretical background

The primitive equations are given by

$$\partial_t \ \boldsymbol{u} + f\underline{\boldsymbol{u}} = -\boldsymbol{\nabla} p + \dot{\boldsymbol{u}} \quad , \quad \partial_t b + wN^2 = \dot{b} \tag{3.1}$$

supplemented by the diagnostic relations $b = \partial_z p$ and $\boldsymbol{\nabla} \cdot \boldsymbol{u} + \partial_z w = 0$, where $\dot{\boldsymbol{u}}$ and $\dot{b}$ contain the advection terms, friction and mixing. The Coriolis frequency f and the stability frequency N are assumed to be constant. The vorticity and divergence equation are given by

$$\partial_t \ \boldsymbol{\nabla}^2 \psi - f \partial_z w = \underline{\boldsymbol{\nabla}} \cdot \dot{\boldsymbol{u}} \quad , \quad \partial_t \ \boldsymbol{\nabla}^2 \phi - f \boldsymbol{\nabla}^2 \psi = -\boldsymbol{\nabla}^2 p + \boldsymbol{\nabla} \cdot \dot{\boldsymbol{u}} \tag{3.2}$$

introducing a streamfunction ψ and potential ϕ with $\underline{\boldsymbol{\nabla}} \cdot \boldsymbol{u} = \boldsymbol{\nabla}^2 \psi$ and $\boldsymbol{\nabla} \cdot \boldsymbol{u} = \boldsymbol{\nabla}^2 \phi$. All vectors are two-dimensional here and $\underline{\boldsymbol{u}}$ denotes anticlockwise rotation of $\boldsymbol{u}$ by 90^o.

Now we neglect the term $\partial_t \ \boldsymbol{\nabla}^2 \phi$ in the divergence equation which eliminates linear gravity waves (e.g. Olbers et al., 2012), and which is equivalent to the quasi-geostrophic assumption. With $p = p^g + p^a$, where p^a is the ageostrophic pressure, which is assumed to be small compared to the geostrophic pressure p^g it follows to leading order

$$\boldsymbol{\nabla}^2 \psi = -\boldsymbol{\nabla}^2 p_g / f \ \rightarrow \ \psi = p_g / f \tag{3.3}$$

and to first order that $\boldsymbol{\nabla}^2 p^a = \boldsymbol{\nabla} \cdot \dot{\boldsymbol{u}}$. The relation involving the ageostrophic pressure is needed below to diagnose the ageostrophic velocities.

Now we express the vertical velocity in the vorticity equation with the remaining equations to obtain the potential vorticity equation. Buoyancy is also split into a dominant geostrophic part b^g and a ageostrophic part for which both the hydrostatic relation $b = \partial_z p$ holds, and thus

$$\partial_t q = \underset{\neg}{\boldsymbol{\nabla}} \cdot \dot{\boldsymbol{u}} + \frac{f}{N}\partial_z \dot{b}^g \quad , \quad \left(\boldsymbol{\nabla}^2 + \frac{f^2}{N^2}\partial_{zz}\right)\psi = q = \underset{\neg}{\boldsymbol{\nabla}} \cdot \boldsymbol{u} + \frac{f}{N}\partial_z b \tag{3.4}$$

where $\dot{b}^g$ denotes the advection and mixing operator applied to b^g. The potential vorticity equation Eq. (3.4) can be integrated in time using suitable vertical and lateral boundary conditions as specified below. The second part of Eq. (3.4) can be used to obtain ψ from a given field of $\boldsymbol{u}$ and b using equivalent boundary conditions. It is not exactly correct since we replaced b^g with b, which needs to be done since b^g is unknown to us before inverting the second part of Eq. (3.4) for ψ. The vertical velocity and ageostrophic velocity components can be diagnosed from the buoyancy and momentum equation as

$$w = w^a = (\dot{b}^g - \partial_t f \partial_z \psi)/N^2 \ , \ \boldsymbol{u}^a = \boldsymbol{\nabla}\phi = (-\underset{\neg}{\boldsymbol{\nabla}} p^a + \underset{\neg}{\dot{\boldsymbol{u}}} + \partial_t \boldsymbol{\nabla}\psi)/f \tag{3.5}$$

where p^a can be diagnosed from $\boldsymbol{\nabla}^2 p^a = \boldsymbol{\nabla}\cdot\dot{\boldsymbol{u}}$, while the time tendency of ψ can be obtained by inverting the first part of Eq. (3.4). Note that the diagnosed $\boldsymbol{u}^a$ from equation Eq. (3.5), however, is incomplete since the rotation part of $\boldsymbol{u}^a$ is not included in the diagnosed $\boldsymbol{u}^a$. This diagnosed $\boldsymbol{u}^a$ might have an affect on the interpretation of the horizontal velocity related to the unbalanced motions. However, in this analysis mostly the vertical component of the velocity is used to interpret the unbalanced motions. The discrepancy related to the diagnosed $\boldsymbol{u}^a$ in Eq. (3.5) is being investigated currently.

At the upper- and lowermost levels we need to implement the surface and bottom boundary conditions, $w = 0$ at $z = 0, -h$ which becomes

$$\partial_t f \ \partial_z \psi = \dot{b} \quad \rightarrow \quad \partial_z \psi = b/f \quad \text{at} \quad z = 0, -h \tag{3.6}$$

This is again not exactly correct since we replaced b^g with b, but b^g remains unknown diagnosing ψ from the model variables. For lateral boundary conditions some auxillary conditions are needed which are given in Appendix A. In Appendix B we detail the consistent discrete version of all diagnostic relations which we use.

We now invert Eq. (3.4) for ψ using $\boldsymbol{u}$ and b from an integration of an idealized primitive equation model. From ψ we diagnose the geostrophic velocities $\boldsymbol{u}^g$, and from the diagnostic

relations Eq. (3.5) we diagnose the ageostrophic velocities $\boldsymbol{u}^a$ and w^a. We then interpret the residual $\boldsymbol{u} - \boldsymbol{u}^g - \boldsymbol{u}^a$ and $w - w^a$ as gravity wave activity.

3.2.2 Demonstration of the method

To demonstrate the operation of the QG filter, we show a simple test case which is essentially an initial value problem. The test case model setup consists of a double periodic domain with a lens of warmer water in the middle of the domain in a constant stratification. To show case how the QG filter works, the test case is initialized in two different ways: first, the test case setup is initialized using only the buoyancy structure of the lens (this test case is referred to as INI_NO); second the model is initialized with quasi-geostrophic balanced conditions obtained by applying Eq. (3.4) to the buoyancy structure of the lens (this test case is referred to as INI_QG). Further, the skill of the QG filter is compared against a third test case in which the initial conditions are balanced using the non-linear initialization procedure of Machenhauer (1977), which has been implemented as a diagnostic for gravity waves in Chapter 2. The third test case is referred to as INI_MH and achieves balance by an iterative procedure that eliminates the high frequency unbalanced motions with each iteration. Presently we have used three iterations, although in Machenhauer's method has been shown to attain quasi-geostrophic balanced conditions after the first iteration (Leith, 1980).

Each of the three model setups have the same domain size and resolution. The zonal and meridional extents are set to 100 km with 199 grid points each, and the domain depth is 1000 m with 100 grid points. The Coriolis frequency and the initial stratification frequency are $f = 10^{-4}$ and $N_0 = 5 \times 10^{-3}$ respectively. Fig. 3.1 shows the initial state of the model for the horizontal velocities with buoyancy contours. Note that the test case INI_NO has only an initially buoyancy structure (no horizontal velocities), similar to the buoyancy contours in Fig. 3.1. The snapshots of vertical velocities at different time intervals are shown in Fig. 3.2 for differently initialized models. The leftmost panel in Fig. 3.2 refers to a model where no balanced initial conditions are imposed (INI_NO). The middle and the rightmost panels show the vertical velocities for a model initially balanced with the quasi-geostrophic balanced conditions (INI_QG) and the balanced condition obtained from the Machenhauer's non-linear initialization procedure (INI_MH) , respectively. As is evident from the figure, the magnitude associated with the vertical velocity for INI_NO is much higher than INI_QG and INI_MH. The emission of waves, as can be seen from the vertical velocity, is similar to the process of geostrophic adjustment. Owing to this geostrophic

adjustment there is a significant amount of wave emission in a state which is not initially balanced.

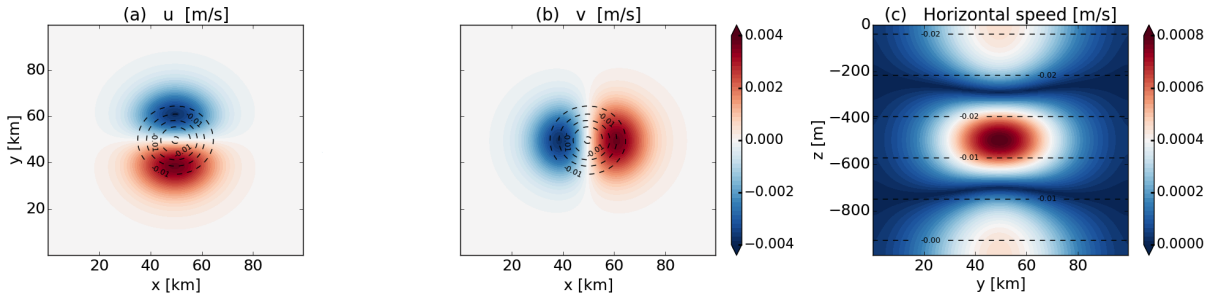


Figure 3.1: Initial state of the model used for INI_NO, INI_QG, and INI_MH. (a) and (b) Snapshots of zonal and meridional velocities respectively at the initial model state with buoyancy contours in black. (c) Horizontal speed variation with depth and buoyancy contours in black. The initial state for INI_NO consists of only buoyancy shown in black contours. The initial states for INI_QG and INI_MH comprises of both buoyancy and horizontal velocities. Note that the vertical axis in (a) and (b) corresponds to the meridional extent and in (c) to the depth.

When the model state is initially in a quasi-geostrophic balance (INI_QG), there is a significant reduction in the emission of the waves. This implies that the flow is already in a quasi-balanced state and emits much smaller amount of waves than the initially non-balanced initial state. Further, the Machenhauer's initial condition of balance imposes a stricter state of balance and allows for a state where the wave emission is almost negligible. Note that the Machenhauer's condition imposed here is calculated for the discrete version of the method with discrete eigenvectors. The use of analytical eigenvectors in computation yields a slight error. The discretized version of the equations used in Machenhauer's method and the discrete eigenvectors are detailed in Appendix C.

Furthermore, we assess the efficiency of the QG filter in frequency wavenumber space. To this purpose, the QG filter has been used in conjunction with an idealized numerical model representative of the Eady model (Eady, 1949) of baroclinic instability. The model setup is similar to the one described in Chapter 2, and consists of a baroclinically unstable flow on an f-plane with a double periodic domain to exclude boundary instabilities. The model setup used here, however, has a higher resolution: $nx = ny = 240$ grid points in the horizontal direction and $nz = 80$ grid points in the vertical direction. The model is forced by the restoring of the zonal mean buoyancy and zonal mean velocity towards its initial state. This restoring is set by the Richardson number (Ri) (as in Chapter 2),

which differs for different dynamical regimes. Small and large Ri represent ageostrophic and quasi-geostrophic regimes respectively. Any other forcing is absent from the model.

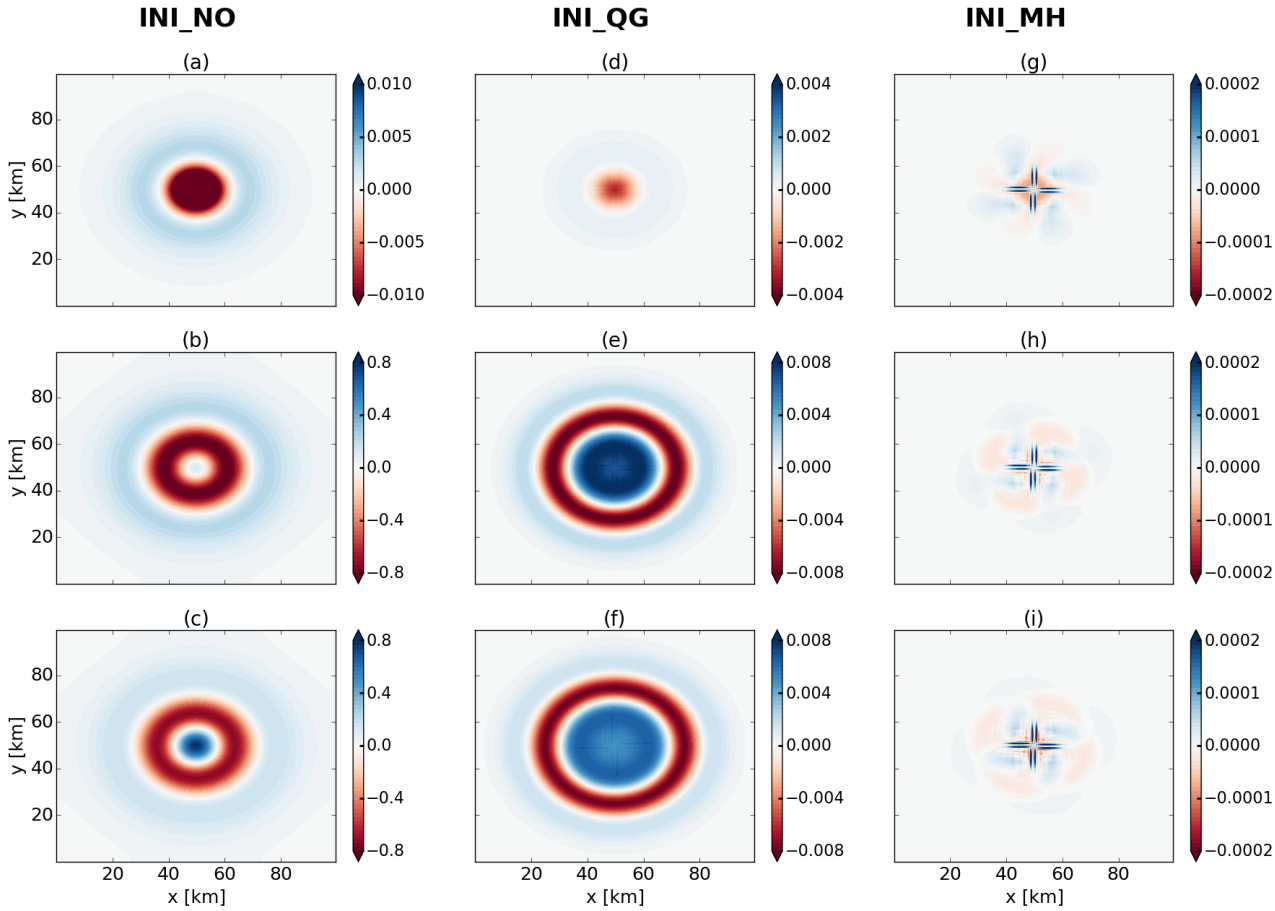


Figure 3.2: Snapshots of vertical velocity w in 10^{-3} ms^{-1} from a simple model configuration initialized with differently balanced conditions as indicated by the experiment name. (a), (b), and (c) Snapshots from a model without any imposed balanced condition (INI_NO). (d), (e) and (f) show snapshots from a model initialized using the quasi-geostrophic balance conditions (INI_QG). (g), (h) and (i) show the snapshots from the model when the model is initialized with the balanced conditions obtained after three iterations from the normal mode initialization technique of Machenhauer (INI_MH). Note that the values of w are shown with a factor of 10^{-3}, such that for Machenhauer's method (INI_MH) the presence of vertical velocity is negligible. Also note that the colorscales are different.

