

# GEOPHYSICAL FLUID MECHANICS

VOLUME 5

WAVES AND INSTABILITIES

Stephen.M.Griffies@gmail.com

Princeton University Atmospheric and Oceanic Sciences

Draft from January 12, 2026



COPYRIGHT ©2025 BY STEPHEN M. GRIFFIES

ALL RIGHTS RESERVED

GRIFFIES, STEPHEN M., 1962-

GEOPHYSICAL FLUID MECHANICS

THIS BOOK WAS TYPESET USING LATEX.

# CONTENTS

	<b>Page</b>
PREFACE	<b>vii</b>
GUIDE TO THIS BOOK	<b>xvii</b>
<b>Part I. Wave mechanics</b>	<b>1</b>
1 PLANE WAVES AND WAVE PACKETS	5
2 WAVES IN A GENTLY VARYING BACKGROUND	43
3 ACOUSTIC WAVES	63
4 INTERFACIAL WAVES ON POTENTIAL FLOW	95
5 INERTIAL WAVES ON THE $f$ -PLANE	147
6 BAROTROPIC VORTICITY WAVES	165
7 SHALLOW WATER WAVES	191
8 SHALLOW WATER WAVES: CASE STUDIES	231
9 INTERNAL INERTIA-GRAVITY WAVES	249
10 INTERNAL GRAVITY WAVES: CASE STUDIES	291
<b>Part II. Flow instabilities</b>	<b>323</b>
11 SYMMETRIC FLOWS	327
12 STABILITY OF FLUID INTERFACES	371
13 SHEAR INSTABILITY	389
14 QUASI-GEOSTROPHIC WAVES AND BAROCLINIC INSTABILITY	431
<b>Part III. End matter</b>	<b>467</b>
A GLOSSARY OF CONCEPTS AND TERMS	469
B LIST OF ACRONYMS	479
C LIST OF SYMBOLS	481
BIBLIOGRAPHY	491
INDEX	499



## CONTENTS



## PREFACE

Geophysical fluid mechanics (GFM) is a branch of theoretical physics concerned with natural fluid motion on a rotating and gravitating planet or star. The subject makes use of concepts from classical continuum mechanics and thermodynamics, along with the corresponding methods of mathematical physics. The primary inspiration for our study comes from the motion of fluids in the earth's atmosphere and ocean, though the principles and methods are also applicable to extra-terrestrial fluid flows. Geophysical fluids are in near rigid-body motion with the rotating planet, thus prompting a description from the rotating (non-inertial) planetary reference frame. Body forces from gravity plus planetary rotation (Coriolis and centrifugal) are fundamental features of the motion, as are contact forces from stresses (pressure and friction). In this book, we limit attention to the motion of a single phase of matter (gas or liquid), with the study of multiphase geophysical fluid mechanics, which is relevant to a moist atmosphere, outside our scope. Electromagnetic forces, important for the study of astrophysical fluid motions, are also ignored. We also ignore chemical reactions (which transform matter from one form to another), and nuclear reactions (which convert between matter and nuclear energy).

Geophysical fluid flows manifest over a huge range of space and time scales, with linear and nonlinear interactions transferring information across these scales. Physical insights into such flows typically result from examining a hierarchy of conceptual models using a variety of methods and perspectives. Some of the models we consider are formulated within the context of a **perfect fluid** comprising a single material constituent with fundamental processes limited to the reversible and mechanical. Other models are posed using a **real fluid** that is comprised of multiple matter constituents exposed to an **irreversible process** such as mixing of momentum through viscous friction, mixing of matter through matter diffusion, and/or the mixing of enthalpy through conduction. Some models consider constant density fluids, as commonly considered in classical **hydrodynamics**. And some models ignore rotation, thus tacitly applying to flows with length scales too short to feel the planetary Coriolis acceleration, whereas others ignore buoyancy to thus focus on the dynamics of a homogeneous fluid in a rotating reference frame.

We develop geophysical fluid mechanics from a mathematical physics perspective, with a grounding in fundamentals offering a robust and versatile framework for exploring a gamut of special cases and approximations. Topics are approached by establishing general principles prior to the examination of case studies. Consistent with this approach, our treatment focuses on developing the mechanics of geophysical fluid motion, with that focus supporting theoretical explorations that often extend beyond that required for phenomenological purposes. Correspondingly, we embrace the opportunity to examine physics through multiple lenses that render a variety of complementary insights. In a nutshell, if a physical system can be formulated and analyzed in more than one way, then we do so if it enhances pedagogy and exposes layers of understanding. As a result, brevity is sacrificed to support exposition and exploration, with this perspective leading to a book with multiple volumes.

---

The presentation is based on the premise that skills in theoretical physics are optimally taught by nurturing physical reasoning, with physical reasoning supported by mathematical precision coupled to the elucidation of concepts using words and pictures. Correspondingly, the presentation is both deductive and descriptive. The deductive approach supports a precise understanding through the use of elementary physical notions that are expressed mathematically. The descriptive approach builds skills in reasoning along with the ability to articulate physical ideas using words and pictures that complement the maths. Readers are supported by development of salient physical concepts and mathematical methods in the process of building understanding. With sufficient study, the material in this book should be accessible to the advanced undergraduate student or entering graduate student in fields such as applied mathematics, astrophysics, atmospheric physics, engineering, geophysics, ocean physics, planetary physics, and theoretical physics.

We generally offer details to mathematical derivations. Doing so nurtures the mathematical skills required for the budding theorist, with the reader strongly encouraged to work through the various derivations and exercises to fully embrace each detail and concept. Exposing mathematical details also helps to unpack many of the physical concepts encapsulated by equations. It is notable that the concepts encountered in this book generally accord with common experience (we are doing classical physics), thus affording a means to check on the validity of the maths. Even so, it does take time to wrap one's head around the physics of large-scale ocean and atmosphere circulations, so patience and persistence are required. Furthermore, as we are studying physics, all mathematical equations must satisfy dimensional consistency, with this constraint offering the physicist a powerful tool for exposing spurious mathematical statements.

We consider this book to be an intellectual journey taken together by the author and reader, thus motivating use of the first person plural pronouns *we* and *us*. Furthermore, we cultivate the deductive and descriptive approaches by embracing the synergism between physics and maths, whereby physics informs the maths and maths reveals the physics. This synergism is facilitated by a presentation style inspired by [Mermin \(1989\)](#), who identified the following characteristics for clear presentations of mathematical physics.

- RULE 1: All displayed equations are given numbers to facilitate cross-referencing. Additionally, any equation supporting another equation or a discussion is itself afforded an equation number.
- RULE 2: Cross-referenced equations are referred to by their equation number as well as descriptive phrases or names (e.g., “as seen by the vector-invariant velocity equation XX.YY” rather than just “as seen by equation XX.YY”). Coupling maths to words supports learning and reduces the need to flip pages to view the cited equation.
- RULE 3: Equations are part of the prose and are thus subject to punctuation.

## Concerning the book’s title

The study of rotating and stratified geophysical fluid motion largely started in the first half of the 20th century. During recent decades, the study has seen particular evolution through deepening physical foundations, refining mathematical formulations, increasing the intellectual and predictive value of numerical simulations, extending applications across terrestrial and planetary systems, and expanding observational and laboratory measurements and techniques.

---

What has emerged is a recognition that a fruitful study of rotating and stratified fluid flows makes use of ideas that go beyond the traditional notions of **geophysical fluid dynamics** (GFD). Rather, the contemporary practitioner develops insights by weaving together concepts and tools from mathematics, Newtonian mechanics, analytical mechanics, fluid mechanics, thermodynamics, classical scalar field theory, numerical simulations, laboratory experiments, field measurements, and data science. Acknowledging this broadening of the practice motivates the term *mechanics* in this book's title, rather than the more focused *dynamics*. It is a minor change in verbiage that reflects a broadening of the perspectives and goals pursued here.

## Two pillars of theoretical geophysical fluid mechanics

We conceive of two pillars to theoretical geophysical fluid mechanics that are synergistic, thus offering lessons, guidance, and feedback to the other. The **elements pillar** of geophysical fluid mechanics comprises the physical and mathematical formulation of conceptual models used to garner insight into rotating and stratified fluid motion. This pillar is concerned with setting the stage by deductively and descriptively exposing how physical concepts are mathematically expressed to describe geophysical fluid flows. We provide a thorough treatment of the element pillar given its foundational importance, and since it is commonly offered only a terse treatment in other presentations. We emphasize that the elements pillar is far more than equation manipulation, although one certainly must become adept at that task. Instead, at its core, the elements pillar allows the physicist to reveal the fundamental physical concepts in a precise mathematical manner. Doing so supports understanding while building the foundations for the **emergent phenomena pillar**. The emergent phenomena pillar of geophysical fluid mechanics studies solutions to equations that describe phenomena, such as waves, instabilities, turbulence, and general circulation, all of which emerge from the fundamental equations. Phenomena can emerge in manners that are far from simple to understand deductively, particularly when considering nonlinear behavior such as turbulence. Our treatment of the emergent pillar is limited to waves and instabilities, whereas turbulence and general circulation are beyond our scope, though we do touch upon these topics where suited to the discussion.<sup>1</sup>

## Some themes found in this book

This multi-volume book covers a number of topics in theoretical geophysical fluid mechanics. Throughout, we encounter a number of themes that appear in various guises, with the following offering a brief survey.

### Causation and budgets

A great deal of this book is concerned with deriving and understanding equations that describe the evolution of fluid properties, with such equations (differential or integral) derived from physical principles such as Newton's laws of motion, Hamilton's principle of stationary action, Noether's theorem, thermodynamic laws, mass conservation, and vorticity mechanics. These **budget equations** form the theoretical foundation of continuum mechanics. As part of this development we often seek information about what *causes* fluid motion, making use of a variety of kinematic and mathematical frameworks. The causality question is nicely posed by Newton's

---

<sup>1</sup>The further one moves along the axis of nonlinearity, the more Sisyphean the task of connecting fundamental processes to emergent phenomena. This perspective is lucidly discussed by [Anderson \(1972\)](#).

---

equation of motion, which says that acceleration (motion) arises from a net force (the cause of motion). Even though seemingly a clear decomposition of cause and effect, this fundamental statement of Newtonian mechanics offers little more than the definition of a force. We break the self-referential loop, and thus make physical progress, after specifying the nature of the force (e.g., gravitational, frictional), as well as by offering properties of these forces as per Newton's third law (the action/reaction law).<sup>2</sup>

In geophysical fluid mechanics, we sometimes refer to a time evolving budget equation as an **evolution equation** or, more commonly, a **prognostic equation**, with each term in the prognostic equation referred to as a **time tendency**.<sup>3</sup> For prognostic equations, knowledge of the processes contributing to the net time tendency enables a prediction of flow properties. The question arises how to practically determine the tendencies acting in the fluid, particularly when tendencies are generally dependent on the flow itself. This question is often very difficult to answer. Such is the complexity and beauty inherent in nonlinear field theories such as fluid mechanics, where cause and effect are intrinsically coupled.

We can sometimes make progress by turning the problem around, whereby kinematic knowledge of the motion offers inferential knowledge of the dynamical processes contributing to the motion. This situation is exemplified by pressure forces acting within a non-divergent flow whereby pressure provides the force that acts, instantaneously and globally, to maintain the constraint that the velocity is non-divergent.<sup>4</sup> We may also make use of constraints that restrict the flow in manners that assist in prediction and understanding.

## Constraints

Determining the forces, either directly or indirectly, provides physical insight into the cause of fluid flow and its changes. This approach is sometimes referred to a **momentum based viewpoint** since it is based on working directly with the momentum equation (i.e., Newton's second law of motion). However, we are commonly unable to deduce the forces due to complexities inherent in nonlinear field theories. Furthermore, there are many occasions when we are simply uninterested in the forces. In these cases, we are motivated to use constraints that can allow us to sidestep forces but still garner insights into the motion.

One example of a constraint concerns the inability of fluid to flow through a solid static material boundary, such as the solid-earth boundary encountered by geophysical flows. To understand how this constraint impacts the macroscopic fluid motion, we do not need to understand details of the atomic forces that underlie the resistance to macroscopic motion. Instead, we simply impose the kinematic boundary condition whereby the component of the velocity that is normal to the boundary vanishes at the boundary. The forces active within the fluid, no matter what flavor they may take, are constrained to respect the kinematic boundary condition. Another example concerns the study of vorticity. A variety of vorticity constraints offer the means to deduce flow properties without determining forces. Indeed, the **vorticity based viewpoint** often provides a framework that is more versatile in practice than the momentum-based approach, thus prompting the importance of vortex mechanics in the study of geophysical fluid flows.

---

<sup>2</sup>For more on this perspective of Newton's laws, see Chapter 1 of *Symon* (1971) or Chapter 2 of *Marion and Thornton* (1988).

<sup>3</sup>This language has its origins in weather forecasting.

<sup>4</sup>For non-divergent flow, pressure acts as the *Lagrange multiplier* enforcing flow non-divergence.

---

## Associations and balances

Besides seeking causal relations pointing toward the future, many basic questions of fluid mechanics arise either instantaneously, as in the constraints maintaining non-divergent flows, or when the flow is steady, in which case properties at each point in space have no time dependence. In steady flows, the net acceleration, and hence the net force, vanish at each point within the fluid, although the fluid itself can still be moving (steady flows are not necessarily static). For steady flows we are unconcerned with causality since time changes have been removed. In this manner, a steady state equation is a **diagnostic equation** rather than a **prognostic equation**. Diagnostic equations thus provide mechanical statements about associations between physical processes that manifest as balances. The **geostrophic balance** is the canonical association in geophysical fluid mechanics, where the horizontal Coriolis force is balanced by the horizontal pressure gradient force. Another balance concerns the vertical pressure gradient and its near balance with the weight of fluid above a point in the fluid, with this **hydrostatic balance** approximately maintained at the large scale even for moving geophysical fluids. Further associations arise when studying steady vorticity balances, with the **Sverdrup balance** a key example that is commonly used in ocean circulation theory.

We summarize the above by saying that diagnostic equations are concerned with the way things are, whereas prognostic equations point to how things will be. So although a predictive theory requires prognostic equations that manifest causal relations, an understanding of how fluid motion appears, and in particular how it is constrained, is revealed by studying diagnostic relations that expose associations through balances.

## Mathematical transformations between kinematic perspectives

Geophysical fluid flows are complex. Hence, it proves useful to avail ourselves of a variety of methods and perspectives that support a mechanistic description of the motion. Many methods are associated with distinct kinematic lenses that reveal particular facets of the flow that might be less visible using alternative lenses. Examples include the Eulerian (spatial) and Lagrangian (material) kinematics used throughout fluid mechanics; the dual position space ( $x$ -space) and wavevector space ( $k$ -space) used for wave mechanics; the variety of vertical coordinates used for vertically stratified flows; and the analysis of motion in property spaces exemplified by watermass or thermodynamic analysis. We make use of these perspectives throughout this book, and offer the mathematical tools (e.g., tensor methods) needed to transform between them.

## Newtonian mechanics and Hamilton's principle

Throughout this book we pursue the maxim

PURSUE ALL WAYS TO FORMULATE AND TO SOLVE A PROBLEM.

A canonical example concerns the complementary perspectives available from Newtonian mechanics and Hamilton's principle of stationary action. Each offers logically consistent results yet approaches mechanics from fundamentally distinct conceptual and operational perspectives. In a Newtonian approach to fluid mechanics, governing differential equations are formulated using a continuum version of Newton's law of motion, in which forces (causes) and accelerations (effects) are articulated as a means to understand and predict the flow. The alternative approach of Hamilton's principle of stationary action approaches mechanics via a variational formulation involving the **action**. Hamilton's principle says that the action functional is extremized by

---

the physically realized system. The action is the space-time integral of the difference between kinetic and potential/internal energies, and by extremizing the action we reveal the governing Euler-Lagrange differential equations. The Euler-Lagrange equations are identical to Newton's equations for those cases where Newton's equations are available,<sup>5</sup> and yet the route to deriving these equations is very distinct. It is by pursuing these distinct paths that we uncover new insights and develop distinct tools for analysis.

Hamilton's principle is not typically covered in fluid mechanics books. This absence contrasts to the ubiquity of Hamilton's principle in other areas of physics. We include facets of Hamilton's principle in this book with the hope that doing so partially remedies the disconnect.<sup>6</sup> Furthermore, we include Hamilton's principle since it provides novel perspectives on the fundamental equations of geophysical fluid mechanics, and renders insights and tools for the study of emergent phenomena such as waves and instabilities. The reader interested in a serious pursuit of theoretical mechanics should, at some point, make friends with Hamilton's principle. The effort is nontrivial as it requires brain muscles not exercised when studying Newtonian mechanics. But the conceptual and technical payoff is significant.

### Non-dimensionalization and scale analysis

Mathematical symbols describing a physical system generally have physical dimensions. Examining the physical dimensions of an equation supports an understanding of the physical content of the equation, and provides a powerful means to identify errors in mathematical manipulations. It is for this reason that we prefer to expose physical dimensions throughout this book, rather than the alternative approach of working predominantly with non-dimensional equations. Even so, scale analysis, as realized through [non-dimensionalization](#), offers an essential tool for deriving mathematical equations used to describe particular flow regimes.

There are two general types of dimensional scales that we use to non-dimensionalize a mathematical physics equation. The first is the [external scale](#), with examples in this book being the gravitational acceleration, Coriolis parameter, and specified properties of the background state such as the buoyancy frequency or prescribed flow. External scales are set by the geophysical parameter regime in which the flow occurs, and as such they are under direct control of the theorist or experimentalist. The second is the [emergent scale](#), which emerges from the flow itself. Emergent scales, such as the length scale and velocity scale of the flow, are specified by the subjective interest of the theorist though these scales are not under direct control. That is, we choose to focus on flows with a particular scale for purposes of examining the corresponding equations that describe that flow regime. A key example concerns our study of planetary geostrophy and quasi-geostrophy, where we choose to focus on flows of a particular scale where the Coriolis acceleration is of leading order importance.

We thus consider the operational aspects of scale analysis to be largely subjective in nature. Namely, we approach the analysis with a subjective bias towards the flow regime of interest, which in turn affects choices for non-dimensional parameters that lead to the corresponding

---

<sup>5</sup>Hamilton's principle yields the Maxwell's equations of electromagnetism, and yet Maxwell's equations are distinct from Newton's equations. Indeed, Hamilton's principle is used throughout modern physics in areas far beyond Newtonian mechanics.

<sup>6</sup>There certainly are examples where Hamilton's principle is discussed in fluid mechanics books, with [Salmon \(1998\)](#), [Olbers et al. \(2012\)](#), and [Badin and Crisciani \(2018\)](#) notable examples that have inspired this author. Even so, these books remain the exception rather than the norm. As a result, the broader geophysical fluid mechanics community, even those pursuing theoretical aspects, are largely unaware of the beauty and power of Hamilton's principle. This situation contrasts to nearly every other area of mechanics, in which Hamilton's principle is central to both theory and application.

---

asymptotic equations that describe the regime. Hence, scale analysis is deductive while being strongly guided by our subjective interests.

## Geophysical Fluid Mechanics and Climate Science

Fluid mechanics has a history of applications that span science and engineering, from blood flow to the evolution of galaxies. A key 21st century application of geophysical fluid mechanics concerns the questions of earth system science associated with the uncontrolled greenhouse gas experiment pursued by industrialized civilization's carbon centered energy use. Leading order science questions about climate warming have been sufficiently addressed to recognize that the planet has reached a crisis point threatening many features of the biosphere. Even so, mechanistic answers to a number of questions remain at the cutting edge of research. What will happen to the atmospheric jet stream and storm tracks in a world without summer Arctic sea ice? Will tropical storms be more powerful in a warmer world? What are the patterns for coastal sea level rise and their connections to large-scale ocean circulation? What are the key processes acting to bring relatively warm ocean waters to the base of high latitude ice shelves? How stable are the ocean and atmosphere's large-scale overturning circulations and their associated heat transport? Are there feasible and sustainable climate intervention options that equitably reduce the negative impacts of climate warming without introducing new problems? These questions, and countless others, constitute key intellectual challenges of climate science in particular and Earth system science more generally.

Numerical circulation models, observational field campaigns (both *in situ* and remote), and laboratory experiments, are core platforms for Earth system science. Many of these platforms have reached a level of maturity allowing them to vividly reveal details of the complex and multi-scaled nature of planetary fluid flow. Geophysical fluid mechanics is key to the design of observational field campaigns and novel laboratory and numerical experiments, and it provides the intellectual framework for developing mechanistic analyses and robust interpretations of measurements and simulations. In this way, geophysical fluid mechanics furthers predictive capability for weather and climate forecast systems and it enhances confidence in projections for future climate. In a world of increasingly large volumes of simulated and measured data, we conjecture that the marriage of fundamental physical theory to data science tools will enable the significant science and engineering advances needed to address key questions of Earth system science.

## About the cover

I took the cover photo of an iceberg, ocean, clouds, and sea bird (can you find the bird?) in the Orkney Passage region of the Southern Ocean during a research cruise from March-May 2017 aboard the British ship James Clark Ross. I am grateful to Alberto Naveira Garabato, the chief scientist on this cruise, for taking me to this amazing part of the planet. Although I largely pursue theoretical research, experiences with seagoing field research have greatly enhanced my scientific viewpoint and profoundly deepened a connection to the natural forces and phenomena that are in part described by geophysical fluid mechanics.

---

## Gratitudes

This book greatly benefited from interactions with students in the Princeton University Atmospheric and Oceanic Sciences Program. In particular, parts of this book served as the basis for my teaching, over many years, a two-semester graduate course, AOS 571 and AOS 572. It also supported a variety of special topic classes (AOS/GEO 585) and lecture series. Further inspiration was offered by students, postdocs, and fellow researchers and scholars encountered on my path. I also thank those who provided specific suggestions, corrections, and comments on various drafts of this book, whose names are too many to list.

I am grateful for having been part of the unique research and learning environment cultivated by three of the world's best examples of scientific enterprises. First and foremost, I am the product of NOAA's Geophysical Fluid Dynamics Laboratory (GFDL), where I worked as a research physicist from 1996 until 2025. As part of my life as a US federal research scientist, I was fortunate to also be associated with Princeton University's Atmospheric and Oceanic Sciences (AOS) program, where I was a postdoc from 1993-1996 and then a faculty member from 2014-2026. As of 2026, I entered the most recent (hopefully not the final!) part of my career journey as a CNRS research scientist in Paris, a position offering an amazing, and humbling, level of intellectual freedom. Throughout my career, I have focused research concerns on ocean physics and the ocean's role in climate, and I have pursued this research from the perspective garnered from the theoretical physics, applied maths, and chemical engineering training of my undergraduate and graduate education.

The communities at GFDL, Princeton AOS, and CNRS provide an ideal setting for those interested in broadening scientific perspectives while diving deep into particular research areas. As part of my research and mentoring in this community, I have encountered thinkers whose style, questions, and insights have taken root in my work. This work has also afforded me the opportunity to travel the world to interact with colleagues whose wisdom and love of the scientific endeavor are infectious and inspiring. Throughout these interactions, I have entered into trusting and non-judgmental spaces where deep learning and understanding arise. Partaking in these spaces, where heart and mind meld, has been among the most fulfilling experiences of my life.

Developing a book of this nature is not a simple endeavor. It starts modestly, grows over time, and eventually becomes a passion and obsession. I was particularly drawn to writing during the COVID-19 pandemic that kept the world largely sequestered at home, and I am grateful that my life situation allowed for this work to safely flourish during what were otherwise very difficult times for civilization. Writing this book has been an exercise in rational thought that exemplifies the maxim "to write is to learn", as articulated by [Zinnser \(1993\)](#). It was furthermore fed by spiritual food from meditation, yoga, family, and community. In particular, each step was supported by my wife, Adi, and our son, Francisco. I am deeply grateful for their patience and trust as I satisfied the goal of writing this book through countless nights, weekends, and holidays. I treasure being part of our family and I dedicate this work to you two amazing human beings.

## Caveats and limitations

This book remains a work in progress that is not yet ready for publication. There are many loose threads detailed at the start of various chapters. In addition, here are items targeted for completion prior to release of this book to a publisher.

- 
- Wave mechanics
    - equatorial shallow water waves
    - Rossby wave packets and motion in non-homogeneous background
    - Shallow water waves on a rotating sphere, including Laplace's tidal equations, Hough functions, and spherical harmonics
    - Ray theory using Hamilton's principle as in [Tracy et al. \(2014\)](#)
  - Flow stability
    - Charney problem of baroclinic instability
    - Arnold's stability theorem
    - Rayleigh-Benard convection
  - Application of Hamilton's principle
    - Referential flow using Hamilton's principle
    - shallow water and Hamilton's principle
    - semi-geostrophy and Hamilton's principle
    - quasi-geostrophy and Hamilton's principle
    - waves and mean flow interactions
  - Mathematical topics
    - Lie derivative following Section F.3 of [Tromp \(2025\)](#)





## GUIDE TO THIS BOOK

No book is an island, with this book generously making use of other books, review articles, research papers, and online tutorials. Many readers find value in studying a subject from a variety of perspectives and voices, thus justifying the proliferation of books with overlapping subject matter. Sometimes it is merely one or two sentences that allow for an idea or concept to click within the reader's brain, whereas other topics require the full gamut of detailed derivations and discussions coming from multiple voices. For these reasons we provide pointers to written and/or video presentations that offer supportive views on material in this book. Many further resources are available through a quick internet search or consultation with artificial intelligence (AI).

There is no pretense that any reader will study all topics in this multi-volume book. This recognition is particularly apparent in a world where research and educational agendas often spread rather than focus attention. Hence, an attempt has been made to facilitate picking up each book at a variety of starting points. For that purpose, each chapter is written in a reasonably self-contained manner and with a brief guide at the start of each chapter listing pre-requisite material. As such, some equations and derivations are reproduced in more than one place, thus obviating the need to back reference. Certainly each chapter cannot be fully self-contained, since this is a book with material building from other chapters across the volumes. We thus make generous use of cross-referencing to point out allied material treated elsewhere. We also make extensive use of the glossary to help define concepts accessed in one volume that might be more thoroughly treated in another volume.



## Organization

This book contains five volumes, each of which comprises parts that have multiple chapters. Parts and chapters start with a brief guide to the material along with pointers to dependencies. The book's end matter includes a glossary of key concepts and terms. Items highlighted within the text identify terms with a glossary entry. The glossary also serves as an annotated index, with page numbers pointing to where the terms and concepts are examined within a particular volume. Indeed, the glossary is an essential means to navigate this multi-volume book, reducing (though not eliminating) the need to have more than one volume open at a time. The glossary is then followed by a list of acronyms<sup>7</sup> and then by a list of symbols. A bibliography follows, with pages listed for where the book or paper is cited. We close the book with an index.

Not all topics are treated equally, with some probed deeply whereas others are given relatively superficial treatment. Indeed, there are many topics omitted that arguably should

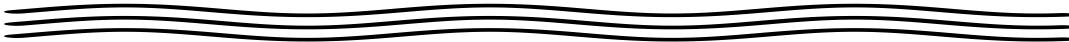
---

<sup>7</sup>We generally try to avoid acronyms, but some are inevitable.

---

find a home here. Each shortcoming reflects on the author's limited energy and experience rather than a judgement of relative importance.

Cross-referencing to specific sections and equations is provided when pointing to material within the same volume. Cross-referencing material in other volumes is less specific. In many cases, a cross-reference concerns an item in the glossary and/or index, which can be consulted across volumes to help make the connection.



## Volume 1

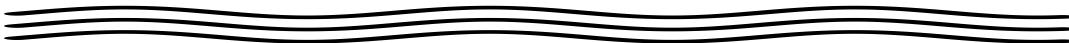
VOLUME 1 establishes foundations in mathematical physics and classical mechanics.

### **Mathematical physics**

We start the book with a suite of mathematical methods chapters. Many readers can skim these chapters without sacrificing too much from later chapters, assuming they have a working knowledge of Cartesian tensors as well as vector differential and integral calculus. Where unfamiliar mathematics topics arise in later chapters, the reader is encouraged to return to this part of the book to help develop the necessary skills.

### **Classical mechanics**

We here survey salient topics from classical mechanics with a geophysical perspective, and in turn develop concepts and methods that have direct relevance to the continuum physics of geophysical fluid flows. Of particular note, this part develops an understanding of physics as viewed from a rotating reference frame. Doing so allows for the sometimes non-intuitive results of rotating physics to be developed within the context of a particle system as a pedagogical preface to later developments for geophysical fluid motion.



## Volume 2

VOLUME 2 treats the fundamentals of fluid mechanics with an emphasis on geophysical fluid mechanics.

### **Kinematics of fluid flow**

Mechanics is comprised of kinematics (the study of intrinsic properties of motion) and dynamics (the study of forces and energies causing motion). In the fluid kinematics part of this book, we initiate a study of fluid mechanics by focusing on the kinematics of fluid flow and matter transported by that flow. Our treatment exposes both the Eulerian and Lagrangian viewpoints and emphasizes the variety of kinematic notions and tools key to describing fluid motion. We also encounter facets of material transport as described by the tracer equation. We emphasize that fluid flow, and the transport of matter within that flow, have many features that are fundamentally distinct from point particle and rigid body motion. It takes practice to intellectually digest these differences.

---

Some kinematic topics can seem esoteric on first encounter, particularly the study of Lagrangian kinematics. However, an incomplete understanding of fluid kinematics can lead to difficulties appreciating facets of fluid dynamics. The reader is thus encouraged to fully study the kinematics chapters, and to revisit the material as the needs arise in later chapters.

## Thermodynamics

We study the rudiments of thermodynamics with a focus on topics arising in the study of geophysical fluids. We pay particular attention to the role of gravity in modifying the treatment of thermodynamic equilibrium states, with gravity an essential facet of geophysical fluids and yet a force that is commonly ignored in standard treatments of thermodynamics. We ignore phase transitions, which is a notable limitation of our treatment, thus making this part of the book a mere introduction to the study of a moist atmosphere or an ocean with sea ice.

### Dynamics of geophysical fluid flow

In this lengthy part of the book, we study how Newton's laws of mechanics and the principles of thermodynamics are applied to continuum fluid motion on a rotating and gravitating planet. We approach the subject by focusing on how forces that act on fluid elements lead to accelerations and thus to motion. In particular, we examine **body forces** that act throughout the volume of a fluid element (e.g., planetary gravity, planetary Coriolis, and planetary centrifugal) as well as **contact forces** that act on the boundary of a fluid element (e.g., pressure and friction).



## Volume 3

### Shallow water mechanics

A shallow water fluid is comprised of hydrostatically balanced homogeneous fluid layers. The layers are also typically assumed to be immiscible, so that interactions between layers occur only via mechanical forces from pressure acting at the layer interfaces. The shallow water fluid allows us to focus on planetary rotation and vertical stratification without the complexities of vertically continuous stratification and thermodynamics. Many physical insights garnered by studying shallow water fluids extend to more realistic fluids, thus making the shallow water model very popular among theorists and teachers. Indeed, [Zeitlin \(2018\)](#) provides an example of just how far one can go in understanding geophysical fluids with shallow water theory.

### Vorticity

**Vorticity** plays a role in the motion of all geophysical fluids since motion on a rotating planet provides a nonzero **planetary vorticity** even to fluids at rest on the planet. This feature of geophysical fluids contrasts to many other areas of fluid mechanics, where irrotational flows are commonly encountered. **Potential vorticity** is a strategically chosen component of the vorticity vector that melds mechanics (vorticity) to thermodynamics (stratification). Material conservation properties of potential vorticity are striking and render important constraints on fluid motion. Indeed, perhaps the most practical reason to study vorticity concerns the various constraints imposed on the flow moving on a rotating and gravitating planet. These constraints provide conceptual insights and predictive power.

---

## Nearly geostrophic balanced flows

Balanced models generally remove the horizontally divergent motions associated with gravity waves, thus allowing a focus on the large-scale vortical motions. Balanced models have a rich history among theoretical geophysical fluid studies, providing insights into both laminar oceanic flows through planetary geostrophy, and wave-turbulent atmospheric and oceanic flows through quasi-geostrophy. When studying balanced models, we focus on the shallow water and continuously stratified versions of quasi-geostrophy and planetary geostrophy.



## Volume 4

### Generalized vertical coordinates

The chapters on **generalized vertical coordinate (GVC)s (GVC)** dive into the maths, kinematics, dynamics, and applications of such coordinates for the study of geophysical fluid mechanics. This material is central to many current research activities, including subgrid scale parameterizations and the design of numerical atmosphere and ocean models. The mathematics in this part leans heavily on the general tensors studied in VOLUME 1. Even so, many readers can make the most of these chapter without the full gamut of general tensors.

### Scalar fields

Many chapters target the mechanics of scalar fields with a focus mostly on the ocean. Here we consider active tracers (temperature and salinity), passive tracers, and buoyancy. Much of this study forms the basis of tracer mechanics, which has proven very important for the ocean since it is generally very difficult to measure vector fields such as velocity and vorticity, whereas tracer distributions are far more readily measured. We also consider facets of sea level analysis in this part of the book.

### Hamilton's principle for geophysical flows

We study the analytical mechanics of geophysical flows using Hamilton's principle. This material forms the heart of field theory, both classical and quantum. It requires a different set of techniques than used in the study of Newtonian fluid mechanics used elsewhere in this book. Hence, it offers complementary insights that deepen our understanding of geophysical fluid flows in particular.



## Volume 5

### Linear wave mechanics

We study a variety of geophysical waves and associated mathematical methods used for their characterization. Notably, we consider waves not commonly included in a book on geophysical fluids, such as sound and capillary waves, with these waves included due to their ubiquity in the natural environment as well as their pedagogical value. Most focus, however, is given to waves arising from the Coriolis acceleration (inertial waves, planetary Rossby waves, topographic

---

Rossby waves) and gravitational acceleration (surface gravity waves, internal gravity waves). Furthermore, we study linear waves and their corresponding wave packets, first studying their behavior in a homogeneous background environment where Fourier methods are available. Thereafter, we introduce methods needed to study linear waves on a gently varying background, including the methods of geometrical optics and wave action, with these methods of use particularly when Fourier methods are not suited.

### Flow instabilities

We study instabilities that arise in geophysical fluid motions, distinguishing two classes of fluid instabilities: [local instabilities](#) or [parcel instabilities](#) versus [global instabilities](#) or [wave instabilities](#). Local instabilities are afforded a local necessary and sufficient condition to determine whether the fluid base state is unstable to perturbations. In contrast, global instabilities arise from the constructive interference of waves and so involve the solution of an eigenvalue problem to determine properties of unstable waves. At most, a necessary condition can be derived to determine whether a global instability exists. Our study of fluid instabilities introduces a suite of case studies that foster analysis and conceptual methods to establish a foundation for further study. Geophysical fluid instability analysis remains an active area of research, with insights into the suite of primary and secondary instabilities providing compelling stories for how the ocean and atmosphere work.



## Pointers on written and spoken communication

To thrive in research and teaching requires one to master elements of both written and spoken communication. Here are a few pointers for the student interested in furthering these skills.

### CLEAR THINKING LEADS TO CLEAR COMMUNICATION

Clear communication is the sign of clear thinking. Some people communicate better in writing, where one has the opportunity to carefully organize thoughts and refine the writing. Others are better at speaking, where spontaneous and interactive reflections and experiences can bolster the clarity of a presentation.

As inspiration for both the clear and obscure, pick up a textbook or lecture notes and analyze the presentation for clarity. Where is the presentation confusing? Where is the material crystal clear? Then pick up a journal article and perform the same analysis. What is appealing? What is unappealing? Then go to the internet and find a science or engineering lecture, old or new. What makes the speaker engaging and clear, or boring and obscure?

### EMPATHY IS KEY

Empathy is a basic facet of effective communication and teaching, where the writer, speaker, or teacher places their mind inside that of an interested and smart reader or listener. Identify with their quest to understand new ideas and to comprehend the foundations and assumptions. Are the assumptions justified based on the audience? How compelling is the scientific story? Are missing steps crucial to understanding or easily dispensed with for streamlining the presentation?

Although poor communication hinders our ability to digest new ideas and concepts, it is also important to appreciate that some material is tough no matter how well it is communicated. We should aim to make a subject matter as simple as possible, but not simpler (paraphrasing Einstein). Furthermore, it sometimes takes a few generations of teaching before some scientific material can be sufficiently digested to allow for the core conceptual nuggets to be revealed. As an example, try reading Newton or Maxwell's original works as compared to a modern presentation of Classical Mechanics or Electromagnetism. So as we strive for clear communication, we cannot presume that clarity is sufficient to remove the struggles everyone experiences when learning.

