SUBMESOSCALE TURBULENCE IN THE UPPER OCEAN

by

Jörn Callies

Dipl.-Met., Universität Hamburg, 2011

Submitted in partial fulfillment of the requirements for the degree of
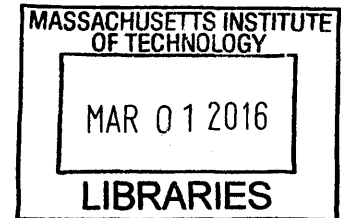
Doctor of Philosophy

at the

Massachusetts Institute of Technology

and the

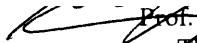Woods Hole Oceanographic Institution

February 2016

Signature redacted

Signature of author: _____

Joint Program in Oceanography/Applied Ocean Science & Engineering
Massachusetts Institute of Technology and Woods Hole Oceanographic Institution
18 September 2015

Signature redacted

Certified by: _____

Prof. Raffaele Ferrari
Thesis supervisor
Massachusetts Institute of Technology

Signature redacted

Accepted by: _____

Dr. Larry J. Pratt
Chair, Joint Committee for Physical Oceanography
Woods Hole Oceanographic Institution

Submesoscale Turbulence in the Upper Ocean

by

Jörn Callies

Submitted on 18 September 2015 to the Joint Program in Oceanography/Applied Ocean Science & Engineering, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

Submesoscale flows, current systems 1–100 km in horizontal extent, are increasingly coming into focus as an important component of upper-ocean dynamics. A range of processes have been proposed to energize submesoscale flows, but which process dominates in reality must be determined observationally. We diagnose from observed flow statistics that in the thermocline the dynamics in the submesoscale range transition from geostrophic turbulence at large scales to inertia–gravity waves at small scales, with the transition scale depending dramatically on geographic location. A similar transition is shown to occur in the atmosphere, suggesting intriguing similarities between atmospheric and oceanic dynamics. We furthermore diagnose from upper-ocean observations a seasonal cycle in submesoscale turbulence: fronts and currents are more energetic in the deep wintertime mixed layer than in the summertime seasonal thermocline. This seasonal cycle hints at the importance of baroclinic mixed layer instabilities in energizing submesoscale turbulence in winter. To better understand this energization, three aspects of the dynamics of baroclinic mixed layer instabilities are investigated. First, we formulate a quasi-geostrophic model that describes the linear and nonlinear evolution of these instabilities. The simple model reproduces the observed wintertime distribution of energy across scales and depth, suggesting it captures the essence of how the submesoscale range is energized in winter. Second, we investigate how baroclinic instabilities are affected by convection, which is generated by atmospheric forcing and dominates the mixed layer dynamics at small scales. It is found that baroclinic instabilities are remarkably resilient to the presence of convection and develop even when rapid overturns keep the mixed layer unstratified. Third, we discuss the restratification induced by baroclinic mixed layer instabilities. We show that the rate of restratification depends on characteristics of the baroclinic eddies themselves, a dependence not captured by a previously proposed parameterization. These insights sharpen our understanding of submesoscale dynamics and can help focus future inquiry into whether and how submesoscale flows influence the ocean's role in climate.

Thesis supervisor: Raffaele Ferrari
Title: Breene M. Kerr Professor of Oceanography

ACKNOWLEDGMENTS

I thank my advisor, Raffaele Ferrari, for his inspiration and unlimited enthusiasm in often long and sometimes late meetings over the past four years. I am looking forward to continuing this enjoyable and productive collaboration in the future.

My thesis committee, Glenn Flierl, Baylor Fox-Kemper, Jim Ledwell, Alan Plumb, and Shafer Smith, offered generous comments, advice, and criticism, which helped improve my work and encouraged me to clarify its context. Carl Wunsch generously agreed to chair the defense. He also offered insightful advice along the way.

I am grateful that I was able to incorporate into my work observations for which many have invested considerable time and effort. Both the Oleander and LatMix teams have generously given me access to their data and provided help in understanding the capabilities and limitations of the instruments. Kathleen Donohue, Andrey Shcherbina, and Jody Klymak were particularly helpful.

Much support was also available for the computational work, in particular that with the MITgcm. Jean-Michel Campin helped implement changes to the code; Chris Hill, Greg Shomo, and Oliver Jahn helped with the logistics of running the code on the cluster. Their expertise prevented countless headaches.

I would also like to thank my fellow students in the Joint Program and EAPS as well as the postdocs at MIT for their companionship and friendship. Many discussions with them, both scientific and not, helped create supportive and stimulating environments in Cambridge and Woods Hole.

Special thanks are due to my friends and family for supporting me both from across the Atlantic and locally.

CONTENTS

CHAPTER 1

INTRODUCTION

Oceanic flows span an enormous ten orders of magnitude in scale, enabled by the ocean's global extent and seawater's minute viscosity. Gyre and overturning circulations fill entire ocean basins, while three-dimensional turbulence reaches down to millimeter scales. Physical oceanography in the 20th century has illuminated major components of this oceanic spectrum, uncovering important interactions across the tremendous range of scales.

Inertia–gravity waves occupy a large part of the oceanic spectrum. The study of their dynamics goes back to the 19th century, but it was the major expeditions of the early 20th century that discovered that the global ocean is host to a pervasive inertia–gravity wave field (e.g. Defant, 1961). Inertia–gravity waves have frequencies between the inertial frequency f and the buoyancy frequency N. A large part of the inertia–gravity wave energy is associated with near-inertial or tidal frequencies, but a continuum of frequencies between f and N bears energy (e.g. Ferrari and Wunsch, 2009). Garrett and Munk (1972) used linear wave theory to synthesize observations in the horizontal, vertical, and time coordinates to characterize the distribution of wave energy across scales. They focused on the frequency continuum and established these waves to dominate motions on horizontal scales smaller than 1 km throughout the thermocline and deep ocean. At small scales, inertia–gravity waves give way to three-dimensional turbulence when they break. The small-scale mixing induced by this occasional wave breaking is thought to play an important role in the global meridional overturning circulation. Wave breaking is thought to mix heat down into cold abyssal waters formed in polar regions, allowing these waters to return to the surface.

The "mesoscale revolution" (W. Munk in von Storch and Hasselmann, 2010) established another crucial component of the oceanic spectrum: geostrophic eddies of order 100 km in horizontal scale. These eddies are strongly constrained by the earth's rotation and the ocean's stratification; they evolve on subinertial time scales, typically weeks to months. They were first hinted at by acoustically tracked deep floats in the late 1950s (Swallow, 1971) and were fully exposed by the seminal MODE experiment (The MODE Group, 1978). Mesoscale eddies dominate the kinetic energy in much of the ocean and nowadays are routinely observed by satellite altimetry (e.g. Wunsch and Stammer, 1998). Fluxes of heat and salt achieved by mesoscale eddies have a profound impact on the gyre circulation (e.g. Rhines and Young, 1982) as well as Southern Ocean and meridional overturning dynamics (e.g. Johnson and Bryden, 1989; Marshall and Radko, 2003), the details of which are still being worked out.

This 20th century progress has mapped out much of the oceanic spectrum, but a wide gap remains between mesoscale eddies at order 100 km and inertia–gravity waves and three-dimensional turbulence at scales smaller than 1 km. This relatively unexplored range, roughly 1–100 km, will be referred to as the "submesoscale range"[1] and has re-

---

[1]Some authors prefer a dynamical definition of "submesoscale," for example requiring Rossby numbers of order 1 (e.g. Thomas et al., 2008). We instead use the term to designate the range of scales between the

cently come into the focus of physical oceanographic research. The dynamics governing flows in the submesoscale range are the topic of this dissertation.

In the main thermocline, balanced submesoscale flows are expected to be weak and energy to be largely trapped at mesoscales (Charney, 1971). The horizontal stirring of tracers in the interior is then dominated by mesoscale eddies, as evidenced by the filamentation of artificial tracers injected into the thermocline (e.g. Ledwell et al., 1998). Modeling studies and satellite observations suggest, however, that the dynamics can be quite different at the sea surface: under certain conditions a convoluted network of energetic submesoscale surface fronts forms (e.g. Capet et al., 2008a,d). These fronts, elongated regions of large surface buoyancy gradients, are associated with strong along-front currents that are to leading order geostrophically balanced.

Submesoscale surface fronts have been suggested to have an impact on larger-scale flows and tracer distributions, similar to what had been discovered for inertia–gravity waves and mesoscale eddies. This impact is thought to be mediated by vigorous exchanges between the surface and interior ocean that occur along these fronts (e.g. Capet et al., 2008a). Energetic submesoscale flows produce much larger vertical velocities than mesoscale eddies (Mahadevan and Tandon, 2006). Modeling studies suggest that these enhanced vertical exchanges can modify the physical properties of waters subducted into the thermocline (Lévy et al., 2010) as well as affect the biogeochemistry of the upper ocean (e.g. Ferrari, 2011; Mahadevan, 2014). These flows bring nutrients up into the surface layer, where they can be used for photosynthesis, and sequester carbon in the interior ocean. The small-scale nature of vertical velocities in the surface ocean parallels that in the midlatitude atmosphere, where precipitation caused by upward motion occurs at scales much smaller than the dominant synoptic eddies.

Besides their effect on exchanges between the surface and interior ocean, energetic near-surface submesoscale flows have also been suggested to strongly affect the stratification of the upper ocean. Submesoscale processes tend to flatten isopycnals, restratifying the surface layers (Lapeyre et al., 2006; Boccaletti et al., 2007). This modifies surface properties and feeds back on the exchange of heat and freshwater with the atmosphere, with the potential to affect both the oceanic general circulation and atmospheric climate. How deep the influence of surface-enhanced submesoscale flows reaches, however, remains unclear. The vertical structure of submesoscale flows is largely unconstrained observationally and not sufficiently understood dynamically.

Another motivation for studying submesoscale dynamics is the attempt to close the energy budget of mesoscale eddies. Mesoscale eddies are energized by thermocline baroclinic and barotropic instabilities, but how energy is drained from them remains unclear (e.g. Ferrari and Wunsch, 2010). Owing to the constraints imposed by rotation and stratification, mesoscale eddies tend to transfer energy to large scales. This prevents mesoscale

---

most energetic mesoscale eddies and about 1 km. This usage is closer to the literal meaning of the term and is more adept to our approach of attempting to infer the dynamics from observations of flows in this range.

energy from being effectively scattered to small scales, where it can be dissipated. A recently proposed solution to this quandary is that surface frontogenesis and ageostrophic submesoscale instabilities overcome the rotational constraint and thereby enable a transfer of mesoscale energy to small dissipation scales (Capet et al., 2008c; Molemaker et al., 2005, 2010). The relevance of this forward route to dissipation remains largely unquantified and can only be assessed once more is known about submesoscale dynamics.

The dynamics of how submesoscale surface fronts sharpen is fairly well understood from theory originally developed for atmospheric frontogenesis (e.g. Stone, 1966a; Hoskins and Bretherton, 1972; Hoskins, 1982). The sea surface plays an important role in the formation of these fronts, because horizontal strain and shear flows can effectively sharpen buoyancy gradients there. In the ocean interior, if horizontal buoyancy gradients are sharpened, an ageostrophic circulation develops in response, which downwells light water on the dense side and upwells dense water on the light side, thereby opposing the increase in horizontal buoyancy gradient. At the surface, on the other hand, the vertical velocity is constrained to be nearly zero, so the ageostrophic circulation cannot effectively oppose the increase in the horizontal buoyancy gradient induced by the horizontal strain or shear flow. On the contrary, horizontal advection by the ageostrophic circulation can accelerate the sharpening of surface buoyancy gradients and may lead to frontal collapse, the formation of buoyancy discontinuities in finite time.

It is not clear, however, what process is most effective in driving frontogenesis in the surface ocean. Two possibilities have been suggested: the straining by mesoscale eddies (e.g. Lapeyre and Klein, 2006) and the horizontal shear induced by baroclinic mixed layer instabilities (e.g. Boccaletti et al., 2007).

Mesoscale eddies have strain fields between their centers, which can act on surface buoyancy gradients and induce frontogenesis. A useful approximation of the submesoscale dynamics produced by mesoscale straining is the surface quasi-geostrophic system. Being quasi-geostrophic, it does not include the ageostrophic advection and thereby does not form any buoyancy discontinuities. But it does capture the fundamental difference in behavior between the surface and the interior, generating sharp surface buoyancy gradients. The surface quasi-geostrophic model can be used to study the dynamics in the strongly nonlinear regime, where filamentary instabilities energize a continuum of scales (Held et al., 1995). Surface quasi-geostrophic turbulence predicts the submesoscale range to be much more energetic near the surface than in the interior (Charney, 1971; Blumen, 1978). These predictions carry over to more complete dynamics, which lead to corrections at small submesoscales (order 1 km) but generally agree with surface quasi-geostrophic dynamics at larger scales (e.g. Hakim et al., 2002; Klein et al., 2008; Capet et al., 2008e).

Studies of mesoscale-driven surface frontogenesis commonly envision an upper ocean with nearly uniform interior potential vorticity. The real upper ocean, however, typically has a well mixed surface boundary layer. The transition from this weakly stratified mixed layer to the strongly stratified thermocline corresponds to a jump in potential vorticity, which can induce additional dynamics. In the presence of horizontal buoyancy gradients,

14

the jump in potential vorticity allows an amplification of baroclinic waves in the mixed layer, which slide dense under light waters (e.g. Haine and Marshall, 1998; Boccaletti et al., 2007). This mixed layer instability generates flows with horizontal shears at submesoscales, inducing frontogenesis not only at the surface but also at the base of the mixed layer.

Both of these processes, mesoscale-driven surface frontogenesis and baroclinic mixed layer instabilities, are from a theoretical point of view equally plausible generators of energetic submesoscale fronts and flows. A first theme of this dissertation is thus to turn to the limited observations of submesoscale flows to identify which of the two processes is more ubiquitous in the real ocean.

A major challenge to this program is that other types of motion can dominate over the balanced dynamics of frontogenesis. It is currently impracticable to collect observations simultaneously in space and time over the relevant scales, so it is typically unclear whether the observed flows are to leading order geostrophically balanced and slowly evolving or dominated by much faster dynamics. Inertia–gravity waves, in particular the near-inertial and tidal components, can have relatively large horizontal scales, so their dominance can reach considerably beyond 1 km and into the submesoscale range. Inertia–gravity waves have very different dynamics than balanced flows, with major implications for the transport of heat and carbon, for example, so care must be taken to not confuse these types of motion in observations.

In Chapter 2, we thus seek to identify the leading-order dynamics of submesoscale flows from in situ observations in two regions.[2] In the thermocline, away from the surface and mixed layer, two distinct submesoscale regimes are identified. In the western North Atlantic, a region with strong mesoscale eddies, submesoscale flows are in geostrophic balance and relatively weak compared to the mesoscale eddies. Energy drops off rapidly in the submesoscale range, consistent with interior quasi-geostrophic turbulence (Charney, 1971). In the eastern Pacific, where mesoscale eddies are less energetic, the submesoscales in the thermocline are instead dominated by inertia–gravity waves, all the way to scales as large as 100 km.

Chapter 2 illustrates how important it is to distinguish between geostrophic flows and inertia–gravity waves. In Chapter 3, a method formalizing the analysis of Chapter 2 is developed, which can decompose one-dimensional current and density measurements into components from these two types of motion.[3] The method's application to ocean observations confirms that in the thermocline large scales are dominated by geostrophic eddies and small scales by inertia–gravity waves—and that the transition between these two types of motion occurs at scales depending strongly on geographic location.

The method developed in Chapter 3 is applied to atmospheric data in Chapter 4, re-

---

[2]This chapter was published as J. Callies and R. Ferrari (2013) Interpreting Energy and Tracer Spectra of Upper-Ocean Turbulence in the Submesoscale Range (1–200 km). *J. Phys. Oceanogr.* 43 (11), 2456–2474.

[3]This chapter was published as O. Bühler, J. Callies, and R. Ferrari (2014) Wave–vortex decomposition of one-dimensional ship-track data. *J. Fluid Mech.* 756, 1007–1026.

vealing an intriguing similarity between midlatitude atmospheric and oceanic dynamics.[4] Similar to some of the ocean observations presented in Chapter 2, the atmospheric data exhibit a conspicuous difference in how energy is distributed across scales at large versus small scales (Nastrom and Gage, 1985). Various explanations based on turbulence theory have been put forward for this transition in the energy spectrum. It is argued in this chapter that the transition instead stems from inertia–gravity waves becoming dominant at small scales.

In Chapter 5, returning to the ocean, we present observational evidence for a seasonal cycle in submesoscale turbulence in the surface layers of the western North Atlantic.[5] In the summertime seasonal thermocline, submesoscale flows are relatively weak and consistent with interior quasi-geostrophic turbulence, as found for the main thermocline in Chapter 2. In the wintertime mixed layer, on the other hand, submesoscale flows are more energetic than in the summertime seasonal thermocline or in the main thermocline below. We attribute this wintertime energization to baroclinic mixed layer instabilities, which grow in deep winter mixed layers and form submesoscale fronts.

The seasonality described in Chapter 5 indicates how important a role baroclinic mixed layer instabilities play in energizing the submesoscales. This is taken as an occasion to further investigate the dynamics of these instabilities, which is the second theme of this dissertation.

Chapter 6 introduces a simple dynamical framework to characterize the process of how baroclinic mixed layer instabilities energize the submesoscale range. We formulate a quasi-geostrophic model that captures the essence of the linear and nonlinear evolution of these instabilities. The model provides a useful starting point for understanding the turbulent submesoscale dynamics in the presence of baroclinic mixed layer instabilities. We show that it successfully reproduces the distribution of energy across scales and depth observed in the wintertime western North Atlantic.

The quasi-geostrophic dynamics employed in Chapter 6 necessarily omit atmospherically forced convection, which keeps the mixed layer unstratified. The interaction between convection and baroclinic mixed layer instabilities is analyzed in Chapter 7. In an idealized setup that allows the concurrent simulation of baroclinic instabilities and convection, we find that baroclinic instabilities are remarkably resilient to the presence of convection. The instabilities develop even when convective overturns are vigorous. This finding helps isolate the mechanism behind the seasonality of submesoscale turbulence described in Chapter 5.

One of the important impacts of baroclinic instabilities is that they restratify the mixed layer. Current global models do not typically resolve this submesoscale instability,

---

[4]This chapter was published as J. Callies, R. Ferrari, and O. Bühler (2014) Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 111 (48), 17033–17038.

[5]This chapter was published as J. Callies, R. Ferrari, J. M. Klymak, and J. Gula (2015) Seasonality in submesoscale turbulence. *Nat. Commun.* 6 (6862).

so a scaling must be used to parameterize the restratification. In Chapter 8, we argue that the pace of restratification depends on characteristics of the baroclinic eddies themselves, a dependence not captured by a previously proposed parameterization (Fox-Kemper et al., 2008).

We summarize the conclusions in Chapter 9, detailing the implications for our understanding of upper-ocean dynamics. We offer an outlook on questions that will need to be addressed in future work.

CHAPTER 2

INTERPRETING ENERGY AND TRACER SPECTRA OF UPPER-OCEAN TURBULENCE IN THE SUBMESOSCALE RANGE

165 m depth
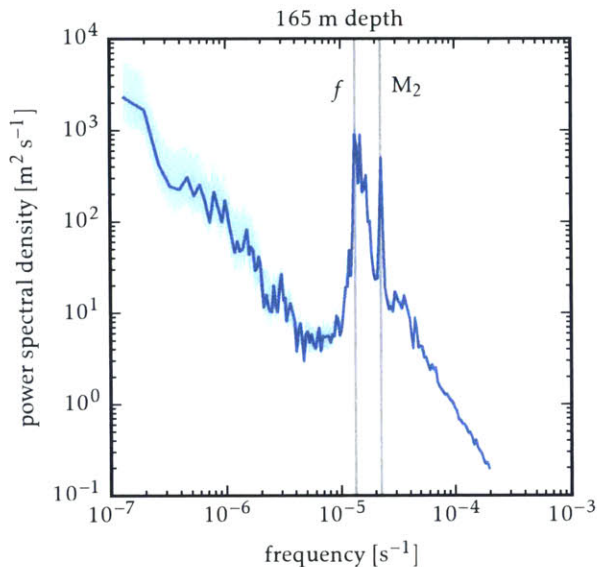
Figure 2.1: Frequency spectrum of kinetic energy calculated from mooring WHOI 794 at 35°N, 152°W (Oct. 1983 to Sept. 1985). Marked are the local inertial frequency f and the M_2 tidal frequency. The shading shows 95% confidence intervals.

2.1 Introduction

Oceanographers have long debated how energy is transferred from large to dissipative scales. Much progress has been made in describing the energy pathways from basin to mesoscales and then from scales[1] of order 1 km down to millimeter scales. But our understanding of the transfer in between, in the submesoscale range[2], is still rudimentary. A major question is whether the energy fluxes in the submesoscale are dominated by internal waves and other unbalanced motions or whether there is an important contribution by geostrophic motions. The geostrophic part of submesoscale motions has received much attention in theoretical studies over the past decade (e.g. Lapeyre and Klein, 2006), but supporting analysis of observations is lacking. Our goal here is to use available observations from the upper midlatitude ocean to test different theories of submesoscale dynamics.

Oceanic motions can easily be separated in the frequency domain. Figure 2.1 shows an example frequency spectrum from a mooring in the subtropical North Pacific (Schmitz, Jr., 1988)[3]. Geostrophically balanced flows span the range of frequencies below the inertial frequency f, inertial oscillations are visible as a broad peak around f, the M_2 lunar

---

[1]All scales are given as wavelengths.

[2]We use the term *submesoscale* to designate the range 1–200 km, roughly the scales below the first deformation radius. There are other uses of the term in the literature, which are based on order-one Rossby and Richardson numbers. We here attempt to understand what the dynamics of these scales are, so we use *submesoscale* to designate the range of scales without presupposing their dynamics.

[3]The data are available at http://cmdac.oce.orst.edu/.
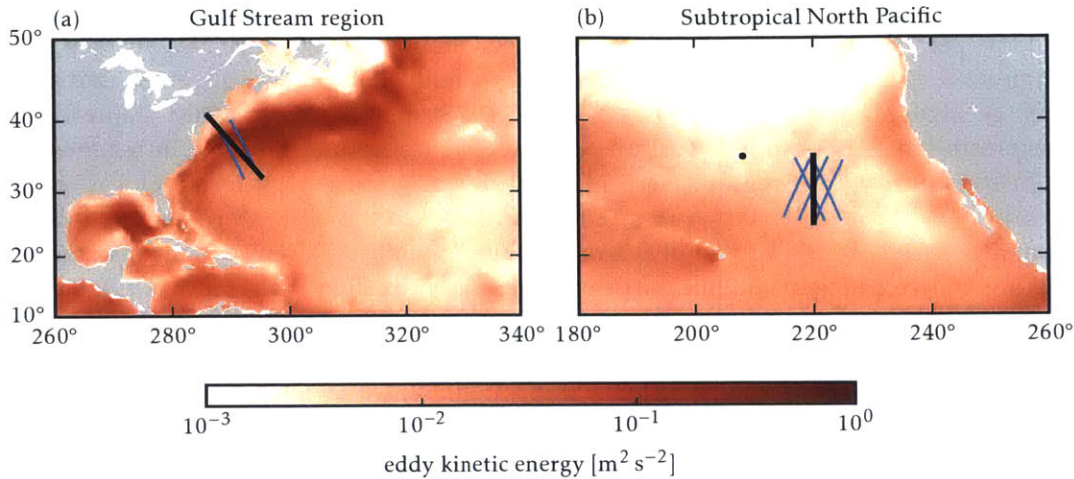
Figure 2.2: Tracks along which measurements were collected in the (a) Oleander and (b) Spice experiments (thick black lines). The satellite tracks used in the analysis are shown as thin blue lines. The site of the mooring used for Fig. 2.1 is marked in (b) with a black dot. The background color is a map of eddy kinetic energy from AVISO gridded altimetry.

tide appears as a sharp superinertial peak, and a broad-band internal-wave continuum spans the frequency range between f and the buoyancy frequency N. It is much less clear how these motions project onto spatial scales. This is a serious limitation to our understanding of submesoscale turbulence because the interaction of modes across spatial scales is at the core of turbulence theories.

It is generally believed that the kinetic energy in the midlatitude upper ocean is dominated by geostrophic eddies at scales larger than the first baroclinic deformation scale—the wavelength associated with the deformation radius—of about 50–200 km and by internal waves at scales below 1 km. We here attempt to understand the range of scales in between, the submesoscale range, using wavenumber spectra of kinetic energy, potential energy, and tracer variance, together with their vertical variations. Is most of the submesoscale variability balanced or unbalanced? Do our theories of geostrophic turbulence successfully describe the balanced part of the flow? Previous attempts to test these theories remain largely inconclusive (e.g. Samelson and Paulson, 1988; Stammer, 1997; Le Traon et al., 2008).

To answer these questions, we analyze two in-situ data sets in two different dynamical regimes: the Oleander data set (Wang et al., 2010) in the Gulf Stream region and the Spice data set (Ferrari and Rudnick, 2000) in the eastern subtropical North Pacific (Fig. 2.2). The Gulf Stream is a strong baroclinic current with an associated deep reversal of the potential vorticity (PV) gradient. This induces deep Phillips-type baroclinic instability and strong mesoscale eddy activity (e.g. Tulloch et al., 2011). In the North Pacific, in the quiescent

21

subtropical gyre, on the other hand, interior PV gradients are weak. A surface buoyancy gradient persists, however, so one might expect shallow instabilities and surface frontal dynamics to play a prominent role there. Lapeyre and Klein (2006) suggested that surface quasi-geostrophic (QG) turbulence (Blumen, 1978) may be the relevant framework for interpreting the submesoscale turbulence that develops in response to frontogenesis.

In the Gulf Stream region, we find spectra consistent with interior QG turbulence (Charney, 1971) at scales larger than 20 km and consistent with internal-wave dynamics at smaller scales. In the subtropical North Pacific, the observations are inconsistent with existing geostrophic-turbulence theories. There is indication that internal tides play an important role at scales as large as 100 km. At small scales, the internal-wave continuum again dominates.

The chapter is organized as follows. In Section 2.2, we review the existing theoretical predictions for the spectra of balanced and unbalanced flows. In Section 2.3, we discuss how these predictions relate to spectra computed from one-dimensional transects. We review previous observations in Section 2.4. In Sections 2.5 and 2.6, we introduce and compare against the theoretical predictions the two sets of observations. We discuss implications in Section 2.7. The appendix examines surface QG turbulence predictions for wavenumber spectra and their vertical dependence in non-constant stratification.

2.2 Theoretical predictions for submesoscale wavenumber spectra

We here review theories that have been proposed to describe submesoscale turbulence. Two variants of geostrophic-turbulence theory have been put forward: interior QG turbulence (Charney, 1971), resulting from deep baroclinic instabilities, and surface QG turbulence (Blumen, 1978), developing in response to frontogenesis. That discussion (e.g. Stammer, 1997; Le Traon et al., 2008; Wang et al., 2010), however, has largely ignored the internal-wave field emphasized in earlier works (e.g. Katz, 1973). The internal-wave continuum, as well as tides, inertial oscillations, and mixed-layer flows, may contribute significantly to submesoscale variability. A substantial theory only exists for the internal-wave continuum. The other unbalanced flows nevertheless have certain characteristics that can be used to assess their relative importance.

2.2.1 Geostrophic turbulence

The theories for submesoscale geostrophic turbulence assume that energy is extracted from the mean flow through baroclinic instability at scales close to the deformation scale, which in the midlatitude ocean generally lies at the large-scale end of what we call the submesoscale range. It is further assumed that dissipation is negligible throughout the submesoscale range, so that nonlinear interactions between modes transfer energy and

enstrophy across a submesoscale inertial range. If the turbulence is horizontally homogeneous and isotropic, Kolmogorov-like dimensional arguments can be used to deduce spectral slopes of horizontal kinetic energy $\mathcal{K} = \frac{1}{2}\langle u^2 + v^2 \rangle$ and potential energy $\mathcal{P} = \frac{1}{2}\langle b^2 \rangle / N^2$ in wavenumber space, where u and v are a pair of orthogonal horizontal velocity components, b is buoyancy, N is the background stratification, and $\langle \cdot \rangle$ is the two-dimensional spatial average. The two-dimensional isotropic spectra will be denoted by $\mathcal{K}_{k_h}$ and $\mathcal{P}_{k_h}$, where $k_h = (k^2 + l^2)^{1/2}$ is the magnitude of the horizontal wavenumber. From the spectral slope of kinetic energy, the spectral slope of the variance of a passive tracer $\mathcal{T} = \frac{1}{2}\langle c^2 \rangle$ can also be inferred, where c denotes the tracer concentration. Because the motion is quasi-geostrophic and therefore along isopycnals, all these spectra are to be understood as along isopycnals.

Any QG flow can be decomposed into a component due to interior PV anomalies and a component due to surface buoyancy anomalies (e.g. Lapeyre and Klein, 2006). Charney (1971) considered QG flows due to interior PV anomalies that are far enough from boundaries that the effect of the surface buoyancy anomalies can be ignored. In the submesoscale inertial range, potential enstrophy is then cascaded down to small scales. It follows that the spectra of kinetic and potential energy scale like[4] $\mathcal{K}_{k_h} \sim \mathcal{P}_{k_h} \sim k_h^{-3}$.

Interior QG turbulence scalings, as already noted by Charney, do not apply at and near the ocean surface. As a complement, Blumen (1978) considered a flow that has a uniform interior PV and is thus entirely associated with surface buoyancy anomalies. This limit is commonly referred to as surface QG. In an ocean with constant stratification, the surface spectra of kinetic and potential energy in the submesoscale inertial range are relatively flat: $\mathcal{K}_{k_h} \sim \mathcal{P}_{k_h} \sim k_h^{-5/3}$.

In constant stratification, surface QG motions of horizontal scale $2\pi/k_h$ have a vertical decay scale f/Nk_h. Modes with horizontal scales smaller than $-Nz/f$ have decayed significantly at a given depth z, while larger-scale modes are essentially unattenuated. Surface QG turbulence thus predicts kinetic- and potential energy spectra to follow the surface spectra above this transition scale and rapidly decay below (see our Fig. 2.3, described in

---

[4]Charney gave the three-dimensional isotropic spectra $\mathcal{K}_K$ and $\mathcal{P}_K$, where $K = (k^2 + l^2 + f^2 m^2 / N^2)^{1/2}$ is the magnitude of the three-dimensional wavenumber with the vertical wavenumber m rescaled by f/N. A three-dimensional isotropic spectrum $\mathcal{S}_K$ is related to the two-dimensional isotropic spectrum $\mathcal{S}_{k_h}$ by (cf., Batchelor, 1953)

$$\mathcal{S}_{k_h} = k_h \int_{k_h}^{\infty} \frac{\mathcal{S}_K}{K(K^2 - k_h^2)^{1/2}} \, dK.$$

If the three-dimensional spectrum locally follows the power law $\mathcal{S}_K = AK^{-n}$, $n > -1$, and has a slope of less than $+1$ at higher wavenumbers, then the integral above is dominated by low wavenumbers:

$$\mathcal{S}_{k_h} \approx Ak_h \int_{k_h}^{\infty} \frac{K^{-n}}{K(K^2 - k_h^2)^{1/2}} \, dK = A'k_h^{-n} \quad \text{with} \quad A' = A \int_1^{\infty} \frac{\kappa^{-(n+1)}}{(\kappa^2 - 1)^{1/2}} \, d\kappa.$$

In this case, the two-dimensional spectrum follows the same power law as the three-dimensional one. The spectral slopes given by Charney therefore directly translate to the two-dimensional spectra.
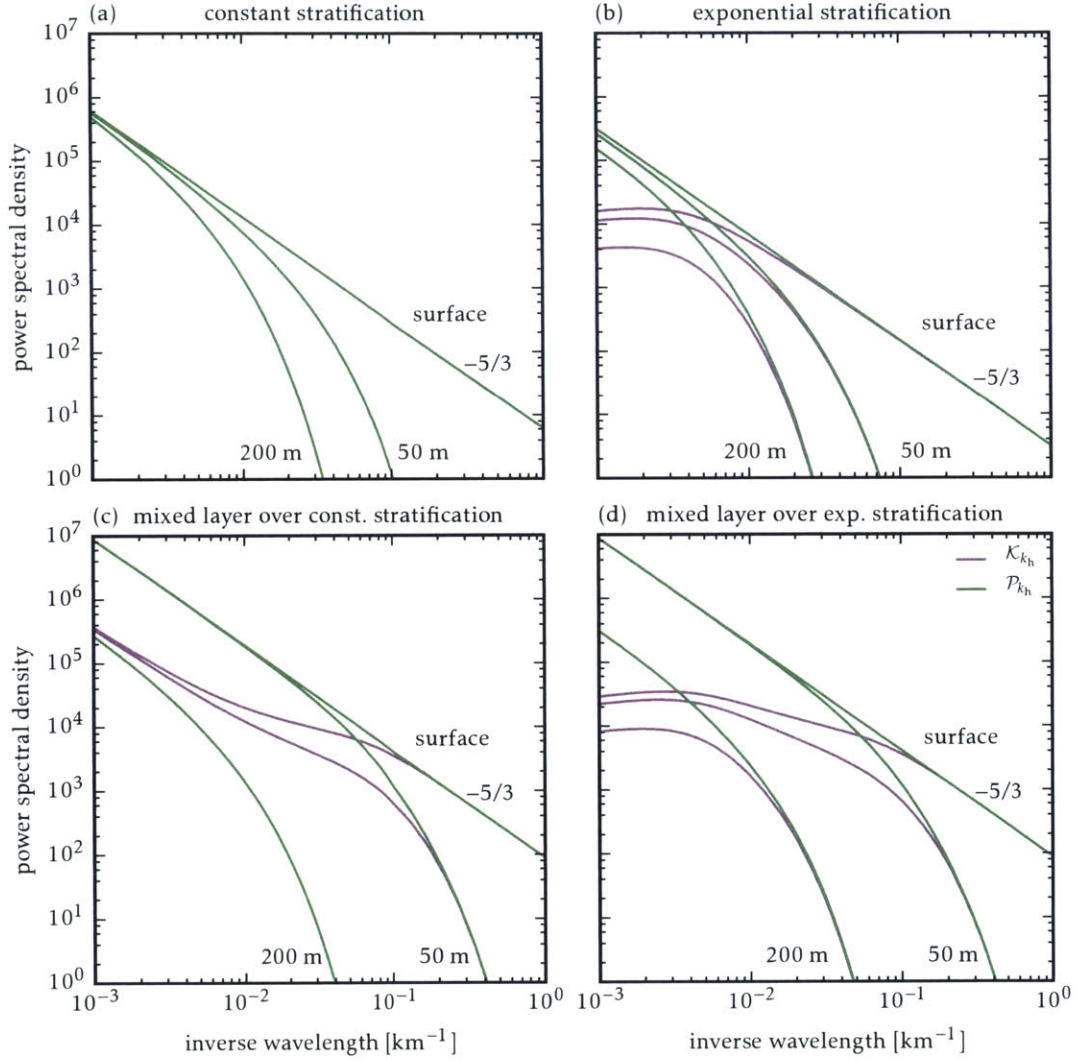
Figure 2.3: Surface QG turbulence predictions of two-dimensional isotropic wavenumber spectra of kinetic and potential energies at the surface and at depths 50 m and 200 m for (a) constant stratification, (b) exponential stratification, (c) a mixed layer of depth 100 m overlying constant stratification, and (d) a mixed layer of depth 100 m overlying exponential stratification

greater detail below, and Scott, 2006). At any given depth, we might expect surface QG turbulence to prevail at large scales, but interior PV anomalies to become important at small scales—with the transition scale depending on depth. This prediction was found to be consistent with idealized primitive-equation model simulations (Klein et al., 2008).

In an ocean with non-constant stratification, the surface QG turbulence predictions for energy spectra significantly differ from the case of constant stratification. In exponential stratification, the kinetic energy spectra are flatter at scales larger than the deformation scale Nd/f associated with the depth scale d of stratification (see Fig. 2.3 and the appendix). An overlying mixed layer further changes the predictions, leading to a flattening of the kinetic energy spectra at scales larger than the deformation scale $N_{ML}h/f$ associated with the depth of the mixed layer h and to drastically decreased energy levels in the thermocline compared to the mixed layer (Fig. 2.3, see Appendix).

In turbulent inertial ranges as described above, there is a relationship between the kinetic energy spectrum and the variance spectrum of a conservative passive tracer stirred by the turbulence (e.g. Vallis, 2006). If the kinetic energy spectrum rolls off as $\mathcal{K}_{k_h} \sim k_h^{-n}$, $n < 3$, the associated passive tracer spectrum rolls off as $\mathcal{T}_{k_h} \sim k_h^{(n-5)/2}$. For kinetic energy spectra steeper than -3, the tracer stirring is non-local and a Batchelor spectrum $\mathcal{T}_{k_h} \sim k_h^{-1}$ emerges. In interior QG turbulence, where the kinetic energy spectrum rolls off as $\mathcal{K}_{k_h} \sim k_h^{-3}$, the passive tracer thus undergoes strong filamentation and develops a spectrum $\mathcal{T}_{k_h} \sim k_h^{-1}$. In surface QG turbulence, on the other hand, where the surface kinetic energy spectrum is $\mathcal{K}_{k_h} \sim k_h^{-5/3}$ or flatter, the more energetic smaller-scale eddies mix the tracer more effectively at smaller scales and the variance spectrum is correspondingly steeper: $\mathcal{T}_{k_h} \sim k_h^{-5/3}$ or steeper. At depth, however, the variance spectrum develops a Batchelor spectrum $\mathcal{T}(k_h) \sim k_h^{-1}$ at the smaller submesoscales, where the surface modes have decayed at the given depth (Scott, 2006).

The surface QG framework is useful to study geostrophic turbulence in the upper ocean. The scaling laws predicted by surface QG turbulence theory, however, need to be revised to account for important ageostrophic corrections (Capet et al., 2008a). The flat kinetic- and potential energy spectra in surface QG turbulence are indicative of the creation of strong horizontal shears and sharp buoyancy gradients, the prelude to frontogenesis (Hoskins and Bretherton, 1972). If one includes advection by the ageostrophic part of the velocity in Blumen's idealized flow (a term ignored in surface QG theory), discontinuities form in both the velocity and buoyancy fields, and the energy spectra become $\mathcal{K}_{k_h} \sim \mathcal{P}_{k_h} \sim k_h^{-2}$ (Boyd, 1992). If a passive tracer is stirred by a discontinuous flow field, it also develops discontinuities and its variance spectrum consequently scales like $\mathcal{T}_{k_h} \sim k_h^{-2}$.

Beyond the changes in spectral slopes, frontogenesis and ageostrophic motion can significantly change the energy cascade. Molemaker et al. (2010) show that ageostrophic flows are responsible for reversing the direction of the energy cascade of a surface QG-like flow from upscale in the QG limit to downscale in a Boussinesq system that allows frontogenesis.

2.2.2 The internal-wave continuum

Starting with the seminal work of Garrett and Munk (1972), it was realized that internal waves are characterized by a continuum spectrum spanning frequencies from f to N, over which a few spectral peaks due to tides and inertial motions are superimposed. The internal-wave continuum spectrum is understood to be set through weak interactions of linear waves (e.g. Lvov et al., 2004). It is remarkably uniform across the ocean and has been shown to be successfully described by the empirical model spectra of Garrett and Munk (GM; e.g. Munk, 1981). The GM spectra predict the spectral densities of the kinetic and potential energies as functions of frequency and vertical wavenumber; they can be converted to horizontal-wavenumber space using the dispersion relation of internal waves and summing over all vertical wavenumbers (e.g. Klymak and Moum, 2007). These wavenumber spectra scale like $\mathcal{K}_{k_h} \sim \mathcal{P}_{k_h} \sim k_h^{-2}$ in the short-wave limit and flatten out at scales larger than about 10 km, as shown, for example, by the dotted purple line in Fig. 2.5.

The GM spectra can also be used to deduce the passive-tracer variance generated by the internal-wave continuum, on both horizontal and isopycnal surfaces. For tracers with vertical gradients much larger than isopycnal gradients, like temperature or salinity, the tracer variance on horizontal surfaces is dominated by distortions of the vertical background gradient. We can then estimate $\hat{c} = (\partial C/\partial z)\hat{\zeta} = (\partial C/\partial z)\hat{b}/N^2$, so that $\mathcal{T} = \frac{1}{2}(\partial C/\partial z)^2 \mathcal{P}/N^2$, where C is the background tracer field and $\zeta = b/N^2$ is the vertical displacement due to the wave. Along isopycnals, this variance is filtered out and we are left with displacements of the much weaker isopycnal gradient. The associated variance can be estimated analogously by considering the isopycnal wave-induced displacements and is typically several orders of magnitude smaller. Both vertical and along-isopycnal displacements generate a tracer spectrum that scales like $\mathcal{T}_{k_h} \sim k_h^{-2}$ in the short-wave limit and flattens out at scales larger than about 10 km.

2.2.3 Tides

Barotropic tides are generated by astronomical forces; baroclinic (internal) tides are forced by barotropic tides flowing over topography. These baroclinic tides are not included in the GM spectra and the associated kinetic energy varies considerably between geographical locations (e.g. Arbic et al., 2012).

Both barotropic and baroclinic tides have a specific frequency that is set by the forcing, so that linear theory can be used to deduce the horizontal wavenumbers baroclinic tides project onto (e.g. Dushaw et al., 1995). Altimetric observations and high-resolution simulations show that baroclinic tides appear as somewhat broadened peaks in wavenumber spectra, to the extent that they may be better described as broadband signals (e.g. Ray and Mitchum, 1997; Ray and Zaron, 2011; Zhao et al., 2012). A theory that explains this

signature in wavenumber spectra is lacking and in-situ observations are difficult because of the tides' large spatial scales and high frequencies.

2.2.4 Inertial oscillations

The ocean is nearly resonant at the inertial frequency. Variable winds therefore easily excite inertial oscillations in the upper ocean, particularly in the mixed layer (e.g. Pollard, 1970). The oscillations are quasi-two-dimensional and have the frequency of the Coriolis parameter f. The internal-wave dispersion relation for waves of frequency f predicts that the waves have an infinite horizontal scale. Some spreading in frequency results from changes in f with latitude and from the background vorticity. As a result, the inertial oscillations are characterized by a broad peak in the frequency spectrum (e.g. Garrett, 2001) and have large horizontal scales, but a theory predicting their signature in wavenumber spectra does not exist. The quasi-two-dimensional nature of inertial oscillations, however, implies a large kinetic to potential energy ratio. This is not what we find in the observations, as discussed in Section 2.6.

2.2.5 Mixed-layer flows

In addition to superinertial waves, the wind forces subinertial variability in the upper ocean. If the wind field is spatially variable, it transfers its variability to the Ekman currents. The resultant wavenumber spectrum depends on the wavenumber spectrum of the winds, which is reddened in the transfer (Frankignoul and Müller, 1979). Additional spatial variability is generated by the vorticity of balanced flows (e.g. Niiler, 1969).

Other sources of surface-intensified turbulence are mixed-layer instabilities. Both symmetric instability (e.g. Emanuel, 1994) and mixed-layer baroclinic instability (e.g. Blumen, 1979) generate turbulent flows in the mixed layer (Haine and Marshall, 1998). The most baroclinically unstable mode has a horizontal scale of order 1–10 km (Boccaletti et al., 2007). Theory (Fox-Kemper et al., 2008), modeling (e.g. Capet et al., 2008e), and observations (Shcherbina et al., 2013) indicate that these mixed layer instabilities can substantially enhance submesoscale energy in the presence of a deep winter mixed layer. Such mixed layer flows have energy spectra close to k_h^{-2}. The dynamics are distinct from surface QG and are better described as Eady-like, with eddies confined between the surface and the strong PV gradient at the mixed layer base.

The mixed layer also features small-scale wind- and buoyancy-driven turbulence (e.g. Kantha and Clayson, 2000). These flows typically have aspect ratios close to one and scales that are generally smaller than what we defined as the submesoscale range.

All these flows are strong in the mixed layer and decay rapidly below. As discussed in Sections 2.5 and 2.6, this is not what we find in the observations: the energy levels in

27

mixed layer and thermocline are very similar.

2.3 Relating two-dimensional spectral theories to one-dimensional spectral observations

All predictions reviewed above are for two-dimensional isotropic spectra. Observations, however, are more typically taken along a one-dimensional track. In such cases, only one-dimensional spectra can be computed as functions of the longitudinal (along-track)[5] wavenumber, say k. The conversion of a two-dimensional isotropic spectrum S_{k_h} to a one-dimensional spectrum S_k is given by (cf., Batchelor, 1953)

$$S_k = \frac{2}{\pi} \int_k^\infty \frac{S_{k_h}}{(k_h^2 - k^2)^{1/2}} \, dk_h. \tag{2.1}$$

If the two-dimensional spectrum locally follows a power law $S_{k_h} \sim k_h^{-n}$, $n > 0$, and remains red at higher wavenumbers, then the one-dimensional spectrum follows the same power law[6] $S_k \sim k^{-n}$. The predicted scalings given above thus directly translate to the one-dimensional spectra of kinetic energy, potential energy, and tracer variance, summarized in Table 2.1.

It is important to realize, however, that for isotropic flow, the one-dimensional spectra of the longitudinal (along-track) and transverse (across-track) components of kinetic energy, $\mathcal{K}^L = \frac{1}{2}\langle u^2 \rangle$ and $\mathcal{K}^T = \frac{1}{2}\langle v^2 \rangle$, are not necessarily equal (u and v are the longitudinal and transverse horizontal velocity components, respectively). Instead, for an isotropic two-dimensional (meaning horizontally nondivergent) flow, they satisfy[7]

$$\mathcal{K}_k^T = -k \frac{d}{dk} \mathcal{K}_k^L, \tag{2.2}$$

so that for a spectrum $\mathcal{K}_{k_h} \sim k_h^{-n}$, the relation between transverse and longitudinal spectra is $\mathcal{K}_k^T = n\mathcal{K}_k^L$ (e.g. Batchelor, 1953; Leith, 1971; Charney, 1971).

---

[5]We use *longitudinal* exclusively in this sense and never in the sense of *zonal*.

[6]The integral in (2.1) is then dominated by low wavenumbers, where $S_{k_h} = Ak_h^{-n}$, and can be approximated by

$$S_k \approx \frac{2}{\pi} \int_k^\infty \frac{Ak_h^{-n}}{(k_h^2 - k^2)^{1/2}} \, dk_h = A'k^{-n} \quad \text{with} \quad A' = \frac{2}{\pi} A \int_1^\infty \frac{\kappa^{-n}}{(\kappa^2 - 1)^{1/2}} \, d\kappa.$$

[7]They are related to the two-dimensional isotropic spectrum $\mathcal{K}_{k_h}$ by

$$\mathcal{K}_k^L = \frac{2}{\pi} \int_k^\infty \frac{(k_h^2 - k^2)^{1/2} \mathcal{K}_{k_h}}{k_h^2} \, dk_h,$$

$$\mathcal{K}_k^T = \frac{2}{\pi} k^2 \int_k^\infty \frac{\mathcal{K}_{k_h}}{k_h^2 (k_h^2 - k^2)^{1/2}} \, dk_h.$$

Table 2.1: Properties of one-dimensional wavenumber spectra ($\mathcal{K}_k^{\mathrm{T}}$ and $\mathcal{K}_k^{\mathrm{L}}$ are the transverse and longitudinal kinetic energy spectra, $\mathcal{P}_k$ is the potential energy spectrum, $\mathcal{T}_k$ is the variance spectrum of a passive tracer)

|	QG turbulence			Internal-wave continuum
	Interior	Surface (constant N)	Surface (non-constant N)	
$\mathcal{K}_k^{\mathrm{T}}$ slope	-3	$-5/3$ (at surface; steep below, see Fig. 2.3)	(see Fig. 2.4)	-2 (in small-scale roll-off)
Relation between $\mathcal{K}_k^{\mathrm{T}}$ and $\mathcal{K}_k^{\mathrm{L}}$	$\mathcal{K}_k^{\mathrm{T}} = 3\mathcal{K}_k^{\mathrm{L}}$	$\mathcal{K}_k^{\mathrm{T}} = 5/3\mathcal{K}_k^{\mathrm{L}}$ (at surface)	(see Fig. 2.4)	$\mathcal{K}_k^{\mathrm{T}} = \mathcal{K}_k^{\mathrm{L}}$
Relation between $\mathcal{P}_k$ and $\mathcal{K}_k^{\mathrm{T}}$, $\mathcal{K}_k^{\mathrm{L}}$	$\mathcal{P}_k = \mathcal{K}_k^{\mathrm{L}}$	$\mathcal{P}_k = \mathcal{K}_k^{\mathrm{L}} + \mathcal{K}_k^{\mathrm{T}}$	(see Fig. 2.4)	(see Fig. 2.6)
$\mathcal{T}_k$ slope	-1	$-5/3$ (at the surface; flatter below)	(steep at the surface, flat below)	-2 (in short-wave roll-off, weak stirring along isopycnals)
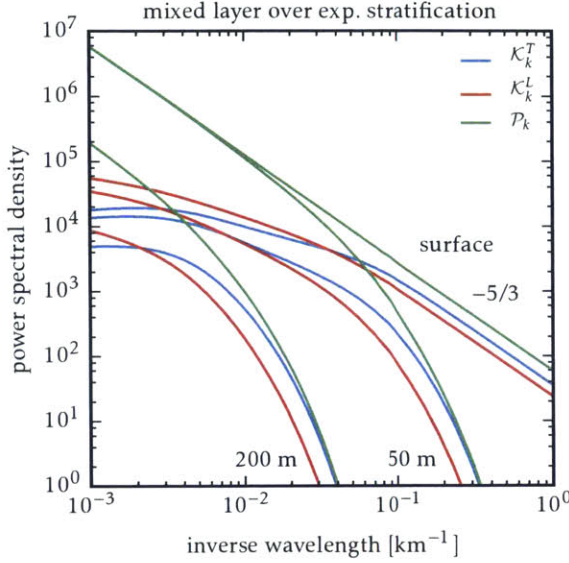
Figure 2.4: Surface QG turbulence predictions of one-dimensional wavenumber spectra of longitudinal and transverse kinetic energy and potential energy at the surface and at depths 50 m and 200 m for a mixed layer of depth 100 m overlying exponential stratification

Relation (2.2) is particularly relevant for geostrophic turbulence, because those flows are horizontally nondivergent at leading order. In addition, QG buoyancy is related to the vertical gradient of the streamfunction and therefore the one-dimensional potential energy spectrum is also related to the components of the one-dimensional kinetic energy spectrum. In interior QG turbulence, the potential energy spectrum equals the longitudinal component of the kinetic energy spectrum, $\mathcal{P}_k = \mathcal{K}_k{}^L$ (Charney, 1971), whereas in surface QG turbulence it equals the total kinetic energy spectrum, $\mathcal{P}_k = \mathcal{K}_k{}^L + \mathcal{K}_k{}^T$, in the case of constant stratification. For more realistic stratification, surface QG turbulence predicts the kinetic-energy levels to be lower than the potential energy levels at scales larger than the deformation scale (Fig. 2.4, see Appendix).

In an internal-wave field, the flow is three-dimensional[8]—so (2.2) does not hold in general. If the wave field is horizontally isotropic, as assumed in Garrett and Munk (1972), we instead expect the one-dimensional longitudinal and transverse kinetic energy spectra to be equal, $\mathcal{K}_k{}^L = \mathcal{K}_k{}^T$. The same is true for an isotropic internal-tide field.

Ageostrophic circulations at fronts, Ekman flows, and mixed-layer turbulence are all

---

[8]For motions of frequency $f < \omega < N$, it follows from the WKB solution of the linear Boussinesq equations that there is significant vertical divergence: the ratio of the sum of the magnitudes of the terms of the horizontal divergence, $\partial u/\partial x$ and $\partial v/\partial y$, to the magnitude of the vertical divergence, $\partial w/\partial z$, satisfies

$$\frac{1}{2}\frac{\omega^2 + f^2}{\omega^2} \leq \frac{k^2|\hat{u}|^2 + l^2|\hat{v}|^2}{m^2|\hat{w}|^2} \leq 1.$$

It must therefore be of order unity, which implies that $|\partial w/\partial z|$ is of the same order as $|\partial u/\partial x|$ and $|\partial v/\partial y|$.

characterized by strong vertical velocities and far from two-dimensional. Hence, (2.2) is not expected to hold for these flows. Due to anisotropies, however, $\mathcal{K}_k{}^T \neq \mathcal{K}_k{}^L$ is likely.

The main point here is that considering the longitudinal and transverse components of the kinetic energy spectrum separately yields additional ways to discriminate between the different regimes of geostrophic turbulence and between two- and three-dimensional motions. These relationships are particularly useful because the spectral slopes are often quite similar for the different regimes.

2.4   Review of previous observations of submesoscale spectra

2.4.1   In-situ observations

Early observations of wavenumber spectra of potential energy in the submesoscale range were often interpreted in terms of internal-wave variability. Katz (1973) reported on tow experiments in the Sargasso Sea thermocline and showed that spectra on scales between about 50 km and 30 m scaled close to $\mathcal{P}_k \sim k^{-2}$. He discussed their relation to the GM model spectra, which were later modified to be consistent with these observation (Garrett and Munk, 1975).

After the Mid-Ocean Dynamics Experiment and the discovery of the mesoscale, sub-mesoscale wavenumber spectra have increasingly been interpreted as manifestations of geostrophic turbulence. Samelson and Paulson (1988) presented wavenumber spectra from the subtropical North Pacific and found that potential energy spectra scaled close to $\mathcal{P}_k \sim k^{-2}$ at scales 10–100 km and close to $\mathcal{P}_k \sim k^{-3}$ at scales 1–10 km in the mixed layer (15 m depth); in the thermocline (70 m depth), they found $\mathcal{P}_k \sim k^{-2}$ at scales 1–100 km. They concluded that the spectral slopes are consistent with interior QG turbulence dominating in the mixed layer and internal-wave variability dominating in the thermocline. The steep spectrum in the mixed layer, however, was not confirmed by an experiment conducted a few degrees further east and on an even wider range of scales (Ferrari and Rudnick, 2000). The potential energy spectra were instead reported to scale close to $\mathcal{P}_k \sim k^{-2}$ in both mixed layer and thermocline.

Wang et al. (2010) analyzed a long-term data set of mixed-layer velocities and temperatures in the Gulf Stream region, along a transect between New York Harbor and Bermuda. They found that the energy spectra fall off approximately like $\mathcal{K}_k \sim \mathcal{P}_k \sim k^{-3}$ on scales between 250 km and 20 km, consistent with interior QG turbulence. They also reported that the power spectra of the zonal and meridional velocities are similar and interpreted this result as evidence of horizontal isotropy. The similarity between the power spectra of kinetic and potential energies was further interpreted as a signature of energy equipartition. As pointed out in the theoretical review, however, a more stringent test of whether the flow is two-dimensional and isotropic relies on the comparison of spectra of

the longitudinal and transverse components of kinetic energy—a test we will explore in this chapter. Furthermore, interior QG turbulence predicts an equipartition between the longitudinal component of kinetic energy and potential energy, $\mathcal{K}_k^L = \mathcal{P}_k$, not between the total kinetic energy $\mathcal{K}_k$ and potential energy $\mathcal{P}_k$.

Wang et al. (2010) ignored scales below 20 km and measurements below the base of the mixed layer, where internal waves are expected to become dominant. A major goal of this chapter will be to provide a comprehensive analysis of spectra in the mixed layer and below on scales from a few hundred kilometers down to 1 km (resolution permitting). Comparing different depths and scales allows a more robust assessment of what dynamics dominate the submesoscale range in different places.

2.4.2 Altimetry observations

The advent of satellite altimetry sparked numerous investigations of surface kinetic energy spectra. The TOPEX/POSEIDON[9] and Jason-1/2 altimeters measured the sea surface height (SSH) along tracks spanning the global oceans from 65°N to 65°S. After correcting for instrument and atmospheric noise and removing the contribution of barotropic tides, geostrophic balance is routinely applied to convert the SSH measurements to transverse geostrophic velocities at the sea surface, from which the transverse component of the geostrophic surface kinetic energy spectrum can be estimated. Stammer (1997) found this spectrum to fall off like $\mathcal{K}_k^T \sim k^{-3}$ throughout the extratropical oceans, consistent with interior QG turbulence. Le Traon et al. (2008) instead found the spectra in major current regions to be closer to $\mathcal{K}_k^T \sim k^{-5/3}$ and concluded that surface QG turbulence is the more relevant dynamical framework.

Xu and Fu (2011, 2012) used Jason-1/2 along-track data to create a global map of spectral slopes of surface kinetic energy on scales between 250 km and 70 km. They found large geographical differences with relatively steep spectra in the major current regions (between k^{-2} and $k^{-2.5}$) and flatter spectra in the rest of the extratropical oceans (between k^0 and $k^{-1.5}$). These slopes are everywhere smaller than those inferred from in-situ observations. One possible explanation is that Xu and Fu's estimates, despite the attempt to remove instrumental noise, are still contaminated by noise at scales that are only marginally resolved by altimeters. This was corroborated by the modeling results of Sasaki and Klein (2012) and Richman et al. (2012), who found steeper spectra than the altimetric ones in regions of low mesoscale activity. A second possibility is that ageostrophic flows, like internal tides, have an important signature in SSH, so that velocities inferred from SSH using geostrophic balance are not accurate (Richman et al., 2012). A third possibility is that in-situ observations are dominated by ageostrophic flows not measured by

---

[9]Acronyms for "Ocean Topography Experiment" and "Premier Observatoire Spatial Étude Intensive Dynamique Océan et Novosphère" (sic) or "Positioning Ocean Solid Earth Ice Dynamics Orbiting Navigator" (Wunsch and Stammer, 1998)

altimeters.

In summary, the surface kinetic energy spectra inferred from altimetry, which only account for the geostrophic part of the flow (possibly contaminated by ageostrophic flow that projects onto SSH), are steeper in major current regions than in quiescent gyres. Spectra inferred from in-situ measurements, which account for the full geostrophic plus ageostrophic flow, confirm this tendency, but the actual slopes differ substantially. Here we try to unravel what dynamics dominate the submesoscale in different geographical regions and at different depths, analyzing in-situ data with a broad set of diagnostics: (i) the longitudinal and transverse components of kinetic energy spectra, (ii) the potential energy spectra, (iii) the variance spectra of passive tracers, and (iv) the changes of the spectra with depth.

2.5 Gulf Stream region

2.5.1 Oleander data set

We revisit the Oleander data set[10], previously analyzed by Wang et al. (2010), to provide a more thorough test of whether the dynamics are consistent with geostrophic-turbulence theories. Furthermore, we investigate whether a transition to an internal-wave regime occurs at scales below 20 km that were not analyzed in Wang et al. (2010).

The ADCP data were collected on repeat transects from New York Harbor to Bermuda (Fig. 2.2) between 1994 and 2009. We use data at 39 m depth, which is within the mixed layer, and at 150 m depth, which is below the base of the mixed layer except during a few occasions of strong convection in winter. We do not choose a deeper level to retain good statistics; data gaps become more frequent at greater depths. We make the common fast-tow approximation that assumes the fields to be frozen in time over the course of a transect. It takes 20 min to cover 10 km and 3.5 hours to cover 100 km at a ship speed of about 8 m s^{-1}. Geostrophic eddies have generally longer time scales (e.g. Stammer, 1997) and are well-resolved. Internal waves have periods larger than $2\pi/N = 20$ min and this high-frequency component projects onto small horizontal scales. Waves with larger horizontal scales have larger periods, so that the frozen-field approximation is also well-justified for internal waves (cf., Garrett and Munk, 1972). The approximation becomes problematic for tidal flows with large horizontal scales, for which aliasing will occur.

We select transects at least 1000 km long and with data gaps no wider than 10 km (311 transects in the mixed layer, 264 in the thermocline) and use cubic splines to interpolate onto a 2.5 km regular grid, which is about the spacing of the original data. Even though the presence of the Gulf Stream renders the velocity field statistically inhomogeneous, we use the full transects; discarding the part of the transects that includes the Gulf Stream

---

[10]The data are available at http://po.msrc.sunysb.edu/Oleander/.

does not give qualitatively different results. We also disregard seasonal variations—the qualitative characteristics of the spectra are independent of season in this data set. More recent data from the same transect, but collected with a different instrument, does show a seasonal cycle with more energetic small scales in the winter mixed layer. The results in this chapter should therefore be regarded as representative of times with no deep mixed layer. We rotate the velocities into the reference frame of the ship track to separate the longitudinal and transverse components. We divide the transects into three segments with a 50% overlap, apply a Hann window, and calculate the discrete Fourier transform of each segment. We average the resulting Fourier transforms over the three segments and ten wavenumber bins per decade to obtain the spectra. Because of the large number of transects, formal error bars are very small.

For comparison, we use altimetric measurements of the SSH anomaly as distributed by AVISO (multimission along-track delayed-time product[11]). The transverse surface velocity is calculated from the along-track SSH gradient assuming geostrophic balance. The power spectra of the transverse velocity presented below are averages over all track segments of the arcs 50 and 126 that lie in a $7° \times 8°$ box centered at 68.5°W, 35°N (years 1993–2002). These track segments nearly parallel the Oleander ship tracks (Fig. 2.2).

2.5.2 Analysis

We analyze kinetic energy spectra at two depths: one in the mixed layer and one in the thermocline. Kinetic energy levels in the mixed layer are slightly higher than in the thermocline (Fig. 2.5). At scales between 200 km and 20 km, both the longitudinal and transverse spectra (solid red and blue in Fig. 2.5) fall off steeply, with a slope close to -3, both in the mixed layer and the thermocline. The transverse spectra are about three times larger than the longitudinal ones, which is consistent with a two-dimensional isotropic flow, because for $\mathcal{K}_k^L \sim k^{-3}$, relation (2.2) yields $\mathcal{K}_k^T = 3\mathcal{K}_k^L$. These characteristics—the slope and the relation between longitudinal and transverse spectra—are consistent with the theoretical predictions for homogeneous, isotropic interior QG turbulence reviewed in Section 2.2. The slight decay of energy levels with depth is consistent with deep vertical modes. Tulloch et al. (2011) found that the scale of the most unstable baroclinic mode in this region is considerably larger than the deformation scale of about 150 km, consistent with the -3 slope extending to scales as large as 200 km.