The gravity wave component obtained from the QG filter is applied to this model for different dynamical regimes. Using the QG filter, the full flow field of interest can be

segregated into the balanced and unbalanced wave counterparts. From the unbalanced horizontal velocity components provided by the QG filter, the kinetic energy (KE) associated with the unbalanced dynamics can be estimated. This can be compared against the KE related to the balanced dynamics to assess the distribution of KE in the balanced and unbalanced components.

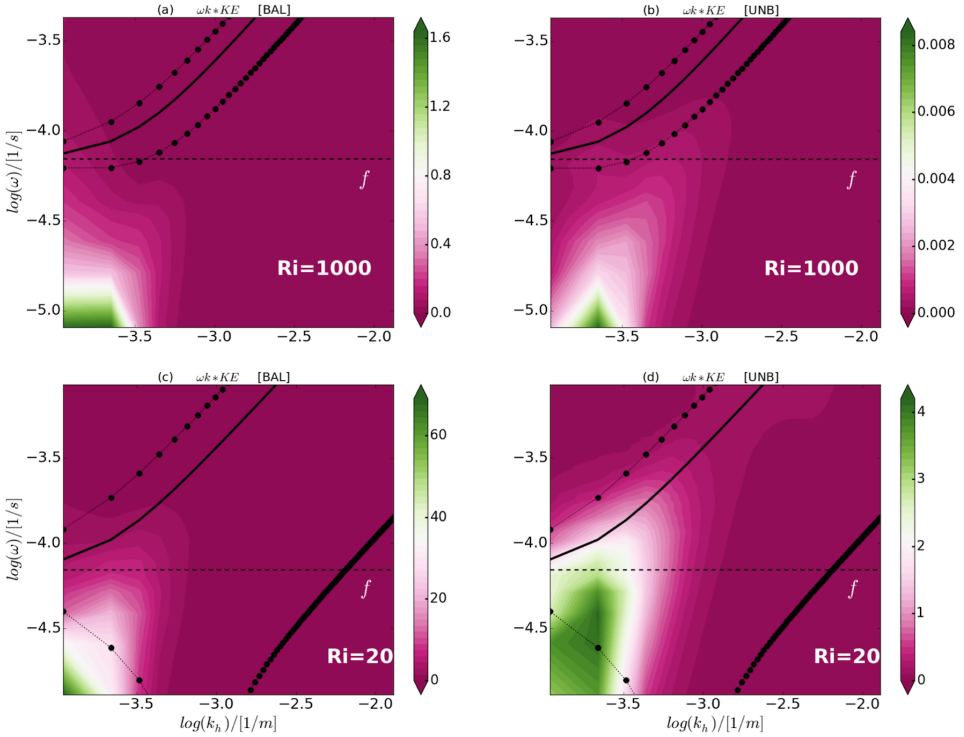


Figure 3.3: Variance preserving frequency-wavenumber spectra of KE for a high and low Ri regime, from the model setup described in Chapter 2 but with a higher resolution. Balanced and unbalanced component obtained from the QG filter are shown for Ri=1000 in (a,b) and for Ri=20 in (c,d). The balanced component is computed as u_g+u_a whereas the unbalanced components are computed as $u-u_g-u_a$ and are interpreted as gravity waves. The spectra shown are depth averaged as well as averaged over 20 chunks from the high resolution simulations for a baroclinically unstable setup. In each subplot: the color shading shows the respective KE as a function of frequency (ω) and horizontal wavenumber (k_h); the dashed black line indicates the Coriolis frequency (f), the solid black curve indicates the gravity wave dispersion relation and lines with black circles show the Doppler shifted dispersion relation. Note that the gravity-wave branch is enclosed by the Doppler-shifted extrema. Notice that the colorscales differ.

An example for two very different regimes with different Ri is shown in Fig. 3.3. The balanced component is denoted by $u_g + u_a$ whereas the unbalanced components are computed as $u-u_g-u_a$ and are interpreted as gravity waves. The spectra shown in Fig. 3.3 are

depth averaged as well as averaged over 20 chunks from the high resolution simulations. A frequency-wavenumber spectra combined with the dispersion relation of gravity wave can help to identify the scales in spectral space with high IGW activity. The gravity wave dispersion relation shown in Fig. 3.3 (solid black lines) is for the first baroclinic mode and takes the form: $\omega^2 = f^2 + c_n^2 k_h^2$ with $n = 1$. Acceleration due to gravity is denoted by g and $c_n = NH/(n\pi)$, where N is the stratificaion frequency and H is the depth. In addition, Fig. 3.3 also shows the Doppler-shift, $\omega \pm U_0 k_h$, where U_0 is the mean flow (see Table 3.2). The possible region for gravity wave activity is assumed to lie approximately within the envelope of the Doppler-shifted extrema, and is called as the gravity-wave branch.

The spectra in Fig. 3.3 is compared for Ri = 20 and Ri = 1000 to indicate gravity wave activity in contrasting (small and large) Ri regimes. The balanced part of the flow is expected to span the lower frequencies and lie outside of the gravity-wave branch in the spectral space. On the contrary, the fast unbalanced motions are expected to lie in the region inside the gravity-wave branch and towards higher frequencies. In line with our expectations, we find that the spectra in Fig. 3.3 shows indeed a reasonable separation of the balanced and unbalanced motions. Also, the unbalanced motions related to gravity wave activity tend to be higher for a small Ri regime than a large Ri regime. This is in accordance with the result presented in Chapter 2.

3.3 Gravity wave emission in different scenarios

We now apply the results from the QG filter to detect gravity wave activity in different scenarios that can generate internal gravity wave in the ocean. Here we present three mechanisms as examples of gravity wave emission from the balanced flow, *viz.* spontaneous emission, convective instability at fronts, and lateral boundary instability. Note that there are several other known mechanisms of IGW generation in the ocean, we simply discuss three of those mechanisms here.

Idealized model setups used to represent these scenarios are spindown simulations of a baroclinic flow initialized with a constant vertical (N_0) and meridional stratification (M_0). In case of double periodic domain, the meridional stratification is sinusoidal with M_0 as amplitude. N_0 and M_0 determine the initial Richardson number (Ri) of the flow defined as the ratio of the vertical density stratification and vertical shear of the horizontal velocity:

$$\mathrm{Ri} = \frac{{N_0}^2 f^2}{{M_0}^4} \tag{3.7}$$

Symbol	Meaning	SPON	CONV	LBI
nx, ny	Grid points in (x, y)	120, 120	240, 240	240, 240
nz	Grid points in z	40	80	80
Ri	Richardson number	500	100	400
α	Aspect ratio	0.02	0.02	0.01
CFL	CFL criteria	0.005	0.004	0.01
Ek	Ekman number	0.01	0.05	0.01
f	Coriolis frequency (s^{-1})	$7e^{-5}$	$7e^{-5}$	$1e^{-4}$
L_x	Length of the domain (m)	$2L_0$	L_0	$2L_0$
Domain	-	Double periodic	Double periodic	Zonally periodic
Lateral boundaries	-	None	None	Meridional

Table 3.1: Model parameters for the setups in different scenarios: spontaneous emission (SPON), convective instability (CONV), and lateral boundary instability (LBI). The aspect ratio α is the ratio of vertical to horizontal scales.

The combination of stratification and shear give rise to perturbations as the baroclinic instability sets in and the most dominant of these perturbations is referred to as the fastest growing mode (Eady, 1949; Stone, 1966). In our model, the length scale L_0 associated with the fastest growing mode and the corresponding wavenumber k_{max} are based on Stone's formulation (Stone, 1966). The corresponding length scale L_0 is:

$$L_0 = \frac{2\pi}{k_{max}} = 2\pi\sqrt{\frac{2}{5}}\sqrt{\frac{\mathrm{Ri}+1}{\mathrm{Ri}}}L_r \tag{3.8}$$

where $L_r = N_0 H/f$ is the Rossby radius of deformation.

The setup is similar to the one described in Chapter 2. The numerical code for the model is identical to the one in Eden (2016). The setup has different values of numerical parameters for different scenarios of gravity wave emission. The differences in the model setup in the three scenarios are listed in Table 3.1, the other set of model parameters listed in Table 3.2 are be derived from the values in Table 3.1.

Symbol	*Meaning*	*Value*	*Unit*
N_0	Brunt-Väisälä frequency	f/α	s^{-1}
M_0	Meridional stratification	$\sqrt{fU_0/H}$	s^{-1}
U_0	Mean flow	$N_0 H/\sqrt{\mathrm{Ri}}$	$\mathrm{m\ s}^{-1}$
dx(=dy)	Horizontal resolution	L_x/nx	m
dz	Vertical resolution	H/nz	m
dt	Time step	dx CFL$/U_0$	s
A_h	Biharmonic horizontal friction	Ekfdx^4	$\mathrm{m^2\ s}^{-1}$
A_v	Harmonic vertical friction	Ekfdz^2	$\mathrm{m^2\ s}^{-1}$

Table 3.2: An overview of the model parameters for all the scenarios derived from values in Table 3.1.

3.3.1 Scenario I: Spontaneous emission

Spontaneous emission refers to the emission of gravity waves from balanced flow spontaneously without any external forcing. It is one of the important mechanisms for the balanced flow to lose its energy towards small scales. Spontaneous emission is more prominent for a Ro $= \mathcal{O}(1)$ regime for which the timescale separation between the balanced and unbalanced motions is gets smaller and gravity wave activity is much enhanced. In contrast, a Ro $\ll 1$ regime shows a very weak to nearly negligible emission of gravity waves. For this reason, such a regime is assumed to be largely balanced.

In such a regime, it becomes particularly challenging to detect gravity wave signals because the signal is faint to detect by traditional means. Here we aim to detect gravity wave signal from a baroclinically unstable flow representing a Ro $\ll 1$ regime. The residual of the vertical velocity $w - w_a$ obtained from the QG filter has the advantage that the balanced ageostrophic part has been further removed from the vertical velocity field, so that the residual vertical velocity sincerely represents the unbalanced part. This is an improvement over the traditional methods to interpret gravity wave signals which rely only on the vertical velocity. As is obvious, in a Ro $\ll 1$ regime the vertical velocity signature would be weak because the dominant balanced part would still be intact in the velocity field, making the weak gravity wave signal extremely hard to detect. With the QG filter diagnostic, we expect to capture even these small signals in a largely balanced regime. The setup in experiment SPON consists of a channel flow with double periodic boundaries as described above. The baroclinic flow is allowed to evolve and dissipate in a spin down run. After a certain time period, i.e. after the baroclinic instability sets in,

the emission of wave packets can be seen. These wave packets appear to move along with the baroclinic eddies, similar to atmospheric numerical experiments (e.g. Plougonven and Snyder, 2007).

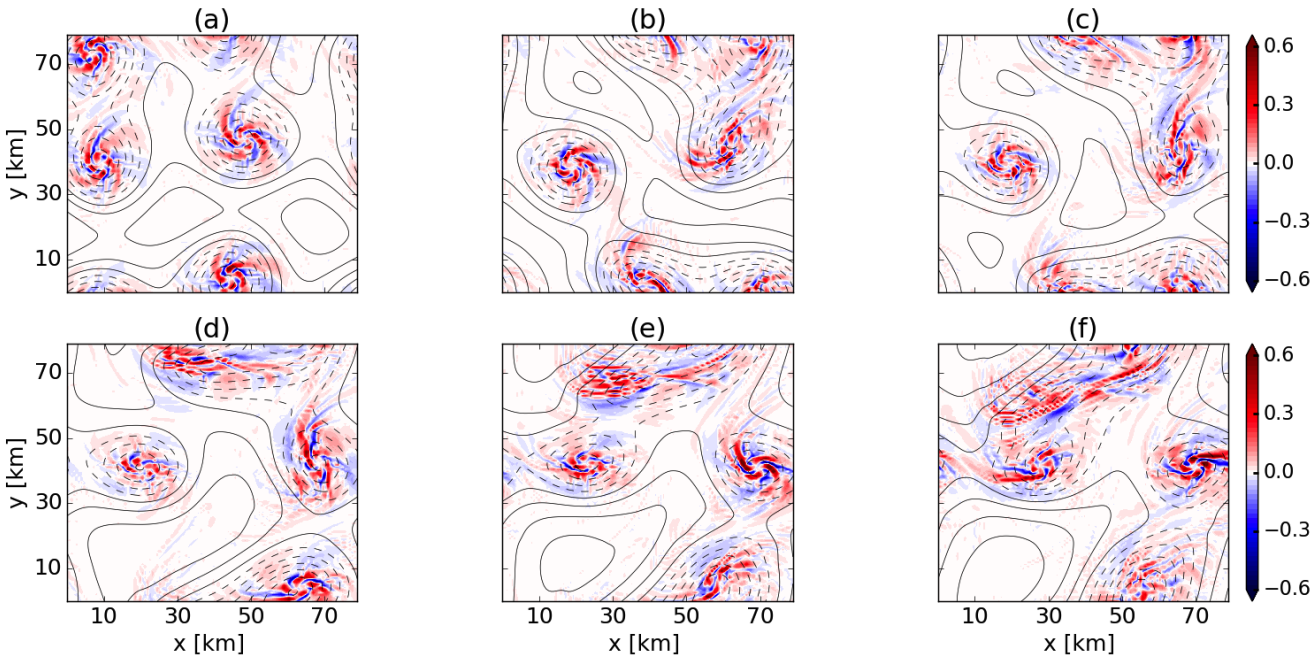


Figure 3.4: Snapshots of $w - w_a$ in 10^{-3} ms^{-1} and streamfunction ψ (black contours) in a baroclinically unstable flow in experiment SPON.

Spontaneous emission in the model setup SPON is depicted in Fig. 3.4 for different time steps. The residual vertical velocity $w - w_a$ represents the wave activity associated with the unbalanced wave field and is shown in Fig. 3.4 with contours of the streamfunction ψ. After the baroclinic instability sets in, it spawns a field of balanced eddies. Balanced eddies then generate unbalanced gravity waves by spontaneous emission as in the example shown in Fig. 3.4. Subsequent snapshots show the evolution of the flow and propagation of the gravity wave packets with the flow. The wave emission in Fig. 3.4a is shown about 85 days after the initial time, but the wave emission by eddies can be seen at an even earlier stage. After the next four days, the eddy field changes to Fig. 3.4b. The snapshots shown thereafter are each set apart by 20 hours (Fig. 3.4(b-f)) to show the evolution of the eddy field and the concurrent wave emission over short time scales. At later stages in Fig. 3.4, a form of secondary instability begins to emerge towards the northern side of the channel (Fig. 3.4(d-f)) (and slight occurrences on the southern side). Snapshots of unbalanced

components of the horizontal velocities obtained from the QG filter, computed as $u-u_g-u_a$, are shown in Fig. 3.5. The horizontal unbalanced velocities provide information about the horizontal components of the wave packets seen in Fig. 3.4. These components also appear encircled by the eddy field like the wave packets in Fig. 3.4, however, the magnitude of horizontal unbalanced components is much larger than the vertical unbalanced components. The horizontal and vertical components can be combined to retrieve other properties of the emitted waves, such as wave propagation from the polarization relation.