Additionally, it is essential to recognize that everyone makes mistakes, either in fundamentals or practices. The toughest part of making mistakes is often the self-imposed shame and embarrassment. However, mistakes offer significant opportunities for learning and advancing, with honesty and humility critical for identifying weaknesses, both in our own work and those of others.



## Pointers on physics problem solving

We conclude each chapter with a suite of exercises. Working through these exercises, in full detail, is an integral part of learning and doing physics. Indeed, there is no replacement for struggle and head-scratching to support the physics problem-solving brain muscle. However, with the advent of AI tools, one can readily access AI generated solutions. It goes without saying that over-reliance on AI tools greatly compromises one's ability to develop the skills necessary to know whether the AI solution is correct. In light of this situation, we provide some worked solutions, which generally go beyond what is expected from a student learning the material for the first time, as we aim for the solutions to be instructional as well as utilitarian. For this reason, we expose many of the intermediate steps needed to derive a solution, further supporting an in-depth learning of how to independently solve physics problems. Additionally, it helps to identify when an incorrect result follows from a physical/conceptual error (e.g., incorrect setup of the problem solution) or a mathematical error (e.g., sign error).

We observe that most people are not born with *a priori* physics problem solving skills. Rather, it takes extensive practice to develop the necessary brain muscle. The student who values the ability to solve physics problems should resist the temptation to quickly flip pages to read solutions. Time pondering an exercise is time well spent learning how to do theoretical physics in a manner needed to pursue novel research. In the remainder of this section we offer specific pointers of use when diving into a physics problem.

### CHECK FOR DIMENSIONAL CONSISTENCY

The symbols we use in mathematical physics correspond to geometrical objects (e.g., points, vectors, tensors) describing a physical concept (e.g., position in space, velocity, temperature, angular momentum, stress). Hence, the symbols generally carry physical dimensions. The physical dimensions we are concerned with in this book are length (L), time (T), mass (M), and temperature. We do not consider electromagnetism or the quantum mechanical world. Physical dimensions of the equations must be self-consistent. For example, if one writes an equation  $A = B$ , where  $A$  and  $B$  have different physical dimensions, then the equation makes no physical

---

sense. Something is wrong. Although not always sufficient to uncover errors, dimensional analysis is an incredibly powerful necessary step in debugging the maths.

#### CHECK FOR TENSORIAL CONSISTENCY

In the same way that mathematical equations in physics need to maintain dimensional consistency, they must also respect tensor rules. For example, the equation  $A = B$  makes mathematical sense if  $A$  and  $B$  are both scalars. Likewise,  $\mathbf{A} = \mathbf{B}$  makes sense if  $\mathbf{A}$  and  $\mathbf{B}$  are both vectors. However, if both  $\mathbf{A}$  and  $\mathbf{B}$  are vectors, then the equation  $\mathbf{A} = \nabla \cdot \mathbf{B}$  does not make sense because the left hand side is a vector and the right hand side is a scalar. A more subtle example is when  $\mathbf{A}$  is a vector yet  $\mathbf{B}$  is an axial vector. In this case,  $\mathbf{A}$  remains invariant under a change from right hand to left hand coordinates whereas  $\mathbf{B}$  flips sign. Maintaining basic tensorial rules can be considered the next level of sophistication beyond dimensional analysis.

#### USE WORDS AND PICTURES

Words and pictures are important elements in explaining a physical concept and/or a problem in physics. Hence, it is good practice to liberally sprinkle sentences in between the key equations for the purpose of explaining what the maths means using clear English. Here are some practical payoffs to the student for this style of presentation.

- The process of explaining the maths using words and pictures requires one to dive deeper into the logic of a physics problem. Doing so often reveals weak points, incomplete or unmentioned assumptions, and errors. This process is a very important learning stage in preparing to stand in front of an audience to present results and to answer questions. It is a key facet of research and teaching.
- Physics teachers are often more forgiving of math errors if you convince the teacher that you have a sensible physical understanding of the problem. Plain English and pictures are very useful means for this purpose.

#### THERE IS OFTEN MORE THAN ONE PATH TO A SOLUTION

In physics, there is often more than one path to a problem solution or to the formulation of a concept. Pursuing distinct paths offers novel physical and mathematical insights, exposes otherwise hidden assumptions, and allows one to double-check the veracity of a solution. Some of the most profound advances in physics came from pursuing distinct formulations. One example concerns the distinct formulation of mechanics offered by Newton (1642-1746), and then later by Lagrange (1736-1813) and then Hamilton (1805-1865). Had Lagrange or Hamilton rested on the merits of their predecessors, we may well have had a very different intellectual evolution of 19th and 20th century physics.





# **Part I**

## **Wave mechanics**

## **Waves and the central role of a dispersion relation**

In this part of the book, we study the mechanics of non-dissipative linear waves appearing in geophysical fluids. This study of [wave mechanics](#) forms a key part of geophysical fluid mechanics, in part since so much of the observed geophysical flows manifest waves in one form or another. Waves in fluids manifest as fluctuations of fluid particles and associated fluid properties, with the fluctuations possessing coherency in both space and time. Linear waves exist dynamically due to a restoring force that provides the means for fluid particles to undergo simple harmonic oscillations. Kinematically, linear waves arise when a symmetry is mildly broken, with coherent fluctuations the collective response to the symmetry breaking.

Linear waves have spatial and temporal properties that are closely linked. This linkage relates the angular frequency,  $\omega$ , which measures the temporal structure of a wave, to the wavevector,  $\mathbf{k}$ , which measures the spatial structure of a wave. The [dispersion relation](#) is the mathematical equation that specifies the linkage between  $\omega$  and  $\mathbf{k}$ . Details of the dispersion relation are determined by the physical forces that give rise to the wave, such as compressibility for acoustic waves (Chapter 3), surface tension for capillary waves (Chapter 4), gravitational acceleration for gravity waves (Chapters 4, 7, and 9), the Coriolis acceleration for inertial waves (Chapters 5 and 7) and differential rotation (beta effect) for planetary Rossby waves (Chapters 6 and 7).

Linear waves satisfy the [superposition principle](#) that is generally respected by linear physical phenomena. As a result, interactions between linear waves are reversible, and consist of constructive and destructive interferences. The superposition principle allows for the solution of a linear wave equation to be constructed from constituent elementary pieces. Linear superposition is the fundamental reason we can usefully examine properties of waves by examining the behavior of a single traveling wave. We use this property of linearity when deriving the dispersion relation for waves.

### **Linearization to derive the wave equations**

Derivation of a wave equation and corresponding dispersion relation requires us to linearize the nonlinear governing equations of fluid motion. After deriving the linearized equations, we study their salient properties to help understand the mechanics of their linear wave solutions and dispersion relations. Each wave system has unique properties depending on the underlying dynamical forces. Even so, many kinematical properties transcend details of the dynamics, with kinematics forming the focus for Chapters 1 and 2.

### **Wave energy transport versus matter transport**

The movement of fluid particles is associated with the movement of matter, including tracers, as well as other properties such as momentum, vorticity, enthalpy, entropy, etc. Although waves can lead to a net transfer of matter, through a process known as [Stokes drift](#), the transfer of wave energy and momentum generally occurs without any net movement of fluid particles, where “net movement” refers to a time average over a wave period (i.e., a [phase average](#)). As such, wave energy propagates at speeds that are unconstrained by fluid particle speeds. Indeed, wave energy in linear waves is generally transmitted many times faster than matter. We thus conceive of waves as organized disturbances on a background fluid state (sometimes called the base state or equilibrium state) that transmit energy from one part of the media to another. The speed and direction of wave energy is determined by the dispersion relation, thus making the dispersion relation a central feature of wave mechanics.

---

## The scope of our study

The study of wave mechanics forms a central part of physics, engineering, and applied mathematics. We target the reader aiming to learn wave mechanics for the first time, while also providing selected material for the experienced physicist interested in geophysical fluid wave mechanics. For these purposes, we survey a suite of geophysical fluid waves and dive deep into particular special topics to support understanding and to develop mathematical methods. Our presentation is not comprehensive. Instead, we aim to offer an intellectual platform for further study and research. Insights and skills working with waves requires practice, and we get plenty from the wide variety of geophysical fluid waves considered here.

### MATHEMATICS IN THIS PART

The mathematical methods in this part of the book rely mostly on the Cartesian tensor analysis and vector calculus from VOLUME 1, along with an appreciation for the methods of linear partial differential equations.



# Chapter 1

## PLANE WAVES AND WAVE PACKETS

Forces acting in a fluid determine the dispersion relation satisfied by linear waves, with the dispersion relation connecting the wave angular frequency,  $\omega$ , to the wavevector,  $\mathbf{k}$ . We write this relation in the functional form<sup>1</sup>

$$\omega = \varpi(\mathbf{k}), \quad (1.1)$$

where  $\varpi$  is generally a nonlinear function of the wavevector. The dispersion relation couples the space and time structure of a wave. Hence, once a wavevector is chosen then the angular frequency is determined by evaluating the dispersion relation. In this chapter we develop the foundations for wave kinematics, which is concerned with wave properties arising from the existence of a dispersion relation yet is unconcerned with the forces that determine this relation.

The wave function plays a fundamental role in describing waves. Example wave functions include the velocity potential (acoustic and surface waves), the streamfunction (Rossby waves), and free surface height (shallow water waves). All other dynamical fields can be generated from the wave function, thus allowing us to focus on characterizing how the wave appears through study of the wave function. The wave function has both an  $x$ -space (geographic/height space) representation as well as a  $\mathbf{k}$ -space (wavevector space) representation. These two representations offer complementary characterizations of wave properties, with the transformation between these representations provided by Fourier's integral theorem (see VOLUME 1). One of the more remarkable aspects of this complementarity arises through the uncertainty relation, which states that it is not possible to simultaneously specify the position of a wave packet with arbitrary precision in both  $x$ -space and  $\mathbf{k}$ -space.

### READER'S GUIDE TO THIS CHAPTER

We focus on plane waves in this chapter and so make use of Cartesian coordinates and Cartesian tensors from VOLUME 1. We restrict attention to a homogeneous background/base state upon which waves are supported, and for which we can make use of Fourier analysis methods, also detailed in VOLUME 1. The presentation shares much with books and reviews covering topics in linear wave mechanics, such as [Bretherton \(1971\)](#), chapter 3 of [Acheson \(1990\)](#), chapters 1 and 2 of [Pedlosky \(2003\)](#), chapter 6 of [Olbers et al. \(2012\)](#), chapter 1 of [Sutherland \(2010\)](#), chapter 7 of [Thorne and Blandford \(2017\)](#), and chapter 6 of [Vallis \(2017\)](#). However, be mindful that not all authors agree on the conventions followed here, which are summarized in Section 1.2.

<sup>1</sup>For waves moving on a space and time dependent background, the dispersion relation picks up space and time dependence. This added complexity is considered in Chapter 2.

<b>1.1</b>	<b>Loose threads</b>	<b>6</b>
<b>1.2</b>	<b>Nomenclature and conventions</b>	<b>7</b>
1.2.1	Types of waves	7
1.2.2	Types of dispersion relations	8
1.2.3	Conventions for space-time description of waves	8
<b>1.3</b>	<b>Two classes of linear waves</b>	<b>9</b>
1.3.1	Non-dispersive waves	9
1.3.2	Dispersive waves	10
1.3.3	Comments about time derivatives	10
<b>1.4</b>	<b>Monochromatic patterns and Fourier analysis</b>	<b>11</b>
<b>1.5</b>	<b>Free plane waves</b>	<b>12</b>
1.5.1	Traveling plane wave	12
1.5.2	Characterizing the wave	13
1.5.3	Wavenumber and factors of $2\pi$	15
1.5.4	Superposition of plane waves	16
1.5.5	Wave ansatz	17
1.5.6	Summary of kinematic wave properties	17
<b>1.6</b>	<b>Free wave packets</b>	<b>17</b>
1.6.1	Phase velocity, group velocity, and the dispersion relation	18
1.6.2	A continuum of traveling plane waves	19
1.6.3	The wave function is real	20
1.6.4	Initializing the wave packet	20
1.6.5	The case of non-dispersive waves	22
1.6.6	Wave trains and wave packets	23
1.6.7	PDE for the wave packet modulation function	24
1.6.8	Wave packets of non-dispersive waves	25
1.6.9	Wave function PDE implied by the dispersion relation	26
1.6.10	Standing wave packets	27
<b>1.7</b>	<b>Wave packets in one-dimension</b>	<b>28</b>
1.7.1	The positive wave packet	28
1.7.2	The negative wave packet	28
1.7.3	The positive-negative wave packet	28
1.7.4	The positive-positive wave packet	29
1.7.5	Initial uncertainty relation for a Gaussian packet	29
1.7.6	Extreme examples of the uncertainty relation	31
1.7.7	Evolution of a non-dispersive Gaussian wave packet	32
1.7.8	Evolution of a dispersive Gaussian wave packet	34
1.7.9	The non-dispersive limit of a dispersive packet	36
1.7.10	Comments and further study	37
<b>1.8</b>	<b>Method of stationary phase</b>	<b>37</b>
1.8.1	Riemann-Legesque lemma and center of the packet	38
1.8.2	Wavenumber intervals with no phase extrema	39
1.8.3	Wavenumber interval including a phase extrema	39
1.8.4	Comments and further study	41
<b>1.9</b>	<b>Exercises</b>	<b>41</b>

---

## 1.1 Loose threads

- Figures
- More on wave packets and Green's functions. Derive the Green's function equation from the packet equation plus the dispersion relation.

## 1.2 Nomenclature and conventions

Many readers have encountered wave kinematics in prior studies. Yet some of the conventions are not universal. We here summarize the choices made in this book.

### 1.2.1 Types of waves

Table 1.1 summarizes the various types of elementary wave patterns considered in this book. We mostly encounter plane traveling waves, though also make use of standing waves (as for waves within a bounded domain) and stationary waves (when a wave has only spatial periodicity and no time dependence). The wave patterns in Table 1.1 can arise from any conceivable context, physical or otherwise (e.g., the wave produced by spectators in a large sporting event). The specifics of a physical system arises when enforcing that a wave function,  $\Phi$ , satisfies a linear partial differential equation, and it is through this equation that we derive the dispersion relation.

NAME	SPACE-TIME STRUCTURE
monochromatic	$\text{Re}[\Phi_0 e^{-i\omega t}]$
traveling	$\text{Re}[\mathcal{A}(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}]$
standing	$\text{Re}[\mathcal{A}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}] \cos(\omega t)$
stationary	$\text{Re}[\mathcal{A}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}]$

TABLE 1.1: Summarizing the variety of waves encountered in this book, with distinctions based on their space-time structure. The operator,  $\text{Re}$ , extracts the real part of its argument, so that, for example,  $\text{Re}[e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}] = \cos(\mathbf{k}\cdot\mathbf{x} - \omega t)$ . The first three patterns are monochromatic, which means they oscillate coherently with a single period,  $2\pi/\omega$ , where  $\omega \geq 0$  is the angular frequency. However, note that  $\text{Re}[\Phi_0 e^{-i\omega t}]$  is monochromatic and yet we cannot conclude whether it is a wave until we have further information about its spatial structure. Spatial dependence is introduced in the form  $e^{i\mathbf{k}\cdot\mathbf{x}}$ , where  $\mathbf{k}$  is the wavevector that determines the direction of the wave, and its magnitude, the wavenumber, determines its wavelength,  $2\pi/|\mathbf{k}|$ . For a plane wave in three-dimensional space,  $\mathbf{k}\cdot\mathbf{x} = k_x x + k_y y + k_z z$ , whereas a plane wave in two dimensions has wavevector  $\mathbf{k}\cdot\mathbf{x} = k_x x + k_y y$ . Standing waves do not travel in space but oscillate in place. A stationary wave has an oscillatory spatial pattern but has no time dependence and so is static. The final three waves are spatially **harmonic waves** since they have a regular spatial pattern.

There are many mathematical functions that give rise to periodic patterns. These patterns correspond to physical waves when they satisfy a wave equation that admits a dispersion relation thus connecting the space and time structure of the pattern. Evidently, there are more mathematical wave-like patterns than there are physical waves. For example, consider the function

$$\Phi(x, y, t) = A \cos(k_x x - \omega t) \cos(\omega t) = (A/2) [\cos(k_x x - 2\omega t) + \cos(k_x x)], \quad (1.2)$$

where  $A$  is a constant amplitude. The second expression reveals the sum of a traveling plane wave plus a stationary wave. It is notable that this function is not in the form of a Fourier component that constitutes the wave patterns in Table 1.1. Rather, it is simply the sum of two trigonometric functions having a particular space and time structure. Is this function a particular solution to the classic wave equation with operator  $\partial_{tt} - c^2 \partial_{xx}$ ? A quick calculation reveals that it is not a solution since

$$\partial_{tt}\Phi = -2A\omega^2 \cos(k_x x - 2\omega t) \quad (1.3a)$$

$$\partial_{xx}\Phi = -k_x^2 \Phi = -(k_x^2 A/2)[\cos(k_x x - 2\omega t) + \cos(k_x x)], \quad (1.3b)$$

revealing that there is no way to relate  $\partial_{tt}\Phi$  and  $c^2 \partial_{xx}\Phi$  at arbitrary space and time points.

### 1.2.2 Types of dispersion relations

The dispersion relation (1.1) is the root of all wave properties since it links the space and time structure of the wave. Some waves encountered in this book have a dispersion relation in the slightly more restricted form

$$\omega = \varpi(|\mathbf{k}|), \quad (1.4)$$

where  $|\mathbf{k}|$  is the **wavenumber**. Evidently, these waves have their angular frequency independent of the wave direction. Examples include acoustic waves (Chapter 3) and interfacial waves (Chapter 4). Inertial waves from Chapter 5 and internal gravity waves from Chapter 9 are the complement, with their dispersion relation independent of the wavenumber yet dependent on the wave direction,

$$\omega = \varpi(\hat{\mathbf{k}}) \quad \text{with} \quad \hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|. \quad (1.5)$$

Finally, the Rossby waves introduced in Chapters 6 and 7 have a dispersion relation that depends on both the wavenumber and wave direction; i.e., on the wavevector,  $\mathbf{k}$ .

### 1.2.3 Conventions for space-time description of waves

The suite of dispersion relations encountered in geophysical fluids motivates the following conventions that form the basis for our space-time description of waves. Readers should be aware that certain of these conventions are not universally followed in the literature.

- **ANGULAR FREQUENCY:** Following [Bretherton \(1971\)](#), the **angular frequency** of a stable linear wave that results from the dispersion relation,  $\omega = \varpi(\mathbf{k})$ , is chosen as a non-negative real number,

$$\omega \geq 0. \quad (1.6)$$

Consequently, the wave period,  $2\pi/\omega$ , is also a non-negative number and the wave advances forward in time with the wave direction determined by the wavevector.

This convention is not universally followed in the literature. One reason to allow for a negative angular frequency is that dispersion relations typically lead to quadratic equations that have a positive and a negative root. For example, the one-dimensional acoustic wave equation has

$$\omega^2(\mathbf{k}) = c_s^2 \mathbf{k}^2, \quad (1.7)$$

which has roots  $\omega = \pm c_s |\mathbf{k}|$  ( $c_s$  is the acoustic wave speed; Section 3.5.1). When choosing both roots, one interprets them as representing oppositely moving waves. Even so, we follow a convention that always chooses the positive root, so that the wave direction is determined by the wavevector (see next point) rather than the sign of the angular frequency. Additionally, we are not generally interested in waves traveling backward in time, which is sometimes used to interpret waves with  $\omega < 0$ .

One encounters a negative angular frequency when conducting time-frequency Fourier analysis in the complex plane (see VOLUME 1). In that context, the negative frequency is not physical. Rather, it is a mathematical expedient that allows us to work with a complex Fourier transform. There is nothing physical implied by the negative frequency in this context. Indeed, one could equivalently work with real Fourier sine and cosine transforms where the angular frequency is non-negative. Furthermore, if using complex

Fourier analysis, then switching the sign of the angular frequency and the wavevector merely leads to the complex conjugate wave function. Since the physical wave is real in classical physics, the complex conjugate wave function is the same wave.

- **WAVE DIRECTION:** The direction of a wave is specified by the wavevector,  $\mathbf{k}$  (Section 1.5), and its wavelength is  $2\pi/|\mathbf{k}|$ , where  $|\mathbf{k}| \geq 0$  is the wavenumber.
- **PHASE SPEED:** The **phase speed**,  $C_p$ , is a non-negative number. It is not a vector so that we do not consider components to the phase speed. Rather, we consider components to the phase velocity vector, which is a vector whose direction is given by the wavevector and whose non-negative magnitude, along with the angular frequency, determine the phase speed (Section 1.5.2).

## 1.3 Two classes of linear waves

We distinguish two types of linear wave phenomena: non-dispersive waves and dispersive waves. In this section we present a mathematical overview of these waves, offering examples of each type of wave along with pointers to the physics and maths to come later.<sup>2</sup>

### 1.3.1 Non-dispersive waves

Canonical **non-dispersive waves** are described by “the” wave equation

$$(\partial_{tt} - c^2 \nabla^2) \Phi = 0, \quad (1.8)$$

where  $\Phi$  is a wave function (e.g., velocity potential, streamfunction, surface height), and  $c^2 > 0$  is the squared wave speed. As discussed when studying partial differential equations in VOLUME 1, the wave equation (1.8) is a **hyperbolic partial differential equation** and its solutions all transmit signals with speed,  $c > 0$ , without distortion. That is, waves satisfying equation (1.8) are **non-dispersive waves** so that arbitrary wavelengths travel at the same speed, so there is no mixing across the spectrum of wavevectors. We find that for non-dispersive waves the angular frequency divided by the wavenumber equals to a constant phase speed shared by all waves. The acoustic waves of Chapter 3 and the shallow water gravity waves of Chapter 7 are example non-dispersive waves considered in this book.

To expose some of the mathematical details, consider the wave equation (1.8) in one space dimension and written in the factored form

$$(\partial_t - c \partial_x) (\partial_t + c \partial_x) \Phi = 0. \quad (1.9)$$

As discussed in VOLUME 1, the solution to this hyperbolic partial differential equation takes the form of two signals moving in opposite directions

$$\Phi(x, t) = F(x - ct) + G(x + ct), \quad (1.10)$$

where the functions  $F$  and  $G$  are specified by boundary and initial conditions. Their respective functional dependencies,  $x \pm ct$ , determine the propagation direction of the signal, with  $F(x - ct)$  a signal propagating in the  $+\hat{x}$  direction whereas  $G(x + ct)$  is a signal propagating in the  $-\hat{x}$

---

<sup>2</sup>Chapter 1 of [Whitham \(1974\)](#) provides a thorough mathematical introduction to the variety of linear and nonlinear waves.

direction. A key point is that the signals all move with the same speed, with  $F$  and  $G$  arbitrary functions that may be built from a single linear wave or any number of waves (including a wave continuum). Also note that the wave equation (1.9) has been factored into the product of two linear operators, so that a simpler hyperbolic wave equation is either of the two first order equations

$$(\partial_t \pm c \partial_x) \Phi = 0. \quad (1.11)$$

In other parts of this book this equation is referred to as the **advection equation** for a one-dimensional signal with constant advection velocity.

It is common to be introduced to wave physics through solutions to the hyperbolic wave equations (1.8) or (1.11). Such waves render an intuition based on the non-dispersive nature of these waves. For example, when listening to a symphony, we hear the sounds from various instruments “in concert”. That is, the speed that the acoustic waves propagate are independent of the wavelength. Indeed, imagine how difficult it would be for an audience spread throughout a concert hall to equally appreciate a symphonic performance if acoustic waves were dispersive!

### 1.3.2 Dispersive waves

Upon studying geophysical waves, we soon realize that non-dispersive waves are the exception rather than the norm. That is, most waves in geophysical fluids are **dispersive waves**, with their governing wave equations sometimes mathematically classified as hyperbolic though many are not. A dispersive wave is one that admits solutions of the form

$$\Phi(\mathbf{x}, t) = A e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} = A e^{i\mathbf{k} \cdot (\mathbf{x} - \hat{\mathbf{k}} C_p t)} \quad (1.12)$$

where the speed of a surface with constant phase,

$$C_p = \varpi(\mathbf{k})/|\mathbf{k}|, \quad (1.13)$$

is a function of the wavevector. As such, the movement of the dispersive wave signal (i.e., the **phase velocity**) depends on the wavelength of the wave. For example, if we prepare a packet of dispersive waves with a particular shape in space, the evolution of the packet is characterized by the dispersion of the packet.

Waves visible by eye on the surface of a pond or surface of the ocean offer a canonical example of dispersive waves. As explored in Section 4.5.3, long surface gravity waves (which also have low frequency) travel faster than short and high frequency surface gravity waves. Indeed, it is this dispersion for surface gravity waves that surfers appreciate when assessing the suitability of coastal surf conditions following an offshore storm.<sup>3</sup>

### 1.3.3 Comments about time derivatives

As noted when discussing the non-dispersive wave equation (1.9), the presence of two time derivatives provides the means for two wave signals to arise from solutions to this equation, where the waves move in opposite directions. This case occurs for acoustic waves, gravity waves, and inertial waves. It manifests as a dispersion relation written for the squared angular frequency. In contrast, when there is only one time derivative in the wave equation, then the dispersion relation is written for the angular frequency rather than its square. Furthermore, the system chooses a preferred direction for the wave signal, with **Rossby waves** offering the

---

<sup>3</sup>See [Butt et al. \(2004\)](#) for a scientific view on surface ocean waves geared towards surfing.

canonical example of such anisotropic waves. Rossby waves also offer a clear example of how the non-negative angular frequency convention (Section 1.2.3) helps to understand the phase velocity.

## 1.4 Monochromatic patterns and Fourier analysis

A [monochromatic pattern](#) is a feature characterized by having all points maintain the same time periodic motion with single angular frequency,  $\omega$ . Hence, all points within a [monochromatic wave](#) have the same angular frequency. The trigonometric functions,  $\cos(\omega t)$  and  $\sin(\omega t)$ , both exhibit time periodic behavior, which explains the factors of  $2\pi$  appearing throughout wave physics. As noted in Section 1.2.3, we choose the convention in which the angular frequency is a non-negative number,  $\omega \geq 0$ , so that the period of the oscillation is  $2\pi/\omega$ .

It is often convenient to make use of the [Euler identity](#)

$$e^{-i\omega t} = \cos(\omega t) - i \sin(\omega t), \quad (1.14)$$

to thus write a time periodic [wave function](#) as<sup>4</sup>

$$\Phi = \text{Re}[\Phi_0 e^{-i\omega t}], \quad (1.15)$$

where  $\Phi_0$  is a complex function further specified below, and  $\text{Re}[\cdot]$  is an operator that returns the real part of its argument.<sup>5</sup> For linear calculations we can safely omit the  $\text{Re}$  operation until the end of the calculations. However, for products of wave fields, such as when performing energetic analyses, more care is needed, with details given in the Fourier analysis chapter of VOLUME 1.<sup>6</sup>

Monochromatic space-time wave patterns, such as those modeled by trigonometric functions, are not physically realizable since all physical patterns encounter inhomogeneities in space and/or time. Nonetheless, the following quote from page 424 of *Elmore and Heald (1969)* serves as a concise and compelling motivation for the study of such patterns through their role in Fourier analysis.

The possibility of spectrum analysis and spectrum synthesis by Fourier series and Fourier integrals again reminds us that the theory of steady sinusoidal waves, which never end in space or time, is far more than an idealized special case. Using the concepts of Fourier analysis and synthesis, the theory of continuous waves constitutes the basis for the description of waves of arbitrary shape without further reference to the wave equation or to particularly non-sinusoidal solutions that it may have. The only basic requirement is that the wave equation be linear, so that the principle of superposition applies. The technique, of course, can be used even though the wave velocity is frequency dependent; i.e., the wave medium is dispersive.

---

<sup>4</sup>The minus sign is motivated by convenience for when we consider a traveling wave and the phase velocity in Section 1.5.2. See in particular Figure 1.1.

<sup>5</sup>We summarize salient points concerning complex numbers in the Fourier analysis chapter in VOLUME 1.

<sup>6</sup>As noted when developing the basics of Fourier analysis in VOLUME 1, we use complex variables solely for mathematical convenience, with all physical fields producing real numbers in this book.

## 1.5 Free plane waves

Plane waves are characterized by a single wavevector and single angular frequency, with the wave exhibiting symmetry in directions perpendicular to the propagation direction (hence the “plane” in its name). Any linear wave, be it acoustic, gravity, Rossby, etc., can be decomposed into a sum or integral of plane waves with a suite of frequencies and wavevectors, and with modulation by an amplitude function.

The general form of a stationary plane wave function is given by

$$\Phi_o = \mathcal{A}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (1.16)$$

where  $\mathcal{A}(\mathbf{k})$  is a complex wave amplitude and  $\mathbf{k}$  is the [wavevector](#) or [wavenumber vector](#). The wavevector has dimensions of inverse length and it characterizes the wavelength of the wave as well as the spatial direction of the wave propagation (as per travelling plane waves in Section 1.5.1). The magnitude of the wavevector,  $|\mathbf{k}|$ , is referred to as the [wavenumber](#) as it measures  $2\pi$  times the number of waves per unit length (we further discuss this notion in Section 1.5.2). We allow the wave amplitude,  $\mathcal{A}(\mathbf{k})$ , to be a function of wavevector so that different plane waves can have distinct amplitudes. We also allow for  $\mathcal{A}(\mathbf{k})$  to be complex so that different plane waves can have their phases shifted relative to one another (we discuss phases in Section 1.5.1).

Given that the plane wave possesses symmetry in planes perpendicular to  $\mathbf{k}$ , it is spatially dependent only on the direction parallel to the wavevector,<sup>7</sup>

$$\mathbf{k} = |\mathbf{k}| \hat{\mathbf{k}}. \quad (1.17)$$

We say that the plane wave function (1.16) is free since it exists throughout all of space and is unaffected by boundaries. Furthermore, we are uninterested in how the plane wave was formed. Indeed, since it is present throughout all of space, we imagine that the plane wave has existed for all time. Clearly these properties are not physically realizable. Even so, as noted in Section 1.4, the mathematical simplicity of the plane wave, along with the superposition principle, affords it a central role in the study of wave mechanics. Furthermore, through the construction of wave packets as studied in Sections 1.6 and 1.7, which are synthesized via the superposition of plane waves, we can build wave forms that are spatially and temporally localized and are thus physically realizable.

### 1.5.1 Traveling plane wave

Allowing for time to evolve renders the traveling monochromatic plane wave function

$$\Phi = \text{Re}[\Phi_o e^{-i\omega t}] = \text{Re}[\mathcal{A}(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}] = |\mathcal{A}(\mathbf{k})| \text{Re}[e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t+\alpha)}]. \quad (1.18)$$

We here wrote complex wave amplitude according to a real amplitude and a phase shift

$$\mathcal{A}(\mathbf{k}) = |\mathcal{A}(\mathbf{k})| e^{i\alpha} \quad \text{with} \quad \tan \alpha = \text{Im}[\mathcal{A}(\mathbf{k})]/\text{Re}[\mathcal{A}(\mathbf{k})]. \quad (1.19)$$

The traveling plane wave [phase](#) function is given by

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t + \alpha, \quad (1.20)$$

---

<sup>7</sup>Caution: many authors write  $\hat{\mathbf{k}}$  for the unit vertical direction, whereas in this book we write  $\hat{\mathbf{z}}$  for the unit vertical direction whereas  $\hat{\mathbf{k}}$  is the direction of a wave.

so that

$$\Phi = |\mathcal{A}(\mathbf{k})| \operatorname{Re}[e^{i\mathcal{P}}] = |\mathcal{A}(\mathbf{k})| \cos \mathcal{P}. \quad (1.21)$$

The traveling plane wave function,  $\Phi$ , depends on the wavevector as well as the space and time point. It is also a function of the angular frequency,  $\omega$ . However, recall that for the study of waves realized by a physical system, once the wavevector is known then the angular frequency is specified through the dispersion relation (1.1).

The phase factor on the wave amplitude in equation (1.19) is generally a function of wavevector,  $\alpha = \alpha(\mathbf{k})$ . However, for most of our applications it is taken as a constant, in which case it is referred to as the standard phase.

### 1.5.2 Characterizing the wave

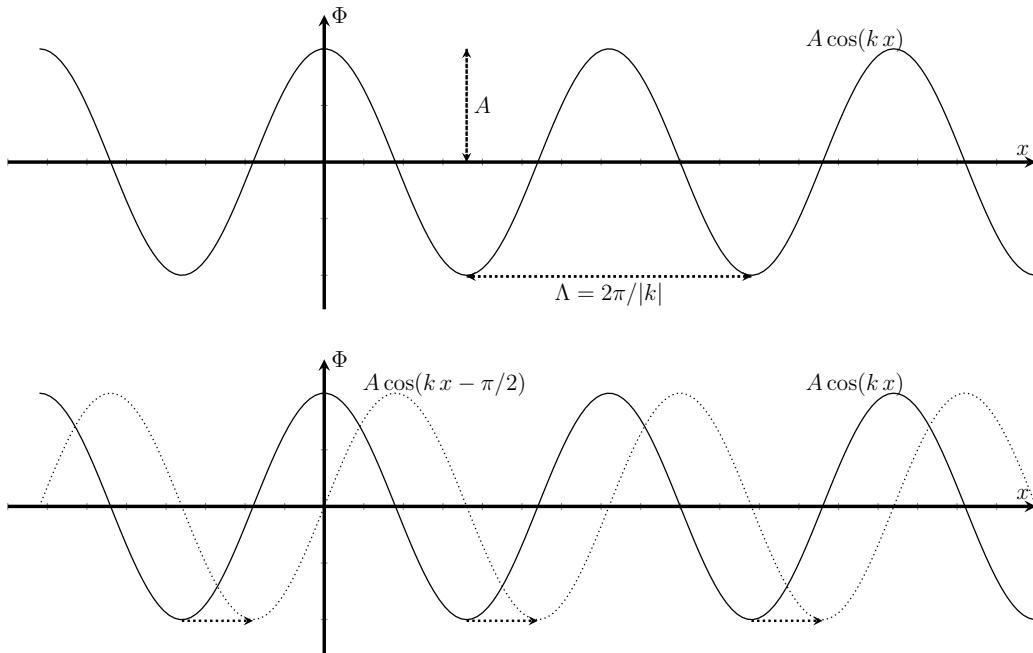


FIGURE 1.1: Top panel: a cosine wave,  $\Phi = A \cos(kx - \omega t)$ , along the  $x$ -axis at  $t = 0$ , with wavelength  $\Lambda = 2\pi/|k|$ , wavevector  $\mathbf{k} = k\hat{x}$ , and real amplitude,  $A$ . Lower panel: two snapshots of a traveling cosine wave, one shown at  $t = 0$  as in the top panel, and another shown a quarter period later, at  $\omega t = \pi/2$ . With  $k > 0$  and  $\omega t > 0$ , as shown in this figure, a point with constant phase moves in the  $+\hat{x}$  direction. Correspondingly, as the phase becomes more negative (as time increases), the wave form moves to the right, in the direction of the wave propagation. If  $k < 0$  then the phase propagates in the  $-\hat{x}$  direction.

#### Wave period

The wave function (1.18) for the traveling plane wave takes on the same value for all space and time points with a phase,  $\mathcal{P}$ , that is shifted by any integer multiple of  $2\pi$ . By fixing a point in space, the plane wave is identical for all times,  $t_n$ , satisfying

$$\omega t_n = \omega t + 2\pi n = \omega(t + 2\pi n/\omega). \quad (1.22)$$

Hence, the wave period is given by

$$\text{wave period} = 2\pi/\omega. \quad (1.23)$$

### Wavelength

By fixing time, we see that the plane wave is identical for all space points,  $\mathbf{x}_n$ , satisfying

$$\mathbf{k} \cdot \mathbf{x}_n = \mathbf{k} \cdot \mathbf{x} + 2\pi n = |\mathbf{k}|(\mathbf{x} \cdot \hat{\mathbf{k}} + 2\pi n/|\mathbf{k}|), \quad (1.24)$$

which allows us to identify  $2\pi/|\mathbf{k}|$  as the **wave length**, which we write as<sup>8</sup>

$$\Lambda = 2\pi/|\mathbf{k}| \implies \mathbf{k} = 2\pi \hat{\mathbf{k}}/\Lambda. \quad (1.25)$$

Figure 1.1 provides an example cosine wave in one-dimension, illustrating the wavelength and traveling nature of the wave.