At scales larger than 100 km, the transverse kinetic energy spectrum in the mixed layer matches the one derived from altimetry (dashed blue in Fig. 2.5). Below 100 km, the altimetric spectrum rolls off much more steeply than the in-situ spectrum, but this range of scales is only marginally resolved by altimeters and strongly affected by the

---

[11]The altimeter product was produced by SSALTO/DUACS and distributed by AVISO, with support from CNES (http://www.aviso.oceanobs.com/duacs/).
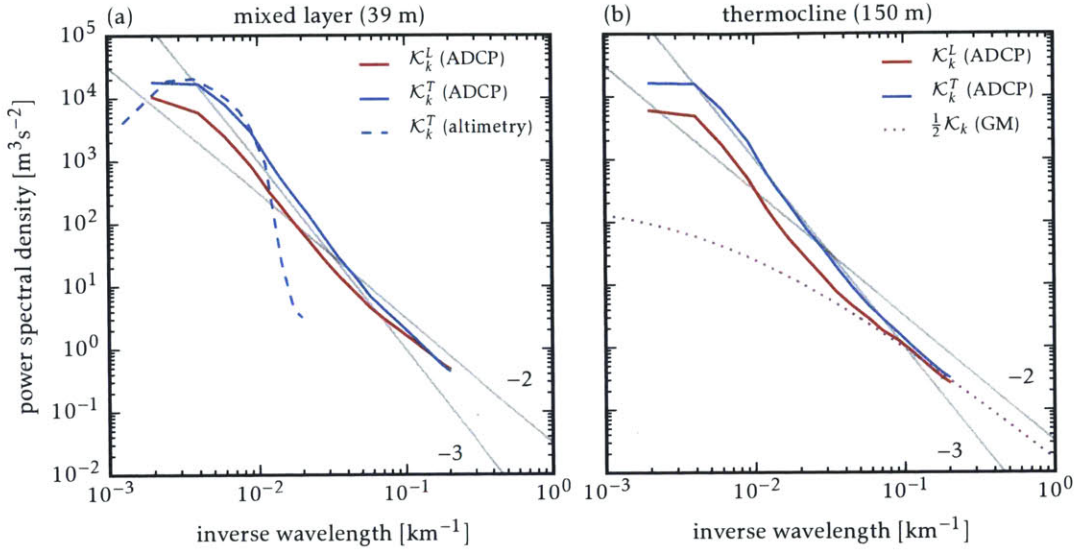
Figure 2.5: Gulf Stream region wavenumber spectra of longitudinal and transverse kinetic energies $\mathcal{K}_k{}^L$ and $\mathcal{K}_k{}^T$ in (a) the mixed layer (39 m depth) and (b) the thermocline (150 m depth) from in-situ observations (ADCP); wavenumber spectrum of surface transverse kinetic energy $\mathcal{K}_k{}^T$ from altimetry; Garrett–Munk (GM) model spectrum for kinetic energy $\mathcal{K}_k$ in the thermocline; lines with slopes -2 and -3 for reference (gray solid lines). Confidence intervals are too small to be visible.

smoothing filter applied to remove noise. The match of the in-situ and the altimetric spectra is consistent with the flow being predominantly geostrophic and further supports the interpretation in terms of geostrophic turbulence.

In the thermocline, at scales below 20 km, the longitudinal and transverse spectra flatten out, converge, and approach the GM spectrum for a stratification $N = 6 \times 10^{-3}$ s^{-1} taken from the ARGO climatology by Roemmich and Gilson (2009) in this region (dotted purple in Fig. 2.5). This convergence and the fact that $\mathcal{K}_k{}^L = \mathcal{K}_k{}^T$ are consistent with internal-wave dynamics and inconsistent with two-dimensional isotropic flow, suggesting that internal waves dominate the kinetic energy at scales below 20 km. The exact match with the GM spectrum may be fortuitous, however, because the interpolation onto a regular grid slightly reduces the spectra at these high wavenumbers, so that they ought to drop below the GM spectrum if only internal-waves were present. A small balanced component may explain the slightly increased spectral amplitudes. Another possible cause of the flattening of the spectra is instrumental noise, but it would be surprising if the instrumental noise just happened to match the amplitude of the internal-wave field. We conclude that the convergence of the longitudinal and transverse spectra and the transition to a flatter slope are a robust result, indicating a change in dynamics. In the mixed layer, the longitudinal and transverse spectra also flatten out and converge, with energy

levels slightly higher than in the thermocline. An enhancement of internal-wave energy in the mixed layer is consistent with the vertical structure of low-mode internal waves: thermocline modes project up and can induce increased kinetic energies in the mixed layer, despite the weak stratification (e.g. D'Asaro, 1978). It is, however, also possible that some of the excess energy is due to Ekman flows and frontal circulations.

The interpretation of the kinetic energy spectra yields a fairly straightforward picture of submesoscale dynamics in the Gulf Stream region. An interior QG turbulence regime exists at large scales and an internal-wave regime at small scales, with the transition occurring at about 20 km. We find no evidence of a surface QG turbulence regime in the upper ocean, which has previously been suggested from analysis of altimetry data from this region (e.g. Le Traon et al., 2008). Enhanced submesoscale energy in the mixed layer, however, may occur in the presence of a deep winter mixed layer, as discussed in Section 2.2.

2.6 Subtropical North Pacific

2.6.1 Spice data set

We revisit the Spice data set presented in Ferrari and Rudnick (2000) to get a better understanding of the submesoscale dynamics in view of recent theoretical developments. The data were collected with a research vessel in successive occupations of the meridian at 140°W between 25° and 35°N (Fig. 2.2) in the subtropical North Pacific during January and February 1997. We use temperature and salinity from four SeaSoar tows along this transect: a sawtooth profile between 5 and 320 dbar with a period of 12 min, two horizontal tows along the 50 and 200 dbar isobars, and a tow along the 25.5 kg m^{-3} isopycnal. Velocity measurements in the upper 300 m are available from a shipboard ADCP. The velocity measurements and the sawtooth temperature and salinity profiles are averaged into bins of 8 m in depth and about 3 km in horizontal extent. For the horizontal and isopycnal tows, temperature and salinity are averaged into 10 m bins. See Ferrari and Rudnick (2000) for more detailed information on the data and the rationale behind the scales chosen for the horizontal and vertical averaging. Again, the fast-tow approximation is made. With a ship speed of 4 m s^{-1}, this is slightly less justified than in the Oleander data set, but geostrophic eddies and the internal-wave continuum remain well-resolved. The 7 hours it takes to cover 100 km, however, are close to the dominant tidal period—large-scale tidal flows are likely aliased.

The latitude–depth section of density available from the sawtooth profile reveals that the mixed layer has a depth ranging from 100 to 150 m. The 50 dbar horizontal tow is thus within the mixed layer, whereas the 200 dbar horizontal tow is well within the thermocline. The isopycnal tow straddles around the 200 dbar tow at a depth of about
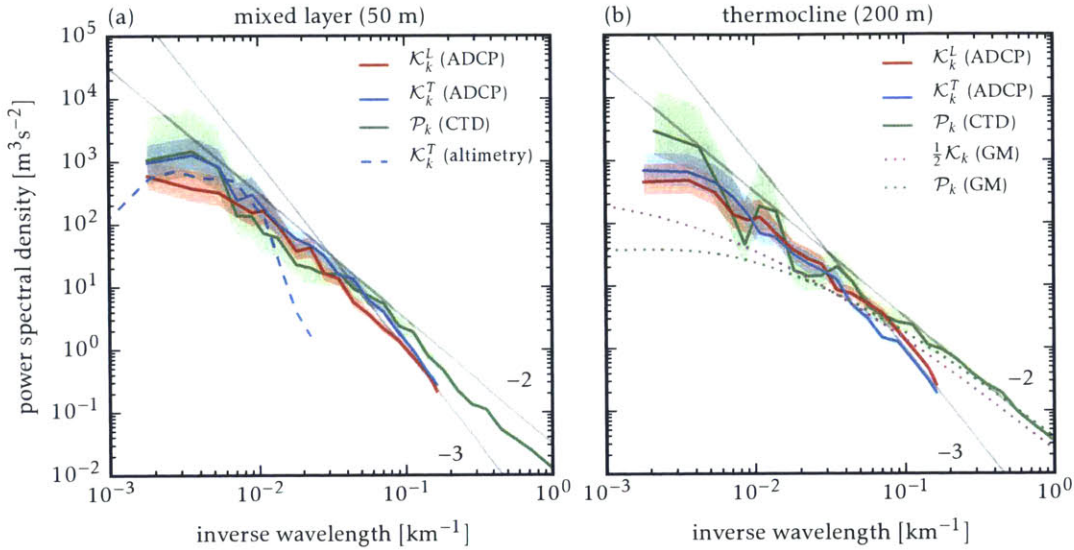
Figure 2.6: Subtropical North Pacific wavenumber spectra of longitudinal and transverse kinetic energies $\mathcal{K}_k{}^L$ and $\mathcal{K}_k{}^T$ in (a) the mixed layer (50 m depth) and (b) the thermocline (200 m depth) from in-situ observations (ADCP, CTD); wavenumber spectrum of surface transverse kinetic energy $\mathcal{K}_k{}^T$ from altimetry; Garrett–Munk (GM) model spectra for kinetic and potential energies $\mathcal{K}_k$ and $\mathcal{P}_k$ in the thermocline; lines with slopes -2 and -3 for reference (gray solid lines). The shadings show 95% confidence intervals.

150 to 250 m, that is below the mixed layer.

We again compare to AVISO along-track altimetry data. Here we use all track segments that were collected in February (years 1993–2002) and lie in a $10° \times 10°$ box centered at 140°W, 30°N, the midpoint of the ship track (arcs 19, 95, 106, 171, 182; Fig. 2.2).

2.6.2 Analysis

The observations in the subtropical North Pacific do not lend themselves to as simple a dynamical interpretation as those in the Gulf Stream region. We analyze kinetic- and potential energy spectra in the mixed layer and the thermocline. The kinetic energy spectra are computed from the ADCP data as before; the potential energy spectra are computed from the horizontal CTD tows, using averaged mixed-layer and thermocline values of the stratification N calculated from the sawtooth profile ($N = 2.3 \times 10^{-3}$ s^{-1} in the mixed layer, $N = 8.7 \times 10^{-3}$ s^{-1} in the thermocline).

The kinetic-energy levels are similar in mixed layer and thermocline (Fig. 2.6). At scales larger than 20 km, both the longitudinal and transverse components of kinetic energy (solid red and blue in Fig. 2.6) are much flatter than in the Gulf Stream region—the

slopes are close to -2.[12] The longitudinal and transverse spectra are similar throughout the observed scales, in both the mixed layer and the thermocline. With a slope of -2, relation (2.2) predicts $\mathcal{K}_k^{\mathrm{T}} = 2\mathcal{K}_k^{\mathrm{L}}$ for two-dimensional isotropic turbulence, which is not consistent with the observed spectra, but cannot be conclusively rejected because of the rather large error bars. If we interpret this as a violation of (2.2), one explanation is that the flow is anisotropic. But an additional transect orthogonal to the ones presented here shows that again transverse and longitudinal kinetic energy spectra are similar—the flow appears to be isotropic. The implication is that the near-failure of (2.2) must stem from a flow that is at least partially horizontally divergent. At scales below 20 km, the kinetic energy spectra are steeper than -2 and fall below the GM spectra (dotted purple in Fig. 2.6). This is due to the interpolation onto a regular grid: an artificial internal-wave field sampled in the same way as the observations shows the same drop at high wavenumbers.

At scales larger than 20 km, the potential energy spectra (solid green in Fig. 2.6) have similar magnitudes in the mixed layer and thermocline. At both depths, they also have similar magnitudes as the kinetic energy spectra and exhibit a slope close to -2. Energy equipartition $\mathcal{P}_k = \mathcal{K}_k^{\mathrm{L}} + \mathcal{K}_k^{\mathrm{T}}$ is only marginally satisfied within error bars; $\mathcal{K}_k^{\mathrm{L}} + \mathcal{K}_k^{\mathrm{T}}$ is consistently larger than $\mathcal{P}_k$. At scales below 20 km, the thermocline spectrum of potential energy follows the GM spectrum (dotted green in Fig. 2.6), while the mixed-layer one is significantly reduced, consistent with the interpretation that the spectrum represents internal waves: free thermocline internal-wave modes of buoyancy are evanescent in the mixed layer and must satisfy a zero boundary condition at the surface (rigid lid), so they must decay in the mixed layer. This is in contrast to the free modes of horizontal velocities that, as mentioned earlier, can be enhanced in the mixed layer, because their surface boundary condition is zero shear.

At scales larger than 100 km, the in-situ spectrum of transverse kinetic energy has the same magnitude as the altimetric spectrum (dashed blue in Fig. 2.6), with slightly more energy in the in-situ spectrum. This may be interpreted as evidence that a significant fraction of the flow is in geostrophic balance, but there is also a significant ageostrophic component. This interpretation, however, ignores the possibility that the altimetric spectrum may be contaminated by ageostrophic flow that projects onto SSH.

The analysis so far does not quite support the inference that motions at scales larger than 20 km are geostrophically balanced. The large error bars, however, preclude definite conclusions. It would still be plausible to interpret the -2 slopes of the kinetic and po-

---

[12]More precisely: Both -2 and $-5/3$ slopes are consistent with the observations to within error bars. A slope of -2 gives a better fit over a wider range of scales, but we do not attempt to distinguish between these two slopes, because this distinction has little impact on the discussion to follow. In contrast, the spectra are clearly inconsistent with a slope of -3. Note, however, that the error bars presented here are formal ones, representing the random error only. Biases due to, for example, a finite tow speed, unresolved scales, and instrument drift are not accounted for. Statements about spectral slopes in what follows are to be understood with these caveats.

tential energy spectra as signatures of surface QG turbulence, as suggested by Klein et al. (2008). The −2 slopes, however, apply only to an ocean with constant stratification; surface QG turbulence theory predicts different slopes in non-constant stratification (LaCasce, 2012). In the appendix, we show that surface QG turbulence theory predicts flat kinetic energy spectra in the submesoscale range for an ocean with a weakly stratified mixed layer and a vertically decaying stratification in the thermocline. In addition, the theory predicts that surface QG modes with scales smaller than 100 km should have decayed at 200 m depth in the thermocline. The observations instead show that the kinetic energy spectrum has a slope of −2 across the whole submesoscale range and that the kinetic- and potential-energy levels are similar in mixed layer and thermocline. The energy must be in deep modes, not in surface-trapped surface QG modes. The deep modes, however, are not the result of interior QG turbulence, because the spectra are flatter than −3.

One possible explanation for the failure of geostrophic-turbulence theories, as alluded to earlier, is that the flow may have a significant ageostrophic component at scales larger than 20 km. Because the energy levels exceed the GM spectra at these large scales, five candidates remain: frontal circulations, Ekman flows, mixed-layer turbulence, near-inertial oscillations, and internal tides. Ferrari and Rudnick (2000) showed that density fronts are weak in the region. We further find that the in-situ vertical shear is much greater than the geostrophic shear in both the mixed layer and the thermocline, which contradicts the interpretation in terms of frontal circulations, because in frontal flows, most of the shear is in the along-front component and thus in geostrophic balance. Ekman flows and mixed-layer turbulence are also unlikely to dominate over these scales, because these flows decay rapidly below the mixed layer, while the energy levels in the thermocline are observed to be close to those in the mixed layer at these scales. Near-inertial oscillations have a large ratio of kinetic to potential energy, while the potential energy is observed to be of the same order as the kinetic energy at all scales. There is instead support for a strong internal-tide field in the region. Richman et al. (2012) ran numerical simulations of the global oceans that resolved both geostrophic eddies and low-mode tides. They found that in regions of low mesoscale and high tidal activity, superinertial variability dominates the kinetic energy spectra from scales of a few kilometers to 100 km and larger. Using model output provided by Richman et al., we find that internal tides do indeed dominate the kinetic energy at scales of order 100 km in the region under consideration here. Richman et al. (2012) also found that the peaks due to internal tides are considerably broadened, which, together with the superposition with subinertial energy, could explain why no distinct peaks are apparent in the observed spectra.

If internal tides dominate at these large scales, the finite sampling speed is problematic and temporal variations of tidal motions alias into the wavenumber spectra. This prevents a straightforward comparison of model results and observations. More work is required to understand how internal tides project onto spatial scales.

To assess whether geostrophic turbulence is simply masked by unbalanced flow in the energy spectra, we now turn to an analysis of temperature spectra along isopycnals.
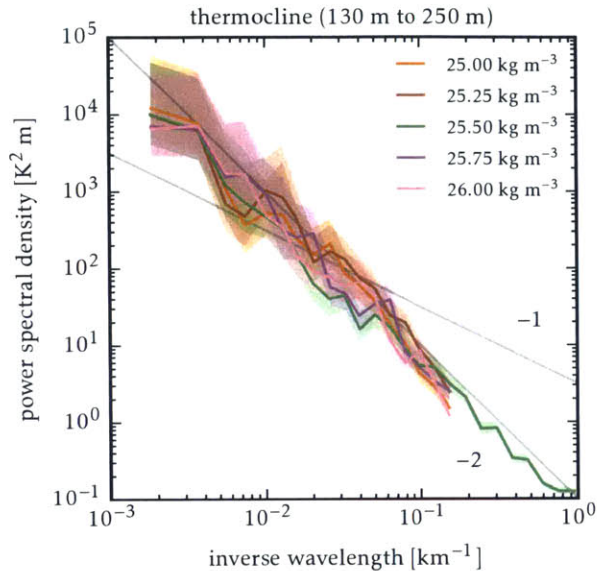
39

thermocline (130 m to 250 m)

Figure 2.7: Subtropical North Pacific temperature spectra along isopycnals interpolated from the sawtooth profile (25.00 kg m^{-3}, 25.25 kg m^{-3}, 25.75 kg m^{-3}, 26.00 kg m^{-3}) and from the isopycnal tow (25.50 kg m^{-3}); lines with slopes −1 and −2 for reference (gray solid lines). The shadings show 95% confidence intervals.

Temperature fluctuations along isopycnals are fully compensated by salinity and thus behave like a passive tracer (e.g. Ferrari and Rudnick, 2000; Smith and Ferrari, 2009). These spectra allow inferences about the balanced part of the flow, because stirring by internal waves, including internal tides, is essentially filtered out along isopycnals: only the along-isopycnal part of internal-wave motion displaces the weak horizontal temperature gradient.

We calculate temperature spectra along different isopycnals from the CTD tow along the 25.50 kg m^{-3} isopycnal and by interpolating temperature from the sawtooth profile onto isopycnals. The tracer spectra along the different isopycnals, which have a depth range of about 130 m to 250 m, all have very similar variance levels and have slopes close to −2, so there is no change with depth (Fig. 2.7). Recent observations with gliders in a nearby region confirm this finding: the slopes of passive-tracer spectra remain close to −2 down to 800 m depth (Cole and Rudnick, 2012). The slopes are again inconsistent with the vertical decay of surface QG modes. One would expect that passive-tracer spectra transition, as one moves deeper into the thermocline, from the surface QG surface spectrum to a flatter Batchelor spectrum at small scales, either due to interior QG turbulence or non-local stirring by large-scale modes (cf., Scott, 2006). Given the spectra estimated in the appendix, this transition should be well-resolved by the observations, but is not observed. Flows associated with frontal dynamics, as described by semi-geostrophy, have a less rapid vertical decay than surface QG flows and generate −2 spectral slopes (Hoskins and Bretherton, 1972; Badin, 2012). This may help explain the vertical structure of the tracer spectra, but further study is needed to clarify what spectra frontal dynamics

produce at different depths.

Another possible explanation for the failure of geostrophic-turbulence theories is that the assumption of a submesoscale inertial range is violated. Tulloch et al. (2011) suggested that baroclinic instability can inject energy and enstrophy throughout the submesoscale range. They performed a linear stability analysis in the subtropical North Pacific and found that there are submesoscale instabilities, in contrast to strong baroclinic current regions, where instabilities are confined to the mesoscale. The reason for the different instability regimes are the different PV gradient structures. There is a deep reversal of PV gradients in strong baroclinic currents, which induces a Phillips-type instability (Phillips, 1954). In quiescent gyres, on the other hand, the instability is more Charney-like, with a surface buoyancy gradient interacting with an interior PV gradient, which allows for the growth of small-scale Charney modes (Charney, 1947; Tulloch and Smith, 2009; Roullet et al., 2012). Capet et al. (2008d) performed high-resolution primitive-equation simulations that exhibit submesoscale instabilities. The kinetic- and potential energy spectra in these simulations transition from k_h^{-2} near the surface to k_h^{-3} at depth, while the spectrum of temperature remains k_h^{-2}. This indicates that subinertial flows with submesoscale instabilities are capable of producing tracer spectra similar to the ones observed, but again, the mechanisms require further study.

A further possibility is that vertical shear enhances vertical diffusion, which can then act on relatively large horizontal scales (Haynes and Vanneste, 2004; Smith and Ferrari, 2009). Further investigation is required to determine whether this mechanism can steepen the spectra from k_h^{-1} to k_h^{-2}.

2.7 Conclusions

The observations indicate that in the Gulf Stream region, interior QG turbulence dominates at scales larger than 20 km, where a transition to internal waves occurs in both mixed layer and upper thermocline. In the eastern subtropical North Pacific, geostrophic-turbulence theories fail. There is some indication of a significant unbalanced contribution to the kinetic energy up to scales of 100 km, most likely due to internal tides.

The results in the Gulf Stream region suggest that in a strong baroclinic current, where there are strong interior PV gradients, a Phillips-type instability creates deep mesoscale modes whose variability is then transferred into the submesoscale through an interior-QG enstrophy cascade. The associated flows dominate the kinetic energy down to scales of tens of kilometers, where internal waves become important. This picture, however, may change substantially in the presence of a deep winter mixed layer. Energetic submesoscale flows can develop and dominate over the interior modes in the mixed layer (e.g. Fox-Kemper et al., 2008; Capet et al., 2008e).

In quiescent-gyre regions, like the eastern subtropical North Pacific, interior PV gradients are much weaker. There is still a surface buoyancy gradient, so one might expect

surface QG dynamics to play a prominent role. But there is no evidence for surface QG turbulence in the observations: both kinetic-energy and tracer spectra are inconsistent with its predictions. Instead, unbalanced flows appear to make a leading-order contribution to the energy throughout the submesoscale range. Passive-tracer spectra along isopycnals, which filter out the effect of internal waves, show a -2 slope independent of depth. These tracer spectra indicate that also the balanced part of the flow does not follow predictions of geostrophic-turbulence theories: neither interior nor surface QG turbulence can explain this behavior. Small-scale Charney-type instabilities may be important in setting these spectra, but how -2 slopes emerge remains obscure. Another possibility is coupling between surface buoyancy anomalies and interior PV anomalies, an aspect of the dynamics considered in neither surface nor interior QG turbulence theory. This also warrants further investigation. A third possibility is that the tracer spectra are modified by shear-enhanced diffusion.

We found that previous claims that kinetic energy spectra with a $-5/3$ slope are telltale signatures of surface-QG dynamics failed to take into account the role of the mixed layer. The presence of a weakly stratified layer in the upper ocean acts to flatten the surface QG kinetic energy spectra in the submesoscale range between the first baroclinic deformation scale of order 100 km and the mixed layer deformation scale of order 10 km. Surface QG theory also ignores the strong PV gradient at the mixed layer base, which may introduce Eady-like dynamics in the mixed layer. Our analysis suggests that -2 spectral slopes in kinetic energy spectra can also be associated with a variety of unbalanced processes like fronts and internal tides. Clearly, the question of what sets the spectra of kinetic and potential energies at the ocean surface has not been fully answered yet.

That ageostrophic flows may be important near the surface in regions of weak mesoscale eddy activity has implications for attempts to retrieve the surface flow with altimeters, especially for the upcoming Surface Water and Ocean Topography (SWOT) mission that is expected to resolve SSH fluctuations down to scales of 10 km (e.g. Fu and Ferrari, 2008). Two problems come to the fore: (i) if the ageostrophic flow can project significantly onto SSH, it will contaminate the geostrophic surface velocity estimates and (ii) if the ageostrophic flow has no significant SSH expression, it will need to be inferred through other instruments, because it apparently contributes significantly to the surface velocities.

2.A   Surface QG for non-constant stratification

In theories of surface QG turbulence (Blumen, 1978; Lapeyre and Klein, 2006), the stratification is typically assumed to be constant. Below the injection scale, which is thought to be the mesoscale with its baroclinic instability, there is an enstrophy cascade, in which the surface potential energy scales like $\mathcal{P}_{k_h}(0) \sim k_h^{-5/3}$. With constant stratification, the surface kinetic energy spectrum then also scales like $\mathcal{K}_{k_h}(0) \sim k_h^{-5/3}$ throughout the enstrophy cascade range. The ocean, however, has highly non-constant stratification,

with an approximately exponential stratification in the thermocline and a weakly strat-ified mixed layer on top. LaCasce (2012) pointed out that in an ocean with exponential stratification and finite depth, the $k_h^{-5/3}$ surface potential energy spectrum translates into a surface kinetic energy spectrum that changes slope. We here consider a similar case with parameters appropriate for comparison with our observations in the subtropical North Pacific. We furthermore consider how the presence of a mixed layer changes the spec-tra. Because our observations span a range of depths below the surface, we also compute the vertical dependence of the spectra predicted by surface QG turbulence. In constant stratification, small-scale modes decay faster in the vertical than large-scale modes, so spectra at depth fall off steeply at small scales (Scott, 2006). We here extend this result to non-constant stratification.

To recapitulate Scott's result, let us start with the case of constant stratification. For simplicity, we consider an infinitely deep ocean, to avoid the introduction of an additional length scale associated with the depth of the ocean. The spectral surface QG stream-function $\hat{\psi}$ is then found by inverting the zero interior PV condition (e.g. Blumen, 1978; Lapeyre and Klein, 2006)

$$- k_h{}^2 \hat{\psi} + \frac{d}{dz}\left(\frac{f^2}{N^2}\frac{d\hat{\psi}}{dz}\right) = 0, \qquad f\frac{d\hat{\psi}}{dz}(0) = \hat{b}(0), \qquad \hat{\psi}(-\infty) = 0, \tag{2.3}$$

which gives, for constant N,

$$\hat{\psi}(z) = \frac{\hat{b}(0)}{Nk_h}e^{\frac{Nk_h z}{f}} \tag{2.4}$$

and the kinetic- and potential energy spectra are

$$\mathcal{K}_{k_h}(z) = \mathcal{P}_{k_h}(z) = \mathcal{P}_{k_h}(0)e^{\frac{2Nk_h z}{f}}. \tag{2.5}$$

These spectra are shown in Fig. 2.3 (top left), using the stratification $N = 8.7 \times 10^{-3}$ s^{-1} observed at 200 m depth in the subtropical North Pacific. There is equipartition between kinetic and potential energy at all scales. At a certain depth z, the spectra follow the surface potential energy spectrum at scales large enough for the modes to penetrate to this depth, that is at scales larger than the deformation scale associated with z, $k_h \ll -f/Nz$. Around $k_h = -f/Nz$, the spectra fall off sharply, because smaller-scale modes do not reach this depth.

For exponential stratification $N = N_0 \exp(z/d)$ and again an infinitely deep ocean, the surface QG streamfunction is (cf., LaCasce, 2012)

$$\hat{\psi}(z) = \frac{\hat{b}(0)}{N_0 k_h} \frac{I_1(\mu e^{\frac{z}{d}})}{I_0(\mu)}e^{\frac{z}{d}}, \tag{2.6}$$

where the I_α are modified Bessel functions of the first kind and $\mu = N_0 k_h d/f$. The kinetic-

and potential energy spectra are

$$\mathcal{K}_{k_h}(z) = \mathcal{P}_{k_h}(0)\left[\frac{I_1(\mu e^{\frac{z}{d}})}{I_0(\mu)}e^{\frac{z}{d}}\right]^2, \qquad \mathcal{P}_{k_h}(z) = \mathcal{P}_{k_h}(0)\left[\frac{I_0(\mu e^{\frac{z}{d}})}{I_0(\mu)}e^{\frac{z}{d}}\right]^2. \tag{2.7}$$

These are shown in Fig. 2.3 (top right) with $d = 600$ m, which is representative of the subtropical North Pacific thermocline, and N_0 chosen such that the stratification matches the observed thermocline value $N = 8.7 \times 10^{-3}$ s^{-1} at 200 m depth. Let us consider the surface spectra first. For scales small compared to the deformation scale $k_h \gg f/N_0 d$ associated with the depth scale of stratification d, the kinetic- and potential energy spectra are equal, because $I_1(\mu)/I_0(\mu) \approx 1$ for $\mu \gg 1$. In this range, the spectra behave in the same way as with constant stratification.[13] At scales larger than the deformation scale, $k_h \ll f/N_0 d$, however, the surface kinetic energy spectrum is

$$\mathcal{K}_{k_h}(0) \approx \frac{N_0{}^2 k_h{}^2 d^2}{4 f^2} \mathcal{P}_{k_h}(0), \tag{2.8}$$

because $I_1(\mu)/I_0(\mu) \approx \mu/2$ for $\mu \ll 1$. The surface kinetic energy spectrum in this range of scales is therefore much smaller than the potential energy spectrum and has a different slope—the equipartition between kinetic and potential energy is broken. Below the surface, as with constant stratification, the spectra fall off at small scales, because modes decay away from the surface.

The presence of a mixed layer also modifies the energy spectra. To demonstrate this, we consider a mixed layer of depth h with constant stratification N_{ML} overlying an infinitely deep ocean, which also has constant but higher stratification N_{TH}. The surface QG streamfunction for this case is

$$\hat{\psi}(z) = \begin{cases} \dfrac{\hat{b}(0)}{N_{ML}k_h} \dfrac{N_{TH}\sinh\frac{N_{ML}k_h(z+h)}{f}+N_{ML}\cosh\frac{N_{ML}k_h(z+h)}{f}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}} & z > -h, \\[20pt] \dfrac{\hat{b}(0)}{k_h} \dfrac{e^{\frac{N_{TH}k_h(z+h)}{f}}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}} & z < -h \end{cases} \tag{2.9}$$

---

[13]This can also be rationalized by WKB arguments: if the modes are shallow enough to not feel the change in stratification, they behave as if the stratification were constant.

and the kinetic- and potential energy spectra are

$$\mathcal{K}_{k_h}(z) = \begin{cases} \mathcal{P}_{k_h}(0)\left[\dfrac{N_{TH}\sinh\frac{N_{ML}k_h(z+h)}{f}+N_{ML}\cosh\frac{N_{ML}k_h(z+h)}{f}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}}\right]^2 & z > -h, \\[4ex] \mathcal{P}_{k_h}(0)\left[\dfrac{N_{ML}e^{\frac{N_{TH}k_h(z+h)}{f}}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}}\right]^2 & z < -h, \end{cases} \tag{2.10}$$

$$\mathcal{P}_{k_h}(z) = \begin{cases} \mathcal{P}_{k_h}(0)\left[\dfrac{N_{TH}\cosh\frac{N_{ML}k_h(z+h)}{f}+N_{ML}\sinh\frac{N_{ML}k_h(z+h)}{f}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}}\right]^2 & z > -h, \\[4ex] \mathcal{P}_{k_h}(0)\left[\dfrac{N_{ML}e^{\frac{N_{TH}k_h(z+h)}{f}}}{N_{TH}\cosh\frac{N_{ML}k_h h}{f}+N_{ML}\sinh\frac{N_{ML}k_h h}{f}}\right]^2 & z < -h. \end{cases} \tag{2.11}$$

These are shown in Fig. 2.3 (bottom left) for $N_{ML} = 2.3\times10^{-3}$ s^{-1} and $N_{TH} = 8.7\times10^{-3}$ s^{-1}, the values observed at 50 m and 200 m depth, and $h = 100$ m, approximately the observed mixed-layer depth. The surface kinetic energy spectrum again follows the potential energy spectrum at small scales, where modes decay within the mixed layer, but transitions to a lower level at large scales, where modes are much deeper than the mixed layer and the thermocline stratification dominates the mode structure. This transition occurs around the deformation scale associated with the depth of the mixed layer, $N_{ML}h/f$. In the thermocline, energy levels are drastically reduced, due to the jump to higher stratification at the base of the mixed layer.

Let us now combine a mixed layer of depth h and constant stratification N_{ML} with exponential stratification $N = N_0\exp(z/d)$ below, yielding a somewhat realistic stratification profile. The surface QG streamfunction for this case is

$$\hat{\psi}(z) = \begin{cases} \dfrac{\hat{b}(0)}{N_{ML}k_h}\,\dfrac{N_0 e^{-\frac{h}{d}}\sinh\frac{N_{ML}k_h(z+h)}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\cosh\frac{N_{ML}k_h(z+h)}{f}I_1(\mu e^{-\frac{h}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})} & z > -h, \\[4ex] \dfrac{\hat{b}(0)}{k_h}\,\dfrac{I_1(\mu e^{\frac{z}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})}e^{\frac{z+h}{d}} & z < -h \end{cases} \tag{2.12}$$

and the kinetic- and potential energy spectra are

$$\mathcal{K}_{k_h}(z) = \begin{cases} \mathcal{P}_{k_h}(0)\left[\dfrac{N_0 e^{-\frac{h}{d}}\sinh\frac{N_{ML}k_h(z+h)}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\cosh\frac{N_{ML}k_h(z+h)}{f}I_1(\mu e^{-\frac{h}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})}\right]^2 & z > -h, \\[4ex] \mathcal{P}_{k_h}(0)\left[\dfrac{N_{ML}I_1(\mu e^{\frac{z}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})}e^{\frac{z+h}{d}}\right]^2 & z < -h, \end{cases} \tag{2.13}$$

$$\mathcal{P}_{k_h}(z) = \begin{cases} \mathcal{P}_{k_h}(0)\left[\dfrac{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h(z+h)}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h(z+h)}{f}I_1(\mu e^{-\frac{h}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})}\right]^2 & z > -h, \\[4ex] \mathcal{P}_{k_h}(0)\left[\dfrac{N_{ML}I_0(\mu e^{\frac{z}{d}})}{N_0 e^{-\frac{h}{d}}\cosh\frac{N_{ML}k_h h}{f}I_0(\mu e^{-\frac{h}{d}})+N_{ML}\sinh\frac{N_{ML}k_h h}{f}I_1(\mu e^{-\frac{h}{d}})}e^{\frac{z+h}{d}}\right]^2 & z < -h. \end{cases} \tag{2.14}$$

The spectra shown in Fig. 2.3 (bottom right), again using parameters appropriate for the subtropical North Pacific, exhibit the combined effect of exponential stratification and a mixed layer: there is a flattening of the kinetic energy spectra at large scales due to the exponential stratification, potential-energy levels in the mixed layer are much higher than in the thermocline, mixed-layer kinetic energy spectra transition from the potential energy spectra at small scales to lower energy levels at large scales, and all spectra at depth fall off steeply at small scales.

To make connection with the observations, we transform the two-dimensional isotropic spectra to one-dimensional spectra using (2.1) and compute both longitudinal and transverse kinetic energy spectra (Fig. 2.4). The overall shape of the spectra is similar to the associated two-dimensional isotropic spectra, with slight modifications where the spectra are flat. The longitudinal and transverse kinetic energy spectra still obey (2.2)— but for non-constant slopes, they are not separated by a constant factor anymore.

This analysis shows that taking both non-constant stratification and the vertical decay of modes into account leads to surface QG turbulence predictions of energy spectra at depths 50 m and 200 m that do not exhibit $k^{-5/3}$ power laws anywhere in the submesoscale range. Even the surface kinetic energy spectrum does not exhibit this scaling throughout the submesoscale range: it is much flatter at scales larger than 10 km. It should also be noted that throughout the submesoscale range, both kinetic- and potential-energy levels are predicted to drop significantly going from the mixed layer into the thermocline.

The discussion in this appendix follows a literal interpretation of surface QG turbulence. The jump in stratification at the base of the mixed layer, however, is associated with a large PV gradient and may act like a surface itself (Smith and Vanneste, 2013). Such a scenario could be represented with a simplified model consisting of two constant-PV layers—the mixed layer and the thermocline—joined by an interface at the base of the mixed layer. This model will produce different modal structures and energy spectra than that considered above, but it remains to be addressed whether these match the observations.

CHAPTER 3

WAVE–VORTEX DECOMPOSITION OF ONE-DIMENSIONAL SHIP-TRACK DATA

3.1 Introduction

The decomposition of a complex flow into various constituents that are distinguished by their intrinsic physical and mathematical properties is a powerful conceptual tool, which is particularly useful in the flows typical for the atmosphere and ocean, where small-scale dispersive waves, quasi-two-dimensional large-scale vortical flows along isopycnal surfaces, and pockets of three-dimensional turbulence all intermingle in a nonlinear jigsaw puzzle. The most elementary of such flow decomposition methods is based on the linearized fluid equations relative to a state of rest, which for a rotating and stratified three-dimensional fluid system such as the Boussinesq model leads to the familiar decomposition into a horizontal flow in geostrophic and hydrostatic balance on the one hand and into unbalanced inertia–gravity waves on the other. However, even this most basic decomposition method in principle requires knowledge of all the flow variables throughout the entire three-dimensional domain, a task that is as straightforward in a numerical model as it is hopeless in observational practice.

Indeed, many observations in the atmosphere and ocean are confined to a fixed location, or to a sequence of locations along a horizontal transect following an airplane flight or a ship track, say. In this latter situation it is possible to compute one-dimensional spectra along the ship track of flow variables such as the horizontal velocities or the buoyancy. By assuming stationarity and homogeneity as needed, this allows the estimation of one-dimensional covariance functions or, equivalently, of one-dimensional power spectra. Of course, such ship track data is highly aliased in the sense that the power spectra at the one-dimensional wavenumber $k > 0$ along the ship track, say, are affected by the multi-dimensional dynamics associated with all wavenumber vectors with magnitude greater or equal to k. The situation improves if one can assume horizontal isotropy at least, in which case one can exploit the link between one-dimensional and isotropic two-dimensional spectra. In particular, one can then exploit the well-known differences between the power spectra of along-track and across-track velocity components (e.g. Batchelor, 1953; Charney, 1971) to gain some insight into the dynamics of the underlying flow, such as its decomposition into waves and other constituents.

These issues are of pressing concern especially in oceanography, where observations are sparse and our understanding of the relevant dynamical processes is poor. For example, the oceanic motions are very energetic in the submesoscale horizontal range between order 100 km, the scale of the Rossby radius of deformation, and order 1 km, the smallest scale at which rotation strongly affects dynamics. These motions play an important role in the overall ocean circulation, because they connect the large scales where the ocean is energized by atmospheric forcing to the small scales at which energy is dissipated (Ferrari and Wunsch, 2010). Despite the explosion of theoretical studies of submesoscale dynamics in the last decade, the dearth of observations on this range of scales has slowed progress.

A major question is to partition the relative contributions of geostrophic eddies or

inertia–gravity waves at the submesoscales, because they have different impact on dynamics and tracer transport (e.g. Chapter 2). A turbulent field of geostrophic eddies tends to transfer energy to larger scales and mix tracers along density surfaces. Internal wave interactions, instead, transfer energy to smaller vertical-scale waves, which break and mix tracers across density surfaces. In situ mooring observations are useful to separate subinertial and super-inertial motions, but the lack of spatial information prevents any conclusive statement about whether the sub-inertial motions are submesoscale geostrophic eddies or larger-scale motions and whether the super-inertial motions are inertia–gravity waves or other forms of stratified turbulence. Two dimensional sections (along ship tracks and depth) of velocity and density provide snapshots of the superposition of all submesoscale motions with no temporal information to separate the relative contributions. Maps of potential vorticity (PV) could be used to distinguish the two classes of motion, because geostrophic eddies are associated with PV anomalies whilst inertia–gravity waves have no PV signature. But PV requires three dimensional sections and it is extremely difficult to measure accurately.

In Chapter 2, we used the fact that the ratio of the power spectra of along and across ship track velocities can be used to determine whether a velocity field is horizontally nondivergent, a necessary condition for flows to be in geostrophic balance. In this sense the geostrophic flow left a clear fingerprint in the observed horizontal velocities that could be tracked down. However, the horizontal velocity field due to inertia–gravity waves has both rotational and divergent components, so its fingerprint in the observed fields is not so immediately apparent.

In this work, we present a two-step method that allows extracting the fingerprint of inertia–gravity waves from ship track spectra. First, we show that measurements of the two horizontal velocity components along ship tracks are sufficient to partition the flow uniquely into rotational (i.e. horizontally nondivergent) and divergent (i.e. horizontally irrotational) components, provided one can assume that the two fields are horizontally homogeneous and isotropic at the measured scales, and that their rotational and divergent components are uncorrelated in a statistical sense. This Helmholtz decomposition of the one-dimensional velocity spectra is purely kinematic in nature and is achieved by solving two simple ODEs in spectral space, an easy numerical task. We have not been able to find a direct prior reference in the literature for this kind of exact Helmholtz decomposition method for one-dimensional spectra, although related theoretical investigations do exist (e.g. Lindborg, 2007).

Second, we show that the one-dimensional inertia–gravity wave energy spectrum can be computed from the aforementioned Helmholtz decomposition in combination with a statement about wave energy equipartition for hydrostatic and vertically homogeneous inertia–gravity waves. If buoyancy spectra are also available then this means that the total observed energy spectrum can be exactly decomposed into its inertia–gravity wave component and its residual vortical component in geostrophic balance. We also show that one can derive additional relationships between the power spectra if one has some infor-

51

mation about the frequency content of the wave field (for example that it is narrow band, or that it follows a simple model spectrum). In this case it is possible to decompose not just the total observed energy spectrum but also the individual observed field spectra into their balanced and unbalanced components. This allows a complete decomposition of the one-dimensional spectra into their wave and vortex components, which is of obvious dynamical significance.

In Section 3.2, we introduce the Helmholtz decomposition method for separating the rotational and divergent components of a two-dimensional velocity field from one dimensional spectra. Then we show how to determine what part of the spectrum is composed of inertia–gravity waves. Throughout, we consider only a very simple fluid set-up, which is a three-dimensional Boussinesq model with constant Coriolis parameter f and buoyancy frequency N. In Section 3.3, the method is applied to ship based measurements of the upper ocean velocity field from two field experiments, one in the eastern subtropical North Pacific and one in the western North Atlantic. Despite noisy data and restrictive assumptions, in both cases the separation of submesoscale geostrophic motions from inertia–gravity waves appears to be fully successful.

3.2 Helmholtz decomposition and wave diagnostics

We begin by assembling some generic facts about one-dimensional velocity spectra that derive from a two-dimensional horizontal flow with homogeneous and isotropic statistics. For horizontally nondivergent flows this is a subset of well-known results from homogeneous turbulence theory, but for inertia–gravity waves we need to accommodate horizontal velocity fields that have both rotational and divergent components, which is a less studied case.

3.2.1 Helmholtz decomposition of one-dimensional spectra

Let u and v be horizontal velocity components defined in the xy-plane with x aligned with the ship track, so u is the along-track, "longitudinal" component and v is the across-track, "transverse" component. The time t and depth z are considered fixed during the measurement, so we may ignore these coordinates at this stage. If the flow is purely rotational, i.e. horizontally nondivergent, then (u,v) derive from a stream function $\psi(x,y)$ in the standard way:

$$u_x + v_y = 0 \quad \Rightarrow \quad u = -\psi_y \quad \text{and} \quad v = +\psi_x. \tag{3.1}$$

Let ψ be a homogeneous and isotropic zero-mean random function such that

$$\mathbb{E}[\psi] = 0 \quad \text{and} \quad C^{\psi}(x,y) = \mathbb{E}[\psi(x_0,y_0)\psi(x_0 + x, y_0 + y)] = F(r) \tag{3.2}$$

52

where $r = (x^2 + y^2)^{1/2}$ and $\mathbb{E}$ denotes taking the expected value. The function $F(r)$ is the covariance of the stream function, which is a function of horizontal distance $r \geq 0$ in the two-dimensional plane and encapsulates all the statistical knowledge that is available for the random velocity field. The power spectrum $\hat{C}^\psi(k,l)$ is the Fourier transform of $C^\psi(x,y)$, i.e.

$$\hat{C}^\psi(k,l) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C^\psi(x,y)e^{-i(kx+ly)}\,dx\,dy = 2\pi \int_0^\infty J_0(k_h r)F(r)\,r\,dr = \hat{F}(k_h) \tag{3.3}$$

where $k_h = (k^2 + l^2)^{1/2}$. The corresponding velocity spectra follow from (3.1) as

$$\hat{C}^u(k,l) = l^2\,\hat{C}^\psi(k,l) = l^2\,\hat{F}(k_h) \quad \text{and} \quad \hat{C}^v(k,l) = k^2\,\hat{C}^\psi(k,l) = k^2\,\hat{F}(k_h). \tag{3.4}$$

These are clearly not isotropic even though $\hat{C}^\psi$ is. For completeness, the cross-spectrum

$$\hat{C}^{uv}(k,l) = -kl\,\hat{C}^\psi(k,l) = -kl\,\hat{F}(k_h). \tag{3.5}$$

Now, along the ship track $y = 0$ and $r = x$, so the relevant one-dimensional covariance functions are given by $C^\psi(x,0) = F(x)$, for example. For the power spectra this corresponds to integrating over the transverse wavenumber l:

$$\hat{C}^\psi(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{C}^\psi(k,l)\,dl = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{F}(k_h)\,dl. \tag{3.6}$$

This integrand is even in l and, at fixed k, we have $l\,dl = k_h\,dk_h$, which allows rewriting (3.6) as

$$\hat{C}^\psi(k) = \frac{1}{\pi}\int_{|k|}^{\infty} \frac{\hat{F}(k_h)}{\sqrt{k_h^2 - k^2}}\,k_h\,dk_h. \tag{3.7}$$

The same steps lead to the one-dimensional velocity spectra

$$\psi \text{ only:} \quad \hat{C}^u(k) = \frac{1}{\pi}\int_{|k|}^{\infty} \hat{F}(k_h)\sqrt{k_h^2 - k^2}\,k_h\,dk_h \tag{3.8}$$

$$\hat{C}^v(k) = \frac{k^2}{\pi}\int_{|k|}^{\infty} \frac{\hat{F}(k_h)}{\sqrt{k_h^2 - k^2}}\,k_h\,dk_h, \tag{3.9}$$

where "ψ only" is added as a reminder that these expressions hold only for horizontally nondivergent flows that can be expressed through a stream function ψ. The corresponding cross-spectrum $\hat{C}^{uv}(k) = 0$, because the relevant symbol kl in (3.5) is odd in l and hence integrates to zero. This will always be the case, so we won't consider the cross-spectrum any further. By inspection, and using Leibniz's rule, we obtain the celebrated formula

$$\psi \text{ only:} \quad \hat{C}^v(k) = -k\frac{d}{dk}\hat{C}^u(k) \tag{3.10}$$

53

for the horizontally nondivergent case (e.g. Charney, 1971). For power-law velocity spectra of the form k^{-n} this yields

$$\psi \text{ only:} \quad \hat{C}^v(k) = n\hat{C}^u(k), \tag{3.11}$$

which, for $n > 1$, means that along a ship track the transverse spectrum dominates the longitudinal spectrum for two-dimensional incompressible flows. This is the fingerprint that was exploited in Chapter 2.

Conversely, we may consider a purely divergent flow, i.e. one that is two-dimensionally irrotational such that (3.1) is replaced by

$$v_x - u_y = 0 \quad \Rightarrow \quad u = \phi_x \quad \text{and} \quad v = \phi_y \tag{3.12}$$

in terms of a homogeneous and isotropic zero-mean random potential $\phi(x,y)$ defined by

$$\mathbb{E}[\phi] = 0 \quad \text{and} \quad C^\phi(x,y) = \mathbb{E}[\phi(x_0,y_0)\phi(x_0+x,y_0+y)] = G(r), \tag{3.13}$$

where $G(r)$ is the covariance of ϕ. Clearly, apart from the sign change, this simply reverses the roles of u and v in (3.1), so all the spectra can be worked out just as before. For the two-dimensional spectra this yields

$$\hat{C}^u(k,l) = k^2 \hat{G}(k_h) \quad \text{and} \quad \hat{C}^v(k,l) = l^2 \hat{G}(k_h) \tag{3.14}$$

and the one-dimensional ship track spectra are

$$\phi \text{ only:} \quad \hat{C}^u(k) = \frac{k^2}{\pi} \int_{|k|}^\infty \frac{\hat{G}(k_h)}{\sqrt{k_h^2 - k^2}} k_h dk_h \tag{3.15}$$

$$\hat{C}^v(k) = \frac{1}{\pi} \int_{|k|}^\infty \hat{G}(k_h)\sqrt{k_h^2 - k^2}\, k_h dk_h, \tag{3.16}$$

where "ϕ only" is added as a reminder that these expressions hold only for irrotational flows that can be expressed through a velocity potential ϕ. The relationship (3.10) is replaced by

$$\phi \text{ only:} \quad \hat{C}^u(k) = -k\frac{\mathrm{d}}{\mathrm{d}k}\hat{C}^v(k) \tag{3.17}$$

and hence now the longitudinal spectrum dominates for power laws with $n > 1$:

$$\phi \text{ only:} \quad \hat{C}^u(k) = n\hat{C}^v(k). \tag{3.18}$$

Note that by definition both $\hat{F}$ and $\hat{G}$ are real and non-negative functions of k_h.

Now, a general two-dimensional flow has a Helmholtz decomposition into rotational and divergent components of the form

$$u = -\psi_y + \phi_x \quad \text{and} \quad v = \psi_x + \phi_y, \tag{3.19}$$

54

which implies the two-dimensional Poisson equations

$$\psi_{xx} + \psi_{yy} = v_x - u_y \quad \text{and} \quad \phi_{xx} + \phi_{yy} = u_x + v_y. \tag{3.20}$$

This determines both ψ and ϕ up to a harmonic function, but with doubly periodic boundary conditions such a harmonic function could only be a physically meaningless constant, so ψ and ϕ are in fact uniquely determined by (3.20). Progress with the statistical theory is then possible if $\psi(x,y)$ and $\phi(x,y)$ are uncorrelated in the sense that

$$\mathbb{E}[\psi(x_0, y_0)\phi(x_0 + x, y_0 + y)] = 0 \tag{3.21}$$

holds for all (x,y). Under this assumption the velocity covariances due to ψ and ϕ simply add up, which yields the one-dimensional ship track spectra

$$\hat{C}^u(k) = \frac{1}{\pi} \int_{|k|}^{\infty} \left[\hat{F}(k_h)\sqrt{k_h^2 - k^2} + \frac{k^2 \hat{G}(k_h)}{\sqrt{k_h^2 - k^2}} \right] k_h dk_h \tag{3.22}$$

$$\hat{C}^v(k) = \frac{1}{\pi} \int_{|k|}^{\infty} \left[\frac{k^2 \hat{F}(k_h)}{\sqrt{k_h^2 - k^2}} + \hat{G}(k_h)\sqrt{k_h^2 - k^2} \right] k_h dk_h. \tag{3.23}$$

These expressions can be substantially simplified if one introduces the auxiliary functions $D^\psi(k)$ and $D^\phi(k)$ defined by

$$D^\psi(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} l^2 \hat{C}^\psi(k,l) \, dl = \frac{1}{\pi} \int_{|k|}^{\infty} \hat{F}(k_h)\sqrt{k_h^2 - k^2} \, k_h dk_h \tag{3.24}$$

$$D^\phi(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} l^2 \hat{C}^\phi(k,l) \, dl = \frac{1}{\pi} \int_{|k|}^{\infty} \hat{G}(k_h)\sqrt{k_h^2 - k^2} \, k_h dk_h, \tag{3.25}$$

The functions D^ψ and D^ϕ are the spectra of ψ_y and ϕ_y, respectively, and they allow rewriting (3.22-3.23) in the succinct form

$$\hat{C}^u(k) = D^\psi(k) - k\frac{d}{dk}D^\phi(k) \quad \text{and} \quad \hat{C}^v(k) = D^\phi(k) - k\frac{d}{dk}D^\psi(k). \tag{3.26}$$

This is the main result of this section, which neatly incorporates both (3.10) and (3.17) as special cases. The functions D^ψ and D^ϕ can be viewed as analogous to ψ and ϕ in the Helmholtz decomposition of the ship track velocity spectra. Correspondingly, the horizontal kinetic energy spectrum can be viewed as the sum of a rotational and a divergent part:

$$\frac{1}{2}\left[\hat{C}^u(k) + \hat{C}^v(k) \right] = \frac{1}{2}\left[D^\psi(k) - k\frac{d}{dk}D^\psi(k) \right] + \frac{1}{2}\left[D^\phi(k) - k\frac{d}{dk}D^\phi(k) \right]. \tag{3.27}$$

55

3.2.2 Numerical method for Helmholtz decomposition

The spectral functions $D^\psi(k)$ and $D^\phi(k)$ are not directly observable from the ship track data, but (3.26) suggests a simple and robust method for computing them from the directly observed $\hat{C}^u$ and $\hat{C}^v$. First, the functions D^ψ and D^ϕ are symmetric in k so we only need to find their values for $k \geq 0$. Second, in the limit $k \to \infty$ we have the robust decay boundary conditions

$$D^\psi(+\infty) = D^\phi(+\infty) = 0 \tag{3.28}$$

and this allows us to compute D^ψ and D^ϕ for $k \geq 0$ by integrating the two ODEs in (3.26) backwards in k, starting from zero values at $k = +\infty$. This is particularly easy in the logarithmic wavenumber

$$s = \ln k \quad \text{such that} \quad \frac{d}{ds} = k\frac{d}{dk}. \tag{3.29}$$

The ODEs in (3.26) can be integrated numerically, but there is also a closed-form solution

$$D^\psi(s) = \int_s^\infty \left[\hat{C}^u(\bar{s})\sinh(s-\bar{s}) + \hat{C}^v(\bar{s})\cosh(s-\bar{s})\right] d\bar{s} \tag{3.30}$$

$$D^\phi(s) = \int_s^\infty \left[\hat{C}^u(\bar{s})\cosh(s-\bar{s}) + \hat{C}^v(\bar{s})\sinh(s-\bar{s})\right] d\bar{s}. \tag{3.31}$$

In this formulation, the functions D^ψ and D^ϕ at wavenumber k only depend on the velocity spectra at wavenumbers larger than k, which is consistent with the aliasing apparent in their definition in (3.24). Notably, the $\sinh(s-\bar{s})$ terms in (3.30) and (3.31) are negative, which can lead to unphysical negative values in D^ψ or D^ϕ. This may occur when either D^ψ or D^ϕ become very small, say comparable to the instrumental noise threshold or to the errors imposed by the limitations of the assumptions of isotropy and homogeneity in the data (cf. Section 3.3.2 below).

3.2.3 Inertia–gravity waves and the wave energy spectrum

We now consider the linear Boussinesq equations with constant f and N in a domain with doubly periodic horizontal boundary conditions. The horizontal velocity field induced by linear inertia–gravity waves then has both a stream function and a velocity potential component, which are related as follows. The vertical vorticity satisfies the linear equation

$$(v_x - u_y)_t = fw_z = -f(u_x + v_y) \quad \Rightarrow \quad \psi_t + f\phi = 0. \tag{3.32}$$

Equation (3.32) shows that ψ and ϕ are in quadrature in time, which we will take to imply that for a stationary and horizontally isotropic field of random waves ψ and ϕ are uncorrelated at any fixed time t, and therefore the key assumption (3.21) indeed holds for linear inertia–gravity waves. The reasoning behind this is described in appendix 3.A.

It is of course possible to define further spectral relationships based on the linear equations, for example (3.32) implies the frequency-dependent relationship

$$\hat{C}_W^{\psi}(k,l,\omega) = \frac{f^2}{\omega^2} \hat{C}_W^{\phi}(k,l,\omega) \tag{3.33}$$

between the three-dimensional wave spectra of ψ and ϕ, which are defined in the usual way as functions of the horizontal wavenumbers and the wave frequency. (Here and in the following we denote wave-related functions by the subscript W.) However, such relationships include the wave frequency ω as a parameter, which depends on (k,l) but also on the vertical wavenumber or some other information about the vertical structure of the waves. This is of limited use for general ship track observations, where the vertical structure is typically not known and therefore (3.33) cannot be reduced to a unique statement for two-dimensional or one-dimensional spectra. We note in passing that the situation would be very different in a two-dimensional fluid system such as the shallow-water equations, where the dispersion relation determines ω^2 as a function of (k,l). In this case (3.33) would indeed hold for the two-dimensional spectra, with ω^2 determined from the dispersion relation.

We will therefore pursue a different course of action under the assumption that the waves are *hydrostatic*, implying that the vertical velocity is negligible in the wave energy budget. Somewhat surprisingly, this assumption will allow us to deduce the exact one-dimensional wave energy spectrum

$$E_W(k) = \frac{1}{2}\left[\hat{C}_W^u(k) + \hat{C}_W^v(k) + \hat{C}_W^b(k)\right] \tag{3.34}$$

solely from ship track observations of u and v! Here $\hat{C}_W^b(k)$ is the spectrum of b/N where b is the linear buoyancy disturbance, which is related to the vertical velocity w by

$$b_t + N^2 w = 0. \tag{3.35}$$

Hence $\frac{1}{2}\hat{C}_W^b$ is the potential energy spectrum. The computation of $E_W(k)$ in (3.34) hinges on the following statement about the energy equipartition for linear hydrostatic inertia–gravity waves that are stationary in time as well as spatially homogeneous in all three directions: the sum of the potential energy plus the rotational horizontal kinetic energy due to ψ then equals the divergent horizontal kinetic energy due to ϕ. This statement is demonstrated in appendix 3.B and using (3.27) it takes the form

$$\hat{C}_W^b(k) + \left[D_W^{\psi}(k) - k\frac{d}{dk}D_W^{\psi}(k)\right] = \left[D_W^{\phi}(k) - k\frac{d}{dk}D_W^{\phi}(k)\right]. \tag{3.36}$$

Substitution in (3.34) then immediately yields the key result

$$E_W(k) = D_W^{\phi}(k) - k\frac{d}{dk}D_W^{\phi}(k). \tag{3.37}$$

57

In a nutshell, this equation asserts that a Helmholtz decomposition of the horizontal velocity spectra along a ship track yields the exact wave energy spectrum of linear hydrostatic inertia–gravity waves, at least under the assumption that the waves can be modeled as spatially homogeneous, including in the vertical. For a wave field dominated by low-order vertical modes this latter assumption would fail, but otherwise the generality of (3.37) is remarkable.

3.2.4 Combination with a geostrophic flow component

If the inertia–gravity wave field is embedded in a quasi-geostrophic flow then the horizontal velocity field can be viewed as the sum of an unbalanced wave part and of a balanced "vortex" part. We allow for this by extending the Helmholtz decomposition to

$$\psi = \psi_W + \psi_V \quad \text{and} \quad \phi = \phi_W, \tag{3.38}$$

where the subscript V denotes the vortex part. The vortex part is horizontally nondivergent and therefore ϕ has no vortex part.

It is reasonable on physical grounds to assume that ψ_V is statistically independent of ψ_W and ϕ_W, in which case the covariances due to ψ_V simply add to the wave covariances we have already considered, i.e.,

$$\hat{C}^u(k) = \hat{C}^u_W(k) + \hat{C}^u_V(k) \quad \text{and} \quad \hat{C}^v(k) = \hat{C}^v_W(k) + \hat{C}^v_V(k). \tag{3.39}$$

The corresponding Helmholtz decomposition leads to

$$D^\psi(k) = D^\psi_W(k) + D^\psi_V(k) \quad \text{and} \quad D^\phi(k) = D^\phi_W(k) \tag{3.40}$$

such that

$$\hat{C}^u_W(k) = D^\psi_W(k) - k\frac{d}{dk}D^\phi_W(k), \quad \hat{C}^v_W(k) = -k\frac{d}{dk}D^\psi_W(k) + D^\phi_W(k) \tag{3.41}$$

$$\hat{C}^u_V(k) = D^\psi_V(k), \quad \hat{C}^v_V(k) = -k\frac{d}{dk}D^\psi_V(k). \tag{3.42}$$

Crucially, $D^\phi_W = D^\phi$ can be computed from the observed velocity spectra exactly as before, i.e., the function D^ϕ_W computed from the velocity observations is unaffected by the quasi-geostrophic flow component. The same is hence true for the hydrostatic wave energy spectrum E_W computed in (3.37). In other words, allowing for the presence of a quasi-geostrophic stream function does not affect the method of computation of the wave energy spectrum at all, because the crucial potential part of the Helmholtz decomposition of the spectra is not affected by the quasi-geostrophic flow.

58

At this point E_W, $D^\phi = D_W^\phi$ and the sum $D^\psi = D_W^\psi + D_V^\psi$ are known, but not D_W^ψ and D_V^ψ individually. So the vortical energy spectrum as well as the individual wave and vortex velocity spectra in (3.41–3.42) are still unknown. Either additional assumptions or additional observations are needed to progress further. We consider two options: either observing b/N along the ship track, or assuming additional information about the frequency content of the wave field.

3.2.5   Observed buoyancy spectrum

If b/N is observed along the ship track then we know the potential energy spectrum $\frac{1}{2}\hat{C}^b$ and hence also the total energy spectrum

$$E(k) = E_W(k) + E_V(k) = \frac{1}{2}\left[\hat{C}^u(k) + \hat{C}^v(k) + \hat{C}^b(k)\right].\tag{3.43}$$

Now, because E_W can be computed from (3.37), the vortical energy spectrum

$$E_V(k) = \frac{1}{2}\left[\hat{C}_V^u(k) + \hat{C}_V^v(k) + \hat{C}_V^b(k)\right] = E(k) - E_W(k)\tag{3.44}$$

simply follows by subtraction, so we now know both E_W and the residual E_V. This provides an exact energy decomposition into wave and vortex parts based solely on observing $(u, v, b/N)$ along a ship track, which is of obvious physical importance.

On the other hand, it is still not possible to compute the spectra of the individual wave and vortex fields from the available data. For example, $\hat{C}^b$ is known but not its constituents $\hat{C}_W^b$ and $\hat{C}_V^b$. At least one further auxiliary assumption would be needed to overcome this. For example, in Charney's conception of energy equipartition in three-dimensional quasi-geostrophic turbulence with isotropic statistics (after rescaling the vertical coordinate by f/N) (Charney, 1971), the vortical buoyancy spectrum is approximately equal to the longitudinal velocity spectrum, which in the present notation would imply

$$\hat{C}_V^b(k) = \hat{C}_V^u(k) = D_V^\psi(k).\tag{3.45}$$

Combining this with (3.44) and (3.42) yields

$$E_V(k) = D_V^\psi(k) - \frac{k}{2}\frac{\mathrm{d}}{\mathrm{d}k}D_V^\psi(k),\tag{3.46}$$

which is an ODE for D_V^ψ in terms of the known E_V. This is again easily solved for $D_V^\psi(k)$ by starting with a zero value at $k = +\infty$ and solving backwards in k for all $k \geq 0$. Thereafter $D_W^\psi = D^\psi - D_V^\psi$ is known as well and hence all fields have been completely decomposed into their wave and vortex constituents. This is an attractive theoretical result, but it must be clearly noted that the underlying heuristic assumption (3.45) based on isotropy in rescaled coordinates may in practice hold only for the smallest scales of the quasi-geostrophic flow, if it holds at all. We will not use this approach in this chapter.

3.2.6 Frequency models for the wave spectrum

Substantial progress can be made if one assumes a more detailed model for the frequency content of the wave spectrum. For example, if the wave field is narrow-banded in frequency then it makes sense to assign a single typical value to the parameter f^2/ω^2 that appears in relationships such as (3.33). If we denote this constant value by f^2/ω_0^2 then (3.33) implies the spectral relationships

$$\hat{C}_W^\psi(k,l) = \frac{f^2}{\omega_0^2}\hat{C}_W^\phi(k,l) \quad \text{and} \quad D_W^\psi(k) = \frac{f^2}{\omega_0^2}D^\phi(k). \tag{3.47}$$

With D_W^ψ and $D_V^\psi = D^\psi - D_W^\psi$ now in hand we can evaluate all the terms in (3.41) and (3.42). Moreover, $\hat{C}_W^b$ follows from the known E_W as

$$\hat{C}_W^b(k) = 2E_W(k) - \hat{C}_W^u(k) - \hat{C}_W^v(k). \tag{3.48}$$

This narrow-band approach for the wave spectrum might be relevant for inertia–gravity wave fields dominated by specific tidal components such as the M2 tide, but such internal tides also tend to be quite anisotropic.