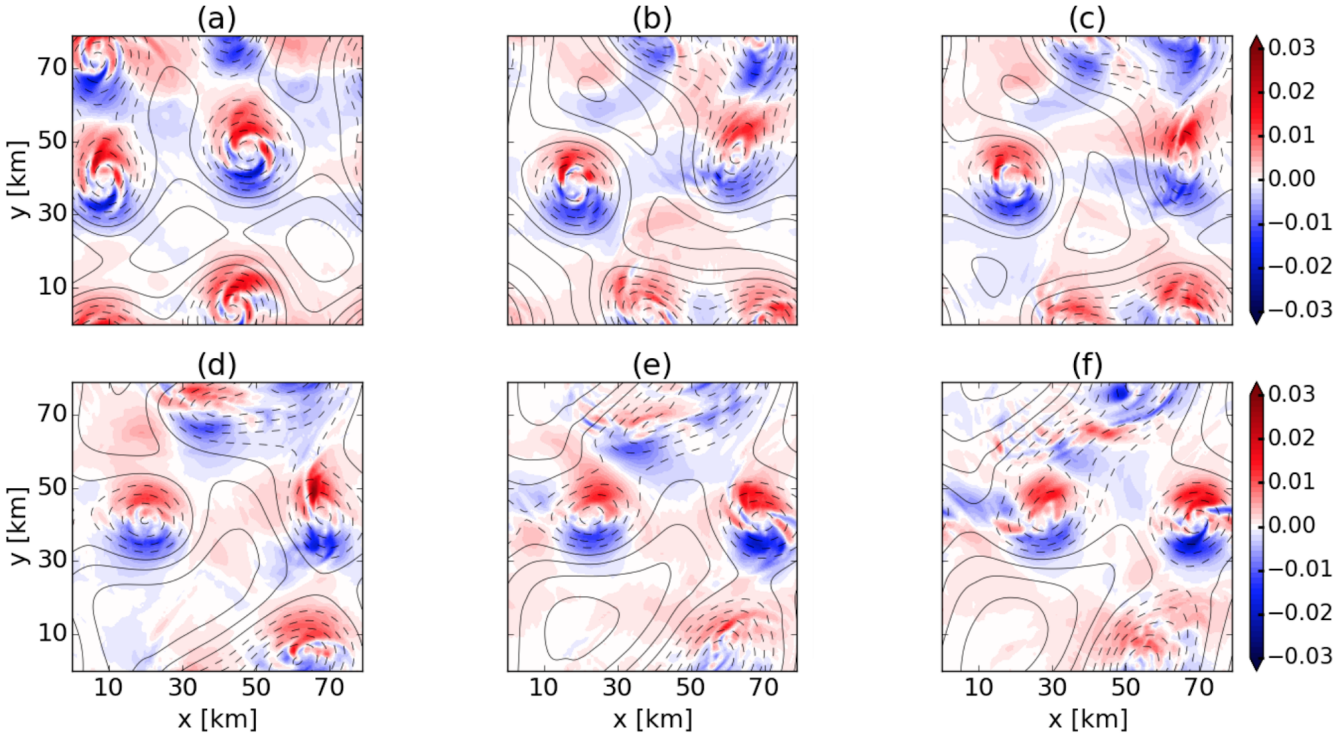


Figure 3.5: Snapshots of $u - u_g - u_a$ in ms^{-1} and streamfunction ψ (black contours) in a baroclinically unstable flow in experiment SPON. The snapshots correspond to the snapshot in Fig. 3.4.

These instabilities manifest themselves in wave trains of much smaller spatial scales than the gravity wave field resulting from spontaneous emission. These secondary instabilities could be related to spontaneous emission but they could also be a consequence of interactions between unbalanced wave components generated by balanced eddies in the first place. It is instructive to also look at the local Rossby number to further characterize these motions. Snapshots of local Rossby number are shown in Fig. 3.6 for the corresponding to the time steps shown in Fig. 3.4 and Fig. 3.5. Clearly, the region close to the eddies

filled by the gravity wave packets in Fig. 3.4 are the regions of $\mathcal{O}(1)$ Rossby number, a regime which is characterized by unbalanced motions.

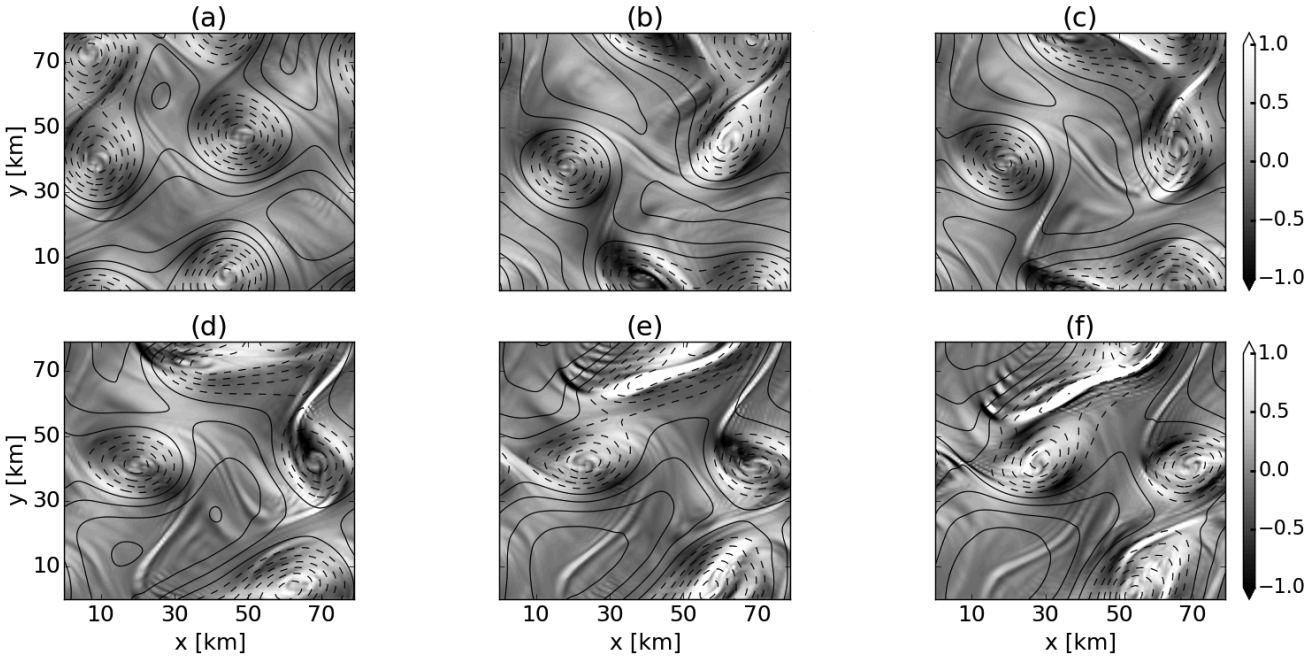


Figure 3.6: Snapshot of local Rossby number, $\mathrm{Ro} = \zeta/f$, for a baroclinically unstable flow and streamfunction ψ (black contours) in experiment SPON. Here, ζ is obtained from the full components of horizontal velocities. The snapshots correspond to the snapshot in Fig. 3.4.

3.3.2 Scenario II: Convective instability

The experiment CONV is similar to SPON but with higher horizontal and vertical resolution (see Table 3.1). The setup represents a baroclinically unstable flow with double periodic boundaries and can be run to depict different dynamical regimes. We discuss here an example for a case with $\mathrm{Ri} = 100$, while the experiment SPON was for $\mathrm{Ri} = 500$.

Convective instability in a stably stratified fluid can arise when denser fluid gets lifted above lighter fluid and this unstable situation can cause overturning of the isopycnals. Such a situation co-occurs with a sharp localized decrease in the stratification frequency as shown in Fig. 3.7, causing the isopycnals to overturn, indicating the onset of convective instability. The gradual sharpening of the density gradients lead to the formation of fronts which can result in such an unstable situation. An example of such a front for a section of the domain

is presented in Fig. 3.8a which shows the residual vertical velocity component $w - w_a$ from the QG filter and streamfunction contours in experiment CONV. The snapshot in Fig. 3.8a is 20 days after the model initial time and the subsequent snapshots (Fig. 3.8(c-d)) are each 10 hours apart. Frontogenesis can further intensify a front by further sharpening the gradients (Fig. 3.8b). The resulting convective instability can eventually lead to a frontal collapse which is accompanied by the generation of unbalanced processes such as internal waves (Fig. 3.8(c,d)).

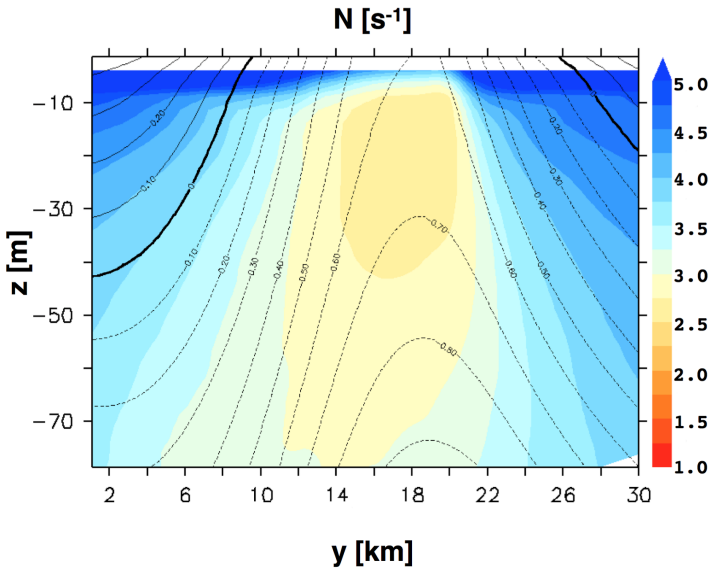


Figure 3.7: A section of the local Brunt-Väisälä frequency $N = \sqrt{\partial_z b}$ in 10^3 ms^{-1} and temperature (black contours) across a typical front. The snapshot shown is after about 20 days from the initial model state in experiment CONV.

Convective instability is generally known to cause the breaking of internal waves, which by breaking generate small-scale turbulence and contribute to diapycnal mixing. Here we show an example of convective instability as a generation mechanism for internal waves in a baroclinically unstable flow. However, the emitted internal waves might then break as they propagate. In Fig. 3.8, IGW emission can be seen at the unstable front, but in the subsequent snapshots waves appear to radiate away from the front and then decay as the front collapses. Thus internal waves generated due to convective instability are likely short-lived as they appears to decay soon after their genesis.

In comparison to the previous scenario of spontaneous emission (experiment SPON), convective instability shows a higher magnitude of wave generation. Unlike spontaneous emission for which wave emission increases for a small Ri regime (Chapter 2), the wave

generation through convective instability is likely to be a local phenomenon and further work in this direction will be undertaken to test the Ri dependence of the mechanism. It might be that the IGWs seen in Chapter 2 are actually generated by convective instability. As the generation occurs at the site of the unstable front the convective instability mechanism is expected to be more important locally for wave generation.

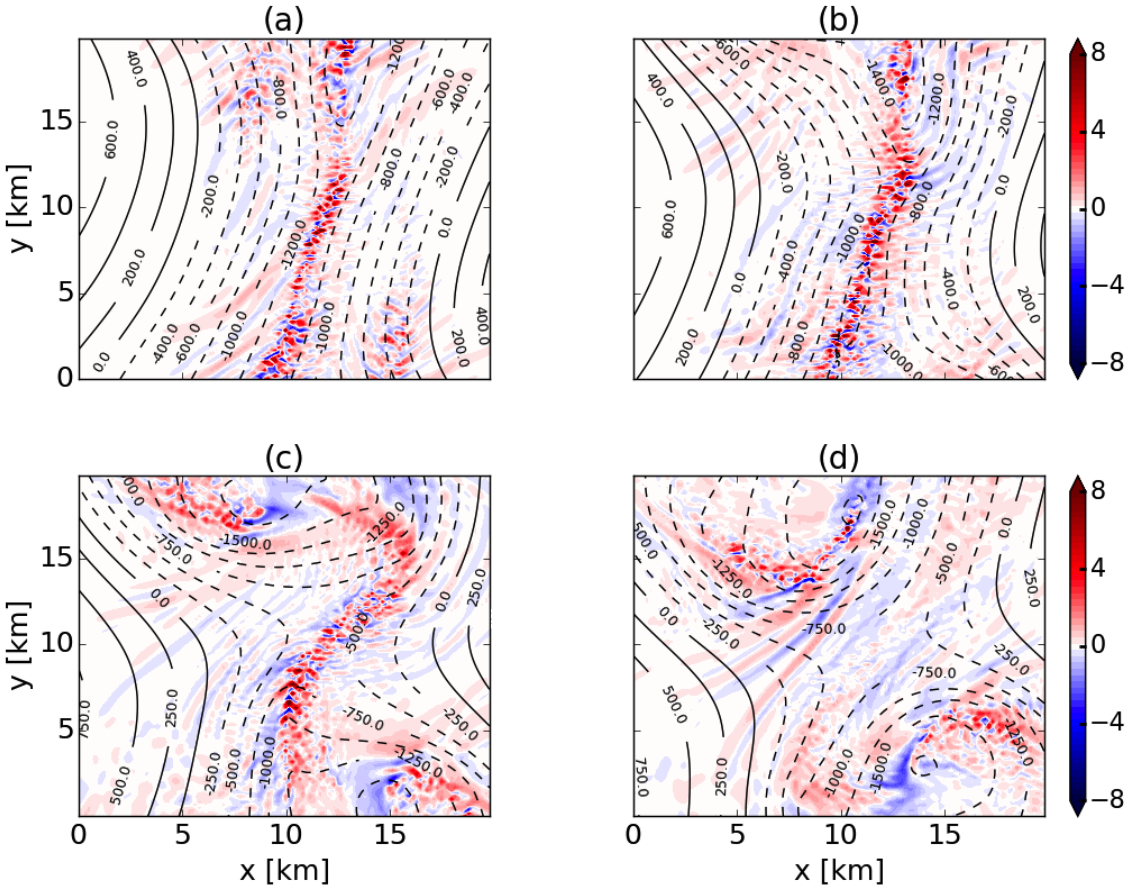


Figure 3.8: Subsequent snapshots of $w - w_a$ in 10^{-3} ms^{-1} at 20 m depth and streamfunction ψ (black contours). Snapshot (a) is the model state after about 20 days from the initial model state from experiment CONV, the following snapshots are 10 hours apart each.

3.3.3 Scenario III: Lateral boundary instability

So far we have seen internal wave generation from balanced flow through internal mechanisms, i.e. without the presence of any external parameters. In addition, to avoid flow interaction with the lateral boundaries, we applied double periodic boundaries in the numerical

model. However, the real ocean is not devoid of boundaries— it has lateral boundaries, sea mounts, ridges, and more complex topography— and interaction between balanced motions and the oceanic topography is a common occurrence. In this section, we explore this scenario of unbalanced wave generation: the interaction of balanced flow with the lateral boundary such as a continental shelf.

Lateral boundary instability can be thought of as the interaction of the balanced flow with a simplified topography such as a lateral wall. To represent this scenario, we introduce meridional boundaries (experiment LBI) to the setup similar to that described in Section 3.3.2. Boundary conditions are that of no-normal flow and free slip. The associated boundary conditions for the QG filter are described in Appendix A.