### Phase velocity and phase speed

When observing a traveling plane wave from a fixed point in space, one wavelength passes by the point within a single wave period. We refer to the **phase velocity**,  $\mathbf{c}_p$ , as the velocity of a point fixed on a phase surface and traveling in the direction of the wavevector. Its magnitude is the **phase speed**,  $C_p$ , which is the speed that the phase moves in the direction of the wavevector,

$$\mathbf{c}_p = C_p \hat{\mathbf{k}} = (\omega/|\mathbf{k}|) \hat{\mathbf{k}} \implies C_p = \mathbf{c}_p \cdot \hat{\mathbf{k}} = (2\pi/|\mathbf{k}|)/(2\pi/\omega) = \omega/|\mathbf{k}| \geq 0, \quad (1.26)$$

which allows us to write the angular frequency as

$$\omega = \mathbf{c}_p \cdot \mathbf{k} = C_p |\mathbf{k}|. \quad (1.27)$$

Figure 1.2 provides an illustration of the phase velocity and phase speed.

An equivalent means to understand the phase velocity is to consider the time derivative,  $D_p/Dt$ , defined to measure time changes within the reference frame of a point on a constant phase surface

$$\frac{D_p \mathcal{P}}{Dt} \equiv \partial_t + \mathbf{c}_p \cdot \nabla, \quad (1.28)$$

in which case

$$\frac{D_p \mathcal{P}}{Dt} = (\partial_t + \mathbf{c}_p \cdot \nabla)(\mathbf{x} \cdot \mathbf{k} - \omega t) = \mathbf{c}_p \cdot \mathbf{k} - \omega = 0. \quad (1.29)$$

We are thus led to the expression for the phase speed and angular frequency

$$C_p = \hat{\mathbf{k}} \cdot \frac{D_p \mathbf{x}}{Dt} \quad \text{and} \quad \omega = \mathbf{c}_p \cdot \mathbf{k} = \mathbf{k} \cdot \frac{D_p \mathbf{x}}{Dt}. \quad (1.30)$$

The phase time derivative (1.28) is directly analogous to the **material time derivative**,

$$\frac{D}{Dt} = \partial_t + \mathbf{v} \cdot \nabla, \quad (1.31)$$

defined in the reference frame following the velocity of a fluid particle. Likewise, in Section

<sup>8</sup>We choose the uppercase,  $\Lambda$ , to denote wavelength rather than the more commonly used  $\lambda$ . We do so in order to distinguish the wavelength,  $\Lambda$ , from the longitude,  $\lambda$ , that commonly appears in this book.

2.3.2 we make use a time derivative following the group velocity,  $\mathbf{c}_g$ , in which case  $\mathbf{c}_p$  in equation (1.28) is replaced by  $\mathbf{c}_g$ .

### Phase distance

We can define the phase speed in another manner by introducing the projection of the position vector along the direction of the wavevector

$$S_{\hat{\mathbf{k}}} \equiv \hat{\mathbf{k}} \cdot \mathbf{x}, \quad (1.32)$$

which we refer to as the phase distance. Note that  $S_{\hat{\mathbf{k}}}$  can be positive or negative, depending on the direction of the wavevector. The wavelength as defined by equation (1.24) takes on the form

$$\mathbf{k} \cdot \mathbf{x}_n = \mathbf{k} \cdot \mathbf{x} + 2\pi n = |\mathbf{k}|(S_{\hat{\mathbf{k}}} + 2\pi n/|\mathbf{k}|). \quad (1.33)$$

As such, the phase speed is the change in the position of the constant phase lines in the direction of the wavevector, so that

$$C_p = \left[ \frac{\partial S_{\hat{\mathbf{k}}}}{\partial t} \right]_{\mathcal{P}} = -\frac{\partial \mathcal{P}/\partial t}{\partial \mathcal{P}/\partial S_{\hat{\mathbf{k}}}} = \omega/|\mathbf{k}|. \quad (1.34)$$

For this equation we made use of some basic mathematics of [generalized vertical coordinates](#) from [VOLUME 4](#). Finally, this notation allows us to write the phase function as

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t + \alpha = |\mathbf{k}|(S_{\hat{\mathbf{k}}} - C_p t) + \alpha. \quad (1.35)$$

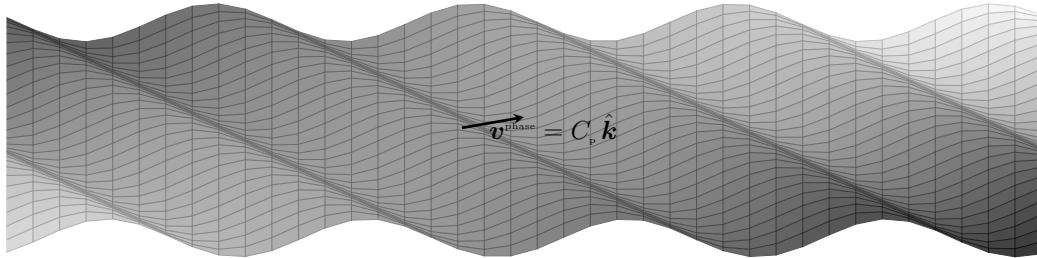


FIGURE 1.2: A mesh plot of a plane wave with phase velocity,  $\mathbf{c}_p = C_p \hat{\mathbf{k}} = (\omega/|\mathbf{k}|) \hat{\mathbf{k}}$ , pointing perpendicular to surfaces of constant phase.

### 1.5.3 Wavenumber and factors of $2\pi$

The magnitude of the wavevector,  $|\mathbf{k}|$ , appears in many places in wave kinematics, such as in the definition (1.25) of the wavelength. As a shorthand terminology, we refer to  $|\mathbf{k}|$  as the **wavenumber**, with the wavenumber related to the wavelength,  $\Lambda$ , via

$$|\mathbf{k}| = 2\pi/\Lambda. \quad (1.36)$$

The wavenumber measures  $2\pi$  times the number of waves per unit length, so that the wavenumber measures the spatial angular frequency of a wave. In some contexts it is useful to introduce the

reduced wavelength

$$\lambda = \Lambda/2\pi, \quad (1.37)$$

so that the wavenumber is given by the inverse reduced wavelength

$$|\mathbf{k}| = \lambda^{-1}. \quad (1.38)$$

Alternatively, it is sometimes convenient to work with the reduced wavenumber

$$\tilde{k} = |\mathbf{k}|/2\pi. \quad (1.39)$$

Hence, the reduced wavenumber is the inverse wavelength

$$\tilde{k} = 1/\Lambda \quad (1.40)$$

so that  $\tilde{k}$  measures the number of waves per unit length.

In our discussion of Fourier analysis in VOLUME 1, we raised concerns about the variety of conventions for placing  $2\pi$  factors. We further those concerns by noting that the  $2\pi$  factor associated with wavenumbers can be unclear in the literature. The ambiguity occurs when a particular length scale,  $L$ , is considered, yet it is not specified whether the length scale refers to a wavelength, in which case the corresponding wavenumber is  $|\mathbf{k}| = 2\pi/L$ , or to a reduced wavelength, in which case the corresponding wavenumber is  $|\mathbf{k}| = 1/L$ . The bottomline is that care is necessary to ensure the proper usage.

#### 1.5.4 Superposition of plane waves

Through the principle of superposition respected by linear waves, realistic linear waves can be decomposed into the sum or integral of modulated plane waves. Indeed, this decomposition forms the mathematical basis for Fourier analysis detailed in VOLUME 1. We here offer an illustration of this decomposition by considering how two plane waves combine, and then how three plane waves combine.

Consider the sum of two plane waves, each with the same real amplitude of unity (in arbitrary units), and with distinct (but close) wavevectors and angular frequencies

$$\mathbf{k}_\pm = \mathbf{k} \pm \Delta\mathbf{k} \quad \text{and} \quad \omega_\pm = \omega \pm \Delta\omega \quad \text{with} \quad |\Delta\mathbf{k}|/|\mathbf{k}| \ll 1 \quad \text{and} \quad |\Delta\omega/\omega| \ll 1. \quad (1.41)$$

The resulting superposition of the two waves is given by

$$\text{Re}[e^{i(\mathbf{x} \cdot \mathbf{k}_+ - t\omega_+)} + e^{i(\mathbf{x} \cdot \mathbf{k}_- - t\omega_-)}] = 2 \cos(\Delta\mathbf{k} \cdot \mathbf{x} - \Delta\omega t) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (1.42)$$

The  $\cos(\Delta\mathbf{k} \cdot \mathbf{x} - \Delta\omega t)$  factor acts as a low wavenumber and low frequency modulation of the second factor,  $\cos(\mathbf{k} \cdot \mathbf{x} - \omega t)$ , which is much more rapidly varying in space and time. We illustrate this superposition of two plane waves in Figure 1.3 for time  $t = 0$ . Adding a third wave, also with unit amplitude,  $\exp[i(\mathbf{k} \cdot \mathbf{x} - \omega t)]$ , renders the superposition

$$\begin{aligned} \text{Re}[e^{i(\mathbf{x} \cdot \mathbf{k}_+ - t\omega_+)} + e^{i(\mathbf{x} \cdot \mathbf{k} - t\omega)} + e^{i(\mathbf{x} \cdot \mathbf{k}_- - t\omega_-)}] \\ = 4 \cos^2[(\Delta\mathbf{k} \cdot \mathbf{x} - \Delta\omega t)/2] \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \end{aligned} \quad (1.43)$$

The third wave serves to double the amplitude through constructive interference, while also broadening the width of the low amplitude destructive interference region. Additionally, the

modulation function,  $\cos^2[(\Delta\mathbf{k} \cdot \mathbf{x} - \Delta\omega t)/2]$ , is rendered non-negative when adding the third wave. We also illustrate this superposition in Figure 1.3.

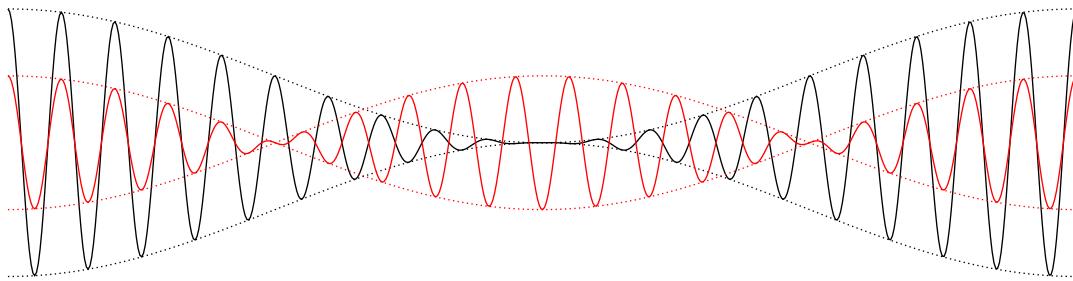


FIGURE 1.3: Superposition of two plane waves (red lines; equation (1.42)) and three plane waves (black lines; equation (1.43)), with all plane waves having equal real amplitudes and here shown for time  $t = 0$ . Notice how in both cases, a high wavenumber wave (short wave) is modulated by a low wavenumber wave (long wave). The units in this figure are arbitrary.

### 1.5.5 Wave ansatz

The principle of superposition allows us to study properties of a linear wave equation, in particular the dispersion relation, by inserting a single plane wave into the wave equation. In this manner we refer to the plane wave as an **ansatz**, which is a German word meaning educated guess, trial form, or assumed functional form.

### 1.5.6 Summary of kinematic wave properties

In Table 1.2, we summarize the variety of kinematic properties of waves. In this book we work mostly with the angular frequency,  $\omega$ , the wavevector,  $\mathbf{k}$ , and the wavenumber,  $|\mathbf{k}|$ . Even so, it is useful to be versed in the alternative choices.

NAME	MATH SYMBOL	DIMENSIONS	MATH RELATION
period	$\tau_{\text{period}}$	T	$\tau_{\text{period}} = 2\pi/\omega$
angular frequency	$\omega$	$\text{T}^{-1}$	$\omega = 2\pi/\tau_{\text{period}}$
frequency	$f$	$\text{T}^{-1}$	$f = \omega/2\pi = 1/\tau_{\text{period}}$
wavenumber	$ \mathbf{k} $	$\text{L}^{-1}$	$ \mathbf{k}  = 2\pi/\Lambda$
wavevector	$\mathbf{k}$	$\text{L}^{-1}$	$\mathbf{k} =  \mathbf{k}  \hat{\mathbf{k}}$
reduced wavevector	$\tilde{\mathbf{k}}$	$\text{L}^{-1}$	$\tilde{\mathbf{k}} = \mathbf{k}/2\pi$
wavelength	$\Lambda$	L	$\Lambda = 2\pi/ \mathbf{k}  = 1/ \tilde{\mathbf{k}} $
reduced wavelength	$\lambda$	L	$\lambda = \Lambda/2\pi = 1/ \mathbf{k} $

TABLE 1.2: Summarizing the variety of terms used to kinematically describe waves. In this book we mostly use the angular frequency,  $\omega \geq 0$ , the wavevector,  $\mathbf{k}$ , and the wavenumber,  $|\mathbf{k}|$ .

## 1.6 Free wave packets

The modulation of waves in Figure 1.3 serve to organize waves into  $\mathbf{x}$ -space regions known as **wave trains**. A **wave packet** is a train of free waves that are organized into a localized (and moving) region of  $\mathbf{x}$ -space. Localization is enabled by a **modulation function** and the plane wave modes that are organized within the packet are known as **carrier waves**. Wave trains and

wave packets are generated by a [wavemaker](#). For example, consider a wavemaker oscillating at a known frequency,  $\omega_0$ . Ramping up the wavemaker amplitude for a period of time, and then ramping it down towards zero, then produces a group of waves that are localized in time and have angular frequency and wavenumber centered on  $\omega_0$  and  $\mathbf{k}_0$ , where  $\mathbf{k}_0$  is determined by the dispersion relation given  $\omega_0$ . In the laboratory, the wavemaker might be an oscillating piston placed in a tank of water, such as considered for acoustic waves in Section 3.8. In the natural environment, a wavemaker might be a storm that passes over the ocean and generates surface gravity waves, or an atmospheric convective event that generates atmospheric internal gravity waves.

Mathematically, wave packets are described by complementary wave functions in  $\mathbf{k}$ -space and  $\mathbf{x}$ -space. The  $\mathbf{k}$ -space representation,  $\mathcal{A}(\mathbf{k}, t)$ , is also called the amplitude function for the  $\mathbf{x}$ -space wave function,  $\Phi(\mathbf{x}, t)$ . In the world of mathematical abstraction, we prepare wave packets with a specified  $\mathbf{k}$ -space initial condition,  $\mathcal{A}(\mathbf{k}, t = 0)$ , that has a corresponding  $\mathbf{x}$ -space initial condition,  $\Phi(\mathbf{x}, t = 0)$ , determined by the inverse Fourier transform. Upon initializing the packets we study their evolution. Since the packets are constructed by linear plane waves, their evolution is determined by the superposed and evolving constituent plane waves. What distinguishes a packet of acoustic waves from a packet of gravity waves, for example, is the dispersion relation, which is specific to the physics describing the particular waves.

### 1.6.1 Phase velocity, group velocity, and the dispersion relation

Examination of Figure 1.3 suggests that the velocity of a wave packet's center is not directly related to the phase velocity of the constituent plane waves. Rather, it is determined by motion of the modulation function. For the one-dimensional examples of Figure 1.3, the packet motion is determined by  $\Delta\omega/\Delta k$ . In this section we generalize this result to the [group velocity](#) determined by<sup>9</sup>

$$\mathbf{c}_g = \nabla_{\mathbf{k}} \varpi(\mathbf{k}). \quad (1.44)$$

In this equation, we introduced the wave dispersion relation (1.1) that specifies how the angular frequency,  $\omega$ , is functionally connected to the wavevector. That is, by specifying a particular wavevector we then determine the corresponding angular frequency

$$\omega = \varpi(\mathbf{k}). \quad (1.45)$$

Equation (1.44) says that the group velocity is determined by the gradient of the dispersion relation in wavevector space, which prompted us to expose a wavevector subscript on the gradient operator. The group velocity points in the  $\mathbf{k}$ -space direction of maximum increase in the angular frequency. Figure 1.4 illustrates two dispersion relations for a one-dimensional wave.

The dispersion relation for non-dispersive waves results in phase speeds that are independent of the wavenumber so that

$$\varpi = C_p |\mathbf{k}|, \quad (1.46)$$

where  $C_p > 0$  is the constant phase speed. Evidently, the angular frequency is identical for all waves having the same wavenumber. That is, the angular frequency only cares about the wavelength of the non-dispersive wave and not about its direction. Furthermore, we see that the continuum of wavenumbers results in a continuum of angular frequencies, with

---

<sup>9</sup>Many dispersion relations are naturally written as  $\varpi^2$ , in which case it is simpler to compute the group velocity according to  $\nabla_{\mathbf{k}} \varpi^2 = 2 \varpi \mathbf{c}_g$ .

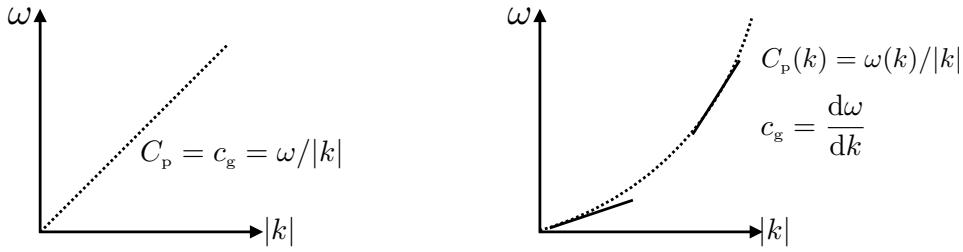


FIGURE 1.4: Example dispersion relations for one-dimensional waves. Left panel: a non-dispersive wave has a phase speed,  $C_p$ , that is a constant so that all waves, regardless their wavelength or angular frequency, move at the same speed. Right panel: a dispersive wave, here shown for a wave whose phase speed increases as the wave number increases, as does the group velocity.

higher wavenumbers rendering higher angular frequencies. This behavior is familiar for both electromagnetic and acoustic waves, which are both non-dispersive, in which waves of smaller wavelength (higher wavenumber) have higher frequencies. We depict this sort of dispersion relation in the left panel of Figure 1.4.

Dispersive waves are characterized by a dispersion relation with the phase speed a function of the wavevector

$$\varpi = C_p(\mathbf{k}) |\mathbf{k}| > 0. \quad (1.47)$$

Again, the phase speed and angular frequency are positive, but now we find the phase speed depends on the wavenumber and possibly the wave direction. We provide an example for this sort of dispersion relation in the right panel of Figure 1.4.

### 1.6.2 A continuum of traveling plane waves

Consider the superposition of a continuum of traveling plane waves given by the integral expression for the wave function

$$\Phi(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t)} d\mathbf{k}. \quad (1.48)$$

The wave function,  $\Phi(\mathbf{x}, t)$ , is synthesized by combining a continuum of traveling plane waves, with the amplitude function,  $\mathcal{A}(\mathbf{k})$ , determining which wavevectors contribute to the packet and by how much they contribute. The integral in equation (1.48) is computed over all the three-dimensional wavevector space ( $\mathbf{k}$ -space). For example, a Cartesian representation of a wavevector volume element is written

$$d\mathbf{k} = dk_x dk_y dk_z, \quad (1.49)$$

with the integration limits  $k_x, k_y, k_z \in (-\infty, \infty)$ , corresponding to waves traveling in all directions and spanning all wavenumbers. Furthermore, the angular frequency in equation (1.48) is specified by the dispersion relation,

$$\omega = \varpi(\mathbf{k}) \geq 0, \quad (1.50)$$

so that the wave function (1.48) is built from a continuum of wavevectors with corresponding angular frequencies.

### 1.6.3 The wave function is real

For brevity we often drop the  $\text{Re}$  operator on the wave function (1.48). Even so, it is critical to note that the wave function is real, which has implications for how we compute the  $\mathbf{k}$ -space amplitude function,  $\mathcal{A}(\mathbf{k})$ . For example, if the amplitude function is real<sup>10</sup>

$$\mathcal{A}(\mathbf{k}) = A(\mathbf{k}) = A^*(\mathbf{k}), \quad (1.51)$$

then the wave function takes the form

$$\Phi(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int A(\mathbf{k}) \cos[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t] d\mathbf{k}. \quad (1.52)$$

However, there are cases where we need the extra phase degree of freedom afforded by a complex amplitude function, in which case the real part of equation (1.48) leads to

$$\Phi(\mathbf{x}, t) = \text{Re} \left[ \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t)} d\mathbf{k} \right] \quad (1.53a)$$

$$= \frac{1}{2(2\pi)^3} \int [\mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t)} + \mathcal{A}^*(\mathbf{k}) e^{-i(\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t)}] d\mathbf{k}. \quad (1.53b)$$

Since the integral is over all of  $\mathbf{k}$ -space, we can write

$$\int \mathcal{A}^*(\mathbf{k}) e^{-i(\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t)} d\mathbf{k} = \int \mathcal{A}^*(-\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} + \varpi(-\mathbf{k}) t)} d\mathbf{k}, \quad (1.54)$$

in which case the wavefunction takes the form

$$\Phi(\mathbf{x}, t) = \frac{1}{2(2\pi)^3} \int e^{i\mathbf{k} \cdot \mathbf{x}} [\mathcal{A}(\mathbf{k}) e^{-i\varpi(\mathbf{k}) t} + \mathcal{A}^*(-\mathbf{k}) e^{i\varpi(-\mathbf{k}) t}] d\mathbf{k}, \quad (1.55)$$

and its time tendency is

$$\partial_t \Phi(\mathbf{x}, t) = \frac{i}{2(2\pi)^3} \int e^{i\mathbf{k} \cdot \mathbf{x}} [-\varpi(\mathbf{k}) \mathcal{A}(\mathbf{k}) e^{-i\varpi(\mathbf{k}) t} + \varpi(-\mathbf{k}) \mathcal{A}^*(-\mathbf{k}) e^{i\varpi(-\mathbf{k}) t}] d\mathbf{k}. \quad (1.56)$$

We will become friendly with these sorts of wave packet manipulations in the following.

### 1.6.4 Initializing the wave packet

We often consider wave packets in the context of initial value problems. Furthermore, we are concerned with freely moving wave packets, so that boundary conditions are not considered here. Thus, we consider three possibilities: initializing the packet's wave function, initializing its time tendency, or initializing both the wave function and its time tendency. To describe these initial conditions requires the use of complex  $\mathbf{k}$ -space amplitudes. Setting  $t = 0$  in equation (1.55) for the wave function and equation (1.56) for the time tendency leads to

$$\Phi_o(\mathbf{x}) \equiv \Phi(\mathbf{x}, 0) = \frac{1}{2(2\pi)^3} \int e^{i\mathbf{k} \cdot \mathbf{x}} [\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k})] d\mathbf{k} \quad (1.57a)$$

$$\dot{\Phi}_o(\mathbf{x}) \equiv \partial_t \Phi(\mathbf{x}, 0) = \frac{i}{2(2\pi)^3} \int e^{i\mathbf{k} \cdot \mathbf{x}} [-\varpi(\mathbf{k}) \mathcal{A}(\mathbf{k}) + \varpi(-\mathbf{k}) \mathcal{A}^*(-\mathbf{k})] d\mathbf{k}. \quad (1.57b)$$

---

<sup>10</sup>The star is the complex conjugate operator.

Inverting the Fourier transform for the initial condition (1.57a) and the initial tendency (1.57b) renders

$$\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k}) = 2 \int \Phi_o(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} \quad (1.58a)$$

$$-\varpi(\mathbf{k}) \mathcal{A}(\mathbf{k}) + \varpi(-\mathbf{k}) \mathcal{A}^*(-\mathbf{k}) = -2i \int \dot{\Phi}_o(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}, \quad (1.58b)$$

which then leads to the  $\mathbf{k}$ -space amplitude

$$\mathcal{A}(\mathbf{k}) = \frac{2}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} \int [\varpi(-\mathbf{k}) \Phi_o(\mathbf{x}) + i \dot{\Phi}_o(\mathbf{x})] e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} \quad (1.59a)$$

$$\mathcal{A}^*(-\mathbf{k}) = \frac{2}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} \int [\varpi(\mathbf{k}) \Phi_o(\mathbf{x}) - i \dot{\Phi}_o(\mathbf{x})] e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}. \quad (1.59b)$$

Notice that

$$\mathcal{A}(\mathbf{k}) = \mathcal{A}^*(-\mathbf{k}) \quad \text{if } \dot{\Phi}_o = 0, \quad (1.60)$$

which is referred to as **conjugate symmetry** when studying Fourier transforms (see VOLUME 1), whereas

$$\mathcal{A}(\mathbf{k}) = -\mathcal{A}^*(-\mathbf{k}) \quad \text{if } \Phi_o = 0, \quad (1.61)$$

which we refer to as conjugate anti-symmetry.

Now substitute the amplitude functions (1.59a) and (1.59b) into the wave function (1.56), and rearrange to find

$$\begin{aligned} \Phi(\mathbf{x}, t) &= \frac{1}{(2\pi)^3} \int \left[ \int \frac{e^{i\mathbf{k}\cdot(\mathbf{x}-\boldsymbol{\xi})} [\varpi(-\mathbf{k}) e^{-i\varpi(\mathbf{k})t} + \varpi(\mathbf{k}) e^{i\varpi(-\mathbf{k})t}]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k} \right] \Phi_o(\boldsymbol{\xi}) d\boldsymbol{\xi} \\ &\quad + \frac{i}{(2\pi)^3} \int \left[ \int \frac{e^{i\mathbf{k}\cdot(\mathbf{x}-\boldsymbol{\xi})} [e^{-i\varpi(\mathbf{k})t} - e^{i\varpi(-\mathbf{k})t}]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k} \right] \dot{\Phi}_o(\boldsymbol{\xi}) d\boldsymbol{\xi}. \end{aligned} \quad (1.62)$$

This equation can be written as the convolution of the initial conditions with two kernel functions

$$\Phi(\mathbf{x}, t) = \int [\dot{G}(\mathbf{x} - \boldsymbol{\xi}, t) \Phi_o(\boldsymbol{\xi}) + G(\mathbf{x} - \boldsymbol{\xi}, t) \dot{\Phi}_o(\boldsymbol{\xi})] d\boldsymbol{\xi}, \quad (1.63)$$

where the kernels are

$$G(\mathbf{x}, t) \equiv \frac{i}{(2\pi)^3} \int \frac{e^{i\mathbf{k}\cdot\mathbf{x}} [e^{-i\varpi(\mathbf{k})t} - e^{i\varpi(-\mathbf{k})t}]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k} \quad (1.64a)$$

$$\dot{G}(\mathbf{x}, t) \equiv \frac{1}{(2\pi)^3} \int \frac{e^{i\mathbf{k}\cdot\mathbf{x}} [\varpi(-\mathbf{k}) e^{-i\varpi(\mathbf{k})t} + \varpi(\mathbf{k}) e^{i\varpi(-\mathbf{k})t}]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k}. \quad (1.64b)$$

In fact, the kernels are related by a time derivative, which can be seen by writing them in the following form<sup>11</sup>

$$G(\mathbf{x}, t) = -\frac{2}{(2\pi)^3} \int \frac{\sin[\mathbf{k}\cdot\mathbf{x} - \varpi(\mathbf{k})t]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k} \quad (1.65a)$$

$$\dot{G}(\mathbf{x}, t) = \frac{2}{(2\pi)^3} \int \frac{\varpi(\mathbf{k}) \cos[\mathbf{k}\cdot\mathbf{x} - \varpi(\mathbf{k})t]}{\varpi(\mathbf{k}) + \varpi(-\mathbf{k})} d\mathbf{k} = \partial_t G(\mathbf{x}, t). \quad (1.65b)$$

<sup>11</sup>To derive equation (1.65b), recall that the integral is over all of  $\mathbf{k}$ -space.

Evidently, the kernels provide a wave mechanism to propagate information about the initial conditions outward in space as time moves forward. In this manner, these functions are [Green's function](#). Indeed, the solution (1.65b) corresponds to the Green's function derived in VOLUME 1 for the initial-boundary value problem with non-dispersive waves. Here, we are just concerned with the initial condition problem for free wave packets where the waves can be either non-dispersive or dispersive. It is satisfying that even dispersive waves allow us to write the wave packet evolution in terms of Green's functions as in equation (1.63).

### 1.6.5 The case of non-dispersive waves

To help further understand the nature of the wave packet equation (1.63), consider a packet of non-dispersive waves for  $t \geq 0$  and in one space dimension.<sup>12</sup> For this case the dispersion relation takes the form

$$\varpi(\mathbf{k}) = C_p |k| \implies \mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t = k x - C_p |k| t, \quad (1.66)$$

where  $C_p > 0$  is the constant phase speed. In this case we make use of the Dirac delta expression from VOLUME 1,

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk = \frac{2}{2\pi} \int_0^{\infty} \cos(kx) dk, \quad (1.67)$$

so that

$$\dot{G}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(kx - C_p |k| t) dk \quad (1.68a)$$

$$= \frac{1}{2\pi} \int_{-\infty}^0 \cos(kx + C_p k t) dk + \frac{1}{2\pi} \int_0^{\infty} \cos(kx - C_p k t) dk \quad (1.68b)$$

$$= \frac{1}{2\pi} \int_0^{\infty} \cos(kx + C_p k t) dk + \frac{1}{2\pi} \int_0^{\infty} \cos(kx - C_p k t) dk \quad (1.68c)$$

$$= \frac{1}{2} [\delta(x + C_p t) + \delta(x - C_p t)] \quad (1.68d)$$

$$= \frac{1}{2} \delta(C_p t - |x|). \quad (1.68e)$$

Hence,  $\dot{G}$  is built from two Dirac delta wave fronts that move in opposite directions at speed  $C_p$ . Its time integral leads to the Green's function as a [Heaviside step function](#)

$$G(x, t) = \mathcal{H}(C_p t - |x|)/(2 C_p). \quad (1.69)$$

To within a constant factor, the Green's function (1.69) is the [causal free space Green's function](#) derived in VOLUME 1 for the one dimensional non-dispersive waves. Making use of these results in the wave packet equation (1.65b) leads to

$$\Phi(x, t) = \frac{1}{2} \int_{-\infty}^{\infty} \delta(C_p t - |x - \xi|) \Phi_o(\xi) d\xi + \frac{1}{2 C_p} \int_{-\infty}^{\infty} \mathcal{H}(C_p t - |x - \xi|) \dot{\Phi}_o(\xi) d\xi \quad (1.70a)$$

$$= [\Phi_o(x - C_p t) + \Phi_o(x + C_p t)]/2 + \frac{1}{2 C_p} \int_{x-C_p t}^{x+C_p t} \dot{\Phi}_o(\xi) d\xi. \quad (1.70b)$$

---

<sup>12</sup>This case corresponds to the infinite string discussed in Section 7.8 of [Stakgold \(2000b\)](#). We also consider packets of non-dispersive waves in Section 1.7.7, with a focus on Gaussian packets.

For the second equality we used

$$\int_{-\infty}^{\infty} \mathcal{H}(C_p t - |x - \xi|) \dot{\Phi}_o(\xi) d\xi = \int_{-\infty}^{\infty} \mathcal{H}(C_p t - |y|) \dot{\Phi}_o(x + y) dy \quad (1.71a)$$

$$= \int_{-C_p t}^{C_p t} \dot{\Phi}_o(x + y) dy \quad (1.71b)$$

$$= \int_{x-C_p t}^{x+C_p t} \dot{\Phi}_o(\xi) d\xi. \quad (1.71c)$$

Evidently, for non-dispersive waves the initial condition,  $\Phi_o(x)$ , splits in half as two signals that propagate without distortion in opposite directions, whereas the initial tendency,  $\dot{\Phi}_o(x)$ , contributes via an integral over the domain of influence, which is given by the spatial region  $[x - C_p t, x + C_p t]$ .

### 1.6.6 Wave trains and wave packets

Consider a wave train initialized according to<sup>13</sup>

$$\Phi_o(\mathbf{x}) = \text{Re}[a(\epsilon \mathbf{x}) e^{i \mathbf{k}_o \cdot \mathbf{x}}] = a(\epsilon \mathbf{x}) \cos(\mathbf{k}_o \cdot \mathbf{x}), \quad (1.72)$$

where  $a$  is a real  $\mathbf{x}$ -space modulation function and  $\epsilon^{-1}$ , is a length scale that is large relative to the wavelength of the carrier wave; i.e.,  $\epsilon^{-1} \ll |\mathbf{k}_o|$ . Figure 1.5 provides an example of such a wave train. A further modulation of the  $\mathbf{x}$ -space wave train into a wave packet can be realized by an amplitude that sets  $\Phi_o(\mathbf{x})$  to zero (or exponentially close to zero) for length scales larger than  $\epsilon^{-1}$ .

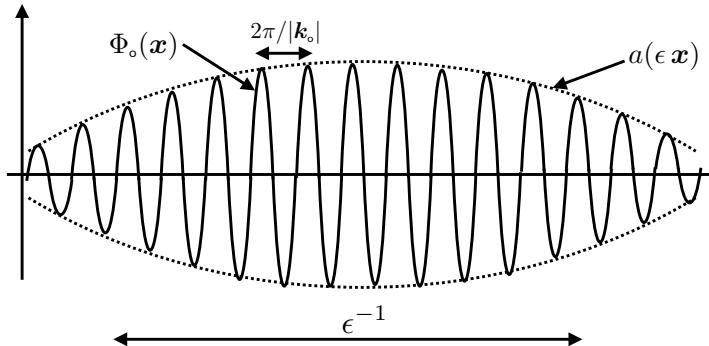


FIGURE 1.5: Example of a wave train comprised of a modulated amplitude function,  $a(\epsilon \mathbf{x})$ , acting on a carrier wave,  $e^{i \mathbf{k}_o \cdot \mathbf{x}}$ . The modulation function has a length scale,  $\epsilon^{-1}$ , which is large relative to that of the wave length,  $2\pi/|\mathbf{k}_o|$ , so that  $2\pi/\epsilon \ll |\mathbf{k}_o|$ .

For either a wave train or wave packet to be dominated by a single carrier wave of wavenumber  $\mathbf{k}_o$ , the  $\mathbf{k}$ -space wave amplitude must vanish for wavevectors outside of an  $\epsilon$  range of  $\mathbf{k}_o$

$$|\mathcal{A}(\mathbf{k})| \approx 0 \quad \text{for } |\mathbf{k} - \mathbf{k}_o| > \epsilon. \quad (1.73)$$

As a result, evolution of the  $\mathbf{x}$ -space wave packet (1.48) is dominated by the dispersion relation

<sup>13</sup>Note that the real operator, Re, will henceforth be dropped for brevity.

centered on  $\mathbf{k}_o$ , thus motivating a Taylor series

$$\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}) t \approx \mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k}_o) t - (k_a - k_{oa}) [\partial_{k_a} + (1/2)(k_b - k_{ob}) \partial_{k_a} \partial_{k_b}] \varpi(\mathbf{k}_o) t \quad (1.74a)$$

$$= \mathbf{k}_o \cdot \mathbf{x} - \omega_o t + (\mathbf{k} - \mathbf{k}_o) \cdot [\mathbf{x} - \mathbf{c}_g t - (t/2)(k_b - k_{ob}) \partial_{k_b} \mathbf{c}_g], \quad (1.74b)$$

where we made use of the summation convention<sup>14</sup> and introduced the angular frequency,  $\omega_o$ , and group velocity,  $\mathbf{c}_g$ , as defined by the dispersion relation evaluated at  $\mathbf{k}_o$

$$\omega_o = \varpi(\mathbf{k}_o) \quad \text{and} \quad \mathbf{c}_g = \nabla_{\mathbf{k}} \varpi(\mathbf{k}_o). \quad (1.75)$$

Plugging the Taylor expansion (1.74b) into the wave packet (1.48) renders

$$\Phi(\mathbf{x}, t) = e^{i(\mathbf{k}_o \cdot \mathbf{x} - \omega_o t)} \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} - \mathbf{k}_o) \cdot \tilde{\mathbf{x}}} d\mathbf{k} = e^{i(\mathbf{k}_o \cdot \mathbf{x} - \omega_o t)} \mathcal{M}(\mathbf{x}, t). \quad (1.76)$$

In this equation we introduced the moving position determined by the group velocity and its derivative

$$\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{c}_g t - (t/2)(k_b - k_{ob}) \partial_{k_b} \mathbf{c}_g \iff \tilde{x}^a = x^a - t \frac{\partial \varpi}{\partial k_a} - \frac{t}{2}(k_b - k_{ob}) \frac{\partial^2 \varpi}{\partial k_a \partial k_b}. \quad (1.77)$$

The evolving packet in equation (1.76) represents a traveling plane carrier wave,  $e^{i(\mathbf{k}_o \cdot \mathbf{x} - \omega_o t)}$ , with the integral term a moving modulation function

$$\mathcal{M}(\mathbf{x}, t) \equiv \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} - \mathbf{k}_o) \cdot \tilde{\mathbf{x}}} d\mathbf{k}. \quad (1.78)$$

We follow this discussion in Section 1.7 by studying one-dimensional Gaussian wave packets, where we explicitly compute the  $\mathbf{x}$ -space modulation function, and find that it is peaked at a point following the group velocity,

$$\mathbf{x}_{\text{peak}} = \mathbf{x} - \mathbf{c}_g t, \quad (1.79)$$

whereas the term,  $\partial_{k_j} \mathbf{c}_g$ , appearing in equation (1.77) gives rise to a spread or dispersion of the  $\mathbf{x}$ -space packet.