A more flexible modeling approach extends the second equation in (3.47) to a function $\omega_*(k)$, say, which is *defined* by

$$D_W^\psi(k) = \frac{f^2}{\omega_*^2(k)}D^\phi(k). \tag{3.49}$$

Clearly, if $\omega_*(k)$ is somehow known then one can again compute D_W^ψ from the observed D^ϕ and hence again obtain the complete decomposition of the spectra into their wave and vortex constituents in (3.41–3.42). For ease of reference, in terms of D^ψ and D^ϕ these relations become

$$\hat{C}_W^u(k) = \frac{f^2}{\omega_*^2(k)}D^\phi(k) - k\frac{\mathrm{d}}{\mathrm{d}k}D^\phi(k), \quad \hat{C}_W^v(k) = -k\frac{\mathrm{d}}{\mathrm{d}k}\left[\frac{f^2}{\omega_*^2(k)}D^\phi(k)\right] + D^\phi(k) \tag{3.50}$$

$$\hat{C}_V^u(k) = D^\psi(k) - \frac{f^2}{\omega_*^2(k)}D^\phi(k), \quad \hat{C}_V^v(k) = -k\frac{\mathrm{d}}{\mathrm{d}k}\left[D^\psi(k) - \frac{f^2}{\omega_*^2(k)}D^\phi(k)\right]. \tag{3.51}$$

Again, $\hat{C}_W^b$ can then be computed from (3.48), which provides a "sight unseen" prediction for the buoyancy spectrum.

Of course, this only works if the function $\omega_*(k)$ is known by some method, for example by using a model spectrum for the wave field to evaluate the definition of $\omega_*(k)$ that follows from the exact three-dimensional spectral relationship between $\hat{C}_W^\psi$ and $\hat{C}_W^\phi$ in (3.33), namely

$$\frac{f^2}{\omega_*^2(k)} = \frac{D_W^\psi(k)}{D^\phi(k)} = \frac{\int\int \frac{f^2}{\omega^2}l^2\hat{C}^\phi(k,l,\omega)\,dl\,d\omega}{\int\int l^2\hat{C}^\phi(k,l,\omega)\,dl\,d\omega}. \tag{3.52}$$

For example, in the next section we will use the standard Garrett–Munk (GM) spectrum (e.g. Garrett and Munk, 1979) for ocean inertia–gravity waves as a basis for modeling the function $\omega_*(k)$. Notably, $\omega_*(k)$ as defined by (3.52) is insensitive to the overall amplitude of the wave spectrum. Alternatively, if the observed spectra are dominated by waves in some wavenumber band (i.e. $D^\psi \approx D_W^\psi$ there), then it is possible to estimate $\omega_*(k)$ in that wavenumber band directly from its definition (3.49), with D_W^ψ approximated by D^ψ. This can provide a useful check on the validity of any assumed model spectrum, as we shall see in Section 3.3.1 below.

3.2.7 Summary of the Helmholtz decomposition method

We can summarize the theoretical results as follows. Observations of the longitudinal and transverse velocity spectra $\hat{C}^u$ and $\hat{C}^v$ allow a unique Helmholtz decomposition in terms of the functions D^ϕ and D^ψ. Moreover, the function D^ϕ then implies the energy spectrum E_W for hydrostatic inertia–gravity waves, and this implication is unaffected by the presence of a quasi-geostrophic vortex flow component, provided only that this component is uncorrelated with the wave component. Additional observations of $\hat{C}^b$ allow computation of the vortex energy spectrum E_V as well, which can then be compared with the wave energy spectrum E_W. Finally, by making additional assumptions either about the structure of the balanced flow or about the frequency content of the wave field, it becomes possible to decompose all measured fields into their wave and vortex components.

3.3 Application to oceanic data sets

We now illustrate how these methods can be used to decompose observed ship-track spectra in the upper thermocline of the eastern subtropical North Pacific and of the Gulf Stream region. These spectra, spanning the mesoscale (about 200–500 km) and submesoscale (about 5–200 km) ranges, are the same as those analyzed in Chapter 2, to which we refer the reader for a more in-depth discussion of the data sets as well as of mesoscale and submesoscale dynamics.

In the eastern subtropical North Pacific, both velocity and buoyancy measurements are available, which allows an exact decomposition of the energy spectra into a balanced vortex part and an inertia–gravity wave part. Moreover, by approximating the function $\omega_*(k)$ using the GM model spectrum, we can decompose all the observed fields into their balanced and wave parts. In the Gulf Stream region, on the other hand, only velocity data are available, but by assuming a GM model for $\omega_*(k)$ a decomposition of the observed fields can be achieved there, too.

The eastern subtropical North Pacific has weak mesoscale eddy activity and in Chapter 2 we argued that much of the submesoscale range is dominated by inertia–gravity

waves. In the Gulf Stream region, on the other hand, where the mesoscale eddy field is strong, the submesoscale is dominated by balanced motions down to a scale of about 20 km, where a transition to inertia–gravity waves occurs. The decompositions performed here confirm these results.

3.3.1 Eastern subtropical North Pacific

Velocity and buoyancy data were collected in successive occupations of the 140°W meridian between 25° and 35°N in January and February 1997 (cf. Ferrari and Rudnick, 2000; Chapter 2). We use shipboard acoustic Doppler current profiler (ADCP) data from four transects at 200 m depth (8 m depth bin), which is below the base of the mixed layer. Buoyancy data are obtained from a transect of a CTD-equipped SeaSoar programmed to stay at constant depth 200 m. An average stratification of $N = 8.7 \times 10^{-3}$ s^{-1} is obtained from a transect with the SeaSoar following a sawtooth profile. We interpolate both velocity and buoyancy data onto a 3 km regular grid, which is about the spacing of the raw velocity data averaged over 12 min bins, and rotate the velocities into a frame of reference aligned with the ship track. Spectra are obtained by dividing all transects into three segments with 50% overlap, applying a Hann window to each segment, computing the discrete Fourier transform, and averaging over all transforms and over 10 wavenumber bins per decade.

The resulting $\hat{C}^u$, $\hat{C}^v$, and $\hat{C}^b$ show relatively low mesoscale energies and are relatively flat in the submesoscale range (Fig. 3.1a). Using equations (3.30–3.31), we perform the Helmholtz decomposition into rotational and divergent components (Fig. 3.1b). The rotational component dominates at scales larger than 200 km, below which the divergent component dominates. From the velocity spectra $\hat{C}^u$ and $\hat{C}^v$, we can diagnose the total energy due to inertia–gravity waves E_W using (3.37). The diagnosed E_W matches the observed total energy E remarkably well below 100 km (Fig. 3.1c). This indicates that the observed signal is consistent with an isotropic, hydrostatic inertia–gravity wave field in this range. At scales larger than 100 km there is a substantial balanced component.

We can exploit the fact that the range below 100 km is dominated by inertia–gravity waves to estimate the frequency content of the wave field expressed by ω_*. The ratio $D^\psi/D^\phi \approx f^2/\omega_*^2$ roughly follows the GM curve between 20 and 100 km (Fig. 3.2). At larger scales the ratio becomes larger than unity, which is incompatible with inertia–gravity waves, for which $\omega^2 \geq f^2$ holds robustly. At smaller scales, the diagnosed ratio is much larger than the GM value, but in this range the effects of both the interpolation onto a regular grid and the cutoff at the Nyquist wavenumber contaminate the estimate. At 100 km, the diagnosed ratio drops below the GM curve and has a value close to that of a monochromatic M$_2$ tidal wave ($f^2/\omega_0^2 = 0.27$ at 30°N), which may be interpreted as evidence for a significant tidal signal at this scale. Nevertheless using the GM curve to perform the decomposition (3.48, 3.50, 3.51), we find that the balanced components $\hat{C}_V^u$, $\hat{C}_V^v$, and $\hat{C}_V^b$ match the observed spectra $\hat{C}^u$, $\hat{C}^v$, and $\hat{C}^b$ at scales larger than 200 km

Figure 3.1: Observations from the eastern subtropical North Pacific: (a) observed transverse and longitudinal kinetic energy and potential energy spectra $\hat{C}^u$, $\hat{C}^v$, and $\hat{C}^b$, (b) decomposition into rotational and divergent components D^ψ and D^ϕ from (3.30–3.31) and (3.27); here $K^\psi = \frac{1}{2}(D^\psi - k\mathrm{d}D^\psi/\mathrm{d}k)$ and $K^\phi = \frac{1}{2}(D^\phi - k\mathrm{d}D^\phi/\mathrm{d}k)$, (c) total observed energy E and total inertia–gravity wave energy E_W from (3.37)

Figure 3.2: Observations from the eastern subtropical North Pacific: ratio D^ψ/D^ϕ compared to ω_* from Garrett–Munk (GM) spectrum and M_2 value for reference

(Fig. 3.3a) and that the inertia–gravity waves components $\hat{C}_W^u$, $\hat{C}_W^v$, and $\hat{C}_W^b$ match the observed spectra $\hat{C}^u$, $\hat{C}^v$, and $\hat{C}^b$ at scales smaller than 100 km (Fig. 3.3b).
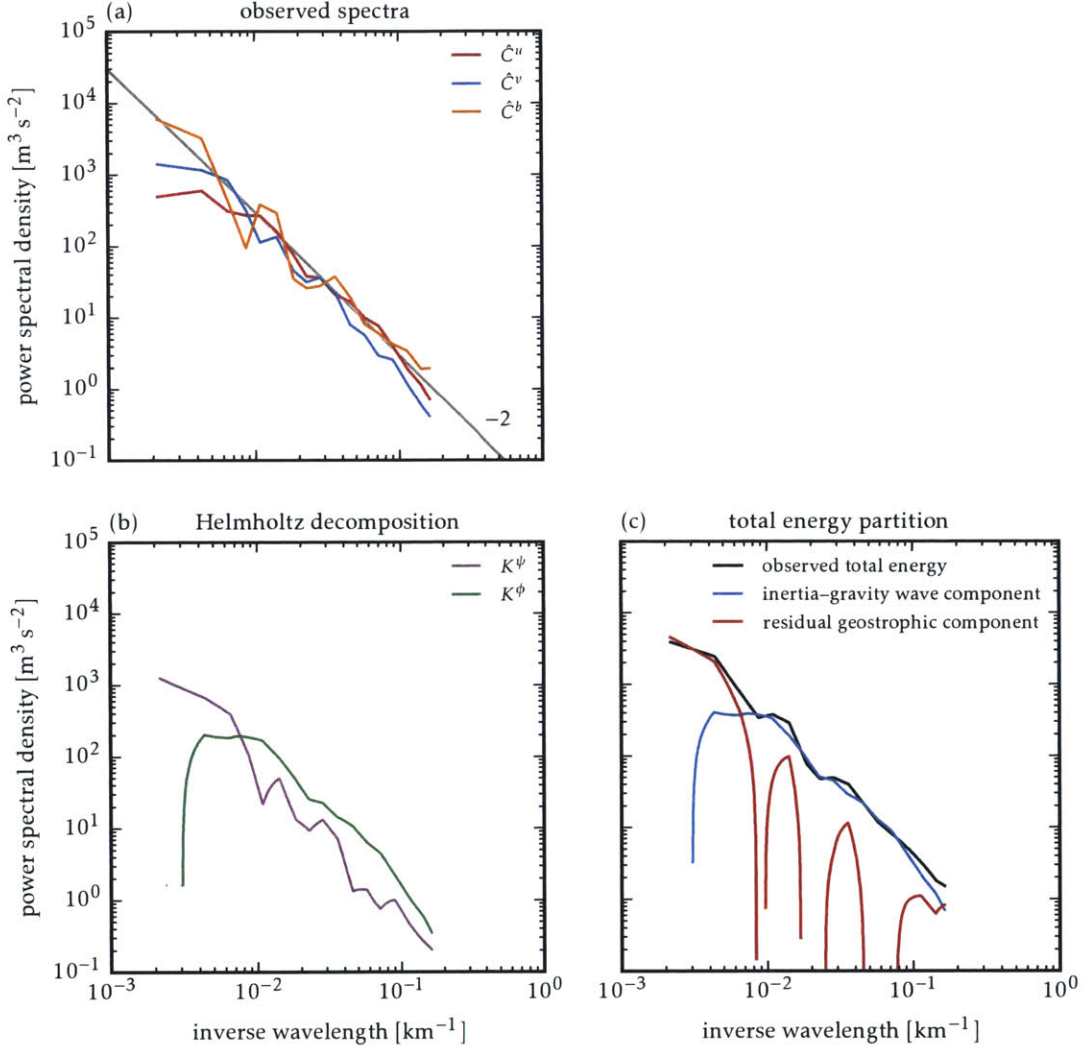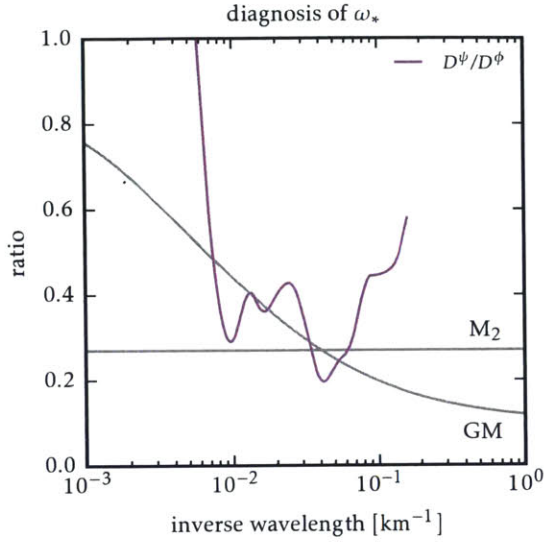
In summary, kinetic and potential energy spectra are dominated by balanced eddies at scales of 200 km and larger. This is consistent with our finding in Chapter 2 that at these scales, the in situ $\hat{C}^v$ also matches that obtained from applying geostrophic balance to along-track altimetric measurements of sea surface height. At scales smaller than 100 km, the energy spectra are dominated by inertia–gravity waves. The inertia–gravity wave field may have a substantial tidal component.

3.3.2 Gulf Stream region

Velocity data in the Gulf Stream region were collected using a 150 kHz shipboard ADCP on repeat transects from New York Harbor to Bermuda between 1994 and 2004 (cf. Wang et al., 2010; Chapter 2). We use data at 150-m depth, which is below the base of the mixed layer in most of the year. We select transects that are at least 1000 km long and have at least 400 data points, resulting in a total of 306 transects. We interpolate onto a 2.5 km regular grid, which is about the spacing of the raw data averaged over 5 min bins, and rotate the velocities into a frame of reference aligned with the ship track. Spectra are obtained by the same procedure as described for the Pacific data set.

The resulting $\hat{C}^u$ and $\hat{C}^v$ show large mesoscale energies and fall off steeply in the submesoscale range (Fig. 3.4a). The spectra exhibit a conspicuous flattening at a scale of about 20 km. Using equations (3.30–3.31), we perform the Helmholtz decomposition into
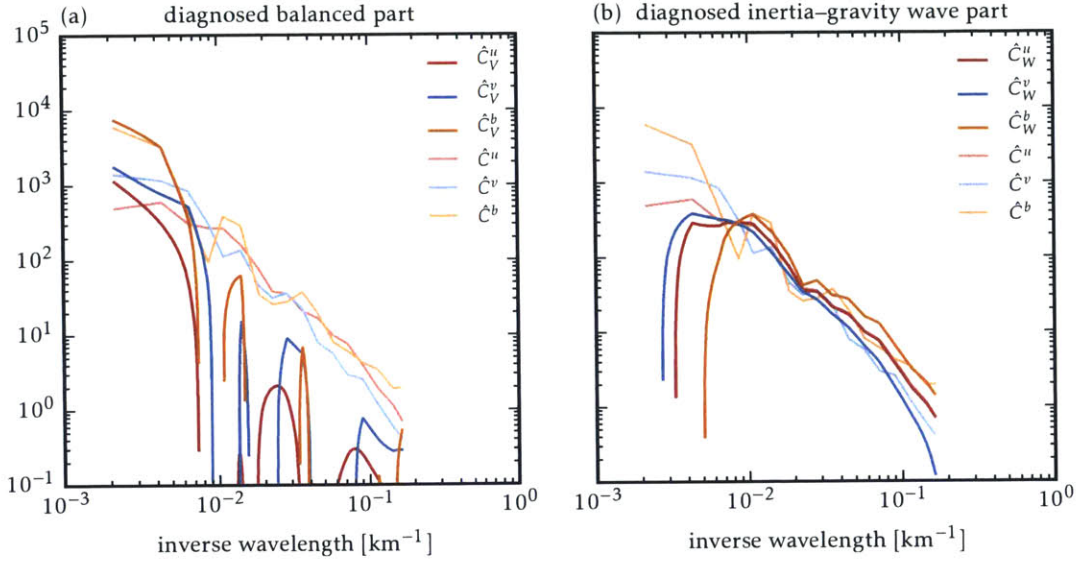
Figure 3.3: Observations from the eastern subtropical North Pacific: (a) diagnosis of the balanced components of the observed spectra $\hat{C}^u_V$, $\hat{C}^v_V$, and $\hat{C}^b_V$, (b) diagnosis of the inertia–gravity wave component of the observed spectra $\hat{C}^u_W$, $\hat{C}^v_W$, and $\hat{C}^b_W$.

rotational and divergent components (Fig. 3.1b). In contrast to the eastern Pacific case, the rotational part D^ψ here vastly dominates over a wide range of scales: only at 20 km does the divergent component D^ϕ become comparable to the rotational component D^ψ. Notably, at large scales the true D^ϕ becomes close to zero, but our computed D^ϕ actually becomes negative, which is of course unphysical. This is the numerical robustness issue discussed at the end of 3.2.2.

Since no buoyancy data are available, the only way to decompose into a balanced part and an inertia–gravity wave part is to make an assumption about the frequency content of the waves. We choose the GM curve to perform the decomposition, since the GM empirical spectrum is largely based on observations collected nearby the North Atlantic region considered here. The diagnosed balanced components $\hat{C}^u_V$ and $\hat{C}^v_V$ show good agreement with the observed spectra $\hat{C}^u$ and $\hat{C}^v$ in the range 50–200 km (Fig. 3.4c). At larger scales, the reconstruction overestimates the longitudinal component $\hat{C}^u$. This is likely the effect of anisotropy in the flow: the Gulf Stream, running mostly transverse to the ship track, has scales of a few hundred kilometers and renders the geostrophic flow field highly anisotropic cf. Wortham et al., 2014, violating the isotropy assumption made throughout the development of the theory. At the small-scale end, the balanced components $\hat{C}^u_V$ and $\hat{C}^v_V$ start deviating from the observed spectra $\hat{C}^u$ and $\hat{C}^v$ at 20–50 km. This scale is coincident with the scale at which the observed spectra $\hat{C}^u$ and $\hat{C}^v$ start flattening out. The balanced components $\hat{C}^u_V$ and $\hat{C}^v_V$ show no sign of a transition and keep falling off

65
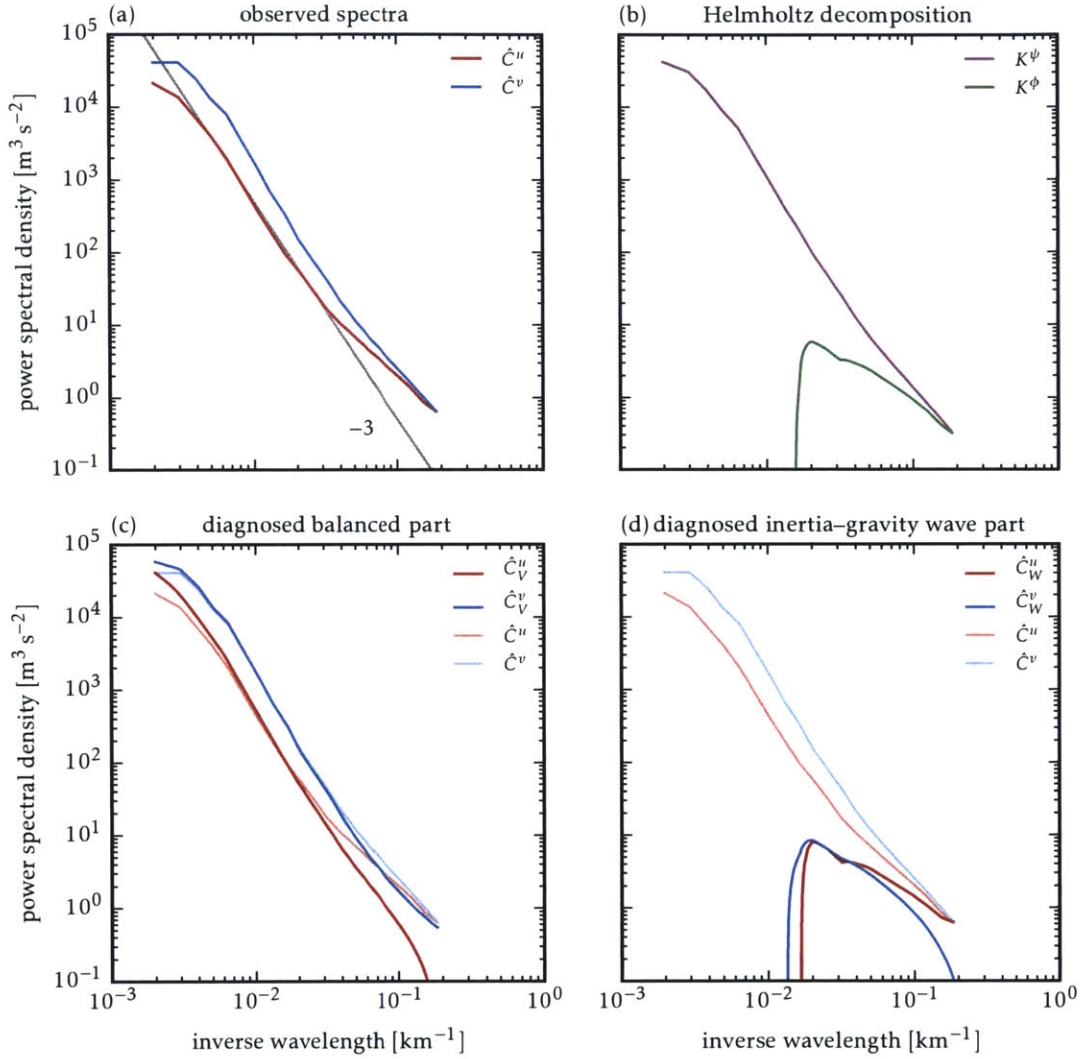
Figure 3.4: Observations from the Gulf Stream region: (a) observed transverse and longitudinal kinetic energy spectra $\hat{C}^u$ and $\hat{C}^v$, (b) decomposition into rotational and divergent components D^ψ and D^ϕ from (3.30–3.31) and (3.27); here $K^\psi = \frac{1}{2}(D^\psi - k\mathrm{d}D^\psi/\mathrm{d}k)$ and $K^\phi = \frac{1}{2}(D^\phi - k\mathrm{d}D^\phi/\mathrm{d}k)$, (c) diagnosis of the balanced components of the observed spectra $\hat{C}_V^u$ and $\hat{C}_V^v$, (d) diagnosis of the inertia–gravity wave component of the observed spectra $\hat{C}_W^u$ and $\hat{C}_W^v$.

steeply. The diagnosed inertia–gravity wave components $\hat{C}_W^u$ and $\hat{C}_W^v$ start contributing substantially at this scale (Fig. 3.4d), indicating that the flattening of the spectra $\hat{C}^u$ and $\hat{C}^v$ is due to a transition to a range dominated by inertia–gravity waves.

In summary, the decomposition into a balanced part and an inertia–gravity wave part shows that in the Gulf Stream region, in contrast to the eastern Pacific case, much of the submesoscale range is dominated by balanced flows. The transition in the slope of the observed spectra at about 20 km is due to inertia–gravity waves.

3.4 Concluding comments

Our proposed method is very easy to implement and gave robust results when applied to real oceanic data sets. Of course, there are restrictions inherent in the derivation of our method, which one needs to bear in mind. The assumptions of stationarity, homogeneity, and horizontal isotropy will have obvious limitations in any practical situation. Another obvious restriction that is hindering in practice is that we needed to assume that the wave field was homogeneous in the vertical in order to exploit the energy equipartition result (3.36) for the wave diagnostic. Strictly speaking, this step disallows considering standing normal modes in the vertical, which is a relevant case for large-scale ocean inertia–gravity waves such as low-mode internal tides. So it would be nice to have an extension of our method that covers that case.

Slightly less obvious will be the impact of nonlinear effects, as our method relied on geostrophic balance and linear waves dynamics. For example, the assumption of uncorrelated ψ and ϕ, which underlies both steps in our method, presumably would fail for vortex motions with a noticeable divergent component (the so-called "ageostrophic" velocities of quasi-geostrophic theory). Submesoscale balanced flows tend to develop a large divergence close to the ocean surface through frontogenesis (Chapter 2). It is not obvious what this would entail for the correlations between the rotational and divergent parts of the balanced velocity field.

Similarly, nonlinear effects would also modify the linear argument for the uncorrelated ψ and ϕ for the wave field. That argument can be viewed as a statement of zero linear PV, or of stretching of background vorticity, in the sense of $v_x - u_y = -f b_z/N^2$. In nonlinear theory one could imagine that nonlinear terms in the definition of PV become important, and that the stretching of background vorticity might reasonably contain a contribution from the vorticity of the geostrophic flow. A model equation for the wave vorticity to study in a weakly nonlinear regime might be

$$(\partial_{xx}^2 + \partial_{yy}^2)\partial_t \psi_W = \left[f + (\partial_{xx}^2 + \partial_{yy}^2)\psi_V \right](\partial_{xx}^2 + \partial_{yy}^2)\phi_W. \tag{3.53}$$

Compared to (3.32) this includes a nonlinear coupling term between ψ_V and the wave-related ψ_W and ϕ_W. Of course, in a strongly nonlinear regime one would also encounter

67

stratified turbulence, a broad subject heading describing three-dimensionally turbulent flows that are significantly modified by, and interact with, the stable stratification (e.g. Smith and Waleffe, 2002; Waite and Bartello, 2004; Riley and Lindborg, 2008). It is an open question whether such stratified turbulence, perhaps also modified by background rotation, would leave a detectable fingerprint in the one-dimensional spectra that would allow distinguishing it from the spectrum due to nearly linear waves, for example.

Finally, our method should naturally be applicable to the substantial body of commercial air flight track data that has been collected near the tropopause in the atmosphere, where attention has been focused for some time on disentangling the dynamical processes that underlie the conspicuous spectral slopes that are observed in the so-called Gage–Nastrom spectrum (e.g. Nastrom and Gage, 1985; Lindborg, 1999; Tulloch and Smith, 2006). We are hoping to report on the results of this application in the near future.

3.A   Horizontally uncorrelated wave stream function and potential

We seek to demonstrate that

$$(3.21): \quad \mathbb{E}[\psi(x_0, y_0, z_0, t_0)\phi(x_0 + x, y_0 + y, z_0, t_0)] = 0$$

holds if the expectation is taken over a stationary, homogeneous, and horizontally isotropic wave spectrum, and if ψ and ϕ are wave fields related by a time derivative, as in the relation $\psi_t = -f\phi$ that holds for inertia–gravity waves. Here we rewrote (3.21) in a form that highlights that the covariance is taken at different horizontal positions but equal altitude z_0 and time t_0. By stationarity, the statement (3.21) is trivially true in the case of $x = y = 0$, but not otherwise.

We illustrate this issue using a slightly simpler model, namely the rotating shallow-water in one spatial dimension x, say, in which horizontal isotropy is replaced by left–right symmetry. The standard variables are (h, u, v) and the linear equations are

$$u_t - fv + gh_x = 0, \quad v_t + fu = 0, \quad \text{and} \quad h_t + Hu_x = 0, \tag{3.54}$$

where g is gravity and H is the basic layer depth. Note that u and v are in quadrature because of the time derivative linking the two fields in (3.54), just as ϕ and ψ were in the inertia–gravity wave case. The fields depend on x and t and we assume stationarity and homogeneity. We then like to find out whether

$$\mathbb{E}[u(x_0, t_0)v(x_0 + x, t_0)] = 0 \tag{3.55}$$

holds for left–right symmetric waves. Now, it is easy to construct a homogeneous and stationary spectrum for which (3.55) fails. Specifically, consider this spectrum of plane-wave solutions corresponding to right-going waves only:

$$u = \beta \cos(k[x - ct] - \alpha), \quad v = \frac{f}{ck}\beta \sin(k[x - ct] - \alpha), \quad \text{and} \quad h = \frac{H}{c}\beta \cos(k[x - ct] - \alpha). \tag{3.56}$$

Here $k > 0$ is a wavenumber, $c = \omega/k > 0$ is the phase speed corresponding to the positive root of the dispersion relation $\omega^2 = f^2 + gHk^2$, $\alpha \in [0, 2\pi]$ is a uniformly distributed random phase shift and β is a random zero-mean amplitude that is independent of α and has variance σ^2, say. It follows that

$$\mathbb{E}[u(x_0, t_0)v(x_0 + x, t_0)] = \frac{f}{ck}\sigma^2 \sin kx, \tag{3.57}$$

where it was crucial to take the expectation over the random phase shift α. Obviously, for this random wave field (3.55) failed. However, if we impose left–right symmetry on the random wave field then we must augment (3.56) by a second wave with identical and uncorrelated statistics, but going in the opposite direction. This corresponds to setting $c = -\omega/k < 0$ and obviously uses the other branch of the dispersion relation. This leads to a second term in (3.57) with equal-and-opposite sign, which cancels the first term, and therefore (3.55) is indeed satisfied.

The conclusion is that for wave systems with equal-and-opposite frequency branches in the dispersion relation a horizontally isotropic stationary random wave field has the property that (3.21) is guaranteed to hold if the two fields ψ and ϕ are related by a simple time derivative and hence in quadrature. This is the case for the rotating Boussinesq equations as well as for the rotating shallow-water equations and therefore holds quite generally for gravity waves.

3.B   Energy equipartition statement (3.36) for plane inertia–gravity waves

We derive

$$(3.36): \quad \hat{C}_W^b(k) + \left[D_W^\psi(k) - k\frac{\mathrm{d}}{\mathrm{d}k}D_W^\psi(k) \right] = \left[D_W^\phi(k) - k\frac{\mathrm{d}}{\mathrm{d}k}D_W^\phi(k) \right]$$

for a spectrum of stationary, homogeneous, and horizontally isotropic three-dimensional plane inertia–gravity waves. For such waves the wavenumber vector (k, l, m) and the intrinsic frequency ω are related by the hydrostatic dispersion relation of the Boussinesq system with constant f and N:

$$\omega^2 = f^2 + N^2 \frac{k_h^2}{m^2}. \tag{3.58}$$

It suffices to derive the following equivalent statement for three-dimensional spectra, from which (3.36) then follows by integration over l and m:

$$\hat{C}_W^b(k, l, m) + k_h^2 \hat{C}_W^\psi(k, l, m) = k_h^2 \hat{C}_W^\phi(k, l, m). \tag{3.59}$$

For a plane wave with frequency ω the linear buoyancy equation $b_t + N^2 w = 0$ and the continuity equation $w_z = -(u_x + v_y)$ imply

$$\omega^2 \hat{C}_W^b = N^2 \hat{C}_W^w \quad \text{and} \quad m^2 \hat{C}_W^w = k_h^4 \hat{C}_W^\phi. \tag{3.60}$$

69

Combining this with $\omega^2 \hat{C}_W^\psi = f^2 \hat{C}_W^\phi$ from (3.33) and substituting in (3.59) yields

$$\left(\frac{N^2}{\omega^2} \frac{k_h^2}{m^2} + \frac{f^2}{\omega^2} - 1 \right) k_h^2 \hat{C}_W^\phi(k,l) = 0. \qquad (3.61)$$

This holds by (3.58), which therefore establishes (3.59) and hence (3.36) after integration over l and m.

This calculation does depend on the assumption of plane waves. For example, if standing vertical modes were considered instead, then the result would not hold, because the spectrum is then not homogeneous in the vertical. Mathematically, this is because the second equation in (3.60) does not hold for standing waves. Physically, this occurs because at nodal horizontal planes, where $b = w = 0$ at all times, there is no potential energy at all and equipartition fails even for non-rotating waves.

CHAPTER 4

TRANSITION FROM GEOSTROPHIC TURBULENCE TO INERTIA–GRAVITY WAVES IN THE ATMOSPHERIC ENERGY SPECTRUM

The mid-latitude high and low pressure systems visible in weather maps are associated with winds and temperature fluctuations that we experience as weather. These fluctuations arise from a baroclinic instability of the mean zonal winds at horizontal scales of a few thousand kilometers, commonly referred to as the synoptic scales (Charney, 1947; Eady, 1949; Phillips, 1954). The combined effects of rotation and stratification constrain the synoptic-scale winds to be nearly horizontal and to satisfy geostrophic balance, a balance between the force exerted by the changes in pressure and the Coriolis force resulting from Earth's rotation. It is an open question whether the same constraints dominate in the mesoscale range, i.e. at scales of 10–500 km, or whether qualitatively different dynamics govern flows at these scales.

The synoptic-scale flows are turbulent in the sense that nonlinear scale interactions, which lie at the core of the difficulty to predict the weather, exchange energy between different scales of motion (Batchelor, 1953; Lorenz, 1963, 1969; Vallis, 2006). Under the constraints of rotation and stratification, the synoptic-scale winds are approximately two-dimensional and nondivergent (Rossby, 1939; Charney, 1971). In two-dimensional flows, nonlinear scale interactions tend to transfer energy to larger scales, i.e. the synoptic-scale pressure anomalies often merge and form larger ones, contrary to nonlinear scale interactions in three-dimensional flows, which tend to transfer energy to smaller scales (Fjørtoft, 1953). Little energy is thus transferred to scales smaller than those at which the synoptic-scale fluctuations are generated through instabilities. Theory and numerical simulations predict that the energy per unit horizontal wavenumber k decays as rapidly as k^{-3} at wavenumbers larger than the wavenumber corresponding to the instability scale (Charney, 1971; Kraichnan, 1967). This predicted kinetic energy spectrum is roughly consistent with synoptic-scale observations (Charney, 1971; Leith, 1971).

Long-range passenger aircraft have been instrumented to collect velocity and temperature measurements as part of the Global Atmospheric Sampling Program in the 1970s and the Measurement of Ozone and Water Vapor by Airbus In-Service Aircraft (MOZAIC) project in the 1990s and 2000s. The resulting dataset, described in the materials section, consists of tens of thousand of flights. Because aircraft travel at altitudes between 9 and 14 km, the data largely reflect the upper troposphere and lower stratosphere, near the tropopause. These measurements confirm that the kinetic energy spectrum drops as k^{-3} in the synoptic wavenumber range, but there is a transition in behavior at a scale of about 500 km (Nastrom and Gage, 1985) (cf. Fig. 4.1a). In the mesoscale range, at scales smaller than 500 km, the kinetic energy spectrum decays more slowly, roughly like $k^{-5/3}$ (Nastrom and Gage, 1985; Gage, 1979; Cho and Lindborg, 2001).

The measured kinetic energy spectrum is intriguing, because it agrees so well with Charney's theory of geostrophic turbulence at the synoptic scales (Charney, 1971), but deviates from that prediction at the mesoscale. The transition to the flatter $k^{-5/3}$ mesoscale spectrum has been interpreted as the signature of small-scale geostrophic flows generated by convective events (Gage, 1979; Lilly, 1983; Vallis et al., 1997), as the development of fronts at the edge of synoptic-scale cyclones and anticyclones at the top of the troposphere

74
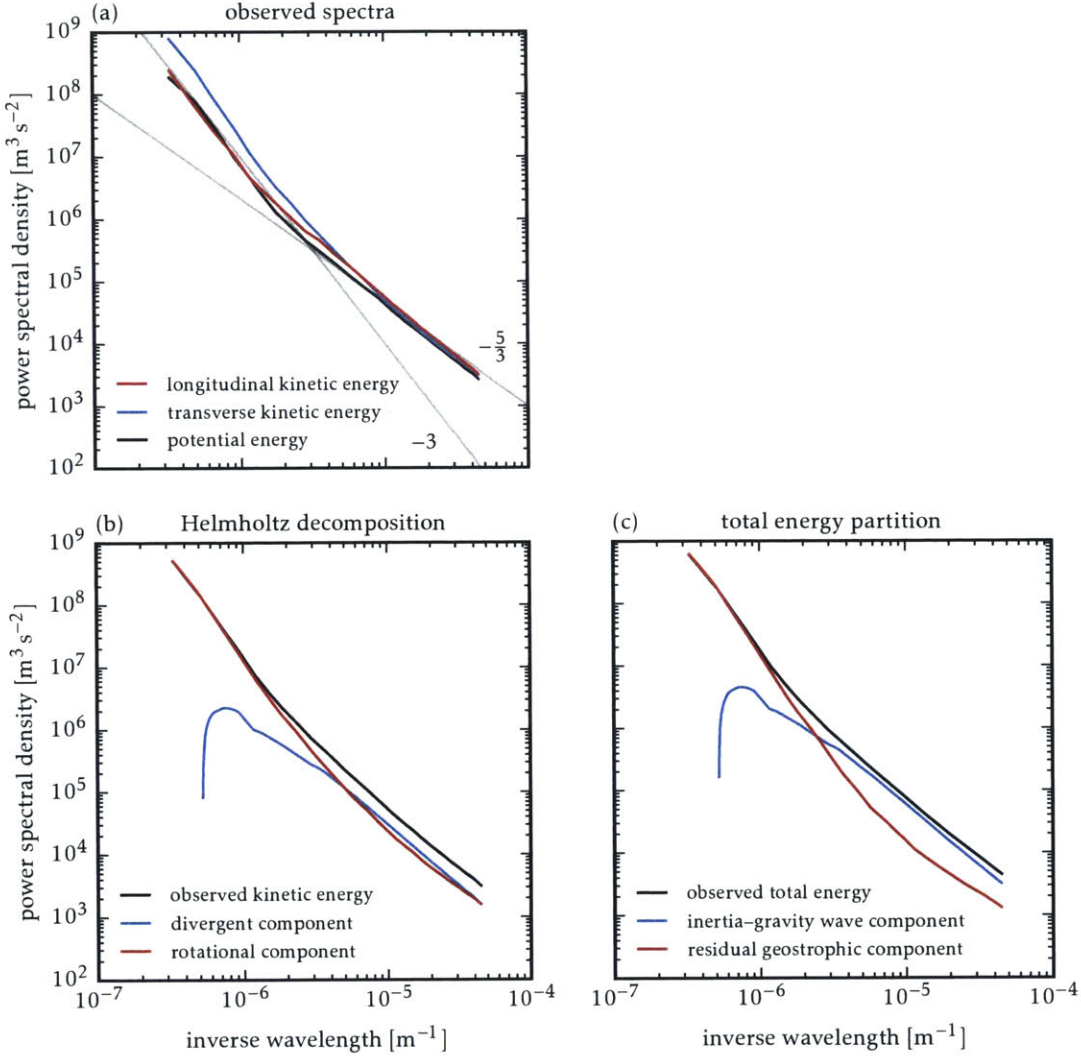
Figure 4.1: Observed wavenumber spectra of near-tropopause midlatitude winds and decomposition into geostrophic component and inertia–gravity wave component. (a) Observed spectra of longitudinal kinetic energy $S^u(k)$, transverse kinetic energy $S^v(k)$, and potential energy $S^b(k)$. (b) Helmholtz decomposition of the observed kinetic energy spectrum $K(k)$ into its rotational and divergent components $K^\psi(k)$ and $K^\phi(k)$. (c) Partitioning of the total energy spectrum $E(k)$ into the diagnosed inertia–gravity wave component $E_w(k)$ and the residual geostrophic component $E_g(k)$.

(equivalent to the warm and cold mesoscale fronts we experience at the Earth's surface) (Tulloch and Smith, 2006), or as the signature of stratified turbulence at scales where the rotational constraints become less important (Lindborg, 2006). These explanations of the synoptic-to-mesoscale transition invoke turbulent dynamics and strong interactions between the synoptic and mesoscale flows.

A rotating and stratified atmosphere, however, supports an additional, much faster set of motions: inertia–gravity waves. These are internal gravity waves, modified by the effect of rotation, that have periods of several minutes to a few hours. In contrast to the strongly nonlinear, turbulent synoptic-scale flow, these motions are wave-like and at small amplitude they are approximately governed by linear dynamics (Pedlosky, 2010). It has been proposed that the mesoscale energy is dominated by inertia–gravity waves (Dewan, 1979; VanZandt, 1982), which are easily excited by any fast fluctuation of the atmospheric flows (Fritts and Alexander, 2003). In this explanation of the mesoscale part of the spectrum, linear inertia–gravity and nonlinear synoptic-scale turbulence coexist with little interaction.

In this chapter, we present an analysis of the MOZAIC data that utilizes the decomposition method developed in Chapter 3. For the first time, this new analysis provides compelling evidence that linear inertia–gravity waves indeed dominate the observations in the mesoscale.

4.1   Theories for the synoptic-to-mesoscale transition

Dewan first suggested that the mesoscale energy is dominated by a continuum of weakly nonlinear inertia–gravity waves (Dewan, 1979). VanZandt showed that the mesoscale spectra of horizontal wind fluctuations as a function of horizontal wavenumber, vertical wavenumber, and frequency are related through the dispersion and polarization relations of inertia–gravity waves (VanZandt, 1982). The dominance of inertia–gravity waves at the mesoscale, however, appeared inconsistent with vertical velocity frequency spectra measured from radar in light-wind conditions (Gage and Nastrom, 1985). More recently Vincent and Eckerman showed that the signature of inertia–gravity waves is recovered after correcting for Doppler-shift effects in radar observations (Vincent and Eckermann, 1990).

In the last two decades, the interpretation of the synoptic-to-mesoscale transition in terms of inertia–gravity wave dynamics has received little attention. The $k^{-5/3}$ power-law dependence of the mesoscale spectrum has instead been interpreted as evidence that the mesoscale is strongly turbulent. The strong nonlinear interactions characteristic of turbulent flows continuously redistribute energy across wavenumbers and are known to result in power-law energy distributions (Batchelor, 1953; Vallis, 2006). A number of competing turbulent theories have been proposed.

The earliest turbulent theory argues that the mesoscale spectrum is due to an inverse

cascade of energy injected at even smaller scales, for example by convective activity (Gage, 1979; Lilly, 1983; Vallis et al., 1997). The hypothesis is that nonlinear mesoscale interactions, much like at synoptic scales, transfer energy to larger scales because of the constraints of rotation and stratification. The kinetic energy spectrum in such a quasi-two-dimensional inverse cascade scales like $k^{-5/3}$ at wavenumbers smaller than the injection scale (Kraichnan, 1967), in contrast to Charney's k^{-3} spectrum, which develops at wavenumbers larger than the injection scale. The theory does not predict at which scale the dynamics switch from the synoptic to the mesoscale regime.

The second turbulent theory proposes that the flat mesoscale spectrum is the signature of sharp temperature fronts, which develop when synoptic-scale flows intersect a rigid boundary, like the Earth's surface, or a strongly stratified layer, like the tropopause (Tulloch and Smith, 2006). Importantly, the winds associated with the temperature fronts are still largely in geostrophic balance. In this view, the $k^{-5/3}$ mesoscale spectrum is a feature of measurements taken at the tropopause level, the cruising altitude of long range commercial aircraft, but should not appear in the mid troposphere. This prediction cannot be tested with available data.

The third turbulent theory proposes that the $k^{-5/3}$ mesoscale spectrum emerges at the scales where the flows escape the rotational constraint and energy can be transferred to smaller scales (Lindborg, 2006). Turbulent flows constrained by stratification, but not rotation, are collectively known as "stratified turbulence". The forward energy cascade is achieved by the overturning of layer-like structures. These flows are not in geostrophic balance, and thus differ from the quasi-two-dimensional dynamics of the previous two theories; and they are strongly nonlinear, and thus differ from approximately linear inertia–gravity waves.

4.2   Inertia–gravity waves and geostrophic flow

The understanding that atmospheric winds are composed of slow flows in approximate geostrophic balance and fast inertia–gravity waves has been the foundation for much progress in atmospheric science. The first numerical weather predictions were based on quasi-geostrophic dynamics, an approximation to the more complete primitive equations that filters out inertia–gravity waves (Charney, 1949; Charney and Eliassen, 1949; Charney et al., 1950). In the troposphere and lower stratosphere, inertia–gravity waves typically have small amplitudes and therefore interact only weakly with geostrophic flows. While there is a growing appreciation of rare instances of inertia–gravity waves directly influencing sensitive weather patterns (Ruppert and Bosart, 2014), strong interactions between inertia–gravity waves and the geostrophic flow are typically confined to the middle and upper atmosphere, where the wave amplitude becomes large enough to allow for breaking of inertia–gravity waves and the concomitant drag force on the geostrophic flow that is well known to be crucial for the global angular momentum budget of the

atmosphere (Fritts and Alexander, 2003).

In the deep ocean, breaking inertia–gravity waves mix heat and carbon. This leading-order effect has led to intensive study of the oceanic inertia–gravity wave field. It is composed of a continuous spectrum of linear waves together with isolated peaks at the inertial and tidal frequencies (Garrett and Munk, 1979). Similar to the lower atmosphere, these linear waves interact only weakly with the geostrophic flow. Only at small vertical scales of a few tens of meters do the waves break.

In Chapters 2 and 3, we showed that the energy spectra of oceanic flows are dominated by geostrophic flows at large scales and by inertia–gravity waves at small scales. The transition between the two classes of motion occurs at scales of 10–100 km, depending on the relative strength of the geostrophic eddies and the waves. In what follows, we show that the synoptic-to-mesoscale transition in the atmospheric kinetic energy spectrum is likely an equivalent transition from geostrophic to inertia–gravity wave dynamics.

4.3   Decomposition

From the MOZAIC aircraft observations of wind near the midlatitude troposphere, we compute power spectra of the longitudinal (along-track) velocity u and transverse (across-track) velocity v, $S^u(k) = \langle |\hat{u}(k)|^2 \rangle$ and $S^v(k) = \langle |\hat{v}(k)|^2 \rangle$, where the caret denotes a Fourier transform and the angle brackets an average over flights. From the temperature observations, we compute the potential energy spectrum $S^b(k) = \langle |\hat{b}(k)|^2 \rangle / N^2$, where $b = g(\theta - \theta_0)/\theta_0$ is buoyancy, $g = 9.81$ m s^{-2} is the gravitational acceleration, $\theta_0 = 340$ K is the reference potential temperature, and N^2 is the average vertical gradient of b. Potential temperature is the temperature of an air parcel corrected for dynamically irrelevant compression effects. We use a typical stratification of the lower stratosphere, $N = 0.02$ s^{-1}, estimated from the ERA-Interim reanalysis (Dee et al., 2011). Fig. 4.1a shows that these MOZAIC spectra display the transition from a steep synoptic range to a flat mesoscale range at about 500 km.

If simultaneous wind and temperature observations were available in space and time, one could directly test whether the dispersion and polarization relations of inertia–gravity waves are satisfied by mesoscale motions. One could further separate out inertia–gravity waves and geostrophic flows, because inertia–gravity waves are restricted to frequencies between the Coriolis frequency f (equal to twice the rotation rate of the Earth multiplied by the sine of the latitude) and the buoyancy frequency N (the frequency at which a vertically displaced parcel of air will oscillate within the stably stratified atmosphere), while geostrophic flows evolve on much longer time scales. But it is extremely difficult to collect simultaneous measurements of mesoscale fluctuations of winds and temperature in space and time.

We showed in Chapter 3 that the decomposition can be achieved from space-only measurements, provided that concurrent observations of horizontal velocities and tempera-

ture are available (Chapter 3). Applying this new decomposition to the MOZAIC data (see Materials and Methods) produces two powerful arguments in support of the hypothesis that the mesoscale spectrum is dominated by inertia–gravity waves.

First, assuming that the flow is an uncorrelated superposition of a geostrophic flow and inertia–gravity waves, we diagnose the inertia–gravity wave component of the total energy, the sum of the kinetic and potential energies, solely based on the observed horizontal velocities. We then show that the thus predicted inertia–gravity wave energy spectrum closely matches the observed total energy spectrum in the mesoscale range. This indicates that the mesoscale potential energy spectrum, predicted by this procedure, is consistent with linear wave theory.

Second, assuming that the geostrophic component of the flow obeys Charney's isotropy relation for geostrophic turbulence (Charney, 1971), we decompose the three observed individual spectra of longitudinal kinetic energy, transverse kinetic energy, and potential energy into a geostrophic component and an inertia–gravity wave component. The diagnosed inertia–gravity wave spectra closely match the observed spectra in the mesoscale range. This is another powerful test of the mesoscale flow's consistency with the dispersion and polarization relations of inertia–gravity waves.

4.3.1 Helmholtz decomposition

Fig. 4.1a shows that in the synoptic range, the spectra approximately satisfy Charney's prediction for geostrophic turbulence: $S^u(k) = S^b(k)$ and $S^v(k) = 3S^u(k)$ (Charney, 1971; Leith, 1971). At the transition to the mesoscale, all three spectra converge. As shown by Lindborg and in the following, this convergence is evidence that the flow is no more in geostrophic balance at leading order (Lindborg, 1999, 2007).

Any horizontal flow field can be decomposed into its rotational and divergent components: $u = -\psi_y + \phi_x$, $v = \psi_x + \phi_y$, where ψ is the streamfunction, ϕ is the velocity potential, x is the along-track coordinate, and y is the across-track coordinate. If the flow is statistically isotropic horizontally and ψ and ϕ are uncorrelated (as is the case for a superposition of geostrophic flow and linear inertia–gravity waves), the spectra $S^u(k)$ and $S^v(k)$ can be written in terms of spectral functions associated with ψ and ϕ (see Material and Methods):

$$S^u(k) = D^\psi(k) - k\frac{\mathrm{d}}{\mathrm{d}k}D^\phi(k), \tag{4.1}$$

$$S^v(k) = -k\frac{\mathrm{d}}{\mathrm{d}k}D^\psi(k) + D^\phi(k). \tag{4.2}$$

The spectral functions $D^\psi(k)$ and $D^\phi(k)$ can easily be computed from the observed $S^u(k)$ and $S^v(k)$ by solving the system of ordinary differential equations (4.1) and (4.2) (see Materials and Methods). Using $D^\psi(k)$ and $D^\phi(k)$, the observed kinetic energy spectrum

$K(k) = \frac{1}{2}[S^u(k) + S^v(k)]$ can be decomposed into its rotational and divergent components,

$$K^\psi(k) = \frac{1}{2}\left(1 - k\frac{d}{dk}\right)D^\psi(k),\tag{4.3}$$

$$K^\phi(k) = \frac{1}{2}\left(1 - k\frac{d}{dk}\right)D^\phi(k).\tag{4.4}$$

The Helmholtz decomposition of the MOZAIC kinetic energy spectrum in Fig. 4.1b shows that the rotational component dominates in the synoptic range, whereas the divergent component becomes of the same order at the transition to the mesoscale range. In the mesoscale range, the divergent component slightly dominates over the rotational component.

The dominance of the rotational component in the synoptic scales is consistent with Charney's geostrophic turbulence, because a geostrophic flow is horizontally nondivergent and thus purely rotational. The significant divergent component in the mesoscale, on the other hand, is inconsistent with the mesoscale theories that rely on a leading-order geostrophic balance, namely the inverse-cascade theory and the frontogenesis theory. Instead, it points to the dominance of ageostrophic dynamics.

Lindborg also found that the rotational and divergent components of the flow are of the same order in the mesoscale range (Lindborg, 2007). In his analysis, based on curve fitting and selective Fourier transforming, the rotational component slightly dominated the divergent component in the mesoscale range. He argued that this was inconsistent with inertia–gravity waves, for which he expected the divergent component to be much larger than the rotational component. If the inertia–gravity wave field is dominated by near-inertial waves, however, as suggested by balloon measurements in the lower stratosphere (Hertzog et al., 2002), the rotational component is expected to be of the same order as the divergent component. In the following section, we show that the mesoscale signal is indeed consistent with linear inertia–gravity wave dynamics, i.e. with the dispersion and polarization relations of hydrostatic inertia–gravity waves.

4.3.2 Decomposition of the total energy spectrum into geostrophic and inertia–gravity wave components

Geostrophic flows are horizontally nondivergent and therefore only have a rotational component, while inertia–gravity waves have both a rotational and a divergent component. To perform the decomposition into these two classes of motion, we note that the component of the total energy spectrum $E(k) = \frac{1}{2}[S^u(k) + S^v(k) + S^b(k)]$ that is due to hydrostatic inertia–gravity waves can be diagnosed from $D^\phi(k)$ alone (see Materials and Methods):

$$E_w(k) = 2K^\phi(k) = \left(1 - k\frac{d}{dk}\right)D^\phi(k),\tag{4.5}$$

where the subscript "w" designates the inertia–gravity wave component. This somewhat surprising result follows directly from linear inertia–gravity wave dynamics, if horizontal isotropy and vertical homogeneity are assumed. Provided there is no additional type of motion, the residual of the observed total energy spectrum can be attributed to a geostrophic flow: $E_g(k) = E(k) - E_w(k)$, where the subscript "g" designates the geostrophic component.

The decomposition into geostrophic flow and inertia–gravity waves of the MOZAIC data is shown in Fig. 4.1c. The mesoscale range is dominated by inertia–gravity waves, which do not contribute much energy at the synoptic scales. The residual spectrum, i.e. the total spectrum minus the inertia–gravity wave component, dominates at the synoptic scales. This component can be confidently attributed to geostrophic flows, because we have shown that the spectrum in the synoptic range is purely rotational and hence horizontally nondivergent. At the transition scale, the geostrophic component of the total energy keeps falling off steeply—the transition appears to be due to inertia–gravity waves becoming dominant in the mesoscale range. This is the main result of this chapter.

Notice that the geostrophic spectrum keeps falling off steeply past the transition at 500 km, but eventually flattens out at smaller scales. This flattening is likely an artifact, because at these scales the geostrophic component makes up a small fraction of the observed signal. It is quite possible that the flattening is due to noise or biases introduced by the interpolation procedure or by truncation errors in the reported wind and temperature data.

4.3.3 Decomposition of kinetic and potential energy spectra

We have shown that the total energy spectrum can be decomposed into its geostrophic and inertia–gravity wave components. To confirm that the observed spectra are consistent with geostrophic dynamics at synoptic scales and with inertia–gravity wave dynamics at mesoscales, we now decompose into its two components each of the atmospheric spectra, the longitudinal and transverse kinetic energy spectra, $S^u(k)$ and $S^v(k)$, and the potential energy spectrum, $S^b(k)$. This can be done if one makes one further assumption. Following Charney and results from numerical simulations of geostrophic turbulence, the geostrophic streamfunction is assumed to be three-dimensionally isotropic, with the vertical coordinate rescaled by f/N (Charney, 1971). This implies $S_g^u(k) = S_g^b(k)$, a relation that we noted is satisfied by the observed spectra in the synoptic range (Fig. 4.1a).

This decomposition—obtained by applying (4.15)–(4.17) and (4.20)–(4.22), given in the Materials and Methods section—confirms the main conclusion that the observed synoptic-scale flow is consistent with geostrophic dynamics and that the observed mesoscale flow is consistent with inertia–gravity wave dynamics. In the synoptic range, the observed spectra match the diagnosed geostrophic components (Fig. 4.2a). At the transition scale, the diagnosed geostrophic spectra start deviating from the observed spectra
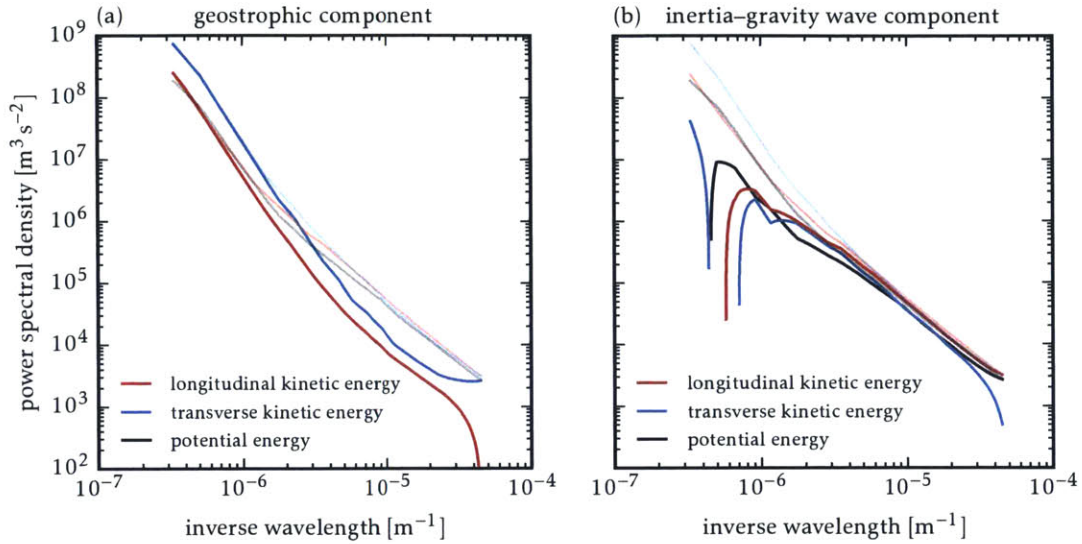
Figure 4.2: Decomposition of observed wavenumber spectra into the geostrophic and inertia–gravity wave components. (a) Diagnosed geostrophic component of the spectra of longitudinal kinetic energy $S_g^u(k)$, transverse kinetic energy $S_g^v(k)$, and potential energy $S_g^b(k)$ (heavy lines) and observed spectra for reference (faint lines). Note that by construction $S_g^u(k) = S_g^b(k)$. (b) Diagnosed inertia–gravity wave component of the spectra of longitudinal kinetic energy $S_w^u(k)$, transverse kinetic energy $S_w^v(k)$, and potential energy $S_w^b(k)$ (heavy lines) and observed spectra for reference (faint lines).

and keep falling off steeply. At this scale, the diagnosed inertia–gravity wave spectra become comparable to the observed spectra and start matching them in the mesoscale range (Fig. 4.2b).

4.4 Discussion

Our analysis shows that the aircraft observations are consistent with a geostrophic flow that dominates the synoptic range and with inertia–gravity waves that dominate the mesoscale range. This conclusion is predicated on the assumption that the total observed flow is a superposition of geostrophic flow and inertia–gravity waves and that these two components are uncorrelated and horizontally isotropic.

In accord with Lindborg's result (Lindborg, 1999), our analysis conclusively shows that mesoscale flows are not in geostrophic balance, thus falsifying previous suggestions that mesoscale spectra represent geostrophic eddies generated by atmospheric convective events or geostrophically balanced fronts at the tropopause. Our analysis then further shows that the observed mesoscale spectra $S^u(k)$, $S^v(k)$, and $S^b(k)$ are consistent with the

dispersion and polarization relations of linear hydrostatic inertia–gravity waves. Presumably, it would be very surprising if the strongly nonlinear ageostrophic flows characteristic of stratified turbulence were to yield the same relations between these mesoscale spectra as the linear waves.

A similar transition between a dominant geostrophic flow at large scales and dominant inertia–gravity waves at small scales is well established in the oceanic spectra (Chapter 2), providing further support for our interpretation of the synoptic-to-mesoscale transition. The ocean, like the atmosphere, is a strongly rotating and stratified fluid. The most energetic large-scale fluctuations are generated by baroclinic instabilities in both fluids, while small-scale inertia–gravity waves are triggered by any fast perturbation. A posteriori, it should not be surprising that both fluids exhibit a transition from geostrophic dynamics at large scales to inertia–gravity wave dynamics at small scales.

That inertia–gravity waves dominate the mesoscale spectrum does not mean that the atmosphere is filled with a nearly uniform and stationary wave field, as appears to be the case in the ocean (Garrett and Munk, 1979). The atmospheric wave field is likely highly intermittent in both space and time (Fritts and Alexander, 2003). Our result merely suggests that on average, inertia–gravity waves dominate the mesoscale range. A good understanding of what sets the shape of the average atmospheric inertia–gravity wave spectrum is lacking. The oceanic inertia–gravity wave spectrum, however, has been shown to be an equilibrium solution of weakly interacting inertia–gravity waves with slopes close to $k^{-5/3}$ (McComas and Müller, 1981; Polzin and Lvov, 2011). While the residence time of waves in the atmosphere is shorter and shears provided by jet streams stronger, weak turbulence theory may similarly yield insight into the atmospheric inertia–gravity wave field.

The spatial patterns of mesoscale energy are also consistent with the dominance of inertia–gravity waves at the mesoscales. The aircraft spectra in the mesoscale range are up to six times larger in mountainous regions than over flat terrain (Nastrom et al., 1987). The inertia–gravity wave activity is expected to be enhanced over mountains, where lee waves are excited by large- and synoptic-scale flows impinging on topography. Stratified turbulence, on the other hand, is fed directly by synoptic-scale energy and therefore not expected to be enhanced over mountainous regions.

High-resolution numerical models reproduce the synoptic-to-mesoscale transition (Koshyk et al., 1999). Model results are consistent with the observation that the synoptic range is dominated by rotational flow, while the rotational and divergent components are of the same order in the mesoscale range (Koshyk and Hamilton, 2001; Hamilton et al., 2008). Models also show a pattern of enhanced mesoscale energies in regions of high topography (Hamilton et al., 2008). The simulations further show that the mesoscale spectrum is not the result of stratified turbulence (Skamarock et al., 2014), in agreement with our conclusion that they are the signature of inertia–gravity waves.

The emergent picture is relatively simple. Geostrophic synoptic-scale baroclinic disturbances force a forward enstrophy cascade that continues through the synoptic-to-

mesoscale transition. The k^{-3} or slightly steeper geostrophic spectrum is masked by inertia–gravity waves at scales smaller than 500 km. This picture does not rule out the possibility that some fraction of the energy in the mesoscale spectra is associated with fronts and stratified turbulence, but these contributions must be small.

The result of this chapter may also have some implications for the theoretical predictability of atmospheric flows. Lorenz argued that a turbulent flow with a $k^{-5/3}$ kinetic energy spectrum has a finite predictability time, which is of the order of the eddy turnover time (Lorenz, 1969). More and more accurate knowledge of the initial state cannot push forecasts beyond that limit—even if we had a perfect model. In contrast, turbulent flows with a k^{-3} kinetic energy spectrum do not have such a predictability limit and ever more accurate initial conditions can lead to ever longer forecasts.