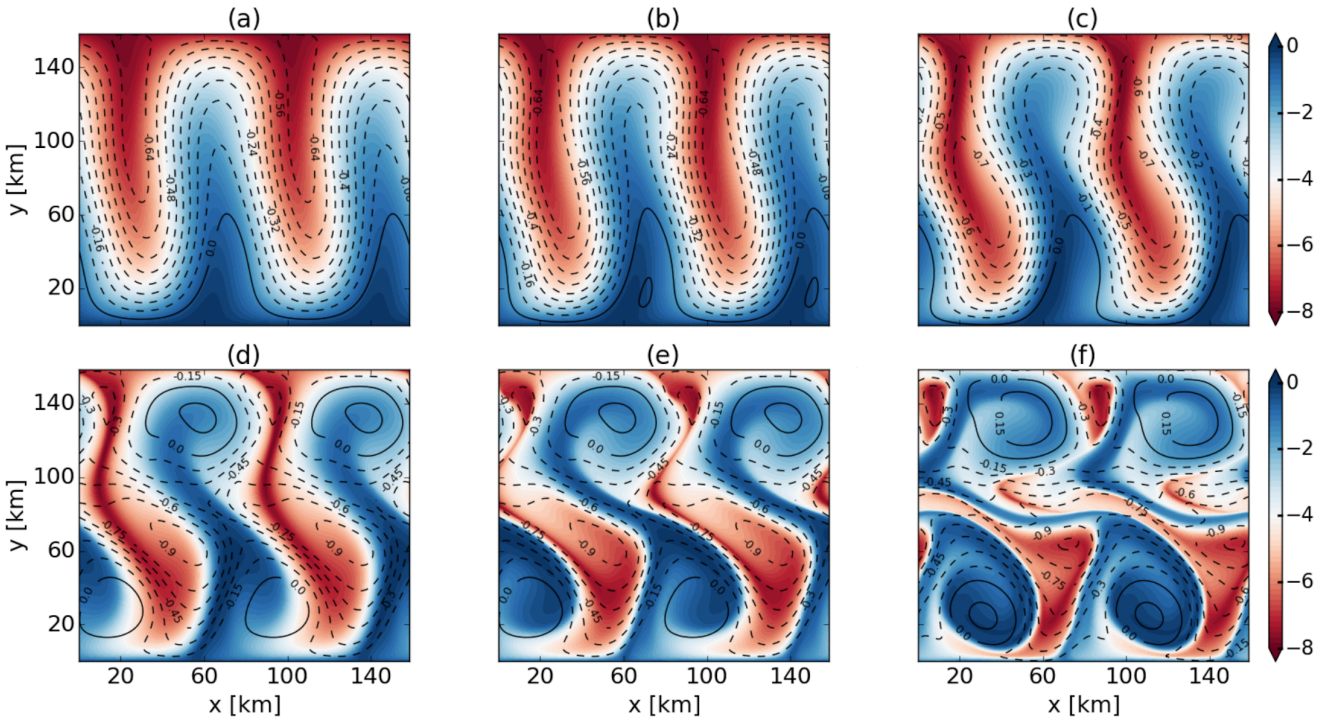


Figure 3.9: Snapshots of full buoyancy b in 10^{-3} ms^2 with pressure field in m^{2}s^2 with (black contours) in baroclinic unstable flow with meridional boundaries in experiment LBI. Snapshots are shown for a depth near the surface. The southern wall is located at $y = 0$ km. Snapshot (a) is the model state after 20 days from the initial model state in experiment LBI, the following snapshots (b-f) are two days apart each.

Balanced baroclinic vortices in the model are allowed to travel and reach the meridional walls, where these vortices interact with the lateral wall and generate unbalanced internal gravity waves. Such a balanced vortex-wall interaction for a case with Ri=400 (equivalent

to Ro $\ll 1$ regime) is shown in Fig. 3.9 near the surface. The southern wall is located at $y = 0$ km and the northern wall at $y = 160$ km. The figure shows the full buoyancy after 20 days of the initial model state of experiment LBI (Fig. 3.9a) followed by snapshots each two days apart (Fig. 3.9(b-f)). The pressure contours (black) and the full buoyancy (color shading) characterise the coherent vortex or eddy structures that hit the wall and generate unbalanced motions ((Fig. 3.10, discussed below). This eddy-wall interaction gets stronger and stronger with later times. The developing meanders followed by eddies hitting the walls is clearly seen in Fig. 3.9.

A similar example of unbalanced wave generation has been studied by Dewar and Hogg (2010) with barotropic and baroclinic vortices interacting with various types of topography. They found that the unbalanced waves generated as a result of the vortex-wall interaction have the characteristics of boundary trapped Kelvin waves. Further, they found that upon hitting the wall the balanced vortex generates both low mode and high mode Kelvin waves, where the latter being slower can be arrested by the balanced vortex. For a wave to be arrested by the balanced vortex, the wave should travel sufficiently slowly. A wave is said to be arrested by a balanced vortex if the wave does not move relative to the vortex. In other words, the balanced vortex and the wave structure move with similar speeds.

In our case, one can get an estimate of the eddy speed from the snapshots in (Fig. 3.9(a-e)). For instance, tracing the course of the eddy near the surface and located at about $x = 70$ km from Fig. 3.9a to Fig. 3.9e (where the eddy is located at $x = 100$ km), the eddy travels about 30 km in 8 days (each snapshot is 2 days apart) giving an estimated eddy speed $c_{eddy} = 0.043\ m/s$. We will get back to this aspect after a discussion of the unbalanced waves generated from the eddy-wall interaction. In what follows, we weigh the characteristics of the detected wave signals to verify if these wave signals are indeed Kelvin wave related structures and if they are arrested in a similar way as the Kelvin wave arrest discussed by Dewar and Hogg (2010).

Unbalanced wave generation after the eddy-wall interaction in our setup is shown in Fig. 3.10. The variation of unbalanced vertical $w - w_a$ at the southern wall of the channel at different times are shown in Fig. 3.10(a-e) and an example from the northern wall is shown in Fig. 3.10f. The snapshots in Fig. 3.10(a-e) are at at the same time steps as in Fig. 3.9(a-e). As shown in the figure, clear signatures of wave trains appear at the wall where the eddy collides with the wall (Fig. 3.9). The wave signatures increase in magnitude with time (Fig. 3.10(a-e)) and decay from the surface into the interior of the channel. Interestingly, the corresponding northern wall snapshot in Fig. 3.10f, appears as

a laterally inverted mirror image of the wave train structure in Fig. 3.10c at the southern wall, and in both the cases the waves decay towards the interior and away from the wall. The structure of the wave train resembles a wave with a very high vertical mode number, for instance, the wave train structure in the middle of the wall in Fig. 3.10c shows a mode number of about 10. A high mode number wave structure also appears in the wave train at the northern wall (Fig. 3.10f). Similar high mode wave structures appear at different instances (not shown) and could be related to Kelvin waves like in Dewar and Hogg (2010).

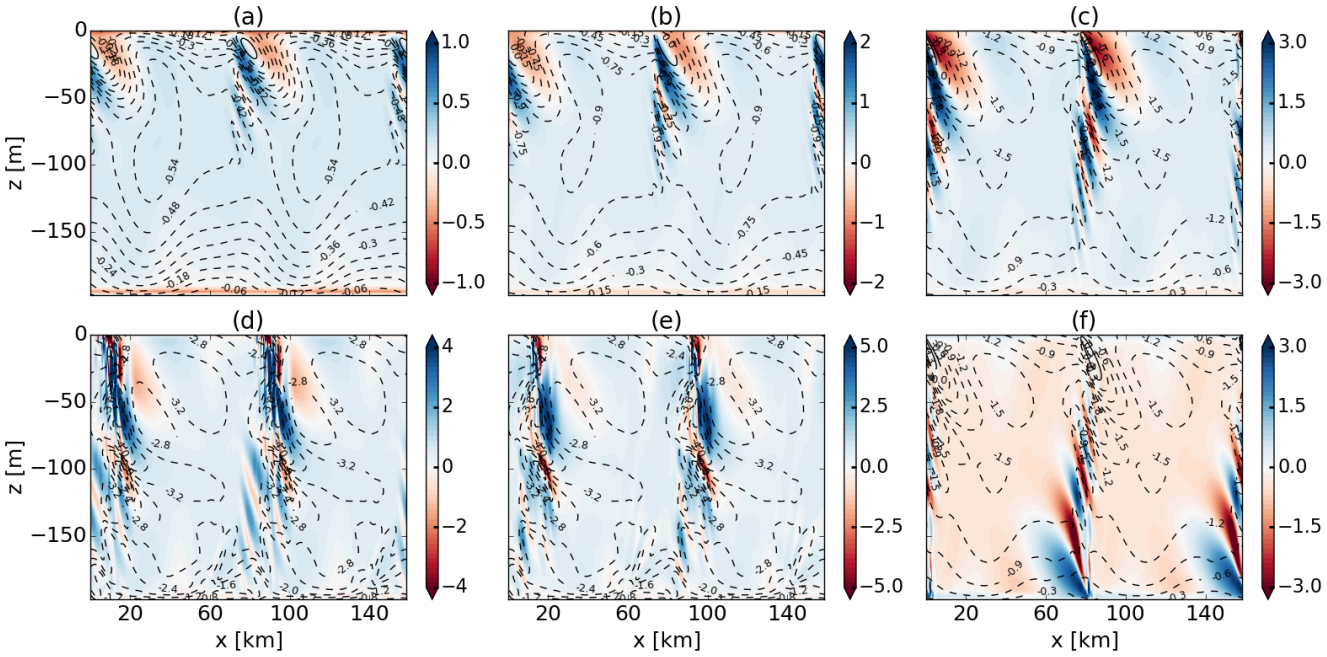


Figure 3.10: Snapshots of residual vertical velocity $w - w_a$ in 10^{-4} ms^{-1} diagnosed from the QG filter and full buoyancy b in 10^{-3} ms^2 (black contours) in baroclinic unstable flow with meridional boundaries in experiment LBI. (a-e) $w - w_a$ at the southern side wall of a channel. (f) $w - w_a$ at the northern wall corresponding to the snapshot in (c). The snapshots (a-e) correspond to the same time steps as in Fig. 3.9(a-e).

From Fig. 3.10(a-e) the speed of travel of the wave trains can be worked out. Beginning from Fig. 3.10(a) and tracking the wave train up to the snapshot Fig. 3.10(e), the wave train at the middle of the wall appears to travel about 30 km in 8 days (each snapshot is 2 days apart). With that the wave train has a speed of $c_{wave} = 0.043$ ms^{-1}. The match between the wave speed and the eddy speed c_{eddy} implies that the eddy and the wave train travel the same distance in about the same time and that the unbalanced wave is arrested by the balanced eddy flow. That the wave appears to move as slow as the balanced eddy

guides towards a high mode number wave structure, in accordance with the wave structures detected in the model (Fig. 3.10).

For the given parameters in our model ($N = f/\alpha$, $H = 200\ m$, $f = 10^{-4}$ ms^{-1}, $\alpha = 0.02$), the Kelvin wave phase speed $c_n = NH = 0.63\ m/s$, with mode number $n = 1$ corresponding to the first baroclinic Rossby radius ($L_r = NH/(n\pi f)$). For a mode number of 10 the phase speed reduces to $c_{10} = 0.063$ ms^{-1}, for an even higher value of n the phase speed reduces further ($c_{11} = 0.057\ m/s$, $c_{12} = 0.053$ ms^{-1}). Although these values are higher than the speed c_{wave} of the wave train as seen from Fig. 3.10, they appear to match reasonably well; for a higher mode number the phase speed gets even more closer to c_{wave}. The similar speeds convey that these wave structures are consistent with Kelvin wave characteristics. The lower value of c_{wave} suggests that the wave structure detected in our model is a Kelvin wave of even higher mode number. However, this could also be an effect of the Doppler shift due to the mean flow, which we have not taken into account.

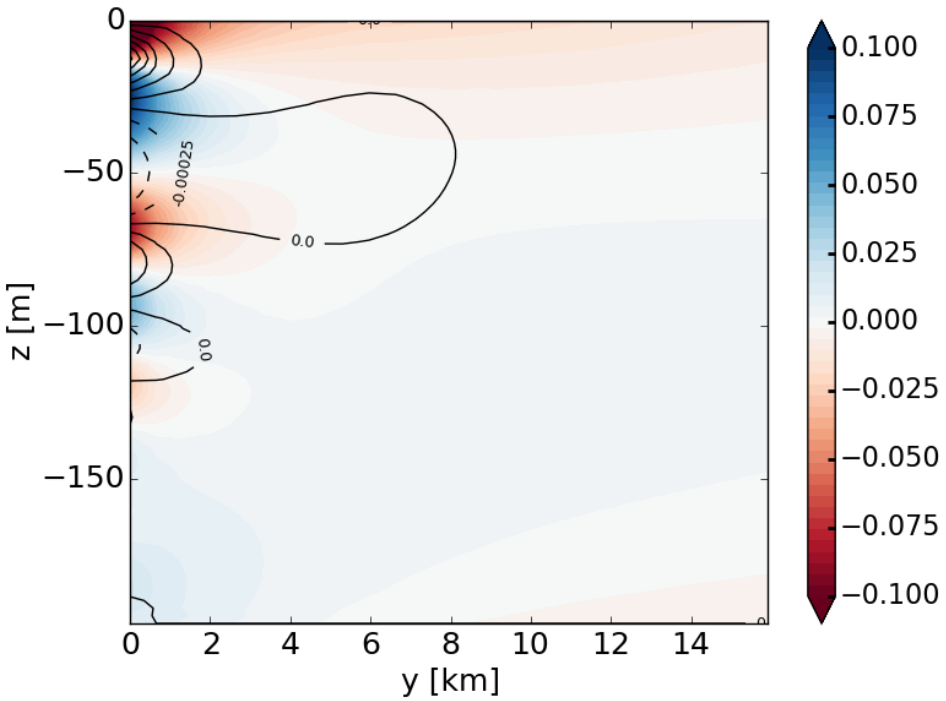


Figure 3.11: Snapshots of $u - u_g - u_a$ in ms^{-1} with contours (black) of $b - b_g$ in baroclinic unstable flow at the southern side wall of the channel in experiment LBI.

Now we look at the residual zonal velocity component $u - u_g - u_a$. A depth section along the location of the wave train (Fig. 3.10(c)) in x direction is shown in (Fig. 3.11) for the residual zonal velocity. The figure shows a region close to the southern wall at $y = 0$

km and the decay of the wave amplitude away from the wall, similar to a Kelvin wave, is clearly visible. The overlaid black contours of the unbalanced buoyancy component $b - b_g$ show a similar pattern. Also, most of the wave signal is concentrated close to the wall, and thereafter it shows a drastic decay away from the wall. The e-folding length scale the decay of the wave signals (less than a km) roughly matches the length scale of the 10^{th} baroclinic deformation radius which is about 636.94 m. The corresponding meridional component $(v - v_g - v_a)$ (not shown) is nought, also consistent with the Kelvin wave property for which the meridional velocity vanishes at the wall due to the no-normal flow condition.