### 1.6.7 PDE for the wave packet modulation function

To further an understanding of wave packet evolution, we derive the partial differential equation satisfied by the modulation function. For this purpose, take the space and time derivatives of  $\mathcal{M}(\mathbf{x}, t)$  from equation (1.78) to find

$$\frac{\partial \mathcal{M}}{\partial t} = \frac{-i}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) (k_a - k_{oa}) (\partial_{k_a} \varpi + (1/2)(k_b - k_{ob}) \partial_{k_b} \partial_{k_a} \varpi) e^{i(\mathbf{k} - \mathbf{k}_o) \cdot \tilde{\mathbf{x}}} d\mathbf{k} \quad (1.80a)$$

$$\frac{\partial \mathcal{M}}{\partial x^a} = \frac{i}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) (k_a - k_{oa}) e^{i(\mathbf{k} - \mathbf{k}_o) \cdot \tilde{\mathbf{x}}} d\mathbf{k} \quad (1.80b)$$

$$\frac{\partial^2 \mathcal{M}}{\partial x^a \partial x^b} = -\frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) (k_a - k_{oa})(k_b - k_{ob}) e^{i(\mathbf{k} - \mathbf{k}_o) \cdot \tilde{\mathbf{x}}} d\mathbf{k}, \quad (1.80c)$$

---

<sup>14</sup>We use the Cartesian version of the summation convention so there is no distinction between upstairs and downstairs indices.

which then leads to

$$(\partial_t + \mathbf{c}_g \cdot \nabla) \mathcal{M} = \frac{i}{2} \frac{\partial^2 \varpi}{\partial k_a \partial k_b} \frac{\partial^2 \mathcal{M}}{\partial x^a \partial x^b}. \quad (1.81)$$

The left hand side reveals the time evolution of the modulation function following the group velocity, with the group velocity evaluated at the carrier wavevector,  $\mathbf{k}_o$ . The right hand side is a phase shifted (by  $\pi/2$ ) transport of the modulation function by the dispersion tensor,  $\mathbb{K}$ , with the dispersion tensor determined by the second derivative of the dispersion relation evaluated at the carrier wavevector,  $\mathbf{k}_o$

$$\mathbb{K}^{ab} = \left. \frac{\partial^2 \varpi}{\partial k_a \partial k_b} \right|_{\mathbf{k}=\mathbf{k}_o}. \quad (1.82)$$

The dispersion tensor is symmetric, as for a diffusion tensor. However, there is no guarantee that it is positive-definite. Hence, dispersion can both broaden the wave packet (as for diffusion, with a positive-definite dispersion tensor) or sharpen the wave packet (as for anti-diffusion).

It is important to emphasize that dispersion is distinct from dissipation. Indeed, dissipation is absent from this chapter. Evidently, modification of the modulation function arises from the dependence of the phase velocity on the wavevector, hence the name **dispersion**. Dispersion is a property inherent in the waves and has nothing to do with dissipation.

### 1.6.8 Wave packets of non-dispersive waves

Packets of non-dispersive waves propagate their initial condition without modification, thus retaining their initial structure. We can understand this result by recalling that all waves comprising the packet move with the same phase velocity, so that the group and phase velocities are identical and constant. Hence, the packet maintains its initial organization of its waves as the packet evolves. We here verify that this result holds when making use of the wave packet formalism of this section.

A vanishing dispersion tensor (1.82) is one signature that a wave packet of non-dispersive waves moves coherently and without modification. As a result, the packet modulation function remains constant following the group velocity,

$$(\partial_t + \mathbf{c}_g \cdot \nabla) \mathcal{M} = 0 \quad \text{non-dispersive waves.} \quad (1.83)$$

The modulation function is given by equation (1.78) with a constant phase velocity

$$\mathcal{M}(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k}-\mathbf{k}_o) \cdot (\mathbf{x} - \mathbf{c}_g t)} d\mathbf{k} = \mathcal{M}(\mathbf{x} - \mathbf{c}_g t), \quad (1.84)$$

which means that the wave packet (1.76) evolves according to

$$\Phi(\mathbf{x}, t) = e^{i(\mathbf{k}_o \cdot \mathbf{x} - \omega_o t)} \mathcal{M}(\mathbf{x} - \mathbf{c}_g t). \quad (1.85)$$

Indeed, returning to the original form (1.76) of the wave packet we find that for non-dispersive waves

$$\Phi(\mathbf{x}, t) = e^{i(\mathbf{k}_o \cdot \mathbf{x} - \omega_o t)} \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k}-\mathbf{k}_o) \cdot (\mathbf{x} - \mathbf{c}_g t)} d\mathbf{k} \quad (1.86a)$$

$$= \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i\mathbf{k} \cdot (\mathbf{x} - \mathbf{c}_g t)} d\mathbf{k} \quad (1.86b)$$

$$= \Phi_o(\mathbf{x} - \mathbf{c}_g t), \quad (1.86c)$$

where we introduced the initial condition,  $\Phi_0$ , from equation (1.57a). Evidently, packets of non-dispersive waves translate their initial condition without any alteration, with translation determined by the group velocity as determined by the carrier wavevector,  $\mathbf{k}_0$ . We make this result compatible with the one dimensional case found in equation (1.70b) by noting that here we only focus on the carrier wavenumber, whereas for the one dimensional packet in equation (1.70b) we considered both the carrier wavenumber,  $k_0$ , and its opposite,  $-k_0$ . However, for a wave packet, we generally focus only on the group velocity determined by the carrier wave since any other wavevectors outside the local neighborhood have an exponentially small amplitude. We consider this result for the specific case of a Gaussian wave packet in Section 1.7.7, with further comment in Section 1.7.9.

### 1.6.9 Wave function PDE implied by the dispersion relation

Typically when studying waves we start with a partial differential equation in space and time, linearize this equation, and then insert a plane wave ansatz to determine a dispersion relation. Here we follow in a manner analogous to that in Section 1.6.7 and seek the partial differential equation satisfied by the wave function, now assuming knowledge of the dispersion relation as might be provided by measurements of wave properties. The method is best illustrated by an example, here taken as that for a planetary Rossby wave for the horizontally non-divergent barotropic model (Section 6.3.4)

$$\varpi(\mathbf{k}) = -\beta k_x / |\mathbf{k}|^2, \quad (1.87)$$

where  $\beta = \partial_y f \geq 0$  is the meridional derivative of the Coriolis parameter, and  $\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}$  is a two-dimensional wavevector. Now build a wave function from such waves, where the wave function is written as in equation (1.48)

$$\Phi(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} d\mathbf{k} = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} + t\beta k_x / |\mathbf{k}|^2)} d\mathbf{k}. \quad (1.88)$$

What is the partial differential equation satisfied by this wave function?

To answer this question we proceed as in Section 1.6.7 for the modulation function by taking partial derivatives in space and time, and noting that the integral for the wave function is over  $\mathbf{k}$ -space so that space and time derivatives commute with the integral. Hence, the Laplacian of the wave function is given by

$$\nabla^2 \Phi = -\frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) |\mathbf{k}|^2 e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} d\mathbf{k}, \quad (1.89)$$

and its time derivative is

$$\partial_t(\nabla^2 \Phi) = \frac{i}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) \varpi |\mathbf{k}|^2 e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} d\mathbf{k} \quad (1.90a)$$

$$= -\frac{i\beta}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) k_x e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} d\mathbf{k} \quad (1.90b)$$

$$= -\beta \partial_x \Phi, \quad (1.90c)$$

which then leads to the linear partial differential equation

$$\partial_t(\nabla^2 \Phi) + \beta \partial_x \Phi = 0. \quad (1.91)$$

As we see in Chapter 6, equation (1.91) is the linearized version of the vorticity equation

(6.29) with zero background flow, and solutions of this equation are planetary Rossby waves. Consequently, equation (1.91) describes the evolution of a packet of planetary Rossby waves in a two-dimensional non-divergent barotropic fluid.

### 1.6.10 Standing wave packets

The wave packet (1.48) can be written in the form

$$\Phi^{(+)}(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} d\mathbf{k} = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i|\mathbf{k}|[S_{\hat{\mathbf{k}}} - C_p t]} d\mathbf{k}, \quad (1.92)$$

where

$$S_{\hat{\mathbf{k}}} = \hat{\mathbf{k}} \cdot \mathbf{x} \quad (1.93)$$

is the phase distance from equation (1.32), and  $C_p = \varpi(\mathbf{k})/|\mathbf{k}| > 0$  is the phase speed. As noted in Section 1.6.1, the phase speed is a constant for non-dispersive waves, whereas for dispersive waves it is a function of the wavevector,  $C_p = C_p(\mathbf{k})$ . We introduce the + notation on the packet (1.92) to distinguish from the oppositely traveling packet,  $\Phi^{(-)}$ , defined according to

$$\Phi^{(-)}(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(-\mathbf{k}) e^{-i[\mathbf{k} \cdot \mathbf{x} + \varpi(\mathbf{k})t]} d\mathbf{k}. \quad (1.94)$$

If the amplitude function has even parity

$$\mathcal{A}(-\mathbf{k}) = \mathcal{A}(\mathbf{k}), \quad (1.95)$$

then  $\Phi^{(-)}(\mathbf{x}, t) = \Phi^{(+)}(-\mathbf{x}, t)$ , in which case the sum of these two traveling packets is a standing wave packet

$$\Phi^{(-)}(\mathbf{x}, t) + \Phi^{(+)}(\mathbf{x}, t) = \frac{2}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{-i\varpi(\mathbf{k})t} \cos(\mathbf{k} \cdot \mathbf{x}) d\mathbf{k} \quad \text{if } \mathcal{A}(-\mathbf{k}) = \mathcal{A}(\mathbf{k}), \quad (1.96)$$

and the initial condition is the inverse cosine transform of  $\mathcal{A}(\mathbf{k})$

$$\Phi^{(-)}(\mathbf{x}, t=0) + \Phi^{(+)}(\mathbf{x}, t=0) = \frac{2}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) \cos(\mathbf{k} \cdot \mathbf{x}) d\mathbf{k} \quad \text{if } \mathcal{A}(-\mathbf{k}) = \mathcal{A}(\mathbf{k}). \quad (1.97)$$

Recall that to ensure the wave function is real, then the amplitude function must satisfy conjugate symmetry via equation (1.95). Hence, with the further assumption of even parity (1.95), then the amplitude function,  $\mathcal{A}(\mathbf{k})$ , is real, in which case we indeed see that  $\Phi^{(-)}(\mathbf{x}, t=0) + \Phi^{(+)}(\mathbf{x}, t=0)$  is a real function.

As seen in Section 1.7.5, we generally consider Gaussian packets that include modulated traveling waves, so that such wave packets have amplitude functions that are not even functions of the wavevector,  $\mathcal{A}(-\mathbf{k}) \neq \mathcal{A}(\mathbf{k})$ . Even so, the present discussion identifies the general properties satisfied by packets of standing waves, which occur, for example, in bounded regions (e.g., see Section 4.8 for standing gravity waves in a closed basin).

## 1.7 Wave packets in one-dimension

In this section we study a variety of packets moving in one space dimension.<sup>15</sup> Doing so provides a pedagogical means to explicitly reveal some of the general ideas developed in Section 1.6. We give particular attention to the Gaussian wave packet since it allows for analytic expressions.<sup>16</sup>

### 1.7.1 The positive wave packet

For a packet built from plane waves traveling in the  $+\hat{x}$  direction, we write the wavevector as  $\mathbf{k} = k \hat{x}$  with wavenumber  $k \geq 0$  so that a positive-traveling packet takes the form

$$\Phi^{(+)}(x, t) = \frac{1}{2\pi} \int_0^\infty \mathcal{A}(k) e^{i(kx - \varpi t)} dk. \quad (1.98)$$

### 1.7.2 The negative wave packet

A packet moving in the negative  $\hat{x}$  direction is built from plane waves with  $\mathbf{k} = -k \hat{x}$ , again with wavenumber  $k \geq 0$ , so that

$$\Phi^{(-)}(x, t) = \frac{1}{2\pi} \int_0^\infty \mathcal{A}(-k) e^{-i[kx + \varpi t]} dk. \quad (1.99)$$

Observe that  $\Phi^{(-)}(x, t)$  can be written in the equivalent form

$$\Phi^{(-)}(x, t) = \frac{1}{2\pi} \int_0^\infty \mathcal{A}(-k) e^{-i[kx + \varpi t]} dk = \frac{1}{2\pi} \int_{-\infty}^0 \mathcal{A}(k) e^{i[kx - \varpi t]} dk, \quad (1.100)$$

where we assumed that the dispersion relation has even parity

$$\varpi(k) = \varpi(-k) > 0. \quad (1.101)$$

This assumption for the dispersion relation does not hold for all waves (e.g., Rossby waves). When it does hold, it means that a wave traveling to the right with wavenumber  $k$  has the same angular frequency as a wave traveling to the left with the same wavenumber. The equality (1.100) allows us to interpret integration over negative wavenumbers to represent a wave packet traveling in a direction opposite to the sense suggested by the phase relation. That is, the phase  $kx - \varpi t$  suggests a positive traveling packet, but the packet is actually a negative packet since the integration extends over negative wavenumbers.

### 1.7.3 The positive-negative wave packet

The sum of the positive packet (1.98) and negative packet (1.100) leads us to define the positive-negative packet

$$\Phi^{(+-)}(x, t) = \Phi^{(+)}(x, t) + \Phi^{(-)}(x, t) = \frac{1}{2\pi} \int_{-\infty}^\infty \mathcal{A}(k) e^{i(kx - \varpi t)} dk. \quad (1.102)$$

---

<sup>15</sup>Section 7.8 of [Stakgold \(2000b\)](#) provides a thorough discussion of one-dimensional waves, such as those propagating on a string.

<sup>16</sup>As a wave packet in one-dimension, the wave function  $\Phi$  has dimensions of its amplitude function,  $\mathcal{A}$ , times inverse length. In three dimensions, in contrast, the wave function has dimensions of  $\mathcal{A}$  times inverse volume as in equation (1.48).

### 1.7.4 The positive-positive wave packet

This final wave packet is designed to travel in a single direction, which defines the positive-positive packet

$$\Phi^{(++)}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{A}(k) e^{ik(x - C_p t)} dk, \quad (1.103)$$

where  $C_p = \varpi/|k| > 0$  is the phase speed. It is revealing to decompose this packet as<sup>17</sup>

$$\Phi^{(++)}(x, t) = \frac{1}{2\pi} \int_{-\infty}^0 \mathcal{A}(k) e^{ik(x - \varpi t/|k|)} dk + \frac{1}{2\pi} \int_0^{\infty} \mathcal{A}(k) e^{ik(x - \varpi t/|k|)} dk \quad (1.104a)$$

$$= \frac{1}{2\pi} \int_{-\infty}^0 \mathcal{A}(k) e^{ik(x + \varpi t)} dk + \frac{1}{2\pi} \int_0^{\infty} \mathcal{A}(k) e^{ik(x - \varpi t)} dk. \quad (1.104b)$$

The second integral equals to  $\Phi^{(+)}(x, t)$ , whereas the first is new. Again, both contributions to  $\Phi^{(++)}(x, t)$  move in the  $+\hat{x}$  direction.

### 1.7.5 Initial uncertainty relation for a Gaussian packet

The initial condition for both  $\Phi^{(+-)}$  and  $\Phi^{(++)}$  are the same, which we write as

$$\Phi_o(x) = \Phi^{(+-)}(x, 0) = \Phi^{(++)}(x, 0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{A}(k) e^{ikx} dk. \quad (1.105)$$

The specific case of a Gaussian wave packet offers insight into the physics while allowing for analytically tractable expressions. Specifically, consider the  $\mathbf{k}$ -space representation of a wave packet with a Gaussian spread around a central wavevector,  $\mathbf{k}_o = k_o \hat{x}$ . We choose the wave amplitude as

$$\mathcal{A}(k) = A e^{-\sigma(k - k_o)^2}, \quad (1.106)$$

where  $A > 0$  is a real constant and  $\sigma > 0$  has dimensions of squared length. If we choose the central wavenumber to be positive,  $k_o > 0$ , then the wave packet is dominated by plane waves moving in the  $+\hat{x}$  direction. Referring to the discussion in Section 1.6.10, note that  $\mathcal{A}(k) \neq \mathcal{A}(-k)$  when  $k_o \neq 0$ . As we will see, if  $k_o = 0$  then we have a Gaussian packet without any modulated plane waves.

#### Width of the initial wave packet in $\mathbf{k}$ -space

The squared modulus of the  $\mathbf{k}$ -space wave packet (1.106) provides a measure of the packet's intensity

$$|\mathcal{A}(k)|^2 = A^2 e^{-2\sigma(k - k_o)^2}, \quad (1.107)$$

with the intensity peaked at the wavenumber  $k = k_o$  and declining to  $e^{-1}$  times the maximum for

$$k = k_o \pm (2\sigma)^{-1/2} \implies \Delta k^{\text{efold}} \equiv 2(2\sigma)^{-1/2}, \quad (1.108)$$

where  $\Delta k^{\text{efold}}$  measures the  $\mathbf{k}$ -space width of the wave packet. We say that the  $\mathbf{k}$ -space width of the packet is narrow band for  $\sigma$  large, in which the packet is concentrated around  $k = k_o$  since  $\Delta k^{\text{efold}} \rightarrow 0$ .

---

<sup>17</sup>In deriving equation (1.104b), it is important to remember that  $\varpi > 0$  for all wavenumbers.

### The initial wave packet in $x$ -space

The inverse Fourier transform of the  $\mathbf{k}$ -space wave function (1.106) leads to the initial condition for the  $x$ -space wave function (1.105)

$$\Phi_o(x) = \frac{A}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma(k-k_o)^2} e^{ikx} dk = \frac{A e^{ik_o x}}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma q^2} e^{iqx} dq, \quad (1.109)$$

with the corresponding initial conditions for the positive and negative wave packets

$$\Phi^{(+)}(x, 0) = \frac{A}{2\pi} \int_0^{\infty} e^{-\sigma(k-k_o)^2} e^{ikx} dk = \frac{A e^{ik_o x}}{2\pi} \int_{-k_o}^{\infty} e^{-\sigma q^2} e^{iqx} dq \quad (1.110a)$$

$$\Phi^{(-)}(x, 0) = \frac{A}{2\pi} \int_{-\infty}^0 e^{-\sigma(k-k_o)^2} e^{ikx} dk = \frac{A e^{ik_o x}}{2\pi} \int_{-\infty}^{-k_o} e^{-\sigma q^2} e^{iqx} dq. \quad (1.110b)$$

With the central wavenumber assumed positive,  $k_o > 0$ , observe that the initial negative wave packet,  $\Phi^{(-)}(x, 0)$ , has an exponentially small amplitude since the integral in equation (1.110b) never samples  $q = k - k_o = 0$ . In contrast,  $\Phi^{(+)}(x, 0)$  samples  $q = 0$  so that

$$\Phi_o(x) = \Phi^{(+)}(x, 0) + \Phi^{(-)}(x, 0) \approx \Phi^{(+)}(x, 0). \quad (1.111)$$

Although there is no exact closed form expression for the initial conditions,  $\Phi^{(+)}(x, 0)$  and  $\Phi^{(-)}(x, 0)$ , we can evaluate the integral for  $\Phi_o(x)$  in equation (1.109). First observe that

$$\int_{-\infty}^{\infty} e^{-\sigma q^2} \sin(qx) dq = 0, \quad (1.112)$$

since the Gaussian is symmetric under  $q \rightarrow -q$ , whereas  $\sin(qx)$  switches sign. Hence, expanding the imaginary exponential in equation (1.109) according to Euler's identity leaves only the cosine contribution, so that<sup>18</sup>

$$\Phi_o(x) = \frac{A e^{ik_o x}}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma q^2} \cos(qx) dq = \frac{A e^{ik_o x}}{2\pi} \sqrt{\frac{\pi}{\sigma}} e^{-x^2/(4\sigma)}, \quad (1.113)$$

where we made use of an integral table for the final equality (e.g., integral 679 of [Beyer \(1973\)](#)).

Evidently, the initial wave packet,  $\Phi_o(x)$ , in equation (1.113) consists of a single plane wave,  $e^{ik_o x}$ , modulated by the Gaussian,  $e^{-x^2/(4\sigma)}$ . In this manner the plane wave, which is defined for all space, has been localized in space by the Gaussian modulation function. In Figure 1.6 we illustrate a Gaussian wave packet of the form (1.113).

If we choose the central wavenumber to be zero,  $k_o = 0$ , then the wavenumber amplitude function has even parity,  $\mathcal{A}(k) = \mathcal{A}(-k)$ . The wave function,  $\Phi_o(x)$ , in equation (1.113) reduces to just the Gaussian modulation function without a carrier wave, so that Figure 1.6 reduces to just the positive Gaussian modulation function. Hence, the  $k_o = 0$  limit results in a Gaussian signal, constructed with a continuum of plane waves, yet without any modulated carrier wave.

### The uncertainty relation

The squared modulus of the  $x$ -space wave function (1.113) is proportional to  $\exp[-x^2/(2\sigma)]$ , which has an e-folding width

$$\Delta x^{\text{efold}} = 2(2\sigma)^{1/2}. \quad (1.114)$$

<sup>18</sup>The integral in equation (1.113) is also encountered in our study of Fourier transforms in VOLUME 1.

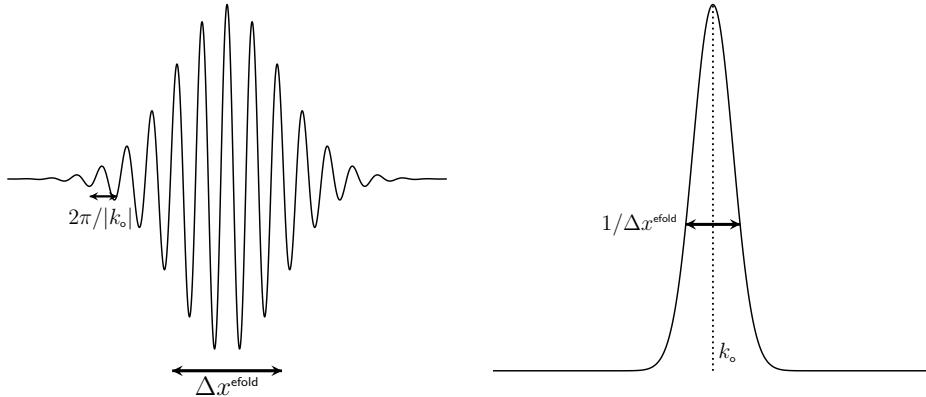


FIGURE 1.6: Left panel: example of a Gaussian wave packet,  $\Phi_0(x)$ , of the form given by the real part of equation (1.113), with  $2\pi/|k_0|$  wavelength for the modulated carrier wave. From equation (1.114) we also show  $\Delta x^{\text{efold}} = 2(2\sigma)^{1/2} \gg 2\pi/|k_0|$ , which determines the e-folding width of the squared modulus of the packet. The plane wave,  $e^{ik_0 x}$ , is referred to as the [carrier wave](#). Right panel: the  $\mathbf{k}$ -space Gaussian amplitude function (1.106). Units are arbitrary in both panels.

This  $x$ -space wave function spread exactly complements its  $\mathbf{k}$ -space spread (1.108), so that their product is a constant

$$\Delta x^{\text{efold}} \Delta k^{\text{efold}} = 4. \quad (1.115)$$

The precise value of the constant is not important since it depends on the somewhat arbitrary choices made for defining  $\Delta x$  and  $\Delta k$ . What is important is that if we narrow the wavenumber band by increasing  $\sigma$ , thus making  $A(\mathbf{k})$  more sharply peaked, then we broaden the  $x$ -space width of the wave function. Evidently, a narrow Gaussian wave packet in  $\mathbf{k}$ -space leads to a broad wave packet in  $x$ -space, with the complement also holding so that a narrow packet in  $x$ -space corresponds to a broad packet in  $\mathbf{k}$ -space. This tradeoff holds for general wave packets, whereby it is not possible to have a narrow packet in both  $x$ -space and  $\mathbf{k}$ -space.

We refer to equation (1.115) as an [uncertainty relation](#). We also encounter this relation when studying Fourier transforms in VOLUME 1. The uncertainty relation expresses a fundamental property of wave kinematics, and it has nothing to do with the accuracy of measurement devices. In quantum mechanics it is known as the [Heisenberg uncertainty principle](#), which forms one of the most important elements of atomic physics (e.g., Chapter 3 of [Bohm \(1951\)](#) or Chapter 2 of [Gasiorowicz \(1974\)](#)).

### 1.7.6 Extreme examples of the uncertainty relation

The uncertainty relation is a bit odd on first encounter. Why is it impossible to know arbitrarily precise information about both the  $\mathbf{k}$ -space location and  $x$ -space location of a wave packet? The answer fundamentally boils down to the dual relation between these two spaces. As a means to further exemplify this relation, and thus to build understanding, we consider the most extreme example of the uncertainty relation.<sup>19</sup>

First assume the  $\mathbf{k}$ -space wave function is a single plane wave with just one wavevector,

<sup>19</sup>We also consider similar examples in VOLUME 1 when studying Fourier transforms.

$\mathbf{k} = k_o \hat{\mathbf{x}}$ . In this case the  $\mathbf{k}$ -space amplitude function is

$$\mathcal{A}(k) = a_o \delta(k - k_o), \quad (1.116)$$

where  $a_o$  is a constant and  $\delta$  is the [Dirac delta](#) studied in [VOLUME 1](#). Note that  $\delta(k - k_o)$  has physical dimensions of length and is normalized so that

$$\int_{-\infty}^{\infty} \delta(k - k_o) dk = 1. \quad (1.117)$$

The sifting property of the Dirac delta yields the  $\mathbf{x}$ -space wave function

$$\Phi_o(x) = \frac{a_o}{2\pi} \int_{-\infty}^{\infty} \delta(k - k_o) e^{ikx} dk = \frac{a_o e^{ik_o x}}{2\pi}. \quad (1.118)$$

So although we know the precise  $\mathbf{k}$ -space position,  $k = k_o$ , we have zero information about the wave function's  $\mathbf{x}$ -space position. That is, the wave function is a pure carrier wave without any modulation function so that the wave function is equally present throughout all space.

We turn the table by assuming precise  $\mathbf{x}$ -space information, in which case the modulation function is a Dirac delta

$$\Phi_o(x) = \phi_o \delta(x - x_o), \quad (1.119)$$

where  $\phi_o$  is a constant and  $\delta(x - x_o)$  is normalized so that

$$\int_{-\infty}^{\infty} \delta(x - x_o) dx = 1. \quad (1.120)$$

We say that the modulation function has exactly specified the spatial position of this packet. Yet the price to pay for this precise  $x$ -space information is that there is zero information concerning the  $\mathbf{k}$ -space location since

$$\mathcal{A}(k) = \phi_o \int_{-\infty}^{\infty} \delta(x - x_o) e^{-ikx} dx = \phi_o e^{-ikx_o}. \quad (1.121)$$

That is, the carrier wave can be any plane wave with arbitrary wavenumber,  $k$ . So although we know the precise  $\mathbf{x}$ -space position,  $x = x_o$ , we have zero information about the wave function's  $\mathbf{k}$ -space position. That is, the wave function carries precise  $\mathbf{x}$ -space information (delta modulated) yet zero  $\mathbf{k}$ -space information.

### 1.7.7 Evolution of a non-dispersive Gaussian wave packet

We now study the evolution of the  $\mathbf{x}$ -space wave functions,  $\Phi^{(+-)}(x, t)$  and  $\Phi^{(++)}(x, t)$ . Starting with the positive-negative wave function, recall that it evolves according to equation (1.102)

$$\Phi^{(+-)}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{A}(k) e^{i(kx - \varpi t)} dk = \frac{A}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma(k - k_o)^2} e^{i[kx - \varpi(k)t]} dk. \quad (1.122)$$

This integral offers a complete space and time specification of  $\Phi^{(+-)}(x, t)$ . Even so, it is useful to massage the integral to garner insight into the physics of this evolving pattern. We do so in this section for the case of non-dispersive waves, where we show that the  $\mathbf{x}$ -space wave packet propagates the initial wave packet without alteration. This result is expected since non-dispersive waves are described by a [hyperbolic partial differential equation](#) that translates

initial patterns without alteration (Section 1.3.1). Furthermore, we already proved this result for a general packet of non-dispersive waves in Section 1.6.8. Even so, it is useful to expose the details in the context of the one dimensional Gaussian packets. Doing so offers insights into certain features of wave packets, and provides practice for the technically more challenging case of dispersive waves in Section 1.7.8.

### Evolution of a Gaussian $\Phi^{(+-)}$ packet of non-dispersive waves

As already encountered in this chapter, a non-dispersive wave is characterized by a dispersion relation with a constant phase speed

$$\omega = \varpi = C_p |k|. \quad (1.123)$$

Such waves are said to be non-dispersive since waves with arbitrary wavenumber travel at the same phase speed. The absolute value sign in the dispersion relation (1.123) means that the relation has a continuous derivative only when  $k \neq 0$ . In that manner, it is a linear dispersion relation only when  $k \neq 0$ . Care must be exercised to account for the sign swap when moving across  $k = 0$ , with this movement corresponding to waves moving in opposite directions as seen earlier in Section 1.6.5.

The  $x$ -space wave packet (1.122) propagates according to

$$\Phi^{(+-)}(x, t) = \frac{A}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma(k-k_0)^2} e^{i(kx - |k|C_p t)} dk \quad (1.124a)$$

$$= \frac{A}{2\pi} \int_0^{\infty} e^{-\sigma(k-k_0)^2} e^{ik(x-C_p t)} dk + \frac{A}{2\pi} \int_{-\infty}^0 e^{-\sigma(k-k_0)^2} e^{ik(x+C_p t)} dk \quad (1.124b)$$

$$= \frac{A}{2\pi} \int_0^{\infty} e^{-\sigma(k-k_0)^2} e^{ik(x-C_p t)} dk + \frac{A}{2\pi} \int_0^{\infty} e^{-\sigma(k+k_0)^2} e^{-ik(x+C_p t)} dk \quad (1.124c)$$

$$= \Phi^{(+)}(x - C_p t, 0) + \Phi^{(-)}(x + C_p t, 0), \quad (1.124d)$$

where we made use of equations (1.110a) and (1.110b) for the  $\Phi^{(+)}$  and  $\Phi^{(-)}$  initial conditions. Evidently, the  $x$ -space wave packet splits from its initial condition into positive and negative propagating wave packets, and the packets move without altering their respective initial conditions. However, as already discussed for the initial conditions leading up to equation (1.111), the wave packets are *not* symmetric reflections of each other if  $k_0 \neq 0$ . In fact,  $\Phi^{(-)}$  is exponentially smaller than  $\Phi^{(+)}$  if  $k_0 > 0$ .

### Evolution of a Gaussian $\Phi^{(++)}$ packet of non-dispersive waves

Now consider evolution of the  $\Phi^{(++)}$  packet (1.103) built from non-dispersive waves. In this case we can perform the integral to produce

$$\Phi^{(++)}(x, t) = \frac{A}{2\pi} \int_{-\infty}^{\infty} e^{-\sigma(k-k_0)^2} e^{ik(x-C_p t)} dk \quad (1.125a)$$

$$= \frac{A}{2\pi} e^{ik_0(x-C_p t)} e^{-(x-C_p t)^2/(4\sigma)} \int_{-\infty}^{\infty} e^{-\sigma[q-i(x-C_p t)/(2\sigma)]^2} dq \quad (1.125b)$$

$$= \frac{A e^{ik_0(x-C_p t)}}{2\pi} \sqrt{\frac{\pi}{\sigma}} e^{-(x-C_p t)^2/(4\sigma)}. \quad (1.125c)$$

The second equality resulted from completing the square in the exponential, and the final equality evaluated the integral using methods from complex analysis.<sup>20</sup> Comparing the expression in equation (1.125c) to the initial wave packet in equation (1.113) reveals that

$$\Phi^{(++)}(x, t) = \Phi^{(++)}(x - C_p t, 0) = \Phi_o(x - C_p t). \quad (1.126)$$

As noted in Section 1.7.4, the initial condition for  $\Phi^{(++)}$  is propagated in the positive direction as a single coherent packet, which contrasts to the splitting found for the packet  $\Phi^{(+-)}$  in equation (1.124d).

### 1.7.8 Evolution of a dispersive Gaussian wave packet

Most geophysical waves are **dispersive waves**. As we see in the following, dispersive waves render a spreading of the  $x$ -space wave packet as it evolves, along with a decay in its amplitude. Furthermore, the packet moves at the **group velocity**, which, for dispersive waves, is distinct from the phase velocity. To reveal these properties analytically, we here characterize evolution of the  $\Phi^{(+-)}$  wave packet when constructed with dispersive waves, following steps taken for the general case in Section 1.6.6. Though straightforward, the maths is more tedious than for the non-dispersive packets discussed in Section 1.7.7.

#### Taylor expanding the phase function

To examine the propagation of a dispersive wave packet, we Taylor expand the dispersion relation around the central wavenumber,  $k = k_o$ , and assume the packet is relatively narrow band in wavevector space. With these assumptions we truncate the Taylor series at second order

$$\varpi(k) \approx \varpi(k_o) + (k - k_o) \left[ \frac{d\varpi}{dk} \right]_{k=k_o} + \frac{(k - k_o)^2}{2} \left[ \frac{d^2\varpi}{dk^2} \right]_{k=k_o} \quad (1.127a)$$

$$\equiv \omega_o + (k - k_o) c_g + (k - k_o)^2 \mu. \quad (1.127b)$$

We here introduced the one-dimensional group velocity evaluated at  $k = k_o$

$$c_g = \hat{x} c_g = \hat{x} \left[ \frac{d\varpi}{dk} \right]_{k=k_o}. \quad (1.128)$$

Recall from its definition in equation (1.26) that the phase speed is positive by definition,  $C_p > 0$ . In contrast,  $c_g$  can be positive or negative, or even zero, depending on the slope of the dispersion relation at  $k = k_o$ . In equation (1.127b) we also introduced one-half the second derivative of the dispersion relation (dimensions of  $L^2/T$ )

$$\mu = \frac{1}{2} \left[ \frac{d^2\varpi}{dk^2} \right]_{k=k_o}. \quad (1.129)$$

As shown below, a nonzero  $\mu$  leads to an  $x$ -space modification of the wave packet shape, and generally leading to a reduction in the packet's amplitude (regardless the sign of  $\mu$ ). It thus acts as a kinematic diffusivity of the wave packet.

---

<sup>20</sup>More precisely, it makes use of the calculus of residues, with Section G<sub>I</sub> in *Cohen-Tannoudji et al. (1977)* providing a discussion and Appendix A9 of *Fetter and Walecka (2003)* providing a tutorial for physicists.