If the flat mesoscale spectrum were due to a turbulent cascade, that would pose a limit on predictability of synoptic systems (Vallis, 2006; Palmer, 2000). Observing and modeling the atmosphere beyond the transition at 500 km would yield rapidly diminishing returns in predictability. If this part of the spectrum is due to inertia–gravity waves, however, improving observational systems and forecast models may not prove as futile. Inertia–gravity waves do not propagate errors in the same way as the turbulent flows discussed by Lorenz. If the geostrophic component of the mesoscale flow were to be observed, despite the dominance of inertia–gravity waves, the forecast times of synoptic systems could potentially be extended considerably.

It should be noted, however, that other processes not considered in this theoretical argument may affect predictability. Moist convective processes, for example, lead to rapid growth of errors that can leak into the geostrophic flow (Zhang et al., 2007). Currently, the practical predictability of the weather is likely limited by inadequate representation of such processes.

4.5 Materials

4.5.1 Aircraft data

The spectra shown in Fig. 4.1a are calculated from wind and temperature measurements obtained by the MOZAIC program, which equipped commercial aircraft with instrumentation to measure trace gases, but also records wind speed and direction from the board computer. The data used here were obtained in 2002–2010 and are restricted to the northern hemisphere midlatitudes (30 to 60 degrees latitude). Great circles are fit to the flight paths and segments are discarded if they are shorter than 6,000 km, the average sample spacing is coarser than 1.2 km, or the deviation from the great circle is greater than 2 degrees. The data are then linearly interpolated onto a regular grid with 1 km spacing. Data at pressures larger than 350 hPa are discarded. Subsequently, for each flight, data

deviating more than 1 km in altitude from the mean altitude are also discarded. Temperature data are adjusted to account for remaining variations in flight altitude, assuming a constant stratification $N = 0.02$ s^{-1}, but this correction is of no consequence for the results discussed in the main part of the chapter. Nastrom and Gage (Nastrom and Gage, 1985) and Cho and Lindborg (Cho and Lindborg, 2001) showed that the spectral shapes are qualitatively the same in the upper troposphere and lower stratosphere, thus we do not separate the data into vertical bins. Spectra are computed by applying a Hann window, compensating for the variance loss, performing a discrete Fourier transform, and averaging over all 458 segments. Spectra at wavelengths smaller than 20 km are discarded because they are potentially affected by the interpolation procedure or by truncation errors in the reported data. Locations were reported in longitude/latitude with an accuracy of 0.01 degrees; zonal and meridional winds were reported with an accuracy of 0.01 m s^{-1}; temperatures were reported with an accuracy of 0.01 K. These noise levels do not affect the spectra on scales larger than 20 km. We also discard the largest resolved wavelength, because the power at this wavelength is reduced artificially by the window.

4.5.2 Details of decomposition

We here give more detail on the decomposition techniques used to analyze the aircraft spectra. We gave a more comprehensive description of these techniques in Chapter 3 and illustrated their skill to analyze oceanic spectra.

Let u and v be horizontal velocity components defined in the xy-plane with x aligned with the aircraft track, so u is the longitudinal (along-track) component and v is the transverse (across-track) component. The time t and altitude z are considered fixed during the measurement, so they will be ignored. Now, a general two-dimensional flow has a Helmholtz decomposition into rotational and divergent components of the form $u = -\psi_y + \phi_x$ and $v = \psi_x + \phi_y$. The functions ψ and ϕ are uniquely determined in terms of the velocity field with doubly periodic boundary conditions.

Progress with the statistical theory is possible if $\psi(x,y)$ and $\phi(x,y)$ are uncorrelated. We can then write the two-dimensional power spectra of u and v as

$$S^u(k,l) = l^2 S^\psi(k,l) + k^2 S^\phi(k,l), \tag{4.6}$$

$$S^v(k,l) = k^2 S^\psi(k,l) + l^2 S^\phi(k,l), \tag{4.7}$$

where l is the across-track wavenumber. Integration over l and some manipulation gives (4.1) and (4.2) with

$$D^\psi(k) = \frac{1}{\pi} \int_k^\infty \sqrt{k_h^2 - k^2} S^\psi(k_h) dk_h, \tag{4.8}$$

$$D^\phi(k) = \frac{1}{\pi} \int_k^\infty \sqrt{k_h^2 - k^2} S^\phi(k_h) dk_h, \tag{4.9}$$

85

where $S^\psi(k_h)$ and $S^\phi(k_h)$ are the two-dimensional isotropic spectra of ψ and ϕ, related to the two-dimensional spectra by $S(k_h) = 2\pi k_h S(k,l)$, $k_h = \sqrt{k^2 + l^2}$ is the magnitude of the horizontal wavenumber vector, and $k \geq 0$ is the along-track wavenumber.

The equations (4.1) and (4.2) can be solved explicitly, given the boundary conditions $D^\psi(\infty) = 0$ and $D^\phi(\infty) = 0$,

$$D^\psi(s) = \int_s^\infty [S^u(\sigma)\sinh(s-\sigma) + S^v(\sigma)\cosh(s-\sigma)]\,d\sigma, \qquad (4.10)$$

$$D^\phi(s) = \int_s^\infty [S^u(\sigma)\cosh(s-\sigma) + S^v(\sigma)\sinh(s-\sigma)]\,d\sigma, \qquad (4.11)$$

where for convenience the coordinate was transformed to $s = \ln k$.

Relation (4.5) follows from the dispersion and polarization relations of hydrostatic inertia–gravity waves. Combining the vorticity and continuity equations of the linearized primitive equations yields $\nabla^2\psi_t = f\nabla^2\phi$ (Pedlosky, 2010), which implies that

$$S^\psi(k,l,\omega) = \frac{f^2}{\omega^2}S^\phi(k,l,\omega) \qquad (4.12)$$

and thus, with the use of (4.6) and (4.7),

$$S^u(k,l,\omega) + S^v(k,l,\omega) = \left(1 + \frac{f^2}{\omega^2}\right)\left(k^2 + l^2\right)S^\phi(k,l,\omega). \qquad (4.13)$$

The linear buoyancy equation is $b_t + N^2 w = 0$ and the potential energy spectrum $S^b(k) = \langle|\hat{b}(k)|^2\rangle/N^2$ can also be related to the spectrum of the velocity potential,

$$S^b(k,l,\omega) = \left(1 - \frac{f^2}{\omega^2}\right)\left(k^2 + l^2\right)S^\phi(k,l,\omega), \qquad (4.14)$$

if uncorrelated plane waves are assumed, with slowly varying background allowed. Adding (4.13) and (4.14) then eliminates the dependence on ω, so that (4.5) follows by integrating over l and ω.

The decomposition of the three individual spectra can be achieved by decomposing $D^\psi(k) = D_g^\psi(k) + D_w^\psi(k)$. The divergent part $D^\phi(k)$ needs no such decomposition, because the geostrophic component of the flow is divergence-free. Using

$$S_g^u(k) = D_g^\psi(k), \qquad (4.15)$$

$$S_g^v(k) = -k\frac{d}{dk}D_g^\psi(k), \qquad (4.16)$$

$$S_g^b(k) = D_g^\psi(k), \qquad (4.17)$$

where Charney's assumption $S_g^b(k) = S_g^u(k)$ was applied, we can write the total energy spectrum of the geostrophic component $E_g(k) = \frac{1}{2}[S_g^u(k) + S_g^v(k) + S_g^b(k)]$ as

$$E_g(k) = \left(1 - \frac{k}{2}\frac{d}{dk}\right)D_g^\psi(k).\qquad(4.18)$$

Since $E_g(k) = E(k) - E_w(k)$ can be diagnosed from the observations and (4.5), this can be solved for $D_g^\psi(k)$:

$$D_g^\psi(s) = 2\int_s^\infty E_g(\sigma)e^{2(s-\sigma)}\,d\sigma,\qquad(4.19)$$

where the boundary condition $D_g^\psi(\infty) = 0$ was used and the coordinate was again transformed to $s = \ln k$ for convenience. The decomposition is now complete. The wave spectra are

$$S_w^u(k) = D_w^\psi(k) - k\frac{d}{dk}D_w^\phi(k),\qquad(4.20)$$

$$S_w^v(k) = -k\frac{d}{dk}D_w^\psi(k) + D_w^\phi(k),\qquad(4.21)$$

$$S_w^b(k) = \left(1 - k\frac{d}{dk}\right)\left[D_w^\phi(k) - D_w^\psi(k)\right],\qquad(4.22)$$

where $D_w^\psi(k) = D^\psi(k) - D_g^\psi(k)$, $D_w^\phi(k) = D^\phi(k)$, and (4.5) was used to deduce the last equation.

CHAPTER 5

SEASONALITY IN SUBMESOSCALE TURBULENCE
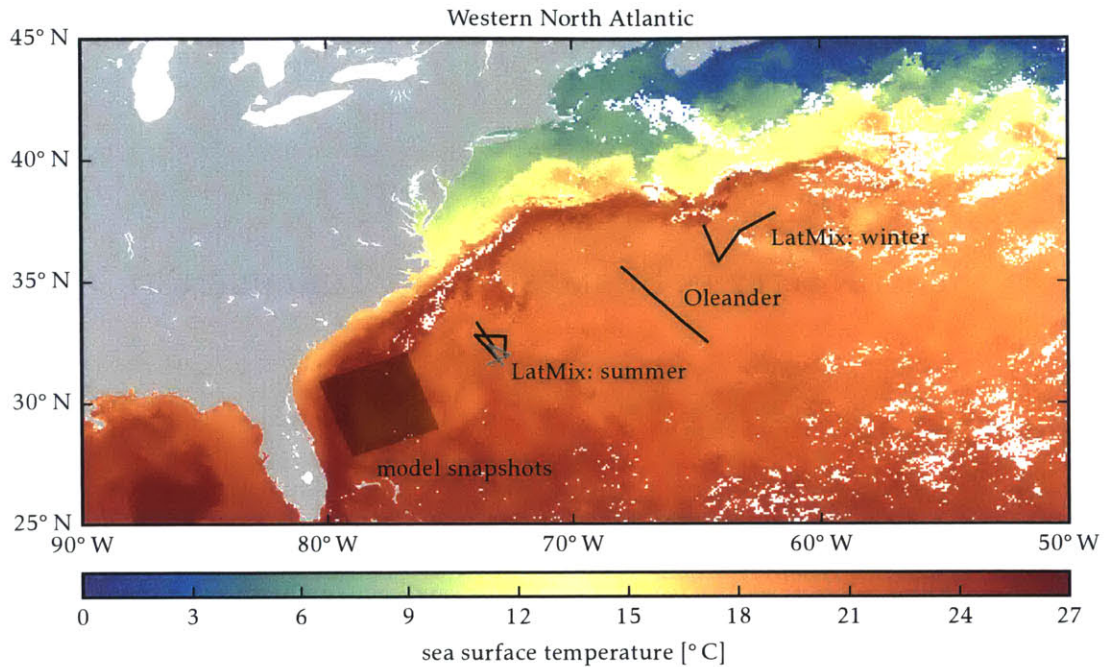
Figure 5.1: Locations of velocity transects (black lines), of additional buoyancy transects (dark gray lines), and of the model snapshots shown in Fig. 5.2 (transparent shading). The color shading shows sea surface temperatures on 13–20 March 2012 (8-day L3 MODIS Aqua composite of 4 μm nighttime temperature). Missing data are indicated by white shading.

5.1 Introduction

The ocean supports motions on a wide range of scales. The most energetic flows are eddies with horizontal scales of a few hundred kilometers, which are the oceanic analogues of the atmospheric cyclones and anticyclones. These mesoscale eddies are dynamically well understood and are routinely observed with satellite altimeters (Stammer, 1997). Classic theories predict that the kinetic energy in the mesoscale eddy field should decay very rapidly at smaller scales (Charney, 1971) and be associated with weak vertical velocities. Recent numerical modeling studies suggest, however, that there is a very dynamic near-surface eddy field on scales 1–100 km, which drives vertical velocities one or two orders of magnitude larger than those of order 1 m day^{-1} associated with the divergence of mesoscale flows (Lapeyre and Klein, 2006; Capet et al., 2008a). These submesoscale flows consist of narrow horizontal currents associated with strong horizontal gradients in buoyancy.

90
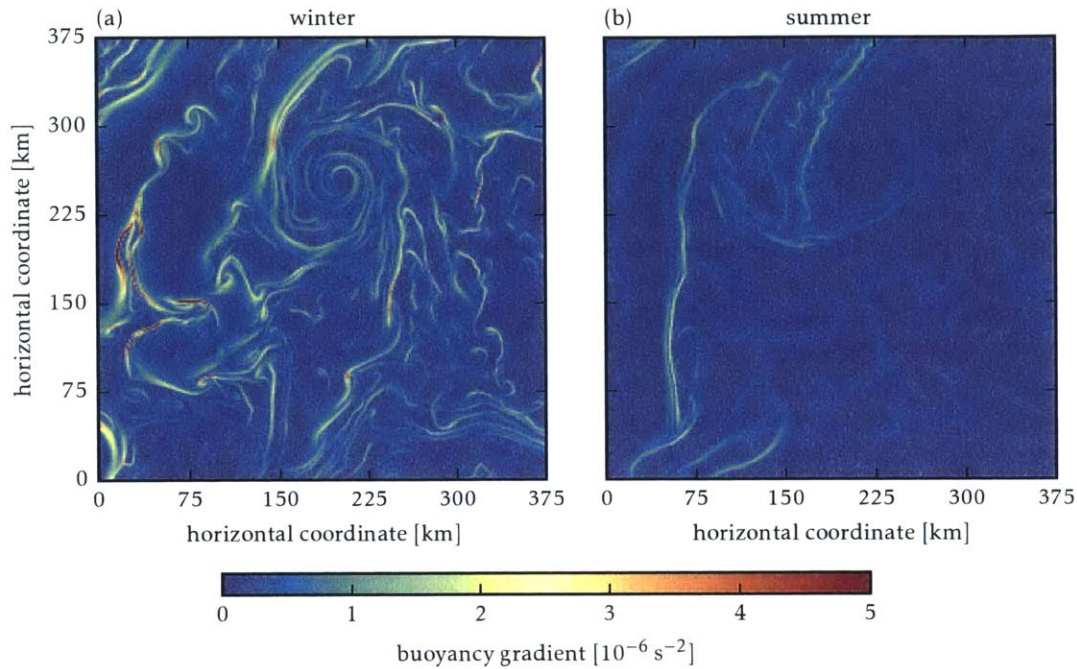
Figure 5.2: Surface buoyancy gradient magnitudes from a numerical simulation in the Gulf Stream region in (a) winter (February 15) and (b) summer (August 15)

High-resolution numerical simulations suggest that submesoscale flows are much stronger in winter than in summer (Capet et al., 2008d; Mensa et al., 2013; Sasaki et al., 2014). This is illustrated here with snapshots from a simulation of the Gulf Stream region with a horizontal resolution of 750 m (Fig. 5.1). More detail on the simulation is given in the *Methods* section. The snapshots of surface buoyancy gradients show that fronts are strong in winter and much less pronounced in summer, except for the Gulf Stream front that persists throughout the year (Fig. 5.2). This seasonality provides important clues to the dynamics of submesoscale turbulence.

One mechanism that has been proposed to energize submesoscale flows is mesoscale-driven surface frontogenesis (Lapeyre and Klein, 2006): the presence of the ocean surface allows strain fields at the edges of mesoscale eddies to sharpen surface buoyancy gradients (Stone, 1966a; Lapeyre et al., 2006). Unlike in the ocean interior, upwelling of dense water and downwelling of light water cannot counteract the increase of lateral buoyancy gradients at the surface, where the vertical transport of waters must vanish. The resulting surface fronts have been shown to break up in secondary roll-up instabilities, which drive a nonlinear cascade of surface buoyancy variance and energize the entire range of scales below the mesoscale (Held et al., 1995). Surface quasi-geostrophic theory has

91

been used to qualitatively illustrate these dynamics and predicts surface kinetic and potential energy spectra to scale like $k^{-5/3}$ (ref. Blumen, 1978), where k is the horizontal wavenumber. Surface quasi-geostrophic flows are predicted to be confined to a shallow surface layer; they decay exponentially away from the surface, with smaller-scale motions decaying more rapidly than larger-scale ones. Energy spectra are therefore expected to fall off with wavenumber more rapidly away from the surface. Surface quasi-geostrophic dynamics are modified by ageostrophic effects that accelerate frontogenesis (Hoskins and Bretherton, 1972), lead to true frontal collapse, and thereby produce k^{-2} energy spectra (Boyd, 1992).

Another mechanism that has been proposed to energize submesoscale flows in the mixed layer is baroclinic instability (Haine and Marshall, 1998; Boccaletti et al., 2007). This instability allows perturbations to extract energy from the lateral buoyancy gradients generated by mesoscale stirring or by spatial variations in atmospheric forcing. Eady edge waves can propagate both along the surface and along the sharp increase in stratification at the base of the mixed layer (Eady, 1949; Blumen, 1979; Rivest et al., 1992; Juckes, 1994; Held et al., 1995). Edge waves with scales close to the mixed layer deformation radius, Nh/f, propagate at the same speed along the surface and the base of the mixed layer and can therefore constructively interfere (N is the density stratification of the mixed layer, h is its depth, f is the Coriolis parameter). This interference results in an amplification of these disturbances on a time scale $N/f\Lambda$ (Eady, 1949; Λ is the geostrophic shear in the mixed layer), which causes water from the denser side of the lateral buoyancy gradient to slide under the more buoyant waters. The most unstable mode in the Eady model, which assumes that N and Λ are constant through the mixed layer and that the base of the mixed layer is rigid, has a wavelength of $4Nh/f$ and a growth rate of $0.3f\Lambda/N$ (Eady, 1949; Vallis, 2006). In deep mixed layers of a few hundred meters and for typical values of $f \sim 10^{-4}$ s^{-1}, $\Lambda \sim 10^{-4}$ s^{-1}, and $N \sim 10^{-3}$ s^{-1}, disturbances therefore grow on horizontal scales of order 1–10 km and time scales of order 1 day (Boccaletti et al., 2007). Once these disturbances have grown to finite amplitude, turbulent scale interactions exchange kinetic energy between scales. If the flow is sufficiently constrained by the Earth's rotation, kinetic energy is preferentially transferred to larger scales, filling in the spectrum between the mixed layer deformation radius and the mesoscale (Boccaletti et al., 2007; Fjørtoft, 1953; Kraichnan, 1967; Sasaki et al., 2014). A major prediction of this scenario, which distinguishes it from mesoscale-driven surface frontogenesis, is that the full depth of the mixed layer is energized and that the energy decays rapidly below the base of the mixed layer.

The energization of the submesoscale by such mixed layer instabilities depends crucially on the mixed-layer depth. Mixed layer baroclinic instabilities are powered by the potential energy stored in lateral buoyancy gradients. The deeper the mixed layer, for a given lateral buoyancy gradient, the more potential energy is released by the slumping isopycnals and the more energetic the resulting submesoscale flows. Given that lateral buoyancy gradients are fairly constant throughout the year, the kinetic energy generated

by the instabilities is thus expected to peak in winter, when mixed layers are deepest. In summer, when mixed layers are shallow, the instabilities can easily be damped out. While the growth rate of mixed layer instabilities is independent of the mixed-layer depth h, the damping of mixed layer motions by atmospheric forcing is faster in shallower mixed layers (Marshall and Schott, 1999). Additionally, the mixed layer deformation radius scales with h, so the instability scale is smaller when the mixed layer is shallow. Since smaller-scale flows can more easily escape the influence of rotation, kinetic energy created in shallow mixed layers may cascade to smaller rather than larger scales, so that the instability fails to energize the range between the mixed layer deformation radius and the mesoscale (Molemaker et al., 2005; Capet et al., 2008c).

The instability mechanism therefore implies a seasonal cycle in submesoscale turbulence. In winter, when the mixed layer is deep, instabilities grow on a time scale of about 1 day and at a horizontal scale between 1–10 km, subsequently energizing the entire submesoscale range through turbulent scale interactions. In summer, the instability is either damped out or fails to energize the submesoscale range because of the lack of a strong inverse cascade, when the mixed layer deformation radius is very small. The mesoscale-driven surface frontogenesis mechanism, on the other hand, only depends on the presence of a mesoscale eddy field, which does not undergo a strong seasonal cycle. The strong seasonality at the submesoscale visible in numerical simulations (Fig. 5.2) supports the instability mechanism.

Observational confirmation of a seasonal cycle in submesoscale turbulence, however, is so far lacking. Numerical models used for studying submesoscale dynamics typically do not explicitly resolve small-scale processes in the mixed layer. In summer, for example, when the mixed layer is shallow, the models do not resolve the instability scale of the baroclinic instability in the mixed layer, which can be as small as 10–100 m. High-resolution satellite observations show a seasonal cycle in sea surface temperature gradients in the Gulf Stream region (Armstrong et al., 2012), but it remains unclear whether the temperature variations are reflected in density or compensated by salinity variations (Ferrari and Rudnick, 2000), whether the temperature signal is accompanied by along-front flows, whether these flows are in geostrophic balance, and what the subsurface structure of these flows is.

In-situ observations of velocity and buoyancy in the upper ocean are used in this study to conclusively show that submesoscale flows undergo a strong seasonal cycle on scales of 1–100 km. An enhancement of the observed submesoscale flows throughout the winter mixed layer and the flows' leading-order horizontal non-divergence are argued to be further evidence for the instability mechanism.

5.2 Results

5.2.1 Data

We present data from the western subtropical North Atlantic, where the mixed layer is known to undergo a strong seasonal cycle (de Boyer Montégut et al., 2004). The data were collected as part of two separate observational programs: the Oleander project along a transect between Elisabeth, New Jersey, and Hamilton, Bermuda, occupied weekly in 2005–2013, and the Lateral Mixing Experiment (LatMix) along several straight transects off Cape Hatteras in June 2011 (summer) and just south of the Gulf Stream extension in March 2012 (winter) (Shcherbina et al., 2013) (Fig. 5.1). We consider only the Oleander data south–west of the Gulf Stream extension and separate the transects into winter ones (January through March, when the mixed layers are deepest, de Boyer Montégut et al., 2004) and summer ones (June through August). In the Oleander region, mixed layers are on average 150–200 m deep in winter and shoal to order 10 m in summer (de Boyer Montégut et al., 2004). In the LatMix winter experiment, the mixed layer was about $h = 220$ m deep and had an average stratification of $N = 2 \times 10^{-3}$ s^{-1}, giving an Eady instability scale $4Nh/f = 20$ km. In the LatMix summer experiment, the mixed layer was no deeper than $h = 5$ m and, while there were no shallow measurements of stratification, the Eady instability scale was most likely no larger than 100 m.

Power spectra are computed at constant depths for the longitudinal (along-track) velocities u, the transverse (across-track) velocities v (both obtained with acoustic Doppler current profilers, ADCP), and the buoyancy b (obtained from towed conductivity–temperature–depth sensors, CTD; available only for the LatMix data):

$$S^u(k) = \langle |\hat{u}(k)|^2 \rangle, \qquad S^v(k) = \langle |\hat{v}(k)|^2 \rangle, \qquad S^b(k) = \langle |\hat{b}(k)|^2 \rangle / N^2, \tag{5.1}$$

where k is the along-track wavenumber, the caret denotes a Fourier transform, and $\langle \cdot \rangle$ denotes the average over transects. More detail on the observations and the computation of these spectra is provided in the *Methods* section. Departing from recent literature, we do not focus solely on spectral slopes, because they do not uniquely determine the dynamics: different dynamics can produce the same spectral slopes. We instead use power spectra primarily to diagnose submesoscale energy levels at different times and depths. Together with the relationships between the spectra of u, v, and b, this allows a more nuanced assessment of submesoscale dynamics.

5.2.2 Seasonality

The near-surface spectra of kinetic and potential energy,

$$K(k) = \tfrac{1}{2}[S^u(k) + S^v(k)], \qquad P(k) = \tfrac{1}{2}S^b(k), \tag{5.2}$$

94

Figure 5.3: (a) Kinetic energy spectrum at 50 m depth for the Oleander winter data. (b) Potential and kinetic energy spectra at 20 m depth for the LatMix winter experiment. (c) Kine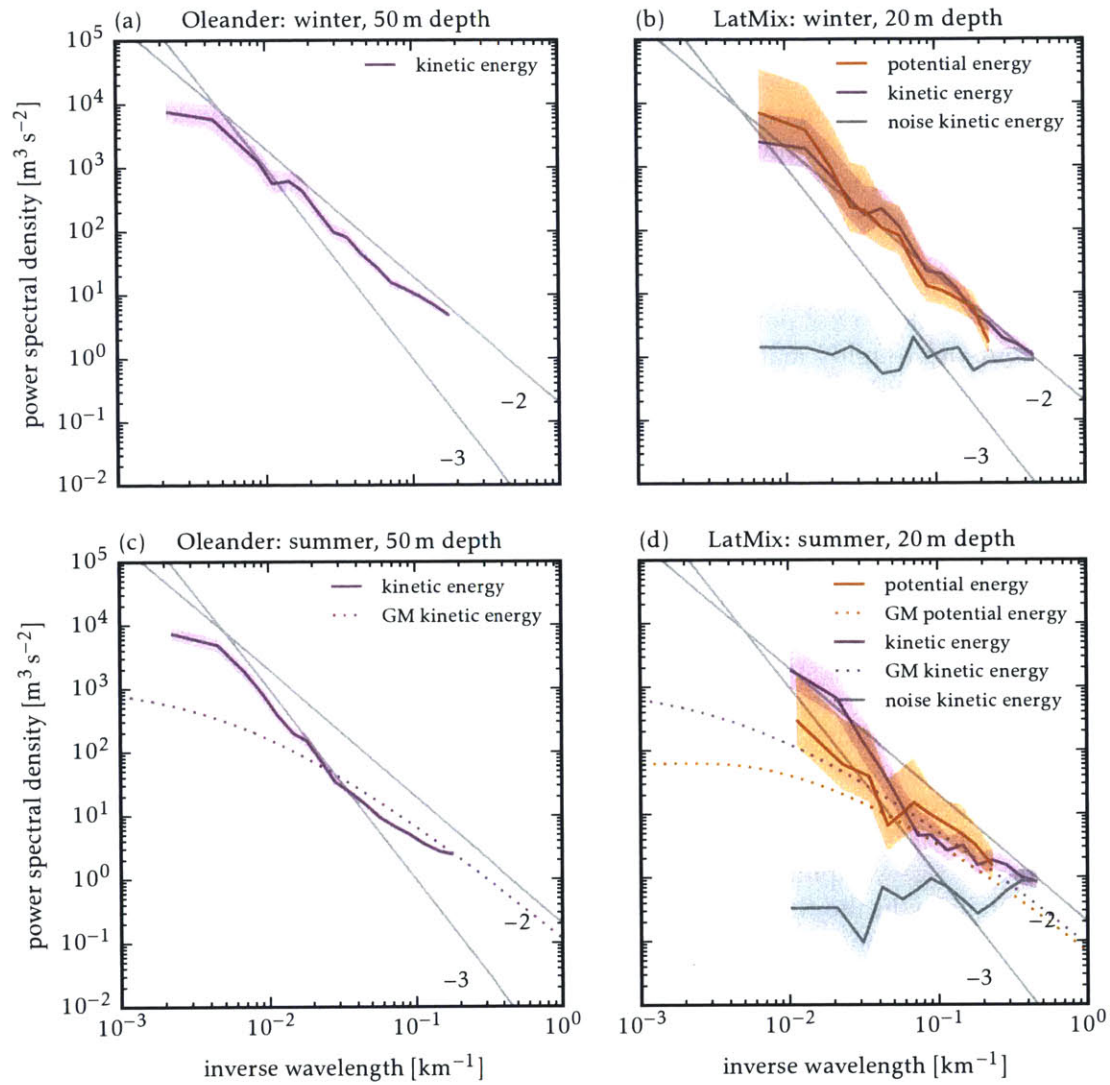tic energy spectrum at 50 m depth for the Oleander summer data. (d) Potential and kinetic energy spectra at 20 m depth for the LatMix summer experiment. The light shadings are 95% confidence intervals. Also shown are Garrett–Munk (GM) model spectra for internal waves in the seasonal thermocline, estimates for the noise level of the LatMix velocity data, and reference lines with slopes −2 and −3.

at 50 m depth for Oleander and 20 m depth for LatMix are shown in Fig. 5.3. The kinetic energy spectra from Oleander and LatMix have similar shapes, suggesting that the dynamics are similar across the different geographical locations. All spectra exhibit marked differences between summer and winter. While the energy levels at the mesoscale (scales larger than 100 km) remain fairly constant across seasons, the submesoscale energy levels (scales of order 10 km) undergo a seasonal transition: they are significantly higher in winter than in summer (Fig. 5.3). This seasonal difference in submesoscale energy levels is reflected in how rapidly the energy falls off with wavenumber in the submesoscale range. In winter, the spectra are relatively flat and approximately follow a k^{-2} power law (Fig. 5.3a,b). In summer, the spectra in the range 20–100 km fall off more rapidly and approximately follow a k^{-3} power law (Fig. 5.3c,d); at scales smaller than 20 km, the spectra flatten out and roughly match the Garrett–Munk (GM) empirical model spectrum of internal waves (Munk, 1981). Potential and kinetic energy spectra are approximately equal in the LatMix winter experiment (Fig. 5.3b). In the LatMix summer experiment, the kinetic energy spectrum is considerably larger than the potential energy spectrum at scales larger than 20 km (Fig. 5.3d). At smaller scales, they are roughly equal and consistent with the GM spectrum for internal waves.

These spectra confirm that submesoscale turbulence is more energetic in winter than in summer. The seasonal transition in submesoscale potential energy levels is consistent with the seasonal transition in frontal strength seen in numerical simulations (Fig. 5.2). In winter, the observed equipartition between kinetic and potential energy, $K(k)/P(k) \simeq 1$, and the k^{-2} spectral slopes are likely the result of the turbulent dynamics induced by mixed layer instabilities and their interaction with the mesoscale straining field, but how these dynamics shape the spectra is currently not understood. The wintertime spectral slopes are also consistent with the mesoscale-driven surface frontogenesis mechanism, but the seasonal transition is not. In summer, the steep energy spectra are consistent with interior quasi-geostrophic turbulence (Charney, 1971). Eddies generated through baroclinic instability in the thermocline transfer very little energy to submesoscales. The observation that the kinetic energy spectrum is substantially larger than the potential energy spectrum in this range is also consistent with the ratio $K(k)/P(k) \simeq 4$ predicted by interior quasi-geostrophic turbulence (Charney, 1971; Chapter 2). The emergence of the GM spectrum at scales below 20 km confirms that internal waves dominate these scales and mask an even more dramatic drop-off of geostrophic submesoscale motions in summer, as pointed out in Chapters 2 and 3.

5.2.3   Wintertime submesoscale flows

We examine the vertical structure of kinetic energy in the LatMix winter experiment, which provides further evidence for the mixed layer instability mechanism. We use the LatMix data because the accompanying buoyancy data provide information on the mixed

96
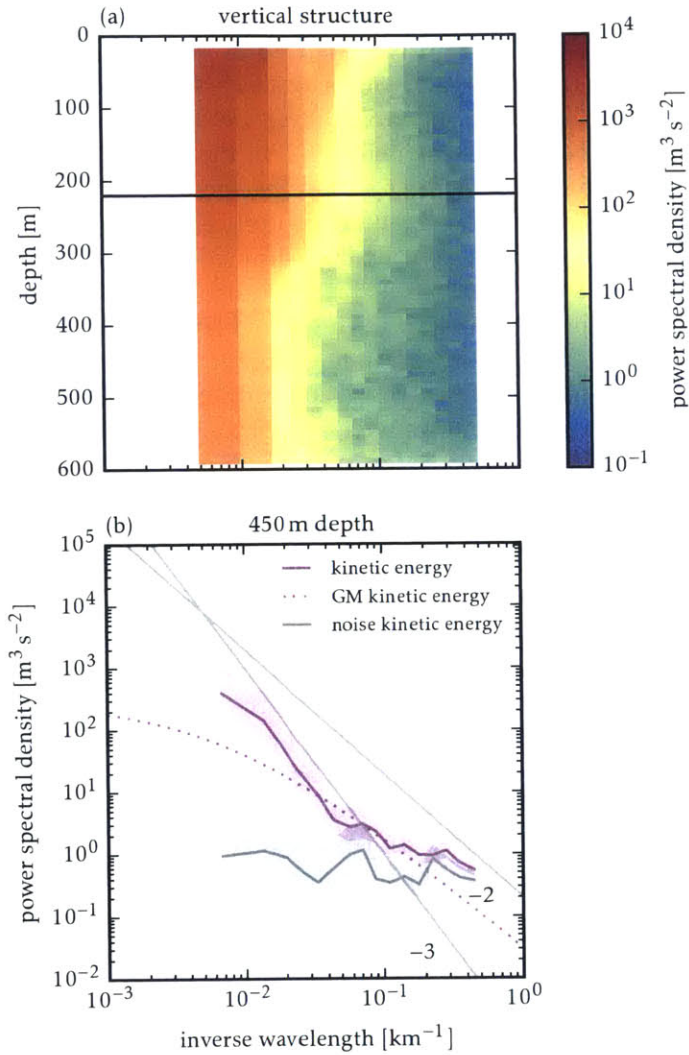
Figure 5.4: Vertical structure in the LatMix winter experiment: (a) Kinetic energy spectrum as a function of depth. The depth of the mixed layer $h = 220$ m is indicated by the black horizontal line. (b) Kinetic energy spectrum at a depth of 450 m. The light shadings are 95% confidence intervals. Also shown is the Garrett–Munk (GM) model spectrum for internal waves in the thermocline, an estimate of the noise level of the LatMix velocity data, and reference lines with slopes −2 and −3.

97

layer depth. It is roughly constant at 220 m across the transects, which facilitates the interpretation of the vertical structure in observed kinetic energy. The submesoscale kinetic energy is fairly constant throughout the mixed layer at scales larger than 20 km (Fig. 5.4a). At smaller scales, the energy is slightly enhanced at the surface and at the base of the mixed layer, as expected from Eady dynamics at scales below the deformation radius. The spectra uniformly scale like k^{-2} in the mixed layer. Below the base of the mixed layer, the energy decays rapidly; the spectra steepen to k^{-3} in the permanent thermocline at scales larger than 20 km (Fig. 5.4b). At smaller scales, the spectra in the thermocline follow the GM model and internal waves mask a steep drop-off. This vertical structure is consistent with submesoscale turbulence induced by baroclinic instabilities in the mixed layer, but inconsistent with the mesoscale-driven surface frontogenesis, for which the vertical decay of submesoscale turbulence would start in the mixed layer (Chapter 2).

The observations further suggest that the energetic submesoscale flows in winter are to leading order in geostrophic balance, as required for an inverse energy cascade. Horizontally nondivergent, isotropic flows have a distinct signature in the power spectra of transverse and longitudinal velocities: for a flow with a k^{-2} kinetic energy spectrum, the ratio $S^v(k)/S^u(k) \simeq 2$ is expected (Charney, 1971; Leith, 1971; Chapters 2 and 3). Error bars are substantial, but a clear trend across scales is evident, which suggests that the observed spectra are roughly consistent with this relation. In the Oleander data, where the large number of transects results in smaller error bars, the relation is roughly satisfied over the entire submesoscale range (Fig. 5.5a). In the LatMix winter experiment, the relation is roughly satisfied throughout the submesoscale range (Fig. 5.5b) except for scales smaller than 5 km, where noise may affect the spectra significantly (cf. Fig. 5.3b). This is consistent with a flow that is to leading order geostrophically balanced. Strongly ageostrophic dynamics, on the other hand, have a leading-order horizontal divergence and thus a different $S^v(k)/S^u(k)$ ratio. Near-inertial waves forced by strong winter winds or wind- and buoyancy-driven three-dimensional turbulence can thus be ruled out to dominate the submesoscale observations in winter.

Calculation of the Rossby and Froude numbers as functions of scale, $Ro(k) = [k^3 K(k)]^{1/2}/f$ and $Fr(k) = [k^3 P(k)]^{1/2}/f$, further suggests that the observed wintertime submesoscale flows follow quasi-geostrophic dynamics. Quasi-geostrophic dynamics describe flows with $Fr(k) \sim Ro(k) \ll 1$ (e.g. Pedlosky, 1987). For k^{-2} kinetic energy spectra, as observed in winter, $Ro(k)$ increases moderately with k, approximately like $k^{1/2}$. Since $Ro(k)$ is small in the mesoscale, it remains fairly small throughout the observed submesoscale range (Fig. 5.5c,d). The condition $Fr(k) \sim Ro(k)$ is satisfied because of the approximate equipartition between kinetic and potential energy (Fig. 5.3b, 5.5d). Extrapolating $Ro(k)$ and $Fr(k)$ to smaller scales, we expect that they reach order 1 at order 1 km. The dynamics are thus expected to become strongly ageostrophic—allowing effective transfer of energy to dissipation scales (Molemaker et al., 2005; Capet et al., 2008c)—at scales smaller than the observed range.

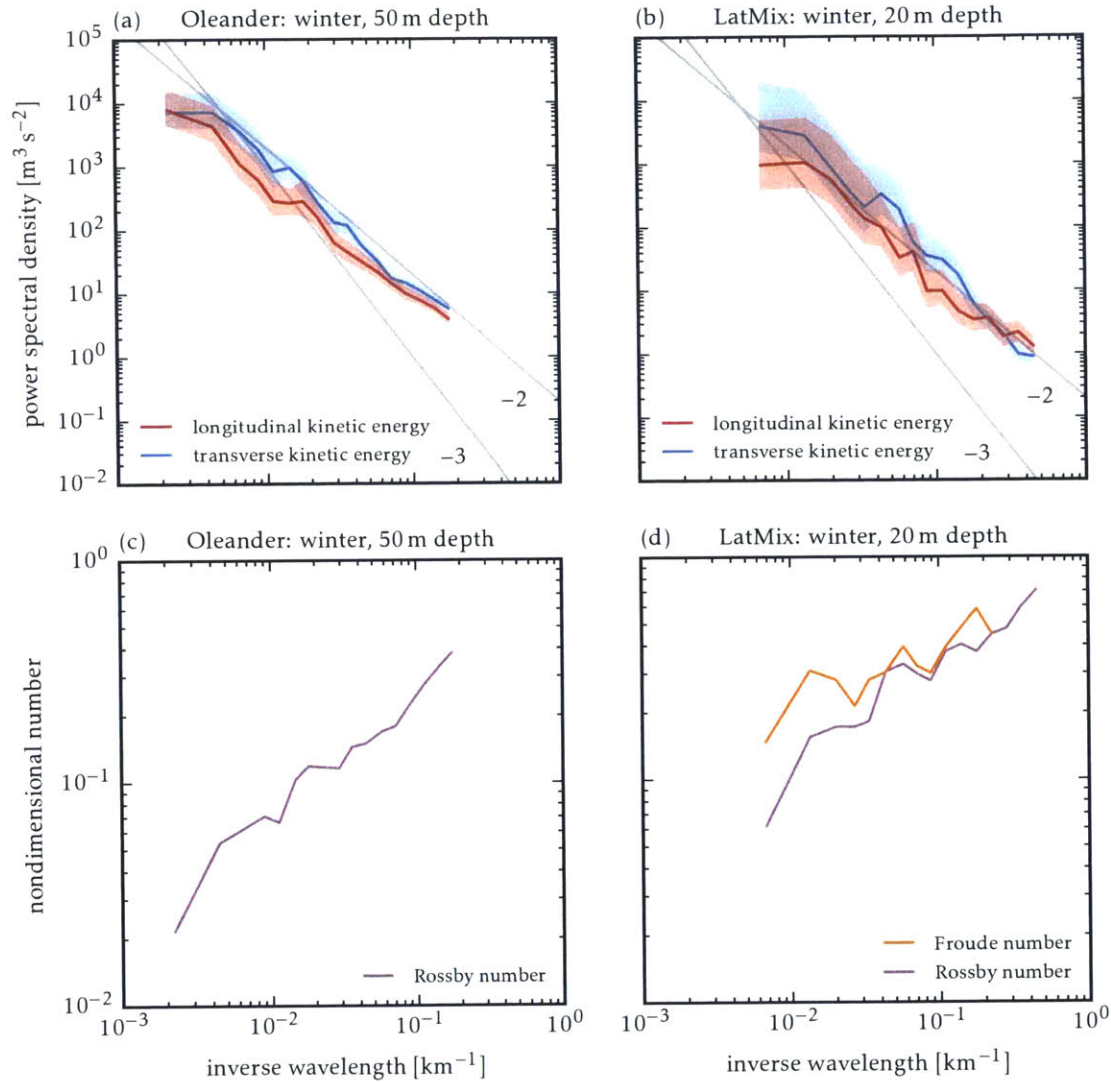Figure 5.5: (a) Power spectra of transverse and longitudinal velocities at 50 m depth for the Oleander winter data. (b) Power spectra of transverse and longitudinal velocities at 20 m depth for the LatMix winter experiment. (c) Rossby number at 50 m depth for the Oleander winter data. (d) Rossby and Froude numbers at 20 m depth for the LatMix winter experiment. The light shadings are 95% confidence intervals.

5.3 Discussion

The near-surface observations are characterized by a pronounced seasonal cycle in kinetic and potential energy: submesoscale flows are much more energetic in deep winter mixed layers than in the seasonal thermocline in summer. In contrast, the mesoscale energy remains fairly constant throughout the year. The near-surface submesoscale spectral roll-off changes from a rapid k^{-3} in summer to a frontal k^{-2} in winter. In winter, submesoscale energies are fairly constant with depth throughout the mixed layer and decay rapidly below its base, transitioning to a k^{-3} regime in the main thermocline.

We can rationalize these results in terms of quasi-geostrophic turbulence theory. A very weak forward transfer of energy below the scale of the deep mesoscale instability prevails in the seasonal thermocline in summer and in the main thermocline throughout the year. In winter, deep mixed layers allow a secondary baroclinic instability in the mixed layer that converts potential to kinetic energy in the submesoscale, over a range of scales around the mixed layer deformation radius between 1–10 km. This injection of kinetic energy throughout the depth of the mixed layer drives an inverse cascade of kinetic energy to larger scales and energizes the entire range between the mesoscale and the mixed layer deformation radius.

The observations are consistent with realistic high-resolution numerical simulations (Mensa et al., 2013; Sasaki et al., 2014) and thereby suggest that these simulations capture the leading-order dynamics of submesoscale turbulence. This consistency further builds confidence in the instability mechanisms, because in the simulations one can directly diagnose a seasonal cycle in the release of available potential energy through mixed layer instabilities and an inverse cascade of kinetic energy thus created (Mensa et al., 2013; Sasaki et al., 2014). While this inverse cascade is not directly diagnosed from data, two crucial ingredients are. First, the lack of a spectral peak at the instability scale indicates that the instability has grown to finite amplitude and become fully turbulent. Second, the diagnosed horizontal non-divergence indicates that the flow is to leading-order in geostrophic balance, which is necessary for turbulent scale interactions to preferentially transfer energy to large scales.

The observed seasonal cycle of submesoscale energy levels, the vertical structure of the flow, and its horizontal non-divergence in winter are all consistent with the instability mechanism. In conjunction with the results of numerical simulations, the observations therefore strongly favor the hypothesis that the submesoscale is energized through a baroclinic instability in the mixed layer. This consistency, of course, does not constitute a proof of the instability mechanism, and further progress can be made by deriving additional falsifiable predictions or by obtaining additional observations. But at the moment, the instability mechanism appears to be the only available explanation for the observations on hand.

Previous studies suggested that mesoscale-driven surface frontogenesis generates the submesoscale kinetic energy in the upper ocean (Lapeyre and Klein, 2006). This process

appears to be weak in the western subtropical North Atlantic. In summer, when the mixed layer instability is not active, the submesoscale spectra roll off like k^{-3}—surface frontogenesis, which can occur year-round, would generate steps in the velocity field and thus produce k^{-2} spectra. It appears that in this region the summertime flow is dominated by interior quasi-geostrophic turbulence, even close to the surface.

The mixed layer instability mechanism is likely to dominate the generation of submesoscale kinetic energy in all parts of the ocean that are characterized by deep mixed layers. For example, the mechanism appears to be at work in the entire Gulf Stream and Kuroshio regions in winter, as found in high-resolution numerical simulations (Mensa et al., 2013; Sasaki et al., 2014). We also expect this mechanism to be at work in parts of the Southern Ocean, where winter mixed layers can reach as deep as 500 m (Dong et al., 2008).

Our results further suggest that a strong seasonal cycle in submesoscale energy is probably characteristic of regions with relatively strong interior potential vorticity gradients reversing sign at depth, like the western subtropical North Atlantic analyzed here (Tulloch et al., 2011). With such potential vorticity profiles, the mesoscale field is generated through a Phillips-type baroclinic instability (Phillips, 1954), which energizes low-mode eddies and transfers little energy to scales smaller than the mesoscale instability scale. Surface buoyancy gradients are relatively weak, rendering any mesoscale-driven surface frontogenesis ineffective. In other regions, where surface buoyancy gradients are larger and deep potential vorticity gradients weaker, Charney-type baroclinic instabilities and mesoscale-driven surface frontogenesis may become a leading-order driver of submesoscale turbulence in the upper ocean (Roullet et al., 2012).

Submesoscale turbulence is associated with strong flows along steep isopycnals. These generate large vertical fluxes of physical and chemical tracers (Capet et al., 2008a) and may impact the exchange of these tracers between the mixed layer and the interior ocean. The submesoscale enhancement of fluxes of heat and salt has been suggested to affect the water mass properties and circulation of the permanent thermocline (Lévy et al., 2010). Furthermore, submesoscale fluxes of nutrients between the nutrient-depleted mixed layer and the nutrient-rich thermocline are believed to play an important role in maintaining primary production in subtropical gyres (Klein and Lapeyre, 2009). It remains to be studied, however, how the seasonality in submesoscale turbulence affects these exchanges. Stommel (Stommel, 1979) argued that the properties of waters subducted from the mixed layer into the permanent thermocline are set in winter, when mixed layers are deepest. Our study has shown that winter is also the time when submesoscale turbulence is most vigorous—just when Stommel's demon opens the trapdoor. An impact of submesoscale turbulence on the physical and chemical properties of interior waters therefore appears plausible.

5.A Methods

5.A.1 Numerical simulation

The snapshots of buoyancy gradients shown in Fig. 5.2 are taken from a numerical simulation of the Gulf Stream region performed with the Regional Ocean Modeling System (Shchepetkin and McWilliams, 2005). This simulation has a horizontal resolution of 750 m and 50 vertical levels. The model domain spans 1050 km by 900 km, which provides generous padding to the domain shown in the snapshots. Boundary conditions are supplied by a lower-resolution simulation that spans the Atlantic basin (Mason et al., 2010). The simulation is forced by daily winds and diurnally modulated surface fluxes. The modeling approach is described in detail in Gula et al. (2015).

5.A.2 In-situ observations

The Oleander data were collected in 2005–2013 with a 75 kHz ADCP, averaged over 8-meter depth bins and 4.5–6-minute intervals, which at a ship speed of about 8 m s^{-1} results in an average spacing of 2.0–2.5 km. Data west of 68°W and north of 36.5°N are discarded, so the inhomogeneities of the Gulf Stream extension do not affect the results (Fig. 5.1). Subsequently, transects with fewer than 190 data points are discarded. The measured velocities are transformed into a coordinate system aligned with the ship track and interpolated with cubic splines onto a regular grid with a spacing of 2.6 km. After selecting for the season, a Hann window is applied and the Fourier transforms are averaged over the different transects to form the spectra. The summer spectrum is an average over 46 transects; the winter spectrum is an average over 11 transects. The spectra are further averaged over ten wavenumber bins per decade. The spectral amplitudes at the smallest resolved scales are affected by the averaging and interpolation procedures, but an assessment of these effects using synthetic signals suggests that the impacts are small and that they do not affect our conclusions, which depend mostly on larger scales.

A comparison to the Oleander data collected with a 150 kHz ADCP in 1994–2004 reveals that while the two data sets are consistent in summer, the 1994–2004 data set exhibits less submesoscale energy in winter than the 2005–2013 data set. More accurate navigational data can be used in the processing of the 2005–2013 data, which makes it more reliable and leads us to suspect that the submesoscale energy is artificially reduced in the 1994–2004 data set during winter.

The LatMix velocity data were collected with 75 kHz ADCPs aboard the RV Endeavor (summer) and RV Atlantis (winter). They are averaged over 8-meter depth bins and 5-minute intervals, which at a ship speed of about 3 m s^{-1} results in an average spacing of about 1 km. Four straight large-scale transects are used from the summer experiment, three from the winter experiment (Fig. 5.1). The winter data were previously analyzed

by Shcherbina et al. (2013). The measured velocities are transformed into a coordinate system aligned with the ship track and interpolated onto a regular grid with a spacing of 1 km. The spectra are obtained following the same procedure as described above for the Oleander data. Estimates of the noise level are calculated as in Shcherbina et al. (2013).

The LatMix buoyancy data were collected using Moving Vessel Profilers (MVP), sampling the water column 5–100 m in summer and 10–250 m in winter. We use three transects occupied by the RV Endeavor in summer (Fig. 5.1) and three transects occupied by the RV Atlantis in winter (same as for velocity). Data from the nearly vertical downcasts are averaged over 1-meter depth bins; data from upcasts are discarded. The data are further interpolated onto a regular along-track grid with a spacing of 2 km. The stratification is obtained by fitting a second-order polynomial to the horizontally averaged buoyancy profile in the range 10–100 m in summer and 20–200 m in winter. The spectra are obtained following the same procedure as described above for the velocities.

CHAPTER 6

THE ROLE OF MIXED LAYER INSTABILITIES IN
SUBMESOSCALE TURBULENCE

6.1 Introduction

The upper ocean is host to energetic flows at scales smaller than the order 100 km mesoscale eddies. Sharp surface fronts associated with strong along-front currents are generated in high-resolution numerical simulations (e.g. Capet et al., 2008a; Klein et al., 2008) and are observed in the wintertime midlatitude ocean (Shcherbina et al., 2013; Chapter 5). These submesoscale flows at scales 1–100 km are associated with large vertical fluxes of both physical and biogeochemical tracers that have been argued to regulate the oceanic heat and carbon uptake in global warming scenarios (Capet et al., 2008a; Klein and Lapeyre, 2009; Ferrari, 2011; Lévy et al., 2012; Mahadevan, 2014). Current global ocean models do not resolve submesoscale flows, so these fluxes must be represented by parameterizations that should be based on physical understanding.

Despite the attention received by submesoscale flows in both the theoretical and observational oceanographic communities, the dynamics that generate them remain a topic of debate. Two mechanisms have been proposed: mesoscale-driven surface frontogenesis (Lapeyre and Klein, 2006; Roullet et al., 2012) and baroclinic mixed layer instabilities (Boccaletti et al., 2007). It is important to understand the differences between these two mechanisms, because they produce—as we shall see—distinct submesoscale flow characteristics and vertical fluxes.

The essential physics of mesoscale-driven surface frontogenesis can be understood with quasi-geostrophic (QG) dynamics (Stone, 1966a). A mesoscale strain field sharpens lateral buoyancy gradients at the surface more effectively than in the interior of the ocean. An ageostrophic circulation develops in response to the increasing lateral buoyancy gradient, as described by the omega equation (e.g. Hoskins et al., 1978). In the interior, this circulation acts to weaken the lateral buoyancy gradient: light water downwells on the dense side and dense water upwells on the light side of the gradient. At the surface, however, the vertical velocity must vanish and the ageostrophic circulation cannot counteract the increase in lateral buoyancy gradient—the mesoscale strain field is left unopposed to create strong submesoscale surface fronts.

Mixed layer instabilities, on the other hand, can energize submesoscale flows by releasing available potential energy stored in large- and mesoscale buoyancy gradients in the surface mixed layer. The weak stratification in deep wintertime mixed layers allows baroclinically unstable modes to rapidly amplify (Haine and Marshall, 1998). Much like deep mesoscale modes in the ocean interior (e.g. Gill et al., 1974), these mixed layer modes slide dense water under light water, but their horizontal scale is 1–10 km and they grow on time scales of order 1 day (Boccaletti et al., 2007).

The presence of a seasonal cycle in submesoscale turbulence suggests that baroclinic mixed layer instabilities are an important aspect of upper-ocean dynamics. Both modeling (Mensa et al., 2013; Sasaki et al., 2014) and observations (Chapter 5) show that submesoscale turbulence is energized in winter and suppressed in summer. Mixed layer instabilities are expected to undergo a strong seasonal cycle, following the seasonal cycle of

106

the mixed layer depth and the associated mixed layer potential energy. Mesoscale-driven surface frontogenesis, on the other hand, is not expected to vary seasonally, because mesoscale eddies do not undergo a strong seasonal cycle (cf. Qiu, 1999; Qiu and Chen, 2004). A full understanding of how submesoscale turbulence is energized by baroclinic mixed layer instabilities, however, is not as well established as that of mesoscale-driven surface frontogenesis.

The simplest model capturing the essence of mesoscale-driven surface frontogenesis is the surface QG model (Blumen, 1978; Held et al., 1995). It assumes an infinitely deep ocean with constant stratification and vanishing interior QG potential vorticity (PV),

$$q = \nabla^2 \psi + \frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial \psi}{\partial z}\right) = 0, \tag{6.1}$$

where q is the PV, ψ is the geostrophic streamfunction, f is the (constant) Coriolis frequency, and N the buoyancy frequency. The streamfunction is related to the horizontal flow by $u = (-\partial \psi/\partial y, \partial \psi/\partial x)$. The evolution of the flow is completely determined by the lateral advection of buoyancy at the surface, which represents the no-normal-flow boundary condition (i.e. the vertical advection of the background stratification vanishes),

$$\frac{\partial b}{\partial t} + J(\psi, b) = 0, \tag{6.2}$$

where $b = f\partial \psi/\partial z$ is buoyancy and

$$J(\psi, b) = \frac{\partial \psi}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial b}{\partial x} \tag{6.3}$$

is the Jacobian operator. The surface buoyancy supplies the boundary condition for the elliptic problem (6.1). Straining by mesoscale eddies creates sharp buoyancy gradients associated with strong flows at the surface. Filamentary instabilities eventually lead to fully turbulent dynamics (Held et al., 1995). Kolmogorov-like dimensional arguments (Kolmogorov, 1941) predict that the fully turbulent surface kinetic and potential energy spectra scale like $K_{k_h} = P_{k_h} \sim k_h^{-5/3}$ in a submesoscale inertial range in which surface potential energy b^2/N^2 is cascaded to small scales (Blumen, 1978)—k_h is the horizontal wavenumber. This prediction implies that the gradients of velocity and buoyancy (i.e. fronts), whose spectra scale like $k_h{}^2 K_{k_h}$ and $k_h{}^2 P_{k_h}$, are stronger at small submesoscales (large k_h) than at large submesoscales (small k_h). The forward cascade of surface potential energy occurs in conjunction with an inverse cascade of surface kinetic energy fed by the release of potential energy through slumping fronts (Capet et al., 2008e). The submesoscale energy generated by surface QG turbulence is surface-trapped: modes decay exponentially in the vertical, with small-scale modes decaying more rapidly than large-scale modes (e.g. Scott,

107

2006).[1]

If non-QG dynamics are taken into account, ageostrophic advection of buoyancy further accelerates frontogenesis and potentially leads to frontal collapse, the formation of true discontinuities in buoyancy (Hoskins and Bretherton, 1972). In this case, the submesoscale energy spectrum is modified to $E_{k_h} \sim k_h^{-2}$ (Boyd, 1992) and the decay in the vertical becomes less rapid (Badin, 2012). An additional modification of the dynamics by non-QG effects is that the release of potential energy leads to near-surface restratification (Hakim et al., 2002; Lapeyre et al., 2006). This effect is neglected in QG dynamics, where stratification is fixed. Non-QG surface frontogenesis also induces a finite forward flux of kinetic from small submesoscales (order 1 km) to dissipation scales, which is not present in QG dynamics (Capet et al., 2008b,a; Klein et al., 2008; Molemaker et al., 2010). Despite these omissions, surface QG turbulence predicts many of the characteristics found in primitive equation simulations of mesoscale-driven surface frontogenesis (e.g. Klein et al., 2008). Our working hypothesis is thus that surface QG turbulence faithfully captures the leading-order dynamics of the balanced flow in mesoscale-driven surface frontogenesis.

The upper ocean does not have a nearly constant PV, however, contrary to what is assumed in surface QG turbulence. Instead, a weakly stratified mixed layer typically overlies a strongly stratified thermocline (Fig. 6.1). There is a sharp step-like increase in stratification at the base of the mixed layer, corresponding to a step-like increase in background PV. This PV jump is dynamically important, because it supports edge waves that have the potential to interact with surface edge waves and thus produce a baroclinic instability in the mixed layer (e.g. Haine and Marshall, 1998). This linear instability is to leading order captured by an Eady model with a rigid interface at the base of the mixed layer (Eady, 1949). Corrections due to ageostrophic effects and a moveable interface at the mixed layer base can be computed (Stone, 1966b; Boccaletti et al., 2007), but for typical wintertime conditions the instability scale and growth rate are qualitatively captured by Eady's QG model. When baroclinic mixed layer instabilities grow to finite amplitude, turbulent scale interactions distribute energy across scales. Because of the rotational constraint, they transfer energy preferentially to larger scales. These nonlinear dynamics have been studied in idealized mixed layer models, where baroclinic mixed layer instabilities grow on a prescribed front (e.g. Boccaletti et al., 2007; Fox-Kemper et al., 2008).

In the real ocean, baroclinic mixed layer instabilities occur in the presence of an energetic mesoscale eddy field, so mixed layer modes can grow on mesoscale buoyancy gradients and can be sheared by mesoscale strain fields. Realistic submesoscale-permitting models capture these dynamics (e.g. Mensa et al., 2013; Sasaki et al., 2014), but the models' complexity makes it hard to distill the essence of the dynamics and even these simula-

---

[1] Heuristic extensions of surface QG ideas have been developed to infer interior flows from surface properties (Lapeyre and Klein, 2006). These extensions are diagnostic in nature and do not attempt to describe the evolution of the flow or make predictions for submesoscale energy levels. We therefore limit our discussion of surface QG dynamics to the case with zero interior PV.
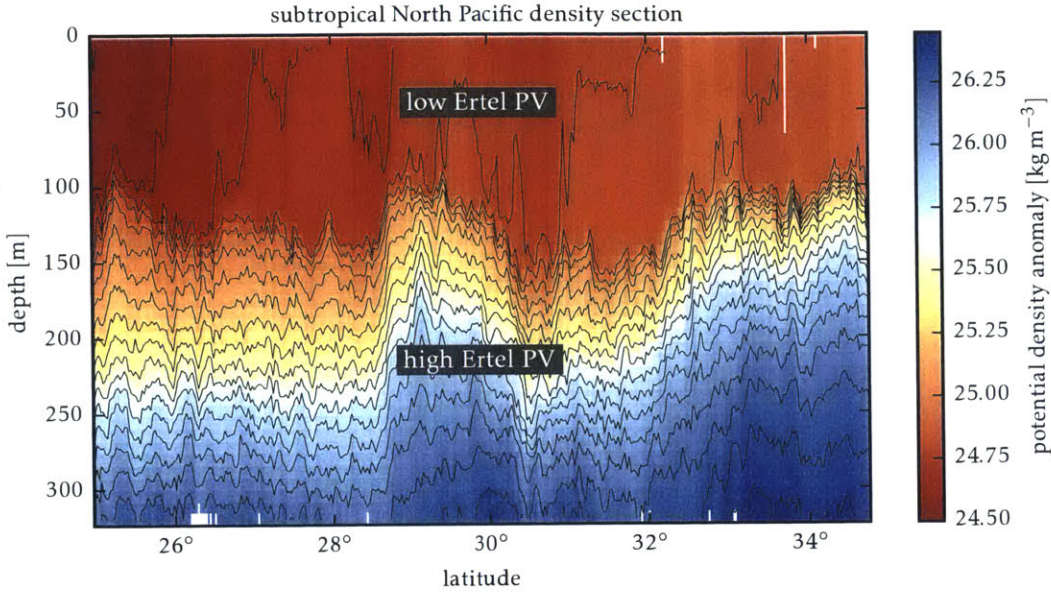
Figure 6.1: Potential temperature section from the wintertime eastern subtropical North Pacific. The data were obtained using a conductivity–temperature–depth sensor towed on a sawtooth profile along 140°W as part of the Spice experiment on Jan. 28 to Feb. 2, 1997. For more information on the data, see Ferrari and Rudnick (2000).

tions only marginally resolve many submesoscale phenomena. In this chapter, we explore submesoscale dynamics by formulating a QG model that allows both mixed layer and thermocline instabilities. This simple model of submesoscale turbulence energized by baroclinic mixed layer instabilities captures salient features of wintertime observations of submesoscale flows. If the mixed layer in this model is eliminated, the submesoscale dynamics dynamics revert to surface QG turbulence, which allows a straightforward comparison of the two mechanisms that can energize submesoscale turbulence.

We use QG scaling to formulate the dynamics of the model, which requires small Rossby and Froude numbers (e.g. Pedlosky, 1987). Typical mesoscale Rossby and Froude numbers are on the order 0.1 and increase slowly with wavenumber if the submesoscales are energetic, reaching order 1 at scales of order 1 km (Chapter 5). While the QG approximation does not apply anymore at these small scales, the QG system can be expected to capture the leading-order dynamics over the 10–100 km range.

A major limitation of QG dynamics in representing mixed layer instabilities is that QG scaling does not allow for restratification. The weak mixed layer stratification will be fixed in our QG model, whereas in reality there is a competition between restratification and atmospherically forced mixed layer turbulence, which tends to keep the mixed layer deep and unstratified. The assumption is that the slower balanced dynamics described

109

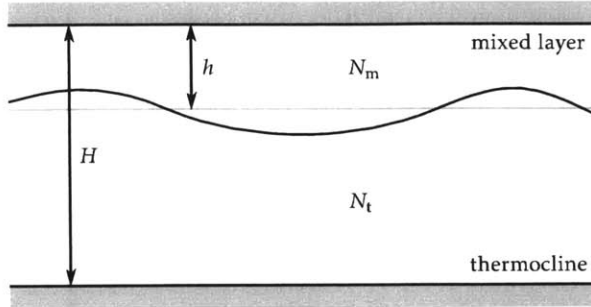Figure 6.2: Schematic of the model setup in a vertical–horizontal plane. There are rigid surfaces at $z = 0$ and $z = -H$ and a deformable interface at the mean depth $z = -h$, separating layers of constant stratifications, N_m in the mixed layer and N_t in the thermocline.

by the QG model develop on top of this background state, which is maintained by fast small-scale turbulence. A full description of mixed layer dynamics will eventually need to consider the interplay of the fast and slow dynamics, a topic we hope to address in a future study (cf. Chapter 7; Hamlington et al., 2014). We take up the discussion of how other non-QG effects may alter the dynamics toward the end of the chapter.

As described above, our inquiry into the dynamics of a weakly stratified mixed layer coupled to a strongly stratified thermocline is motivated by the study of the submesoscale upper ocean. The model we present and the dynamics we describe, however, have relevance for the atmosphere as well, where a weakly stratified troposphere is capped by a strongly stratified stratosphere (e.g. Eady, 1949). The dynamics may also apply to the atmospheres of gas giants (e.g. Seiff et al., 1998) or other geophysical fluids that have layers of different stratification.