These consistencies shed light towards the point that the unbalanced wave structures generated in response to the eddy-wall interaction are slowly propagating Kelvin wave structures, and like in Dewar and Hogg (2010), these Kelvin waves are arrested by the balanced eddy flow. In addition, the unbalanced wave emission can be compared against the two scenarios discussed in the previous sections. The magnitude of $w - w_a$ for the eddy-wall interaction appears to be of comparable magnitude with spontaneous emission, but less than that for emission due to convective instability at fronts.

3.4 Summary and conclusions

We introduce a new balance procedure called the Quasi-Geostrophic (QG) filter to diagnose internal gravity waves in different dynamical scenarios. The QG filter is based on quasi-geostrophic equations and is a potential vorticity inversion method with the additional capability to diagnose the horizontal ageostrophic components. We have implemented the QG filter in a fully discrete version consistent on the discrete grid level. Quasi-geostrophic horizontal velocities are obtained from the inversion of potential vorticity equation while the ageostrophic horizontal and vertical velocities are obtained from the diagnostic relations. When a model based on the primitive equations is diagnosed by the QG filter, the residual obtained by subtracting the geostrophic and ageostrophic components from the full component can be interpreted as the gravity wave component. Alternatively, the QG filter can be used to initialize a model with QG balanced conditions to eliminate the unbalanced gravity waves.

The QG filter is used to diagnose gravity waves in different scenarios and the key conclusions are:

1. The new diagnostic, QG filter, is a competent diagnostic tool to diagnose internal gravity waves from a balanced flow field. The QG filter can be applied as a diagnostic

as well as a balance procedure. When use to balance initial conditions, the QG filter eliminates gravity waves similar to the normal mode initialization procedure by Machenhauer (1977). The Machenhauer method in its discrete version is, however, superior to the QG filter in balancing the initial conditions.

2. Spontaneous emission of gravity waves from baroclinic eddies shows that there is a significant amount of gravity wave activity and gravity waves appear to emanate from the eddies. Although gravity wave emission is weak for a Ro $\ll 1$ regime, the new diagnostic method is capable of detecting the gravity wave signals.

3. Convective instability during frontogenesis arising from an unstable density stratification and resulting into overturning of isopycnals excites gravity waves at fronts. The waves excited at the frontal instability eventually propagate away from the front and dissipate.

4. Lateral boundary instability results in gravity wave emission by eddy-wall interaction. In this setting, a baroclinic eddy upon collision with a lateral wall emits unbalanced gravity waves. The emitted gravity waves bear the characteristics of a high mode number Kelvin wave structures which appear arrested by the balanced eddy.

The additional capability of the QG filter provides us with the horizontal velocities associated with the unbalanced motions. This feature has backed a deeper insight into some of the internal wave generation mechanisms addressed in this paper. Among these mechanisms of internal gravity wave generation, convective instability, though a local event, appears with a larger magnitude of wave activity than spontaneous emission and lateral boundary instability. The wave generation in the latter two scenarios appears comparable in magnitude although the processes are entirely different. Though different, in some situations these processes can co-exist and can contribute together to internal wave generation. Further, the emitted unbalanced wave motions can interact with other unbalanced processes and generate further unbalanced wave motions through resonant interactions. In the present work, however, we have focused on the wave generation from the balanced flow. In our next steps we intend to further expand the capabilities of QG filter to explore more complex interactions.

The results have strong implications not only for gravity wave dynamics but also for the energetics of balanced motions as well as the ocean's energy budget. Gravity wave emission from balanced flows provides an important source of the internal wave field in the

ocean, and breaking of internal gravity wave is a primary contributor to diapycnal mixing in the ocean. In this way, the internal gravity wave pathway links the balanced flow to the scale of molecular mixing, acting as an energy sink of the balanced flow.

3.5 Appendix A: QG filter: Lateral boundary conditions

For a closed domain without islands the no-normal-flow boundary condition is $\psi = \mu_0(t,z)$ along the boundary. The boundary value μ_0 is obtained by insuring zero net vertical mass flux $\int_A w dA = 0$, which implies

$$\partial_t f \; \partial_z \int_A \psi dA = \int_A \dot{b} dA \quad \rightarrow \quad f \; \partial_z \int_A \psi dA = \int_A b dA \tag{3.9}$$

at all depths, where A is the surface area of the domain. This constrain also needs to be satisfied for a double periodic domain.To satisfy that condition we decompose ψ as

$$\psi(t, \boldsymbol{x}, z) = \psi_0(t, \boldsymbol{x}, z) + \mu_0(t, z) \tag{3.10}$$

with $f\partial_z \psi_0 = b$ and $\partial_z \mu_0 = 0$ at $z = 0, -h$ to satisfy the upper and lower boundaries, and with $\psi_0 = 0$ along the lateral boundaries (if there are any). The condition yields

$$\partial_t \; \partial_z \mu_0 = \frac{1}{A} \int_A \dot{b}/f dA - \partial_t \; \partial_z \frac{1}{A} \int_A \psi_0 dA \quad \rightarrow \quad \partial_z \mu_0 = \frac{1}{A} \int_A b/f dA - \; \partial_z \frac{1}{A} \int_A \psi_0 dA \tag{3.11}$$

which yields as differential equation for ψ_0

$$\boldsymbol{\nabla}^2 \psi_0 + \frac{f^2}{N^2} \partial_{zz} \; \left(\psi_0 - \frac{1}{A} \int_A \psi_0 dA \right) = q - \frac{f^2}{N^2} \partial_z \frac{1}{A} \int_A b/f dA \tag{3.12}$$

For a zonally reentrant channel configuration no normal flow boundary conditions need to be established by the additional constraint of global momentum conservation $\int \dot{u} dA$. We decompose the streamfunction as

$$\psi(t, \boldsymbol{x}, z) = \psi_0(t, \boldsymbol{x}, z) + \mu_0(t, z) + \mu_1(t, z)\psi_1(\boldsymbol{x}) \quad , \quad \boldsymbol{\nabla}^2 \psi_1 = 0 \tag{3.13}$$

Lateral boundary conditions are $\psi_0 = 0$ at all boundaries and $\psi_1 = 0$ at the southern boundary at $y = 0$ but $\psi_1 = 1$ at the northern boundary at $y = L$, and with $f\partial_z \psi_0 = b$

and $\partial_z \mu_0 = 0$, $\partial_z \mu_1 = 0$ at $z = 0, -h$ to satisfy the upper and lower boundaries. The momentum conservation condition becomes

$$\int \dot{u} dA = \int \partial_t u dA = -\int \partial_t \partial_y \psi dA \quad \rightarrow \quad \int u dA = -\int \partial_y \psi dA \quad \rightarrow \quad \mu_1 = -\frac{1}{B} \int u dA \tag{3.14}$$

where B is the zonal extent of the channel. The zero mass flux condition becomes

$$\partial_z \mu_0 = \frac{1}{A} \int_A b/f dA - \frac{1}{A} \int_A \partial_z \psi_0 dA - \partial_z \mu_1 \frac{1}{A} \int_A \psi_1 dA \tag{3.15}$$

and the differential equation for ψ_0 becomes

$$\boldsymbol{\nabla}^2 \psi_0 + \frac{f^2}{N^2} \partial_{zz} \left(\psi_0 - \frac{1}{A} \int_A \psi_0 dA \right) = q - \frac{f^2}{N^2} \frac{1}{A} \int_A \partial_z b/f dA - \frac{f^2}{N^2} \partial_{zz} \mu_1 \left(\psi_1 - \frac{1}{A} \int_A \psi_1 dA \right) \tag{3.16}$$

The northern boundary acts like an islands. If there are more than one island line integrals along the boundaries of the island need to be solved in addition.

$$\oint_C \boldsymbol{\nabla}^\perp \partial_t \psi \cdot d\boldsymbol{\ell} = \oint_C \dot{\boldsymbol{u}} \cdot d\boldsymbol{\ell} \quad \rightarrow \quad \oint_C \boldsymbol{\nabla}^\perp \psi \cdot d\boldsymbol{\ell} = \oint_C \boldsymbol{u} \cdot d\boldsymbol{\ell} \tag{3.17}$$

where $d\boldsymbol{\ell}$ is the increment parallel to the coastline C.

3.6 Appendix B: QG filter: Discrete equations

The discrete version for $f = const$ on an equidistant C-grid is

$$\begin{aligned}
\partial_t \, u_{i,j,k} - f \overline{\overline{v_{i,j,k}}^{i+}}^{j-} &= -\delta_x^+ p_{i,j,k} + \dot{u}_{i,j,k} && (3.18)\\
\partial_t \, v_{i,j,k} + f \overline{\overline{u_{i,j,k}}^{i-}}^{j+} &= -\delta_y^+ p_{i,j,k} + \dot{v}_{i,j,k} && (3.19)\\
\delta_z^+ p_{i,j,k} &= \overline{b_{i,j,k}}^{k+} && (3.20)\\
\partial_t b_{i,j,k} + N^2 \overline{w_{i,j,k}}^{k-} &= \dot{b}_{i,j,k} && (3.21)\\
\delta_x^- u_{i,j,k} + \delta_y^- v_{i,j,k} + \delta_z^- w_{i,j,k} &= 0 && (3.22)
\end{aligned}$$

with the finite differencing operators

$$\delta_x^+ p_{i,j,k} = (p_{i+1,j,k} - p_{i,j,k})/\Delta_x \ , \ \delta_y^+ p_{i,j,k} = (p_{i,j+1,k} - p_{i,j,k})/\Delta_y \ , \ \delta_z^+ p_{i,j,k} = (p_{i,j,k+1} - p_{i,j,k})/\Delta_z \tag{3.23}$$

$$\delta_x^- p_{i,j,k} = (p_{i,j,k} - p_{i-1,j,k})/\Delta_x \ , \ \delta_y^- p_{i,j,k} = (p_{i,j,k} - p_{i,j-1,k})/\Delta_y \ , \ \delta_z^- p_{i,j,k} = (p_{i,j,k} - p_{i,j,k-1})/\Delta_z \tag{3.24}$$

and the averaging operators

$$\overline{v_{i,j,k}}^{i+} = 0.5(v_{i,j,k} + v_{i+1,j,k}) \ , \ \overline{v_{i,j,k}}^{i-} = 0.5(v_{i,j,k} + v_{i-1,j,k}) \tag{3.25}$$

$$\overline{v_{i,j,k}}^{j+} = 0.5(v_{i,j,k} + v_{i,j+1,k}) \ , \ \overline{v_{i,j,k}}^{j-} = 0.5(v_{i,j,k} + v_{i,j-1,k}) \tag{3.26}$$

$$\overline{v_{i,j,k}}^{k+} = 0.5(v_{i,j,k} + v_{i,j,k+1}) \ , \ \overline{v_{i,j,k}}^{k-} = 0.5(v_{i,j,k} + v_{i,j,k-1}) \tag{3.27}$$

Since all differencing and averaging operators commute the vorticity and divergence equations become

$$\partial_t(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)\psi_{i,j,k} - f\delta_z^-\overline{\overline{(w_{i,j,k})}^{i+}}^{j+} = \delta_x^+ \ \dot{v}_{i,j,k} - \delta_y^+\dot{u}_{i,j,k} \tag{3.28}$$

$$\partial_t(\delta_x^+\delta_x^- \ + \delta_y^+\delta_y^-)\phi_{i,j,k} - f(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = -(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)p_{i,j,k} \tag{3.29}$$

$$+\delta_x^- \ \dot{u}_{i,j,k} + \delta_y^-\dot{v}_{i,j,k} \tag{3.30}$$

introducing a streamfunction ψ and potential ϕ with $u_{i,j,k} = \delta_x^+\phi_{i,j,k} - \delta_y^-\psi$ and $v_{i,j,k} = \delta_y^+\phi_{i,j,k} + \delta_x^-\psi$ and where the vorticity equation was combined with $(\delta_x^+\delta_x^- + \delta_y^+\delta_y^-)\phi_{i,j,k} + \delta_z^- w_{i,j,k} = 0$.

3.6.1 Quasi-geostrophic approximation

Implementing the quasi-geostrophic assumption as in the continuous case yields

$$f(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = (\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)p^g_{i,j,k} \ \rightarrow \ \overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = p^g_{i,j,k}/f \tag{3.31}$$

and

$$(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)p^a_{i,j,k} = \delta_x^- \ \dot{u}_{i,j,k} + \delta_y^-\dot{v}_{i,j,k} \tag{3.32}$$

Buoyancy is also split into a dominant geostrophic part b^g and a ageostrophic part for which both the hydrostatic relation holds, and thus

$$\delta_z^+ p^g_{i,j,k} = \overline{b^g_{i,j,k}}^{k+} \ \rightarrow \ f\delta_z^+\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = \overline{b^g_{i,j,k}}^{k+} \tag{3.33}$$

The leading order buoyancy equation becomes

$$\partial_t b^g_{i,j,k} + N^2\overline{w_{i,j,k}}^{k-} = \dot{b}^g_{i,j,k} \ \rightarrow \ \frac{f^2}{N^2}\partial_t\delta_z^-\delta_z^+\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = -f\delta_z^-\overline{\overline{w_{i,j,k}}^{k-}}^{k+} + \frac{f}{N^2}\delta_z^-\overline{\dot{b}^g_{i,j,k}}^{k+} \tag{3.34}$$

The potential vorticity equation becomes

$$\partial_t \left((\delta_x^- \delta_x^+ + \delta_y^- \delta_y^+) \overline{\overline{\psi_{i,j,k}}^{k+}}^{k-} + \frac{f^2}{N^2} \delta_z^- \delta_z^+ \overline{\overline{\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-}}^{i+}}^{j+} \right) = \overline{\overline{(\delta_x^+ \dot{v}_{i,j,k} - \delta_y^+ \dot{u}_{i,j,k})}^{k+}}^{k-} + \frac{f}{N^2} \delta_z^+ \overline{\overline{\overline{\overline{\dot{b}^g_{i,j,k}}^{i+}}^{j+}}^{k-}} \quad (3.35)$$

thus the quasi-geostrophic potential vorticity on a C grid is given by

$$(\delta_x^- \delta_x^+ + \delta_y^- \delta_y^+) \overline{\overline{\psi_{i,j,k}}^{k+}}^{k-} + \frac{f^2}{N^2} \delta_z^- \delta_z^+ \overline{\overline{\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-}}^{i+}}^{j+} = q_{ijk} = \overline{\overline{(\delta_x^+ v_{i,j,k} - \delta_y^+ u_{i,j,k})}^{k+}}^{k-} + \frac{f}{N^2} \delta_z^+ \overline{\overline{\overline{\overline{b_{i,j,k}}^{i+}}^{j+}}^{k-}} \quad (3.36)$$

Eq. (**??**) corresponds to Eq. (3.4) in the continuous case.