## Performing the integral

Making use of the Taylor expansion (1.127b) yields the approximate form of the phase function appearing in the  $\mathbf{x}$ -space wave packet (1.122)

$$k x - \varpi(k) t = (k - k_0) x + k_0 x - t [\omega_0 + c_g (k - k_0) + \mu (k - k_0)^2] \quad (1.130a)$$

$$= (k_0 x - \omega_0 t) + (x - c_g t) (k - k_0) - t \mu (k - k_0)^2. \quad (1.130b)$$

As a result, the  $\mathbf{x}$ -space wave packet takes on the approximate form

$$\Phi^{(+)}(x, t) \approx \frac{A e^{i(k_0 x - \omega_0 t)}}{2\pi} \int_{-\infty}^{\infty} e^{-(\sigma + i\mu t)(k - k_0)^2 + i(k - k_0)(x - c_g t)} dk \quad (1.131a)$$

$$= \frac{A e^{i(k_0 x - \omega_0 t)}}{2\pi} \int_{-\infty}^{\infty} e^{-(\sigma + i\mu t)q^2 + iq(x - c_g t)} dq, \quad (1.131b)$$

where the second equality made use of the same substitution,  $q = k - k_0$ , used when evaluating the integral for the initial value wave packet in Section 1.7.5. Introducing the shorthand

$$\alpha = x - c_g t \quad \text{and} \quad \beta = \sigma + i\mu t, \quad (1.132)$$

proves useful for completing the square in the integral exponential

$$-(\sigma + i\mu t)q^2 + iq(x - c_g t) = -\beta [q - i\alpha/(2\beta)]^2 - \alpha^2/(4\beta), \quad (1.133)$$

which brings the  $\mathbf{x}$ -space wave packet (1.131b) to the form<sup>21</sup>

$$\Phi^{(+)}(x, t) = \frac{A e^{i(k_0 x - \omega_0 t)} e^{-\alpha^2/(4\beta)}}{2\pi} \int_{-\infty}^{\infty} e^{-\beta[q - i\alpha/(2\beta)]^2} dq \quad (1.134a)$$

$$= \frac{A e^{i(k_0 x - \omega_0 t)}}{2\pi} \sqrt{\frac{\pi}{\beta}} e^{-\alpha^2/(4\beta)}. \quad (1.134b)$$

Using the definition (1.132) for  $\alpha$  and  $\beta$  exposes a complex exponential multiplied by a real exponential

$$\Phi^{(+)}(x, t) = \frac{A e^{i(k_0 x - \omega_0 t)}}{2\sqrt{\pi}} \frac{e^{-[(x - c_g t)^2/(4(\sigma + i\mu t))]} }{\sqrt{\sigma + i\mu t}} \quad (1.135a)$$

$$= \frac{A e^{i(k_0 x - \omega_0 t)}}{2\sqrt{\pi}} \frac{e^{i[(x - c_g t)^2 \mu t / 4(\sigma^2 + (\mu t)^2)]} e^{-\frac{(x - c_g t)^2 \sigma}{4(\sigma^2 + (\mu t)^2)}}}{\sqrt{\sigma + i\mu t}} \quad (1.135b)$$

$$= \frac{A e^{i(k_0 x - \omega_0 t)}}{2\sqrt{\pi}} \frac{e^{i[\frac{(x - c_g t)^2 \mu t}{4(\sigma^2 + (\mu t)^2)} + \varphi/2]} e^{-\frac{(x - c_g t)^2 \sigma}{4(\sigma^2 + (\mu t)^2)}}}{[\sigma^2 + (\mu t)^2]^{1/4}}, \quad (1.135c)$$

where  $\tan \varphi = -\mu t/\sigma$ .

---

<sup>21</sup>The integral in equation (1.134b) is evaluated using complex analysis as discussed in Section G<sub>I</sub> of [Cohen-Tannoudji et al. \(1977\)](#) as well as Section 54 and Appendix A9 of [Fetter and Walecka \(2003\)](#).

### Interpreting the wave packet

The phase factor in the  $\mathbf{x}$ -space wave packet (1.135c)

$$\mathcal{P} = (k_0 x - \omega_0 t) + \frac{(x - c_g t)^2 \mu t}{4(\sigma^2 + (\mu t)^2)} + \varphi/2, \quad (1.136)$$

equals to that for a non-dispersive wave,  $k_0 x - \omega_0 t$ , plus a space and time dependent phase shift that is nonzero for dispersive waves ( $\mu \neq 0$ ). The phase shift simplifies for points following the group velocity,  $x = c_g t$ , and it is here that the Gaussian exponential is maximized

$$\Phi^{(+)}(x = c_g t, t) = \frac{A e^{i(k_0 x - \omega_0 t + \varphi/2)}}{2\sqrt{\pi} [\sigma^2 + (\mu t)^2]^{1/4}}. \quad (1.137)$$

Evidently, as the  $\mathbf{x}$ -space wave packet moves with the group velocity (which can be either signed), its amplitude declines according to  $|\mu t|^{-1/2}$ . The  $t^{-1/2}$  decay of the packet amplitude is also found from the method of stationary phase for dispersive packets considered in Section 1.8.3.

### Time dependent uncertainty relation

As for the  $\mathbf{k}$ -space packet in Section (1.7.5), we consider the squared modulus of the  $\mathbf{x}$ -space packet (1.135c) as a measure of its intensity

$$|\Phi^{(+)}(x, t)|^2 = \frac{A^2 e^{-\frac{(x - c_g t)^2 \sigma}{2(\sigma^2 + (\mu t)^2)}}}{4\pi [\sigma^2 + (\mu t)^2]^{1/2}}. \quad (1.138)$$

The e-folding width of  $|\Phi(x, t)|^2$  is revealed by setting the decaying exponential to unity, which leads to

$$x = c_g t \pm (2\sigma)^{1/2} [1 + (\mu t/\sigma)^2]^{1/2} \implies \Delta x^{\text{efold}} = 2(2\sigma)^{1/2} [1 + (\mu t/\sigma)^2]^{1/2}. \quad (1.139)$$

Multiplying by the time-independent  $\mathbf{k}$ -space packet width (1.108) leads to the time dependent uncertainty relation

$$\Delta x^{\text{efold}} \Delta k^{\text{efold}} = 2(2\sigma)^{1/2} [1 + (\mu t/\sigma)^2]^{1/2} 2(2\sigma)^{-1/2} = 4 [1 + (\mu t/\sigma)^2]^{1/2}. \quad (1.140)$$

The time dependent uncertainty relation starts from its initial condition (1.115) at  $t = 0$ , and then grows as  $t^{1/2}$ . For non-dispersive waves ( $\mu = 0$ ), the uncertainty relation is time-independent, which is expected since non-dispersive waves translate the initial packet without changing the properties of the packet.

### 1.7.9 The non-dispersive limit of a dispersive packet

We set the spreading parameter,  $\mu$ , to zero for non-dispersive waves, in which case  $\mu = 0 \implies \varphi = 0$ . In this limit the wave packet (1.135c) takes on the form

$$\mu = 0 \implies \Phi^{(+)}(x, t) = \frac{A e^{i k_0 (x - c_g t)}}{2\pi} \sqrt{\frac{\pi}{\sigma}} e^{-(x - c_g t)^2 / (4\sigma)}. \quad (1.141)$$

If we set  $c_g = C_p$  then this result corresponds to equation (1.125c), which is the non-dispersive form of the packet  $\Phi^{(++)}(x, t)$ .

So why did the non-dispersive limit not reduce to equation (1.124d), which is the non-dispersive form of  $\Phi^{(+-)}(x, t)$ ? The reason is that when performing the Taylor series expansion for the dispersive packet, we only picked out the group velocity at the single wavenumber,  $k = k_o$ . Hence, we can only get one of the two packets comprising  $\Phi^{(+-)}(x, t)$  in equation (1.124d). Namely, we ignore the exponentially small packet using the Taylor series approach. Furthermore, note that  $c_g$  can be positive or negative for dispersive waves (indeed, it can even be zero). In contrast,  $C_p$  is the phase speed and that is always positive. So in the case of  $c_g = -C_p < 0$  then equation (1.141) is a negative moving wave packet.

So in summary, the non-dispersive limit of a dispersive packet does not exactly correspond to the purely non-dispersive packet due to a few subtleties. In particular, if one cares about the exponentially small packet exposed with the non-dispersive analysis in Section 1.7.7, then it is necessary to follow the approach taken in that section rather than taking the non-dispersive limit of dispersive waves as considered here.

### 1.7.10 Comments and further study

The study of wave packets in this section revealed properties that appear throughout the study of waves. First, there is the [uncertainty relation](#), whereby a packet that is narrow banded in wavevector space is broad banded in position space, and conversely. Second, dispersive wave packets have a modulation function that is modified with time, thus producing a time dependent uncertainty relation. Regardless the sign of the dispersion coefficient,  $\mu$  (equation (1.129)), the packet amplitude decays in time and the uncertainty grows. We see this behavior for generic wave packets when studying the stationary phase method in Section 1.8. Third, the center of the packet moves with the group velocity rather than phase velocity, with the two velocities distinct for dispersive waves. We also encountered this property in Section 1.6.7 when studying how the wave packet modulation function evolves. It will appear again in Chapter 2 when studying how energy (or more generally the wave action) propagates with the wave field.

The study of wave packets and their evolution is a central concern of quantum mechanics. Most books on the subject have a discussion of quantum wave packets at the level discussed here. In particular, we made use in this section of Chapter 3 of [Bohm \(1951\)](#), Chapter 2 of [Gasiorowicz \(1974\)](#), and Section G<sub>I</sub> of [Cohen-Tannoudji et al. \(1977\)](#). We also followed Section 54 of [Fetter and Walecka \(2003\)](#), who consider wave packets built from surface gravity waves. However, it is notable that the distinctions we made here between positive and negative moving packets in Section 1.7.7 are not considered by this literature. They are, however, consistent with the treatment in [Stakgold \(2000b\)](#).

## 1.8 Method of stationary phase

As shown in Section 1.6.8, a packet of non-dispersive waves holds the spacial structure of its initial condition unchanged as it propagates the signal at the phase speed. This evolution is exact. Hence, we know everything about a packet of non-dispersive waves for all time, given the phase velocity and the initial conditions.

The situation is more complicated for packets of dispersive waves, whose modulation function changes its shape due to wave dispersion. We encountered such behavior when introducing wave trains and wave packets in Section 1.6.6, as well as when studying Gaussian wave packets

of dispersive waves in Section 1.7. In the present section we study the long-time behavior for a packet of dispersive waves using the method of stationary phase. From the Gaussian wave packet study in Section 1.7.8, we expect the packet's amplitude to decay according to  $t^{-1/2}$ , as revealed by the exact Gaussian packet results in Section 1.7. Here we show that this  $t^{-1/2}$  behavior is generic for packets of dispersive waves.

For analytical simplicity we focus on a wave function in one space dimension, written as

$$\Phi(x, t) = \frac{1}{2\pi} \int_{k_a}^{k_b} \mathcal{A}(k) e^{i h(k) t} dk, \quad (1.142)$$

where  $\mathcal{A}(k)$  is the amplitude function,  $h(k)$  is a real phase function, and we focus on a wavenumber interval  $k \in [k_a, k_b]$ . For example, with a plane wave we have

$$h(k) = k x/t - \varpi(k), \quad (1.143)$$

where  $\varpi(k)$  is the dispersion relation. Note that for the integral in equation (1.142),  $x$  and  $t$  are considered parameters, so we only expose the  $k$  dependence to the phase function,  $h(k)$ . Finally, with a focus on dispersive waves we assume

$$\varpi''(k) \neq 0. \quad (1.144)$$

### 1.8.1 Riemann-Legesque lemma and center of the packet

As  $t \rightarrow \infty$ , the integrand,  $\mathcal{A}(k) e^{i h(k) t}$  of the wave packet (1.142) oscillates faster. Consequently, terms in the integral cancel since the oscillations dominate any behavior of the amplitude (which is assumed smooth). Indeed, the Riemann-Legesque lemma states that the integral has a zero limit as  $t \rightarrow \infty$ , so long as the amplitude function is integrable and finite.

Contributions to the wave packet are maximized when two elements of the integrand align. First we want to maximize the amplitude, which for a wave packet is assumed to be maximized in a small interval surrounding a wavenumber,  $k_o$ . Next we want the phase function to be in a small neighborhood of an extrema, which is where  $h'(k) = 0$ . Aligning this phase extrema with the amplitude maximum means we want space-time locations where  $h'(k_o) = 0$ . For a packet built from plane waves,  $h'(k_o) = 0$  occurs at a specific space-time point determined by the group velocity

$$h'(k_o) = x/t - c_g = 0 \implies x/t = c_g, \quad (1.145)$$

where

$$c_g = c_g(k_o) = \varpi'(k_o) \quad (1.146)$$

is the group velocity for the packet as determined at the wavenumber,  $k_o$ , where the amplitude has its maximum. Evidently, the center of the wave packet (i.e., where the packet has its maximum amplitude) is located at

$$x_{\text{center}} = t c_g, \quad (1.147)$$

so that the center moves at the group velocity. This result for the packet center was previously found using different approaches in Sections 1.6.6, 1.7.7, and 1.7.8.

In addition to wanting information about the packet center, it is useful to know about its amplitude which, according to Riemann-Legesque, decays to zero as time increases. To get an expression for the amplitude modulation requires some work, which is the topic of the remainder of this section.

### 1.8.2 Wavenumber intervals with no phase extrema

Consider the case whereby the phase has no extrema within the chosen wavenumber interval,  $k \in [k_a, k_b]$ , so that  $h'(k) \neq 0$ . This assumption allows us to use integration by parts in the form

$$\int_{k_a}^{k_b} v \, du = [uv] \Big|_{k_a}^{k_b} - \int_{k_a}^{k_b} u \, dv, \quad (1.148)$$

where

$$v = \mathcal{A}/h' \quad \text{and} \quad u = e^{ih t}/(i t), \quad (1.149)$$

so that

$$v \, du = (\mathcal{A}/h') d[e^{ih t}/(i t)] = \mathcal{A} e^{ih t} \, dk, \quad (1.150)$$

and

$$u \, dv = [e^{ih t}/(i t)] d(\mathcal{A}/h'). \quad (1.151)$$

We can thus write the integral as

$$2\pi \Phi = \frac{1}{i t} \left[ (\mathcal{A}/h') e^{ih t} \Big|_{k_a}^{k_b} - \int_{k_a}^{k_b} e^{ih t} d(\mathcal{A}/h') \right] \sim \mathcal{O}(t^{-1}). \quad (1.152)$$

The key point is that the integral decays as  $t^{-1}$  for wavenumber intervals,  $k \in [k_a, k_b]$ , where the phase,  $h(k)$ , has no extrema.

### 1.8.3 Wavenumber interval including a phase extrema

We expect that contributions from regions near an extrema decay slower in time, since near those regions the phase does not oscillate so rapidly.<sup>22</sup> In particular, as noted in Section 1.8.1, we expect that the packet center follows a point in space determined by the group velocity. We thus consider the case where the phase has an extrema at the wavenumber  $k_o$ , which is now included in the interval:  $k_o \in [k_a, k_b]$ , in which case we Taylor expand the phase

$$h(k) = h(k_o) + h'(k_o)(k - k_o) + (k - k_o)^2 h''(k_o)/2 + \dots = h(k_o) + (k - k_o)^2 h''(k_o)/2 + \dots, \quad (1.153)$$

where we set

$$h'(k_o) = 0 \quad (1.154)$$

since it is an extrema. We showed in Section 1.8.2 that wavenumber regions where there is no extrema contribute terms of order  $\mathcal{O}(t^{-1})$  to the integral. We now show that regions including an extrema decay like  $\mathcal{O}(t^{-1/2})$ , thus allowing us to focus on the region surrounding the packet center,

$$k_o - 1/\sqrt{2\sigma} \leq k \leq k_o + 1/\sqrt{2\sigma}. \quad (1.155)$$

We introduced  $1/\sqrt{2\sigma}$  as a measure of the packet width, such as used for the Gaussian packet in equation (1.108). We thus have the integral

$$2\pi \Phi = e^{it h(k_o)} \int_{k_o - 1/\sqrt{2\sigma}}^{k_o + 1/\sqrt{2\sigma}} \mathcal{A}(k) e^{it(k - k_o)^2 h''(k_o)/2} \, dk + \mathcal{O}(t^{-1}). \quad (1.156)$$

<sup>22</sup>According to Section 11.3 of [Whitham \(1974\)](#), this insight is originally due to Lord Kelvin.

Changing variables to a shifted wavenumber  $\ell = k - k_{\circ}$  gives the expression

$$2\pi\Phi = e^{it h(k_{\circ})} \int_{-1/\sqrt{2\sigma}}^{1/\sqrt{2\sigma}} \mathcal{A}(\ell + k_{\circ}) e^{it \ell^2 h''(k_{\circ})/2} d\ell + \mathcal{O}(t^{-1}). \quad (1.157)$$

One more change in variables to  $p = \ell \sqrt{t}$  renders the integral

$$2\pi\Phi = t^{-1/2} e^{it h(k_{\circ})} \int_{-\sqrt{t/2\sigma}}^{\sqrt{t/2\sigma}} \mathcal{A}(pt^{-1/2} + k_{\circ}) e^{ip^2 h''(k_{\circ})/2} dp + \mathcal{O}(t^{-1}). \quad (1.158)$$

We simplify the integrand by performing a Taylor expansion

$$\mathcal{A}(pt^{-1/2} + k_{\circ}) = \mathcal{A}(p_0) + \mathcal{O}(t^{-1/2}), \quad (1.159)$$

so that, keeping only terms of order  $t^{-1/2}$ , brings about the integral

$$2\pi\Phi = t^{-1/2} \mathcal{A}(k_{\circ}) e^{it h(k_{\circ})} \int_{-\sqrt{t/2\sigma}}^{\sqrt{t/2\sigma}} e^{ip^2 h''(k_{\circ})/2} dp + \mathcal{O}(t^{-1}). \quad (1.160)$$

For a dispersive packet with non-zero  $h''(k_{\circ})$ , the  $t \rightarrow \infty$  limit allows us to extend the limit on the integral (1.160) to infinity so that

$$2\pi\Phi = t^{-1/2} \mathcal{A}(k_{\circ}) e^{it h(k_{\circ})} \int_{-\infty}^{\infty} e^{ip^2 h''(k_{\circ})/2} dp + \mathcal{O}(t^{-1}). \quad (1.161)$$

The integral is in the form of a [Fresnel integral](#) and it can be done using methods from complex analysis to find

$$\int_{-\infty}^{\infty} e^{ip^2 h''(k_{\circ})/2} dp = \sqrt{\frac{2\pi}{|h''(k_{\circ})|}} e^{\pm i\pi/4}, \quad (1.162)$$

where the  $\pm$  sign corresponds to the sign of  $h''(k_{\circ})$ . We thus have the stationary phase expression for the wave function

$$\Phi(x, t) = \frac{\mathcal{A}(k_{\circ}) e^{it h(k_{\circ}) \pm i\pi/4}}{[2\pi t |h''(k_{\circ})|]^{1/2}} + \mathcal{O}(t^{-1}). \quad (1.163)$$

As noted at the start of this section, a packet built from plane waves has

$$h(k_{\circ}) t = k_{\circ} x - \varpi(k_{\circ}) t, \quad (1.164)$$

so that the long-time behavior of the wave packet is given by

$$\Phi(x, t) = \mathcal{A}(k_{\circ}) e^{i(k_{\circ} x - \varpi(k_{\circ}) t)} \frac{e^{\pm i\pi/4}}{[2\pi t |\varpi''(k_{\circ})|]^{1/2}} + \mathcal{O}(t^{-1}). \quad (1.165)$$

This wave function is built from a plane wave with wavenumber,  $k_{\circ}$ , that is modulated by a function whose amplitude is decaying according to  $t^{-1/2}$ . The  $t^{-1/2}$  decay accords with the exact solution of the dispersive Gaussian packet in Section 1.7.8. Furthermore, note how the strength of the  $t^{-1/2}$  decay is affected by the size of  $|\varpi''(k_{\circ})|$ .

As noted in Section 1.8.1, the packet is a maximum when sampled at the center, which is given by  $x_{\text{center}} = c_g t$  (equation (1.147)). At  $x_{\text{center}}$ , the phase of the wave function (1.165) takes

the form

$$k_o x_{\text{center}} - \varpi(k_o) t = k_o c_g t - \varpi(k_o) t = k_o (c_g - c_p) t, \quad (1.166)$$

where we introduced the phase velocity for the carrier wave

$$c_p = \varpi(k_o)/k_o. \quad (1.167)$$

As a result, the wave function as evaluated at the packet center is given by

$$\Phi(x_{\text{center}}, t) = \mathcal{A}(k_o) e^{i k_o (c_g - c_p) t} \frac{e^{\pm i \pi/4}}{[2 \pi t |\varpi''(k_o)|]^{1/2}} + \mathcal{O}(t^{-1}). \quad (1.168)$$

#### 1.8.4 Comments and further study

If both the first and second derivatives vanish,  $h'(k_o) = 0$  and  $h''(k_o) = 0$ , then the same procedure as used above must be pursued but to the next higher order in the Taylor series expansion. Also if there are multiple extrema, then each will add a contribution of the form given by equation (1.163).

Variations on the derivation given in this section can be found in section 11.3 of [Whitham \(1974\)](#), section 3.7 of [Lighthill \(1978\)](#), section 55 of [Fetter and Walecka \(2003\)](#) and section 1.C.2 of [Cohen-Tannoudji et al. \(1977\)](#).



## 1.9 Exercises

### EXERCISE 1.1: PARSEVAL'S IDENTITY FOR WAVE PACKETS

Consider the expression (1.48) for a wave function

$$\Phi(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int \mathcal{A}(\mathbf{k}) e^{i(\mathbf{k} \cdot \mathbf{x} - \varpi t)} d\mathbf{k}. \quad (1.169)$$

Prove the following form of Parseval's identity

$$\int |\nabla \Phi|^2 d\mathbf{x} = \frac{1}{(2\pi)^3} \int |\mathbf{k}|^2 |\mathcal{A}(\mathbf{k})|^2 d\mathbf{k}. \quad (1.170)$$

Notice that the right hand side is time-independent, so that the left hand side must be likewise.  
Hint: make use of the following representation of the Dirac delta

$$\delta(\mathbf{k} - \mathbf{q}) = \frac{1}{(2\pi)^3} \int e^{i(\mathbf{k} - \mathbf{q}) \cdot \mathbf{x}} d\mathbf{x}. \quad (1.171)$$

### EXERCISE 1.2: SQUARE WAVE PACKET

Consider a wave packet in one space dimension with real  $k$ -space amplitude function

$$A(k) = \begin{cases} 0 & k < -K \\ N & -K < k < K \\ 0 & k > K, \end{cases} \quad (1.172)$$

where  $K > 0$ .

- (a) Find the  $\mathbf{x}$ -space wave function, which in one space dimension takes the form

$$\Phi_o(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk. \quad (1.173)$$

Hint:  $\Phi_o(x)$  is the inverse Fourier transform of  $A(k)$ . The integral is simple to do.

- (b) Find the value of  $N$  for which

$$\int_{-\infty}^{\infty} |\Phi_o(x)|^2 dx = 1. \quad (1.174)$$

Hint: massage the integral until it looks like one found in a standard integral table.

- (c) Relate the above choice for  $N$  to one that makes

$$\int_{-\infty}^{\infty} |A(k)|^2 dk = 1. \quad (1.175)$$

- (d) Show that a reasonable definition for  $\Delta x$  in part (a) yields the uncertainty relation

$$\Delta x \Delta k > 1. \quad (1.176)$$

Hint: this uncertainty relation holds independently of  $K$ .

#### EXERCISE 1.3: INVERSE SQUARED WAVE PACKET

Consider a wave packet in one space dimension with real  $k$ -space amplitude function

$$A(k) = \frac{N}{k^2 + \alpha^2} \quad (1.177)$$

with  $\alpha > 0$ .

- (a) Find the  $\mathbf{x}$ -space wave function, which in one space dimension takes the form

$$\Phi_o(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(k) e^{ikx} dk \quad (1.178)$$

Hint:  $\Phi_o(x)$  is the inverse Fourier transform of  $A(k)$ . Massage the integral until it looks like one found in a standard integral table.

- (b) Show that a reasonable definition for  $\Delta x$  in part (a) yields the uncertainty relation

$$\Delta x \Delta k > 1. \quad (1.179)$$

Hint: this relation holds independently of  $\alpha$ .

#### EXERCISE 1.4: WAVE FUNCTION PDE DERIVED FROM THE DISPERSION RELATION

Follow the method from Section 1.6.9 to derive the partial differential equation satisfied by a wave function,  $\Phi(\mathbf{x}, t)$ , whose constituent waves satisfy the following dispersion relations, where for each case  $\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}$  a three-dimensional wavevector.

- (a)  $\varpi^2 = c^2 |\mathbf{k}|^2$ , with  $c$  a constant with dimensions  $L T^{-1}$ .
- (b)  $\varpi^2 = B^2 k_z^2 / |\mathbf{k}|^2$ , with  $B$  a constant with dimensions  $T^{-1}$ .



## Chapter 2

### WAVES IN A GENTLY VARYING BACKGROUND

We here study linear waves propagating through a prescribed gently varying (in space and time) background environment, such as for waves moving through a stratified ocean or atmosphere, or waves moving through a mean flow. We assume the relevance of space and time coherent wave patterns so that we can consider a wave phase function. However, the wavevector and angular frequency, as well as the wave amplitude, are here functions of space and time and as such we cannot use traditional Fourier analysis. Instead, asymptotic methods are needed, and we develop the leading order theory referred to in various contexts as *ray theory*, *geometric optics*, *eikonal approximation* or the *WKB approximation*.<sup>1</sup> Notably, we do *not* consider the back-reaction effects of waves on the background state, which is the subject matter of *waves and mean flow interaction* theory.

Given the more general background state, we must consider a wave ansatz that is more general than the plane wave Fourier ansatz from Chapter 1. We refer to the new ansatz as the *WKB wave ansatz* or equivalently the *eikonal wave ansatz* (the specific form is given by equation (2.54)). Plugging in this wave ansatz to the wave equation then leads to a suite of asymptotic equations that are used to build the wave function, ray equations, energy equation, etc. In principle this approach is straightforward, yet in practice it is tedious and uninspired. *Whitham's variational principle* provides an alternative that is both elegant and powerful. In brief, Whitham's principle is based on Hamilton's principle from classical continuum mechanics (Part ?? of this book), only now applied to the leading order (in WKB expansion) phase averaged action. The resulting Euler-Lagrange equations render the dispersion relation and conservation equations for an energy-momentum-stress tensor. As part of these conservation equations, we encounter the *wave action*, which offers a generalization of wave energy, and the flux of wave action is determined by the group velocity.

#### READER'S GUIDE TO THIS CHAPTER

We assume familiarity with the material from Chapter 1, which considered wave kinematics for traveling plane waves on a static and homogeneous background state. We here make use of Cartesian coordinates and Cartesian tensors (Chapters ?? and ??). The wave energetics/action discussion makes use of Hamilton's principle for a continuum from Part ?? of this book, in particular the material in Chapter ?. Much of the material here is inspired by the review chapter from [Bretherton \(1971\)](#), as well as [Bretherton and Garrett \(1969\)](#), chapter 11 of [Whitham \(1974\)](#), Part 2 in the Epilogue of [Lighthill \(1978\)](#), [Andrews and McIntyre \(1978\)](#),

<sup>1</sup>WKB stands for Wentzel, Kramers, and Brillouin, who are scientists promoting the method in the 20th century for use in quantum mechanics. Often this method is referred to as WKBJ in reference to the additional work of Jeffreys. Notably, the method was developed in the 19th century by Liouville and Green.

chapter 9 of *Olbers et al.* (2012), and *Tracy et al.* (2014).

<b>2.1</b>	<b>Loose threads</b>	<b>44</b>
<b>2.2</b>	<b>General phase functions</b>	<b>44</b>
2.2.1	Path independence of phase difference	45
2.2.2	Conservation of wave crests	46
2.2.3	Phase velocity and phase speed	47
2.2.4	Angular frequency and wavelength	47
<b>2.3</b>	<b>Kinematics of rays</b>	<b>48</b>
2.3.1	Eikonal equation	48
2.3.2	Rays are integral curves of the group velocity	49
2.3.3	Evolution of $\mathbf{k}$ along a ray	49
2.3.4	Evolution of $\omega$ along a ray	50
2.3.5	Changes in the phase following a ray	51
2.3.6	Summary of the ray equations	51
2.3.7	Comments and further study	52
<b>2.4</b>	<b>Hamilton's principle and the Euler-Lagrange equation</b>	<b>52</b>
2.4.1	Stationary action $\iff$ Euler-Lagrange equation	52
2.4.2	Hamiltonian density and energy of the continuum	53
2.4.3	Stress-energy-momentum tensor	53
<b>2.5</b>	<b>Whitham's variational principle</b>	<b>53</b>
2.5.1	Some motivation	54
2.5.2	The stretched string	54
2.5.3	Space and time scale separation	55
2.5.4	The leading order phase averaged action	56
2.5.5	Whitham's variational principle for the string	57
2.5.6	Phase averaged Hamiltonian	58
2.5.7	A general statement of Whitham's variational principle	58
2.5.8	Interpreting the string's wave action conservation equation	59
2.5.9	Comments	60
<b>2.6</b>	<b>Variational methods for self-adjoint wave operators</b>	<b>60</b>
2.6.1	Self-adjoint linear wave operators	60
2.6.2	Varying the action	61
2.6.3	Comments on the method	62

## 2.1 Loose threads

- Figures
- Adiabatic invariants as per Section 9.2 of *Olbers et al.* (2012) and Section 49 of *Landau and Lifshitz* (1976). See also *José and Saletan* (1998).

## 2.2 General phase functions

Throughout Chapter 1, the phase of the wave function takes the linear plane wave form given by equation (1.20) (here written with zero phase shift)

$$\mathcal{P}(\mathbf{x}, t) = \mathbf{k} \cdot \mathbf{x} - \omega t = k_a x^a - \omega t, \quad (2.1)$$

where the wavevector,  $\mathbf{k}$ , and angular frequency,  $\omega$ , are parameters of the wave that are independent of space and time. The second equality made use of the summation convention, with the wavevector components written with lower indices to accord with the upstairs spatial indices. The relatively simple form (2.1) of the phase function is suited only for background states that are homogeneous and static. In this chapter we consider a more general phase function that allows us to study waves moving through a background state that is inhomogeneous and/or time dependent. The key restriction to our approach is that we retain the notion of a locally defined wavevector and angular frequency, which are now considered to be functions of space and time

$$\nabla \mathcal{P} \equiv \mathbf{k}(\mathbf{x}, t) \quad \text{and} \quad -\partial_t \mathcal{P} \equiv \omega(\mathbf{x}, t). \quad (2.2)$$

These expressions tacitly assume that the base state fluid properties are changing slowly in space and time relative to the wave phase, thus allowing us to generalize much of the wave kinematics holding for homogeneous/static media while locally considering the waves to be planar. For this assumption to hold, it is sufficient to make the following space and time scale separation, as assumed throughout this chapter

$$|\mathbf{k}| = |\nabla \mathcal{P}| \gg L^{-1} \quad \text{and} \quad \omega = -\partial_t \mathcal{P} \gg T^{-1}, \quad (2.3)$$

where  $L$  is a characteristic length scale defined by spatial variations in the background state, and  $T$  is the corresponding time scale for changes in the background state. In terms of the local measure of the wavelength,  $\Lambda = 2\pi/|\mathbf{k}|$ , and period,  $2\pi/\omega$ , we have

$$\Lambda \ll 2\pi L \quad \text{and} \quad 2\pi/\omega \ll T. \quad (2.4)$$

Values for  $L$  and  $T$  depend on details of the physical system defining the background state. For example, in Section 2.5.2 we consider  $L$  and  $T$  for a stretched string with time dependent tension and space dependent mass density, and in Section 3.9.5 we consider  $L$  for the case of acoustic waves moving through a spatially inhomogenous yet static background.

### 2.2.1 Path independence of phase difference

Given that the wavevector is defined as the gradient of the phase as per equation (2.2), it must satisfy the consistency condition

$$\nabla \times \mathbf{k} = \nabla \times \nabla \mathcal{P} = 0. \quad (2.5)$$

This property of the wavevector means that there are the same number of wave crests between any two points in the fluid at any particular time instance, no matter what path is taken to connect the two points. This property is trivially maintained by plane waves in a homogeneous media since  $\mathbf{k}$  is a space-time constant vector. To prove it holds for the more general phase function, consider the difference in phase (at a fixed time) between points  $A$  and  $B$  within the fluid (see Figure 2.1). Compute this phase difference via a path,  $\mathcal{C}_1$ , that goes from point  $A$  to point  $B$ , and then via an alternative path,  $\mathcal{C}_2$ , that also goes from point  $A$  to point  $B$

$$\Delta \mathcal{P}_{\mathcal{C}_1} = \int_{\mathcal{C}_1} \mathbf{k} \cdot d\mathbf{x} \quad \text{and} \quad \Delta \mathcal{P}_{\mathcal{C}_2} = \int_{\mathcal{C}_2} \mathbf{k} \cdot d\mathbf{x}. \quad (2.6)$$

The path  $\mathcal{C}_1 - \mathcal{C}_2$  is a closed loop that is oriented in the counterclockwise direction. We can thus make use of Stokes' theorem to find

$$\Delta\mathcal{P}_{\mathcal{C}_1} - \Delta\mathcal{P}_{\mathcal{C}_2} = \oint_{\mathcal{C}_1 - \mathcal{C}_2} \mathbf{k} \cdot d\mathbf{x} = \oint (\nabla \times \mathbf{k}) \cdot \hat{\mathbf{n}} dS = 0, \quad (2.7)$$

which proves the path independence of the phase difference.

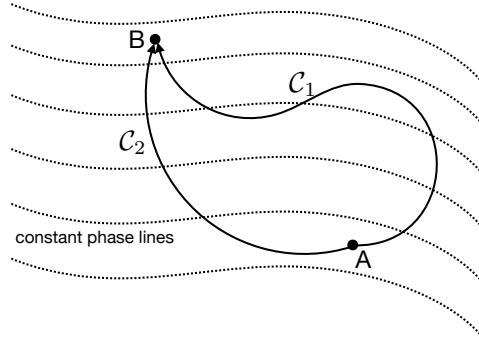


FIGURE 2.1: This figure illustrates the conservation of wave crests embodied by equation (2.9). Here we consider two paths in space,  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , that go from point  $A$  to point  $B$ . Both paths move across surfaces of constant phase, illustrated by the dotted lines and shown for a particular time instance. As shown by equation (2.7), the difference in phase computed along the two different paths is the same, and this property follows from  $\nabla \times \mathbf{k} = 0$  and Stokes' theorem.

## 2.2.2 Conservation of wave crests

Because the mixed partial derivatives of the phase function commute, we have

$$\frac{\partial \omega}{\partial x^a} = -\frac{\partial^2 \mathcal{P}}{\partial t \partial x^a} \quad \text{and} \quad \frac{\partial k_a}{\partial t} = \frac{\partial^2 \mathcal{P}}{\partial x^a \partial t}, \quad (2.8)$$

which leads to the vector equation

$$\frac{\partial \mathbf{k}}{\partial t} + \nabla \omega = 0. \quad (2.9)$$

This equation says that the time change in wavevector,  $\partial_t \mathbf{k}$ , is exactly compensated by the spatial change in angular frequency,  $\nabla \omega$ . This self-consistency condition is referred to as the *conservation of wave crests*. Motivation for this name follows since the wavenumber,  $|\mathbf{k}|$ , is the number of wave crests per unit length at a fixed time. Likewise, the angular frequency,  $\omega$ , is the number of wave crests passing a fixed location per unit time. Having their respective space and time derivatives match is a self-consistency condition for a coherent wave pattern to exist.

As a further means to understand the balance equation (2.9), and the name “conservation of wave crests”, consider an integral between two points fixed in space taken along a fixed path in space, such as the path  $\mathcal{C}_1$  in Figure 2.1. The time tendency of the phase difference is given by

$$\partial_t(\Delta\mathcal{P}_{\mathcal{C}_1}) = \frac{\partial}{\partial t} \int_{\mathcal{C}_1} \mathbf{k} \cdot d\mathbf{x} = \int_{\mathcal{C}_1} \partial_t \mathbf{k} \cdot d\mathbf{x} = - \int_{\mathcal{C}_1} \nabla \omega \cdot d\mathbf{x} = \omega(A) - \omega(B). \quad (2.10)$$

If the angular frequency is greater at point  $A$  than at point  $B$ , then that means that there is an accumulation of wave crests entering the region at point  $A$  relative to those leaving at point  $B$ ; i.e., there is a convergence of wave crests between the two points. This convergence is associated with an increase in the wavenumber between the two points.

### 2.2.3 Phase velocity and phase speed

Consider an observer moving on a smooth trajectory through space and time defined by a fixed point on a constant phase surface, such as when the observer remains fixed on the crest of a traveling wave.<sup>2</sup> As such, we assume that the trajectory is aligned in the direction of  $\nabla \mathcal{P}$ , which is the local direction of the wave

$$\hat{\mathbf{k}} \equiv \nabla \mathcal{P} / |\nabla \mathcal{P}| \equiv \mathbf{k} / |\mathbf{k}|, \quad (2.11)$$

with the local wavevector defined by equation (2.2).