We formulate the model and give some physical intuition for its behavior in Section 6.2. In Section 6.3, we investigate the linear dynamics of the model to understand its stability properties. These linear dynamics are suggestive of the fully nonlinear turbulent dynamics that we address in Section 6.4, where we analyze the energy spectra and fluxes for cases with and without baroclinic mixed layer instabilities. We compare the results to observations in Section 6.5 and conclude in Section 6.6.

6.2 Model formulation

Consider two layers with constant stratification and constant mean shear on an f-plane, so that each layer has constant PV (Fig. 6.2). The upper layer represents the mixed layer, which has a mean depth h, stratification N_m, and mean zonal shear Λ_m that is in thermal wind balance with the mean meridional buoyancy gradient $-f\Lambda_m$. The lower layer represents the thermocline and has stratification N_t and mean zonal shear Λ_t that

is in thermal wind balance with the mean meridional buoyancy gradient $-f\Lambda_t$. The total depth is H. The layers are coupled through a deformable interface; flat rigid boundaries are assumed at both the surface and the bottom. The presence of a rigid bottom at the base of the thermocline is not realistic, but we will show that the bottom layer still captures the key thermocline physics relevant to our study. A weakly stratified abyssal layer could be included but is omitted here for simplicity, because it does not significantly affect the surface submesoscale dynamics of interest here. The approximation that the stratification is discontinuous at the base of the mixed layer is appropriate at horizontal scales larger than the deformation radius Nd/f associated with the transition depth d between the mixed layer and the thermocline (Smith and Bernard, 2013). The transition at the base of the mixed layer is typically quite sharp (Fig. 6.1), so this deformation radius is much smaller than the submesoscales we are interested in here.

The assumption of a uniform PV within the two layers greatly simplifies the dynamics. PV conservation within the layers is trivial, as in the classic Eady (1949) problem. The flow in the interior of the layers is obtained by solving (6.1), with the boundary conditions supplied by the distribution of buoyancy at the surface and bottom and by matching conditions at the interface between the mixed layer and the thermocline.

In QG, the buoyancy anomaly b is governed by the horizontal advection of buoyancy anomalies by the geostrophic flow and by the vertical advection of the background buoyancy field,

$$\frac{\partial b}{\partial t} + J(\psi, b) + wN^2 = 0. \tag{6.4}$$

At the surface and bottom, where the vertical velocity w vanishes, buoyancy anomalies are conserved under horizontal advection and (6.4) reduces to (6.2). To ensure that pressure is continuous at the interface, we require that the streamfunction ψ is continuous. Mass conservation requires that the vertical velocity w also is continuous. These conditions are applied at $z = -h$, consistent with QG scaling. The conservation equations for buoyancy just above the interface at $z = -h$,

$$\frac{\partial b^+}{\partial t} + J(\psi_1, b^+) + wN_m^2 = 0, \qquad b^+ = f\frac{\partial \psi}{\partial z}(-h^+), \tag{6.5}$$

and just below the interface,

$$\frac{\partial b^-}{\partial t} + J(\psi_1, b^-) + wN_t^2 = 0, \qquad b^- = f\frac{\partial \psi}{\partial z}(-h^-), \tag{6.6}$$

can then be combined to eliminate w, where ψ_1 denotes the streamfunction at $z = -h$. This gives a conservation law for the quantity

$$\theta_1 = f\left(\frac{b^+}{N_m^2} - \frac{b^-}{N_t^2}\right), \tag{6.7}$$

111

which is simply advected by the horizontal flow at the interface,

$$\frac{\partial \theta_1}{\partial t} + J(\psi_1, \theta_1) = 0. \tag{6.8}$$

It should be noted that this does not ensure buoyancy to be continuous at $z = -h$. Instead, there is an implied interface displacement and buoyancy is continuous across the displaced interface. The displacement is small and, consistent with QG scaling, the matching conditions are applied at $z = -h$.

The quantity θ_1 is nothing but the integrated PV associated with the interface displacement, as can be verified by integrating

$$q = \nabla^2 \psi + f \frac{\partial}{\partial z} \left(\frac{b}{N^2} \right) \tag{6.9}$$

across the interface. (The relative vorticity term vanishes because ψ is continuous across the interface.) While there are no PV anomalies within the two layers, the displacement of the interface between the layers induces a PV anomaly that, according to (6.8) and consistent with QG dynamics, is advected by the geostrophic flow at $z = -h$. The conservation equation (6.8) has been used to study the dynamics of the tropopause, which is similarly an interface between the weakly stratified fluid in the troposphere and the strongly stratified fluid in the stratosphere (Eady, 1949; Rivest et al., 1992; Juckes, 1994; Held et al., 1995).

The two-layer model can equivalently be interpreted as consisting of three PV sheets:

$$q = \theta_0 \delta(z) + \theta_1 \delta(z + h) + \theta_2 \delta(z + H), \tag{6.10}$$

where δ is Dirac's delta function and $\theta_0 = -fb/N_m^2$ at $z = 0$, $\theta_2 = fb/N_t^2$ at $z = -H$, and θ_1 at $z = -h$ is given in (6.7). PV is advected by the geostrophic flow, so

$$\frac{\partial \theta_j}{\partial t} + J(\psi_j, \theta_j) = 0, \tag{6.11}$$

where $j = 0, 1, 2$ and ψ_j is the streamfunction at the level corresponding to θ_j. This formulation is simply an extension of Bretherton's (1966) representation of boundary conditions to include an interior PV sheet due to the deflection of an interface between layers of different stratification.

Note that even though θ_1 is only advected by the geostrophic flow, this does not imply that the vertical velocity vanishes at the interface, much like the fact that interior PV anomalies in the QG system are only advected by the geostrophic flow does not imply that the vertical velocity vanishes. The vertical velocity is implicit in the dynamics and can be solved for using the omega equation (e.g. Hoskins et al., 1978).

To complete the dynamics, we require an inversion relation that allows us to obtain the streamfunctions ψ_j from the conserved quantities θ_j. For simplicity, we consider a
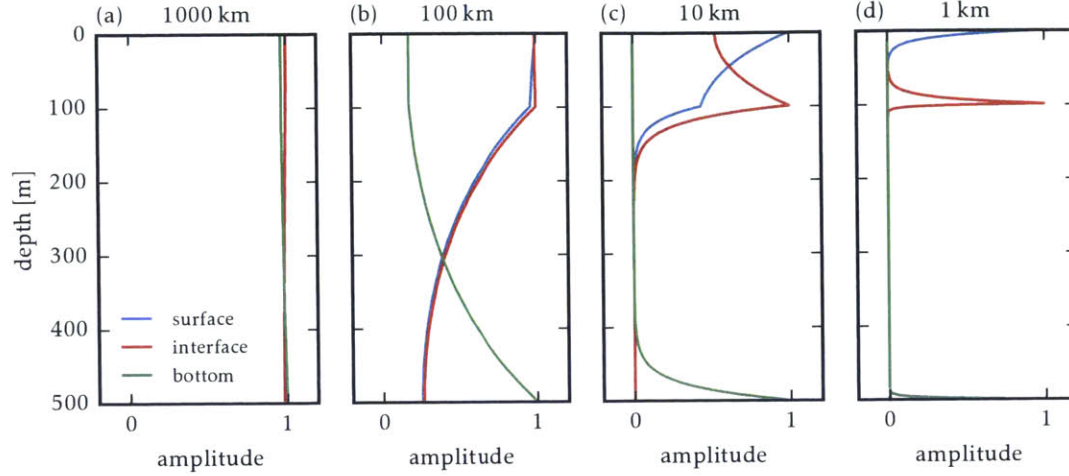
Figure 6.3: Vertical structure of streamfunction amplitude associated with anomalies of θ_0 (surface), θ_1 (interface), and θ_2 (bottom). Shown are the vertical profiles for θ_j anomalies with different horizontal wavenumbers $k_h = 2\pi/\lambda$. The wavelength λ is given in the respective panel title.

doubly-periodic domain and express the inversion relation as a linear equation for Fourier coefficients of the variables θ_j and ψ_j:

$$\hat{\boldsymbol{\theta}} = \mathbf{L}\hat{\boldsymbol{\psi}}, \qquad \boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)^\mathsf{T}, \qquad \boldsymbol{\psi} = (\psi_0, \psi_1, \psi_2)^\mathsf{T}. \tag{6.12}$$

where Fourier transforms are denoted by carets. The matrix $\mathbf{L}$, which depends on the zonal and meridional wavenumbers k and l, is determined by solving

$$-k_h{}^2 \hat{\psi} + \frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial \hat{\psi}}{\partial z}\right) = 0 \tag{6.13}$$

in each layer, where $k_h = (k^2 + l^2)^{1/2}$ is the magnitude of the horizontal wavenumber vector. The first column of $\mathbf{L}$ is determined by setting $\hat{\boldsymbol{\psi}} = (1,0,0)^\mathsf{T}$, solving (6.13) for $\hat{\psi}(z)$, and subsequently calculating θ_0, θ_1, and θ_2. Repeating for $\hat{\boldsymbol{\psi}} = (0,1,0)^\mathsf{T}$ and $\hat{\boldsymbol{\psi}} = (0,0,1)^\mathsf{T}$ gives

$$\mathbf{L} = fk_h \begin{pmatrix} -\dfrac{\coth\mu_m}{N_m} & \dfrac{\operatorname{csch}\mu_m}{N_m} & 0 \\[2ex] \dfrac{\operatorname{csch}\mu_m}{N_m} & -\dfrac{\coth\mu_m}{N_m} - \dfrac{\coth\mu_t}{N_t} & \dfrac{\operatorname{csch}\mu_t}{N_t} \\[2ex] 0 & \dfrac{\operatorname{csch}\mu_t}{N_t} & -\dfrac{\coth\mu_t}{N_t} \end{pmatrix}, \tag{6.14}$$

where $\mu_m = N_m k_h h/f$ and $\mu_t = N_t k_h (H-h)/f$ are nondimensional wavenumbers. This 3×3 matrix can easily be inverted.

This model can be generalized to an arbitrary number of layers of constant stratification and shear, which may be a useful way to approximate more realistic stratification

and shear profiles. This is discussed in Appendix 6.A. The model can also be extended to include a density jump at the interface, as is sometimes present at the base of the mixed layer. The formulation is given in Appendix 6.B. Here we consider only the case of a continuous density profile, which is simpler and captures the essential physics of the submesoscale ocean.

To build intuition for the dynamics of the model, we illustrate the vertical structure of the flow associated with anomalies of θ_j at the surface, the interface, and the bottom. Here and throughout the chapter, we use the parameters given in Table 6.1, which are typical of the wintertime midlatitude ocean.[2] At the largest scales, for θ_j anomalies with wavelength $\lambda = 1000$ km or $k_h \ll f/N_t H$, the flow is nearly depth-independent, irrespective of which level the anomaly is at (Fig. 6.3a). Around the thermocline deformation radius, at $\lambda = 100$ km or $k_h \sim f/N_t H \ll f/N_m h$, flow anomalies significantly decay in the thermocline, while the flow is nearly uniform across the mixed layer (Fig. 6.3b). Surface (θ_0) and interface (θ_1) anomalies still induce significant flow at the bottom and vice versa. Around the mixed layer deformation radius, at $\lambda = 10$ km or $k_h \sim f/N_m h \gg f/N_t H$, surface ($\theta_0$) and interface ($\theta_1$) anomalies induce very little flow at the bottom and vice versa (Fig. 6.3c). The flow now varies significantly across the mixed layer, but surface anomalies (θ_0) still induce significant flow at the interface and vice versa. At $\lambda = 1$ km or $k_h \gg f/N_m h$, all levels are decoupled: θ_j anomalies on any one of the levels induce very little flow at the other levels (Fig. 6.3d).

The dependence of the vertical flow structure on the horizontal scale of the anomalies illuminates the qualitative dynamics of the model. At the largest scales, the flow is essentially depth-independent and follows two-dimensional dynamics. At scales $k_h \sim f/N_t H$, around the thermocline deformation radius, surface or interface anomalies can interact with bottom anomalies, allowing phase locking and a thermocline instability. At scales $k_h \sim f/N_m h$, around the mixed layer deformation radius, surface and interface anomalies can interact, enabling an instability in the mixed layer. Bottom anomalies, on the other hand, are decoupled, so there is no thermocline instability at these scales. At the smallest scales, all three levels are independent and follow surface QG dynamics.

---

[2]The values listed imply a mixed layer Richardson number of $N_m^2/\Lambda_m^2 = 400$, which is larger than the order-1 Richardson numbers typically considered (e.g. Boccaletti et al., 2007). The relatively large Richardson number is the result of a relatively weak shear, which is chosen such that realistic energy levels are reached in the nonlinear simulations described below. QG dynamics overestimate the baroclinic growth rate for small Richardson numbers (Stone, 1966b), which would result in unrealistically strong mixed layer instabilities if a larger shear was chosen. It should also be noted that the leading-order QG dynamics can be rescaled to different Richardson numbers.
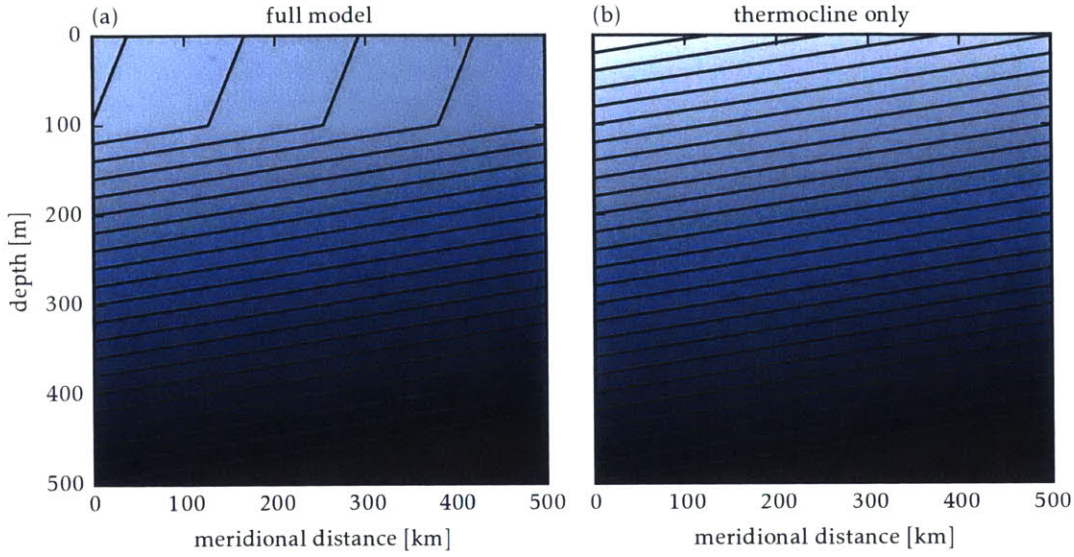
Figure 6.4: Mean buoyancy structure in the meridional–vertical plane for (a) the full model and (b) the thermocline-only model. The contours show isopycnals; light shading indicates more buoyant fluid.

6.3 Linear stability analysis

We now analyze the linear stability of the model formulated above. This linear analysis reveals the nature of the instabilities that fuel the nonlinear turbulence, which we describe in the next section.

Blumen (1979) analyzed short-wave instabilities in the atmosphere using a model consisting of two coupled constant-PV layers. He performed a linear stability analysis equivalent to what will be presented here. For completeness, we repeat the analysis in the context of upper ocean dynamics to emphasize the aspects most relevant for the nonlinear regime.

We consider the linear stability of normal-mode perturbations to a zonal flow with constant vertical shear Λ_m in the mixed layer and Λ_t in the thermocline (Fig. 6.4). The linearized conservation equation for the perturbations from this mean state, written in Fourier space, is

$$\frac{\partial \hat{\theta}}{\partial t} + ik\mathbf{U}\hat{\theta} + ik\Gamma\hat{\psi} = 0, \tag{6.15}$$

where the mean zonal flows and mean meridional PV gradients at the respective levels

115

Table 6.1: Parameters used throughout this chapter unless otherwise noted. These are typical of the wintertime midlatitude ocean.

Parameter	Symbol	Value
Mixed layer depth	h	100 m
Total depth	H	500 m
Mixed layer stratification	N_m	2×10^{-3} s^{-1}
Thermocline stratification	N_t	8×10^{-3} s^{-1}
Mixed layer shear	Λ_m	10^{-4} s^{-1}
Thermocline shear	Λ_t	10^{-4} s^{-1}
Coriolis frequency	f	10^{-4} s^{-1}
Domain size	a	500 km
Numerical resolution	Δx	~ 1 km

are represented by the diagonal elements of the matrices U and Γ:

$$\mathsf{U} = \mathrm{diag}\left(0, -\Lambda_m h, -\Lambda_m h - \Lambda_t(H - h)\right), \tag{6.16}$$

$$\Gamma = \mathrm{diag}\left(f^2 \Lambda_m / N_m^2, -f^2 \Lambda_m / N_m^2 + f^2 \Lambda_t / N_t^2, -f^2 \Lambda_t / N_t^2\right). \tag{6.17}$$

The system is Galilean invariant, so we are free to set the mean zonal flow to zero at the surface. Using the inversion relation (6.12), we can replace the $\hat{\psi}$ and obtain an equation for the $\hat{\theta}$ coefficients only,

$$\frac{\partial \hat{\theta}}{\partial t} + ik\mathsf{U}\hat{\theta} + ik\Gamma\mathsf{L}^{-1}\hat{\theta} = 0. \tag{6.18}$$

Substituting $\hat{\theta} = \tilde{\theta}e^{i\omega t}$, with complex frequency ω, turns this equation for $\hat{\theta}$ into the eigenvalue problem

$$(\mathsf{U} + \Gamma\mathsf{L}^{-1})\tilde{\theta} = c\tilde{\theta}, \tag{6.19}$$

where the eigenvalue is $c = \omega/k$. The real part of c is the zonal phase speed; the imaginary part gives the growth rate $\sigma = k \,\mathrm{Im}\, c$.

Being a third-order system, (6.19) can be solved analytically, but the solutions are rather complicated and give little useful insight. We instead explore the characteristics of the solutions numerically for the set of parameters given in Table 6.1. We then explain the stability properties and parameter dependencies by considering special limit cases.

6.3.1 Full model

We start by considering the growth rate σ as a function of horizontal wavenumber. The eigenvalue problem (6.19) only depends on $k_h = (k^2 + l^2)^{1/2}$, so the eigenvalue c is a
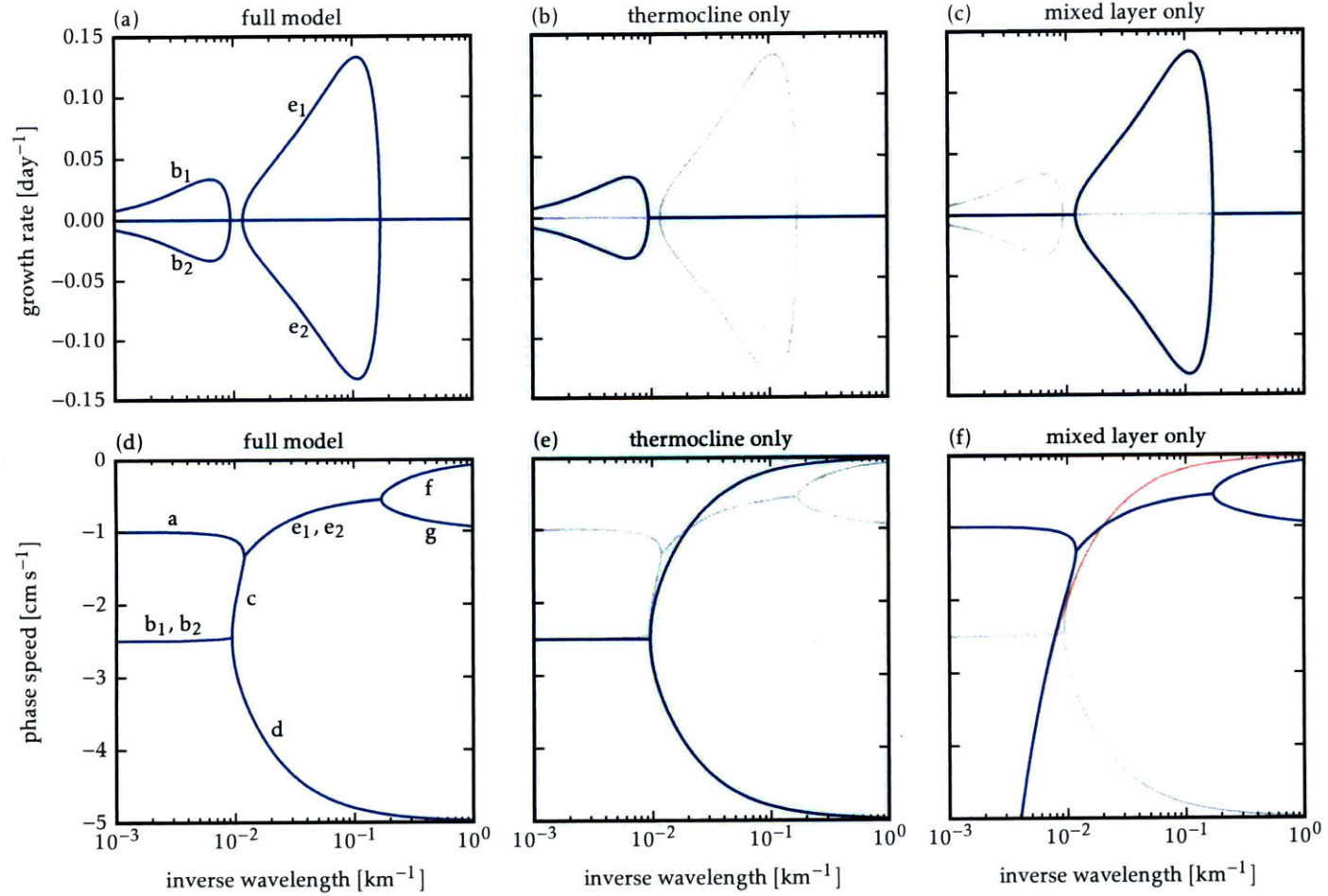
Figure 6.5: Linear stability analysis of the model equations. (a) growth rates and (d) phase speeds of the full model, (b) growth rates and (e) phase speeds of the thermocline-only model, (c) growth rates and (f) phase speeds of the mixed-layer-only model. Growth rates and phase speeds are shown in blue; the growth rates and phase speeds of the full model are overlaid for reference in gray. The phase speed of a surface edge wave is given in faint red in (f).

function of k_h only. For a given k_h, the maximum growth rate $\sigma = k \operatorname{Im} c$ hence occurs at $l = 0$. We therefore consider disturbances with no meridional dependence only.

Plotting the growth rate σ as a function of zonal wavenumber k reveals that there are two lobes of instability: one at the mesoscale and one at the submesoscale (Fig. 6.5a, branches 'b' and 'e'). The maximum growth rates occur at zonal wavelengths of about 160 km (mesoscale) and 9 km (submesoscale). The two lobes can overlap, for example if the mixed layer is deeper or if a density jump at the base of the mixed layer is included (not shown). The submesoscale instability has a peak growth rate much larger than the mesoscale instability. The growth rates are similar to what Boccaletti et al. (2007) found in a linear QG stability analysis of a realistic mean state of the wintertime eastern subtropical North Pacific. The magnitudes are slightly smaller here, because the shear is slightly weaker. But the similarity of the instabilities supports that this model, despite being highly idealized, captures the essential physics of mesoscale and submesoscale instabilities. Whether it also captures the essential physics in the nonlinear regime will be discussed in Sections 6.4–6.6.

The mesoscale and submesoscale instabilities have very different vertical structures, as also noted by Boccaletti et al. (2007). The perturbation streamfunctions—derived from the eigenvectors of (6.19)—show that the fastest-growing mesoscale mode is deep and spans the entire water column (Fig. 6.6a), whereas the fastest-growing submesoscale mode is almost completely confined to the mixed layer, with only weak penetration into the thermocline below 100 m depth (Fig. 6.6b). Both modes exhibit the familiar pattern of baroclinically unstable modes with streamfunction perturbations tilted into the shear, which is necessary to extract potential energy from the mean flow.

Further insight into the dynamics of the model can be gained by considering the propagation speeds of the linear modes in conjunction with their growth rates (Fig. 6.5a,d). Being a third-order system, the model has three normal modes at each wavenumber. In both lobes of instability, the growing modes are conjugate to decaying modes, which have the same phase speeds. This is the familiar phase locking of counter-propagating waves in baroclinic instability (branches 'b' and 'e'). In these unstable wavenumber ranges, there is an additional neutral mode (branches 'a' and 'd'). At wavenumbers with no instability, all three modes have distinct phase speeds—no phase locking occurs. We will discuss the dynamics of the various branches by considering approximations to the full model.

6.3.2 Thermocline only

We start by examining the deep mesoscale instability of the full model. As discussed in the model formulation (Section 6.2), the mesoscale modes are deep and only slightly modified by the presence of the mixed layer. We can understand the mesoscale instability by eliminating the mixed layer altogether and consider a thermocline-only model (Fig. 6.4b). This amounts to setting $h = 0$ (or $N_m = N_t$ and $\Lambda_m = \Lambda_t$) in the full model.

118
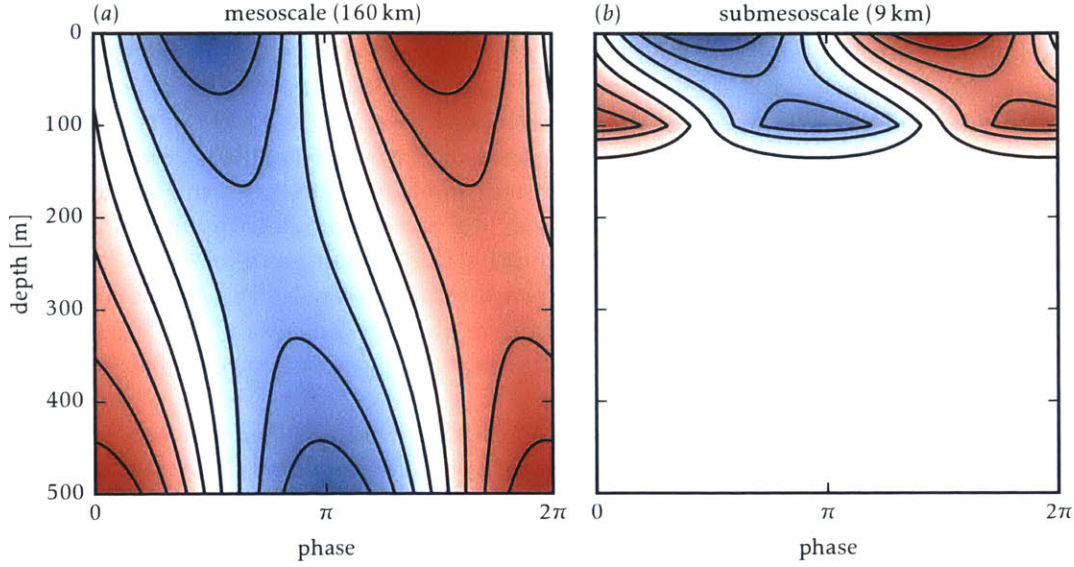
Figure 6.6: Perturbation streamfunction of the most unstable (a) mesoscale and (b) submesoscale modes of the full model. Red and blue shading represents positive and negative values, respectively.

In this limit, the model reduces to one layer with the dynamics controlled by buoyancy advection at the surface and bottom only—the classic Eady (1949) model. The inversion matrix (6.14) reduces to

$$\mathbf{L} = f k_{\mathrm{h}} \begin{pmatrix} -\dfrac{\coth \mu_{\mathrm{t}}}{N_{\mathrm{t}}} & \dfrac{\operatorname{csch} \mu_{\mathrm{t}}}{N_{\mathrm{t}}} \\ \dfrac{\operatorname{csch} \mu_{\mathrm{t}}}{N_{\mathrm{t}}} & -\dfrac{\coth \mu_{\mathrm{t}}}{N_{\mathrm{t}}} \end{pmatrix}, \tag{6.20}$$

and the matrices representing the mean flow are

$$\mathbf{U} = \operatorname{diag}\left(0, -\Lambda_{\mathrm{t}} H\right), \tag{6.21}$$

$$\Gamma = \operatorname{diag}\left(f^2 \Lambda_{\mathrm{t}} / N_{\mathrm{t}}^2, -f^2 \Lambda_{\mathrm{t}} / N_{\mathrm{t}}^2\right). \tag{6.22}$$

Solving the eigenvalue problem (6.19) with these matrices, we find the eigenvalues

$$c = -\frac{\Lambda_{\mathrm{t}} H}{2} \pm \frac{i \Lambda_{\mathrm{t}} H}{\mu_{\mathrm{m}}} \left(\mu_{\mathrm{t}} \coth \mu_{\mathrm{t}} - 1 - \frac{\mu_{\mathrm{t}}^2}{4}\right)^{\frac{1}{2}}, \tag{6.23}$$

where $\mu_{\mathrm{t}} = N_{\mathrm{t}} k_{\mathrm{h}} H / f$ is the nondimensional wavenumber (Eady, 1949).

The solution (6.23) shows that this thermocline-only model has a baroclinic instability near the thermocline deformation radius $N_{\mathrm{t}} H / f$. The maximum growth rate $\sigma = 0.31 f \Lambda_{\mathrm{t}} / N_{\mathrm{t}}$ occurs at $\mu_{\mathrm{t}} = 1.6$ and $l = 0$, which corresponds to a zonal wavelength $\lambda =$

119

$3.9N_tH/f$. The growth curve for this thermocline-only model approximates the mesoscale lobe of the full model very well (Fig. 6.5b). The short-wave cutoff in the Eady model at $\lambda = 2.6N_tH/f$ nearly coincides with the short-wave cutoff of the mesoscale instability in the full model. The phase speed of the phase-locked waves $-\Lambda_tH/2$ very nearly matches the phase speed of the unstable mesoscale mode of the full model (Fig. 6.5e). The split at the short-wave cutoff into surface and bottom modes also features in the full model. In the thermocline-only model, these surface and bottom modes are very nearly Eady edge waves that do not sense the other boundary. The bottom mode of the thermocline-only model very nearly matches that of the full model (branch 'd'). The surface mode of the thermocline-only model traces out branch 'c' of the full model, but then the full model transitions to dynamics associated with the mixed layer that are not present in the thermocline-only model.

This comparison shows that the mesoscale instability of the full model very nearly follows Eady dynamics. The presence of the mixed layer only modifies the characteristics of the instability slightly. At submesoscales, on the other hand, the thermocline-only model has surface QG dynamics, as opposed to the mixed layer dynamics of the full model.

6.3.3 Mixed layer only

Turning our attention to the submesoscale instability, we note that the submesoscale instability peaks around the deformation radius of the mixed layer N_mh/f. Based on the discussion in the model formulation (Section 6.2) and the vertical structure of this instability (Fig. 6.6b), we anticipate that this instability arises from the interaction between anomalies at the surface and at the interface between mixed layer and thermocline.

In a first attempt to isolate the submesoscale instability, we disregard the possibility that surface and interface anomalies induce flow in the thermocline and assume a rigid bottom at the base of the mixed layer. This reduces the full model to an Eady model for the mixed layer, which is the limit of infinite thermocline stratification. This Eady model reasonably approximates the location and magnitude of the peak growth rate with $\lambda = 3.9N_mh/f = 8$ km and $\sigma = 0.31f\Lambda_m/N_m = 0.13$ day^{-1}. The Eady model captures the short-wave cutoff of the full model, but misses the long-wave cutoff. This suggests that the fastest-growing mode approximately follows Eady dynamics as if the thermocline acted like a rigid bottom, but also that larger-scale modes are significantly modified by reaching into the thermocline.

All features of the submesoscale instability are captured if flow in the thermocline is allowed. To still isolate the submesoscale instability, we consider again the layered model but let the thermocline be infinitely deep. That eliminates bottom edge waves, so no mesoscale instability occurs. Eady (1949) considered the upside-down atmospheric analogue to this system, relaxing the assumption that a rigid lid is placed at the tropopause.

120

Our system with no bottom again reduces to two variables; the inversion matrix is

$$\mathbf{L} = fk_{\mathrm{h}} \begin{pmatrix} -\dfrac{\coth \mu_{\mathrm{m}}}{N_{\mathrm{m}}} & \dfrac{\operatorname{csch} \mu_{\mathrm{m}}}{N_{\mathrm{m}}} \\ \dfrac{\operatorname{csch} \mu_{\mathrm{m}}}{N_{\mathrm{m}}} & -\dfrac{\coth \mu_{\mathrm{m}}}{N_{\mathrm{m}}} - \dfrac{1}{N_{\mathrm{t}}} \end{pmatrix} \tag{6.24}$$

and the mean flow is represented by

$$\mathbf{U} = \operatorname{diag}\left(0, -\Lambda_{\mathrm{m}} h\right), \tag{6.25}$$

$$\boldsymbol{\Gamma} = \operatorname{diag}\left(f^2 \Lambda_{\mathrm{m}} / N_{\mathrm{m}}^2, -f^2 \Lambda_{\mathrm{m}} / N_{\mathrm{m}}^2 + f^2 \Lambda_{\mathrm{t}} / N_{\mathrm{t}}^2\right). \tag{6.26}$$

In the case $\Lambda = \Lambda_{\mathrm{m}} = \Lambda_{\mathrm{t}}$, the solution to the eigenvalue problem (6.19) is

$$c = -\frac{\Lambda h}{2}\left(1 + \frac{\alpha}{\mu_{\mathrm{m}}}\right) \pm \frac{i\Lambda h}{\mu_{\mathrm{m}}}\left[\frac{(1 - \alpha^2)(\mu_{\mathrm{m}} - \tanh \mu_{\mathrm{m}})}{\tanh \mu_{\mathrm{m}} + \alpha} - \frac{1}{4}(\mu_{\mathrm{m}} - \alpha)^2\right]^{\frac{1}{2}}, \tag{6.27}$$

where $\alpha = N_{\mathrm{m}}/N_{\mathrm{t}}$ (Eady, 1949; Blumen, 1979). This converges to the classic Eady solution if $\alpha \ll 1$ and $\alpha \ll \mu_{\mathrm{m}}$, which is equivalent to $N_{\mathrm{t}} \gg N_{\mathrm{m}}$ and $k_{\mathrm{h}} \gg f/N_{\mathrm{t}}h$. This shows that large thermocline stratification acts like a rigid bottom, but only for scales that are not too large, as alluded to above. Modes of large horizontal scale penetrate into the thermocline and their dynamics are altered.

The growth rates and phase speeds of this reduced model very nearly match the growth rates and phase speeds of the full model at scales smaller than about 100 km (Fig. 6.5c and 6.5f). This model now captures the long-wave cutoff of the submesoscale instability. At large scales, where $\mu_{\mathrm{m}} \ll \alpha$ and $\mu_{\mathrm{m}} \ll 1$ or equivalently $k_{\mathrm{h}} \ll f/N_{\mathrm{t}}h$ and $k_{\mathrm{h}} \ll f/N_{\mathrm{m}}h$, the dynamics split into modes that are barotropic and baroclinic in the mixed layer. The barotropic mode behaves like a surface edge wave, which has a phase speed $-f\Lambda/N_{\mathrm{t}}k_{\mathrm{h}}$ and does not sense the mixed layer (Fig. 6.5f). The baroclinic mode is baroclinic in the mixed layer and remains shallow for large scales—its critical level is the base of the mixed layer and its phase speed is $-\Lambda h$. The vastly different phase speeds of these two modes prevent phase-locking, so no instability occurs at large scales. This stabilization is analogous to that by the β-effect (Phillips, 1954; Lindzen, 1994; Vallis, 2006). Note that no tilt in the interface is required for this long-wave cutoff (cf. Boccaletti et al., 2007). For the unstable modes, the reduced model with no bottom also captures the deepening of the critical level as the scale gets larger, $-\Lambda h(1 + f/N_{\mathrm{t}}k_{\mathrm{h}}h)/2$, which is due to the increasing penetration of the unstable mode into the thermocline.

The location of the long-wave cutoff in this constant-shear case depends on the ratio $N_{\mathrm{m}}/N_{\mathrm{t}}$. In the more general case $\Lambda_{\mathrm{m}} \neq \Lambda_{\mathrm{t}}$, it also depends on the ratio $\Lambda_{\mathrm{m}}/\Lambda_{\mathrm{t}}$. No long-wave cutoff occurs if $\Lambda_{\mathrm{t}} = 0$, as found by Rivest et al. (1992), who considered the atmospheric case with no shear in the stratosphere. There is also no long-wave cutoff if $N_{\mathrm{m}}/N_{\mathrm{t}} \to 0$, which is the Eady limit. The instability itself requires a reversal of the PV gradient, so the condition for instability is $\Lambda_{\mathrm{m}}/N_{\mathrm{m}}^2 > \Lambda_{\mathrm{t}}/N_{\mathrm{t}}^2$. This condition is typically

satisfied in the ocean, because the thermocline stratification is much larger than the mixed layer stratification and horizontal buoyancy gradients are typically of the same order in the mixed layer as in the thermocline, if not larger.

6.3.4 Summary

We are now in a position to understand all branches in the phase speed diagram of the full model. Branch 'a' is a mode that is baroclinic in the mixed layer and does not penetrate much into the thermocline. It does not sense the bottom. Branch 'b_1' is the unstable branch corresponding to the Eady-like thermocline instability; branch 'b_2' is the conjugate decaying branch. Branch 'c' is a mode that is nearly barotropic in the mixed layer and behaves like a surface edge wave in the thermocline. It does not interact much with the bottom. Branch 'd' is a bottom edge wave that is independent of the surface and interface. Branch 'e_1' is the unstable branch corresponding to the mixed layer instability; branch 'e_2' is the conjugate decaying branch. The instability is significantly modified by the modes' penetration into the thermocline, but the size and growth rate of the most unstable mode still scale with the mixed layer deformation radius and the Eady growth rate. Branches 'f' and 'g' are edge waves propagating on the surface and the interface that do not interact with any of the other edge waves.

6.4   Nonlinear dynamics

We now turn to the nonlinear dynamics that arise when perturbations are amplified by the instabilities and grow to finite amplitude. We solve numerically the full nonlinear equations

$$\frac{\partial \theta}{\partial t} + \mathsf{U}\frac{\partial \theta}{\partial x} + \mathsf{\Gamma}\frac{\partial \psi}{\partial x} + J(\psi, \theta) = -r\nabla^{-2}\theta - \nu\nabla^{2n}\theta, \tag{6.28}$$

where the Jacobian operator is understood to act element-wise:

$$J(\psi, \theta) = (J(\psi_0, \theta_0), J(\psi_1, \theta_1), J(\psi_2, \theta_2))^{\mathsf{T}}. \tag{6.29}$$

These are the evolution equations for perturbations from the prescribed mean zonal flow, which appears in form of the diagonal matrices U and $\mathsf{\Gamma}$. We consider flows that are doubly periodic in the perturbations, so no modification of the prescribed mean can occur. We introduce hypoviscosity with coefficient r, which provides a drag to remove energy from large scales, and hyperviscosity with coefficient ν and order n, which helps ensure numerical stability and absorbs enstrophy at small scales. Hypoviscosity is a convenient but somewhat unphysical choice. We introduce it to halt the inverse cascade and allow for mesoscale equilibration. Hypoviscosity appears in the dynamical equations for the conserved quantities at the surface, interface, and bottom, but it can be thought of as acting

throughout the layers. If applied to buoyancy and momentum, it does not affect PV and PV conservation within the layers remains trivial. Linear drag, which may appear as an obvious choice, cannot prevent an inverse cascade to the domain scale without significantly damping the instabilities.

We integrate these equations on a 500 km × 500 km domain using a fully dealiased pseudo-spectral code with a resolution 512×512. The time derivatives are discretized using a forth-order Runge-Kutta scheme. The hypoviscosity coefficient is $r = 10^{-16}$ m^{-2} s^{-1}; the hyperviscosity is of order $n = 10$ and the coefficient is $\nu = 2.5 \times 10^{46}$ m^{20} s^{-1}. All calculations are initialized with white noise of small amplitude in θ_j.

Before considering the combined effect of mesoscale and mixed layer instabilities, we first consider them separately. We start with the thermocline-only model, which allows only the mesoscale thermocline instability while submesoscale flows follow surface QG dynamics. We subsequently contrast this case with the mixed-layer-only model, which allows only the submesoscale mixed layer instability. We finally consider the full model, in which both instabilities occur.

6.4.1   Thermocline only

We start by studying surface QG turbulence generated by mesoscale eddies, one of the proposed mechanisms to energize submesoscale flows. Surface QG flows cannot themselves extract energy from the mean flow, so they must be forced at the mesoscale. Instead of prescribing external forcing (e.g. Pierrehumbert et al., 1994; Scott, 2006), we use the thermocline-only Eady model to generate mesoscale eddies that in turn generate the submesoscale flows. As we saw in the linear stability analysis, the dynamics of surface buoyancy anomalies in this model decouple from the bottom at scales smaller than the thermocline deformation radius, so flows very nearly follow surface QG dynamics at the submesoscales, i.e. the scales smaller than the thermocline deformation radius. The mesoscale instability is an obviously crude representation of the real mesoscale instability, with no interior PV gradients and the presence of an artificial rigid interface at the base of the thermocline. But the instability does generate mesoscale eddies of roughly the right scale, which is sufficient to drive the submesoscale surface QG flows (cf. Roullet et al., 2012).

Since the dissipative terms are weak in the linear regime, the instability grows until it reaches finite amplitude, when the nonlinear terms become important. Secondary instabilities set in and the flow quickly evolves into a fully turbulent regime. The perturbations grow in scale until they reach a scale where hypoviscosity is significant. Thereby, the flow comes into statistical equilibrium, which is the time period considered in what follows.

A snapshot from the equilibrated state exhibits a patchy surface buoyancy field with strong buoyancy gradients (Fig. 6.7b). The largest eddies are about 200 km in scale. The strongest coherent vortices have a scale of about 50 km. Smaller-scale vortices are present, but weaker the smaller the scale. They result from a roll-up instability that features
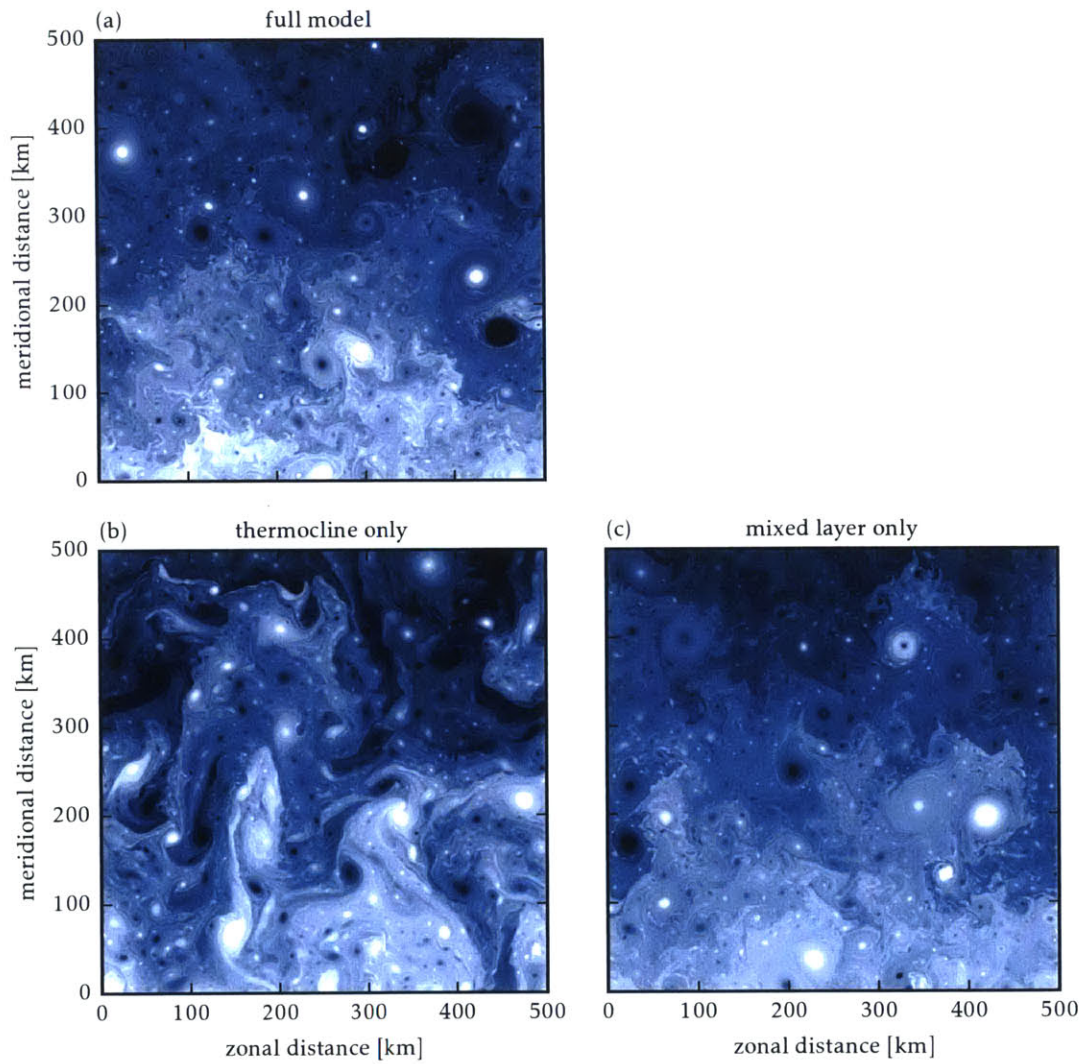
123

Figure 6.7: Snapshots of surface buoyancy (mean plus anomalies) from the equilibrated states of the (a) the full model, (b) the thermocline-only model, and (c) the mixed-layer-only model. The color scale extends from white (more buoyant) through blue to black (less buoyant) and extends between $\pm f \Lambda a$.

prominently in the evolution of the flow (Held et al., 1995).

As typical for turbulent flows, a continuum of scales is energized. This is quantified by the kinetic and potential energy spectra in statistical equilibrium, $\langle K_{k,l} \rangle$ and $\langle P_{k,l} \rangle$, which are defined by

$$K_{k,l} = \frac{1}{2}\left(|\hat{u}|^2 + |\hat{v}|^2\right), \qquad P_{k,l} = \frac{1}{2}\frac{|\hat{b}|^2}{N^2}. \tag{6.30}$$

The angle brackets denote an average in time, performed over the statistical equilibrium, and u and v denote the leading-order zonal and meridional geostrophic velocity components. Isotropic spectra $\langle K_{k_h} \rangle$ and $\langle P_{k_h} \rangle$ are computed by averaging $\langle K_{k,l} \rangle$ and $\langle P_{k,l} \rangle$ over circles of constant k_h in wavenumber space—the statistics are very nearly isotropic.

The surface spectra of both kinetic and potential energy peak at a wavelength of about 200 km and fall off roughly like $\langle K_{k_h} \rangle \sim \langle P_{k_h} \rangle \sim k_h^{-5/3}$ (Fig. 6.8), as predicted by surface QG turbulence theory for scales smaller than the scales at which mesoscale instabilities inject energy into the system (Blumen, 1978). Since small-scale modes decay more rapidly in the vertical than large-scale modes, the spectra are steeper in the interior (e.g. Scott, 2006). At 100 m depth, the mesoscale energy levels are similar to those at the surface, but submesoscales energy levels are much lower.

A useful diagnostic of turbulent dynamics is the spectral energy budget (e.g. Larichev and Held, 1995; Roullet et al., 2012). While the dynamics are completely determined by the advection of conserved quantities at the surface and bottom, we first consider the energy budget over the entire depth range. We will take into account the reduced nature of these models below, where we present a vertically integrated energy budget for the mixed-layer-only case.

The equations for the spectral perturbation potential and kinetic energies are

$$\frac{\partial P_{k,l}}{\partial t} = \mathrm{Re}\left[\frac{f\Lambda}{N^2}\hat{v}^*\hat{b} - \hat{w}^*\hat{b} - \frac{1}{N^2}\hat{b}^*\hat{J}(\psi,b)\right] - \left(rk_h^{-2} + \nu k_h^{2n}\right)P_{k,l} \tag{6.31}$$

$$\frac{\partial K_{k,l}}{\partial t} = \mathrm{Re}\left[-f\frac{\partial}{\partial z}\left(\hat{w}^*\hat{\psi}\right) + \hat{w}^*\hat{b} + \hat{\psi}^*\hat{J}(\psi,\nabla^2\psi)\right] - \left(rk_h^{-2} + \nu k_h^{2n}\right)K_{k,l} \tag{6.32}$$

where the asterisks denote complex conjugates and Re denotes taking the real part. The Fourier transforms in the square bracket are all understood to be evaluated at the wavenumbers k and l. The first term on the right-hand side of the potential energy equation represents the extraction of potential energy from the mean flow. The second term represents the conversion from potential to kinetic energy. This term appears as a source term in the kinetic energy budget. The third term in the potential energy budget represents spectral transfer by triadic interactions. The sum of this term over all wavenumbers vanishes. An equivalent spectral transfer term appears in the kinetic energy budget (third term). Kinetic energy can also be distributed vertically by pressure fluxes, represented by the first term in the kinetic energy budget. The vertical integral of this term vanishes.
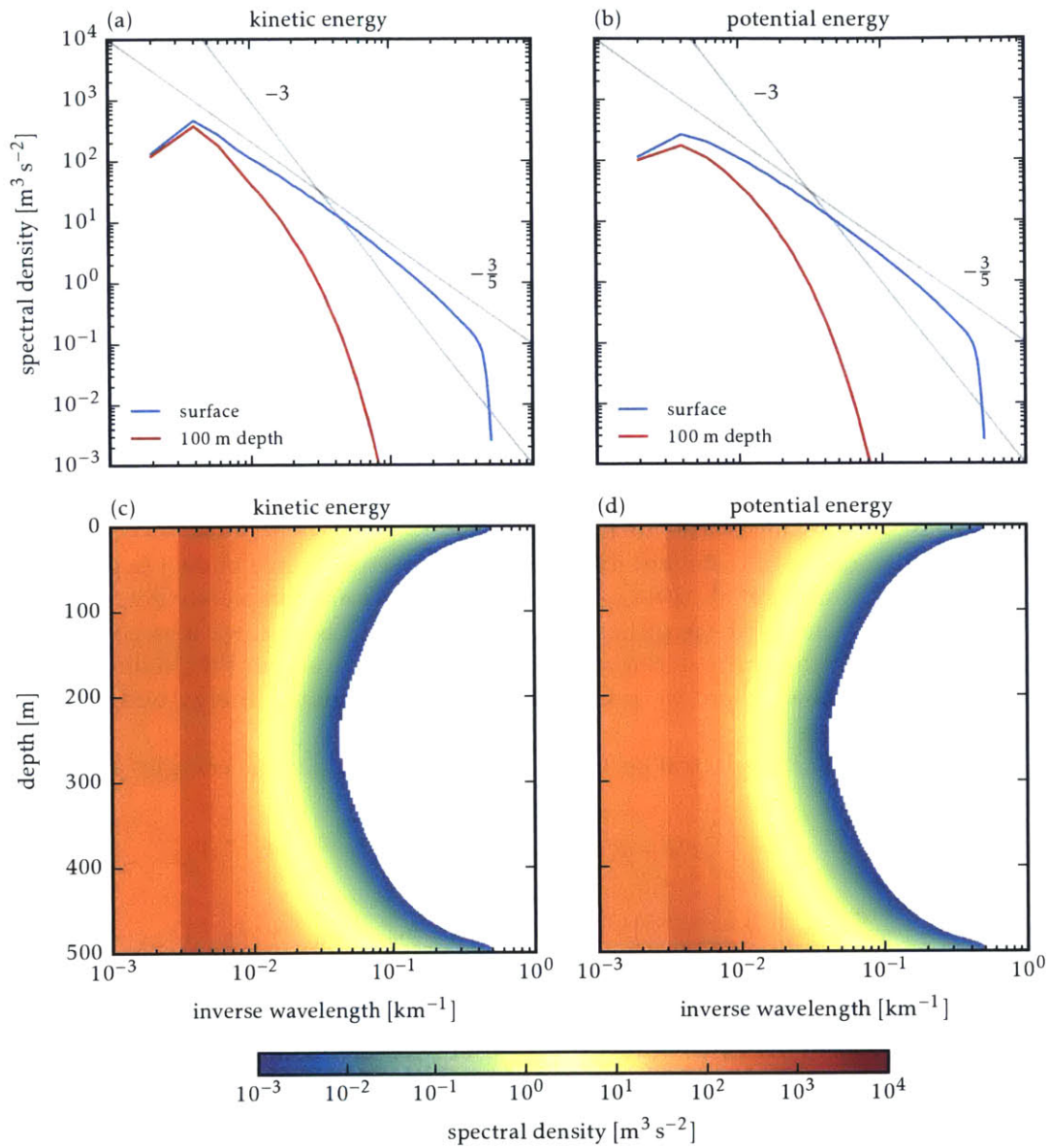
Figure 6.8: Wavenumber spectra of kinetic and potential energy from the thermocline-only simulation. (a) Kinetic and (b) potential energy spectra at the surface and 100 m depth, spectral density of (c) kinetic and (d) potential energy in the wavenumber–depth plane. In panels (c) and (d), no values below 10^{-3} m^3 s^{-2} are shown.

126

The viscosity terms act as sinks for both potential and kinetic energy—hypoviscosity acting at large scales, hyperviscosity at small scales. We present these budgets averaged azimuthally in wavenumber space and over time.

The extraction of potential energy from the mean is dominated by the largest, most energetic eddies (Fig. 6.9a). The extraction is independent of depth, because $q = 0$ and therefore

$$0 = \operatorname{Re} \hat{v}^* \hat{q} = \operatorname{Re} \frac{\partial}{\partial z}\left(\frac{f}{N^2} \hat{v}^* \hat{b}\right),$$

(6.33)

where it was used that the term involving advection of relative vorticity vanishes. Potential energy is transferred downscale by triadic interactions and deposited near the deformation radius as well as in wedges near the surface and the bottom that reach to much smaller scales (Fig. 6.9b). Where potential energy is deposited by scale interactions, it is converted into kinetic energy (Fig. 6.9c). Near the mesoscale deformation radius, this conversion is due to the mesoscale instability that produces vertical buoyancy fluxes. In the wedges near the surface and bottom, the conversion is due to frontogenesis and secondary instabilities present in the surface QG cascades, which occur independently at the surface and the bottom (Roullet et al., 2012). The kinetic energy thus created is transferred back to large scales (Fig. 6.9d). The bulk of the energy is dissipated through hypoviscosity at the scales of the largest, most energetic eddies (Fig. 6.9e). The energy dissipation through hyperviscosity is small, which reflects the fundamental property of geostrophic turbulence that energy is trapped at large scales and viscous energy dissipation vanishes in the limit of infinite resolution and zero (hyper-)viscosity (Kraichnan, 1967; Charney, 1971).

6.4.2 Mixed layer only

We now turn our attention to the nonlinear dynamics of the submesoscale mixed-layer instability and compare its turbulent dynamics to the surface QG turbulence of the thermocline-only case. We study the case with an infinitely deep thermocline, which allows an accurate representation of the submesoscale instability, while eliminating the thermocline instability (Fig. 6.5c,f).

The submesoscale instability grows to finite amplitude and the flow becomes turbulent. There is a turbulent spin-up phase, in which the eddies, which are initially of the size of the instability, grow larger until they reach a statistical equilibrium with hypoviscosity. The flow is host of numerous coherent vorticies embedded in a filamentary sea with strong buoyancy gradients (Fig. 6.7c). A snapshot of surface buoyancy appears quite different from the thermocline-only case, but this visual difference is due mostly to the smaller size of the most energetic eddies.

The energy spectra reflect the nearly frontal structure at the surface (Fig. 6.10). The kinetic energy spectra fall off slightly more steeply than $\langle K_{k_h} \rangle \sim k_h^{-5/3}$ in the scale range of the linear instability and like $\langle K_{k_h} \rangle \sim k_h^{-5/3}$ at scales smaller than the linear short-wave

Figure 6.9: Spectral energy budget for the thermocline-only simulation. The terms are (a) potential energy extraction from the mean, (b) spectral potential energy flux divergence, (c) potential to kinetic energy conversion, (d) kinetic energy flux divergence, including spectral flux and pressure flux, (e) hypoviscosity on both kinetic and potential energy, and (f) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.
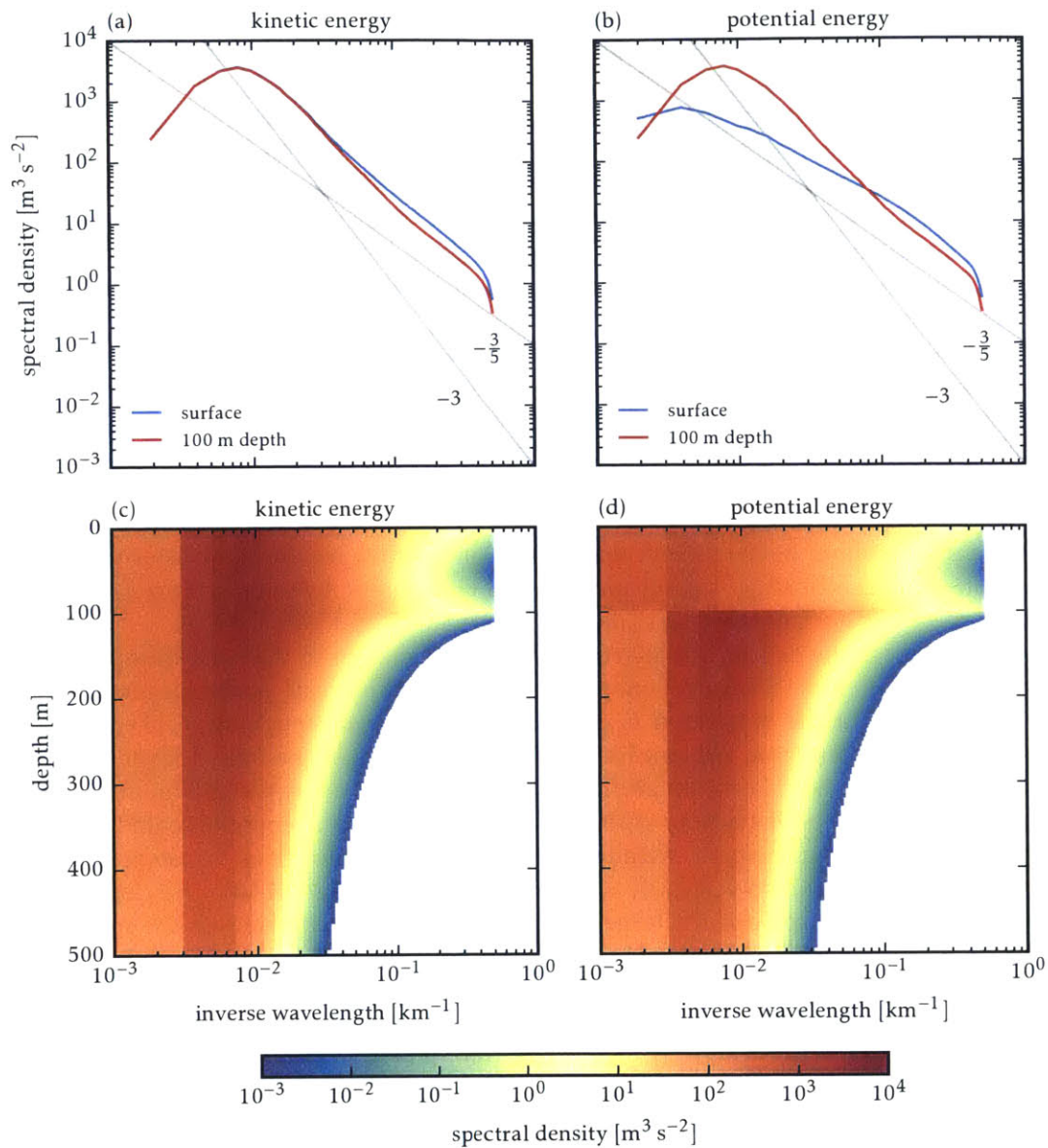
Figure 6.10: Wavenumber spectra of kinetic and potential energy from the mixed-layer-only simulation. (a) Kinetic and (b) potential energy spectra at the surface and 100 m depth (just below mixed layer base), spectral density of (c) kinetic and (d) potential energy in the wavenumber–depth plane. In panels (c) and (d), no values below 10^{-3} m^3 s^{-2} are shown.

cutoff, both at the surface and at the base of the mixed layer at 100 m depth (Fig. 6.10a). The mixed layer instabilities energize the entire depth of the mixed layer. This is in sharp contrast to the thermocline-only simulation, in which surface QG turbulence energizes a thin wedge close to the surface only.

The equilibrated flow in the mixed-layer-only case is much more energetic than in the thermocline-only case and more energetic than is realistic. While the equilibration by hypoviscosity is unrealistic, we will see that the enhanced energy levels are due to more efficient extraction of mean potential energy in the weakly stratified mixed layer, which is a dynamical property of the system that does not depend on how the flow is equilibrated. We will discuss possible reasons for these unrealistically high energy levels in Section 6.5.

Below the base of the mixed layer, the potential energy spectra are the same as the kinetic energy spectra (Fig. 6.10c). In the mixed layer, the potential energy spectra are significantly flatter than the kinetic energy spectra. This is in contrast to observations that show rough equipartition between kinetic and potential energy (Chapters 2 and 5). We currently do not understand the reason for this difference.

The vertical structure of energy shows that the mixed layer instabilities also energize the thermocline below (Fig. 6.10b,d). At the instability scale, the flow does not reach much into the thermocline. But as the horizontal scale of the flow increases, so does the vertical scale. The flow exhibits the familiar property of geostrophic turbulence and barotropizes as it increases its horizontal scale (Charney, 1971; Smith and Vallis, 2001).

The energy transfer into the thermocline is best examined through the spectral energy budget (Fig. 6.11). Potential energy is extracted at the scale of the largest, most energetic eddies, but the extraction is confined to the mixed layer (Fig. 6.11a). Potential energy is transferred from the extraction scale to the scale of the mixed layer instability (Fig. 6.11b). The mixed layer instability converts potential energy into kinetic energy in the mixed layer, at the instability scale (Fig. 6.11c). The kinetic energy created by the instability undergoes an inverse cascade, in which energy is not only transferred to large horizontal scales, but also vertically into the thermocline (Fig. 6.11d). The deposition of kinetic energy at the scale of the largest eddies is well distributed across the mixed layer and upper thermocline. The vertical distribution of the energy sink through hypoviscosity confirms that the flow extends below the mixed layer at the scale of the largest eddies, where hypoviscosity acts (Fig. 6.11e). Hyperviscosity acts only at the smallest resolved scales (Fig. 6.11f). While small, it does affect the other terms in the budget. We do not discuss its effects any further, because they are expected to disappear if the resolution is increased and the hyperviscosity coefficient decreased.