Diagnostic relations for ageostrophic velocities

The vertical velocity can then be diagnosed from the buoyancy equation as

$$\overline{\overline{w_{i,j,k}}^{k-}}^{k+} = \overline{\dot{b}^g_{i,j,k}}^{k+} / N^2 - \frac{f}{N^2} \partial_t \delta_z^+ \overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} \quad (3.37)$$

The divergent velocities $u_a = \partial_x \phi$ and $v_a = \partial_y \phi$ can be obtained from the momentum equation as

$$f \overline{\overline{\overline{\overline{u^a_{i,j,k}}^{i-}}^{j+}}^{i+}}^{j-} = \overline{\overline{(\dot{v}_{i,j,k} - \partial_y p^a_{i,j,k})}^{i+}}^{j-} - \partial_t \overline{\overline{\delta_x^- \psi_{i,j,k}}^{i+}}^{j-} \quad (3.38)$$

$$f \overline{\overline{\overline{\overline{v^a_{i,j,k}}^{i+}}^{j-}}^{i-}}^{j+} = -\overline{\overline{(\dot{u}_{i,j,k} - \partial_x p^a_{i,j,k})}^{i-}}^{j+} - \partial_t \overline{\overline{\delta_y^- \psi_{i,j,k}}^{i-}}^{j+} \quad (3.39)$$

3.6.2 Vertical boundary condition

From the diagnostic relation it follows that

$$\delta_z^+ p_{i,j,k} = \overline{b_{i,j,k}}^{k+} \ , \ f \overline{\overline{\psi_{i,j,k}}^{i-}}^{j-} = p_{i,j,k} \quad \rightarrow \quad f \delta_z^+ \overline{\overline{\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-}}^{i+}}^{j+} = \overline{\overline{b_{i,j,k}}^{k+}}^{i+}{}^{j+} \quad (3.40)$$

Assume that $\delta_z^+ b_k = 0$ at the surface and bottom such that

$$\delta_z^+ \overline{\overline{\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-}}^{i+}}^{j+} = \overline{\overline{b_{i,j,k}}^{i+}}^{j+} / f \quad at \quad k = N_z \quad , \quad \delta_z^- \overline{\overline{\overline{\overline{\psi_{i,j,k}}^{i-}}^{j-}}^{i+}}^{j+} = \overline{\overline{b_{i,j,k}}^{i+}}^{j+} / f \quad at \quad k = 1 \quad (3.41)$$

or the corresponding relation for $\partial_t f \ \partial_z \psi = \dot{b}$.

Lateral boundary condition for closed domain

The discrete version of the zero mass flux condition for a closed domain is

$$\delta_z^+(\mu_0)_k = \frac{1}{A}\int_A \overline{\overline{\overline{b_{i,j,k}/f}^{k+}}^{i+}}^{j+} dA - \delta_z^+ \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+} dA \tag{3.42}$$

Note that the averaging in i and j is not necessary in the surface integrals for a double periodic domain. The differential equation for ψ^0 for the closed domain is then given by

$$(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)\overline{(\psi_0)_{i,j,k}}^{k+}{}^{k-} + \frac{f^2}{N^2}\delta_z^-\delta_z^+\left(\overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+} - \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+}\right) \tag{3.43}$$

$$= \overline{(\delta_x^+\ v_{i,j,k} - \delta_y^+ u_{i,j,k})}^{k+}{}^{k-} + \frac{f}{N^2}\delta_z^+\overline{\left(\overline{b_{i,j,k}}^{i+}{}^{j+} - \frac{1}{A}\int_A \overline{b_{i,j,k}}^{i+}{}^{j+} dA\right)}^{k-} \tag{3.44}$$

3.6.3 Lateral boundary condition for channel configuration

The discrete version of the mass flux constrain for a channel configuration is

$$\delta_z^+(\mu_0)_k = \frac{1}{A}\int_A \overline{\overline{\overline{b_{i,j,k}/f}^{k+}}^{i+}}^{j+} dA - \delta_z^+ \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+} dA - \delta_z^+(\mu_1)_k \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_1)_{ijk}}^{i-}}^{j-}}^{i+}{}^{j+} dA \tag{3.45}$$

and the equation for ψ_0 for a channel configuration becomes

$$(\delta_x^-\delta_x^+ + \delta_y^-\delta_y^+)\overline{(\psi_0)_{i,j,k}}^{k+}{}^{k-} + \frac{f^2}{N^2}\delta_z^-\delta_z^+\left(\overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+} - \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_0)_{i,j,k}}^{i-}}^{j-}}^{i+}{}^{j+}\right) \tag{3.46}$$

$$= \overline{(\delta_x^+\ v_{i,j,k} - \delta_y^+ u_{i,j,k})}^{k+}{}^{k-} + \frac{f}{N^2}\delta_z^+\overline{\left(\overline{b_{i,j,k}}^{i+}{}^{j+} - \frac{1}{A}\int_A \overline{b_{i,j,k}}^{i+}{}^{j+} dA\right)}^{k-} \tag{3.47}$$

$$-\frac{f^2}{N^2}\left(\delta_z^-\delta_z^+\mu_k^1\right)\left(\overline{\overline{\overline{(\psi_1)_{ijk}}^{i-}}^{j-}}^{i+}{}^{j+} - \frac{1}{A}\int_A \overline{\overline{\overline{(\psi_1)_{ijk}}^{i-}}^{j-}}^{i+}{}^{j+} dA\right) \tag{3.48}$$

3.7 Appendix C: Machenhauer initialization: Discrete equations

Discrete system for freely propagating waves

The Fourier transform of the discrete linear system is done using the definitions

$$i\hat{k}_x^+(k_x) = \frac{e^{ik_x\Delta x}-1}{\Delta x} \stackrel{\Delta x\to 0}{=} ik_x \quad , \quad i\hat{k}_x^-(k_x) = \frac{1-e^{-ik_x\Delta x}}{\Delta x} \stackrel{\Delta x\to 0}{=} ik_x \tag{3.49}$$

$$\hat{1}_x^+(k_x) = \frac{e^{ik_x\Delta x}+1}{2} \stackrel{\Delta x\to 0}{=} 1 \quad , \quad \hat{1}_x^-(k_x) = \frac{e^{-ik_x\Delta x}+1}{2} \stackrel{\Delta x\to 0}{=} 1 \tag{3.50}$$

and similar for $\hat{k}_y^+$, $\hat{k}_y^-$, etc, with $(\hat{1}_x^+)^* = \hat{1}_x^-$, $(\hat{k}_x^+)^* = \hat{k}_x^-$, $\hat{k}_x^+(-k_x) = -\hat{k}_x^-(k_x)$ and $\hat{1}_x^+(-k_x) = \hat{1}_x^-(k_x)$. Then

$$u_{i,j,k} = \hat{u}e^{i(k_x x+k_y y+k_z z)} \ , \ v_{i,j,k} = \hat{v}e^{i(k_x x+k_y y+k_z z)} \ , \ b_{i,j,k} = \hat{b}e^{i(k_x x+k_y y+k_z z)} \tag{3.51}$$

yields ignoring the non-linear terms in the discrete equations

$$\begin{aligned} \partial_t\ \hat{u} - f1_x^+1_y^-v + i\hat{k}_x^+p = 0 \quad , \quad \partial_t\ \hat{v} + f1_x^-1_y^+u + i\hat{k}_y^+p = 0 \\ \partial_t\hat{b} + N^2 1_z^- w = 0 \quad \rightarrow \quad \partial_t\hat{p} + ic^2(k_z)1_z^+1_z^-(\hat{k}_x^-u + \hat{k}_y^-v) = 0 \end{aligned} \tag{3.52}$$

with $i\hat{k}_z^+p = 1_z^+b$ and $\hat{k}_x^-u + \hat{k}_y^-v + \hat{k}_z^-w = 0$ and $c^2(k_z) = N^2/(\hat{k}_z^-\hat{k}_z^+)$. This can be written as

$$\partial_t\boldsymbol{z} = i\boldsymbol{A}\cdot\boldsymbol{z} \quad , \quad \boldsymbol{A} = \begin{pmatrix} 0 & -i1_x^+1_y^-f & -\hat{k}_x^+ \\ i1_y^+1_x^-f & 0 & -\hat{k}_y^+ \\ -c^2(k_z)\hat{k}_x^- & -c^2(k_z)\hat{k}_y^- & 0 \end{pmatrix} \quad , \quad \boldsymbol{z} = \begin{pmatrix} \hat{u} \\ \hat{v} \\ \hat{p} \end{pmatrix} \tag{3.53}$$

Note that c^2 is real as the eigenvalues of the matrix. The eigenvalues of $\boldsymbol{A}$ are

$$\omega = 0 \ , \ \omega^2 = c^2(k_y^-k_y^+ + \hat{k}_x^-\hat{k}_x^+) + 1_x^+1_y^-1_x^-1_y^+f^2 \tag{3.54}$$

and the eigenvectors are

$$\boldsymbol{Q}^s = \begin{pmatrix} \frac{if1_x^+1_y^-k_y^+ - k_x^+s|\omega|}{s^2\omega^2 - 1_x^+1_y^-1_y^+1_x^-f^2} \\ \frac{-if1_y^+1_x^-k_x^+ - k_y^+s|\omega|}{s^2\omega^2 - 1_x^+1_y^-1_y^+1_x^-f^2} \\ 1 \end{pmatrix} \quad , \quad \boldsymbol{P}^s = p^s\left(\frac{-if1_y^+1_x^-k_y^- - k_x^-s|\omega|}{s^2\omega^2 - 1_y^+1_x^-1_x^+1_y^-f^2} \ , \ \frac{if1_y^-1_x^+k_x^- - k_y^-s|\omega|}{s^2\omega^2 - 1_y^+1_x^-1_x^+1_y^-f^2} \ , \ c^{-2}\right)$$

with the normalisation

$$p_s = \frac{|s^2\omega^2 - 1_x^+ 1_y^- 1_y^+ 1_x^- f^2|}{(1+s^2)(k_y^- k_y^+ + \hat{k}_x^- \hat{k}_x^+ + 1_x^+ 1_y^- 1_x^- 1_y^+ L^{-2})} \tag{3.55}$$

Note that this is all identical to the analytical eigenvectors for $\Delta x, \Delta y, \Delta z \to 0$.

Discrete system for vertical modes

The discrete vertical relations with respect to the vertical coordinate are given by

$$\delta_z^+ p = \overline{b}^{k+} \quad \to \quad -p_n \frac{2\sin m\Delta z}{\Delta z(\cos m\Delta z + 1)} = b_n \tag{3.56}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} + \delta_z^- w = 0 \quad \to \quad \boldsymbol{\nabla}\cdot\boldsymbol{u}_n + w_n \frac{2\sin(m\Delta z/2)}{\Delta z} = 0 \tag{3.57}$$

$$\partial_t b + \overline{w}^{k-1} N^2 = 0 \quad \to \quad \partial_t b_n + w_n N^2 \cos(m\Delta z/2) = 0 \tag{3.58}$$

$$\partial_t p_n + c^2 \boldsymbol{\nabla}\cdot\boldsymbol{u}_n = 0 \quad , \quad c^2 = N^2 \frac{\Delta z^2}{4} \frac{(\cos m\Delta z + 1)}{\sin m\Delta z \tan(m\Delta z/2)} \tag{3.59}$$

with $m = n\pi/H$ and the operators δ_z^+, etc as defined above. The horizontal part stays the same as for the free waves.

Chapter 4

Conclusions

The results obtained from the research undertaken in the thesis are compiled in this chapter. The main results are discussed followed by prospects that look at the results from a different light. Perspectives for the present and forthcoming research wind up the thesis and raise new questions to embark upon a new quest.

4.1 Disentangling gravity waves from balanced flow

The first and foremost aim of this thesis has been to disentangle internal gravity waves (IGWs) from balanced flows. IGWs, which are unbalanced high-frequency motions, are coupled with low-frequency balanced motions in a complex manner which makes the separation of these motions equally complex. The separation becomes particularly complicated for regimes with Richardson number $\mathrm{Ri} = \mathscr{O}(1)$ (or equivalently Rossby number $\mathrm{Ro} < 1$) because of the strong coupling between balanced and unbalanced motions. To diagnose IGWs from balanced flows two different methods have been implemented: modal decomposition discussed in Chapter 2 and quasi-geostrophic filter (QG filter) discussed in Chapter 3.

4.1.1 Performance of modal decomposition based on Machenhauer (1977)

The linear modal decomposition gives the linear balanced and unbalanced modes of the system that appear well separated as seen in physical space (Fig 2.6) and further in the frequency-wavenumber spectra (Fig 2.7) of the kinetic energy (KE) associated with the

decomposed modes for a Ri $\gg$ 1. For the linear balanced mode, the associated KE lies predominantly outside the gravity wave branch and towards smaller frequencies. On the other hand, the linear unbalanced mode show energy inside the gravity wave branch but also some energy outside the branch towards smaller frequencies. This discrepancy is attributed to the fact that the inherent non-linearity of the flow is not included in the linear decomposition.

To account for the non-linearity, a non-linear decomposition based on the non-linear initialization technique proposed by Machenhauer (1977) is applied, which gives the non-linearly decomposed modes. A distinct separation between non-linear balanced and non-linear unbalanced modes is evident from the frequency-wavenumber spectra of KE of the decomposed modes (Fig 2.8). For Ri $\gg$ 1 (Ro $\ll$ 1), which exhibits large timescale separation between balanced and unbalanced modes, the separation of modes is less difficult than for Ri = $\mathscr{O}(1)$ (Ro $<$ 1) which exhibits strong coupling between the balanced and unbalanced modes. Accordingly, for Ri = $\mathscr{O}(1)$ (Ri = 3 in Fig 2.8) non-linear unbalanced mode shows a significant KE within the gravity wave branch, but KE lies within the gravity wave branch also for the non-linear balanced mode. This is because the strongly coupled co-existing balanced and unbalanced motions for a Ri = $\mathscr{O}(1)$ i.e. an ageostrophic regime, make the decomposition much more difficult. For higher Ri the decomposition clearly separates the balanced and unbalanced motions. Therefore, the non-linear modal decomposition comes forth as an efficient tool to decompose the flow field into balanced and unbalanced modes and is superior to the linear modal decomposition but becomes less efficient for small Ri.

4.1.2 Performance of the QG filter

The new diagnostic, QG filter, emerges as a competent diagnostic tool to extract IGW signal from a balanced flow field. The additional capability of the QG filter makes available the horizontal velocity components of the unbalanced motions, which are utilized to extract further information of the diagnosed IGWs that aid in a robust characterization of the signals. The frequency-wavenumber spectra of KE calculated from the unbalanced horizontal velocities (Fig. 3.3) gives a first indication of the effective separation of the balanced and unbalanced components. For both small (Ri = 20) and high (Ri = 1000) regimes, the KE in the unbalanced components follows the curve of IGW dispersion relation, whereas most of the balanced KE lies outside the gravity wave branch. Further,

the QG filter efficiently diagnoses IGWs generated by different mechanisms: spontaneous emission, convective instability, and lateral boundary instability.

Additionally, the QG filter can be applied as a balance procedure to balance initial conditions in a model with the aim to eliminate IGWs. To test this, the initial balanced conditions provided by the QG filter is compared on the one hand against a model without balanced initial conditions and on the other hand against the reformulated discrete version of the non-linear initialization technique by Machenhauer (1977) which in this case is implemented with three iterations. The QG filter (Fig. 3.2 d,e,f) performs remarkably well in suppressing the unbalanced IGW motions, which are up to two orders of magnitude smaller than the IGW signals present in the model without balanced initial conditions (Fig. 3.2 a,b,c). The discrete version of Machenhauer (1977) method (Fig. 3.2 g,h,i), however, performs even better resulting in further minimization of IGW signals which are an order of magnitude smaller than the ones obtained from the QG filter.

Prospects

The discrete consistent version of Machenhauer (1977) method (Chapter 3) appears to be better suited than its analytical version (Chapter 2), and so the results of IGW detection presented in Chapter 2 might further improve using the discrete version instead of the analytical one. Moreover, the method is applied in Chapter 2 only up to the first iteration but up to three iterations in Chapter 3. Higher iterations tend to extract more and more IGWs, hence providing a robust interpretation of the IGW signals present in the flow.