To determine the velocity of the fixed-phase observer, consider an infinitesimal spatial increment,  $\delta \mathbf{x}^{\text{phase}}$ , that occurs over a small time increment,  $\delta t$ . Assuming this space increment follows the fixed-phase observer leads to the identity

$$\mathcal{P}(\mathbf{x} + \delta \mathbf{x}^{\text{phase}}, t + \delta t) = \mathcal{P}(\mathbf{x}, t). \quad (2.12)$$

Truncating a Taylor series expansion of this identity leads to the differential equation satisfied by the phase

$$(\partial_t + \mathbf{c}_p \cdot \nabla) \mathcal{P} = 0, \quad (2.13)$$

where we defined the *phase velocity*

$$\mathbf{c}_p = \delta \mathbf{x}^{\text{phase}} / \delta t. \quad (2.14)$$

The partial differential equation (2.13) specifies  $\mathbf{c}_p \cdot \nabla \mathcal{P} = \mathbf{c}_p \cdot \mathbf{k}$  in terms of the time derivative of the phase

$$\mathbf{c}_p \cdot \mathbf{k} = \mathbf{c}_p \cdot \nabla \mathcal{P} = -\partial_t \mathcal{P}. \quad (2.15)$$

Indeed, since the observer is assumed to move along the direction of the wavevector,  $\hat{\mathbf{k}}$ , then  $\mathbf{c}_p \cdot \hat{\mathbf{k}}$  is the only component of the phase velocity. We thus write the phase velocity as in our discussion of plane waves in Section 1.5.2

$$\mathbf{c}_p = (\mathbf{c}_p \cdot \hat{\mathbf{k}}) \hat{\mathbf{k}} \equiv C_p \hat{\mathbf{k}}, \quad (2.16)$$

where  $C_p > 0$  is the *phase speed*, which is the magnitude of the phase velocity. Figure 1.2 provides an illustration.

### 2.2.4 Angular frequency and wavelength

The time derivative,  $\partial_t \mathcal{P}$ , measures the time change of the phase at a fixed point in space, which we use to define the local angular frequency of the wave as per equation (2.2)

$$\omega = -\partial_t \mathcal{P}. \quad (2.17)$$

The differential equation (2.13) for the phase can thus be written in the equivalent manners

$$(\partial_t + \mathbf{c}_p \cdot \nabla) \mathcal{P} = 0 \iff \omega = \mathbf{c}_p \cdot \mathbf{k}, \quad (2.18)$$

---

<sup>2</sup>Our formulation in this section emulates that used in Section ?? for the Lagrangian time derivative, which is concerned with measuring fluid properties following a fluid particle. Here, we are following a wave as defined by a surface of constant phase.

which also leads to the relations

$$\mathbf{c}_p = C_p \hat{\mathbf{k}} = \frac{-\partial_t \mathcal{P}}{|\nabla \mathcal{P}|} \frac{\nabla \mathcal{P}}{|\nabla \mathcal{P}|}, \quad (2.19)$$

where  $C_p \geq 0$  is the phase speed. We likewise identify the wavelength

$$\Lambda = 2\pi/|\mathbf{k}| \implies C_p = \omega \Lambda / 2\pi. \quad (2.20)$$

Note that the relations (2.19) and (2.20) also hold for free plane waves moving through a homogeneous media, as discussed in Section 1.5.

## 2.3 Kinematics of rays

Recall that when studying wave packets in Section 1.6, we found the group velocity to be fundamental to the wave packet evolution. We show later in this chapter that the group velocity defines the paths along which wave action (wave energy divided by wave angular frequency) propagates. Such paths are referred to as *rays*. Given the prominence of rays, we find it useful to develop evolution equations for wave properties along rays. We here start that process, focusing on the wave phase, wavevector, and angular frequency. This material forms elements to *geometric optics*, which is a subject commonly taught in introductory physics courses by presenting a series of rules for how light reflects and refracts. In this section we derive those rules in the form of partial differential equations following from basic principles.

### 2.3.1 Eikonal equation

The dispersion relation provides the local value of the angular frequency as a function of space, time, and wavevector, where the wavevector itself is a function of space and time

$$\omega = \varpi(\mathbf{x}, t, \mathbf{k}(\mathbf{x}, t)) = \varpi(\mathbf{x}, t, \nabla \mathcal{P}). \quad (2.21)$$

As such, the dispersion relation is an explicit function of  $(\mathbf{x}, t)$ , as well as a function of a function,  $\mathbf{k} = \mathbf{k}(\mathbf{x}, t)$ , which makes  $\varpi$  also an implicit function of  $(\mathbf{x}, t)$ . This dependence is reminiscent of the Lagrangian and Hamiltonian densities encountered in Chapter ???. Just like for the Lagrangian density, it is crucial to account for this functional dependency when computing space and time derivatives.<sup>3</sup> A useful rule to remember is that the angular frequency,  $\omega$ , is a function of space and time,  $\omega = \omega(\mathbf{x}, t) = -\partial_t \mathcal{P}(\mathbf{x}, t)$ . Yet when connected to a dispersion relation,  $\varpi(\mathbf{x}, t, \mathbf{k}(\mathbf{x}, t))$ , which links a wavevector to an angular frequency, then we must treat the dispersion relation as an explicit function of  $\mathbf{x}, t$ , as well as an implicit function through the wavevector dependence,  $\mathbf{k} = \mathbf{k}(\mathbf{x}, t)$ .

Recall from Section 2.2 that there is a relation between the wave phase,  $\mathcal{P}$ , the local angular frequency,  $\omega = -\partial_t \mathcal{P}$ , and the local wavevector,  $\mathbf{k} = \nabla \mathcal{P}$ . Inserting these identities into the local dispersion relation (2.21) leads to the *eikonal equation*, which is a nonlinear partial differential equation for the phase<sup>4</sup>

$$\partial_t \mathcal{P} + \varpi(\mathbf{x}, t, \nabla \mathcal{P}) = 0. \quad (2.22)$$

---

<sup>3</sup>In Section ?? we provide a detailed discussion of these derivatives, with that discussion suited to the present discussion of geometric optics.

<sup>4</sup>As per page 362 of [Thorne and Blandford \(2017\)](#).

This equation is formally the same as the *Hamilton-Jacobi* equation of classical mechanics, yet with  $\varpi$  playing the role of the Hamiltonian (e.g., [Marion and Thornton \(1988\)](#) or [Goldstein \(1980\)](#)). That analogy offers suggestions for how to make use of this equation. Note that on [Olbers et al. \(2012\)](#) (see their page 168) write the eikonal equation in the alternative form

$$(\partial_t \mathcal{P})^2 = C_p^2 (\nabla \mathcal{P})^2, \quad (2.23)$$

which follows from the definition of the local phase speed in equation (2.19).

### 2.3.2 Rays are integral curves of the group velocity

A *ray* is an integral curve of the group velocity. Hence, the ray trajectory,  $\mathbf{X}(t)$ , follows a ray and is determined by solving the ordinary differential equation

$$\frac{D_r \mathbf{X}}{Dt} = \mathbf{c}_g. \quad (2.24)$$

The time derivative,  $D_r/Dt$ , determines time changes when following a ray, and so it is defined by this equation. Indeed, our specification of  $D_r/Dt$  and  $\mathbf{c}_g$  correspond precisely to how we defined the trajectory of a fluid particle as the integral curve of the fluid velocity as per equation (??). Introducing the dispersion relation leads to

$$\frac{D_r \mathbf{X}}{Dt} = \mathbf{c}_g = \nabla_{\mathbf{k}} \varpi(\mathbf{x}, t, \mathbf{k}). \quad (2.25)$$

Writing this equation in component form

$$\frac{D_r X^a}{Dt} = c_g^a = \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial k_a} \right]_{\mathbf{x}, t} \quad (2.26)$$

emphasizes that the derivative is computed while holding the space and time point fixed while varying the wavevector.

### 2.3.3 Evolution of $\mathbf{k}$ along a ray

To determine the evolution of the wavevector and angular frequency along a ray, we start with the following identity that holds for the phase function,  $\mathcal{P} = \mathcal{P}(\mathbf{x}, t)$ , merely since the partial derivatives commute

$$\nabla(\partial_t \mathcal{P}) = \partial_t(\nabla \mathcal{P}). \quad (2.27)$$

The derivatives in this equation are taken with their complementary variables fixed, and this point is important when inserting the dispersion relation. Indeed, before working through the following manipulations it can be useful to reread Section 2.3.1 to be reminded of the various functional relationships.

Exposing subscripts for clarity, we start from equation (2.27) to write

$$\frac{\partial}{\partial t} \Big|_{\mathbf{x}} \left[ \frac{\partial \mathcal{P}}{\partial x^a} \right]_t = \left[ \frac{\partial k_a}{\partial t} \right]_{\mathbf{x}} \quad \text{equation (2.2)} \quad (2.28a)$$

$$= - \left[ \frac{\partial \omega(\mathbf{x}, t)}{\partial x^a} \right]_t \quad \text{equation (2.27)} \quad (2.28b)$$

$$= - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial x^a} \right]_t \quad \text{equation (2.21)} \quad (2.28c)$$

$$= - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial k_b} \right]_{\mathbf{x}, t} \left[ \frac{\partial k_b}{\partial x^a} \right]_t - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial x^a} \right]_{\mathbf{k}, t} \quad \text{chain rule} \quad (2.28d)$$

$$= -c_g^b \left[ \frac{\partial k_b}{\partial x^a} \right]_t - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial x^a} \right]_{\mathbf{k}, t} \quad \text{equation (2.25).} \quad (2.28e)$$

We can massage the first right hand side term in equation (2.28e) by commuting partial derivatives

$$\frac{\partial k_b}{\partial x^a} = \frac{\partial^2 \mathcal{P}}{\partial x^b \partial x^a} = \frac{\partial k_a}{\partial x^b}, \quad (2.29)$$

so that

$$\left[ \frac{\partial k_a}{\partial t} \right]_{\mathbf{x}} = -(\mathbf{c}_g \cdot \nabla) k_a - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial x^a} \right]_{\mathbf{k}, t}. \quad (2.30)$$

Introducing the time derivative following a ray renders the wavevector evolution equation

$$\frac{D_r k_a}{Dt} = - \left[ \frac{\partial \varpi(\mathbf{x}, t, \mathbf{k})}{\partial x^a} \right]_{\mathbf{k}, t}. \quad (2.31)$$

Evidently, if the dispersion relation has no explicit dependence on  $x^a$ , then the corresponding component of the wavevector remains constant when following a ray. However, if there is a dependence, then the wavevector evolves along the ray, with this evolution known as *refraction*.

### 2.3.4 Evolution of $\omega$ along a ray

Being a bit more streamlined than in Section 2.3.3, we compute the time derivative of the angular frequency

$$-\frac{\partial^2 \mathcal{P}}{\partial t^2} = \left[ \frac{\partial \omega}{\partial t} \right]_{\mathbf{x}} \quad (2.32a)$$

$$= \left[ \frac{\partial \varpi}{\partial t} \right]_{\mathbf{x}} \quad (2.32b)$$

$$= \left[ \frac{\partial \varpi}{\partial k_b} \right]_{\mathbf{x}, t} \left[ \frac{\partial k_b}{\partial t} \right]_{\mathbf{x}} + \left[ \frac{\partial \varpi}{\partial t} \right]_{\mathbf{x}, \mathbf{k}} \quad (2.32c)$$

$$= -c_g^b \left[ \frac{\partial \omega}{\partial x^b} \right]_t + \left[ \frac{\partial \varpi}{\partial t} \right]_{\mathbf{x}, \mathbf{k}}, \quad (2.32d)$$

where the final step used

$$\left[ \frac{\partial k_b}{\partial t} \right]_{\mathbf{x}} = \frac{\partial^2 \mathcal{P}}{\partial t \partial x^b} = \frac{\partial}{\partial x^b} \frac{\partial \mathcal{P}}{\partial t} = - \left[ \frac{\partial \omega}{\partial x^b} \right]_t. \quad (2.33)$$

We are thus lead to the evolution equation for the angular frequency along a ray

$$\frac{D_r \omega}{Dt} = \left[ \frac{\partial \varpi}{\partial t} \right]_{\mathbf{x}, \mathbf{k}}. \quad (2.34)$$

Evidently, if the dispersion relation has no explicit dependence on time, then the angular frequency is a constant following a ray. For example, this is the situation when gravity waves approach a beach, assuming the slope of the beach is static (considered in Section 4.7.1). For this example, the angular frequency of gravity waves remains constant following a ray, whereas

the wavevector changes according to equation (2.31).

### 2.3.5 Changes in the phase following a ray

Considering the phase function to be a function of space and time,  $\mathcal{P}(\mathbf{x}, t)$ , we compute its time derivative along a ray according to

$$\mathrm{D}_r \mathcal{P} / \mathrm{D}t = (\partial_t + \mathbf{c}_g \cdot \nabla) \mathcal{P} = -\omega + \mathbf{c}_g \cdot \mathbf{k} = \mathbf{k} \cdot (-\mathbf{c}_p + \mathbf{c}_g), \quad (2.35)$$

where  $\omega = -\partial_t \mathcal{P} = \mathbf{c}_p \cdot \mathbf{k}$ , as per equations (2.17) and (2.18), and  $\mathbf{k} = \nabla_{\mathbf{x}} \mathcal{P}$  as per equation (2.2). Non-dispersive have a phase that remains unchanged when following a ray, which holds so long as

$$\mathbf{k} \cdot \mathbf{c}_p = \mathbf{k} \cdot \mathbf{c}_g \quad \text{non-dispersive waves.} \quad (2.36)$$

Otherwise, for dispersive waves the phase changes according to the source term,  $\mathbf{k} \cdot (-\mathbf{c}_p + \mathbf{c}_g)$ .

### 2.3.6 Summary of the ray equations

We here summarize the evolution equations for wave properties derived thus far in this chapter, with these equations constituting *Hamilton's equations for rays*

$$\frac{\mathrm{D}_r \mathbf{X}}{\mathrm{D}t} = \mathbf{c}_g \quad \text{trajectory on a ray} \quad (2.37a)$$

$$\frac{\mathrm{D}_r \mathcal{P}}{\mathrm{D}t} = \mathbf{k} \cdot (-\mathbf{c}_p + \mathbf{c}_g) \iff \partial_t \mathcal{P} + \varpi(\mathbf{x}, t, \nabla \mathcal{P}) = 0 \quad \text{eikonal equation for phase} \quad (2.37b)$$

$$\frac{\mathrm{D}_r \mathbf{k}}{\mathrm{D}t} = - \left[ \frac{\partial \varpi}{\partial \mathbf{x}} \right]_{t, \mathbf{k}} \quad \mathbf{k} \text{ evolution on a ray} \quad (2.37c)$$

$$\frac{\mathrm{D}_r \omega}{\mathrm{D}t} = \left[ \frac{\partial \varpi}{\partial t} \right]_{\mathbf{x}, \mathbf{k}} \quad \omega \text{ evolution on a ray} \quad (2.37d)$$

$$\frac{\mathrm{D}_r}{\mathrm{D}t} = \partial_t + \mathbf{c}_g \cdot \nabla \quad \text{time derivative on a ray} \quad (2.37e)$$

$$\omega = -\partial_t \mathcal{P} \quad \omega \text{ defined} \quad (2.37f)$$

$$\mathbf{k} = \nabla \mathcal{P} \quad \mathbf{k} \text{ defined} \quad (2.37g)$$

$$\nabla \times \mathbf{k} = 0 \quad \mathbf{k} \text{ consistency condition} \quad (2.37h)$$

$$\partial_t \mathbf{k} + \nabla \omega = 0 \quad \text{conservation of wave crests.} \quad (2.37i)$$

Equation (2.37a) defines the rays as integral lines of the group velocity,  $\mathbf{c}_g$ . Equation (2.37b) is the eikonal equation that connects the local angular frequency to the dispersion relation

$$\omega = -\partial_t \mathcal{P} = \varpi(\mathbf{x}, t, \nabla \mathcal{P}). \quad (2.38)$$

Each of the evolution equations (2.37b)–(2.37d) is hyperbolic, which is notable since this property holds even though the wave equation describing the evolution of the wave function is not generally hyperbolic (particularly for dispersive waves).

By definition, the phase remains constant when following along a constant wave phase so that

$$(\partial_t + \mathbf{c}_p \cdot \nabla) \mathcal{P} = 0. \quad (2.39)$$

However, for dispersive waves, equation (2.37b) indicates that the phase does not stay fixed when following rays as defined by the group velocity. Hence, wave crests pass through a point

following the group velocity, such as a point fixed within a wave packet. Conversely, equations (2.37c) and (2.37d) show that for dispersive waves, the wavevector and angular frequency are not constant when following a constant wave phase. Instead, they remain fixed when following the group velocity in media where the dispersion relation is independent of space and time. For more general media, both the wavevector and angular frequency evolve even when following the group velocity.

### 2.3.7 Comments and further study

Knowledge of the dispersion relation is sufficient to solve the ray equations (typically using numerical methods), thus mapping the rays and paths of wave packets, and determining the wavevector, angular frequency, and phase following a ray. This procedure works quite well to describe waves moving through smoothly varying media, and is familiar from the rays of light bending through water or glass. Failure of the method occurs when the background media no longer satisfies the “gently varying” assumptions that are formalized in Section 2.5.3 for the case of a stretched string. Breaking these assumptions often results in the intersection of rays which, in many fluid applications, signals a nonlinear process such as a fluid instability and corresponding turbulent mixing. Chapter 11 of [Whitham \(1974\)](#), Chapters 3 and 4 of [Bühler \(2014\)](#), and the bulk of [Tracy et al. \(2014\)](#) provide more details and insights.

## 2.4 Hamilton's principle and the Euler-Lagrange equation

The geometric optics discussion in Section 2.3 is largely kinematic, with the only piece of dynamical information arising from the dispersion relation. To develop a theory for the energetics of waves and wave packets moving through gently varying background states, we make use of Hamilton's principle, or more specifically Whitham's variational principle applying to the phase averaged equations. We pursue that study in Section 2.5, yet here first summarize salient points from Lagrangian field theory from Chapter ??.

### 2.4.1 Stationary action $\iff$ Euler-Lagrange equation

The *action*,  $\mathcal{S}$ , for a continuous scalar field,  $\psi(\mathbf{x}, t)$ , is given by equation (??)

$$\mathcal{S} = \int_R \mathcal{L}(\psi, \partial_t \psi, \nabla \psi, \mathbf{x}, t) g d^3x dt, \quad (2.40)$$

where  $R$  is a space-time domain,  $g$  is the square root of the metric tensor determinant, and  $\mathcal{L}$  is the *Lagrangian density*. The Lagrangian density is a functional of the field and its space-time derivatives (i.e.,  $\mathcal{L}$  is a function of a function). There can also be additional explicit dependencies on the space and time position, which occurs, in particular, for waves moving through a media that is inhomogeneous in space and/or non-stationary in time, in which case  $\mathcal{L}$  is an explicit function of  $(\mathbf{x}, t)$ . *Hamilton's principle* as stated by equation (??) says that variation of the action is stationary (i.e., variation of the action vanishes) for the physically realized field. Hamilton's principle then leads to the *Euler-Lagrange field equation* (??)

$$\frac{\partial \mathcal{L}}{\partial \psi} - \frac{1}{g} \frac{\partial}{\partial x^\alpha} \left[ g \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \psi)} \right] = 0, \quad (2.41)$$

where  $\alpha = 0, 1, 2, 3$  is the space-time label. In performing the partial derivative with respect to  $\psi$  and its derivatives,  $\partial_\alpha \psi$ , each of the other variables in the Lagrangian density are held fixed. However, when performing the space and time partial derivatives,  $\partial_\alpha$ , we only maintain the complement space and time variable fixed. This technical point is very important when taking derivatives of the Lagrangian density (see Section ??), and it is directly analogous to how we take derivatives of the dispersion relation in Section 2.3.

### 2.4.2 Hamiltonian density and energy of the continuum

In Section ?? we introduced the *generalized momentum density*,  $\mathcal{P}$ , and the *Hamiltonian density*,

$$\mathcal{P} = \frac{\partial \mathcal{L}}{\partial(\partial_t \psi)} \quad \text{and} \quad \mathcal{H} = \mathcal{P} \partial_t \psi - \mathcal{L}. \quad (2.42)$$

We then showed that the Hamiltonian density satisfies the budget equation

$$\partial_t \mathcal{H} + \nabla \cdot \mathcal{F} = -(\partial \mathcal{L} / \partial t)_{\psi, \partial_\alpha \psi, \mathbf{x}}, \quad (2.43)$$

where the time derivative on the right hand side is computed while fixing  $\psi$ , its space and time derivatives, and the space position,  $\mathbf{x}$ . Equation (2.43) also introduced the flux vector,  $\mathcal{F}$ , with components

$$\mathcal{F}^a = \frac{\partial \mathcal{L}}{\partial(\partial_a \psi)} \frac{\partial \psi}{\partial t}. \quad (2.44)$$

In those cases where the Lagrangian has no explicit time dependence, so that it has the functional dependence

$$\mathcal{L} = \mathcal{L}(\psi, \partial_t \psi, \nabla \psi, \mathbf{x}), \quad (2.45)$$

then the Hamiltonian equation (2.43) becomes a conservation equation for energy that follows from time symmetry and Noether's theorem

$$\partial_t \mathcal{H} + \nabla \cdot \mathcal{F} = 0. \quad (2.46)$$

### 2.4.3 Stress-energy-momentum tensor

As detailed in Section ??, the energy equation (2.43) is but one piece of the equation satisfied by the stress-energy-momentum tensor,  $T^\alpha_\beta$ . This equation is given by equation ??

$$\partial_\alpha T^\alpha_\beta = - \left[ \frac{\partial \mathcal{L}}{\partial x^\beta} \right]_{\psi, \partial_\alpha \psi, x^\alpha \neq \beta}. \quad (2.47)$$

where the *stress-energy-momentum tensor* is

$$T^\alpha_\beta = -\delta^\alpha_\beta \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \alpha \psi} \frac{\partial \psi}{\partial x^\beta}. \quad (2.48)$$

## 2.5 Whitham's variational principle

The Euler-Lagrange field equation (2.41) and associated energy equation (2.43) provide the foundation for the field theory of continuous classical matter. In this section we specialize the field theory formalism for the purpose of describing waves and the movement of wave energy through a prescribed background environment. Much of this formalism was proposed

by Whitham (e.g., see Chapter 11 of [Whitham \(1974\)](#)), so that we refer to the method as *Whitham's variational principle*. In other treatments, such as Section 2.2 of [Tracy et al. \(2014\)](#), it is referred to as the *reduced variational principle*.

We anticipated much of the material in this section when studying a pendulum with variable length in Section ?? . In that section we showed that the pendulum energy divided by its frequency is an *adiabatic invariant*; i.e., it is nearly constant. In the present section we show that the *wave action* is the corresponding adiabatic invariant for waves moving on a gently varying background, with the wave action also equal to the energy (or Hamiltonian) divided by the angular frequency.

### 2.5.1 Some motivation

One practical motivation for developing Whitham's variational principle is to describe the energetics of waves propagating on a non-homogeneous and non-stationary background, though the formalism is also quite useful for homogeneous and stationary backgrounds, particularly for dispersive waves. In the presence of a space dependent background properties, there are no simple sinusoidal (plane) wave solutions. That is, Fourier methods are insufficient. Even so, there are waves that are quite close to plane waves *if* the space and time scales for changes in the background state are well separated from those describing the wave properties (frequency and wavelength). This space-time scale separation constitutes a “gently varying” background state, which in turn enables analytical progress. We introduced the scaling for this gentle background at the start of Section 2.2, and follow up in Section 2.5.3 with more details.

Whitham's method starts by inserting a wave ansatz into the Lagrangian density. Phase averaging then produces the *phase averaged Lagrangian* and corresponding *phase averaged action*. We maintain only the leading order term in this phase average Lagrangian, and Whitham's variational principle then states that the physically realized wave produces a stationary wave averaged action. There are two associated Euler-Lagrange fields, one arising from varying the wave amplitude and the other from varying the wave phase.

### 2.5.2 The stretched string

Following [Brettherton and Garrett \(1969\)](#) and [Brettherton \(1971\)](#), we use a stretched string as a case study for exploring Whitham variational method. In particular, recall from Section ?? the Lagrangian density for the stretched string whose tension<sup>5</sup> is a function of time,  $\tau = \tau(t)$ , and whose mass density (mass per length) is a function of space,  $\sigma = \sigma(x)$ , is given by

$$\mathcal{L} = (1/2)[\sigma(\partial_t \psi)^2 - \tau(\partial_x \psi)^2], \quad (2.49)$$

and whose Hamiltonian density (2.42) is

$$\mathcal{H} = (1/2)[\sigma(\partial_t \psi)^2 + \tau(\partial_x \psi)^2]. \quad (2.50)$$

Note that  $\psi(x, t)$  measures transverse displacements of the string from its equilibrium position at  $\psi = 0$ . We furthermore assume that there are no longitudinal waves along the string, thus focusing exclusively on transverse motion.

---

<sup>5</sup>The dimensions of  $\tau$  are force,  $M L T^{-2}$ , but we refer to it as a tension since it is an internal force within the string.

The action for the stretched string is

$$\mathcal{S} = \int \mathcal{L} dx dt = \int (1/2) [\sigma (\partial_t \psi)^2 - \tau (\partial_x \psi)^2] dx dt, \quad (2.51)$$

and the Euler-Lagrange equation resulting from Hamilton's principle is the wave equation

$$\delta \mathcal{S} = \int (\delta \mathcal{L}) dx dt = 0 \iff (\partial_{tt} - c^2 \partial_{xx}) \psi = 0, \quad (2.52)$$

with the squared wave speed

$$c^2(x, t) = \tau(t)/\sigma(x). \quad (2.53)$$

Because the wave speed is a function of space and time, we cannot use traditional Fourier methods to find a wave solution. Hence, the mathematical goal of this section is to develop methods for use when there are space and time variations of the background state, in particular when such variations are “gentle”. In this case we are afforded the *eikonal wave ansatz*<sup>6</sup>

$$\psi(x, t) = A(x, t) \cos[\mathcal{P}(x, t)], \quad (2.54)$$

where  $A > 0$  is a space-time dependent real amplitude function, and  $\mathcal{P}$  is the phase function introduced in equation (2.2). We have more comments on this form for the wave ansatz in Section 2.5.4.

### 2.5.3 Space and time scale separation

To make analytical progress requires us to detail the space and time scale separation between the linear waves and the background state, following from our introduction to this scaling at the start of Section 2.2. Although we here focus on the string, the same sorts of assumptions must be realized for other waves systems in order to make use of the methods of this section. For the time scale, we assume that changes in the string tension occur over time scales that are much longer than the period of the linear waves supported by the string, so that

$$|\partial_t \tau / \tau| \ll |\partial_t \mathcal{P}| / 2\pi = \omega / 2\pi. \quad (2.55)$$

Correspondingly, temporal changes in the angular frequency are assumed to be on the same scale as temporal changes to the string tension

$$|\partial_t \omega / \omega| \sim |\partial_t \tau / \tau| \implies |\partial_t \omega| \ll \omega^2 / 2\pi, \quad \text{or equivalently} \quad |\partial_{tt} \mathcal{P}| \ll (\partial_t \mathcal{P})^2 / 2\pi. \quad (2.56)$$

Likewise, time changes to the wave amplitude are defined by the background state so that

$$|\partial_t A / A| \sim |\partial_t \tau / \tau|. \quad (2.57)$$

For the length scale, we assume spatial changes in the mass density occur over scales that are large compared to the wavelength

$$|\partial_x \sigma / \sigma| \ll |\partial_x \mathcal{P}| / 2\pi = |k| / 2\pi, \quad (2.58)$$

---

<sup>6</sup>The ansatz (2.54) is commonly used with the WKB asymptotic method, and so is sometimes referred to as the WKB wave ansatz.

and that spatial changes to the wavenumber and wave amplitude are on the same scale as changes to string mass density

$$|\partial_x k/k| \sim |\partial_x \sigma/\sigma| \implies |\partial_x k| \ll k^2/2\pi \quad \text{or equivalently} \quad |\partial_{xx} \mathcal{P}| \ll (\partial_x \mathcal{P})^2/2\pi. \quad (2.59)$$

Likewise, space changes to the wave amplitude are defined by the background state so that

$$|\partial_x A/A| \sim |\partial_x \sigma/\sigma|. \quad (2.60)$$

#### 2.5.4 The leading order phase averaged action

The space and time scale separation from Section 2.5.3 motivate the wave ansatz (2.54). Note that a more general ansatz might also consider a space-time dependent phase shift. We do not consider a phase shift since we are concerned with leading order evolution of the wave amplitude, and the extra phase degree of freedom is not directly tied to the amplitude. Furthermore, we could have expanded the amplitude function into an asymptotic series. But again, we are only interested in the leading order, with the ansatz (2.54) sufficient for that purpose.<sup>7</sup>

##### Leading order contribution to $\sigma (\partial_t \psi)^2$

To determine the leading order terms contributing to the Lagrangian (2.49), we start with the squared time tendency of the wave function

$$(\partial_t \psi)^2 = (\partial_t A \cos \mathcal{P})^2 + (A \partial_t \mathcal{P} \sin \mathcal{P})^2 - 2(A \partial_t \mathcal{P} \partial_t A \sin \mathcal{P} \cos \mathcal{P}), \quad (2.61)$$

which has a phase average

$$\langle \sigma (\partial_t \psi)^2 \rangle = \sigma [(\partial_t A)^2 + (A \partial_t \mathcal{P})^2]/2. \quad (2.62)$$

To compute the phase average we assumed that  $\sigma$ ,  $\partial_t A$ , and  $\partial_t \mathcal{P} = -\omega$  are roughly constant over the course of a  $2\pi$  change in the phase, with this assumption following from the scaling in Section 2.5.3. Further use of the scaling in Section 2.5.3 leads to the leading order contribution

$$\langle \sigma (\partial_t \psi)^2 \rangle \approx \sigma (A \partial_t \mathcal{P})^2/2 = \sigma A^2 \omega^2/2. \quad (2.63)$$

##### Leading order contribution to $\tau (\partial_x \psi)^2$

Proceeding just as above, the squared space derivative term in the Lagrangian (2.49) is given by

$$(\partial_x \psi)^2 = (\partial_x A \cos \mathcal{P})^2 + (A \partial_x \mathcal{P} \sin \mathcal{P})^2 - 2(A \partial_x \mathcal{P} \partial_x A \sin \mathcal{P} \cos \mathcal{P}), \quad (2.64)$$

whose phase average is

$$\langle \tau (\partial_x \psi)^2 \rangle = \tau [(\partial_x A)^2 + (A \partial_x \mathcal{P})^2]/2. \quad (2.65)$$

To compute the phase average we assumed that  $\tau$ ,  $\partial_x A$ , and  $\partial_x \mathcal{P} = k$  are roughly constant over the course of a  $2\pi$  change in the phase, with this assumption following from the scaling in Section 2.5.3. Further using the scaling in Section 2.5.3 we are led to the leading order contribution

$$\langle \tau (\partial_x \psi)^2 \rangle \approx \tau (A \partial_x \mathcal{P})^2/2 = \tau A^2 k^2/2. \quad (2.66)$$

---

<sup>7</sup>See Section 4 of [Bretherton \(1971\)](#) for more details of the phase shift and amplitude expansion.

**Leading order phase averaged action**

The above discussion provides the leading order phase averaged action

$$\langle S \rangle = \frac{1}{4} \int A^2 (\sigma \omega^2 - \tau k^2) dx dt = \int \langle \mathcal{L} \rangle dx dt. \quad (2.67)$$

This phase averaged Lagrangian is directly analogous to the phase averaged Lagrangian (??) for the pendulum whose length is slowly varying.

### 2.5.5 Whitham's variational principle for the string

Whitham's variational principle states that the phase averaged action,  $\langle S \rangle$ , is stationary when  $\psi = A \cos \mathcal{P}$  is the physically realized wave function. To find the associated Euler-Lagrange equations requires computing the variation of the action, which in turn requires varying the wave function. Variations in the wave function arise from arbitrary smooth and independent variations (with compact support) of the amplitude,  $A$ , and the phase,  $\mathcal{P}$ . Hence, the phase averaged action must be stationary with respect to independent variations in both  $A$  and  $\mathcal{P}$ . As we see in the following, the dispersion relation is the Euler-Lagrange equation resulting from  $\delta_A \langle S \rangle = 0$ , whereas *wave action* conservation results from  $\delta_{\mathcal{P}} \langle S \rangle = 0$ .

**Vanishing variation with respect to  $A$  is equivalent to the dispersion relation**

Variation of the phase averaged action (2.67) under variations in the wave amplitude is given by

$$\delta_A \langle S \rangle = \frac{1}{2} \int (\sigma \omega^2 - \tau k^2) A \delta A dx dt, \quad (2.68)$$

and with a zero variation leading to

$$\delta_A \langle S \rangle = 0 \iff \omega^2 = (\sigma/\tau)^2 k^2 = c^2 k^2. \quad (2.69)$$

Evidently, satisfying the dispersion relation is equivalent to a zero variation of the phase averaged action with respect to the wave amplitude. Furthermore, since  $\langle \mathcal{L} \rangle$  is proportional to the squared wave amplitude, then satisfying the dispersion relation (2.69) means that the phase averaged Lagrangian vanishes when evaluated with the physical solution,

$$\langle \mathcal{L} \rangle = A^2 (\sigma \omega^2 - \tau k^2)/4 = 0. \quad (2.70)$$

**Vanishing variation with respect to  $\mathcal{P}$** 

Variation of the phase averaged action (2.67) under variations in the wave phase function is given by

$$\delta_{\mathcal{P}} \langle S \rangle = \frac{1}{2} \int A^2 (\sigma \omega \delta \omega - \tau k \delta k) dx dt = \frac{1}{2} \int A^2 [\sigma \partial_t \mathcal{P} \delta(\partial_t \mathcal{P}) - \tau \partial_x \mathcal{P} \delta(\partial_x \mathcal{P})] dx dt, \quad (2.71)$$

with rearrangement giving

$$\begin{aligned} \delta_{\mathcal{P}} \langle S \rangle &= \frac{1}{2} \int A^2 [\partial_t (A^2 \sigma \partial_t \mathcal{P} \delta \mathcal{P}) - \partial_x (A^2 \tau \partial_x \mathcal{P} \delta \mathcal{P})] dx dt \\ &\quad - \frac{1}{2} \int A^2 [\partial_t (A^2 \sigma \partial_t \mathcal{P}) \delta \mathcal{P} - \partial_x (A^2 \tau \partial_x \mathcal{P}) \delta \mathcal{P}] dx dt. \end{aligned} \quad (2.72)$$

Assuming  $\delta\mathcal{P}$  vanishes on the space-time boundaries, or that it has compact support in space and time, eliminates the first integral to leave

$$\delta_{\mathcal{P}}\langle\mathcal{S}\rangle = -\frac{1}{2} \int A^2 [\partial_t(A^2 \sigma \partial_t \mathcal{P}) - \partial_x(A^2 \tau \partial_x \mathcal{P})] \delta\mathcal{P} dx dt, \quad (2.73)$$

with Whitham's principle  $\delta_{\mathcal{P}}\langle\mathcal{S}\rangle = 0$  producing the conservation law

$$\delta_{\mathcal{P}}\langle\mathcal{S}\rangle = 0 \implies \partial_t(A^2 \sigma \omega) + \partial_x(A^2 \tau k) = 0. \quad (2.74)$$

Interpretation of this conservation law follows from the discussion in the remainder of this Section.

### 2.5.6 Phase averaged Hamiltonian

Making use of the eikonal wave ansatz (2.54) renders the leading order phase averaged Hamiltonian (2.50)

$$\langle\mathcal{H}\rangle = A^2 (\sigma \omega^2 + \tau k^2)/4, \quad (2.75)$$

with the dispersion relation (2.69) yielding

$$\langle\mathcal{H}\rangle = A^2 \sigma \omega^2/2 = A^2 \tau k^2/2. \quad (2.76)$$

We can use these expressions to write the conservation law (2.74) as

$$\partial_t(\langle\mathcal{H}\rangle/\omega) + \partial_x(\langle\mathcal{H}\rangle/k) = 0. \quad (2.77)$$

This equation is nearly ready for interpretation, but it aided by the discussion in Section 2.5.7.

### 2.5.7 A general statement of Whitham's variational principle

To help interpret the conservation law (2.77), it is useful to provide a generic expression of Whitham's variational principle, which we write as

$$\delta \int \langle\mathcal{L}\rangle(A, \omega, k; \sigma, \tau) dx dt = \delta \int \langle\mathcal{L}\rangle(A, -\partial_t \mathcal{P}, \partial_x \mathcal{P}; \sigma, \tau) dx dt = \int \delta \langle\mathcal{L}\rangle dx dt = 0. \quad (2.78)$$

The resulting Euler-Lagrange equations follow through variations of the wave amplitude,  $A$ , and phase function,  $\mathcal{P}$ .