These energy pathways are reminiscent of the phenomenology of two-layer baroclinic turbulence. The turbulent dynamics of a two-layer system can be understood in terms of a dual cascade (Rhines, 1977; Salmon, 1978). Baroclinic energy is extracted from the mean at the scale of the largest, most energetic eddies. The barotropic flow dominates at these scales and transfers the baroclinic energy downscale. The baroclinic mode behaves like a passive tracer at these scales. Around the deformation radius, the instability converts
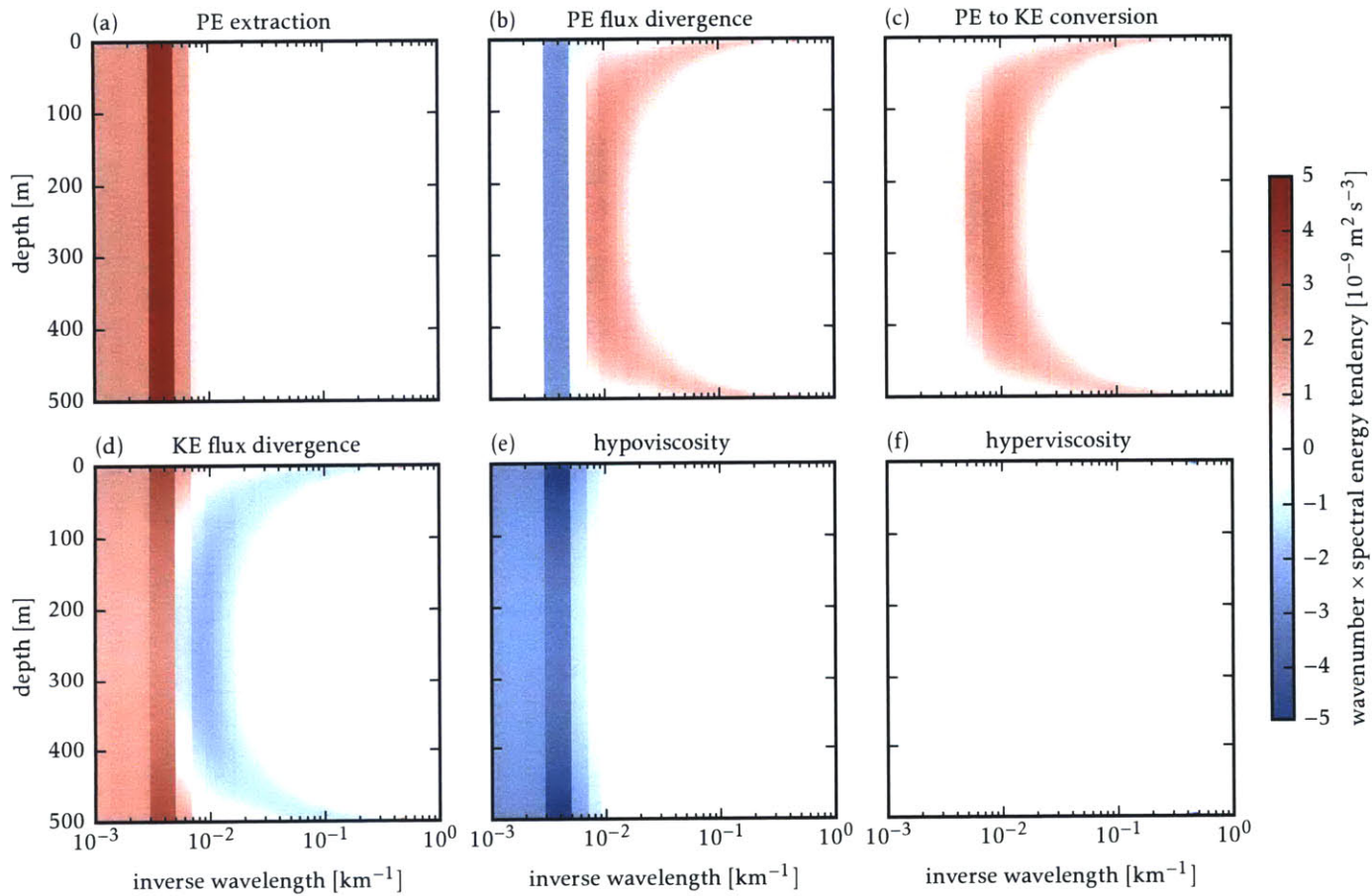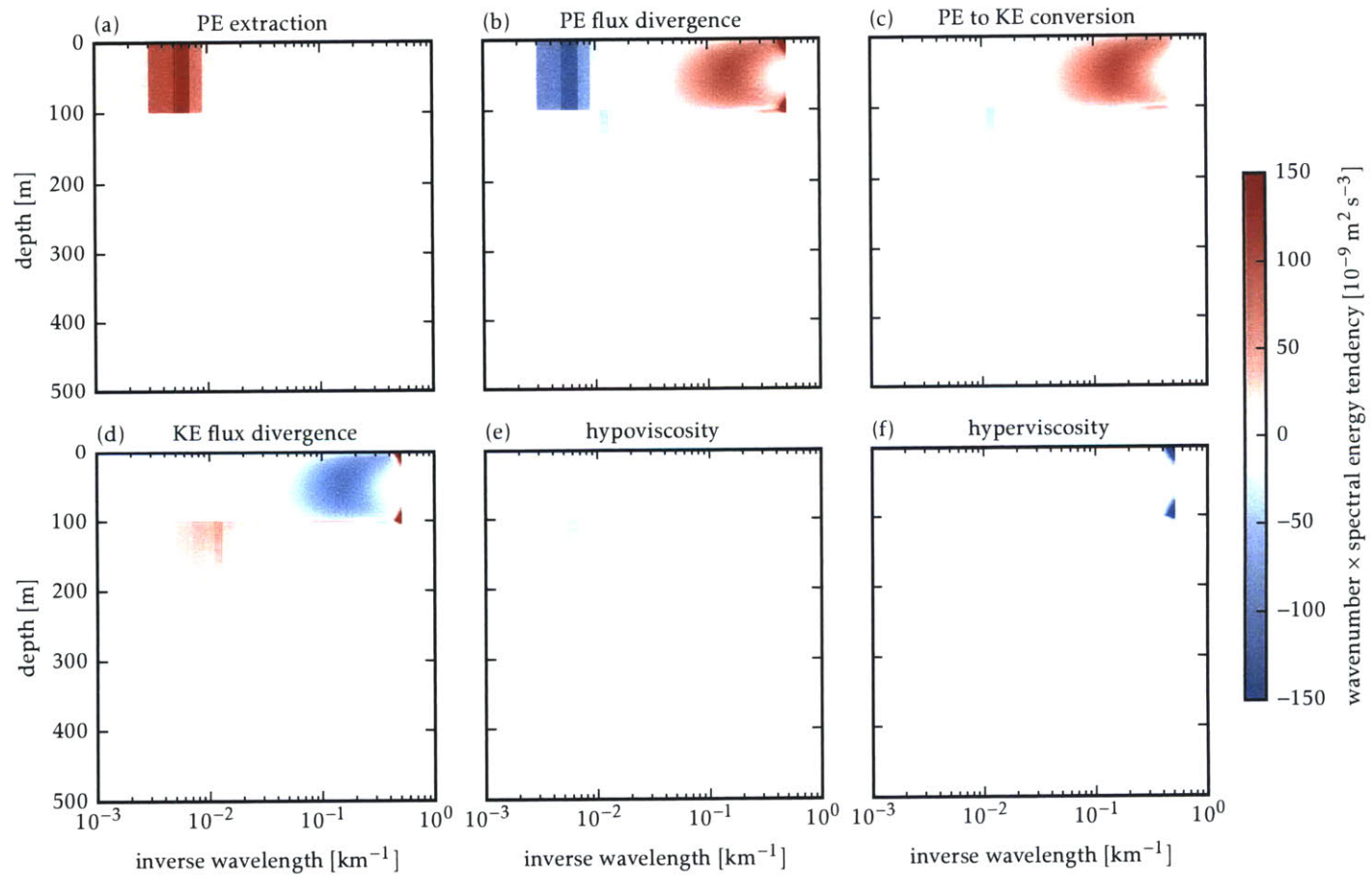
Figure 6.11: Spectral energy budget for the mixed-layer-only simulation. The terms are (a) potential energy extraction from the mean, (b) spectral potential energy flux divergence, (c) potential to kinetic energy conversion, (d) kinetic energy flux divergence, including spectral flux and pressure flux, (e) hypoviscosity on both kinetic and potential energy, and (f) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

baroclinic energy into barotropic energy. The barotropic energy then enters an inverse cascade, which is arrested at some scale by drag or hypoviscosity. The forward cascade of baroclinic energy is compensated by the inverse cascade of barotropic energy, such that no spectral transfer of total energy occurs. This is consistent with the phenomenology that all sources and sinks of total energy occur at the scale of the largest, most energetic eddies—no energy is dissipated at small scales.

Can the turbulent dynamics induced by mixed layer instabilities be understood in similar terms? To pursue the analogy, we must first introduce a modal decomposition of the eddy energy. In our system, the vertically integrated total energy can be written entirely in terms of the quantities at the surface and interface:

$$E_{k,l} = -\frac{1}{2}\hat{\psi}^{\dagger}\hat{\theta} = -\frac{1}{2}\hat{\psi}^{\dagger}\mathbf{L}\hat{\psi}, \tag{6.34}$$

where the conjugate transpose is denoted with a dagger. Since $\mathbf{L}$ is real and symmetric, it can be diagonalized through a unitary matrix $\mathbf{S}$,

$$\mathbf{L} = \mathbf{S}^{\dagger}\mathbf{DS}, \tag{6.35}$$

where $\mathbf{D}$ is diagonal and consists of the real eigenvalues of $\mathbf{L}$, $\mathbf{D}_{jj} = \lambda_j$. The energy can now be written as

$$E_{k,l} = -\frac{1}{2}(\mathbf{S}\hat{\psi})^{\dagger}\mathbf{D}(\mathbf{S}\hat{\psi}) = -\frac{1}{2}\sum_j \lambda_j|(\mathbf{S}\hat{\psi})_j|^2. \tag{6.36}$$

This defines the modes $(\mathbf{S}\hat{\psi})_j$ that are orthogonal with respect to the energy norm, i.e. the energy can be partitioned into contributions $E_{k,l}^j$ from these modes. The structure of the modes depends on wavenumber, because $\mathbf{L}$ and therefore $\mathbf{S}$ does.

For the mixed-layer-only case, with $\mathbf{L}$ given by (6.24), the eigenvalues of $\mathbf{L}$ are

$$\lambda_{0,1} = fk_h\left(\frac{\coth\mu_m}{N_m} + \frac{1}{2N_t} \pm \sqrt{\frac{\mathrm{csch}^2\mu_m}{N_m^2} + \frac{1}{4N_t^2}}\right) \tag{6.37}$$

and we obtain the eigenvectors as the columns of $\mathbf{S}$,

$$\mathbf{S} = \begin{pmatrix} \dfrac{1}{\left[1+\left(\cosh\mu_m+\frac{N_m\lambda_0}{fk_h}\sinh\mu_m\right)^2\right]^{\frac{1}{2}}} & \dfrac{1}{\left[1+\left(\cosh\mu_m+\frac{N_m\lambda_1}{fk_h}\sinh\mu_m\right)^2\right]^{\frac{1}{2}}} \\[4mm] \dfrac{\cosh\mu_m+\frac{N_m\lambda_0}{fk_h}}{\left[1+\left(\cosh\mu_m+\frac{N_m\lambda_0}{fk_h}\sinh\mu_m\right)^2\right]^{\frac{1}{2}}} & \dfrac{\cosh\mu_m+\frac{N_m\lambda_1}{fk_h}}{\left[1+\left(\cosh\mu_m+\frac{N_m\lambda_1}{fk_h}\sinh\mu_m\right)^2\right]^{\frac{1}{2}}} \end{pmatrix}. \tag{6.38}$$

For large scales, $k_h \ll f/N_t h$ or $\mu_m \ll N_m/N_t$, this reduces to

$$\lambda_0 = -\frac{fk_h}{2N_t}, \qquad \lambda_1 = -\frac{2f^2}{N_m^2 h}. \tag{6.39}$$

and simply

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{6.40}$$

Equation (6.40) indicates that at large scales the first mode is barotropic in the mixed layer. It behaves like a surface QG mode penetrating into the thermocline. The streamfunction is proportional to k_h times the conserved quantity (Held et al., 1995),

$$(S\hat{\psi})_0 = -\frac{fk_h}{2N_t}(S\hat{\theta})_0. \tag{6.41}$$

The second mode at large scales is baroclinic in the mixed layer. The relation between the streamfunction and the conserved quantity is

$$(S\hat{\psi})_1 = -\frac{2f^2}{N_m^2 h}(S\hat{\theta})_1, \tag{6.42}$$

which is independent of k_h, as expected for a baroclinic mode. These are the same modes as those found in the linear stability analysis for large scales (Fig. 6.5f).

This description of the orthogonal modes as barotropic and baroclinic mixed layer modes only applies at large scales. At smaller scales, the modes have a more complicated vertical structure (Fig. 6.12). At scales smaller than the mixed layer deformation radius, they morph into modes that are decoupled and localized in the vertical at the surface and at the interface. But for the cascade dynamics to be discussed, the mode structure at large scales is what is most important.

We can now consider the energy budget of these modes. We start from the vertically integrated spectral energy budget, written in terms of the conserved quantities and corresponding streamfunctions at the surface and interface:

$$\frac{\partial E_{k,l}}{\partial t} = -\mathrm{Re}\,\hat{\psi}^\dagger \frac{\partial \hat{\theta}}{\partial t}. \tag{6.43}$$

Using the unitary matrix S, we can rewrite this as

$$\frac{\partial E_{k,l}}{\partial t} = -\mathrm{Re}\,(S\hat{\psi})^\dagger \frac{\partial}{\partial t}(S\hat{\theta}) = -\mathrm{Re}\sum_j (S\hat{\psi})_j^* \frac{\partial}{\partial t}(S\hat{\theta})_j \tag{6.44}$$

and thus split the energy budget into its modal components:

$$\frac{\partial E_{k,l}^j}{\partial t} = -\mathrm{Re}\,(S\hat{\psi})_j^* \frac{\partial}{\partial t}(S\hat{\theta})_j. \tag{6.45}$$

The terms on the right-hand side of this budget can be obtained by substituting in the spectral form of the evolution equation (6.28). To separate out the advective interactions
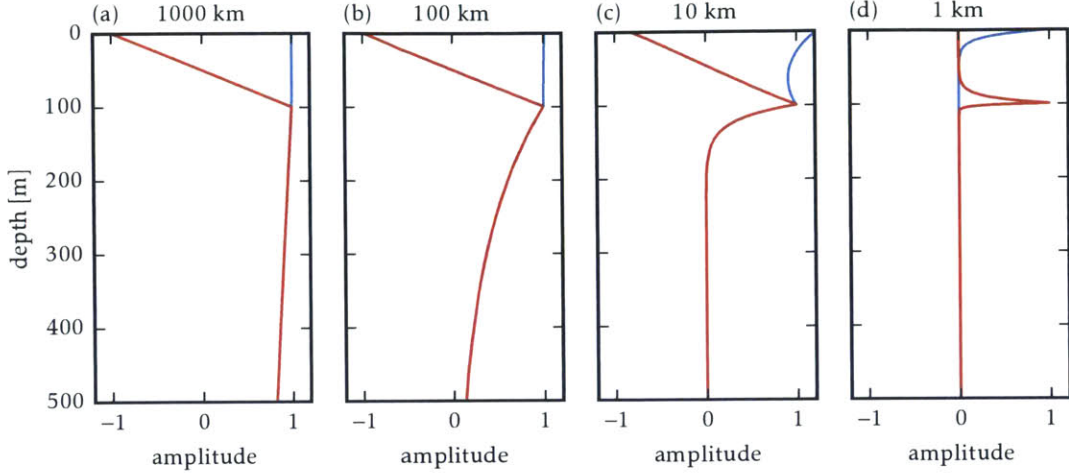
133

Figure 6.12: Vertical structure of the streamfunction corresponding to orthogonal modes in the mixed-layer-only case for different wavenumbers $k_h = 2\pi/\lambda$, with the wavelength λ given in the panel titles. For panels (a), (b), and (c), the modes are normalized to unity at the interface at 100 m depth; for panel (d), the modes are normalized to have a maximum value of unity. Mode 0 is shown in blue, mode 1 in green. In panels (a), (b), and (c), the two modes coincide below the interface.

of the modes with themselves and with each other, we further expand the nonlinear terms in (6.28), using the distributive property of the Jacobian operator, into

$$J(\psi,\theta) = J(\psi^0,\theta^0) + J(\psi^0,\theta^1) + J(\psi^1,\theta^0) + J(\psi^1,\theta^1). \tag{6.46}$$

Here we split the vectors holding the streamfunction and the conserved quantities at the surface and interface into their modal components,

$$\hat{\psi}^j = S^\dagger P_j S\hat{\psi}, \qquad \hat{\theta}^j = S^\dagger P_j S\hat{\theta}, \qquad \psi = \sum_j \psi^j, \qquad \theta = \sum_j \theta^j, \tag{6.47}$$

where P_j are the projections onto the respective modes,

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{6.48}$$

The first term in (6.46) represents the advection of the barotropic mode by the barotropic mode, to use the naming convention introduced above. The second term represents the advection of the baroclinic mode by the barotropic mode, and so on.

In terms of the orthogonal modes, the energy budget is very similar to that of a baroclinic two-layer system (Fig. 6.13, cf. Larichev and Held, 1995). The extraction of potential
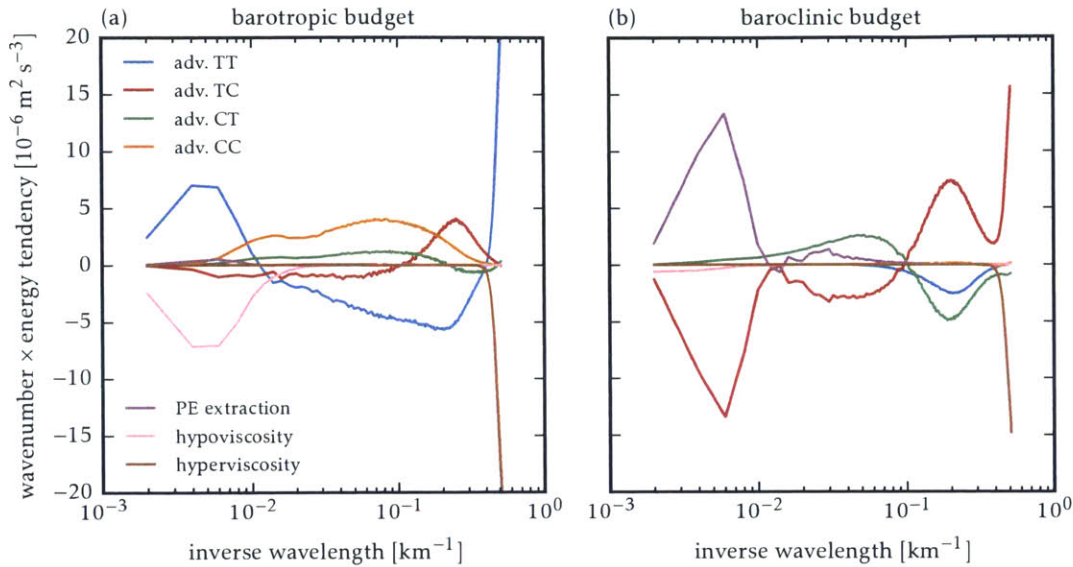
134

Figure 6.13: Modal energy budget for the mixed-layer-only case. The advective terms correspond to the contributions from the four terms in (6.46). The energy tendencies are multiplied by wavenumber to compensate for logarithmic shrinking.

energy from the mean flow is concentrated at the scale of the largest, most energetic eddies and creates mostly baroclinic energy (Fig. 6.13b). The dominant sink is by hypodiffusion, which also acts on the largest, most energetic eddies. Barotropic energy dominates at these scales, so hypodiffusion takes out mostly barotropic energy (Fig. 6.13a). The transfer of energy from the baroclinic mode to the barotropic mode occurs through a dual cascade in the submesoscale range. The baroclinic and barotropic energy components are cascaded in opposite directions so as to yield a vanishing spectral transfer of total energy. Baroclinic energy is transferred down to the instability scale, achieved by the advection of the baroclinic mode by the barotropic mode (Fig. 6.13b). The energy deposited around the instability scale is transferred to the barotropic mode by interactions between the two modes, which represents baroclinic instability (Fig. 6.13b). This energy enters the barotropic budget rather less localized in wavenumber space (Fig. 6.13a). An upscale spectral transfer of barotropic energy closes the budget, taking energy from the instability scale to the scale of the largest, most energetic eddies, where hypodiffusion acts (Fig. 6.13a). Energy loss by hyperdiffusion again enters the budget, but is neglected in this discussion, because it is an artifact of finite resolution.

This model thus exhibits a dual cascade analogous to the classic two-layer system. Baroclinic energy is transferred downscale through advection by the barotropic mode, baroclinic instability converts baroclinic into barotropic energy, and barotropic energy is

135

transferred back upscale in an inverse cascade. The difference is that the barotropic mode at large scales behaves like a surface QG mode, instead of a truly barotropic or two-dimensional mode. The inverse cascade is therefore expected to yield a $\langle E_{k_h} \rangle \sim k_h^{-1}$ surface energy spectrum (Blumen, 1978), which we find to emerge if the inertial range is wide enough (not shown). More importantly, the surface-QG-like behavior implies that in the inverse cascade, energy is transferred to successively larger vertical scales. This provides a pathway for mixed layer instabilities to energize the thermocline below.

6.4.3 Full model

We now consider the case with both mesoscale and submesoscale instabilities present. This full model allows us to address how mesoscale thermocline instabilities modify the energy cycle induced by submesoscale mixed layer instabilities. Furthermore, we pursue a one-to-one comparison between surface-QG dynamics and the dynamics modified by mixed layer instabilities, with a focus on vertical energy distribution and vertical velocities.

The linear growth rate of the mixed layer instability is much larger than that of the thermocline instability (Fig. 6.5a), so during the initial transient of the nonlinear simulations the mixed layer instability grows to finite amplitude first. The evolution in the mixed layer is very similar to that of the mixed-layer-only case: the eddies grow in size until they come into statistical equilibrium with hypoviscosity.

The equilibrated state of the full model is also very similar to that of the mixed-layer-only case in the mixed layer and upper thermocline (Fig. 6.14). The energy levels and spectra at the surface and the base of the mixed layer are very similar. Near the bottom, a wedge in wavenumber–depth space is energized in the full model, just like in the thermocline-only case. This is due to surface QG turbulence at the bottom level.

The energy budget is similar to the mixed-layer-only case (Fig. 6.15). The main energy pathway is again extraction of potential energy in the mixed layer, transfer to the mixed layer instability scale, conversion to kinetic energy, transfer back to large scales and into the thermocline, and dissipation by hypoviscosity. There is additional energy extraction in the thermocline, but that is weak compared to the extraction in the mixed layer. The dominant dynamics are therefore those described for the mixed-layer-only case. Interaction with the bottom level is possible, but of secondary importance in the parameter regime of relevance.

A different picture emerges when the horizontal buoyancy gradient and the associated geostrophic shear in the mixed layer is (unrealistically) reduced. We choose the mixed layer shear such that the growth rates of the two instabilities are comparable, which from Eady scaling occurs if $\Lambda_m/N_m = \Lambda_t/N_t$ and we set $\Lambda_m = 2.5 \times 10^{-5}$ s^{-1}. The horizontal scales of the instabilities and the overall structure of the dispersion curves are the same as those in Fig. 6.5a,d.

136
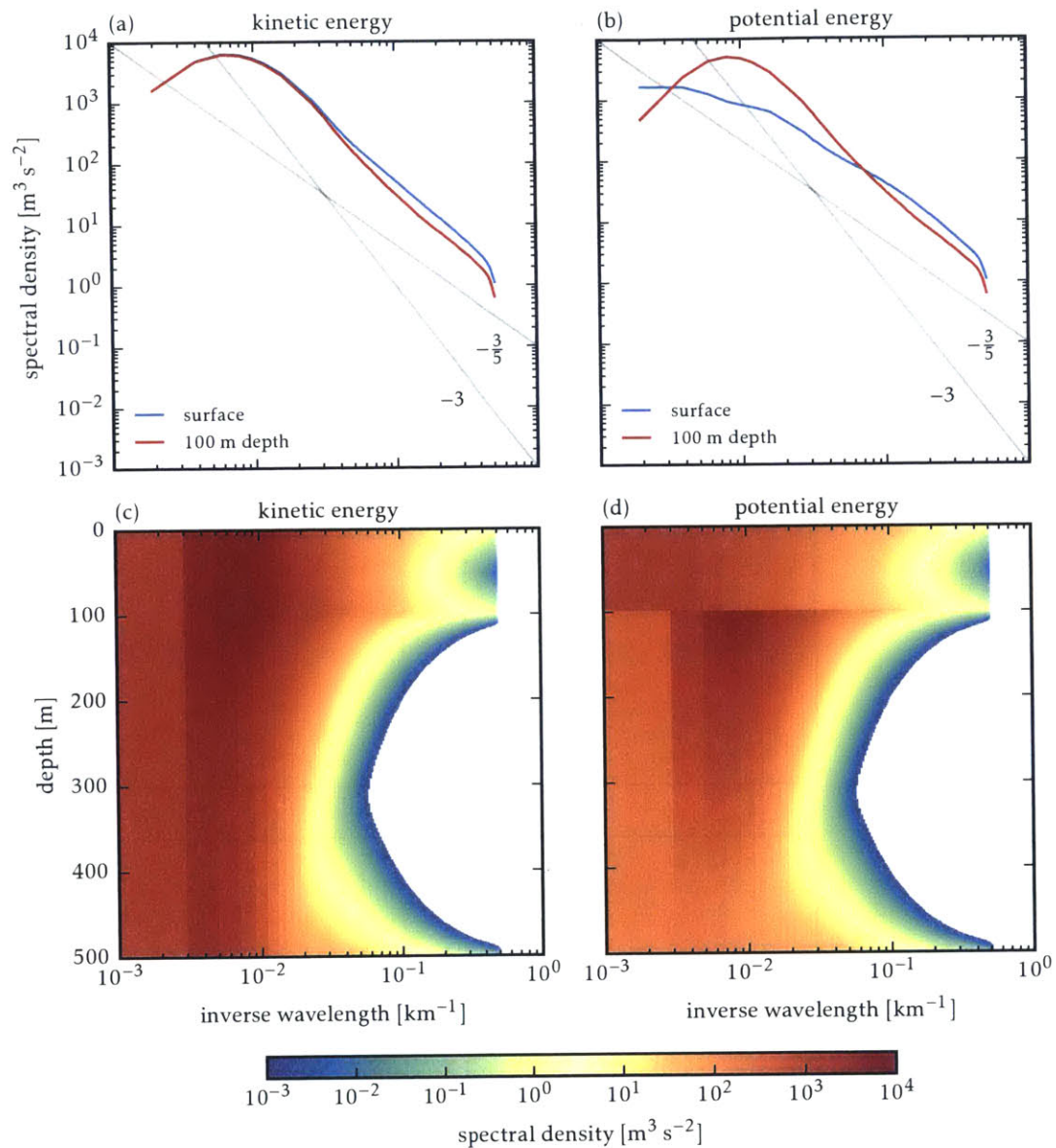
Figure 6.14: Wavenumber spectra of kinetic and potential energy from the full model simulation. (a) Kinetic and (b) potential energy spectra at the surface and 100 m depth (just below mixed layer base), spectral density of (c) kinetic and (d) potential energy in the wavenumber–depth plane. In panels (c) and (d), no values below 10^{-3} m^3 s^{-2} are shown.
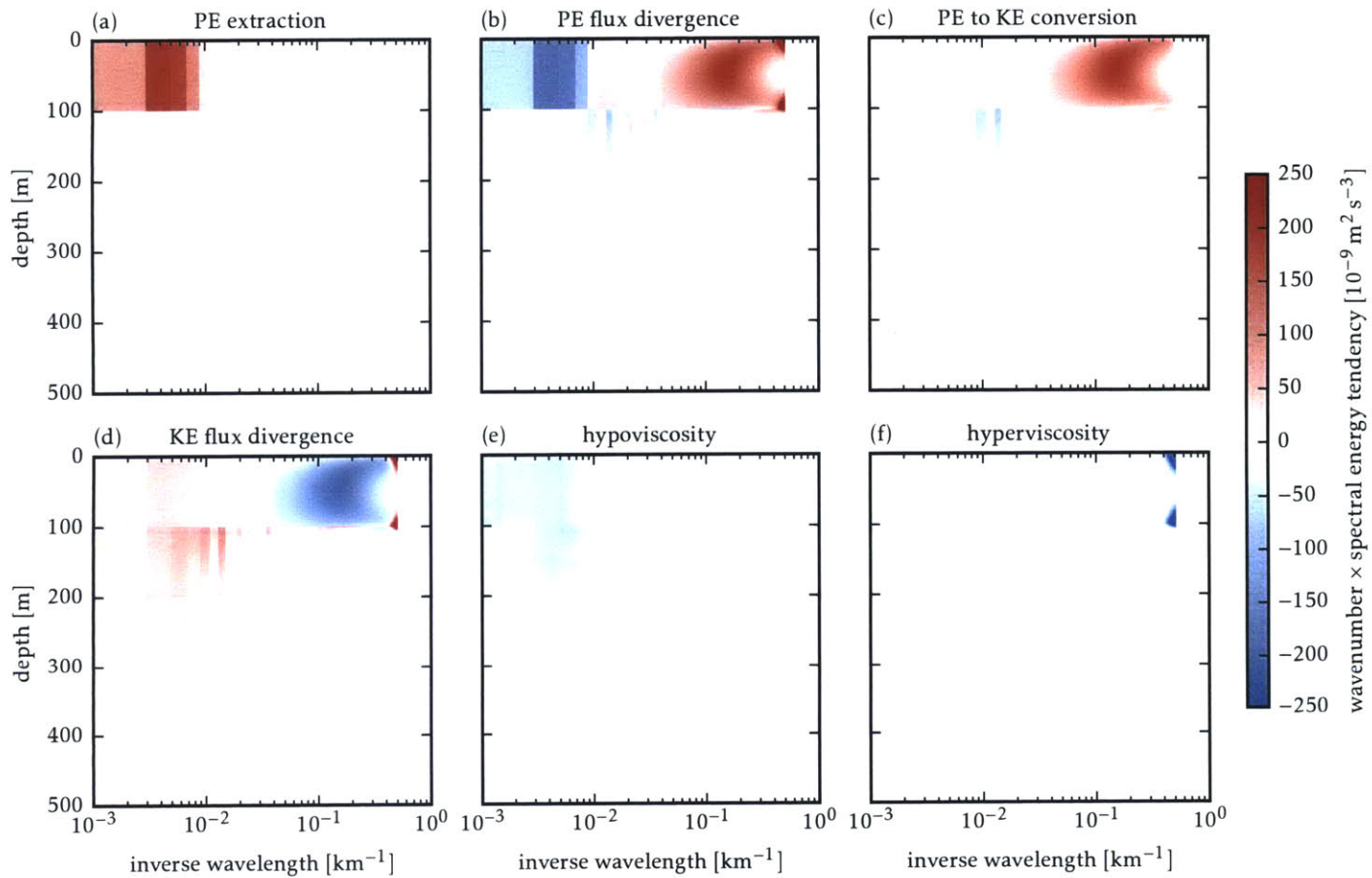
Figure 6.15: Spectral energy budget for the full model simulation. The terms are (a) potential energy extraction from the mean, (b) spectral potential energy flux divergence, (c) potential to kinetic energy conversion, (d) kinetic energy flux divergence, including spectral flux and pressure flux, (e) hypoviscosity on both kinetic and potential energy, and (f) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.
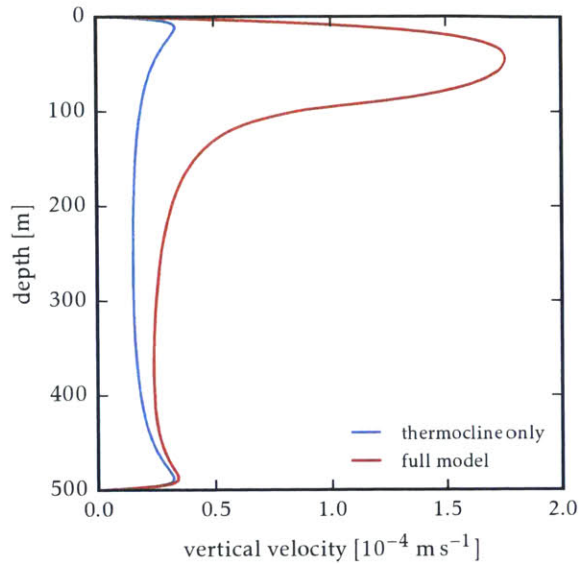
Figure 6.16: Profiles of root mean square vertical velocity for the thermocline-only simulation and the full model simulation with reduced mixed layer shear.

This system with reduced mixed layer shear equilibrates to much lower energy levels than the constant shear case. The energy levels are comparable to the thermocline-only case and thus allow a one-to-one comparison of the dynamics with and without a mixed layer. The vertical structure of energy in this case with a mixed layer is quite different from the thermocline-only case, because the energy pathway enabled by the mixed layer instability is still present—mixed layer instabilities are—on average—not suppressed by the mesoscale strain field (cf. Bishop, 1993a,b; Spall, 1997; McWilliams et al., 2009). The mixed layer instability, while not significantly increasing the mesoscale energy levels, does energize the mixed layer at the submesoscales.

This difference between mixed layer dynamics and surface QG turbulence is also reflected in vertical velocities that are produced by the instabilities (Fig. 6.16). While the available potential energies are the same and the resulting surface energy levels comparable between this reduced mixed layer shear case and the thermocline-only case, there are much larger vertical velocities in the presence of a mixed layer. These enhanced vertical velocities extend significantly below the base of the mixed layer. The largest vertical velocities are located near fronts in the filamentary sea (Fig. 6.17). Coherent vortices, while associated with the large buoyancy gradients, induce relatively weak vertical motion. The large vertical velocities appear instead to be associated with the filamentary structure generated by mixed layer instabilities.

The enhancement of vertical velocities in the presence of mixed layer instabilities can

139

be understood by considering the omega equation (Hoskins et al., 1978):

$$N^2 \nabla^2 w + f^2 \frac{\partial^2 w}{\partial z^2} = -2\nabla \cdot \boldsymbol{Q},$$ (6.49)

where

$$\boldsymbol{Q} = \left(\frac{\partial \boldsymbol{u}}{\partial x} \cdot \nabla b, \frac{\partial \boldsymbol{u}}{\partial y} \cdot \nabla b \right).$$ (6.50)

The vertical velocities can be written as a convolution of the Green's function of (6.49) with the forcing term on the right of (6.49). While the forcing term is not changing much between the cases with and without the mixed layer, the Green's functions do. The reduced stratification in the mixed layer enhances the response to the forcing term there (Thomas et al., 2008). The enhancement of vertical velocities in our simulation with mixed layer can thus be attributed to the reduced stratification. Such a reduction of stratification is always associated with a mixed layer instability, however, so the enhanced vertical velocities are inextricably linked to mixed layer instabilities.

The forcing term on the right of (6.49) is of the same order in our two cases, because by design the mean states have the same available potential energy and similar submesoscale energy levels are produced. It should be kept in mind, however, that the forcing term does likely increase in the real ocean when mixed layers become deep and mixed layer instabilities energize the submesoscale range. The more energetic submesoscale turbulence in the wintertime mixed layer is expected to be associated with stronger submesoscale strains than are present in summer (Chapter 5). In the wintertime mixed layer, vertical velocities are then enhanced by both a decreased stratification and an increased forcing term on the right of (6.49). Mixed layer instabilities in the real ocean thus most likely drive an even more dramatic increase in vertical velocities in winter than is present in our simulations with and without mixed layer.

The root mean square vertical velocities (Fig. 6.16) are similar in structure to those found in primitive equation models (Capet et al., 2008a). A careful comparison is necessary to establish whether the QG dynamics described here reproduce the magnitude of the vertical velocities or whether non-QG effects significantly enhance or reduce them (cf. Mahadevan and Tandon, 2006). Such a comparison is beyond the scope of this chapter.

The interface between mixed layer and thermocline is located at $z = -h + \eta$, where the interface displacement η is determined by requiring the total buoyancy field to be continuous at the interface,

$$N_{\mathrm{m}}^2 \eta + b(-h^+) = N_{\mathrm{t}}^2 \eta + b(-h^-),$$ (6.51)

where the total buoyancy was linearized around $z = -h$, consistent with QG scaling. The interface is material in the sense that

$$\frac{\partial \eta}{\partial t} + J(\psi, \eta) = w.$$ (6.52)
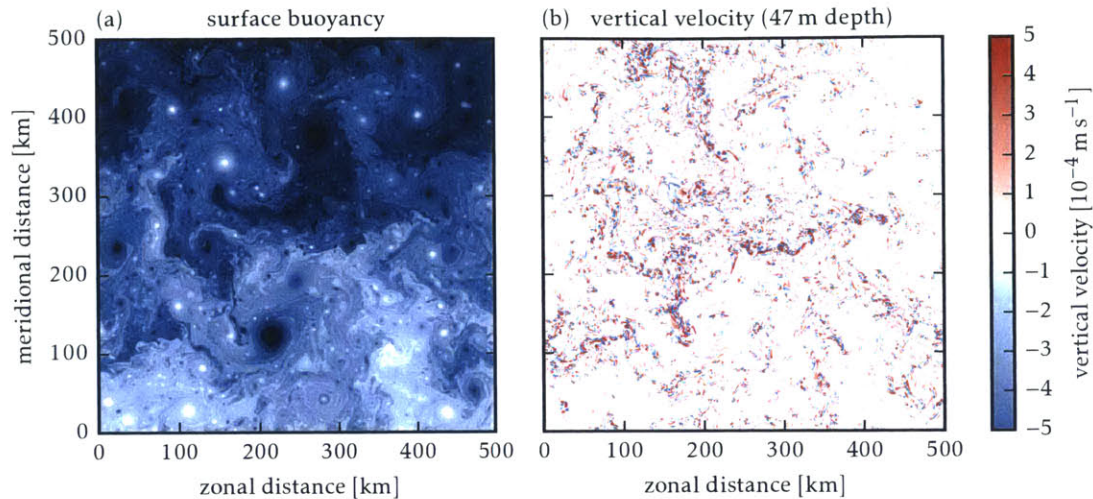
Figure 6.17: Concurrent surface buoyancy and vertical velocity snapshots from the full model simulation with reduced mixed layer shear: (a) surface buoyancy equivalent to Fig. 6.7 and (b) vertical velocity at 47 m, the depth of the maximum root mean square vertical velocity.

To leading order, there is therefore no exchange of fluid between mixed layer and thermocline. If non-QG effects are taken into account, however, thermocline waters can be folded into the mixed layer (Garner et al., 1992), where atmospherically forced small-scale turbulence can transform them into mixed layer waters. It thus seems likely that mixed layer instabilities enhance the exchange between mixed layer and thermocline, but future work will have to investigate what sets the rate of exchange.

6.5   Comparison to observations

A prominent feature of observed (Chapter 5) and modeled (Mensa et al., 2013; Sasaki et al., 2014) submesoscale flows is their seasonal modulation: they are much more energetic in winter than in summer. This is consistent with an energization of the submesoscale by baroclinic mixed layer instabilities. Given that these instabilities grow on time scales of order 1 day, they can quickly release large amounts of available potential energy stored in lateral buoyancy gradients in deep winter mixed layers and energize the submesoscale range. Mixed layer instabilities are instead weak in summer, when mixed layers are shallow and little potential energy is available for release. Mesoscale-driven surface frontogenesis, on the other hand, is not expected to drive a seasonal cycle in submesoscale turbulence, because the mesoscale eddies that generate submesoscale filaments through frontogenesis are about as strong in winter as they are in summer (or even slightly

141

stronger in summer, Qiu, 1999; Qiu and Chen, 2004; Sasaki et al., 2014).

Submesoscale flows observed in the wintertime Gulf Stream region (Chapter 5) are energetic throughout the deep mixed layer and decay rapidly below. The energy spectra roll off roughly like k^{-2} in the mixed layer and transition to roughly k^{-3} below. This spectral and vertical distribution of energy resembles that produced by our simple model of baroclinic mixed layer instabilities, which similarly has spectra that roll off roughly like k^{-2} in the mixed layer and more steeply in the thermocline (Fig. 6.14). The equilibrated turbulent flow produced by baroclinic mixed layer instabilities in this model is thus qualitatively consistent with the energetic wintertime submesoscale flows observed in the Gulf Stream region. In summer, on the other hand, submesoscale flows in the Gulf Stream region are weak. The spectral roll-off is a rapid k^{-3}, even close to the surface. This lack of submesoscale energy is expected from the lack of energy input from baroclinic mixed layer instabilities.

There is so far no observational evidence for submesoscale flows that are governed by mesoscale-driven surface frontogenesis as described by surface QG turbulence. The spectral slope of surface kinetic energy may be consistent with the predictions of surface QG theory (e.g. Le Traon et al., 2008), but the subsurface structure is not. The observations from the Gulf Stream region (Chapter 5) show that the spectral and vertical distribution of submesoscale energy is different from that produced by surface QG turbulence (Fig. 6.8), in both summer and winter. In winter, submesoscale flows are observed to be energetic throughout the mixed layer—not just in a thin surface layer as predicted by surface QG turbulence. In summer, when baroclinic mixed layer instabilities are not active and mesoscale-driven surface frontogenesis could dominate, submesoscale flows are observed to be weak—there is no surface-trapped enhancement as predicted by surface QG turbulence. This suggests that these weak summertime submesoscale flows are instead dominated by deep modes generated by mesoscale thermocline instabilities. Observations from Drake Passage also show no signature of mesoscale-driven surface frontogenesis (Rocha et al., 2015). Whether it dominates the energization of submesoscale flows elsewhere remains an open question. It may be expected to dominate outside the major current systems, where thermocline instabilities depend on surface buoyancy gradients and can more effectively drive a surface QG cascade (Charney, 1947; Tulloch et al., 2011; Roullet et al., 2012), but observations are lacking. Balanced submesoscale flows can also be masked by internal waves, especially where mesoscale eddies—and consequently any balanced submesoscale flows—are weak (Richman et al., 2012; Rocha et al., 2015; Chapters 2 and 3).

Spectra that are observed to fall off like k^{-3} in the seasonal thermocline in summer and in the permanent thermocline throughout the year (Chapter 5) cannot be reproduced by the model formulated in this study. These steep spectra are likely the result of deep Phillips-type instabilities (Phillips, 1954) and the potential enstrophy cascade of interior quasi-geostrophic turbulence (Charney, 1971). Such a cascade is not present in our model, which collapses the interior PV gradients into delta sheets at the interface between mixed

layer and thermocline and at the rigid bottom at the base of the thermocline. Our model does capture, however, the steepening of energy spectra below the winter mixed layer—it only lacks the appropriate thermocline dynamics to yield k^{-3} spectra.

The cascade dynamics of our simple QG representation of baroclinic mixed layer instabilities show that the mesoscale can effectively be energized by baroclinic mixed layer instabilities. This would suggest that not only submesoscale but also mesoscale eddies are more energetic in winter than in summer. Altimetry observations (Qiu, 1999; Qiu and Chen, 2004) and realistic model simulations (Sasaki et al., 2014), however, suggest that there is not a large modulation of mesoscale energy levels—mesoscale eddies are even slightly stronger in summer than in winter. This lack of wintertime energization of the mesoscale may result from a coupling with unbalanced motions that is not captured by QG dynamics.

The QG model makes the baroclinic mixed layer instabilities very effective in energizing the entire turbulent flow, because there is no forward energy cascade resulting in small-scale energy dissipation. If non-QG effects were allowed, a fraction of the energy extracted from the mean in the mixed layer would be dissipated at small scales (Capet et al., 2008c). This energy leak to small scales is likely as Rossby and Froude numbers become order 1 at scales of order 1 km. The possibility of an energy leak to small scales was demonstrated by Molemaker et al. (2010), who studied an Eady instability with Ro = Fr = 0.5, using the full Boussinesq equations. While much of the energy extracted from the mean is still trapped at large scales, as predicted by QG dynamics, some is lost to dissipation at small scales. A small leak of energy in the instability may make a big difference in the cascade dynamics, because that energy is not transferred back to mesoscales, where it would further enhance the extraction of potential energy from the mean. Such an effect could be parameterized in our QG model, but is beyond the scope of this study.

An additional sink occurs if the balanced flow interacts with ageostrophic instabilities in the mixed layer. In the presence of geostrophic shear, convective motions forced by the atmosphere are the result of symmetric rather than gravitational instabilities and are slantwise rather than upright (e.g. Emanuel, 1994; Haine and Marshall, 1998; Thomas and Lee, 2005). Symmetric instabilities can extract kinetic energy from the geostrophic shear, so they can drain energy from the balanced flow and increase dissipation (Taylor and Ferrari, 2010; Thomas et al., 2013). This is another way to render the inverse cascade less effective.

Additional sinks for energy in balanced flows in the mixed layer can be the interaction with externally forced near-inertial waves (e.g. Whitt and Thomas, 2015; Xie and Vanneste, 2015) and the interaction with surface gravity waves (McWilliams and Fox-Kemper, 2013; Hamlington et al., 2014). Both may drain enough energy out of the balanced flow to prevent an effective inverse cascade of submesoscale kinetic energy to mesoscales.

QG dynamics further do not allow for a feedback of eddies on the mean stratification

143

of the mixed layer (e.g. Fox-Kemper et al., 2008), so that no restratification can occur. In the real ocean, the restratification through baroclinic mixed layer instabilities is opposed by convective and mechanical mixing that is driven by atmospheric forcing. The fixed mean stratification in the QG model assumes that this forced vertical mixing is in balance with restratification. It is unclear whether and how vertical mixing alters the energy budget of the balanced flow in the mixed layer. It is also unclear what the effect of an unsteady atmospheric forcing is, which upsets the balance between vertical mixing and restratification.

Another possible explanation for the lack of a wintertime energization of the mesoscale by baroclinic mixed layer instabilities is that it takes a few months for the kinetic energy injected by submesoscale mixed layer instabilities to arrive at the mesoscale (Sasaki et al., 2014). This could also explain the (weak) summer maximum in mesoscale energy (Qiu, 1999; Qiu and Chen, 2004). This time dependence is not addressed in our simple model of baroclinic mixed layer instabilities, in which we prescribe perpetual winter conditions. If the time scale of turbulent equilibration is not much shorter than the seasonal time scale, it is likely that submesoscale—and possibly mesoscale—flows are not in statistical equilibrium. It remains to be investigated how a seasonally modulated mixed layer modifies the inverse cascade and the energization of the mesoscale by baroclinic mixed layer instabilities.

Realistic high-resolution simulations of the type employed by Shcherbina et al. (2013), Sasaki et al. (2014), and Gula et al. (2015) may be able to address these caveats and bridge the gap between observations and the idealized QG dynamics described here. These primitive equation and Boussinesq simulations can provide insights into how non-QG effects modify the submesoscale dynamics induced by mixed layer instabilities. An exploration of non-QG effects in more idealized setups may also prove useful. Primitive equation and Boussinesq simulations can also address how non-QG effects modify mesoscale-driven surface frontogenesis and allow an estimate of the importance of corrections to surface QG turbulence (cf. Hakim et al., 2002; Capet et al., 2008b; Klein et al., 2008; Roullet et al., 2012; Badin, 2012), providing another stepping stone for understanding observations.

6.6 Conclusions

The simple model formulated in this chapter sharpens our understanding of how baroclinic mixed layer instabilities can energize submesoscale turbulence and how this mechanism differs from mesoscale-driven surface frontogenesis. Our analysis suggests that the presence of a mixed layer has a profound effect on submesoscale turbulence. Lateral buoyancy gradients, combined with the low stratification in the mixed layer, provide a large amount of available potential energy that can be extracted by baroclinic instabilities in the mixed layer. The extraction of available potential energy from the large-scale mean is dominated by mesoscale eddies, but potential energy is subsequently transferred

downscale to the deformation radius of the mixed layer, where baroclinic instability converts it into kinetic energy. In the QG dynamics considered here, the energy lost to small scales is negligible. The entire energy extracted from the mean in the mixed layer is converted to kinetic energy around the deformation radius of the mixed layer and subsequently transferred back to larger scales in an inverse cascade that also energizes the thermocline below. Through this process, baroclinic mixed layer instabilities can energize the submesoscale range and even the mesoscale eddy field.

These turbulent dynamics follow a dual cascade similar to that present in two-layer QG flow (Rhines, 1977; Salmon, 1978; Larichev and Held, 1995) and in idealized continuously stratified QG flows (Smith and Vallis, 2002)—but with mixed layer modes. The energy in the baroclinic mode, which is baroclinic in the mixed layer and does not reach much into the thermocline, is transferred downscale through advection by the barotropic mode, which is barotropic in the mixed layer and decays surface-QG-like in the thermocline. Around the deformation radius of the mixed layer, baroclinic instability transforms baroclinic into barotropic energy, which then enters an inverse cascade.

The dynamics resulting from baroclinic mixed layer instabilities substantially differ from mesoscale-driven surface frontogenesis, as described by surface QG turbulence forced by a mesoscale eddy field, which is often invoked to explain energetic submesoscales. Surface QG turbulence can only energize a thin surface layer. Mixed layer instabilities, instead, energize the entire depth of the mixed layer. Vertical velocities are drastically enhanced in the presence of baroclinic mixed layer instabilities compared to surface QG flows of similar energy levels.

The enhancement of submesoscale energy throughout the mixed layer and the decay below its base, as generated by baroclinic mixed layer instabilities, are consistent with wintertime observations from the Gulf Stream region (Chapter 5). These observations as well as models (Mensa et al., 2013; Sasaki et al., 2014) also show that submesoscale flows are most energetic in winter, when baroclinic mixed layer instabilities are active. This evidence points to the importance of baroclinic mixed layer instabilities in energizing the submesoscale.

It remains to be investigated how deviations from QG dynamics affect submesoscale flows when a mixed layer is present. The restratification by mixed layer instabilities, the formation of buoyancy discontinuities, ageostrophic instabilities, and forced mixed layer turbulence all have the potential to modify the leading-order balanced dynamics described here. For example, it remains unclear how much of the submesoscale kinetic energy generated by baroclinic mixed layer instabilities is cascaded to mesoscales, how vertical velocities are modified by non-QG effects, and how the enhancement of vertical velocities by mixed layer instabilities translates into an exchange of fluid between mixed layer and thermocline.

It is hoped that the model and dynamics discussed here in an oceanographic context are of interest in a broader geophysical fluid dynamics context. Atmospheres often display layers of different stratification, which likely induce similar dynamics. These have partly

been explored for the terrestrial atmosphere and it seems likely that similar dynamics also occur in the atmospheres of other planets.

6.A Multi-layer model formulation

A QG system of n layers of constant PV, of thickness h_j and stratification N_j, consists of $n + 1$ conserved quantities that are advected by the geostrophic flow at their respective levels. Compared to the two-layer model considered in the main text, additional interface quantities analogous to θ_1 are present. The linear operator in the inversion relation (6.12) has tridiagonal structure:

$$
\mathbf{L} = f k_{\mathrm{h}}
\begin{pmatrix}
-\dfrac{\coth \mu_0}{N_0} & \dfrac{\operatorname{csch} \mu_0}{N_0} & & & \\[2ex]
\dfrac{\operatorname{csch} \mu_0}{N_0} & -\dfrac{\coth \mu_0}{N_0} - \dfrac{\coth \mu_1}{N_1} & \dfrac{\operatorname{csch} \mu_1}{N_1} & & \\[2ex]
 & \ddots & \ddots & \ddots & \\[2ex]
 & & \dfrac{\operatorname{csch} \mu_{n-1}}{N_{n-1}} & -\dfrac{\coth \mu_{n-1}}{N_{n-1}} - \dfrac{\coth \mu_n}{N_n} & \dfrac{\operatorname{csch} \mu_n}{N_n} \\[2ex]
 & & & \dfrac{\operatorname{csch} \mu_n}{N_n} & -\dfrac{\coth \mu_n}{N_n}
\end{pmatrix},
\tag{6.53}
$$

where $\mu_j = N_j k_{\mathrm{h}} h_j / f$. It may be more efficient to solve the inversion relation numerically instead of calculating the inverse of this matrix, which will generally be full.

One can also include a PV gradient due to differential rotation. This can be done using a trick described by Lindzen (1994): instead of using linear shear and constant stratification in the layers, one can use parabolic shear or a modified stratification profile, which allows cancellation of the contribution from the β-effect and retaining constant PV within the layers. The PV gradient due to β is then included in the PV sheets at the interfaces.

6.B Model formulation with buoyancy jump

If there is a buoyancy jump g' at the interface, the matching conditions must be modified. To ensure a continuous pressure at the interface at $z = -h + \eta$, we require

$$
\psi(-h^+) - \psi(-h^-) = -\frac{g'}{f}\eta.
\tag{6.54}
$$

Here, η is the perturbation of the interface between the constant-PV layers. The buoyancy equations (6.5) and (6.6) can be combined with the kinematic condition

$$
w = \frac{\partial \eta}{\partial t} + J(\psi, \eta),
\tag{6.55}
$$

146

applied at $z = -h^+$ and $z = -h^-$, to give

$$\frac{\partial \theta_1}{\partial t} + J(\psi_1, \theta_1) = 0, \qquad \frac{\partial \theta_2}{\partial t} + J(\psi_2, \theta_2) = 0, \tag{6.56}$$

where

$$\theta_1 = f\frac{b(-h^+)}{N_m^2} + f\eta, \qquad \theta_2 = f\frac{b(-h^-)}{N_t^2} + f\eta, \qquad \psi_1 = \psi(-h^+), \qquad \psi_2 = \psi(-h^-). \tag{6.57}$$

Together with the conservation of surface and bottom buoyancy,

$$\theta_0 = -f\frac{b(0)}{N_m^2}, \qquad \theta_3 = f\frac{b(-H)}{N_t^2}, \tag{6.58}$$

and the inversion relation obtained by solving (6.13) with the matching conditions above, the model is complete. It now consists of four conserved quantities.

CHAPTER 7

BAROCLINIC MIXED LAYER INSTABILITIES IN THE PRESENCE OF CONVECTION

7.1 Motivation

Atmospheric cooling and surface winds frequently mix the surface layer of the ocean. The resulting mixed layer mediates the transfer of heat and momentum between the atmosphere and ocean and thereby affects both the atmospheric climate and the oceanic general circulation. The evolution of the ocean mixed layer has traditionally been understood column by column: atmospheric cooling and wind forcing leads to mixing and deepening of the mixed layer into the thermocline below. It is becoming increasingly clear, however, that lateral exchanges contribute crucially to shaping the mixed layer.

Baroclinic instability in the mixed layer, one such agent of lateral exchange, is thought to be one of the main agents to restore stratification in the mixed layer after the atmospheric forcing ceases (e.g. Spall, 1995; Haine and Marshall, 1998; Boccaletti et al., 2007; Fox-Kemper et al., 2008). This restratification modifies the surface properties and thereby feeds back on the surface fluxes of heat and momentum. Baroclinic eddies can also achieve exchanges between the mixed layer and the thermocline below—a process that may be important in subducting heat and atmospheric gases like carbon into the thermocline as well as bringing nutrients up into the mixed layer, where they can be used for photosynthesis (e.g. Thomas et al., 2008).

Baroclinic instabilities in the mixed layer have recently been suggested to drive a seasonal cycle in submesoscale turbulence (Mensa et al., 2013; Sasaki et al., 2014; Chapter 5). In winter, strong baroclinic instabilities in deep mixed layers are thought to create strong buoyancy gradients at horizontal scales of order 10 km. This suggests that baroclinic instabilities can tap into the large reservoir of available potential energy in the deep winter mixed layer, despite the presence of forced convection that keeps the mixed layer deep and unstratified. In summer, in contrast, variability at horizontal scales of order 10 km is weak. One possible explanation for this absence of submesoscale energy is that baroclinic instabilities are damped out by convection that has short time scales in shallow summer mixed layers. Another possibility is that baroclinic instabilities do develop, but the small amount of potential energy that can be released is not sufficient to energize the submesoscale range.

This motivates our attempt to understand under what conditions baroclinic instabilities develop. In particular, we aim to investigate under what conditions convection arrests the growth of baroclinic disturbances and when instead baroclinic instability develop and manage to restratify the mixed layer against the concomitant action of convective mixing. Mahadevan et al. (2010) argued that restratification occurs when vertical buoyancy fluxes due to baroclinic eddies are larger than the fluxes induced by the surface forcing, but their simulations had too coarse a resolution to explicitly represent convection.

Using a numerical model that allows both convective overturns and baroclinic instabilities, we here confirm that restratification occurs when baroclinic fluxes are larger than those induced by surface fluxes. When the fluxes generated by convection are larger, however, we find that baroclinic instabilities still develop. Baroclinic eddies, while too weak

to restratify, organize the convective turbulence and modify the bulk properties of the mixed layer.

7.2 Approach

To study baroclinic instabilities in the presence of convection, we depart from previous studies of baroclinic mixed layer instabilities that considered the transient spin down of a mixed layer front. In such simulations, the front slumps under the effect of baroclinic eddies and the mixed layer quickly restratifies (e.g. Haine and Marshall, 1998; Boccaletti et al., 2007). There is a succession from upright or slantwise convection to baroclinic eddies as the stratification and hence the Richardson number increases, which is expected from linear stability analysis (Stone, 1966b). Convection and baroclinic eddies do not coexist in these simulations, because the restratification by baroclinic eddies shuts off convection.

Mahadevan et al. (2010) attempted to achieve an equilibrium between the destabilizing effect of surface forcing and the restratifying effect of baroclinic eddies by blowing a wind down a mixed layer front. The Ekman transport of such a down-front wind pushes dense over light water, destabilizing the surface layers (Thomas and Lee, 2005). The destabilizing effect of this forcing evolves with the flow, however, because the Ekman buoyancy flux depends on the orientation of the front—which under the influence of baroclinic instability starts to meander. The simulations described in Mahadevan et al. (2010) were thus necessarily transient; an equilibrium was not achieved, because the increasing misalignment of front and wind stress reduced the effective buoyancy forcing. In the setup described below, we instead use direct buoyancy forcing, which allows us to control the forcing strength and achieve a statistical equilibrium. We can thus study the competition between convection and baroclinic instability in a more controlled environment and reveal that baroclinic instability develops even in the presence of strong convection.

In Chapter 6, we took an extreme approach and studied the evolution of baroclinic instabilities in the ocean mixed layer using a quasi-geostrophic model. Under quasigeostrophic scaling, the restratification by baroclinic eddies is neglected and the mixed layer stratification remains fixed. This implicitly assumes an equilibrium between the restratification by baroclinic eddies and the destabilizing effects of surface forcing. One of the goals here is to understand under what conditions such an equilibrium can exist and whether or not baroclinic eddies can evolve unaffected by convective flows, being influenced only through the mean stratification.

We study a setup that is designed to provide a simple example of coexisting baroclinic and convective flows in a statistical equilibrium. We impose a fixed meridional buoyancy gradient in thermal wind balance with a zonal shear (e.g. Taylor and Ferrari, 2010), which provides baroclinicity for the instability. In addition, we apply a destabilizing buoyancy flux at the surface and bottom of the domain. For simplicity, there is no representation of

the thermocline, so any evolution of the mixed layer depth is neglected. The restratification by baroclinic eddies for our mixed layer with a rigid bottom, however, is expected to be similar to the case with a thermocline, because instead of folding thermocline water with high potential vorticity (PV) into the mixed layer, such high-PV waters are generated at the bottom boundary (Garner et al., 1992). The artificial buoyancy forcing at the bottom is necessary to allow the system to reach equilibrium. In the real ocean, surface-forced convection can deepen the mixed layer into the thermocline and an equilibrium is never reached. It seems worthwhile to study the equilibrium first, however, before the more realistic but complicated transient problem is addressed.

In this setup, we study the flow under different forcing conditions. The key parameter determining the strength of the forcing is $\varepsilon = F/f\Lambda^2 H^2$, where F is the buoyancy forcing, f the Coriolis parameter, Λ the background shear, and H the depth. We vary this parameter over three orders of magnitude. Over this range, we cover the transition from a weakly forced regime, in which baroclinic instability restratifies the mixed layer and shuts off convection, to a very strongly forced regime, in which convection is essentially upright.

7.3 Setup

We study the full non-hydrostatic Boussinesq equations between two horizontal solid plates at $z = -H$ and $z = 0$. We solve for the perturbations that develop on top of an imposed background flow with zonal shear Λ and meridional buoyancy gradient $-f\Lambda$ (Taylor and Ferrari, 2010). The perturbations satisfy periodic boundary conditions in the horizontal coordinates x and y in a domain of zonal extent L_x and meridional extent L_y. The dynamics are given by the perturbation equations:

$$u_t + \Lambda z u_x + \Lambda w + \boldsymbol{u} \cdot \nabla u - fv = -\phi_x + \kappa \nabla^2 u, \tag{7.1}$$

$$v_t + \Lambda z v_x + \boldsymbol{u} \cdot \nabla v + fu = -\phi_y + \kappa \nabla^2 v, \tag{7.2}$$

$$w_t + \Lambda z w_x + \boldsymbol{u} \cdot \nabla w = b - \phi_z + \kappa \nabla^2 w, \tag{7.3}$$

$$b_t + \Lambda z b_x - f\Lambda v + \boldsymbol{u} \cdot \nabla b = \kappa \nabla^2 b, \tag{7.4}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{7.5}$$

where u, v, and w are the zonal, meridional, and vertical velocities, ϕ is the density-normalized pressure, b is buoyancy, κ is the viscosity/diffusivity (unity Prandtl number), and t is time. At the boundaries at $z = -H$ and $z = 0$, we prescribe the buoyancy flux F,

$$-\kappa b_z = F, \tag{7.6}$$

and impose no-normal-flow and free-slip conditions: $u_z = 0$, $v_z = 0$, and $w = 0$.

This setup constitutes an energetically consistent system, in which the only energy source is the potential energy supplied by the buoyancy forcing (plus small sources due to diffusion and viscosity) and the only sink is viscous dissipation. The potential energy equation is

$$\langle -z\overline{b}\rangle_t = FH - f\Lambda\langle z\overline{v}\rangle - \langle \overline{w'b'}\rangle + \kappa[\overline{b}(0) - \overline{b}(-H)]. \tag{7.7}$$

Angle brackets denote a vertical integral over depth of the domain, the overline denotes a horizontal average over the domain, and primes denote deviations from this horizontal average. Potential energy is forced by the buoyancy flux at the boundaries (first term on the right). It can be converted to mean kinetic energy by gravitational slumping (second term on the right) and to eddy kinetic energy by buoyancy production (third term on the right). Downward diffusion of buoyancy also increases the potential energy (fourth term on the right), but this term is generally small for the values of κ we use.