At the same time, the non-linear initialization procedure of Machenhauer (1977) is not entirely free from issues. By setting the initial time derivatives of the IGW modes to zero, and hence imposing a non-linear balance condition devoid of IGW motions, Machenhauer's method implies that the obtained balanced state is a strict balanced manifold, which does not exist as discussed earlier in the thesis. The slow evolving balanced manifold will generate small but non-zero amount of fast motions. The non-linear balance condition is therefore not exactly consistent with such a manifold (cf. Leith, 1980; Theiss and Mohebalhojeh, 2009). This correction was taken into account by Baer (1977) by including a correction term beginning from the second iteration. This correction, however, does not matter for the first iteration (Theiss and Mohebalhojeh, 2009) which is the case in Chapter 2 of this thesis. In Chapter 3, however, the reformulated Machenhauer's version with three iterations might further improve with this correction term taken into account, and is one of the points for exploration hereafter. With that said, Machenhauer's method has its own

advantage of being economical w.r.t. computational time (e.g. Temperton and Williamson, 1981).

Another aspect related to the modal decomposition (not only Machenhauer's method) is that it becomes complicated for Ri = $\mathcal{O}(1)$ owing to the strong coupling between balanced and unbalanced motions. The same applies for the QG filter. For regimes with Ri < 1 where baroclinic instability cannot exist anymore, the onset of other instabilities would introduce new complications to diagnose IGWs. However, for regimes that do allow the separation of the balanced and unbalanced motions, IGWs can be effectively diagnosed by the diagnostic tools presented in this thesis.
The efficient detection of IGWs based on Machenhauer's method and the QG filter is encouraging and can potentially be utilized for forthcoming studies.

4.2 Gravity wave activity in different dynamical regimes

IGW activity in different dynamical regimes is discussed in Chapter 2, where an idealized numerical model is used to mimic baroclinic instability ranging from ageostrophic (Ri = $\mathcal{O}(1)$) to quasi-geostrophic (Ri $\gg 1$). IGW activity in an ageostrophic regime is much more pronounced than in a quasi-geostrophic regime. For a quasi-geostrophic regime most of the energy is contained in the non-linear balanced mode and negligibly small energy in the non-linear unbalanced mode, indicating weak IGW activity.

Further, the dissipation associated with the balanced and unbalanced mode indicates that the unbalanced mode related to IGWs always dissipates through small-scale dissipation. The dissipation of balanced mode is predominantly through large-scale dissipation for a quasi-geostrophic regime, but as one proceeds from quasi-geostrophic to ageostrophic regime small-scale dissipation occupies larger and larger share of the dissipation of the balanced mode. Therefore, we can conclude that an ageostrophic regime features the downscale energy transfer primarily through IGWs.

Prospects
As suggested by the results, IGW emission depends on the flow parameters such as Ri, and thus a similar analysis for a range of Ri for other IGW emission mechanisms would provide more insight into the energy transfer from balanced flow to IGWs. Further, the IGWs discussed here might be generated by both spontaneous emission and convective instability. Distinguishing the source of IGW emission can be achieved in one way by identifying the

characteristics of IGWs emitted by a particular mechanism. This task depends upon the precise detection of emitted IGW signals, and for this purpose enhancing the capabilities of the diagnostic tools employed in this thesis is yet another subject of further research.

For $\mathrm{Ri} < 1$, instabilities such as symmetric instability and Kelvin-Helmholtz instability cast a regime where the dynamics are primarily unbalanced, such regimes might have even stronger IGW activity. Such a postulate is difficult to make though, because a $\mathrm{Ri} < 1$ regime leads to further escalation in the complexity of separating IGWs from a pool of instabilities with similar scales, such that the diagnostic tools based on the concept of balance cannot segregate these co-existing motions. Due to much more pronounced IGW activity such regimes might also serve as a transition regime between baroclinic instability, small-scale turbulence, and molecular scales where viscous dissipation occurs.

4.3 Gravity wave emission by different mechanisms

IGW emission by different mechanisms is discussed in Chapter 2, where an idealized model setup is used to simulate different scenarios of IGW emission by baroclinic instability: spontaneous emission, convective instability, and lateral boundary instability.

Spontaneous emission of IGWs is shown for $\mathrm{Ri} \gg 1$ representing a quasi-geostrophic regime. The residual vertical velocity obtained from the QG filter shows wave packets encircled by the baroclinic eddy field, with local Rossby number of $\mathscr{O}(1)$ in the regions of high wave activity, confirming that the wave signals are indeed IGWs. Although a quasi-geostrophic regime exhibits weak IGW emission, as discussed previously, the QG-filter is efficient in detecting the IGW signals.

Convective instability is a special case of spontaneous emission arising as a consequence of frontogenesis. Baroclinic instability can lead to sharpening of density gradients resulting in a density front. The gradual steepening of isopycnals in the front ultimately leads to an convectively unstable situation resulting in overturning of isopycnals. This frontal collapse is accompanied by emission of IGWs which propagate away from the front and eventually decay. This scenario of wave emission via frontogenesis could also be related to geostrophic adjustment, where the flow system tries to regain balance after the frontal collapse by emitting unbalanced IGWs.

Lateral boundary instability leads to emission of IGWs when balanced flow encounters a wall. This scenario is akin to an eddy colliding with a lateral boundary such as a continental shelf in the ocean. When a balanced eddy collides with meridional walls in

the model, unbalanced IGWs are generated. These IGWs exhibit the characteristics of a high-mode number Kelvin wave which appear to be arrested by the eddies.

Prospects

The magnitude of IGWs generated by convective instability (discussed for Ri = 100) is higher than that for spontaneous emission (discussed for Ri = 500) and lateral boundary instability (discussed for Ri = 400). This is presumably due to the characteristic weak IGW emission for higher Ri and an enhanced emission for smaller Ri. Therefore, an analysis of these mechanisms of IGW generation in the light of Ri dependence will provide useful aid in their quantification and relative importance.

Furthermore, other mechanisms of IGW generation from balanced flows as well as the extension of the present research work to more realistic scenarios is a plan of action for further research. For instance, the lateral boundary instability studied here with simplified topography implemented as meridional walls, can be further extended by including more realistic and complex topography to further examine the characteristics of emitted IGWs, which forms a prospective case for further study.

4.4 Role of gravity waves in the downscale transfer of energy

The question of the role of IGWs in the downscale transfer of energy addressed in the beginning of the thesis is answered in the light of the results presented. IGW activity examined in different settings supports and strengthens the role of IGWs in the downscale energy transfer. Baroclinic instability in all regimes ranging from Ri = $\mathcal{O}(1)$ to Ri $\gg$ (1) exhibits a predominant small-scale dissipation related to the unbalanced mode. For Ri = $\mathcal{O}(1)$ regime, even the balanced mode dissipation occurs primarily by small-scale dissipation, implying that in such a regime the energy transfer from balanced flow *en route* to molecular scale dissipation occurs mainly by IGWs. For Ri $\gg$ (1) regime, which shows weaker IGW activity, also shows small-scale dissipation through IGWs. The emission of IGWs by different mechanisms further reinforces the link between oceanic balanced flows and IGWs which act as a sink to balanced energy.

Prospects

In addition to the vital energy pathway between balanced flows and IGWs, downscale energy transfer in the ocean can occur through several different pathways. The unbalanced motions generated from balanced motions can form resonant triads with balanced as well as unbalanced motions. Resonant triad interactions between slow balanced and fast unbalanced motions, in particular the slow-fast-fast triad, can transfer energy from the balanced motions towards smaller scales (Bartello, 1995).

In addition, unbalanced IGWs which are generated from balanced flows can further interact to form wave-wave resonant triads and transfer energy to other waves or to small-scale turbulence, for instance by parametric subharmonic instability or induced diffusion (e.g. McComas and Bretherton, 1977; Müller et al., 1986; Olbers et al., 2012, also Section 1.2.2). Moreover, the generated IGWs can interact with the balanced flow and draw its energy by gravity wave drag (Eden and Olbers, 2017). Subsequently, the energy transfer from balanced flows to IGWs and from IGWs to small-scale dissipation can become far more complicated than the mechanisms demonstrated in the thesis.

One or more of these IGW related processes may simultaneously be at play and extract energy from the balanced flows. This makes the detection, characterization, and quantification of these processes and their contribution to the downscale energy transfer a challenging task. To include and identify such complex processes the diagnostic tools presented need further refinement to tackle the disentangling of such complex interactions, which is currently being explored.

In consideration of the foregoing, the precise detection of IGWs is a prerequisite, for which the diagnostic tools presented in this thesis appear proficient. There is much more to explore, but the journey has just begun.

4.5 Perspectives

The results presented in this thesis further foster the candidature of IGWs in the dissipation of balanced flow during the downscale transfer of energy in the ocean, and also give way to several new questions.

In this thesis, IGW emission is examined in an idealized numerical environment and the question of how these scenarios translate to the realistic ocean is an obvious one. In a realistic scenario wherein IGWs co-exist with many other processes of diverse scales, all of which exhibit complex non-linear interactions, identifying the energy transfers to and

from IGWs becomes inextricably tricky and even more complicated. IGWs interact and exchange energy with motions of diverse scales, whereas the idealized scenarios considered here include only a subset of these interactions.

In order to perceive the intricacies of the real ocean, new studies are required that include more complex flow configurations, realistic model setups, as well as other forcing such as winds, tides, and topography. This further leads to the question of how the global picture looks like for different scenarios of IGW emission: how can these processes of IGW emission be extrapolated to the global ocean; what are their quantitative estimates; what is the relative importance of these processes and their role in the ocean's energy cycle. Further, the diagnostic tools put forward in this thesis appear promising and would benefit from further refinement to facilitate understanding of other complex interactions such as the resonant triad interactions. Furthermore, the diagnostic procedures can be extended to other model setups such as a β-plane model. Even further, IGWs exist and influence the dynamics of the rotating flows in the Earth's interior as well as the dynamics of other rotating planets beyond the realms of our own blue planet.

The implications of the results contained in this thesis range from the instantaneous mixing to the climate that spans millenia, therefore a familiarity with the IGWs and their energetics is a prerequisite to perceive the subtleties of the overarching picture. IGWs and their interactions are largely parameterized in ocean and climate models as of now, and they are likely to remain unresolved in the face of the complex interactions they present and the computational power available at present. The picture might change in the future, but as a Danish proverb says, it is hard to make predictions, especially about the future.

Acknowledgments

First and foremost, I am sincerely thankful to my supervisor Carsten Eden for the splendid opportunity to bring this work to fruition. I owe my deepest gratitude to him for always pushing my horizons and inspiring me when I doubted my own abilities.

I thank the members of my advisory panel, Jin-Song von Storch and Johanna Baehr for their time and helpful discussions during my panel meetings.

I am thankful to Ulrich Achatz for the insightful discussions and for taking the time to serve as a referee for this thesis.

I express my thanks to Dirk Olbers for the lively discussions and for answering my questions promptly.

I have had wonderful moments with my working group: Alexa, Nils, Philipp, Friederike, Martina, Meike, Jennifer, Lars, Brenda, Rachael, and Thomas. Thanks to Martina for her gentle humor and warm hugs.

Special thanks to Nils for helping me to get into the world of the model I began with and for all the prolific discussions that helped me grasp complex concepts. I deeply thank Alexa, Nils, Jennifer, and Thomas Pohlmann for the helpful discussions and Philipp and Friederike for providing their help related to the German content in this thesis.

I thank IMPRS-ESM for giving me the opportunity to realize this work. A sincere thanks to Antje, Connie, and Wiebke for their incredible support right from the beginning and all throughout. You made life so much easier for a young student flying miles away from home.

Most of all, to my dear parents I express my heartfelt gratefulness. An ocean of love to my lovely mother who inspired me to never stop believing in my capabilities. My hearty thanks to my parents for their love and patience during these years. These cold grey windy wet days all four seasons in Hamburg would not have been comforting but for their warmth.

I immensely thank my dear sister, Ashwita, for being here and brightening up the otherwise rainy Hamburg days as well as for cooking delicious food for me.

To my dear friends who happily go insane with me: Fei, Jörg, Rohit. And for all the cooking evenings! Thanks Mariia, Melinda and Cathy for the all the laughter and more. Thank you all of you for being so, or the world would have been too sane for me.

Many thanks to all the people who have been a part of this joyous journey.

It is a pleasure to thank all the beautiful human beings who kept me going all these wonderful years which now weave together in this thesis.

Thank you!

O' misty Hamburg, coldethst town

My heart it freezeth, the sun is down

The air no breezeth, the river ice

I work on theezith, but people nice!

— Penned by Prof. Dirk Olbers

Bibliography

Alford, M. H. and Pinkel, R. (2000). Observations of overturning in the thermocline: The context of ocean mixing. *Journal of Physical Oceanography*, 30(5):805–832.

Arbic, B. K., Shriver, J. F., Hogan, P. J., Hurlburt, H. E., McClean, J. L., Metzger, E. J., Scott, R. B., Sen, A., Smedstad, O. M., and Wallcraft, A. J. (2009). Estimates of bottom flows and bottom boundary layer dissipation of the oceanic general circulation from global high-resolution models. *Journal of Geophysical Research: Oceans*, 114(C2).

Baer, F. (1977). Adjustment of initial conditions required to suppress gravity oscillations in nonlinear flows. *Beitraege zur Physik der Atmosphaere*, 50(3):350–366.

Baer, F. and Tribbia, J. J. (1977). On complete filtering of gravity modes through nonlinear initialization. *Monthly Weather Review*, 105(12):1536–1539.

Barkan, R., Winters, K. B., and McWilliams, J. C. (2017). Stimulated imbalance and the enhancement of eddy kinetic energy dissipation by internal waves. *Journal of Physical Oceanography*, 47(1):181–198.

Bartello, P. (1995). Geostrophic adjustment and inverse cascades in rotating stratified turbulence. *Journal of the atmospheric sciences*, 52(24):4410–4428.

Batchelor, G. K. (1969). Computation of the energy spectrum in homogeneous two-dimensional turbulence. *The Physics of Fluids*, 12(12):II–233.

Bell, T. (1975). Topographically generated internal waves in the open ocean. *Journal of Geophysical Research*, 80(3):320–327.

Blumen, W. (1972). Geostrophic adjustment. *Reviews of Geophysics*, 10(2):485–528.

Borchert, S., Achatz, U., and Fruman, M. D. (2014). Gravity wave emission in an atmosphere-like configuration of the differentially heated rotating annulus experiment. *Journal of Fluid Mechanics*, 758:287–311.

Brüggemann, N. and Eden, C. (2015). Routes to dissipation under different dynamical conditions. *Journal of Physical Oceanography*, 45(8):2149–2168.

Capet, X., McWilliams, J. C., Molemaker, M. J., and Shchepetkin, A. (2008). Mesoscale to submesoscale transition in the California Current System. Part III: Energy balance and flux. *Journal of Physical Oceanography*, 38(10):2256–2269.

Charney, J. G. (1947). The dynamics of long waves in a baroclinic westerly current. *Journal of Meteorology*, 4(5):136–162.

Charney, J. G. (1971). Geostrophic turbulence. *Journal of the Atmospheric Sciences*, 28(6):1087–1095.

Condie, S. and Condie, R. (2016). Retention of plankton within ocean eddies. *Global Ecology and Biogeography*, 25(10):1264–1277.