#### Varying the wave amplitude and the dispersion relation

The Euler-Lagrange equation resulting from varying the amplitude is written (see equation (2.41))

$$\frac{\delta \mathcal{S}}{\delta A} = \frac{\partial \langle\mathcal{L}\rangle}{\partial A} - \frac{\partial}{\partial t} \left[ \frac{\partial \langle\mathcal{L}\rangle}{\partial (\partial_t A)} \right] - \frac{\partial}{\partial x} \left[ \frac{\partial \langle\mathcal{L}\rangle}{\partial (\partial_x A)} \right] = 0. \quad (2.79)$$

Notably, for linear waves the wave amplitude appears in  $\langle\mathcal{L}\rangle$  only via its square,  $A^2$ , so that the second and third terms in equation (2.79) vanish identically. As a result, the Euler-Lagrange equation resulting from varying  $A$  is

$$\frac{\partial \langle\mathcal{L}\rangle}{\partial A} = 0, \quad (2.80)$$

which, as we saw for the stretched string in Section 2.5.5, is equivalent to the dispersion relation connecting the wave angular frequency to the wavevector. Another way to write the dispersion relation is to note that the phase averaged Lagrangian for linear waves satisfies

$$A \partial \langle \mathcal{L} \rangle / \partial A = 2 \langle \mathcal{L} \rangle, \quad (2.81)$$

so that

$$\langle \mathcal{L} \rangle = 0 \iff \text{dispersion relation.} \quad (2.82)$$

### Varying the phase function, group velocity, and conservation of wave action

For linear waves, the phase function only appears in terms of its space and time derivatives, so that  $\partial \langle \mathcal{L} \rangle / \partial \mathcal{P} = 0$ , in which case the Euler-Lagrange equation resulting from varying the phase function is given by

$$\frac{\delta \mathcal{S}}{\delta \mathcal{P}} = \frac{\partial}{\partial t} \left[ \frac{\partial \langle \mathcal{L} \rangle}{\partial (\partial_t \mathcal{P})} \right] + \frac{\partial}{\partial x} \left[ \frac{\partial \langle \mathcal{L} \rangle}{\partial (\partial_x \mathcal{P})} \right] = 0 \implies -\frac{\partial}{\partial t} \left[ \frac{\partial \langle \mathcal{L} \rangle}{\partial \omega} \right] + \frac{\partial}{\partial x} \left[ \frac{\partial \langle \mathcal{L} \rangle}{\partial k} \right] = 0. \quad (2.83)$$

We can write this conservation equation in a slightly different form by introducing the group velocity. For this purpose, consider the implications for a zero variation of the phase averaged Lagrangian (another way to state Whitham's principle as in equation (2.78)) so that

$$0 = \delta \langle \mathcal{L} \rangle = \frac{\partial \langle \mathcal{L} \rangle}{\partial A} \delta A + \frac{\partial \langle \mathcal{L} \rangle}{\partial \omega} \delta \omega + \frac{\partial \langle \mathcal{L} \rangle}{\partial k} \delta k = \frac{\partial \langle \mathcal{L} \rangle}{\partial \omega} \delta \omega + \frac{\partial \langle \mathcal{L} \rangle}{\partial k} \delta k, \quad (2.84)$$

where  $\partial \langle \mathcal{L} \rangle / \partial A = 0$  follows from equation (2.80). Rearrangement allows us to write the group velocity in terms of derivatives of the phase averaged Lagrangian

$$c_g = \delta \omega / \delta k = -\frac{\partial \langle \mathcal{L} \rangle / \partial k}{\partial \langle \mathcal{L} \rangle / \partial \omega}, \quad (2.85)$$

so that the Euler-Lagrange equation (2.83) can be written in the form of a traditional conservation law

$$\partial_t \mathcal{A} + \partial_x (c_g \mathcal{A}) = 0, \quad (2.86)$$

where we defined the *wave action*

$$\mathcal{A} \equiv \frac{\partial \langle \mathcal{L} \rangle}{\partial \omega}. \quad (2.87)$$

Note that for some applications it is more useful to write the wave action equation (2.86) following the group velocity, so that

$$\frac{D \mathcal{A}}{Dt} = -\mathcal{A} \partial_x c_g. \quad (2.88)$$

Hence, the wave action evolves following a ray when moving through regions where the group velocity has a nonzero convergence.

### 2.5.8 Interpreting the string's wave action conservation equation

For the string, we find

$$\partial \langle \mathcal{L} \rangle / \partial \omega = A^2 \sigma \omega / 2 = \langle \mathcal{H} \rangle / \omega = \mathcal{A}, \quad (2.89)$$

so that the wave action conservation law (2.87) can be written in terms of the Hamiltonian

$$\partial_t(\langle \mathcal{H} \rangle / \omega) + \partial_x(c_g \langle \mathcal{H} \rangle / \omega) = 0, \quad (2.90)$$

which accords with equation (2.77) given that the string is non-dispersive so that the group velocity is  $c_g = \omega/k$ .

So in summary, the Euler-Lagrange equation arising from  $\delta_{\mathcal{P}}\langle \mathcal{S} \rangle = 0$  leads to the wave action conservation law (2.90). This conservation law is associated with a symmetry of the phase averaged Lagrangian; namely, it has no explicit dependence on the phase so that  $\partial\langle \mathcal{L} \rangle / \partial \mathcal{P} = 0$ . Whereas mechanical energy satisfies a conservation law when the background state is time-independent (i.e., the Lagrangian has no explicit time dependence), the wave action satisfies a conservation law since the phase averaged Lagrangian has no explicit dependence on the phase.

### 2.5.9 Comments

Although the presentation in this section used the relatively simple case of a stretched string, the underlying theory holds for all linear waves. As such, we expand our understanding of the theory when considering waves of more complexity in the following chapters. In particular, the theory proves quite useful for organizing our thinking about group velocity and wave energetics, particularly for dispersive waves whether on a stationary and homogeneous background or more generally.

## 2.6 Variational methods for self-adjoint wave operators

Our use of variational methods for waves is concerned with the propagation of wave packets through a gently varying and prescribed background state. We assume the packet is prepared at some initial time and fully known at the final time, so that variations of the wave function (as per Hamilton's principle) are nonzero only at intermediate times. Likewise, we assume the packet is specified at the spatial boundaries, so that its variation vanishes there as well. With these assumptions, we here detail a variational method that makes use of the self-adjoint nature of the wave operator. By exploiting the self-adjoint property, we can reformulate Whitham's variational principle from Section 2.5.

### 2.6.1 Self-adjoint linear wave operators

Assuming Cartesian coordinates, the Euler-Lagrange equations for the acoustic and Klein-Gordon wave equations from Section ?? can be written

$$\frac{\delta \mathcal{S}}{\delta \psi} = \frac{\partial \mathcal{L}}{\partial \psi} - \frac{\partial}{\partial x^\alpha} \left[ \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \psi)} \right] = \hat{D}\psi = 0, \quad (2.91)$$

where  $\hat{D}$  is the linear partial differential operator,

$$\hat{D} = -(\partial_{tt} - c^2 \nabla^2 + \Gamma^2), \quad (2.92)$$

for the Klein-Gordon wave equation with constant  $c$ , and with  $\Gamma = 0$  for the acoustic wave equation. This result allows us to write the action for this, and other, linear non-dissipative

waves in the bilinear form

$$\mathcal{S} = \frac{1}{2} \int_{\mathbb{R}} \psi (\hat{D}\psi) d^3x dt. \quad (2.93)$$

Notably, this action vanishes identically when  $\psi$  is a solution to the wave equation, as per equation (2.91), yet the action is non-zero for a general function.

### 2.6.2 Varying the action

Variation of the action (2.93) can be written

$$2\delta\mathcal{S} = \int_{\mathbb{R}} (\delta\psi) \hat{D}\psi d^3x dt + \int_{\mathbb{R}} \psi (\hat{D}\delta\psi) d^3x dt, \quad (2.94)$$

with integration by parts bringing the second right hand side term into the form

$$\begin{aligned} \int_{\mathbb{R}} \psi (\hat{D}\delta\psi) d^3x dt &= \int_{\mathbb{R}} \delta\psi (\hat{D}\psi) d^3x dt \\ &\quad - \int_{\mathbb{R}} \partial_t(\psi \partial_t \delta\psi - \delta\psi \partial_t \psi) d^3x dt + c^2 \int_{\mathbb{R}} \delta^{ab} \partial_b(\psi \partial_a \delta\psi - \delta\psi \partial_a \psi) d^3x dt. \end{aligned} \quad (2.95)$$

For Hamilton's principle, we assume that  $\delta\psi = 0$  at the temporal endpoints, so that

$$\int_{\mathbb{R}} \partial_t(\delta\psi \partial_t \psi) d^3x dt = 0. \quad (2.96)$$

Likewise, the term

$$\int_{\mathbb{R}} \delta^{ab} \partial_b(\delta\psi \partial_a \psi) d^3x dt = \int_{\mathbb{R}} \nabla \cdot (\delta\psi \nabla \psi) d^3x dt, \quad (2.97)$$

vanishes if  $\delta\psi = 0$  on the spatial boundaries, or if the physically realized field satisfies the natural boundary conditions

$$\hat{\mathbf{n}} \cdot \nabla \psi = 0 \quad \text{for } \mathbf{x} \in \partial\mathcal{R}. \quad (2.98)$$

We are thus led to the variation

$$\int_{\mathbb{R}} \psi (\hat{D}\delta\psi) d^3x dt = \int_{\mathbb{R}} \delta\psi (\hat{D}\psi) d^3x dt + \int_{\mathbb{R}} [-\partial_t(\psi \partial_t \delta\psi) + c^2 \delta^{ab} \partial_b(\psi \partial_a \delta\psi)] d^3x dt. \quad (2.99)$$

We can eliminate the remaining boundary terms as follows. The temporal boundary term vanishes if we assume either  $\psi$  or  $\delta(\partial_t \psi)$  vanishes at temporal boundaries. We assume either of these conditions holds. Likewise, the spatial boundary term vanishes if  $\psi$  or  $\delta(\partial_a \psi)$  vanishes at the spatial boundaries.<sup>8</sup> Assuming the boundary conditions indeed vanish makes  $\hat{D}$  a self-adjoint wave operator, thus bringing about the action variation

$$\delta\mathcal{S} = \int_{\mathbb{R}} (\delta\psi) \hat{D}\psi d^3x dt, \quad (2.100)$$

with Hamilton's principle,  $\delta\mathcal{S} = 0$ , again leading to the wave equation,  $\hat{D}\psi = 0$ .

---

<sup>8</sup>In considering these boundary conditions when, recall that  $\delta$  for Hamilton's principle commutes with space and time derivatives, as detailed in Section ??.

### 2.6.3 Comments on the method

The above presentation is a bit circular since we compute the wave operator,  $\hat{D}$ , from the Euler-Lagrange equation, and the Euler-Lagrange equation requires the Lagrangian density,  $\mathcal{L}$ . However, the discussion suggests a complementary means to connect the variational framework of Hamilton's principle to the study of linear waves. Namely, all we need is the linear self-adjoint wave operator,  $\hat{D}$ , rather than the Lagrangian density,  $\mathcal{L}$ . This approach is useful particularly for those cases where the Lagrangian is tricky to determine. For example, the Lagrangian for Rossby waves requires the introduction of auxiliary fields (see page 293 of [Olbers et al. \(2012\)](#)). So by writing the action in the form of equation (2.93), we directly connect the variational framework to the study of linear waves via knowledge of the wave operator rather than the wave Lagrangian. We offer examples throughout the chapters in this part of the book. One important caveat is that the action for some wave systems cannot be written in the bilinear form (2.93), such as the interface waves from Chapter 4, in which case distinct methods are needed (see Section 4.2.9).

We also note the need to assume boundary conditions that ensure the wave operator is self-adjoint. These boundary conditions are somewhat restrictive. Even so, they allow us to focus on the form of the waves in the interior of the space-time domain, rather than be concerned with initial conditions.



## Chapter 3

# ACOUSTIC WAVES

In this chapter we study the physics of *acoustic waves* in fluids. When reaching the human ear we interpret acoustic waves as sound, hence the synonymous term *sound waves*. The pressure fluctuations associated with standard acoustic waves (those not damaging to the human ear) are a tiny fraction of those arising from, say, a weather disturbance or an atmospheric gravity wave (see Section 3.5.3 for details). Hence, acoustic waves play a negligible role in the geophysical fluid flows forming the focus of this book. Nonetheless, the relative simplicity of acoustic waves, and their ubiquitous presence in the natural environment, make them an ideal pedagogical introduction to the somewhat more complex geophysical waves considered in later chapters. Here, we study acoustic waves in a perfect compressible fluid where the only force arises from pressure, thus ignoring gravity, Coriolis, and viscous friction. We derive the acoustic wave equation using both Lagrangian and Eulerian viewpoints, and thereafter study various properties of acoustic waves, including their energetics.

### READER'S GUIDE TO THIS CHAPTER

Development of the linear equations for acoustic waves relies on the momentum dynamics from Chapter ??, and the acoustic energetics makes use of thermodynamics from Chapters ??, ??, and ?. Furthermore, we assume an understanding of wave kinematics from Chapter 1. As a useful conceptual preface, see Sections ?? and ?? where we study simple harmonic oscillators, as well as Section ?? where we take the continuum limit of the oscillators to reveal the acoustic wave equation.

There are many treatments of acoustic waves in the literature, and we made use of Chapter VIII in [Landau and Lifshitz \(1987\)](#), Chapter 9 in [Fetter and Walecka \(2003\)](#), Section 15.2 of [Kundu et al. \(2016\)](#), and Section 16.5 in [Thorne and Blandford \(2017\)](#). Additionally, the [second half of this video](#) offers a pedagogical introduction to acoustic waves.

<b>3.1</b>	<b>Loose threads</b>	<b>64</b>
<b>3.2</b>	<b>Conceptual introduction to acoustic waves</b>	<b>64</b>
<b>3.3</b>	<b>Lagrangian perspective</b>	<b>65</b>
3.3.1	Mass conservation	66
3.3.2	Momentum conservation	66
3.3.3	Equation of state and acoustic wave equation	67
3.3.4	Sound speed	68
3.3.5	Comments and further study	68
<b>3.4</b>	<b>Eulerian perspective</b>	<b>69</b>
3.4.1	Lagrangian wave equation	69
3.4.2	Inhomogeneous background state	70

3.4.3	Acoustic wave equation . . . . .	70
3.4.4	The velocity potential and acoustic wave properties . . . . .	72
<b>3.5</b>	<b>Dispersion relation and flow properties . . . . .</b>	<b>73</b>
3.5.1	Dispersion relation . . . . .	73
3.5.2	Flow properties for acoustic plane waves . . . . .	74
3.5.3	Example acoustic pressure perturbations . . . . .	75
<b>3.6</b>	<b>Energetics of the wave disturbance . . . . .</b>	<b>77</b>
3.6.1	Expressions for the wave energies . . . . .	77
3.6.2	Equipartition of energy . . . . .	78
3.6.3	Energy budget equation . . . . .	80
<b>3.7</b>	<b>Wave momentum . . . . .</b>	<b>80</b>
<b>3.8</b>	<b>Acoustic waves radiated from a piston . . . . .</b>	<b>81</b>
3.8.1	Setting up the physics problem . . . . .	81
3.8.2	Solution in terms of the Green's function . . . . .	82
3.8.3	Method of images for the Green's function . . . . .	83
3.8.4	Velocity potential . . . . .	84
3.8.5	Further study . . . . .	85
<b>3.9</b>	<b>Acoustic waves in a gently varying background . . . . .</b>	<b>85</b>
3.9.1	Scalar potential for density-weighted velocity . . . . .	85
3.9.2	Energetics . . . . .	86
3.9.3	Stress-energy-momentum tensor . . . . .	86
3.9.4	Alternative expression for the action . . . . .	87
3.9.5	Space and time scale separation . . . . .	88
3.9.6	The WKBJ asymptotic solution . . . . .	89
3.9.7	Phase averaged energy . . . . .	91
3.9.8	Whitham's variational principle . . . . .	92
3.9.9	Further study . . . . .	93
<b>3.10</b>	<b>Exercises . . . . .</b>	<b>93</b>

---

## 3.1 Loose threads

- Need the solution to Exercise 3.3.
- Introduce particle displacement vector as per Section 2.1.1 of *Böhler (2014)* for purposes of energetics and equipartition.

## 3.2 Conceptual introduction to acoustic waves

### Physical processes giving rise to acoustic waves

Consider a static background fluid state that is in mechanical and thermodynamic equilibrium. Acoustic waves involve four physical processes within this background state: (i) the density in a compressible fluid changes according to flow divergences and convergences; (ii) density fluctuations lead to pressure fluctuations; (iii) pressure fluctuations lead to fluid particle fluctuations; (iv) for small amplitude fluctuations, fluid particle displacements exhibit space-time coherent oscillatory motion known as linear acoustic waves. Small fluctuations correspond to a fluid particle speed that is much smaller than the acoustic wave speed, with the ratio known as the *Mach number* (we see this relation in Section 3.5.2). Hence, we are concerned only with small Mach number flow in this chapter. As such, acoustic wave properties such as the wave momentum and wave energy are transmitted at a much greater speed than the

transport of properties arising from fluid particle motion (e.g., enthalpy and mass transport). Furthermore, we find that fluid particles feeling the passage of an acoustic wave oscillate in the direction of the wave, with the alignment of the fluid velocity and wave direction characterizing *longitudinal waves*.

The alternating compression and rarefaction of fluid elements within an acoustic wave give rise to alternating pressure work that affects the internal energy. For small Mach number flow, acoustic waves can be assumed to be isentropic, so that the pressure work is reversible. We make use of perfect fluid mechanics throughout this chapter, thus ignoring mixing and heating so that fluid elements maintain constant specific entropy. Indeed, we generally assume the fluid has uniform specific entropy to further isolate the core physical features accounting for acoustic waves.

#### Comments on compressible, incompressible, and the Boussinesq ocean

All real fluids are compressible and thus support acoustic waves. When compressibility is reduced towards zero, so that the continuous media becomes more rigid, the acoustic wave speed increases. Hence, we generally find acoustic waves travel faster through solids than liquids, and faster through liquids than gases. Taking the mathematical limit of a fully incompressible fluid (where the fluid density is uniform and constant) results in an infinite acoustic wave speed, in which case acoustic waves are absent from incompressible fluid dynamics.

There are occasions where we study incompressible fluids in this book. For example, the shallow water models in Part ?? of this book are comprised of incompressible fluid layers with three-dimensional motion, whereas the two-dimensional non-divergent barotropic model in Chapter ?? considers just a two-dimensional velocity (zero vertical motion). The Boussinesq ocean in Chapter ?? offers an important step towards a more realistic fluid. Namely, the Boussinesq ocean is quasi-compressible since the Boussinesq velocity is non-divergent even though the fluid density varies in space and time. The varying density in the Boussinesq ocean gives rise to a divergent velocity that supports acoustic waves. Even so, we commonly ignore the divergent velocity when working with the Boussinesq ocean since this velocity, and the associated acoustic waves, never couple to the Boussinesq ocean dynamics that are the concern of the theory.

### 3.3 Lagrangian perspective

Consider a continuum fluid system whose motion is constrained to one space dimension,  $\hat{x}$ . Assume that all fluid properties and flow properties are independent of the  $y$  and  $z$  directions. We assume the only forces acting on fluid elements arise from pressure, so that gravity, electromagnetism, and friction are ignored. Furthermore, all fluid motion remains close to a static background state, where the background state density and pressure are everywhere uniform with values  $\rho_e$  and  $p_e$ :

$$\rho_e = \text{background state density} \quad \text{and} \quad p_e = \text{background state pressure}. \quad (3.1)$$

Derivation of the equation for the acoustic waves proceeds in three steps. The first concerns the equation for mass conservation; the second arises from momentum conservation; and the third concerns the equation of state relating pressure, density, and specific entropy. Throughout the derivation in this section, we make use of the Lagrangian displacement field,  $\xi(x, t)$ , used in Section ?? when taking the continuum limit of coupled harmonic oscillators. Recall that this

function measures the displacement, at time  $t$ , of a fluid element whose equilibrium position is  $x$ . We here also introduce the density,  $\rho(x, t)$ , as the density of a fluid element whose equilibrium position is  $x$ , and we maintain the same interpretation for the pressure field,  $p(x, t)$ .

### 3.3.1 Mass conservation

In Figure 3.1 we find two configurations for a fluid element, in which the first assumes the fluid is in mechanical equilibrium with constant density,  $\rho_e$ , and mass per unit area

$$M/A = \rho_e \delta x, \quad (3.2)$$

where  $A$  is the horizontal cross sectional area. The second configuration has been displaced by  $\xi(x, t)$  a small amount so that the fluid element is out of mechanical equilibrium. In this case its density,  $\rho(x, t)$ , is found through mass conservation. Namely, the mass in the original rectangle equals to that in the displaced rectangle

$$M/A = \rho_e \delta x \quad (3.3a)$$

$$= \rho [x + \delta x + \xi(x + \delta x, t + \delta t) - x - \xi(x, t + \delta t)] \quad (3.3b)$$

$$= \rho [\delta x + \xi(x + \delta x, t + \delta t) - \xi(x, t + \delta t)] \quad (3.3c)$$

$$= \rho \delta x \left[ 1 + \frac{\partial \xi(x^*, t)}{\partial x} \right] \quad (3.3d)$$

$$\approx \rho \delta x \left[ 1 + \frac{\partial \xi(x, t)}{\partial x} \right]. \quad (3.3e)$$

The penultimate equality made use of the mean value theorem from differential calculus, where  $x^*$  is a point between  $x$  and  $x + \delta x$ , whereas the final step involves the approximation that arises when taking the infinitesimal limit allowing us to evaluate  $\partial \xi / \partial x$  at the position  $x$ . We are thus led to the relation<sup>1</sup>

$$\rho(x, t) = \rho_e (1 + \partial \xi / \partial x)^{-1} \approx \rho_e (1 - \partial \xi / \partial x). \quad (3.4)$$

This equation says that the density of a fluid element,  $\rho(x, t)$ , whose equilibrium position is  $x$ , differs from its equilibrium density,  $\rho_e$ , according to whether the fluid element is expanded ( $\partial \xi / \partial x > 0$ ) or contracted ( $\partial \xi / \partial x < 0$ ).

### 3.3.2 Momentum conservation

At equilibrium the fluid element experiences a pressure,  $p_e$ , that is assumed to be spatially uniform.<sup>2</sup> If fluid elements are displaced in a manner that produces nonzero density perturbations, then the pressure field is likewise modified, in which case fluid moves. Recall that we interpret  $p(x, t)$  as the pressure acting on a fluid element whose equilibrium position is  $x$ . Hence, the net pressure force per area acting on the displaced fluid element is given by the gradient

$$\mathbf{F}^{\text{press}} = [-p(x + \delta x, t) + p(x, t)] \hat{\mathbf{x}} \approx -(\partial p(x, t) / \partial x) \delta x \hat{\mathbf{x}}. \quad (3.5)$$

---

<sup>1</sup>By assuming small perturbations relative to the background state, we disallow the extreme case where  $\partial \xi / \partial x = -1$ , in which the fluid becomes so rarefied (vanishing mass density) that we can no longer make use of the continuum description (see Chapter ?? for more on the continuum approximation).

<sup>2</sup>Recall we are ignoring gravity, so there is no hydrostatic pressure that would give rise to a pressure gradient.

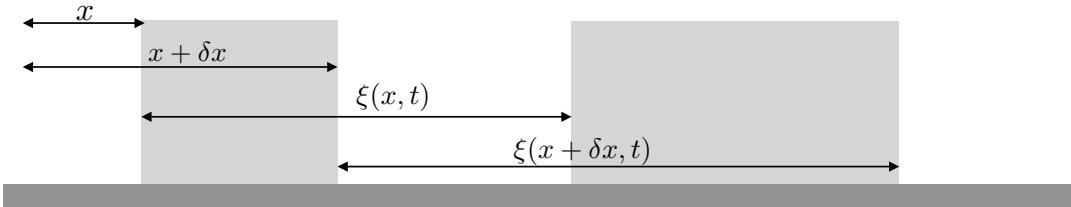


FIGURE 3.1: Schematic of a rectangular fluid element constrained to move just in the  $\hat{x}$  direction with uniform properties in the  $\hat{y}$  and  $\hat{z}$  directions. The fluid element has a constant mass,  $M$ , and with density uniform within the element. The equilibrium configuration in the left rectangle has a mass per unit area  $M/A = \rho_e \delta x$ , where  $A$  is the fixed cross sectional area, and  $\rho_e$  is the equilibrium density. The right rectangle shows the fluid element at a displaced non-equilibrium configuration where the density deviates from its equilibrium value. This density,  $\rho(x, t)$ , is the density of the displaced fluid element (specified by the displacement field  $\xi(x, t)$ ) whose equilibrium position is  $x$ . The mass of the fluid element is identical for both the equilibrium and non-equilibrium configurations, which leads to  $\rho(x, t) = \rho_e (1 + \partial \xi / \partial x)^{-1}$ . This relation means, for example, if  $\partial \xi / \partial x > 0$  (fluid element expands relative to equilibrium), then  $\rho(x) < \rho_e$ .

The position of the fluid element is fully specified by the deviation function,  $\xi(x, t)$ , so that the acceleration is its second time derivative,  $\partial_{tt} \xi(x, t)$ , in which case Newton's equation of motion is given by

$$\frac{\partial}{\partial t} \left[ \rho \frac{\partial \xi}{\partial t} \right] = - \frac{\partial p}{\partial x}. \quad (3.6)$$

### 3.3.3 Equation of state and acoustic wave equation

We now assume an equation of state whereby density is a function of pressure and specific entropy<sup>3</sup>

$$\rho = \rho(p, \mathcal{S}). \quad (3.7)$$

Assuming the fluctuations occur with constant specific entropy,  $\mathcal{S}$ , the equation of motion (3.6) takes the form

$$\frac{\partial}{\partial t} \left[ \rho \frac{\partial \xi}{\partial t} \right] = - \left[ \frac{\partial p}{\partial \rho} \right]_{\mathcal{S}} \frac{\partial \rho}{\partial x}. \quad (3.8)$$

Use of the continuity equation (3.4) allows us to eliminate density in favor of the displacement field

$$\frac{\partial}{\partial t} \left[ \rho \frac{\partial \xi}{\partial t} \right] = \frac{\rho_e c_s^2}{(1 + \partial \xi / \partial x)^2} \frac{\partial^2 \xi}{\partial x^2}, \quad (3.9)$$

where we introduced the inverse squared speed

$$c_s^{-2} = \left[ \frac{\partial \rho}{\partial p} \right]_{\mathcal{S}}, \quad (3.10)$$

where  $c_s$  is interpreted as the sound speed. We can linearize the acoustic equation (3.9) by dropping all terms with the products of the displacement field and assuming density is approximated by its equilibrium value,  $\rho_e$ , in which case we recover the acoustic wave equation

$$(\partial_{tt} - c_s^2 \partial_{xx}) \xi = 0. \quad (3.11)$$

This equation says that the displacements of the fluid elements relative to their equilibrium position satisfy the linear wave equation, and the displacements travel with the sound speed.

<sup>3</sup>We study equations of state, including the ideal gas equation, in Chapters ?? and ??.

### 3.3.4 Sound speed

It is important to observe that the pressure derivative in the sound speed equation (3.10) is computed with a fixed entropy,  $S$ . We are thus treating acoustic waves as reversible adiabatic waves, which accords with our use of a perfect fluid throughout this chapter.<sup>4</sup> This approach is suitable for those cases where the speed of acoustic waves is much larger than the speed of fluid particles, so that an acoustic wave moves through a fluid far faster than the time for enthalpy to be transferred by fluid particles. In this case, acoustic waves are accurately treated as reversible adiabatic waves.

In Section ?? we consider the sound speed in an ideal gas, in which

$$c_s^2 = (p/\rho)(c_p/c_v) = T R^M (c_p/c_v), \quad (3.12)$$

where  $R^M$  is the specific gas constant given by equation (??),  $c_v$  is the specific heat capacity holding specific volume fixed (equation (??)), and  $c_p = c_v + R^M$  is the specific heat capacity holding pressure fixed. For air we have  $c_s \approx 350 \text{ m s}^{-1}$  for  $T = 300 \text{ K}$ . We identify these waves as acoustic (sound) waves due to the agreement of the wave speed (3.12) with the speed of sound measured in the laboratory. Note that LaPlace discovered the relevance of the specific heats ratio,  $c_p/c_v$ , in the expression (3.12) for the sound speed. This ratio arises when recognizing acoustic waves to be constant entropy waves, whereas Newton incorrectly assumed they were isothermal, in which case the specific heat ratio does not appear.

A more compressible media, such as the atmosphere, has a smaller sound speed ( $c_s \approx 350 \text{ m s}^{-1}$ ) than a less compressible media such as the ocean ( $c_s \approx 1500 \text{ m s}^{-1}$ ). Indeed, the sound speed is infinite when the media is fully incompressible, with the infinite speed a signature that the hyperbolic wave system has converted to an elliptic system (see Chapter ??). We offer further discussion of the sound speed in Sections ?? and ?? as part of our study of thermodynamics.

The *Mach number* is the ratio of the fluid particle speed to the sound speed. If a fluid is moving with Mach number greater than unity (supersonic), then there can be discontinuities (shocks) that break the continuum approximation (Chapter ??). In this case, the continuous fluid equations must be supplemented by other physical conditions such as those afforded by molecular dynamics. We have no occasion to study supersonic flow in this book.

### 3.3.5 Comments and further study

The original calculation of sound speed computed the density derivative holding temperature fixed rather than entropy. This mistake, originally made by Newton in his studies of sound, was corrected by LaPlace by noting that acoustic waves more closely maintain adiabatic conditions, which means they preserve specific entropy. It is perhaps a testament to the genius of Newton that even his mistakes took decades to centuries to correct, and often only after being considered in light of new areas of physics (e.g., thermodynamics) that were totally undeveloped during Newton's time.

Elements of this section were taken from Sections 2.1, 2.2, and 2.3 of [Towne \(1967\)](#) and Section 47-3 in Volume I of [Feynman et al. \(1963\)](#). See also Section 1.2 of [Lighthill \(1978\)](#) for more on the sound speed calculation.

---

<sup>4</sup>Matter concentration is also held fixed when considering sound speeds in a fluid with multiple matter constituents.

## 3.4 Eulerian perspective

We here derive the acoustic wave equation using an Eulerian approach, thus offering a complement to the Lagrangian treatment in Section 3.3. Furthermore, we no longer restrict motion to one dimension, so that here the resulting acoustic waves travel in three space dimensions. We also generalize the background state, here allowing for the background density and specific entropy to be static functions of space.

### 3.4.1 Lagrangian wave equation

Consider a single component perfect fluid that does not feel gravity, Coriolis, or friction, so that the only acceleration acting on a fluid element arises from pressure gradients

$$\frac{D\mathbf{v}}{Dt} = -\rho^{-1} \nabla p. \quad (3.13)$$

This expression of the *Euler equation* is coupled to the mass continuity equation (??)

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}, \quad (3.14)$$

that reflects the material constancy of mass following a fluid element. Furthermore, density and pressure are coupled through the equation of state (3.7), so that the density and pressure material time derivatives are related by

$$\frac{D\rho}{Dt} = \left[ \frac{\partial \rho}{\partial p} \right]_s \frac{Dp}{Dt} + \left[ \frac{\partial \rho}{\partial S} \right]_p \frac{DS}{Dt} = c_s^{-2} \frac{Dp}{Dt}, \quad (3.15)$$

where we set  $DS/Dt = 0$  due to the isentropic nature of the perfect fluid, and introduced the squared sound speed

$$c_s^{-2} = \left[ \frac{\partial \rho}{\partial p} \right]_s. \quad (3.16)$$

Combining the continuity equation (3.14) and the material time derivative of the equation of state (3.15) renders

$$\frac{1}{\rho c_s^2} \frac{Dp}{Dt} + \nabla \cdot \mathbf{v} = 0. \quad (3.17)$$

We can relate the velocity divergence to pressure by taking  $D/Dt$  of this equation and then using the Euler equation (3.13)

$$-\frac{D(\nabla \cdot \mathbf{v})}{Dt} = -(\partial_t + v^n \partial_n) (\partial_m v^m) \quad (3.18a)$$

$$= -\nabla \cdot \frac{D\mathbf{v}}{Dt} + \partial_m v^n \partial_n v^m \quad (3.18b)$$

$$= \nabla \cdot (\rho^{-1} \nabla p) + S^m{}_n S^n{}_m - R^m{}_n R^n{}_m. \quad (3.18c)$$

In the final equality we introduced components to the strain rate tensor,  $\mathbf{S}$ , and the rotation tensor,  $\mathbf{R}$ , both of which were introduced in Section ?? when studying the velocity gradient tensor. Furthermore, note that the combination,  $S^m{}_n S^n{}_m - R^m{}_n R^n{}_m$ , also appears in the elliptic pressure equation for the Boussinesq ocean in Section ???. Use of equation (3.18c) along

with  $D/Dt$  of equation (3.17) leads to the pressure equation

$$\frac{D}{Dt} \left[ \frac{1}{\rho c_s^2} \frac{Dp}{Dt} \right] - \nabla \cdot (\rho^{-1} \nabla p) = S^m_n S^n_m - R^m_n R^n_m. \quad (3.19)$$

The left hand side is a *Lagrangian wave equation*, which, when linearized, forms the more familiar acoustic wave equation to be described below. Equation (3.19) describes pressure fluctuations relative to the moving flow, with the nonlinear source on the right hand side arising from strain and rotation within the fluid flow. In Exercise 3.2 we consider the special case of pressure fluctuations when the background flow is a uniform constant.

The pressure equation (3.19) is nonlinear since pressure, density, and velocity are coupled. We garner insight into certain of the physical processes captured by this equation by linearizing around a static background state and examining small amplitude fluctuations, to which we now turn our attention.

### 3.4.2 Inhomogeneous background state

In Section 3.3 we assumed the background state has zero velocity, uniform density, uniform specific entropy, uniform pressure, and uniform sound speed. This trivial state is itself an exact solution to the perfect fluid equations of motion, thus serving as a suitable state from which to study linear fluctuations. Here we introduce a slightly less trivial background state that is also an exact solution to the equations of motion.

We continue to assume a trivial solution of the Euler equation with vanishing velocity (background  $\mathbf{v} = 0$ ) and spatially uniform pressure ( $p_e$  constant). In this case, the continuity equation (3.14) can be satisfied by a background density that is time independent yet spatially inhomogeneous,  $\rho_e = \rho_e(\mathbf{x})$ . This inhomogeneous density is consistent with the equation of state (3.7) if the specific entropy is itself inhomogeneous,  $S_e = S_e(\mathbf{x})$ , in which case

$$\rho_e(\mathbf{x}) = \rho[p = p_e, S = S_e(\mathbf{x})]. \quad (3.20)$$

By allowing  $\rho_e$  and  $S_e$  to be spatially dependent functions, we extend the applicability of the resulting wave equation to the study of acoustic waves propagating in a static inhomogeneous media. A geophysically relevant example concerns an ocean or atmosphere in exact hydrostatic balance (Section ??), with vanishing flow yet density and specific entropy that are functions of geopotential coordinate,  $z$ . [Pierce \(1990\)](#) considers the even more realistic case with a space and time dependent background flow, and a correspondingly nonuniform pressure field, in which case the acoustic wave equation is modified relative to that considered in this chapter.

### 3.4.3 Acoustic wave equation

We now linearize the Euler equation (3.13), along with the mass continuity equation in the form of equation (3.17), and perform the linearization relative to a background state of zero motion yet inhomogeneous density. We thus write pressure, density, and velocity as

$$p = p_e + p' \quad \text{and} \quad \rho = \rho_e(\mathbf{x}) + \rho' \quad \text{and} \quad \mathbf{v} = 0 + \mathbf{v}', \quad (3.21)$$

where the pressure and density perturbations are small relative to their background values  $|p'| \ll p_e$  and  $|\rho'| \ll \rho_e(\mathbf{x})$ , and where the background density is generally a function of space. A positive  $p'$  arises from a local *compression* in the fluid, whereas a negative  $p'$  is a local expansion or *rarefaction*.