To derive the mean kinetic energy budget, we first form the momentum equations of the horizontal mean velocities. These are free to evolve as they are forced by momentum flux convergences:

$$(\Lambda z + \overline{u})_t + (\overline{u'w'})_z - f\overline{v} = \kappa\overline{u}_{zz}, \tag{7.8}$$

$$\overline{v}_t + (\overline{v'w'})_z + f\overline{u} = \kappa\overline{v}_{zz}. \tag{7.9}$$

The horizontal mean vertical velocity vanishes, as required by continuity. The mean kinetic energy budget then reads

$$\frac{1}{2}\langle(\Lambda z + \overline{u})^2 + \overline{v}^2\rangle_t = f\Lambda\langle z\overline{v}\rangle + \langle(\Lambda + \overline{u}_z)\overline{u'w'}\rangle + \langle\overline{v}_z\overline{v'w'}\rangle$$
$$- \kappa\langle(\Lambda + \overline{u}_z)^2 + \overline{v}_z^2\rangle + \kappa\Lambda[\Lambda H + \overline{u}(0) - \overline{u}(-H)]. \tag{7.10}$$

Gravitational slumping appears as a source of mean kinetic energy (first term on the right). Shear production (second and third terms on the right) converts mean to eddy kinetic energy. Dissipation of mean kinetic energy (fourth term on the right) is generally small, as is the input of kinetic energy by the surface and bottom stress implied by prescribing a background shear that does not vanish at the boundaries (fifth term on the right). Finally, the eddy kinetic energy budget reads

$$\frac{1}{2}\langle\overline{u'^2} + \overline{v'^2} + \overline{w'^2}\rangle_t = \langle\overline{w'b'}\rangle - \langle(\Lambda + \overline{u}_z)\overline{u'w'}\rangle - \langle\overline{v}_z\overline{v'w'}\rangle - \kappa\langle\overline{|\nabla u'|^2} + \overline{|\nabla v'|^2} + \overline{|\nabla w'|^2}\rangle \tag{7.11}$$

Buoyancy production (first term on the right) and shear production (second and third terms on the right) appear as sources here. The dissipation term (fourth term on the right) generally dominates the dissipation of energy.

Unlike an analogous quasi-geostrophic system (e.g. Bretherton and Karweit, 1975; Salmon, 1980; Haidvogel and Held, 1980; Larichev and Held, 1995), which we used in Chapter 6 to study baroclinic mixed layer instabilities, this Boussinesq system can exhaust the potential energy available for release by baroclinic instability. While the horizontal buoyancy gradient is fixed, the stratification is free to evolve. Under the effect of

153

baroclinic eddies, the stratification increases until no unstable baroclinic mode fits into the finite domain. The potential energy available for release can thus be exhausted and has to be supplied by buoyancy forcing. In quasi-geostrophic dynamics, on the other hand, one cannot impose a buoyancy flux at the boundaries. The stratification is instead held fixed, as required by quasi-geostrophic scaling, which emulates the supply of potential energy by boundary forcing. The fixed stratification, together with the fixed horizontal buoyancy gradient, however, provides an infinite reservoir of potential energy available for release by baroclinic eddies.

It should be noted that our setup is different from that studied by Molemaker et al. (2010). They removed the horizontal mean momentum flux divergences in (7.1–7.3) and introduced an evolution equation for the horizontal buoyancy gradient. The setup described above is well suited for our purposes, because it has a straightforward energy budget and no direct comparison to quasi-geostrophic dynamics (though not impossible) is attempted.

The adiabatic and frictionless form of equations (7.1–7.5) conserves the Ertel PV of the full flow:

$$q = [f\mathbf{z} + \nabla \times (\Lambda z\mathbf{x} + \mathbf{u})] \cdot (-f\Lambda\mathbf{y} + \nabla b) \tag{7.12}$$

$$= (w_y - v_z)b_x + (\Lambda + u_z - w_x)(-f\Lambda + b_y) + (f + v_x - u_y)b_z, \tag{7.13}$$

where x, y, and z are the Cartesian unit vectors. In our interpretation of simulation results, we will make use of two fundamental properties of PV dynamics. First, baroclinic instabilities increase the bulk PV by injecting positive PV at the boundaries and folding it into the interior of the fluid (Nakamura and Held, 1989). Analogously, baroclinic instability in the real ocean lifts high-PV water from the thermocline into the mixed layer (Garner et al., 1992). Second, convection occurs when $q < 0$ and rapidly restores PV to the marginally stable state $q = 0$ (Hoskins, 1974; Emanuel, 1994; Haine and Marshall, 1998). In the presence of baroclinic shear, this convection is slantwise and the result of symmetric rather than gravitational instabilities. Slantwise convection can produce positive buoyancy stratification ($b_z > 0$) by homogenizing buoyancy along tilted absolute momentum surfaces, but it cannot increase the bulk PV above zero. The telltale signs of such slantwise convective adjustment, regions of positive stratification and near-zero PV, were recently observed at strong ocean fronts (D'Asaro et al., 2011) and have long been known to occur in midlatitude atmospheric fronts (Emanuel, 1988).

We here want to emphasize the difference between convecting and restratifying states, rather than between upright and slantwise convection. We diagnose restratification using PV: when baroclinic instabilities are strong enough, they increase the PV to $q > 0$, which shuts off convection—including the slantwise kind. Convecting states, on the other hand, have a PV close to zero. We will see, however, that a $q = 0$ state does not mean that no baroclinic modes are present. It only means that they are too weak to overcome the destruction of PV by the buoyancy forcing.

The system of equations (7.1–7.5) is solved numerically using the MITgcm (Marshall et al., 1997), modified to include the additional terms arising from the interaction of the perturbations with the background flow. All simulations have a depth $H = 100$ m, a grid resolution $\Delta x = \Delta y = \Delta z = 2$ m, and $\kappa = 10^{-3}$ m s^{-2}. We choose harmonic viscosity/diffusivity over a Smagorinsky (1963) closure, because it is easier to close the energy budget and it remains unclear whether a Smagorinsky scheme is appropriate for flows with strong baroclinic fronts. We do not expect the subgrid closure to affect the results discussed in the following.

To reveal the parameter dependence of our setup, we nondimensionalize the system (7.1–7.5) using the following scales:

$$t \sim f^{-1}, \qquad x,y \sim \Lambda H/f, \qquad z \sim H, \qquad u,v \sim \Lambda H, \qquad w \sim fH, \qquad (7.14)$$

$$b \sim \Lambda^2 H, \qquad \phi \sim \Lambda^2 H^2. \qquad (7.15)$$

This follows Stone (1971)—except for the scale of buoyancy and pressure, for which Stone had the imposed stratification as an available scale. The nondimensional system reads

$$u_t + zu_x + w + \boldsymbol{u} \cdot \nabla u - v = -\phi_x + \mathcal{D}u, \qquad (7.16)$$

$$v_t + zv_x + \boldsymbol{u} \cdot \nabla v + u = -\phi_y + \mathcal{D}v, \qquad (7.17)$$

$$\delta^2(w_t + zw_x + \boldsymbol{u} \cdot \nabla w) = b - \phi_z + \delta^2 \mathcal{D}w, \qquad (7.18)$$

$$b_t + zb_x - v + \boldsymbol{u} \cdot \nabla b = \mathcal{D}b, \qquad (7.19)$$

$$\nabla \cdot \boldsymbol{u} = 0. \qquad (7.20)$$

We abbreviate the viscosity/diffusion operator by $\mathcal{D}A = \gamma[\delta^2(A_{xx}+A_{yy})+A_{zz}]$. The boundary conditions are $-\gamma b_z = \varepsilon$, $u_z = 0$, $v_z = 0$, and $w = 0$ at $z = -1$ and $z = 0$. The nondimensional parameters of the problem are the ratio of the forcing to the scaling for the baroclinic flux proposed by Fox-Kemper et al. (2008), $\varepsilon = F/f\Lambda^2 H^2$, and Stone's (1971) non-hydrostatic parameter, $\delta = f/\Lambda$. Further parameters measure the strength of viscosity/diffusion, $\gamma = \kappa/fH^2$, which is assumed to be unimportant for the bulk behavior as long as it is sufficiently small, and the nondimensional domain size, $\lambda_x = fL_x/\Lambda H$ and $\lambda_y = fL_y/\Lambda H$. The nondimensional potential energy budget reads

$$\langle -z\overline{b}\rangle_t = \varepsilon - \langle z\overline{v}\rangle - \langle \overline{w'b'}\rangle + \gamma[\overline{b}(0) - \overline{b}(-1)], \qquad (7.21)$$

the mean kinetic energy budget

$$\tfrac{1}{2}\langle(z + \overline{u})^2 + \overline{v}^2\rangle_t = \langle z\overline{v}\rangle + \langle(1 + \overline{u}_z)\overline{u'w'}\rangle + \langle \overline{v}_z\overline{v'w'}\rangle$$
$$- \gamma\langle(1 + \overline{u}_z)^2 + \overline{v}_z^2\rangle + \gamma[1 + \overline{u}(0) - \overline{u}(-1)], \quad (7.22)$$

and the eddy kinetic energy budget

$$\tfrac{1}{2}\langle\overline{u'^2} + \overline{v'^2} + \delta^2\overline{w'^2}\rangle_t = \langle\overline{w'b'}\rangle - \langle(1 + \overline{u}_z)\overline{u'w'}\rangle - \langle\overline{v}_z\overline{v'w'}\rangle$$
$$- \gamma\langle\delta^2(\overline{u_x'^2} + \overline{u_y'^2}) + \overline{u_z'^2} + \delta^2(\overline{v_x'^2} + \overline{v_y'^2}) + \overline{v_z'^2} + \delta^2[\delta^2(\overline{w_x'^2} + \overline{w_y'^2}) + \overline{w_z'^2}]\rangle. \quad (7.23)$$
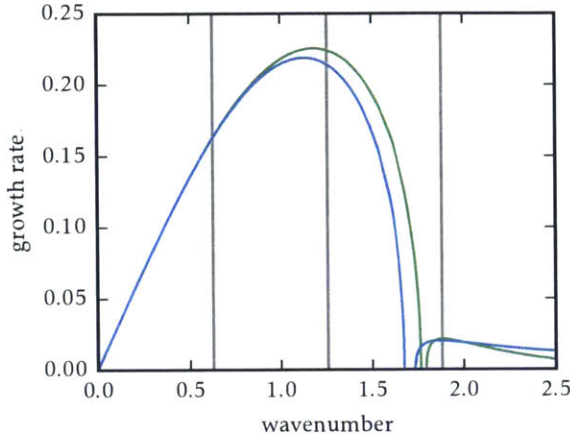
155

Figure 7.1: Growth rates from the linear stability analysis for $\delta = 1$ (——) and $\delta = 0$ (——) for a background state with $q = 0$. The vertical lines denote resolved wavenumbers in a domain that is $\lambda_x = 10$ wide.

The nondimensional form of PV is

$$q = (\delta^2 w_y - v_z)b_x + (1 + u_z - \delta^2 w_x)(-1 + b_y) + (1 + v_x - u_y)b_z. \qquad (7.24)$$

Note that a state with no stratification and no perturbation to the background flow has a background PV of $q = -1$.

The scales chosen above are appropriate when rotation and baroclinicity play an important role in the dynamics. In the limit of strong forcing ($\varepsilon \to \infty$), however, the dynamics revert to the classic problem of non-rotating upright convection between two plates (e.g. Emanuel, 1994). In this limit, more appropriate scales can be formed by the forcing F and depth H (e.g. Marshall and Schott, 1999):

$$t \sim (H^2/F)^{1/3}, \qquad x, y, z \sim H, \qquad u, v, w \sim (FH)^{1/3}, \qquad (7.25)$$

$$b \sim (F^2/H)^{1/3}, \qquad \phi \sim (FH)^{2/3}. \qquad (7.26)$$

This nondimensionalization will prove useful when considering strongly forced simulations.

7.4 Experiments

We focus on a series of experiments with $\delta = 1$, for two reasons. First, linear stability analysis suggests that with this parameter, the baroclinic instability is already close to the hydrostatic limit $\delta \to 0$ (Fig. 7.1, Stone, 1971) that is of particular interest for strong ocean fronts, where shears can greatly exceed f (e.g. D'Asaro et al., 2011). Second, this parameter choice is practical, because the most unstable baroclinic mode then has an aspect ratio close to unity (Stone, 1966b), which allows us to conveniently resolve the baroclinic mode

156

as well as upright convection. The increasing aspect ratio of the baroclinic instability as $\delta \rightarrow 0$ makes the concurrent simulation of baroclinic and convective motions increasingly challenging.

We recapitulate the linear properties of the system by performing a linear stability analysis for the baroclinic axis (no meridional variations) following Stone (1971). We consider a mean state with stratification $b_z = 1$, such that $q = 0$, instead of a basic state with zero stratification. This should be taken as representative of the weakly stratified state near which strong convection will keep the system. The growth rates given in Fig. 7.1 show a low-wavenumber geostrophic mode and a high-wavenumber ageostrophic mode (Stone, 1970; Nakamura, 1988; Molemaker et al., 2005). We focus on the geostrophic mode, because it represents the balanced dynamics whose interaction with convection we seek to investigate. This mode is most unstable at a zonal wavenumber $k = 1.13$ and has a short-wave cutoff at $k = 1.68$. Because of the finite domain, we only resolve a discrete set of wavenumbers (Fig. 7.1). The domain size $\lambda_x = 10$ is chosen such that roughly two wavelengths of the most unstable mode fit into the domain. For comparison, we also perform simulations in narrow domains of zonal extent $\lambda_x = 2$, in which no baroclinic instability can develop (cf. Taylor and Ferrari, 2010). Symmetric instabilities, which occur for $q < 0$, only require across-shear variations, so they are well represented in both domains.

We perform both unforced ($\varepsilon = 0$) and forced simulations with the forcing parameter ε varied over three orders of magnitude. All forced simulations are performed in both a narrow and a wide domain, to isolate the effect of the baroclinic mode. The simulation parameters are listed in Table 7.1.

7.4.1 Unforced simulation

We start with an unforced simulation ($\varepsilon = 0$) to estimate the (dimensionless) buoyancy flux $B = \langle \overline{w'b'} \rangle$ generated by baroclinic instability in the absence of convection. The expectation is, following Mahadevan et al. (2010), that restratification occurs for weak forcing $\varepsilon < B$ and not for strong forcing $\varepsilon > B$.

We determine the flux B empirically instead of using the Fox-Kemper et al. (2008) scaling, because that scaling does not capture a dependence of the flux on the eddy scale (see Chapter 8). As we will see, the eddy scale in our setup is set by the size of the domain. The dependence of the flux B on the eddy scale then means that B depends on the size of the domain. What sets the eddy size in the real ocean is a separate question, which we circumvent by determining B empirically for the domain size used.

We initialize the simulation with a uniform stratification $b_z = 1$ and add small random perturbation to u, v, and b. This state is marginally stable to symmetric instability, so the development is dominated by the baroclinic mode (Fig. 7.1). The evolution, once the system has reached finite amplitude and nonlinear terms have become impor-

Table 7.1: Parameter values for the numerical simulations. We list both nondimensional parameters and the dimensional parameters used in the MITgcm.

| δ	ε	λ_x	λ_y	F [m² s⁻³]	f [s⁻¹]	Λ [s⁻¹]	H [m]	L_x [m]	L_y [m]	Δt [s]
1.0	0	10	10	0	1.00×10^{-4}	1.00×10^{-4}	100	1000	1000	10
	0.01	10	10	10^{-9}	2.15×10^{-4}	2.15×10^{-4}	100	1000	1000	5
		2	10	10^{-9}	2.15×10^{-4}	2.15×10^{-4}	100	200	1000	5
	0.1	10	10	10^{-9}	1.00×10^{-4}	1.00×10^{-4}	100	1000	1000	10
		2	10	10^{-9}	1.00×10^{-4}	1.00×10^{-4}	100	200	1000	10
	1	10	10	10^{-9}	4.64×10^{-5}	4.64×10^{-5}	100	1000	1000	20
		2	10	10^{-9}	4.64×10^{-5}	4.64×10^{-5}	100	200	1000	20
	10	10	10	10^{-9}	2.15×10^{-5}	2.15×10^{-5}	100	1000	1000	20
		2	10	10^{-9}	2.15×10^{-5}	2.15×10^{-5}	100	200	1000	20
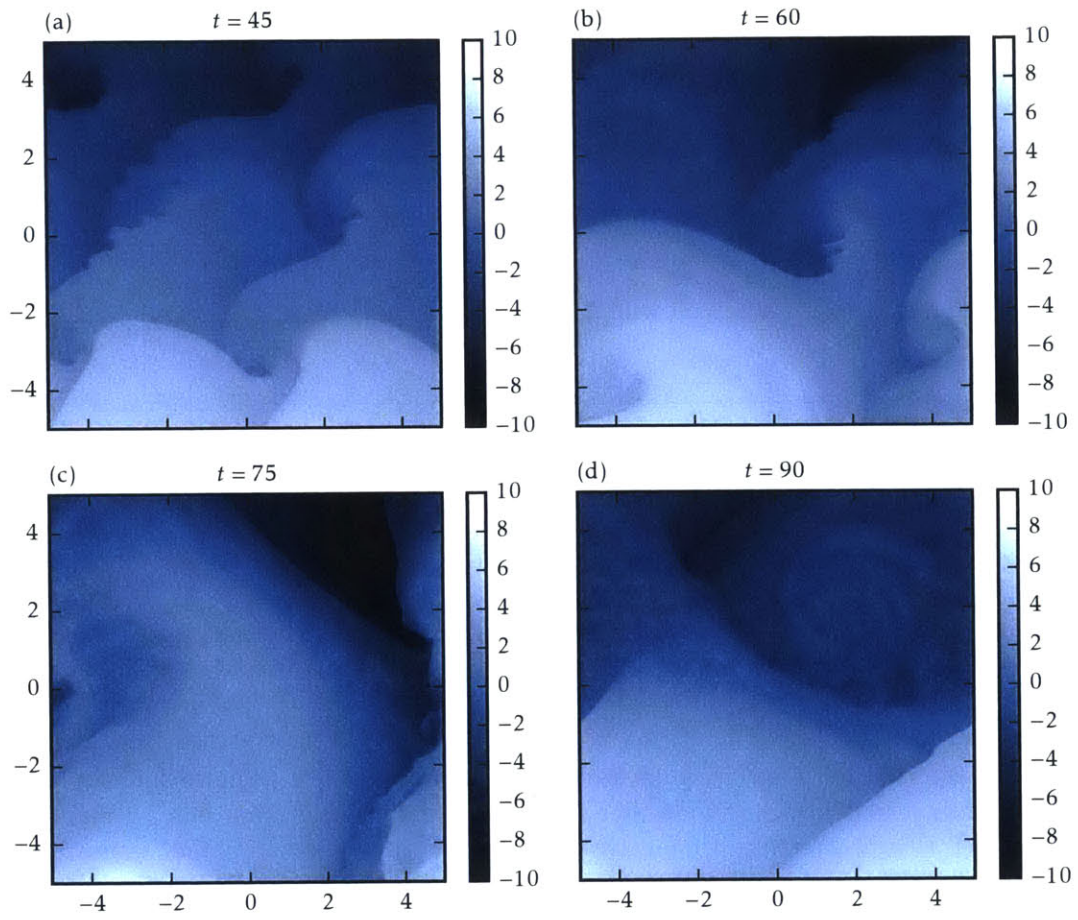
Figure 7.2: Snapshots of $b' - y$ at $z = 0$ for unforced simulation at different stages of the baroclinic instability
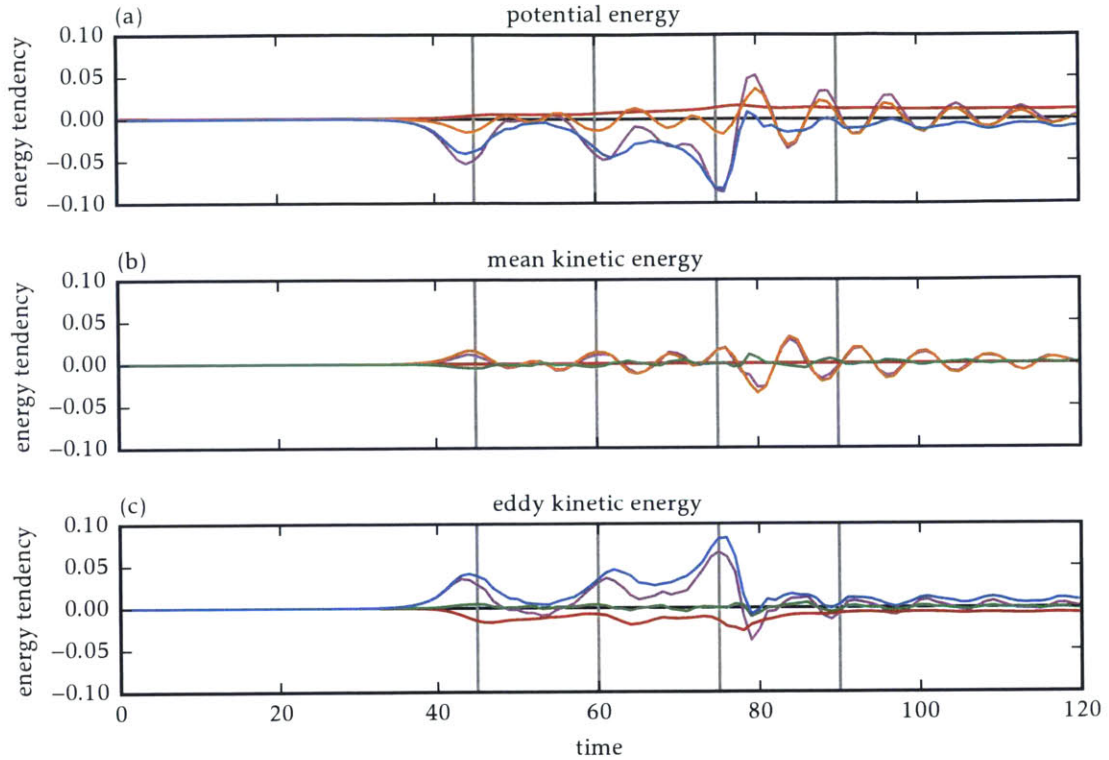
159

Figure 7.3: Time evolution of terms in the (a) potential, (b) mean kinetic, and (c) eddy kinetic energy budget for the unforced simulation ($\varepsilon = 0$): buoyancy production (——), gravitational slumping (——), shear production (——), diffusion/dissipation (——), and tendency (——)

tant, is illustrated with surface buoyancy snapshots in Fig. 7.2. In accordance with the linear stability analysis, the mode that reaches finite amplitude first has a zonal wavelength of 5 (Fig. 7.2a). This first wave perturbation then grows in size, spanning the whole domain size at wavelength 10 (Fig. 7.2b,c), both due to the linear growth of the wavelength-10 mode and nonlinear energy transfer from the wavelength-5 mode. The domain-filling mode reaches finite amplitude, breaks, and rolls up into a domain-filling vortex (Fig. 7.2c,d). These baroclinic perturbations drive a vertical buoyancy flux that restratifies the system. The restratification shifts the short-wave cutoff of the instability to larger scales and eventually renders the configuration stable to all modes that fit into the domain. When this occurs, the domain-filling vortex starts decaying.

This evolution can also be traced in the energy budget shown in Fig. 7.3. We show all term in (7.21–7.23), but here focus on the buoyancy production $\langle \overline{w'b'} \rangle$ that dominates the generation of eddy kinetic energy. This term has three distinct peaks at $t = 44$, $t = 62$,

and $t = 76$. The first peak in buoyancy production is associated with the wavelength-5 mode (Fig. 7.2a) and the third peak is associated with the wavelength-10 mode (Fig. 7.2c). The second peak has both modes contributing and represents the transition from the smaller to the larger mode (Fig. 7.2b). The third peak is largest and reaches $B = 0.08$. After this peak, the buoyancy production drops, as the restratification has rendered the configuration stable.

7.4.2 Weak forcing

We now turn to forced simulations and start with a weak forcing $\varepsilon = 0.01$, which is much smaller than the $B = 0.08$ estimated from the unforced simulation. The simulation is initialized as before with the $q = 0$ state with $b_z = 1$ and no flow anomalies. Convection quickly sets in, but the baroclinic instability develops and restratifies the system. The flow reaches a statistical equilibrium around $t = 100$, characterized by a domain-filling baroclinic vortex (Fig. 7.4a) and strong stratification greatly exceeding the initial $b_z = 1$ (Fig. 7.5a). The equilibrium is achieved by a balance between the destabilization by the forcing and the restratification by the baroclinic mode. Only in the center of the cyclonic vortex, where the stratification is weakest, there is evidence of convection. The convective plumes reaching the surface can be discerned in the surface buoyancy field, which is otherwise smooth (Fig. 7.4a).

In the transient development, the peak buoyancy production of about 0.04 is larger than the imposed forcing (Fig. 7.6). This large buoyancy production draws on the potential energy of the initial state, but cannot be sustained by the forcing. The system restratifies, which stabilizes the baroclinic modes and leads to a decay in buoyancy production. The baroclinic instability is not shut off entirely like in the unforced simulation, however, because the destabilizing forcing keeps the system slightly unstable to baroclinic instability. In statistical equilibrium the buoyancy production thus settles to a value of about 0.01, corresponding to the forcing ε. The energy cycle of this equilibrated state is straightforward: potential energy is created by the forcing, potential energy is converted to eddy kinetic energy by buoyancy production, and eddy kinetic energy is dissipated by viscosity. The mean kinetic energy plays no significant role in the energetics.

To further characterize the statistical equilibrium, we show the vertical profiles of the horizontal mean PV averaged in time over the equilibrated state (Fig. 7.7a). PV has increased from the initial $q = 0$ and is significantly positive over the bulk of the domain. It is negative only near the boundaries, where the boundary conditions force it to be $\bar{q} = -(1 + \varepsilon/\gamma) = -23$ with our choice of $\gamma = 4.64 \times 10^{-4}$.

This equilibrium state is very different from that achieved in a narrow domain (cf. Taylor and Ferrari, 2010). In a domain of width $\lambda_x = 2$, no baroclinic instability can develop, because no unstable mode fits into the domain (Fig. 7.1). The imposed buoyancy flux is carried by slantwise convection, rather than by a baroclinic vortex. The system adjusts to
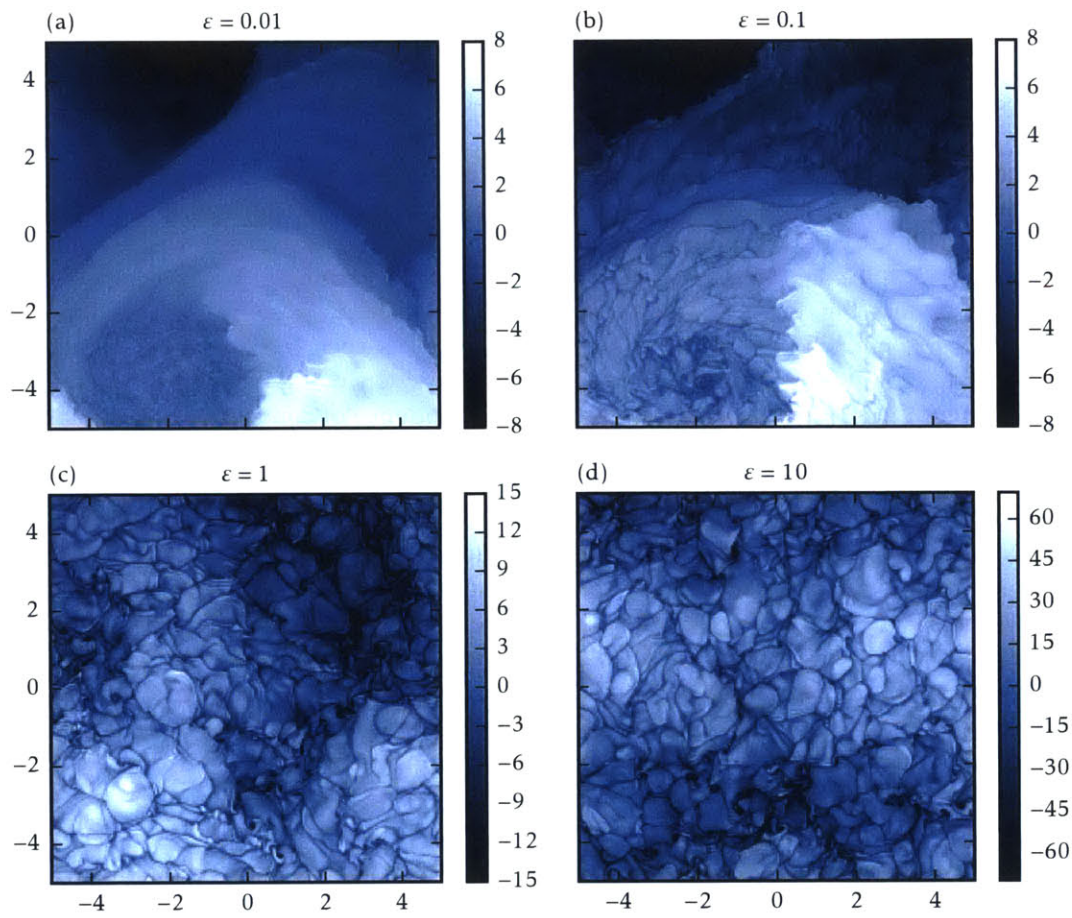
161

Figure 7.4: Snapshots of $b' - y$ at $z = 0$ and $t = 150$ (in the equilibrated stage) for different forcing strengths
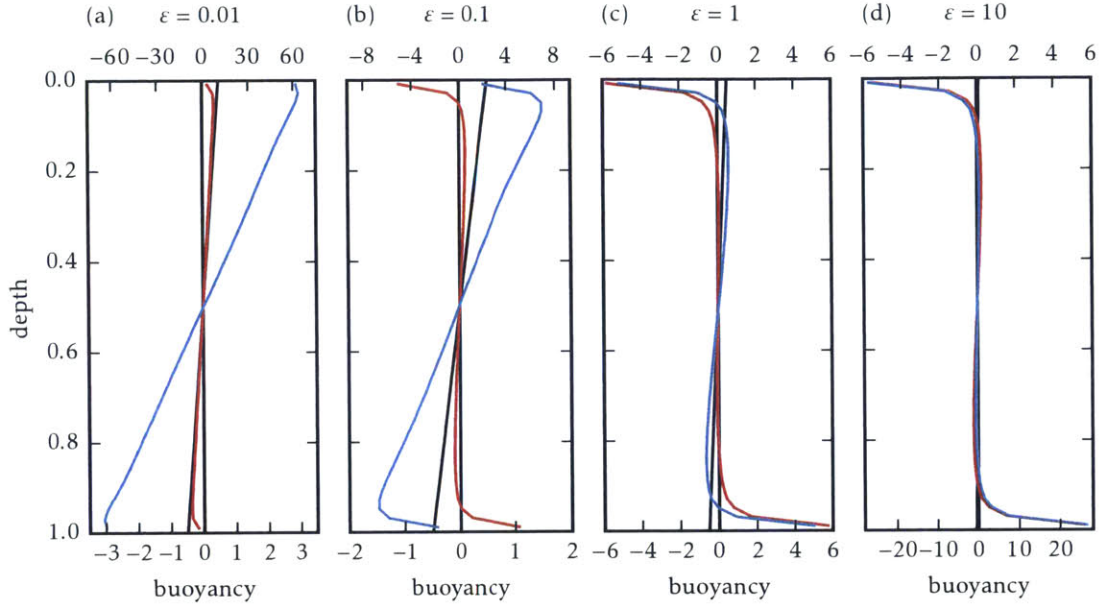
162

Figure 7.5: Horizontally averaged buoyancy profiles for wide (——) and narrow (——) domains and for different forcing strengths as indicated in the titles. The lower buoyancy axis is given in the standard nondimensionalization (7.14–7.15), the upper in the alternative nondimensionalization (7.25–7.26). All profiles are averaged in time over the equilibrated range $100 < t < 200$.

a state that is marginally stable to symmetric instability, that is one with $\overline{q} = 0$ (Fig. 7.7a). The stratification is determined by the imposed geostrophic shear and settles at about $\overline{b} = 1$. This is much weaker a stratification than achieved by baroclinic restratification in the wide domain (Fig. 7.5a).

7.4.3   Moderate forcing

Next, we consider a moderately forced case, in which the forcing $\varepsilon = 0.1$ is only slightly larger than the $B = 0.08$ estimated from the unforced simulation. Again, convection sets in very quickly. A baroclinic mode develops and appears after $t = 50$. The system settles into a statistical equilibrium that exhibits a coexistence of a domain-filling baroclinic vortex and smaller-scale convection (Fig. 7.4b).

The interplay between the baroclinic and convective flows leads to an interesting state that exhibits strong buoyancy stratification but a PV close to zero (Fig. 7.5b, 7.7b). PV is restored to zero by convection, but the convective flows are strongly modified by fronts generated by the finite amplitude baroclinic instability. At these fronts, the convection is
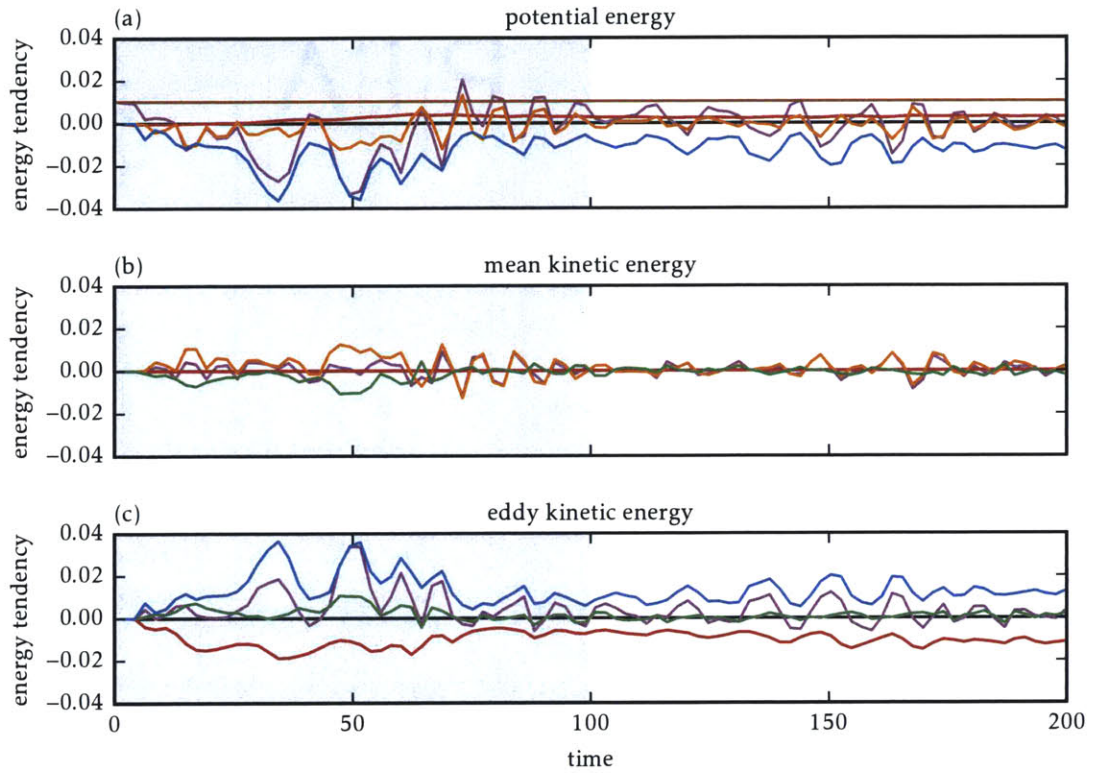
163

Figure 7.6: Time evolution of terms in the (a) potential, (b) mean kinetic, and (c) eddy kinetic energy budget for the weakly forced simulation ($\varepsilon = 0.01$): forcing (——), buoyancy production (——), gravitational slumping (——), shear production (——), diffusion/dissipation (——), and tendency (——)

slantwise rather than upright, giving rise to significant buoyancy stratification.

The buoyancy stratification is stronger than the $\overline{b}_z = 1$ expected from slantwise convection operating on the background shear (Fig. 7.5b). In fact, pure convective flows for this forcing—as simulated in a narrow domain—lead to an even smaller stratification, because momentum fluxes eliminate part of the geostrophic shear. This confirms that the baroclinic vortex significantly modifies the convection.

The horizontal mean PV is close to zero throughout the bulk of the domain, as expected for slantwise convection (Fig. 7.7b). There is significant cancellation between a positive $\overline{b}_z$ and the remaining terms in (7.24), showing the influence of the baroclinic vortex. This cancellation renders the total PV profile very similar to that of the narrow domain, where the stratification is much weaker. This similarity in $\overline{q}$ suggests that the flows in both the wide and narrow domains are equally convecting and restoring to a convectively neutral state $\overline{q} = 0$. The baroclinic mode does not manage to restratify in the
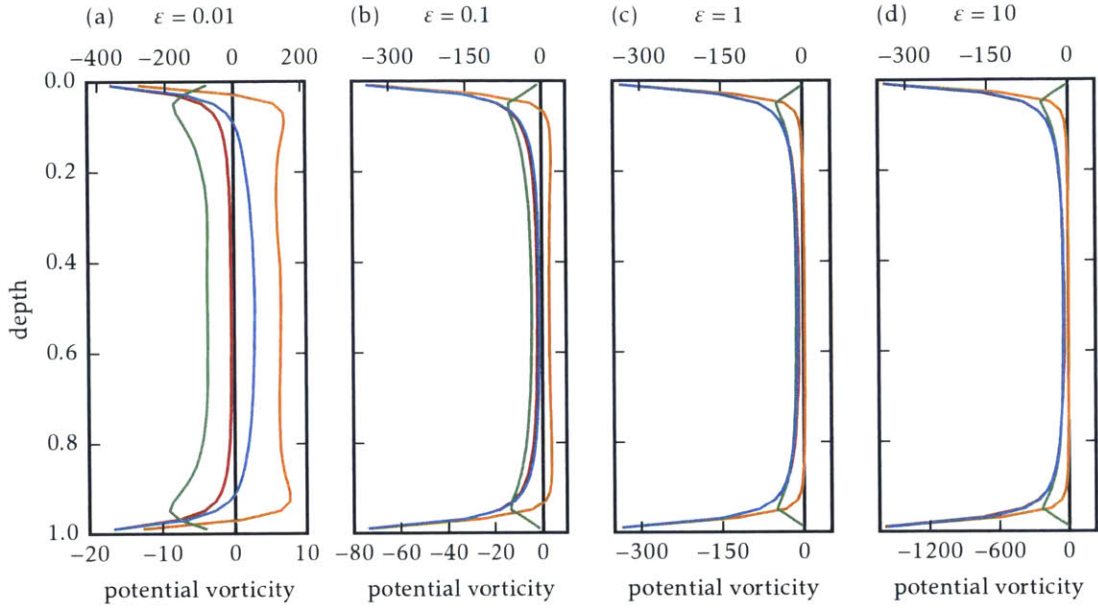
Figure 7.7: Horizontally averaged PV profiles for wide (——) and narrow (——) domains and for different forcing strengths as indicated in the titles. The profiles for the wide domains are split into contributions from b_z (——) and the remaining terms (——). The lower PV axis is given in the standard nondimensionalization (7.14–7.15), the upper in the alternative nondimensionalization (7.25–7.26). All profiles are averaged in time over the equilibrated range $100 < t < 200$.

sense of increasing the bulk PV above zero and thereby shutting off (slantwise) convection, but it does provide additional shear that modifies the convection and allows some buoyancy stratification.

Another piece of evidence for slantwise convection influenced by the baroclinic vortex comes from the energy budget (Fig. 7.8). The bulk of eddy kinetic energy is generated by buoyancy production, but there is a significant contribution from shear production. This is expected for slantwise convection, which can draw on the kinetic energy of the mean flow (Haine and Marshall, 1998; Thomas et al., 2013). The mean kinetic energy is supplied by gravitational slumping that attempts to restore geostrophic balance, out of which the mean flow is pushed by the momentum fluxes.

7.4.4 Strong forcing

As the forcing is increased to $\varepsilon = 1$, much larger than $B = 0.08$, convection is expected to dominate the flow. We find, however, that the strong convective flows are still modified
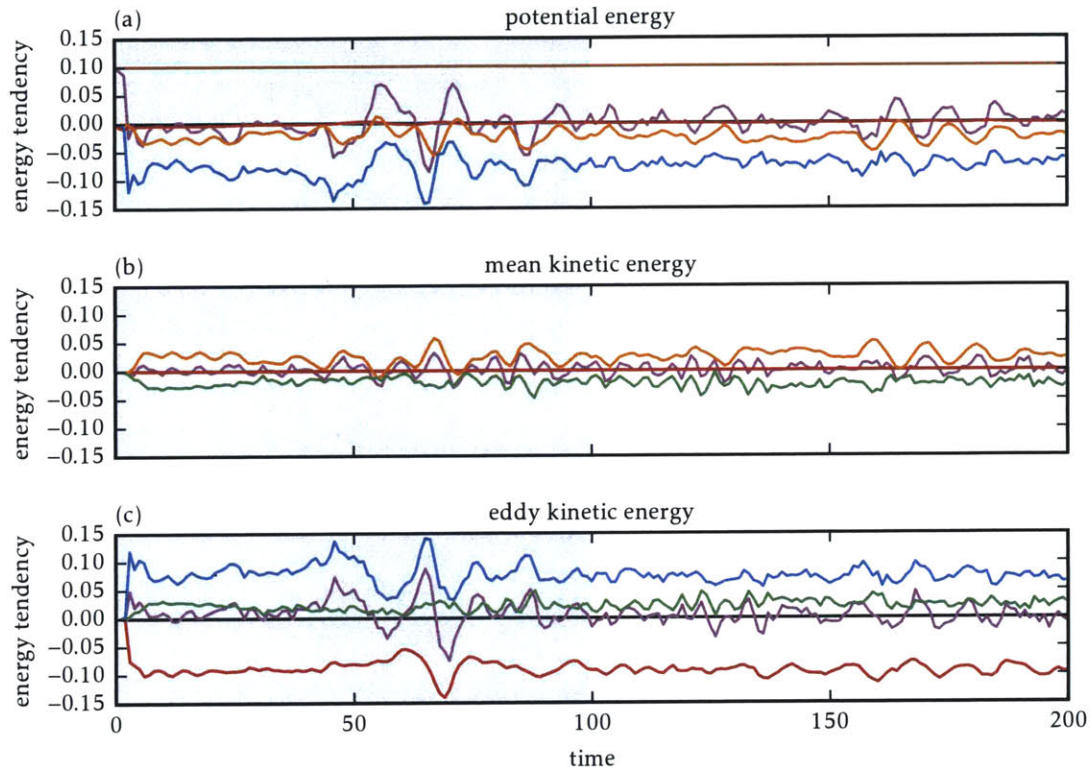
Figure 7.8: Time evolution of terms in the (a) potential, (b) mean kinetic, and (c) eddy kinetic energy budget for the moderately forced simulation ($\varepsilon = 0.1$): forcing (——), buoyancy production (——), gravitational slumping (——), shear production (——), diffusion/dissipation (——), and tendency (——)

detectably by baroclinic flows (Fig. 7.4c). Baroclinic instabilities develop on top of the strong convection, increase the shear at fronts, and render the convection more slantwise.

The buoyancy stratification shows this effect of the baroclinic instability very clearly (Fig. 7.5c). The stratification that results in this interplay of the baroclinic flow with convection is increased above the weak stratification resulting from pure convection in the narrow domain. As expected from $\varepsilon \gg B$, the baroclinic instability is not able to achieve true restratification, in the sense of increasing the mean PV to above zero (Fig. 7.7c). In fact, the PV profiles for the wide and narrow domains are indistinguishable—both are restored to zero by strong convection.

The eddy kinetic energy production is now strongly dominated by buoyancy production, but a significant contributing from shear production is still detectable (Fig. 7.9). Slantwise convection, while weak, thus still manifests itself in the energy budget.
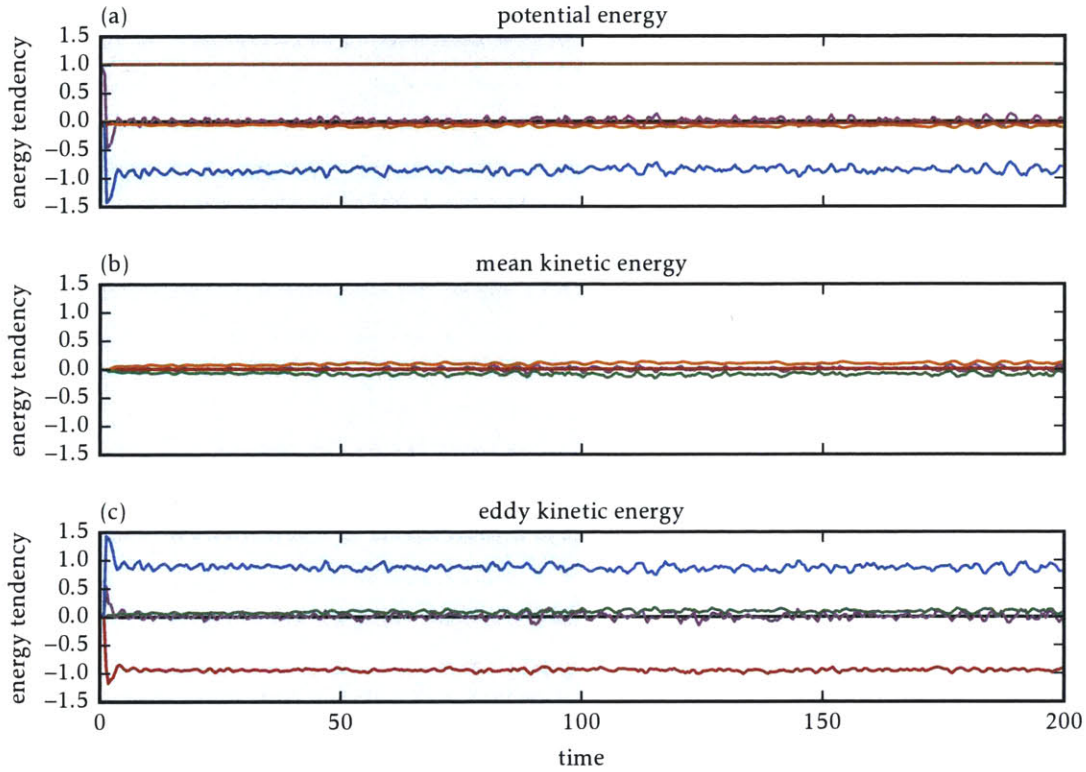
166

Figure 7.9: Time evolution of terms in the (a) potential, (b) mean kinetic, and (c) eddy kinetic energy budget for the strongly forced simulation ($\varepsilon = 1$): forcing (——), buoyancy production (——), gravitational slumping (——), shear production (——), diffusion/dissipation (——), and tendency (——)

7.4.5 Very strong forcing

At a very strong forcing of $\varepsilon = 10$, the flow is finally in the limit of no rotation and no shear and dominated by strong upright convection (Fig. 7.4d). The effect of the baroclinic instability is undetectable and the characteristics of the flow in the wide domain indistinguishable from that in the narrow domain.

The buoyancy profile is now unstratified over the bulk of the domain (Fig. 7.5d). The profiles are indistinguishable between the wide and narrow domains. The profiles are also very similar to that found in the narrow domain with $\varepsilon = 1$ if the comparison in this strongly convective limit is done in the alternative nondimensionalization (7.25–7.26) given in the upper buoyancy axis (Fig. 7.5c). The PV profiles are also indistinguishable between the narrow and wide domains as well as with the narrow domain with $\varepsilon = 1$ (Fig. 7.7c,d).
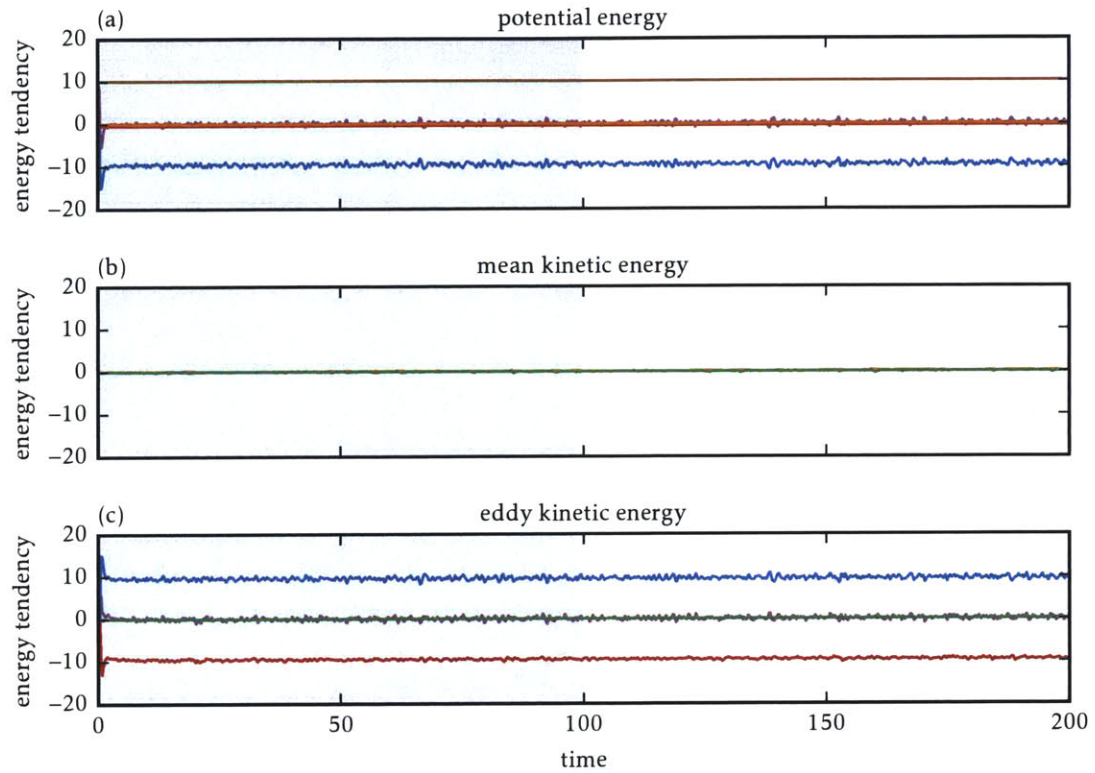
167

Figure 7.10: Time evolution of terms in the (a) potential, (b) mean kinetic, and (c) eddy kinetic energy budget for the very strongly forced simulation ($\varepsilon = 10$): forcing (——), buoyancy production (——), gravitational slumping (——), shear production (——), diffusion/dissipation (——), and tendency (——)

The energy budget is now a straightforward conversion of potential energy into eddy kinetic energy by buoyancy production, balanced by dissipation of eddy kinetic energy, as expected for pure upright convection (Fig. 7.10). There is no significant contribution from shear production.

7.5 Discussion

Our simulations suggest that baroclinic instabilities are remarkably resilient to the presence of convection. When the forcing is weak, the instability's buoyancy flux is larger than the imposed flux and restratification occurs. Positive PV is injected at the boundaries and the bulk PV of the fluid is increased above zero, shutting off convection. If the forcing is stronger, exceeding the buoyancy flux that can be generated by the baroclinic instabil-

ity, the instability does not increase the bulk PV to above zero and convection persists. The convection is modified by the baroclinic instability, however, as shears are intensified along fronts, where convection becomes more slantwise. For very strong forcing, this effect is small and convection becomes indistinguishable from the limit of upright convection. This limit appears to be approached gradually, with no qualitative transition at which the baroclinic mode is shut off.

These results were obtained for a background flow with moderate shear, $\delta = f/\Lambda = 1$. We expect the results to carry over to the hydrostatic limit $\delta \to 0$. As shown by Stone (1971) and in Fig. 7.1, the linear dynamics are similar for $\delta = 1$ and $\delta = 0$. In the presence of convection, we expect the baroclinic instability to survive in the $\delta \to 0$ regime just as well as in the $\delta = 1$ case analyzed here. The increasing scale separation between the instability and upright convection as $\delta \to 0$, which makes the concurrent simulation in this limit so challenging, suggests that strong energy transfer between the processes becomes more difficult. Future work should test this speculation.

Non-hydrostatic effects become important in the linear dynamics for $\delta \gg 1$ (Stone, 1971). We do not consider this limit, because baroclinic instability unlikely plays any significant role in the dynamics when $\delta \gg 1$ occurs in the ocean. The potential energy available for release by baroclinic eddies is small and the dynamics are more likely dominated by upright or rotating convection (e.g. Julien et al., 1996; Marshall and Schott, 1999).

In order to parameterize the restratification due to baroclinic mixed layer instabilities in coarse-resolution ocean models, it is necessary to understand under what atmospheric forcing conditions such restratification occurs. Our analysis shows that the magnitude of the baroclinic flux compared to the forcing distinguishes between convecting and restratifying conditions. As shown in Chapter 8, however, the baroclinic flux depends on the eddy scale, for which we have no good prediction. Understanding this dependence is crucial for the parameterization effort, because it impacts both whether or not restratification occurs and how strong it is when it does occurs.

The finding that baroclinic instabilities can develop even in the presence of strong convection can help us narrow down the mechanism responsible for the seasonal cycle of submesoscale turbulence. The persistence of baroclinic instabilities may explain how they can develop and produce energetic submesoscale flows at order 10 km in winter, despite the strong atmospheric forcing. The coexistence of baroclinic and convective flows in moderate to strong forcing conditions is consistent with the presence of baroclinic instabilities and the lack of restratification of deep winter mixed layers. That baroclinic instabilities can grow in the presence of convection also means, however, that they are unlikely to be damped out in summer, even if mixing time scales are short in shallow mixed layers.

The lack of energization of order 10 km flows in summer must then be explained by a lack of energy transfer to these scales from the small instability scale. A possible scenario is that the small amount of potential energy available for release in shallow summer mixed layers is insufficient to significantly energize order 10 km flows. Somewhat

counter-intuitively, baroclinic instabilities may be more successful in growing and restratifying the mixed layer in summer than in winter—but they quickly exhaust the energy fueling their growth. The energy input from summertime atmospheric forcing is weak, leaving the energy throughput too feeble for the instability to transfer significant energy into order 10 km flows. In winter, in contrast, strong atmospheric forcing produces copious amounts of potential energy available for release by baroclinic instabilities, leading to the strong submesoscale turbulence observed. This scenario should be made quantitative and its consistency tested in a setup that allows for the evolution of the mixed layer depth.

CHAPTER 8

SCALING OF THE BAROCLINIC EDDY FLUX IN BROAD AND NARROW FRONTS

8.1 The Fox-Kemper et al. (2008) scaling

Baroclinic instabilities in the ocean mixed layer are important in restoring stratification after atmospherically forced mixing events (e.g. Haine and Marshall, 1998; Boccaletti et al., 2007). Baroclinic mixed layer eddies are also thought to effect important exchanges of heat, carbon, nutrients, and other tracers between the surface and interior ocean (e.g. Capet et al., 2008a; Lévy et al., 2010, 2012; Mahadevan, 2014).

Mixed layer baroclinic instabilities slide dense waters under light waters, a process that flattens mean isopycnals. The increase in stratification is achieved by a positive eddy buoyancy flux $\overline{w'b'}$, where w' and b' are the vertical velocity and buoyancy anomalies associated with the instability. The overline is an appropriately defined spatial average and will denote a domain average in what follows.

Since baroclinic mixed layer instabilities occur on small horizontal scales of order 0.1–10 km, they are not typically resolved by current global ocean models. The eddies' effects on stratification and transport must be parameterized. Fox-Kemper et al. (2008) proposed a scaling for the eddy buoyancy flux, which can be used to mimic the slumping of baroclinic mixed layer fronts in coarse ocean models. They advance physical arguments that lead to

$$\overline{w'b'} \sim f\Lambda^2 H^2, \tag{8.1}$$

where f is the Coriolis parameter, Λ is the geostrophic shear associated with the lateral buoyancy gradient of the mixed layer front, and H is the mixed layer depth. A parameterization based on this scaling was implemented in global ocean models and shown to improve upper-ocean properties (Fox-Kemper et al., 2011; Gent et al., 2011).

Fox-Kemper et al. (2008) tested this scaling in a suite of spin-down experiments. Their mixed layer front is initially broader than the baroclinic eddies generated by the instability. The eddies grow in scale, reach the frontal scale, and eventually broaden the front. This transition from a broad frontal zone to a front that has the same scale as the eddies complicates the analysis. Considering the broad- and narrow-front regimes separately, we here uncover a dependence of the baroclinic flux $\overline{w'b'}$ on the characteristics of the eddies themselves, a dependence that is not captured by the scaling (8.1).

8.2 Broad front

In a frontal zone that is much broader than the eddies generated by the baroclinic instability, there is no horizontal eddy flux divergence, so the mean lateral buoyancy gradient remains unchanged. The parameters in the scaling (8.1) therefore remain fixed and the baroclinic flux is predicted to be constant.

This prediction is tested in a set of doubly-periodic simulations, in which the lateral buoyancy gradient is imposed and remains fixed (e.g. Taylor and Ferrari, 2010). The im-

posed background meridional buoyancy gradient is $-f\Lambda$ and the associated geostrophic zonal shear Λ. The Boussinesq equations for perturbations from this background flow are

$$u_t + \Lambda z u_x + \Lambda w + \boldsymbol{u} \cdot \nabla u - fv = -\phi_x + \mathcal{D}u \tag{8.2}$$

$$v_t + \Lambda z v_x + \boldsymbol{u} \cdot \nabla v + fu = -\phi_y + \mathcal{D}v \tag{8.3}$$

$$w_t + \Lambda z w_x + \boldsymbol{u} \cdot \nabla w = b - \phi_z + \mathcal{D}w \tag{8.4}$$

$$b_t + \Lambda z b_x - f\Lambda v + \boldsymbol{u} \cdot \nabla b = \mathcal{D}b \tag{8.5}$$

$$\nabla \cdot \boldsymbol{u} = 0 \tag{8.6}$$

where u, v, and w are the zonal, meridional, and vertical velocities, ϕ is the density-normalized pressure, and $\mathcal{D}A = \kappa_h \nabla_h^2 A + \kappa_v A_{zz}$ defines the diffusion operator. The boundary conditions are no normal flow ($w = 0$), no stress on the perturbations ($u_z = v_z = 0$), and no buoyancy flux ($b_z = 0$) at $z = 0$ and $z = -H$. For simplicity, we replace the strongly stratified thermocline by a solid bottom boundary. As long as the thermocline stratification is strong enough, this is a good approximation (cf. Garner et al., 1992). All perturbations are doubly periodic on a square domain of zonal and meridional extent L.

The meridional buoyancy gradient $-f\Lambda$ is constant, so the scaling (8.1) involves only prescribed parameters. Using the MITgcm (Marshall et al., 1997), we perform a suite of experiments with a fixed set of these parameters—we only vary the domain size. The scaling (8.1) predicts no dependence on the domain size, as long as the instability is well-resolved.

We choose $f = \Lambda = 10^{-4}$ s^{-1}, $H = 40$ m, and initialize the flow with $b_z = \Lambda^2$ and zero perturbation velocities. This state is stable to symmetric instability, so the most unstable mode is baroclinic (Stone, 1966b). A numerical linear stability analysis yields a wavelength of 220 m for the most unstable mode (Stone, 1970, 1971). The instability is kicked off by small random initial perturbations in b of magnitude 10^{-7} m s^{-2}. The horizontal and vertical viscosities/diffusivities are $\kappa_h = 10^{-3}$ m^2 s^{-1} and $\kappa_v = 2 \times 10^{-4}$ m^2 s^{-1}. The domain sizes and time steps are $L = 200$ m with $\Delta t = 100$ s, $L = 400$ m with $\Delta t = 50$ s, $L = 800$ m with $\Delta t = 25$ s, and $L = 1600$ m with $\Delta t = 12.5$ s. The grid spacing is $\Delta x = \Delta y = \Delta z = 2$ m.

In all simulations, the instability develops first at the scale of the most unstable mode. Larger eddies are subsequently energized until they reach the size of the domain (Fig. 8.1). This increase in eddy scale is quantified by diagnosing the dominant wavelength λ_0:

$$\frac{2\pi}{\lambda_0} = \frac{\int\int k_h S(k,l)\,dk\,dl}{\int\int S(k,l)\,dk\,dl}, \tag{8.7}$$

where k and l are the zonal and meridional wavenumbers, S is the buoyancy spectrum, and the integration is over the entire wavenumber space. The dominant eddy scale increases and then saturates at a scale that depends on the domain size (Fig. 8.2a). The increase in eddy scale is expected for a number of reasons: larger-scale modes reach finite
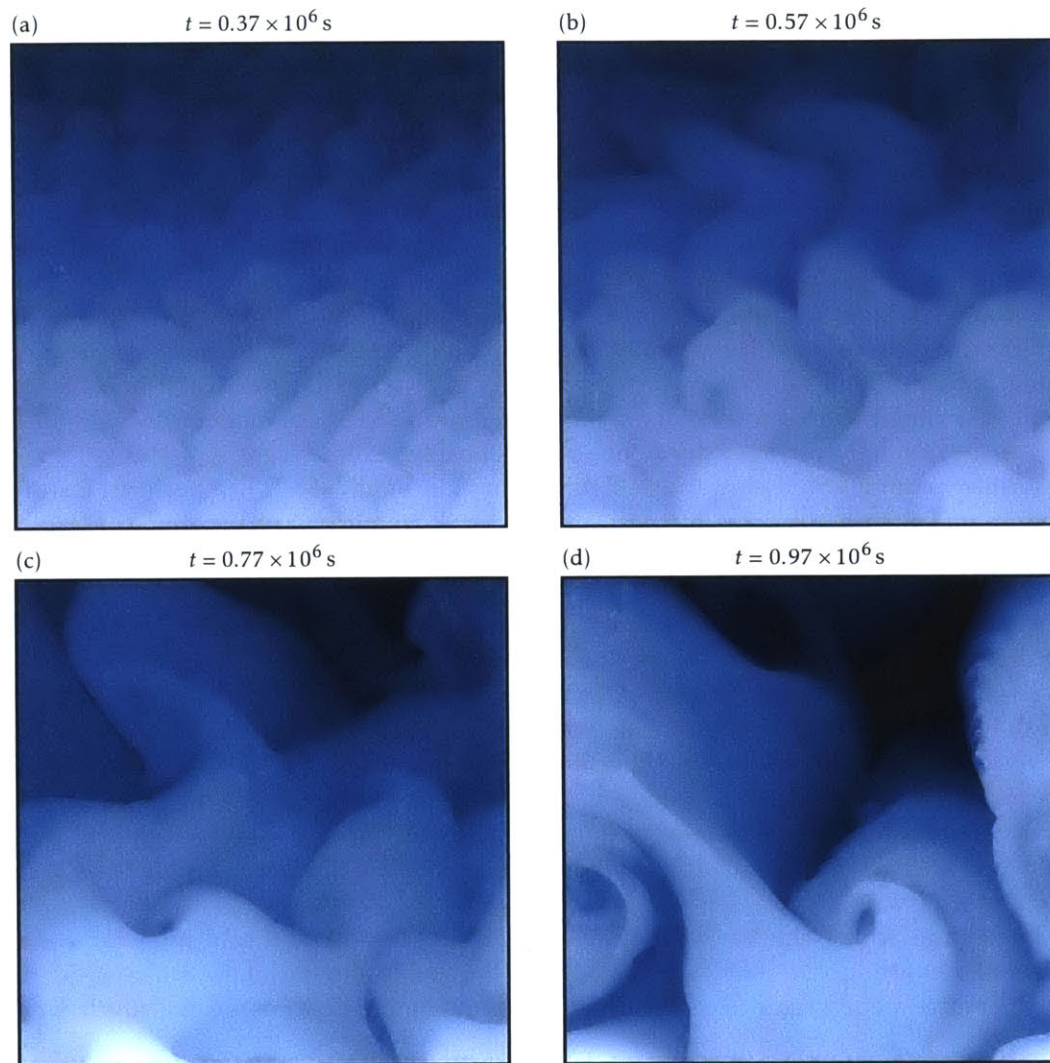
175

(a) $t = 0.37 \times 10^6 \, \text{s}$ (b) $t = 0.57 \times 10^6 \, \text{s}$

(c) $t = 0.77 \times 10^6 \, \text{s}$ (d) $t = 0.97 \times 10^6 \, \text{s}$

Figure 8.1: Snapshots of surface buoyancy $b - f \Lambda y$ showing the evolution of the broad front in the 1600-m domain. The color scale ranges between $\pm f \Lambda L$ from white through blue to black.

amplitude later as they have smaller growth rates, smaller-scale modes become stable as restratification occurs and the cutoff scale for the instability increases, and larger eddies can be energized through a nonlinear transfer of energy across scales.