Cummins, P. F., Holloway, G., and Gargett, E. (1990). Sensitivity of the gfdl ocean general circulation model to a parameterization of vertical diffusion. *Journal of Physical Oceanography*, 20(6):817–830.

Danioux, E., Vanneste, J., Klein, P., and Sasaki, H. (2012). Spontaneous inertia-gravity-wave generation by surface-intensified turbulence. *Journal of Fluid Mechanics*, 699:153–173.

Dewar, W. K. and Hogg, A. M. (2010). Topographic inviscid dissipation of balanced flow. *Ocean Modelling*, 32(1):1–13.

Eady, E. T. (1949). Long waves and cyclone waves. *Tellus*, 1(3):33–52.

Eden, C. (2016). Closing the energy cycle in an ocean model. *Ocean Modelling*, 101:30–42.

Eden, C. and Olbers, D. (2014). An energy compartment model for propagation, nonlinear interaction, and dissipation of internal gravity waves. *Journal of Physical Oceanography*, 44(8):2093–2106.

Eden, C. and Olbers, D. (2017). A closure for internal wave-mean flow interaction. P]art B: Wave drag. *Journal of Physical Oceanography*, (2017).

Ferrari, R. and Wunsch, C. (2009). Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annual Review of Fluid Mechanics*, 41.

Ford, R., McIntyre, M. E., and Norton, W. A. (2000). Balance and the slow quasimanifold: some explicit results. *Journal of the Atmospheric Sciences*, 57(9):1236–1254.

Fox-Kemper, B., Ferrari, R., and Hallberg, R. (2008). Parameterization of mixed layer eddies. part i: Theory and diagnosis. *Journal of Physical Oceanography*, 38(6):1145–1165.

Gent, P. R. and McWilliams, J. C. (1983). Consistent balanced models in bounded and periodic domains. *Dynamics of Atmospheres and Oceans*, 7(2):67–93.

Gertz, A. and Straub, D. N. (2009). Near-inertial oscillations and the damping of midlatitude gyres: A modeling study. *Journal of Physical Oceanography*, 39(9):2338–2350.

Gill, A. E. (1982). *International Geophysics, 30: Atmosphere-ocean Dynamics*. Elsevier.

Görtler, H. (1943). Über eine schwingungserscheinung in flüssigkeiten mit stabiler dichteschichtung. *Zeitschrift Angewandte Mathematik und Mechanik*, 23:65–71.

Gregg, M. (1987). Diapycnal mixing in the thermocline: A review. *Journal of Geophysical Research: Oceans*, 92(C5):5249–5286.

Haine, T. W. and Marshall, J. (1998). Gravitational, symmetric, and baroclinic instability of the ocean mixed layer. *Journal of physical oceanography*, 28(4):634–658.

Heathershaw, A. (1985). Some observations of internal wave current fluctuations at the shelf-edge and their implications for sediment transport. *Continental Shelf Research*, 4(4):485–493.

Holton, J. R. and Hakim, G. J. (2012). *An introduction to dynamic meteorology*, volume 88. Academic press.

Hoskins, B., Draghici, I., and Davies, H. (1978). A new look at the ω-equation. *Quarterly Journal of the Royal Meteorological Society*, 104(439):31–38.

Hoskins, B., Pedder, M., and Jones, D. W. (2003). The omega equation and potential vorticity. *Quarterly Journal of the Royal Meteorological Society*, 129(595):3277–3303.

Houghton, D. D., Baumhefner, D. P., and Washington, W. M. (1971). On global initialization of the primitive equations: Part ii. the divergent component of the horizontal wind. *Journal of Applied Meteorology*, 10(4):626–634.

James, I. (2002). Modelling pollution dispersion, the ecosystem and water quality in coastal waters: a review. *Environmental Modelling & Software*, 17(4):363–385.

Kafiabad, H. A. and Bartello, P. (2016). Balance dynamics in rotating stratified turbulence. *Journal of Fluid Mechanics*, 795:914–949.

Kafiabad, H. A. and Bartello, P. (2017). Rotating stratified turbulence and the slow manifold. *Computers & Fluids*, 151:23–34.

Kolmogorov, A. N. (1941). Dissipation of energy in locally isotropic turbulence. In *Dokl. Akad. Nauk SSSR*, volume 32, pages 16–18.

Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S. (2007). On the driving processes of the atlantic meridional overturning circulation. *Reviews of Geophysics*, 45(2).

Leith, C. (1980). Nonlinear normal mode initialization and quasi-geostrophic theory. *Journal of the Atmospheric Sciences*, 37(5):958–968.

Lorenz, E. N. (1955). Available potential energy and the maintenance of the general circulation. *Tellus*, 7(2):157–167.

Lorenz, E. N. (1980). Attractor sets and quasi-geostrophic equilibrium. *Journal of the Atmospheric Sciences*, 37(8):1685–1699.

Lorenz, E. N. (1986). On the existence of a slow manifold. *Journal of the Atmospheric Sciences*, 43(15):1547–1558.

Lorenz, E. N. (1992). The slow manifold-What is it? *Journal of the atmospheric sciences*, 49(24):2449–2451.

Lorenz, E. N. and Krishnamurthy, V. (1987). On the nonexistence of a slow manifold. *Journal of the Atmospheric Sciences*, 44(20):2940–2950.

Lynch, P. (2001). The swinging spring: a simple model of atmospheric balance.

Machenhauer, B. (1977). On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitr. Phys. Atmos.*, 50:253–271.

MacKinnon, J. A., Alford, M. H., Ansong, J. K., Arbic, B. K., Barna, A., Briegleb, B. P., Bryan, F. O., Buijsman, M. C., Chassignet, E. P., Danabasoglu, G., et al. (2017). Climate process team on internal-wave driven ocean mixing. *Bulletin of the American Meteorological Society*, (2017).

Marshall, J. and Speer, K. (2012). Closure of the meridional overturning circulation through southern ocean upwelling. *Nature Geoscience*, 5(3):171–180.

McComas, C. H. and Bretherton, F. P. (1977). Resonant interaction of oceanic internal waves. *Journal of Geophysical Research*, 82(9):1397–1412.

McKiver, W. J. and Dritschel, D. G. (2008). Balance in non-hydrostatic rotating stratified turbulence. *Journal of Fluid Mechanics*, 596:201–219.

Molemaker, M. J., McWilliams, J. C., and Capet, X. (2010). Balanced and unbalanced routes to dissipation in an equilibrated Eady flow. *Journal of Fluid Mechanics*, 654:35–63.

Molemaker, M. J., McWilliams, J. C., and Yavneh, I. (2005). Baroclinic instability and loss of balance. *Journal of Physical Oceanography*, 35(9):1505–1517.

Müller, P., Holloway, G., Henyey, F., and Pomphrey, N. (1986). Nonlinear interactions among internal gravity waves. 24:493–536.

Munday, D., Johnson, H., and Marshall, D. (2014). Impacts and effects of mesoscale ocean eddies on ocean carbon storage and atmospheric pco2. *Global Biogeochemical Cycles*, 28(8):877–896.

Munk, W. and Wunsch, C. (1998). Abyssal recipes ii: energetics of tidal and wind mixing. *Deep Sea Research Part I: Oceanographic Research Papers*, 45(12):1977–2010.

Nagai, T., Tandon, A., Kunze, E., and Mahadevan, A. (2015). Spontaneous generation of near-inertial waves by the Kuroshio Front. *Journal of Physical Oceanography*, 45(9):2381–2406.

Nansen, F. (1897). Farthest north vol i.

Nikurashin, M. and Ferrari, R. (2011). Global energy conversion rate from geostrophic flows into internal lee waves in the deep ocean. *Geophysical Research Letters*, 38(8).

Olbers, D. and Eden, C. (2013). A global model for the diapycnal diffusivity induced by internal gravity waves. *Journal of Physical Oceanography*, 43(8):1759–1779.

Olbers, D. and Eden, C. (2016). Revisiting the generation of internal waves by resonant interaction with surface waves. *Journal of Physical Oceanography*, 46(8):2335–2350.

Olbers, D. and Eden, C. (2017). A closure for internal wave-mean flow interaction. part a: Energy conversion. *Journal of Physical Oceanography*, (2017).

Olbers, D., Willebrand, J., and Eden, C. (2012). *Ocean dynamics*. Springer Science & Business Media.

Pedlosky, J. (1987). *Geophysical fluid dynamics*. Springer Science & Business Media.

Phillips, N. A. (1954). Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus*, 6(3):274–286.

Pinot, J.-M., Tintoré, J., and Wang, D.-P. (1996). A study of the omega equation for diagnosing vertical motions at ocean fronts. *Journal of marine research*, 54(2):239–259.

Plougonven, R., Muraki, D. J., and Snyder, C. (2005). A baroclinic instability that couples balanced motions and gravity waves. *Journal of the atmospheric sciences*, 62(5):1545–1559.

Plougonven, R. and Snyder, C. (2007). Inertia–gravity waves spontaneously generated by jets and fronts. Part I: Different baroclinic life cycles. *Journal of the atmospheric sciences*, 64(7):2502–2520.

Plougonven, R. and Zhang, F. (2014). Internal gravity waves from atmospheric jets and fronts. *Reviews of Geophysics*, 52(1):33–76.

Pollmann, F., Eden, C., and Olbers, D. (2017). Evaluating the global internal wave model IDEMIX using finestructure methods. *Journal of Physical Oceanography*, (2017).

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., et al. (2007). Climate models and their evaluation. In *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR)*, pages 589–662. Cambridge University Press.

Rhines, P. B. (1977). The dynamics of unsteady currents. *The sea*, 6:189–318.

Rossby, C.-G. (1938). On the mutual adjustment of pressure and velocity distributions in certain simple current systems, II. *J. mar. Res*, 1(3):239–263.

Saenko, O. A. (2006). The effect of localized mixing on the ocean circulation and time-dependent climate change. *Journal of physical oceanography*, 36(1):140–160.

Sallée, J.-B., Matear, R. J., Rintoul, S. R., and Lenton, A. (2012). Localized subduction of anthropogenic carbon dioxide in the southern hemisphere oceans. *Nature Geoscience*, 5(8):579–584.

Schlösser, F. and Eden, C. (2007). Diagnosing the energy cascade in a model of the north atlantic. *Geophysical research letters*, 34(2).

Schmidt, G. M., Wall, M., Taylor, M., Jantzen, C., and Richter, C. (2016). Large-amplitude internal waves sustain coral health during thermal stress. *Coral Reefs*, 35(3):869–881.

Schmitz Jr, W. (1996). On the world ocean circulation: Volume i: some global features/north atlantic circulation. woods hole oceanogr. Technical report, Inst. Tech. Rept., WHOI-96-03.

Scott, R. B. and Wang, F. (2005). Direct evidence of an oceanic inverse kinetic energy cascade from satellite altimetry. *Journal of Physical Oceanography*, 35(9):1650–1666.

Shakespeare, C. J. and Taylor, J. (2015). The spontaneous generation of inertia–gravity waves during frontogenesis forced by large strain: numerical solutions. *Journal of Fluid Mechanics*, 772:508–534.

Shakespeare, C. J. and Taylor, J. R. (2016). Spontaneous wave generation at strongly strained density fronts. *Journal of Physical Oceanography*, 46(7):2063–2081.

Shulzitski, K., Sponaugle, S., Hauff, M., Walter, K. D., and Cowen, R. K. (2016). Encounter with mesoscale eddies enhances survival to settlement in larval coral reef fishes. *Proceedings of the National Academy of Sciences*, 113(25):6928–6933.

Simmons, H. L., Jayne, S. R., Laurent, L. C. S., and Weaver, A. J. (2004). Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Modelling*, 6(3):245–263.

Staquet, C. and Sommeria, J. (2002). Internal gravity waves: from instabilities to turbulence. *Annual Review of Fluid Mechanics*, 34(1):559–593.

Stommel, H. (1961). Thermohaline convection with two stable regimes of flow. *Tellus*, 13(2):224–230.

Stone, P. H. (1966). On non-geostrophic baroclinic stability. *Journal of the Atmospheric Sciences*, 23(4):390–400.

Stone, P. H. (1971). Baroclinic stability under non-hydrostatic conditions. *Journal of Fluid Mechanics*, 45(4):659–671.

Sugimoto, N. and Plougonven, R. (2016). Generation and backreaction of spontaneously emitted inertia-gravity waves. *Geophysical Research Letters*, 43(7):3519–3525.

Sutherland, B. R. (2010). *Internal gravity waves*. Cambridge University Press.

Talley, L. D. (2011). *Descriptive physical oceanography: an introduction*. Academic press.

Talley, L. D. (2013). Closure of the global overturning circulation through the indian, pacific, and southern oceans: Schematics and transports. *Oceanography*, 26(1):80–97.

Temperton, C. and Williamson, D. L. (1981). Normal mode initialization for a multilevel grid-point model. part i: Linear aspects. *Monthly Weather Review*, 109(4):729–743.

Theiss, J. and Mohebalhojeh, A. R. (2009). The equatorial counterpart of the quasi-geostrophic model. *Journal of Fluid Mechanics*, 637:327–356.

Thomas, L. N., Tandon, A., and Mahadevan, A. (2008). Submesoscale processes and dynamics. *Ocean modeling in an Eddying Regime*, pages 17–38.

Thorpe, S. A. (2007). *An introduction to ocean turbulence*. Cambridge University Press.

Toggweiler, J. and Samuels, B. (1993). New radiocarbon constraints on the upwelling of abyssal water to the ocean?s surface. In *The global carbon cycle*, pages 333–366. Springer.

Toggweiler, J. and Samuels, B. (1995). Effect of drake passage on the global thermohaline circulation. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(4):477–500.

Toggweiler, J. and Samuels, B. (1998). On the ocean?s large-scale circulation near the limit of no vertical mixing. *Journal of Physical Oceanography*, 28(9):1832–1852.

Vallis, G. K. (1996). Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Quarterly Journal of the Royal Meteorological Society*, 122(529):291–322.

Vallis, G. K. (2006). *Atmospheric and oceanic fluid dynamics*. Cambridge University Press.

Vanneste, J. (2013). Balance and spontaneous wave generation in geophysical flows. *Annual Review of Fluid Mechanics*, 45.

Vanneste, J. and Yavneh, I. (2004). Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *Journal of the atmospheric sciences*, 61(2):211–223.

Viúdez, Á. and Dritschel, D. G. (2004). Optimal potential vorticity balance of geophysical flows. *Journal of Fluid Mechanics*, 521:343–352.

Warn, T. (1997). Nonlinear balance and quasi-geostrophic sets. *Atmosphere-Ocean*, 35(2):135–145.

Williams, P. D., Haine, T. W., and Read, P. L. (2008). Inertia-gravity waves emitted from balanced flow: Observations, properties, and consequences. *Journal of the atmospheric sciences*, 65(11):3543–3556.

Wunsch, C. and Ferrari, R. (2004). Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.*, 36:281–314.

Xie, J.-H. and Vanneste, J. (2015). A generalised-Lagrangian-mean model of the interactions between near-inertial waves and mean flow. *Journal of Fluid Mechanics*, 774:143–169.

Zeitlin, V. (2008). Decoupling of balanced and unbalanced motions and inertia–gravity wave emission: Small versus large Rossby numbers. *Journal of the Atmospheric Sciences*, 65(11):3528–3542.

This ocean seemingly tranquil

Wraps wild waves within

Serene as a work of art

Yet restless as a young heart