In all simulations, the flux eventually reduces to zero. This is due to the increased stratification that shuts off baroclinic instability by pushing the short-wave cutoff beyond the domain scale and no unstable mode fits into the domain anymore. The domain size thus imposes a limit on the size of the baroclinic eddies.

Irrespective of the process that energizes larger eddies in the growth phase, the scaling (8.1) predicts the baroclinic flux to remain constant once the eddies have reached finite amplitude. Specifically, the Fox-Kemper et al. (2008) parameterization predicts

$$\overline{w'b'} = 0.06 f \Lambda^2 H^2 \overline{\mu(z)} = 0.04 f \Lambda^2 H^2, \tag{8.8}$$

given the vertical structure

$$\mu(z) = \left[1 - \left(\frac{2z}{H} + 1\right)^2\right]\left[1 + \frac{5}{12}\left(\frac{2z}{H} + 1\right)^2\right]. \tag{8.9}$$

The overline denotes a volume average over the domain.

The simulations are inconsistent with this prediction. While (8.8) captures the magnitude of the baroclinic flux initially, there is a significant increase beyond this value as the eddies grow in scale—if their growth is not limited by the domain size (Fig. 8.2b). The increase of the flux is intermittent and occurs in bursts associated with the growth at discrete low wavenumbers, but a clear pattern stands out: growth of the eddy scale is followed by an increase in baroclinic flux. The maximum flux is achieved when eddies have reached the domain scale. As a consequence, there is an increase in maximum baroclinic flux as the domain size is increased. The increase above the prediction (8.8) is largest for the largest domain.

The increase in eddy scale goes hand in hand with an increase in eddy velocities (Fig. 8.2c). The root mean square meridional eddy velocity increases beyond the maximum velocity of the mean flow ΛH, which violates the assumption $v \sim \Lambda H$ made in the derivation of (8.1). This was also noted by Bachman and Fox-Kemper (2013), who showed that the baroclinic flux can be better captured if the eddy velocity scale is known.

A baroclinic flux that increases in time leads to accelerating restratification, which is clearly visible in the largest domain (Fig. 8.2d). This acceleration is not captured by the scaling (8.1) and the Fox-Kemper et al. (2008) parameterization, which predicts a linear increase in stratification:

$$\overline{b_{zt}} = -0.06 f \Lambda^2 H^2 \overline{\mu_{zz}(z)} = 0.6 f \Lambda^2. \tag{8.10}$$

The disparity in the restratification rate is largest when eddies have grown largest. For even larger domain sizes, the stratification rate would surpass the prediction (8.10) even more dramatically.
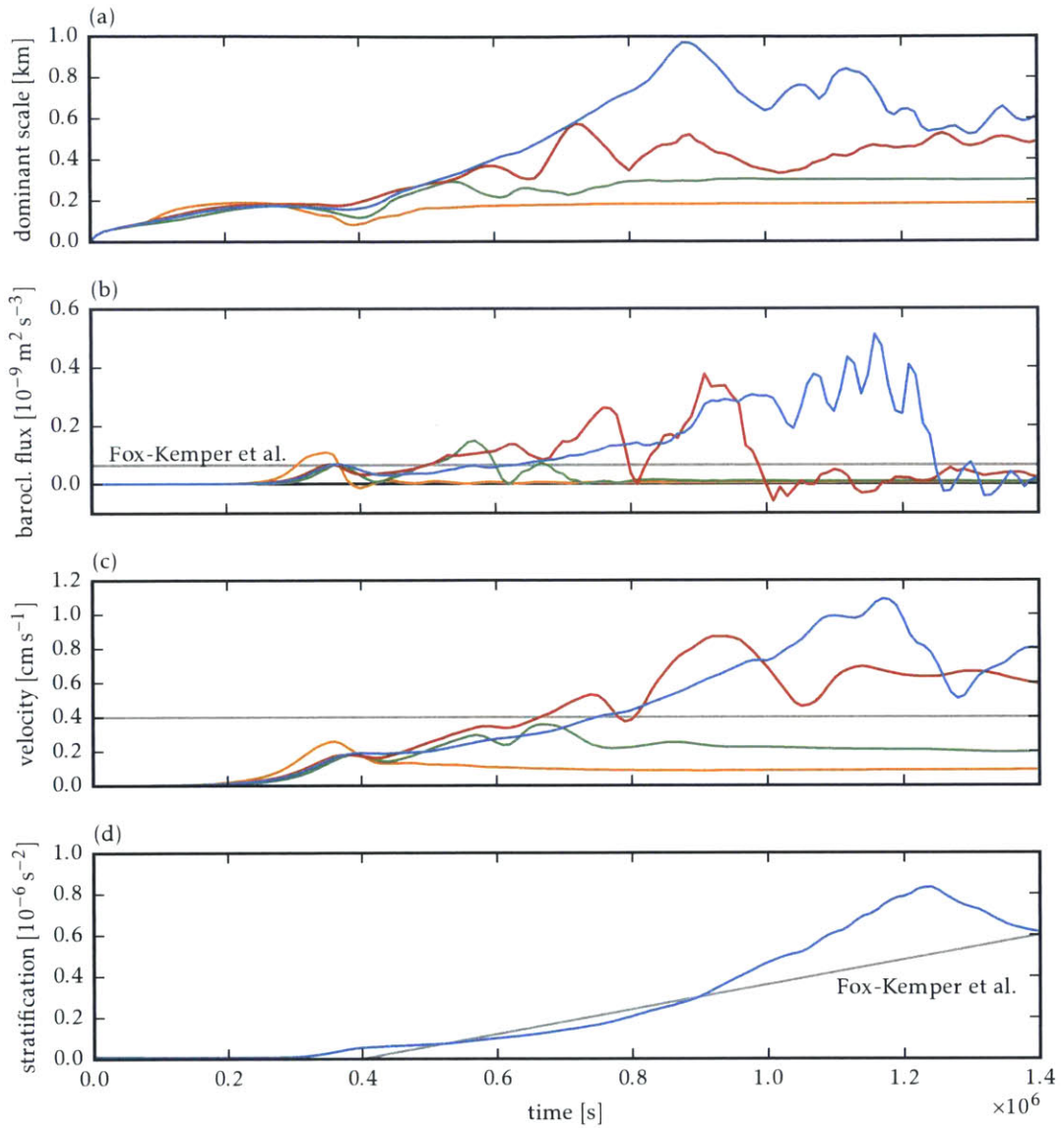
177

Figure 8.2: Time series of (a) the baroclinic flux $\overline{w'b'}$, (b) the dominant wavelength λ_0, (c) root mean square meridional velocity, and (d) average stratification $\overline{b}_z$ for the broad front simulations with the domain sizes $L = 1600$ m (——), $L = 800$ m (——), $L = 400$ m (——), and $L = 200$ m (——). Also shown in (b) and (d) are the predictions of Fox-Kemper et al. (2008) given by (8.8) and (8.10) with arbitrary intercept (——). Also shown in (c) is the maximum mean zonal velocity ΛH (——).

178

The domain size dependence is artificial, but it reveals a dependence of the eddy flux on eddy characteristics. The situation in the real ocean is more complicated, because mixed layer eddies grow in the presence of a turbulent mesoscale eddy field. What sets the relevant eddy scale and kinetic energy in this case is an open question.

Fox-Kemper and Ferrari (2008) test the Fox-Kemper et al. (2008) parameterization by employing it in a two-dimensional simulation. They compare the parameterized evolution to a full three-dimensional setup with explicit baroclinic eddies. They show reasonable correspondence for the time interval considered, over most of which the eddies are smaller than the frontal zone. But there are discrepancies that can be traced back to the increase of the baroclinic flux with eddy scale. The full simulation exhibits an increase in the overturning stream function not captured by the parameterized version, though considerable time variability may obscure this increase in snapshots (Fig. 3 in Fox-Kemper and Ferrari, 2008). This causes accelerating restratification in the full simulation, whereas the parameterized version exhibits constant restratification (Fig. 5 in Fox-Kemper and Ferrari, 2008). The acceleration is not as dramatic as in our largest simulation, because the eddies soon reach the size of the frontal zone, at which point restratification is slowed down. This allows the linear fit (8.8) to work reasonably, but it is clear that it would have failed more dramatically if the front had been broader (cf. Fig. 8.2d).

8.3 Narrow front

The simulations presented in Fox-Kemper et al. (2008) pass from the broad-front to the narrow-front regime as the eddies grow in scale and reach the width of the frontal zone. The goal of this section is to test whether there are dynamical differences between these two regimes or whether there is an increase in baroclinic flux beyond the scaling (8.1) in the narrow front regime as well.

We perform an experiment with a front that is initialized to be narrower than the instability scale. We use the same doubly-periodic setup as above, but choose an initial buoyancy perturbation that cancels out the background buoyancy gradient $-f\Lambda$ in the bulk of the domain and confines buoyancy variations to a narrow front of width L_f:

$$b = N^2\left(z + \frac{H}{2}\right) + f\Lambda y - \frac{\Delta b}{2}\tanh\frac{2y}{L_f} \qquad (8.11)$$

The buoyancy jump across the front is $\Delta b = f\Lambda L$, such that the buoyancy anomalies at $y = \pm L/2$ very nearly vanish ($L_f \ll L$) and periodicity is ensured (Fig. 8.3). The initial zonal flow perturbation is in thermal wind balance:

$$u_z = -\Lambda + \frac{\Delta b}{fL_f}\mathrm{sech}^2\frac{2y}{L_f}. \qquad (8.12)$$

Both the lateral buoyancy gradient and the zonal flow are very weak outside the frontal zone. This setup behaves like one with walls at $y = \pm L/2$ as long as there is no flow there.
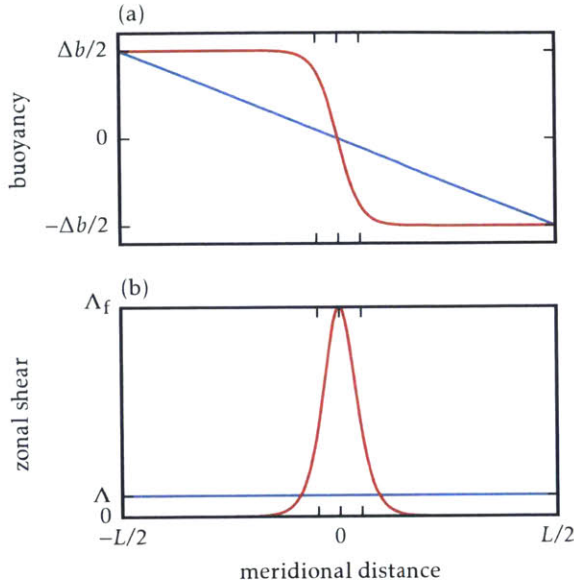
179

Figure 8.3: Illustration of the initial meridional profiles of (a) total buoyancy $-f\Lambda y + b$ at $z = -H/2$ and (b) total zonal shear $\Lambda + u_z$ for the broad-front (——) and narrow-front (——) setups. Note that the anomaly b vanishes in the broad-front case and is periodic in the narrow-front case. The unlabeled tick marks on the horizontal axis are at $\pm L_f/2$.

The initial frontal width is chosen to be $L_f = H = 40$ m. The domain size is $L = 3200$ m. We choose Δb such that the initial shear in the front is $\Lambda_f = \Delta b/f L_f = f = 10^{-4}$ s^{-1}, giving $\Delta b = 4 \times 10^{-7}$ m s^{-2} and $\Lambda = \Delta b/fL = 1.25 \times 10^{-6}$ s^{-1}. The initial stratification is chosen such that the maximum Richardson number in the front is unity: $N = 10^{-4}$ s^{-1}. The horizontal and vertical viscosities/diffusivities are $\kappa_h = 10^{-4}$ m^2 s^{-1} and $\kappa_v = 10^{-5}$ m^2 s^{-1}. The time step is $\Delta t = 100$ s and the grid spacing is again $\Delta x = \Delta y = \Delta z = 2$ m.

The initial instability occurs at a wavelength of about 150 m (Fig. 8.4). Barotropic instabilities are possible, but the phenomenology of the flow suggests that dominant instability is baroclinic in nature. The eddies subsequently increase in scale as they flatten the front. The eddies increase the frontal scale L_f, which is of the same order as the eddy scale (cf. Manucharyan and Timmermans, 2013). We stop the simulation when the eddies fill the domain and the periodicity of the setup would affect the further development of the front. The analysis is restricted to the period in which the eddies widen the front into undisturbed fluid.

We diagnose the domain-average baroclinic flux $\overline{w'b'}$. The flux increases as the instability grows and then remains roughly constant (Fig. 8.5a). As explained in the following, this again implies that the flux increases beyond what is predicted by the scaling (8.1) as the eddies grow in size.

The scaling (8.1) predicts

$$\overline{w'b'}^f \sim f\Lambda_f^2 H^2, \tag{8.13}$$

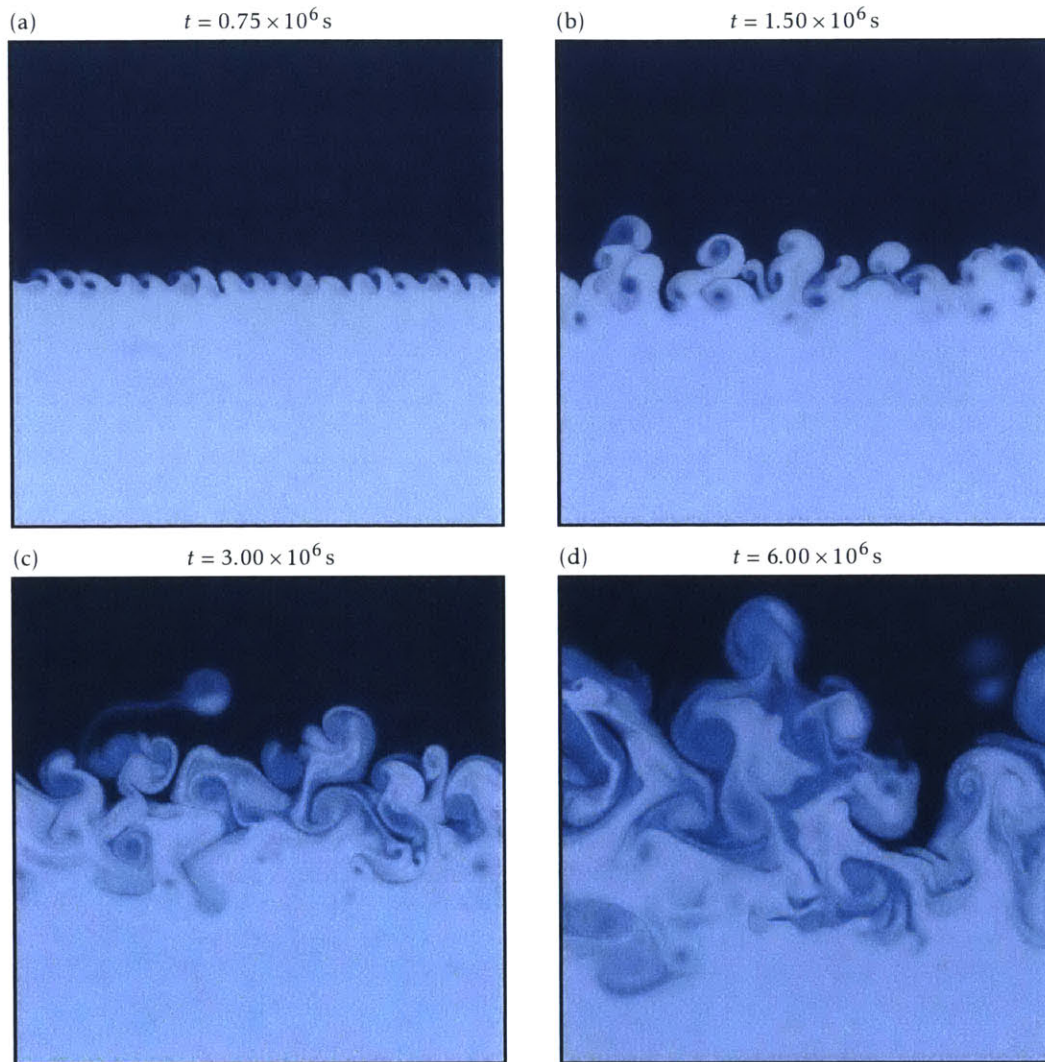where the overline with a superscript 'f' denotes a volume average restricted to the frontal

180

Figure 8.4: Snapshots of surface buoyancy $b - f\Lambda y$ showing the evolution of the initially narrow front. The color scale ranges between $\pm f\Lambda L$ from white through blue to black.
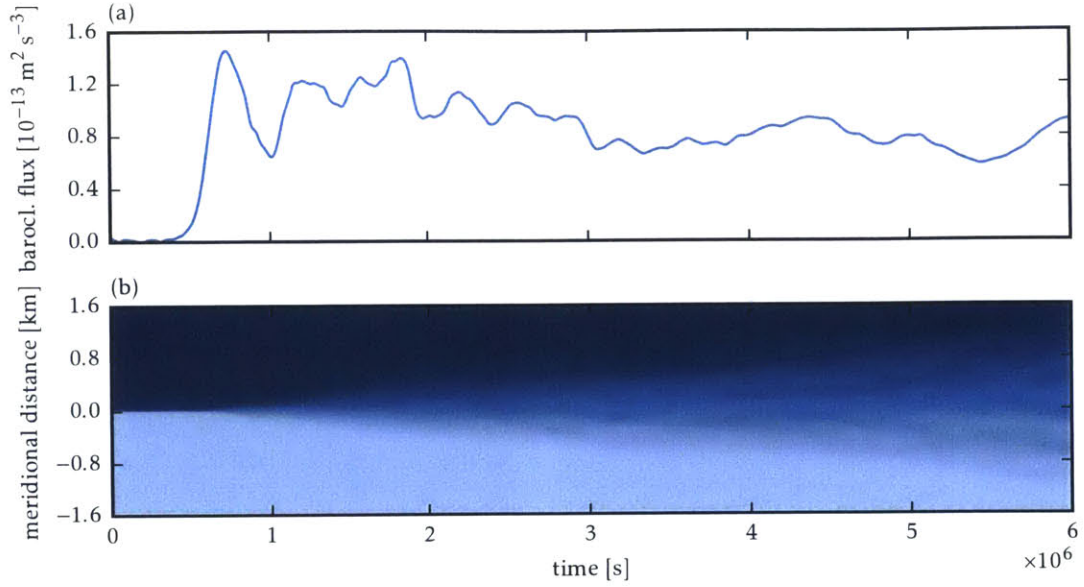
Figure 8.5: Time series of (a) the domain-averaged baroclinic flux $\overline{w'b'}$ and (b) the zonally and vertically averaged buoyancy for the initially narrow front. The color scale in (b) ranges between $\pm f \Lambda L$ from white through blue to black.

zone of width L_f, which increases as the flow evolves. The scaling employs the evolving meridional buoyancy gradient averaged vertically and over the frontal zone $-f \Lambda_f$. Since the buoyancy difference Δb across the front remains fixed, the buoyancy gradient in the frontal zone

$$f \Lambda_f \sim \frac{\Delta b}{L_f}. \tag{8.14}$$

decreases as the front broadens (Fig. 8.5b). The baroclinic flux is confined to the frontal zone, so the domain-average flux is roughly

$$\overline{w'b'} \sim \frac{L_f}{L} \overline{w'b'}^f. \tag{8.15}$$

Together with the scaling (8.13), this yields the prediction

$$\overline{w'b'} \sim \frac{\Delta b^2 H^2}{f L_f L}. \tag{8.16}$$

The domain-average baroclinic flux is thus predicted to decrease like L_f^{-1} as the front broadens. In our simulation, the frontal width increases by over an order of magnitude from the initial instability scale to the domain scale (Fig. 8.5b). The diagnosed baroclinic

182

flux, remaining roughly constant (Fig. 8.5a), thus increasingly exceeds this prediction as the eddies grow larger and the front wider. This increase of the baroclinic flux beyond the prediction (8.1) is qualitatively consistent with that found in the broad-front regime.[1]

8.4 Concluding remarks

Our analysis suggests that the baroclinic flux $\overline{w'b'}$ increases with the eddy scale or equivalently with the eddy kinetic energy, which is not captured by the scaling (8.1). The simulations show an increase of the baroclinic flux beyond the scaling, but what exactly the additional dependence is and how it can be understood physically should be studied in more detail in the future.

The dependence of the buoyancy flux on eddy characteristics implies that a parameterization of mixed layer baroclinic instabilities requires an additional input besides the mixed layer depth and the lateral buoyancy gradient. What sets the relevant eddy characteristics and how they can be determined from large-scale quantities remain open questions. It is likely that nonlinear dynamics are essential (cf. Larichev and Held, 1995; Fox-Kemper et al., 2008). Furthermore, future work will need to investigate whether and how the evolution and restratification is affected by the presence of the thermocline, which is expected to become increasingly important as mixed layer eddies grow larger (cf. Chapter 6; Ramachandran et al., 2014).

Including the restratification by mixed layer baroclinic instabilities in global ocean models has already improved upper-ocean properties. It can be hoped that developing an understanding for the additional dependence on eddy characteristics will improve the parameterization of baroclinic restratification and help further decrease biases in global ocean models.

---

[1] The slight decrease of $\overline{w'b'}$ exhibited by Fig. 8.5a is, if statistically significant, much smaller than the decrease by over an order of magnitude predicted by (8.16).

Chapter 9

Conclusions and Outlook

Submesoscale turbulence has recently come into focus as an important component of upper-ocean dynamics. This dissertation contributes to deciphering what process generates surface fronts and what dynamics determine the characteristics of submesoscale flows.

Observations suggest that energetic submesoscale flows arise from baroclinic mixed layer instabilities, which release large amounts of available potential energy in winter, when mixed layers are deep. The observed spectral and vertical distribution of energy in the near-surface layers of the wintertime western North Atlantic is consistent with a simple quasi-geostrophic model of submesoscale turbulence energized by baroclinic mixed layer instabilities. Outside the winter mixed layer, i.e. in the seasonal thermocline in summer and the main thermocline throughout the year, submesoscale turbulence is observed to be weak: the energy of balanced flows falls off quickly with wavenumber, consistent with interior quasi-geostrophic turbulence.

In winter, atmospheric forcing destabilizes the upper ocean, generating deep mixed layers. In conjunction with lateral buoyancy gradients, this provides copious amounts of potential energy available for release by baroclinic instabilities. These instabilities can develop even in the presence of strong convection. The restratification associated with the release of potential energy is countered by convective motions that keep the mixed layer unstratified. The baroclinic mixed layer instabilities create fronts at the surface as well as the base of the mixed layer and through turbulent scale interactions energize the submesoscale range. The nonlinear dynamics follow expectations from geostrophic turbulence theory: kinetic energy generated by the instabilities is transferred preferentially to large scales. In equilibrium, the dynamics can be described as a dual cascade of barotropic and baroclinic mixed layer modes, but it remains unclear what determines the submesoscale energy levels in such an equilibrium.

It is possible that submesoscale turbulence never reaches an equilibrium and must be viewed in the context of the seasonal cycle of the atmospheric forcing. It has recently been suggested that the kinetic energy produced by baroclinic mixed layer instabilities is a significant source for mesoscale eddies (Sasaki et al., 2014). In this scenario, the submesoscale energy input induces a mesoscale seasonal cycle, which is shifted in phase from the submesoscale cycle because it takes a few months for energy to percolate from the submesoscale instability to order 100 km scales. If this is the case, the submesoscale energy levels are determined by the transient energization by baroclinic mixed layer instabilities in winter, but the submesoscale energy budget is linked inseparably to the mesoscale equilibration problem. An extension of the analysis presented in this dissertation is required in order to understand mesoscale and submesoscale dynamics that evolve and interact on annual time scales.

A transient view of submesoscale dynamics is necessary at least for the summer, as suggested by the weakness of submesoscale flows then. If the mixed layer depth were fixed in time, baroclinic instabilities would release infinite amounts of potential energy, because they can grow even in the presence of fast convection. The instabilities would

eventually energize the submesoscale range also in summer, despite the mixed layer's shallowness. In reality, the instabilities instead exhaust their energy source by restratifying the mixed layer and submesoscale flows remain weak. This scenario should be cast in more quantitative terms, but it suggests taking a closer look at the energy budget of submesoscale flows and linking their energization to the supply of energy by surface forcing.

A related question is what sets the length and energy scales of the baroclinic eddies dominating mixed layer restratification. The rate of restratification depends on these eddy characteristics, so the parameterization of baroclinic mixed layer instabilities in coarse-resolution ocean models should take them into account. It seems likely that nonlinear dynamics play an important role in setting these characteristics, and a comprehensive energy budget of submesoscale flows may provide an avenue to progress in this regard as well.

The energization of submesoscale flows by baroclinic mixed layer instabilities is quite different in character from mesoscale-driven surface frontogenesis, as described by surface quasi-geostrophic dynamics. These dynamics have previously been proposed to be an important energization mechanism, but their fingerprints are not found in the observations analyzed here. Surface quasi-geostrophic dynamics do not predict the seasonal cycle found in the western North Atlantic and the observed distribution of energy across scales and depth does not match what is predicted.

It may be expected that mesoscale-driven surface frontogenesis is more relevant for regions that have interior potential vorticity gradients weaker than those in the western North Atlantic. Surface buoyancy gradients then play a more important role in the dynamics. A problem faced in such regions, however, is that if mesoscale eddies are relatively weak, inertia–gravity waves can dominate and mask balanced flows in the submesoscale range, as in the East Pacific observations analyzed here. The balanced flows, even though they are less energetic than inertia–gravity waves, may still be important in the transfer of tracers between the surface and interior ocean. In such situations one may be able to learn about balanced submesoscale flows from the distribution of tracers stirred by these motions. Our understanding of the stirring by balanced flows, inertia–gravity waves, and combinations thereof, however, appears to be incomplete. It remains unexplained why there is much less submesoscale salinity and temperature variance along isopycnals than expected from stirring by balanced flows alone (Cole and Rudnick, 2012).

A global view of submesoscale flows is anticipated to emerge from the upcoming Surface Water and Ocean Topography (SWOT) mission. The wide swath altimeter is expected to measure sea surface height to much higher accuracy than the nadir-looking altimeters currently employed. The use of geostrophic balance to infer surface currents, however, must be justified much more carefully in the submesoscale range. A major challenge will be to distinguish between sea surface height anomalies associated with balanced flows and those induced by inertia–gravity waves. The repeat cycle will be too long to identify inertia–gravity waves based on their high frequencies. The method we developed to distinguish between geostrophic flows and inertia–gravity waves in one-dimensional ob-

187

servations of currents is not applicable to altimetry data, because only sea surface height is measured, rather than two flow components. But progress may be possible by taking advantage of the two-dimensional sea surface height field measured by SWOT as well as the sea surface temperature distribution concurrently measured by other satellites.

The focus of this dissertation is on the local dynamics of submesoscale turbulence. While many questions remain, the impact of submesoscale flows on the global ocean circulation is even less clear. Modeling studies are beginning to probe in this direction, but the mechanisms of interaction between submesoscales and the gyres, the Antarctic Circumpolar Current, and the meridional overturning circulation are largely unexplored. There is ample room for future inquiry and surprises are inevitably waiting.

BIBLIOGRAPHY

Arbic, B. K., J. G. Richman, J. F. Shriver, P. G. Timko, E. J. Metzger, and A. J. Wallcraft (2012) Global modeling of internal tides within an eddying ocean general circulation model. *Oceanography* 25 (2), 20–29.

Armstrong, E. M., G. Wagner, J. Vazquez-Cuervo, and T. M. Chin (2012) Comparisons of regional satellite sea surface temperature gradients derived from MODIS and AVHRR sensors. *Int. J. Remote Sens.* 33 (21), 6639–6651.

Bachman, S. and B. Fox-Kemper (2013) Eddy parameterization challenge suite I: Eady spindown. *Ocean Model.* 64, 12–28.

Badin, G. (2012) Surface semi-geostrophic dynamics in the ocean. *Geophys. Astrophys. Fluid Dyn.* 107 (5), 526–540.

Batchelor, G. K. (1953) *The Theory of Homogeneous Turbulence*. London: Cambridge University Press, p. 197.

Bishop, C. H. (1993a) On the behaviour of baroclinic waves undergoing horizontal deformation. I: The 'RT' phase diagram. *Q. J. R. Meteorol. Soc.* 119 (510), 221–240.

Bishop, C. H. (1993b) On the behaviour of baroclinic waves undergoing horizontal deformation. II: Error-bound amplification and Rossby wave diagnostics. *Q. J. R. Meteorol. Soc.* 119 (510), 241–267.

Blumen, W. (1978) Uniform Potential Vorticity Flow: Part I. Theory of Wave Interactions and Two-Dimensional Turbulence. *J. Atmos. Sci.* 35 (5), 774–783.

Blumen, W. (1979) On Short-Wave Baroclinic Instability. *J. Atmos. Sci.* 36 (10), 1925–1933.

Boccaletti, G., R. Ferrari, and B. Fox-Kemper (2007) Mixed Layer Instabilities and Restratification. *J. Phys. Oceanogr.* 37 (9), 2228–2250.

Boyd, J. P. (1992) The Energy Spectrum of Fronts: Time Evolution of Shocks in Burgers' Equation. *J. Atmos. Sci.* 49 (2), 128–139.

Bretherton, F. P. and M. Karweit (1975) Mid-Ocean Mesoscale Modeling. In: *Numerical Models of Ocean Circulation*. Ed. by R. O. Reid, A. R. Robinson, and K. Bryan. Washington, D.C.: National Academy of Sciences, 237–249.

Bretherton, F. P. (1966) Critical layer instability in baroclinic flows. *Q. J. R. Meteorol. Soc.* 92 (393), 325–334.

Bühler, O., J. Callies, and R. Ferrari (2014) Wave–vortex decomposition of one-dimensional ship-track data. *J. Fluid Mech.* 756, 1007–1026.

Callies, J. and R. Ferrari (2013) Interpreting Energy and Tracer Spectra of Upper-Ocean Turbulence in the Submesoscale Range (1–200 km). *J. Phys. Oceanogr.* 43 (11), 2456–2474.

Callies, J., R. Ferrari, and O. Bühler (2014) Transition from geostrophic turbulence to inertia-gravity waves in the atmospheric energy spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 111 (48), 17033–17038.

Callies, J., R. Ferrari, J. M. Klymak, and J. Gula (2015) Seasonality in submesoscale turbulence. *Nat. Commun.* 6 (6862).

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin (2008a) Mesoscale to Submesoscale Transition in the California Current System. Part I: Flow Structure, Eddy Flux, and Observational Tests. *J. Phys. Oceanogr.* 38 (1), 29–43.

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin (2008b) Mesoscale to Submesoscale Transition in the California Current System. Part II: Frontal Processes. *J. Phys. Oceanogr.* 38 (1), 44–64.

Capet, X., J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin (2008c) Mesoscale to Submesoscale Transition in the California Current System. Part III: Energy Balance and Flux. *J. Phys. Oceanogr.* 38 (10), 2256–2269.

Capet, X., E. J. Campos, and A. M. Paiva (2008d) Submesoscale activity over the Argentinian shelf. *Geophys. Res. Lett.* 35 (15), L15605.

Capet, X., P. Klein, B. L. Hua, G. Lapeyre, and J. C. McWilliams (2008e) Surface kinetic energy transfer in surface quasi-geostrophic flows. *J. Fluid Mech.* 604, 165–174.

Charney, J. G., R. Fjörtoft, and F. von Neumann (1950) Numerical Integration of the Barotropic Vorticity Equation. *Tellus A* 2 (4), 237–254.

Charney, J. G. (1947) The Dynamics of Long Waves in a Baroclinic Westerly Current. *J. Meteorol.* 4 (5), 135–162.

Charney, J. G. (1949) On the physical basis for numerical prediction of large-scale motions in the atmosphere. *J. Meteorol.* 6 (6), 371–385.

Charney, J. G. (1971) Geostrophic Turbulence. *J. Atmos. Sci.* 28 (6), 1087–1095.

Charney, J. G. and A. Eliassen (1949) A Numerical Method for Predicting the Perturbations of the Middle Latitude Westerlies. *Tellus* 1 (2), 38–54.

Cho, J. Y. N. and E. Lindborg (2001) Horizontal velocity structure functions in the upper troposphere and lower stratosphere: 1. Observations. *J. Geophys. Res.* 106 (D10), 10223–10232.

Cole, S. T. and D. L. Rudnick (2012) The spatial distribution and annual cycle of upper ocean thermohaline structure. *J. Geophys. Res.* 117 (C2), C02027.

D'Asaro, E., C. Lee, L. Rainville, R. Harcourt, and L. Thomas (2011) Enhanced Turbulence and Energy Dissipation at Ocean Fronts. *Science* 332, 318–322.

D'Asaro, E. A. (1978) Mixed Layer Velocities Induced by Internal Waves. *J. Geophys. Res.* 83 (C5), 2437–2438.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109 (C12), C12003.

Dee, D. P. et al. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137 (656), 553–597.

Defant, A. (1961) *Physical Oceanography: Vol. II.* New York: Pergamon Press, p. 598.

Dewan, E. M. (1979) Stratospheric Wave Spectra Resembling Turbulence. *Science* 204 (4395), 832–835.

Dong, S., J. Sprintall, S. T. Gille, and L. Talley (2008) Southern Ocean mixed-layer depth from Argo float profiles. *J. Geophys. Res.* 113 (C6), C06013.

Dushaw, B. D., B. D. Cornuelle, P. F. Worcester, B. M. Howe, and D. S. Luther (1995) Barotropic and Baroclinic Tides in the Central North Pacific Ocean Determined from Long-Range Reciprocal Acoustic Transmissions. *J. Phys. Oceanogr.* 25 (4), 631–647.

Eady, E. T. (1949) Long Waves and Cyclone Waves. *Tellus* 1 (3), 33–52.

Emanuel, K. A. (1988) Observational Evidence of Slantwise Convective Adjustment. *Mon. Weather Rev.* 116 (9), 1805–1816.

Emanuel, K. A. (1994) *Atmospheric Convection.* Oxford University Press, p. 580.

Ferrari, R. (2011) A Frontal Challenge for Climate Models. *Science* 332 (6027), 316–317.

Ferrari, R. and D. L. Rudnick (2000) Thermohaline variability in the upper ocean. *J. Geophys. Res.* 105 (C7), 16857–16883.

Ferrari, R. and C. Wunsch (2009) Ocean Circulation Kinetic Energy: Reservoirs, Sources, and Sinks. *Annu. Rev. Fluid Mech.* 41 (1), 253–282.

Ferrari, R. and C. Wunsch (2010) The distribution of eddy kinetic and potential energies in the global ocean. *Tellus A* 62 (2), 92–108.

Fjørtoft, R. (1953) On the Changes in the Spectral Distribution of Kinetic Energy for Twodimensional, Nondivergent Flow. *Tellus* 5 (3), 225–230.

Fox-Kemper, B., G. Danabasoglu, R. Ferrari, S. Griffies, R. Hallberg, M. Holland, M. Maltrud, S. Peacock, and B. Samuels (2011) Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Model.* 39 (1-2), 61–78.

Fox-Kemper, B. and R. Ferrari (2008) Parameterization of Mixed Layer Eddies. Part II: Prognosis and Impact. *J. Phys. Oceanogr.* 38 (6), 1166–1179.

Fox-Kemper, B., R. Ferrari, and R. W. Hallberg (2008) Parameterization of Mixed Layer Eddies. Part I: Theory and Diagnosis. *J. Phys. Oceanogr.* 38 (6), 1145–1165.

Frankignoul, C. and P. Müller (1979) Quasi-Geostrophic Response of an Infinite β-Plane Ocean to Stochastic Forcing by the Atmosphere. *J. Phys. Oceanogr.* 9 (1), 104–127.

Fritts, D. C. and M. J. Alexander (2003) Gravity Wave Dynamics and Effects in the Middle Atmosphere. *Rev. Geophys.* 41 (1), 1003.

193

Fu, L.-L. and R. Ferrari (2008) Observing Oceanic Submesoscale Processes From Space. *Eos* 89 (48), 489–499.

Gage, K. S. (1979) Evidence for a $k^{-5/3}$ Law Inertial Range in Mesoscale Two-Dimensional Turbulence. *J. Atmos. Sci.* 36 (10), 1950–1954.

Gage, K. S. and G. D. Nastrom (1985) On the spectrum of atmospheric velocity fluctuations seen by MST/ST radar and their interpretation. *Radio Sci.* 20 (6), 1339–1347.

Garner, S. T., N. Nakamura, and I. M. Held (1992) Nonlinear Equilibration of Two-Dimensional Eady Waves: A New Perspective. *J. Atmos. Sci.* 49 (21), 1984–1996.

Garrett, C. (2001) What is the "Near-Inertial" Band and Why Is It Different from the Rest of the Internal Wave Spectrum? *J. Phys. Oceanogr.* 31 (4), 962–971.

Garrett, C. and W. H. Munk (1972) Space-Time scales of Internal Waves. *Geophys. Fluid Dyn.* 3 (1), 225–264.

Garrett, C. and W. H. Munk (1975) Space-Time Scales of Internal Waves: A Progress Report. *J. Geophys. Res.* 80 (3), 291–297.

Garrett, C. and W. H. Munk (1979) Internal waves in the ocean. *Annu. Rev. Fluid Mech.* 11, 339–369.

Gent, P. R. et al. (2011) The Community Climate System Model Version 4. *J. Clim.* 24 (19), 4973–4991.

Gill, A. E., J. S. A. Green, and A. J. Simmons (1974) Energy partition in the large-scale ocean circulation and the production of mid-ocean eddies. *Deep Sea Res.* 21 (7), 499–528.

Gula, J., M. J. Molemaker, and J. C. McWilliams (2015) Gulf Stream Dynamics along the Southeastern U.S. Seaboard. *J. Phys. Oceanogr.* 45 (3), 690–715.

Haidvogel, D. B. and I. M. Held (1980) Homogeneous Quasi-Geostrophic Turbulence Driven by a Uniform Temperature Gradient. *J. Atmos. Sci.* 37 (12), 2644–2660.

Haine, T. W. N. and J. Marshall (1998) Gravitational, Symmetric, and Baroclinic Instability of the Ocean Mixed Layer. *J. Phys. Oceanogr.* 28 (4), 634–658.

Hakim, G. J., C. Snyder, and D. J. Muraki (2002) A New Surface Model for Cyclone–Anticyclone Asymmetry. *J. Atmos. Sci.* 59 (16), 2405–2420.

Hamilton, K., Y. O. Takahashi, and W. Ohfuchi (2008) Mesoscale spectrum of atmospheric motions investigated in a very fine resolution global general circulation model. *J. Geophys. Res.* 113 (D18), D18110.

Hamlington, P. E., L. P. Van Roekel, B. Fox-Kemper, K. Julien, and G. P. Chini (2014) Langmuir–Submesoscale Interactions: Descriptive Analysis of Multiscale Frontal Spindown Simulations. *J. Phys. Oceanogr.* 44 (9), 2249–2272.

Haynes, P. H. and J. Vanneste (2004) Stratospheric Tracer Spectra. *J. Atmos. Sci.* 61 (2), 161–178.

Held, I. M., R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson (1995) Surface quasi-geostrophic dynamics. *J. Fluid Mech.* 282, 1–20.

Hertzog, A., F. Vial, C. R. Mechoso, C. Basdevant, and P. Cocquerez (2002) Quasi-Lagrangian measurements in the lower stratosphere reveal an energy peak associated with near-inertial waves. *Geophys. Res. Lett.* 29 (8), 70–74.

Hoskins, B. J. (1974) The role of potential vorticity in symmetric stability and instability. *Q. J. R. Meteorol. Soc.* 100 (425), 480–482.

Hoskins, B. J., I. Draghici, and H. C. Davies (1978) A new look at the ω-equation. *Q. J. R. Meteorol. Soc.* 104 (439), 31–38.

Hoskins, B. J. (1982) The Mathematical Theory of Frontogenesis. *Annu. Rev. Fluid Mech.* 14, 131–151.

Hoskins, B. J. and F. P. Bretherton (1972) Atmospheric Frontogenesis Models: Mathematical Formulation and Solution. *J. Atmos. Sci.* 29 (1), 11–37.

Johnson, G. C. and H. L. Bryden (1989) On the size of the Antarctic Circumpolar Current. *Deep Sea Res. Part A. Oceanogr. Res. Pap.* 36 (1), 39–53.

Juckes, M. (1994) Quasigeostrophic Dynamics of the Tropopause. *J. Atmos. Sci.* 51 (19), 2756–2768.

Julien, K., S. Legg, J. McWilliams, and J. Werne (1996) Rapidly rotating turbulent Rayleigh-Bénard convection. *J. Fluid Mech.* 322, 243–273.

Kantha, L. H. and C. A. Clayson (2000) *Small Scale Processes in Geophysical Fluid Flows.* San Diego: Academic Press, p. 750.

Katz, E. J. (1973) Profile of an Isopycnal Surface in the Main Thermocline of the Sargasso Sea. *J. Phys. Oceanogr.* 3 (4), 448–457.

Klein, P. and G. Lapeyre (2009) The Oceanic Vertical Pump Induced by Mesoscale and Submesoscale Turbulence. *Ann. Rev. Mar. Sci.* 1, 351–375.

Klein, P., B. L. Hua, G. Lapeyre, X. Capet, S. Le Gentil, and H. Sasaki (2008) Upper Ocean Turbulence from High-Resolution 3D Simulations. *J. Phys. Oceanogr.* 38 (8), 1748–1763.

Klymak, J. M. and J. N. Moum (2007) Oceanic Isopycnal Slope Spectra. Part I: Internal Waves. *J. Phys. Oceanogr.* 37 (5), 1215–1231.

Kolmogorov, A. N. (1941) The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers (in Russian). *Dokl. Akad. Nauk SSSR* 30, 301–305. Translated into English, *Proc. R. Soc. A* 434 (1890), 9–13.

Koshyk, J. N. and K. Hamilton (2001) The Horizontal Kinetic Energy Spectrum and Spectral Budget Simulated by a High-Resolution Troposphere–Stratosphere–Mesosphere GCM. *J. Atmos. Sci.* 58 (4), 329–348.

Koshyk, J. N., K. Hamilton, and J. D. Mahlman (1999) Simulation of the $k^{-5/3}$ mesoscale spectral regime in the GFDL SKYHI general circulation model. *Geophys. Res. Lett.* 26 (7), 843–846.

Kraichnan, R. H. (1967) Inertial Ranges in Two-Dimensional Turbulence. *Phys. Fluids* 10 (7), 1417.

LaCasce, J. H. (2012) Surface Quasigeostrophic Solutions and Baroclinic Modes with Exponential Stratification. *J. Phys. Oceanogr.* 42 (4), 569–580.

Lapeyre, G. and P. Klein (2006) Dynamics of the Upper Oceanic Layers in Terms of Surface Quasigeostrophy Theory. *J. Phys. Oceanogr.* 36 (2), 165–176.

Lapeyre, G., P. Klein, and B. L. Hua (2006) Oceanic Restratification Forced by Surface Frontogenesis. *J. Phys. Oceanogr.* 36 (8), 1577–1590.

Larichev, V. D. and I. M. Held (1995) Eddy Amplitudes and Fluxes in a Homogeneous Model of Fully Developed Baroclinic Instability. *J. Phys. Oceanogr.* 25 (10), 2285–2297.

Le Traon, P.-Y., P. Klein, B. L. Hua, and G. Dibarboure (2008) Do Altimeter Wavenumber Spectra Agree with the Interior or Surface Quasigeostrophic Theory? *J. Phys. Oceanogr.* 38 (5), 1137–1142.

Ledwell, J. R., A. J. Watson, and C. S. Law (1998) Mixing of a tracer in the pycnocline. 103 (C10), 21499–21499.

Leith, C. E. (1971) Atmospheric Predictability and Two-Dimensional Turbulence. *J. Atmos. Sci.* 28 (2), 145–161.

Lévy, M., P. Klein, A.-M. Tréguier, D. Iovino, G. Madec, S. Masson, and K. Takahashi (2010) Modifications of gyre circulation by sub-mesoscale physics. *Ocean Model.* 34, 1–15.

Lévy, M., D. Iovino, L. Resplandy, P. Klein, G. Madec, A.-M. Tréguier, S. Masson, and K. Takahashi (2012) Large-scale impacts of submesoscale dynamics on phytoplankton: Local and remote effects. *Ocean Model.* 43-44, 77–93.

Lilly, D. K. (1983) Stratified Turbulence and the Mesoscale Variability of the Atmosphere. *J. Atmos. Sci.* 40 (3), 749–761.

Lindborg, E. (1999) Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.* 388, 259–288.

Lindborg, E. (2006) The energy cascade in a strongly stratified fluid. *J. Fluid Mech.* 550, 207–242.

Lindborg, E. (2007) Horizontal Wavenumber Spectra of Vertical Vorticity and Horizontal Divergence in the Upper Troposphere and Lower Stratosphere. *J. Atmos. Sci.* 64 (3), 1017–1025.

Lindzen, R. S. (1994) The Eady Problem for a Basic State with Zero PV Gradient but $\beta \neq 0$. *J. Atmos. Sci.* 51 (22), 3221–3226.

Lorenz, E. N. (1963) Deterministic Nonperiodic Flow. *J. Atmos. Sci.* 20 (2), 130–141.

Lorenz, E. N. (1969) The predictability of a flow which possesses many scales of motion. *Tellus A* 21 (3), 289–307.

Lvov, Y. V., K. L. Polzin, and E. G. Tabak (2004) Energy Spectra of the Ocean's Internal Wave Field: Theory and Observations. *Phys. Rev. Lett.* 92 (12), 128501.

Mahadevan, A. (2014) Eddy effects on biogeochemistry. *Nature* 506, 168–169.

Mahadevan, A. and A. Tandon (2006) An analysis of mechanisms for submesoscale vertical motion at ocean fronts. *Ocean Model.* 14 (3-4), 241–256.

Mahadevan, A., A. Tandon, and R. Ferrari (2010) Rapid changes in mixed layer stratification driven by submesoscale instabilities and winds. *J. Geophys. Res.* 115 (C3), C03017.

Manucharyan, G. E. and M.-L. Timmermans (2013) Generation and Separation of Mesoscale Eddies from Surface Ocean Fronts. *J. Phys. Oceanogr.* 43 (12), 2545–2562.

Marshall, J. and T. Radko (2003) Residual-Mean Solutions for the Antarctic Circumpolar Current and Its Associated Overturning Circulation. *J. Phys. Ocean.* 33 (11), 2341–2354.

Marshall, J. and F. Schott (1999) Open-ocean convection: Observations, theory, and models. *Rev. Geophys.* 37 (1), 1–64.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey (1997) A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.* 102 (C3), 5753.

Mason, E., J. Molemaker, A. F. Shchepetkin, F. Colas, J. C. McWilliams, and P. Sangrà (2010) Procedures for offline grid nesting in regional ocean models. *Ocean Model.* 35 (1–2), 1–15.

McComas, C. H. and P. Müller (1981) The Dynamic Balance of Internal Waves. *J. Phys. Oceanogr.* 11 (7), 970–986.

McWilliams, J. C. and B. Fox-Kemper (2013) Oceanic wave-balanced surface fronts and filaments. *J. Fluid Mech.* 730, 464–490.

McWilliams, J. C., M. J. Molemaker, and E. I. Olafsdottir (2009) Linear Fluctuation Growth during Frontogenesis. *J. Phys. Oceanogr.* 39 (12), 3111–3129.

Mensa, J. A., Z. Garraffo, A. Griffa, T. M. Özgökmen, A. Haza, and M. Veneziani (2013) Seasonality of the submesoscale dynamics in the Gulf Stream region. *Ocean Dyn.* 63 (8), 923–941.

Molemaker, M. J., J. C. McWilliams, and I. Yavneh (2005) Baroclinic Instability and Loss of Balance. *J. Phys. Oceanogr.* 35 (9), 1505–1517.

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

Munk, W. (1981) Internal Waves and Small-Scale Processes. In: *Evolution of Physical Oceanography.* Ed. by B. A. Warren and C. Wunsch. Cambridge: The MIT Press. Chap. 9, 264–291.

Nakamura, N. (1988) Scale Selection of Baroclinic Instability—Effects of Stratification and Nongeostrophy. *J. Atmos. Sci.* 45 (21), 3253–3267.

Nakamura, N. and I. M. Held (1989) Nonlinear Equilibration of Two-Dimensional Eady Waves. *J. Atmos. Sci.* 46 (19), 3055–3064.

Nastrom, G. D. and K. S. Gage (1985) A Climatology of Atmospheric Wavenumber Spectra of Wind and Termperature Observed by Commercial Aircraft. *J. Atmos. Sci.* 42 (9), 950–960.

Nastrom, G. D., D. C. Fritts, and K. S. Gage (1987) An Investigation of Terrain Effects on the Mesoscale Spectrum of Atmospheric Motions. *J. Atmos. Sci.* 44 (20), 3087–3096.

Niiler, P. P. (1969) On the Ekman Divergence in an Oceanic Jet. *J. Geophys. Res.* 74 (28), 7048–7052.

Palmer, T. N. (2000) Predicting uncertainty in forecasts of weather and climate. *Reports Prog. Phys.* 63 (2), 71–116.

Pedlosky, J. (1987) *Geophysical Fluid Dynamics.* 2nd ed. New York: Springer, p. 710.

197

Pedlosky, J. (2010) *Waves in the Ocean and Atmosphere.* Berlin: Springer-Verlag, p. 260.

Phillips, N. A. (1954) Energy Transformations and Meridional Circulations associated with simple Baroclinic Waves in a two-level, Quasi-geostrophic Model. *Tellus* 6 (3), 273–286.

Pierrehumbert, R. T., I. M. Held, and K. L. Swanson (1994) Spectra of Local and Nonlocal Two-dimensional Turbulence. *Chaos, Solitons & Fractals* 4 (6), 1111–1116.

Pollard, R. T. (1970) On the generation by winds of inertial waves in the ocean. *Deep Sea Res. Oceanogr. Abstr.* 17 (4), 795–812.

Polzin, K. L. and Y. V. Lvov (2011) Toward Regional Characterizations of the Oceanic Internal Wavefield. *Rev. Geophys.* 49 (RG4003).

Qiu, B. (1999) Seasonal Eddy Field Modulation of the North Pacific Subtropical Countercurrent: TOPEX/Poseidon Observations and Theory. *J. Phys. Oceanogr.* 29 (10), 2471–2486.

Qiu, B. and S. Chen (2004) Seasonal Modulations in the Eddy Field of the South Pacific Ocean. *J. Phys. Oceanogr.* 34 (7), 1515–1527.

Ramachandran, S., A. Tandon, and A. Mahadevan (2014) Enhancement in vertical fluxes at a front by mesoscale-submesoscale coupling. *J. Geophys. Res. Ocean.* 119 (12), 8495–8511.

Ray, R. D. and G. T. Mitchum (1997) Surface manifestation of internal tides in the deep ocean: observations from altimetry and island gauges. *Prog. Oceanogr.* 40 (1–4), 135–162.

Ray, R. D. and E. D. Zaron (2011) Non-stationary internal tides observed with satellite altimetry. *Geophys. Res. Lett.* 38 (17), L17609.

Rhines, P. B. (1977) The dynamics of unsteady currents. In: *Sea, vol. VI.* Ed. by E. Goldberg. Wiley, 189–318.

Rhines, P. B. and W. R. Young (1982) Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.* 122, 347–367.

Richman, J. G., B. K. Arbic, J. F. Shriver, E. J. Metzger, and A. J. Wallcraft (2012) Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides. *J. Geophys. Res.* 117 (C12), C12012.

Riley, J. J. and E. Lindborg (2008) Stratified Turbulence: A Possible Interpretation of Some Geophysical Turbulence Measurements. *J. Atmos. Sci.* 65 (7), 2416–2424.

Rivest, C., C. A. Davis, and B. F. Farrell (1992) Upper-Tropospheric Synoptic-Scale Waves. Part I: Maintenance as Eady Normal Modes. *J. Atmos. Sci.* 49 (22), 2108–2119.

Rocha, C. B., T. K. Chereskin, S. T. Gille, and D. Menemenlis (2015) Drake Passage kinetic energy and sea-surface height horizontal wavenumber spectra in the mesoscale to submesoscale range (10-200 km). *Submitted.*

Roemmich, D. and J. Gilson (2009) The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Prog. Oceanogr.* 82 (2), 81–100.

Rossby, C.-G. (1939) Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. Mar. Res.* 2 (1), 38–55.

Roullet, G., J. C. McWilliams, X. Capet, and M. J. Molemaker (2012) Properties of Steady Geostrophic Turbulence with Isopycnal Outcropping. *J. Phys. Oceanogr.* 42 (1), 18–38.

Ruppert, J. H. and L. F. Bosart (2014) A Case Study of the Interaction of a Mesoscale Gravity Wave with a Mesoscale Convective System. *Mon. Weather Rev.* 142 (4), 1403–1429.

Salmon, R. (1978) Two-layer quasi-geostrophic turbulence in a simple special case. *Geophys. Astrophys. Fluid Dyn.* 10 (1), 25–52.

Salmon, R. (1980) Baroclinic instability and geostrophic turbulence. *Geophys. Astrophys. Fluid Dyn.* 15 (1), 167–211.

Samelson, R. M. and C. A. Paulson (1988) Towed Thermistor Chain Observations of Fronts in the Subtropical North Pacific. *J. Geophys. Res.* 93 (C3), 2237–2246.

Sasaki, H. and P. Klein (2012) SSH Wavenumber Spectra in the North Pacific from a High-Resolution Realistic Simulation. *J. Phys. Oceanogr.* 42 (7), 1233–1241.

Sasaki, H., P. Klein, B. Qiu, and Y. Sasai (2014) Impact of oceanic-scale interactions on the seasonal modulation of ocean dynamics by the atmosphere. *Nat. Commun.* 5 (5636).

Schmitz, Jr., W. J. (1988) Exploration of the Eddy Field in the Midlatitude North Pacific. *J. Phys. Oceanogr.* 18 (3), 459–468.

Scott, R. K. (2006) Local and nonlocal advection of a passive scalar. *Phys. Fluids* 18 (11).

Seiff, A. et al. (1998) Thermal structure of Jupiter's atmosphere near the edge of a 5-μm hot spot in the north equatorial belt. *J. Geophys. Res. Planets* 103 (103), E10.

Shchepetkin, A. F. and J. C. McWilliams (2005) The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* 9 (4), 347–404.

Shcherbina, A. Y., E. A. D'Asaro, C. M. Lee, J. M. Klymak, M. J. Molemaker, and J. C. McWilliams (2013) Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. *Geophys. Res. Lett.* 40 (17), 4706–4711.

Skamarock, W. C., S.-H. Park, J. B. Klemp, and C. Snyder (2014) Atmospheric Kinetic Energy Spectra from Global High-Resolution Nonhydrostatic Simulations. *J. Atmos. Sci.* 71 (11), 4369–4381.

Smagorinsky, J. (1963) General circulation experiments with the primitive equations I. The basic experiment. *Mon. Weather Rev.* 91 (3), 99–164.

Smith, K. S. and E. Bernard (2013) Geostrophic turbulence near rapid changes in stratification. *Phys. Fluids* 25 (4), 046601.

Smith, K. S. and R. Ferrari (2009) The Production and Dissipation of Compensated Thermohaline Variance by Mesoscale Stirring. *J. Phys. Oceanogr.* 39 (10), 2477–2501.

Smith, K. S. and G. K. Vallis (2001) The Scales and Equilibration of Midocean Eddies: Freely Evolving Flow. *J. Phys. Oceanogr.* 31 (2), 554–571.

Smith, K. S. and G. K. Vallis (2002) The Scales and Equilibration of Midocean Eddies: Forced–Dissipative Flow. *J. Phys. Oceanogr.* 32 (6), 1699–1720.

Smith, K. S. and J. Vanneste (2013) A Surface-Aware Projection Basis for Quasigeostrophic Flow. *J. Phys. Oceanogr.* 43 (3), 548–562.

Smith, L. M. and F. Waleffe (2002) Generation of slow large scales in forced rotating stratified turbulence. *J. Fluid Mech.* 451, 145–168.

Spall, M. A. (1995) Frontogenesis, subduction, and cross-front exchange at upper ocean fronts. *J. Geophys. Res.* 100 (C2), 2543.

Spall, M. A. (1997) Baroclinic Jets in Confluent Flow. *J. Phys. Oceanogr.* 27 (6), 1054–1071.

Stammer, D. (1997) Global Characteristics of Ocean Variability Estimated from Regional TOPEX/POSEIDON Altimeter Measurements. *J. Phys. Oceanogr.* 27 (8), 1743–1769.

Stommel, H. (1979) Determination of water mass properties of water pumped down from the Ekman layer to the geostrophic flow below. *Proc. Natl. Acad. Sci. U. S. A.* 76 (7), 3051–5.

Stone, P. H. (1966a) Frontogenesis by Horizontal Wind Deformation Fields. *J. Atmos. Sci.* 23 (5), 455–465.

Stone, P. H. (1966b) On Non-Geostrophic Baroclinic Stability. *J. Atmos. Sci.* 23 (4), 390–400.

Stone, P. H. (1970) On Non-Geostrophic Baroclinic Stability: Part II. *J. Atmos. Sci.* 27 (5), 721–726.

Stone, P. H. (1971) Baroclinic stability under non-hydrostatic conditions. *J. Fluid Mech.* 46, 659–671.

Swallow, J. C. (1971) The Aries current measurements in the western North Atlantic. *Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci.* 270 (1206), 451–460.

Taylor, J. R. and R. Ferrari (2010) Buoyancy and Wind-Driven Convection at Mixed Layer Density Fronts. *J. Phys. Oceanogr.* 40 (6), 1222–1242.

The MODE Group (1978) The Mid-Ocean Dynamics Experiment. *Deep Sea Res.* 25 (10), 859–910.

Thomas, L. N. and C. M. Lee (2005) Intensification of Ocean Fronts by Down-Front Winds. *J. Phys. Oceanogr.* 35 (6), 1086–1102.

Thomas, L. N., A. Tandon, and A. Mahadevan (2008) Submesoscale Processes and Dynamics. In: *Ocean Modeling in an Eddying Regime.* American Geophysical Union, 17–38.

Thomas, L. N., J. R. Taylor, R. Ferrari, and T. M. Joyce (2013) Symmetric instability in the Gulf Stream. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 91, 96–110.

Tulloch, R. and K. S. Smith (2006) A theory for the atmospheric energy spectrum: depth-limited temperature anomalies at the tropopause. *Proc. Natl. Acad. Sci. U. S. A.* 103 (40), 14690–14694.

Tulloch, R. and K. S. Smith (2009) Quasigeostrophic Turbulence with Explicit Surface Dynamics: Application to the Atmospheric Energy Spectrum. *J. Atmos. Sci.* 66 (2), 450–467.

Tulloch, R., J. Marshall, C. Hill, and K. S. Smith (2011) Scales, Growth Rates, and Spectral Fluxes of Baroclinic Instability in the Ocean. *J. Phys. Oceanogr.* 41 (6), 1057–1076.

Vallis, G. K. (2006) *Atmospheric and Oceanic Fluid Dynamics.* New York: Cambridge University Press, p. 745.

Vallis, G. K., G. J. Shutts, and M. E. B. Gray (1997) Balanced mesoscale motion and stratified turbulence forced by convection. *Q. J. R. Meteorol. Soc.* 123 (542), 1621–1652.

VanZandt, T. E. (1982) A universal spectrum of buoyancy waves in the atmosphere. *Geophys. Res. Lett.* 9 (5), 575–578.

Vincent, R. A. and S. D. Eckermann (1990) VHF radar observations of mesoscale motions in the troposphere: Evidence for gravity wave Doppler shifting. *Radio Sci.* 25 (5), 1019–1037.

von Storch, H. and K. Hasselmann (2010) *Seventy Years of Exploration in Oceanography.* Berlin: Springer-Verlag, p. 137.

Waite, M. L. and P. Bartello (2004) Stratified turbulence dominated by vortical motion. *J. Fluid Mech.* 517, 281–308.

Wang, D.-P., C. N. Flagg, K. Donohue, and H. T. Rossby (2010) Wavenumber Spectrum in the Gulf Stream from Shipboard ADCP Observations and Comparison with Altimetry Measurements. *J. Phys. Oceanogr.* 40 (4), 840–844.

Whitt, D. B. and L. N. Thomas (2015) Resonant generation and energetics of wind-forced near-inertial motions in a geostrophic flow. *J. Phys. Oceanogr.* 45 (1), 181–208.

Wortham, C., J. Callies, and M. G. Scharffenberg (2014) Asymmetries between Wavenumber Spectra of Along- and Across-Track Velocity from Tandem Mission Altimetry. *J. Phys. Oceanogr.* 44 (4), 1151–1160.

Wunsch, C. and D. Stammer (1998) Satellite altimetry, the marine geoid, and the oceanic general circulation. *Annu. Rev. Fluid Mech.* 26, 219–253.

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

Xu, Y. and L.-L. Fu (2011) Global Variability of the Wavenumber Spectrum of Oceanic Mesoscale Turbulence. *J. Phys. Oceanogr.* 41 (4), 802–809.

Xu, Y. and L.-L. Fu (2012) The Effects of Altimeter Instrument Noise on the Estimation of the Wavenumber Spectrum of Sea Surface Height. *J. Phys. Oceanogr.* 42 (12), 2229–2233.

Zhang, F., N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio (2007) Mesoscale Predictability of Moist Baroclinic Waves: Convection-Permitting Experiments and Multistage Error Growth Dynamics. *J. Atmos. Sci.* 64 (10), 3579–3594.

Zhao, Z., M. H. Alford, and J. B. Girton (2012) Mapping Low-Mode Internal Tides from Multisatellite Altimetry. *Oceanography* 25 (2), 42–51.