### Wave equation with an inhomogeneous background state

Inserting the perturbations (3.21) into the Euler equation (3.13) leads to

$$(\rho_e + \rho') \partial_t \mathbf{v}' + (\rho_e + \rho') (\mathbf{v}' \cdot \nabla) \mathbf{v}' = -\nabla p', \quad (3.22)$$

and dropping all products of perturbation fields leads to the linearized Euler equation

$$\rho_e \partial_t \mathbf{v}' = -\nabla p'. \quad (3.23)$$

Likewise, the linearized continuity equation (3.17) takes the form

$$\partial_t p' + \rho_e c_s^2 \nabla \cdot \mathbf{v}' = 0, \quad (3.24)$$

where the squared sound speed is here determined by compressibility of the equilibrium state

$$c_s^{-2} = \left[ \left( \frac{\partial \rho}{\partial p} \right)_s \right]_{p=p_e}. \quad (3.25)$$

Taking the divergence of the velocity equation (3.23) and making use of the continuity equation (3.24) renders the acoustic wave equation for the anomalous pressure

$$\frac{1}{\rho_e c_s^2} \frac{\partial^2 p'}{\partial t^2} - \nabla \cdot (\rho_e^{-1} \nabla p') = 0. \quad (3.26)$$

This linear wave equation compares to its fully nonlinear analog in equation (3.19). We can perform analogous manipulations to determine the following wave equation satisfied by the velocity divergence

$$\partial_{tt} \mathcal{D}' - \nabla \cdot [\rho_e^{-1} \nabla (\rho_e c_s^2 \mathcal{D}')] = 0 \quad \text{with } \mathcal{D}' = \nabla \cdot \mathbf{v}'. \quad (3.27)$$

Finally, note that from the linearized Euler equation (3.23) that the density weighted velocity fluctuation has a static curl

$$\partial_t [\nabla \times (\rho_e \mathbf{v}')] = 0. \quad (3.28)$$

### Wave equation with a homogeneous background state

Equation (3.26) for the pressure, equation (3.27) for the velocity divergence, and equation (3.28) for the curl of the density weighted velocity constitute the suite of equations for acoustic waves in a static yet inhomogeneous background state. In the remainder of this section we study properties of the wave solutions to these equations, yet restricted to the homogeneous and static background state with  $\rho_e$  a uniform constant and with the sound speed also uniform, in which case we have the linear acoustic wave equations

$$(\partial_{tt} - c_s^2 \nabla^2) p' = 0 \quad \text{and} \quad (\partial_{tt} - c_s^2 \nabla^2) \mathcal{D}' = 0 \quad \text{and} \quad \partial_t (\nabla \times \mathbf{v}') = 0 \quad \text{with} \quad \mathcal{D}' = \nabla \cdot \mathbf{v}'. \quad (3.29)$$

The more realistic, yet complex, case of waves propagating through an inhomogeneous background state requires extra mathematical technology that we develop in Section 3.9.

### 3.4.4 The velocity potential and acoustic wave properties

The locally static curl (equation (3.29)) means that acoustic waves do not alter vorticity. Hence, if the linear system is initialized with zero vorticity then it stays that way.

#### Velocity potential for acoustic waves

Assuming zero initial vorticity, as appropriate when considering fluctuations around a state of rest, allows us to introduce a velocity potential (dimensions squared length per time),  $\psi$ , so that

$$\mathbf{v}' = -\nabla\psi. \quad (3.30)$$

The velocity equation (3.23) thus implies

$$\nabla(p' - \rho_e \partial_t \psi) = 0 \implies p' = \rho_e (\partial_t \psi + K), \quad (3.31)$$

where  $K(t)$  is an arbitrary function of time that is independent of space. The velocity potential is itself arbitrary up to a function of time. Hence, we can choose to work with a modified velocity potential,  $\Psi$ , that absorbs the function  $K$

$$\Psi = \psi + \int^t K(t') dt', \quad (3.32)$$

so that pressure is determined by the time derivative of  $\Psi$  and velocity is determined by the gradient

$$p' = \rho_e \partial_t \Psi \quad \text{and} \quad \mathbf{v}' = -\nabla\Psi. \quad (3.33)$$

Note that the process of transforming from  $\psi$  to  $\Psi$  is referred to as a *gauge transformation*. The same approach is used in Section 4.2.2 when studying surface waves.

Making use of the relation (3.33) in the linearized continuity equation (3.24) renders the wave equation for the velocity potential

$$(\partial_{tt} - c_s^2 \nabla^2) \Psi = 0. \quad (3.34)$$

Furthermore, we can take the gradient of this equation to find that each of the three velocity components satisfies the acoustic wave equation

$$-(\partial_{tt} - c_s^2 \nabla^2) \nabla\Psi = (\partial_{tt} - c_s^2 \nabla^2) \mathbf{v}' = 0. \quad (3.35)$$

#### Density fluctuations in an acoustic wave

We determine the density perturbation within an acoustic wave by linearizing the equation of state (3.7) around the background state<sup>5</sup>

$$\rho = \rho(p, S) \approx \rho_e + c_s^{-2} (p - p_e), \quad (3.36)$$

so that

$$\rho' = \rho - \rho_e = c_s^{-2} p' = \rho_e c_s^{-2} \partial_t \Psi. \quad (3.37)$$

---

<sup>5</sup>Recall that  $S$  is assumed constant for the background state, so that it plays a passive role in the Taylor expansion (3.36).

Taking the time derivative and using the wave equation (3.34) reveals the self-consistency of this result with the linearized mass continuity equation (3.24)

$$\partial_t \rho' = \rho_e c_s^{-2} \partial_{tt} \Psi = \rho_e \nabla^2 \Psi = -\rho_e \nabla \cdot \mathbf{v}'. \quad (3.38)$$

### Temperature fluctuations in an acoustic wave

For an acoustic wave, a differential temperature increment arises just from changes to the pressure while holding entropy fixed

$$dT = \left[ \frac{\partial T}{\partial p} \right]_s dp. \quad (3.39)$$

In Section ?? we referred to this temperature partial derivative as the *adiabatic lapse rate*. Equation (??) provides a practical form of the lapse rate

$$\left[ \frac{\partial T}{\partial p} \right]_s = \frac{T_e \alpha_T}{\rho_e c_p}, \quad (3.40)$$

with  $T_e$  the background temperature,  $c_p$  the specific heat capacity (??), and  $\alpha_T$  the thermal expansion coefficient (??) defined in terms of the *in situ* temperature ( $\alpha_T$  and  $c_p$  are computed for the background state). We thus have a fluctuation of temperature given by

$$T - T_e = T' = \frac{T_e \alpha_T p'}{\rho_e c_p} = (T_e \alpha_T / c_p) \partial_t \Psi. \quad (3.41)$$

Hence, as an acoustic wave propagates it has an associated oscillation of the temperature field due to oscillations in pressure.

## 3.5 Dispersion relation and flow properties

We now characterize physical properties of acoustic waves, focusing on the properties as realized by a traveling acoustic plane wave in a homogeneous media (i.e.,  $\rho_e$  is a uniform constant). These properties are determined largely through the acoustic wave *dispersion relation*, which provides a functional relation between the wave frequency and the wavevector. After determining the dispersion relation, we derive the pressure, density, and velocity fluctuations in an acoustic wave.

### 3.5.1 Dispersion relation

Not every traveling plane wave is a solution to the acoustic wave equation (3.34). Rather, the wavevector and angular frequency must be related in a specific manner that is dependent on physics of the particular wave. This relation is known as the *dispersion relation*. Plugging in the traveling plane wave (1.18) into the acoustic wave equation (3.34) renders the acoustic wave dispersion relation

$$(\omega^2 - c_s^2 \mathbf{k}^2) \Psi = 0 \implies \omega = c_s |\mathbf{k}| \implies C_p = \omega / |\mathbf{k}| = c_s. \quad (3.42)$$

Note that we only consider the positive sign for the angular frequency since  $\omega \geq 0$  corresponds to our convention for wave frequencies (Section 1.4). Traveling acoustic plane waves of a given wavevector have their angular frequency specified by the dispersion relation (3.42). The

dispersion relation (3.42) reveals that the phase speed for acoustic waves,  $C_p$ , equals to the sound speed,  $c_s$ .

The sound speed is a function of the background fluid state; it is not a function of wave properties such as the wavelength or wave frequency. Hence, all acoustic waves, regardless their wavelength, travel at the same phase speed,  $C_p = c_s$ . Correspondingly, we say that acoustic waves are *non-dispersive*. This property accords with common experience, whereby the variety of sound waves with different frequencies from, say, an orchestra are heard together since all sound frequencies travel with the same speed. Furthermore, it is certainly possible for any particular point in space to be comprised of multiple acoustic waves. Since each wave satisfies the wave equation (3.34), and since the wave equation is linear, acoustic waves satisfy the *principle of superposition*. That is, the sum of multiple traveling acoustic waves is also an acoustic wave that satisfies the same wave equation with same speed,  $c_s$ .

### 3.5.2 Flow properties for acoustic plane waves

We here establish expressions for flow fields, such as velocity, pressure, and density, in the presence of an acoustic plane wave. We start with equation (3.33), which relates pressure and velocity to derivatives of the velocity potential. Following our discussion of traveling plane wave kinematics in Section 1.5.1, write the velocity potential for an acoustic plane wave as

$$\Psi(\mathbf{x}, t) = A \operatorname{Re}[e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha)}] = A \cos(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha), \quad (3.43)$$

which leads to the corresponding pressure fluctuation

$$p' = \rho_e \partial_t \Psi = \rho_e \operatorname{Re}[-i\omega \Psi] = \rho_e \omega A \sin(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha). \quad (3.44)$$

Hence, higher frequency acoustic plane waves have higher magnitude for their fluctuating pressure. Likewise, the fluctuating velocity of fluid particles takes on the form

$$\mathbf{v}' = -\nabla \Psi = -\operatorname{Re}[i\mathbf{k} \Psi] = A \mathbf{k} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha). \quad (3.45)$$

This relation means that fluid particles moving in an acoustic plane wave have their velocity aligned with the wave direction

$$\mathbf{v}' \times \mathbf{k} = 0. \quad (3.46)$$

Furthermore, this relation follows since the vorticity of the linear fluctuations vanish (Section 3.4.4). The alignment of particle velocity and wave vector is a defining feature of *longitudinal waves*.<sup>6</sup> It also means that the squared magnitude of the particle velocity equals to the squared projection of this velocity onto the wave direction

$$|\mathbf{v}'|^2 = (\mathbf{v}' \cdot \hat{\mathbf{k}})^2 = [A |\mathbf{k}| \sin(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha)]^2. \quad (3.47)$$

Finally, the fluctuating density within an acoustic wave is given by the linearized equation of state (3.37), which takes the following form for an acoustic plane wave

$$\rho' = c_s^{-2} p' = c_s^{-2} \rho_e \omega A \sin(\mathbf{k} \cdot \mathbf{x} - \omega t + \alpha). \quad (3.48)$$

---

<sup>6</sup>We later encounter waves in non-divergent flows,  $\nabla \cdot \mathbf{v} = 0$ , which are *transverse waves* whereby  $\mathbf{k} \cdot \mathbf{v} = 0$ . That is, for transverse waves the fluid particle motion is perpendicular to the wave direction, whereas for longitudinal waves the particle motion is parallel to the wave direction.

Equations (3.44), (3.45), and (3.48) imply the following relations between the fluctuating pressure, velocity, and density in an acoustic plane wave

$$\rho_e \omega \mathbf{v}' = p' \mathbf{k} \implies p' = \rho_e c_s \mathbf{v}' \cdot \hat{\mathbf{k}} \implies \rho' = \rho_e \mathbf{v}' \cdot \hat{\mathbf{k}} / c_s. \quad (3.49)$$

where we used the dispersion relation (3.42). These relations reveal that pressure, velocity, and density fluctuations are in-phase; i.e., they have the same phase. The density fluctuation in equation (3.49) reveals that its ratio with the background density equals to the ratio of the fluid particle speed to the sound speed. This ratio is the *Mach number*, so that for an acoustic plane wave we have

$$|p'|/(\rho_e c_s^2) = |\rho'|/\rho_e = |\mathbf{v}' \cdot \hat{\mathbf{k}}|/c_s = Ma \ll 1. \quad (3.50)$$

Hence, small Mach number corresponds to small density fluctuation relative to the background density. We made use of this relation when introducing the Boussinesq ocean at the start of Chapter ??.

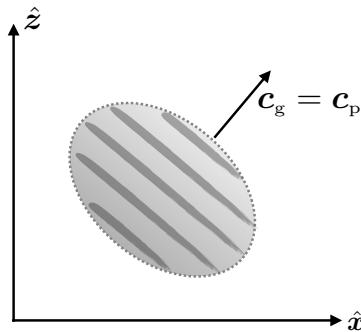


FIGURE 3.2: Example packet of acoustic waves, with the gray shaded lines denoting constant phase surfaces. Acoustic waves are non-dispersive, so that the group and phase velocities are equal,  $\mathbf{c}_p = \mathbf{c}_g$ , and all acoustic waves have phase speed given by the speed of sound,  $C_p = c_s$ . The fluid particle motion is parallel to the phase velocity according to equation (3.45), with such particle motion characterizing longitudinal waves.

### 3.5.3 Example acoustic pressure perturbations

We noted in the introduction to this chapter that acoustic waves, and their associated pressure perturbations, play a negligible role in large-scale geophysical fluid flows, such as those associated with atmospheric weather patterns. To support that contention, consider some common sounds and examine their pressure perturbations. For that purpose, we make use of the decibel scale (dB) for the sound pressure level (SPL)

$$\text{SPL} = 20 \log_{10}(|p'|/p_{\text{ref}}) \implies |p'| = p_{\text{ref}} 10^{\text{SPL}/20}. \quad (3.51)$$

It is conventional to choose the reference pressure,  $p_{\text{ref}}$ , so that 0 dB is a quiet sound at the threshold of human hearing, with

$$p_{\text{ref}} = 20 \times 10^{-6} \text{ Pa} = 20 \mu\text{Pa} \quad (3.52)$$

serving as the international convention. Following equation (3.50), we can determine the Mach number associated with a particular sound pressure level (assuming acoustic plane waves) according to

$$Ma = |p'|/(\rho_e c_s^2) = |p'|/[p(c_p/c_v)], \quad (3.53)$$

where we made use of the ideal gas relation (??) for the sound speed

$$\rho_e c_s^2 = p c_p / c_v. \quad (3.54)$$

We evaluate the sound speed at the standard sea level atmospheric pressure,  $p = p_{\text{stand}} = 101.325 \times 10^3 \text{ Pa}$ , and assume the specific heat capacity ratio,  $c_p/c_v = 7/5$ , which holds for an ideal diatomic gas, in which case

$$\rho_e c_s^2 = p_{\text{stand}} (c_p/c_v) = 140 \times 10^3 \text{ Pa}. \quad (3.55)$$

Table 3.1 tabulates the SPL, pressure perturbation, and Mach number for some common sounds. We see that even extremely loud sounds, relative to human hearing, have very small Mach numbers, thus justifying the use of linear acoustic wave dynamics for their description.

To gauge the size of these acoustic pressure fluctuations relative to typical atmospheric pressure fluctuations, consider a middle latitude geostrophic wind, in which case equation (??) gives

$$\rho f \hat{\mathbf{z}} \times \mathbf{u}_{\text{geostrophic}} = -\nabla p_{\text{geostrophic}} \implies |p'_{\text{geostrophic}}| \sim \rho f U L, \quad (3.56)$$

where  $U$  is the scale of the geostrophic wind speed, and  $L$  is the horizontal length scale over which the winds vary. Taking  $f = 10^{-4} \text{ s}^{-1}$ ,  $U = 10 \text{ m s}^{-1}$ ,  $\rho = 1 \text{ kg m}^{-3}$ , and  $L = 10^6 \text{ m}$  we find a typical pressure fluctuation

$$|p'_{\text{geostrophic}}| \approx 10^3 \text{ Pa}. \quad (3.57)$$

This pressure fluctuation is on the order of that found inside of an automobile exhaust system (see Table 3.1). However, the atmospheric fluctuation associated with this geostrophic wind extends over thousands of kilometers, whereas the sound waves inside of an exhaust system extend over a fraction of a meter. Hence, the energy contained in the atmospheric weather pattern is many orders larger than that for even the loudest sounds extending over typical human length scales. This example further emphasizes the irrelevance of acoustic waves for large-scale geophysical fluid flows.

SOUND	SPL (dB)	$ p'  (\text{Pa})$	$ p' / p'_{\text{geostrophic}} $	Ma
soft whisper	30	$6.3 \times 10^{-4}$	$6.3 \times 10^{-7}$	$4.5 \times 10^{-9}$
normal conversation	60	$2 \times 10^{-2}$	$2 \times 10^{-5}$	$1.4 \times 10^{-7}$
noisy factory	90	0.63	$6.3 \times 10^{-4}$	$4.5 \times 10^{-6}$
rock concert	115	11	$1.1 \times 10^{-2}$	$8 \times 10^{-5}$
aircraft engine	130	63	$6.3 \times 10^{-2}$	$4.5 \times 10^{-4}$
automobile exhaust system	160	$2 \times 10^3$	2	$1.4 \times 10^{-2}$

TABLE 3.1: Acoustic properties of common sounds, following Example 15.2 from [Kundu et al. \(2016\)](#). The first column lists the sound, the second column the sound pressure level in dB, the third column the corresponding pressure fluctuation in Pa, the fourth column the ratio of the pressure fluctuation to a pressure fluctuation associated with a geostrophic atmospheric fluctuation,  $|p'_{\text{geostrophic}}| \approx 10^3 \text{ Pa}$  (equation (3.57)), and the fifth column the Mach number assuming standard atmospheric sea level pressure,  $p_{\text{stand}} = 1.01 \times 10^5 \text{ Pa}$ . Since the Mach number is far smaller than unity even for the loudest sound, we are justified in using linear acoustic wave theory to describe the propagation of these sounds.

## 3.6 Energetics of the wave disturbance

We here specialize the energetic analysis from Section ?? and ?? to study the energetics of an acoustic wave in a homogeneous background state. For an acoustic wave, the total energy is the sum of the kinetic energy of the oscillating fluid particles, plus the changes to internal energy of the fluid that arise from pressure work (recall we are ignoring gravity). We thus write the total energy per mass,  $\mathcal{E}$ , as the sum of the internal energy per mass,  $\mathcal{I}$ , plus kinetic energy per mass,  $\mathcal{K}$ ,

$$\mathcal{E} = \mathcal{I} + \mathcal{K}. \quad (3.58)$$

The Eulerian form of the total energy equation is given by equation (??), which here takes the form

$$\frac{\partial(\rho\mathcal{E})}{\partial t} + \nabla \cdot [\rho \mathbf{v} (\mathcal{I} + \mathcal{K} + p/\rho)] = 0, \quad (3.59)$$

where  $\mathcal{I} + p/\rho$  is the enthalpy per mass (Section ??)

$$\mathcal{H} = \mathcal{I} + p/\rho, \quad (3.60)$$

and we set the thermal and chemical fluxes to zero as per a perfect fluid.

In the remainder of this section we specialize this energy equation to the case of linear acoustic wave fluctuations. It is notable that the terms appearing in the leading order energy equation are quadratic in wave disturbance amplitude, and they are each evaluated using wave properties evaluated to first order.

### 3.6.1 Expressions for the wave energies

To linearize terms in the energy equation (3.59) requires us to drop third order products of fluctuating quantities while keeping second, first, and zeroth order terms. This procedure is further supported by our ability to develop a self-consistent and closed energy budget for acoustic waves.

#### Kinetic energy of an acoustic wave

Since velocity is a first order quantity, the kinetic energy per volume is

$$\rho\mathcal{K} = (\rho_e + \rho') \mathbf{v}' \cdot \mathbf{v}' / 2 \approx \rho_e \mathbf{v}' \cdot \mathbf{v}' / 2. \quad (3.61)$$

#### Internal (potential) energy of an acoustic wave

For the internal energy, recall from Section ?? that its natural functional dependence for a single-component fluid is

$$\mathcal{I} = \mathcal{I}(\mathcal{S}, \rho). \quad (3.62)$$

Since entropy is held fixed in an acoustic wave, consider the following Taylor series approximation for the internal energy per volume as computed around the background state

$$\rho\mathcal{I} \approx \rho_e \mathcal{I}_e + (\rho - \rho_e) \left[ \frac{\partial(\rho\mathcal{I})}{\partial\rho} \right]_{\rho=\rho_e} + \frac{(\rho - \rho_e)^2}{2} \left[ \frac{\partial^2(\rho\mathcal{I})}{\partial\rho^2} \right]_{\rho=\rho_e}. \quad (3.63)$$

The first partial derivative is given by

$$\frac{\partial(\rho \mathcal{I})}{\partial \rho} = \mathcal{I} + \rho \frac{\partial \mathcal{I}}{\partial \rho} = \mathcal{I} + p/\rho = \mathcal{H}, \quad (3.64)$$

where we used identity (??) in the penultimate step. The second partial derivative is thus given by

$$\frac{\partial^2(\rho \mathcal{I})}{\partial \rho^2} = \frac{\partial \mathcal{H}}{\partial \rho} = \frac{\partial \mathcal{H}}{\partial p} \frac{\partial p}{\partial \rho} = c_s^2/\rho, \quad (3.65)$$

where we used the identity (??) for the final step. We also used the equation of state (3.7) to convert the density derivative to a pressure derivative in the penultimate step.

Bringing these results together leads to the approximate internal energy per volume

$$\rho \mathcal{I} \approx \rho_e \mathcal{I}_e + \rho' \mathcal{H}_e + (c_s \rho')^2 / (2 \rho_e). \quad (3.66)$$

The first term on the right hand side is a constant measuring the internal energy per volume of the background state, and it has no relation to the acoustic wave. The second term is the background enthalpy per mass times the fluctuating density. When integrating over the full domain, the fluctuating density vanishes due to mass conservation. That is, we assume the mass in the domain is the same in the background state as well as when there are acoustic waves, so that

$$\int_{\mathcal{R}} \rho dV = \int_{\mathcal{R}} \rho_e dV \implies \int_{\mathcal{R}} \rho' dV = 0. \quad (3.67)$$

Since the  $\rho' \mathcal{H}_e$  term drops out from a domain volume integral, it is common to drop this term when studying the energy density for an acoustic wave. However, we choose not to drop it at this point in the discussion. Rather, we find in Section 3.6.3 that  $\rho' \mathcal{H}_e$  naturally cancels from the flux-form acoustic energy budget equation, at which point we can confidently conclude that it has no significance to the energetic transfers within an acoustic wave. The third term in equation (3.66),  $(c_s \rho')^2 / (2 \rho_e)$ , is a potential energy per volume in the acoustic wave that arises from fluid compressibility.

### Total energy per volume of an acoustic wave

We conclude that the total energy per volume, accurate to second order in fluctuating acoustic wave fields, is given by

$$\rho \mathcal{E} = \rho' \mathcal{H}_e + (\rho_e/2) [(c_s \rho'/\rho_e)^2 + \mathbf{v}' \cdot \mathbf{v}']. \quad (3.68)$$

As a final step, we make use of equations (3.33) and (3.37) to write the wave energy per volume in terms of the velocity potential

$$\rho \mathcal{E} = \rho' \mathcal{H}_e + (\rho_e/2) [(c_s \rho'/\rho_e)^2 + \mathbf{v}' \cdot \mathbf{v}'] = \frac{\rho_e}{2 c_s^2} [2 \mathcal{H}_e \partial_t \Psi + (\partial_t \Psi)^2 + c_s^2 \nabla \Psi \cdot \nabla \Psi]. \quad (3.69)$$

## 3.6.2 Equipartition of energy

Following our study of energy for a simple harmonic oscillator in Section ??, we here study how energy in an acoustic wave is partitioned between kinetic energy and potential energy. Recall that the simple harmonic oscillator equally partitions energy when time averaging over an oscillation period; i.e., performing the phase average of Section ?? . That result follows from

the *virial theorem* considered in Section ???. We here verify energy equipartition also holds for acoustic waves.

As for the oscillator, we expect energy equipartition for acoustic waves, thus reflecting the alternating exchange of energy between kinetic and potential. Mathematically, equipartition arises since the potential energy is a homogeneous function of degree two, with details of this property discussed as part of the virial theorem from Section ???. Here, we show that energy equipartition holds instantaneously for a monochromatic traveling acoustic wave. However, for a general linear acoustic fluctuation, equipartition holds only when averaging over a wave period plus integrating over the closed spatial domain.

### Equipartition for a monochromatic traveling acoustic wave

For an acoustic plane wave, the energy density simplifies through use of equation (3.49) to render

$$(c_s \rho' / \rho_e)^2 = (\mathbf{v}' \cdot \hat{\mathbf{k}})^2 = \mathbf{v}' \cdot \mathbf{v}', \quad (3.70)$$

where the second equality holds since the fluid particle velocity within the plane wave is aligned with the plane wave phase velocity. We thus find that the potential energy per volume and kinetic energy per volume contribute an equal amount to the acoustic plane wave's energy per volume

$$\rho \mathcal{E} = \rho' \mathcal{H}_e + (\rho_e/2) [(c_s \rho' / \rho_e)^2 + \mathbf{v}' \cdot \mathbf{v}'] = \rho' \mathcal{H}_e + 2 \rho_e \mathcal{K}. \quad (3.71)$$

### Equipartition for an arbitrary periodic linear fluctuation

The general expression (3.69) for the energy in a linear fluctuation does not render energy equipartition at each point in space and time. Rather, being inspired by the oscillator in Section ??, we here show that a phase and space averaged energy does possess equipartition. For the phase average we integrate the energy density over a single period, and doing so eliminates the background enthalpy term,  $2 \mathcal{H}_e \partial_t \Psi$ , thus leaving

$$\int_0^{2\pi/\omega} \rho \mathcal{E} dt = \frac{\rho_e}{2 c_s^2} \int_0^{2\pi/\omega} [(\partial_t \Psi)^2 + c_s^2 \nabla \Psi \cdot \nabla \Psi] dt. \quad (3.72)$$

Now integrate by parts and make use of the wave equation (3.34) to find

$$\int_0^{2\pi/\omega} \rho \mathcal{E} dt = \frac{\rho_e}{2 c_s^2} \int_0^{2\pi/\omega} [-\Psi \partial_{tt} \Psi + c_s^2 \nabla \Psi \cdot \nabla \Psi] dt \quad (3.73a)$$

$$= \frac{\rho_e}{2} \int_0^{2\pi/\omega} [-\Psi \nabla^2 \Psi + \nabla \Psi \cdot \nabla \Psi] dt \quad (3.73b)$$

$$= \frac{\rho_e}{2} \int_0^{2\pi/\omega} [-\nabla \cdot (\Psi \nabla \Psi) + 2 \nabla \Psi \cdot \nabla \Psi] dt. \quad (3.73c)$$

We see that for an arbitrary linear and periodic fluctuation, energy equipartition is realized by time integrating over a wave period and then integrating over a spatially closed or spatially periodic domain

$$\int_{\mathcal{R}} \left[ \int_0^{2\pi/\omega} \rho \mathcal{E} dt \right] dV = \rho_e \int_{\mathcal{R}} \left[ \int_0^{2\pi/\omega} \nabla \Psi \cdot \nabla \Psi dt \right] dV = 2 \rho_e \int_{\mathcal{R}} \left[ \int_0^{2\pi/\omega} \mathcal{K} dt \right] dV. \quad (3.74)$$

This equation says that the phase and domain averaged flow contains an equal amount of kinetic energy as internal energy. This result holds for any periodic acoustic fluctuation, and is not specific to plane acoustic waves.

### 3.6.3 Energy budget equation

Taking the partial time derivative of the energy per volume, (3.69), leads to

$$\partial_t(\rho \mathcal{E}) = \frac{\rho_e}{c_s^2} [\partial_{tt}\Psi (\mathcal{H}_e + \partial_t\Psi) + c_s^2 \nabla\Psi \cdot \nabla\partial_t\Psi] \quad (3.75a)$$

$$= \rho_e [\nabla^2\Psi (\mathcal{H}_e + \partial_t\Psi) + \nabla\Psi \cdot \nabla\partial_t\Psi] \quad (3.75b)$$

$$= \rho_e \nabla \cdot [(\mathcal{H}_e + \partial_t\Psi) \nabla\Psi] \quad (3.75c)$$

$$= -\rho_e \nabla \cdot [\mathbf{v}' (\mathcal{H}_e + p'/\rho_e)]. \quad (3.75d)$$

We can thus write the acoustic energy equation in terms of the velocity potential

$$\partial_t [2\mathcal{H}_e \partial_t\Psi + (\partial_t\Psi)^2 + c_s^2 \nabla\Psi \cdot \nabla\Psi] = 2c_s^2 \nabla \cdot [(\mathcal{H}_e + \partial_t\Psi) \nabla\Psi]. \quad (3.76)$$

The two  $\mathcal{H}_e$  terms cancel identically through use of the wave equation (3.34), thus leaving the acoustic energy equation, here written in two equivalent forms

$$(1/2) \partial_t [(\partial_t\Psi)^2 + c_s^2 \nabla\Psi \cdot \nabla\Psi] = c_s^2 \nabla \cdot (\partial_t\Psi \nabla\Psi) \quad (3.77a)$$

$$(\rho_e/2) \partial_t [(c_s p'/\rho_e)^2 + \mathbf{v}' \cdot \mathbf{v}'] = -\nabla \cdot (\mathbf{v}' p'). \quad (3.77b)$$

The energy equation (3.77b) identifies  $\mathbf{v}' p'$  as the energy flux for acoustic waves, whose convergence affects a local time change to the wave energy per volume. In this manner we have established the budget equations for acoustic energy, as summarized by<sup>7</sup>

$$\partial_t(\rho_e \mathcal{E}) = -\nabla \cdot \mathbf{J}_{\mathcal{E}} \quad (3.78a)$$

$$\mathcal{E} = [(c_s p'/\rho_e)^2 + \mathbf{v}' \cdot \mathbf{v}']/2 = [c_s^{-2} (\partial_t\Psi)^2 + \nabla\Psi \cdot \nabla\Psi]/2 \quad (3.78b)$$

$$\mathbf{J}_{\mathcal{E}} = \mathbf{v}' p' = -\rho_e \partial_t\Psi \nabla\Psi. \quad (3.78c)$$

Finally, for an acoustic plane wave, use of the relations (3.49) allows us to write the wave energy flux in terms of the wave energy density

$$\mathbf{J}_{\mathcal{E}} = \mathbf{v}' p' = \mathbf{v}' (\rho_e c_s \mathbf{v}' \cdot \hat{\mathbf{k}}) = c_s \rho_e (\mathbf{v}' \cdot \hat{\mathbf{k}})^2 \hat{\mathbf{k}} = c_s \rho_e \mathcal{E} \hat{\mathbf{k}} = \rho_e \mathcal{E} \mathbf{c}_p \iff \text{plane wave.} \quad (3.79)$$

That is, for the acoustic plane wave, the flux of energy moves with the plane wave phase velocity. For dispersive waves in later chapters, we find that the energy flux moves with the group velocity rather than the phase velocity.

## 3.7 Wave momentum

The linear momentum in a fluid region,  $\mathcal{R}$ , is given by (see Section ??)

$$\mathbf{P} = \int_{\mathcal{R}} \rho \mathbf{v} dV, \quad (3.80)$$

---

<sup>7</sup>As shown following equation (3.76), the term  $p' \mathcal{H}_e$  plays no role in energy transformations. Hence, we drop it from  $\mathcal{E}$  in equation (3.78b).

so that the linear momentum of acoustic waves is given by

$$\mathbf{P} = \int_{\mathcal{R}} (\rho_e + \rho') \mathbf{v}' dV \quad (3.81a)$$

$$= -\rho_e \int_{\mathcal{R}} \nabla \Psi dV + \int_{\mathcal{R}} \rho' \mathbf{v}' dV \quad (3.81b)$$

$$= -\rho_e \oint_{\partial\mathcal{R}} \Psi \hat{\mathbf{n}} dV + c_s^{-2} \int_{\mathcal{R}} \mathbf{J}_{\mathcal{E}} dV. \quad (3.81c)$$

If the waves are localized to a finite region, such as in a wave packet rather than a plane wave, then we can drop the boundary integral so long as the boundary of the domain extends outside the region where the packet is located. In this case, the integrated energy flux equals to the squared wave speed times the linear momentum

$$c_s^2 \mathbf{P} = \int_{\mathcal{R}} p' \mathbf{v}' dV = \int_{\mathcal{R}} \mathbf{J}_{\mathcal{E}} dV. \quad (3.82)$$

## 3.8 Acoustic waves radiated from a piston

In this section we introduce the study of acoustic wave radiation, here for the specific case of a circular piston in a flat wall (see Figure 3.3). In addition to exemplifying some of the physical ideas presented earlier in the chapter, this section illustrates the use of the Green's function for the Helmholtz equation that we originally studied in Section ???. We do not pursue the radiation problem to its entirety as doing so is quite an extensive exercise that is lucidly and thoroughly presented in section 51 of [Fetter and Walecka \(2003\)](#). Instead, our goal is mostly to illustrate the style of problems that can be solved using Green's functions for acoustics.

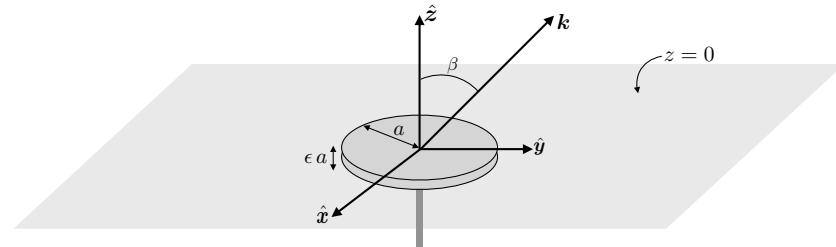


FIGURE 3.3: Illustrating the geometry of a piston radiating acoustic waves as studied in Section 3.8. The piston is circular with radius  $r = a$  and situated within an infinite rigid wall along the  $z = 0$  plane. We make use of cylindrical-polar coordinates from Section ??, with origin at the center of the piston and vertical axis through the origin. The piston oscillates along the vertical axis with a vertical position  $z(t) = \epsilon a e^{-i\omega t}$ , where  $\epsilon \ll 1$  is small and  $\omega$  is the angular frequency of the oscillations. The piston generates acoustic waves with angular frequency  $\omega$  in the surrounding fluid. We depict a single acoustic wavevector,  $\mathbf{k}$ , making an angle,  $\beta$ , with respect to the vertical axis, though note that waves are radiated throughout space. Given the assumed small amplitude of the piston motion, we evaluate boundary conditions on the  $z = 0$  plane, which is sufficiently accurate for linear analysis. Our focus concerns waves within the region  $z > 0$ .

### 3.8.1 Setting up the physics problem

We are interested in properties of acoustic waves radiated in the  $z > 0$  half-plane as depicted in Figure 3.3. If the piston was at rest ( $\omega = 0$ ) at  $z = 0$ , then we would merely be studying acoustic waves in the  $z > 0$  half-plane as described by the boundary value problem for the

velocity potential

$$(\partial_{tt} - c_s^2 \nabla^2) \Psi = 0 \quad z > 0 \quad (3.83a)$$

$$\hat{\mathbf{z}} \cdot \nabla \Psi = 0 \quad z = 0. \quad (3.83b)$$

Indeed, if there is no wavemaker (i.e., no piston motion), then  $\Psi$  would be a space-time constant and so there would be no waves.

Now turn on the piston and write its vertical position as

$$z(t) = \begin{cases} \epsilon a e^{-i\omega t} & r \leq a \\ 0 & r > a, \end{cases} \quad (3.84)$$

where  $r$  is the radial distance from the vertical axis and  $\epsilon \ll 1$  is a small positive and non-dimensional number. The corresponding vertical velocity of the piston is

$$\dot{z}(t) = \begin{cases} -i\epsilon a \omega e^{-i\omega t} & r \leq a \\ 0 & r > a. \end{cases} \quad (3.85)$$

We are thus led to the boundary value problem for the velocity potential

$$(\partial_{tt} - c_s^2 \nabla^2) \Psi = 0 \quad z > 0 \quad (3.86a)$$

$$\hat{\mathbf{z}} \cdot \nabla \Psi = 0 \quad z = 0 \text{ and } r > a \quad (3.86b)$$

$$\hat{\mathbf{z}} \cdot \nabla \Psi = i\epsilon a \omega e^{-i\omega t} \quad z = 0 \text{ and } r \leq a. \quad (3.86c)$$

The time dependent boundary condition generates acoustic waves, and these waves manifest in the velocity potential. We restrict attention to times long enough after the initial setup that allows the wave field to be present throughout the fluid, in which case the velocity potential can be written as a monochromatic field

$$\Psi(\mathbf{x}, t) = \Phi(\mathbf{x}) e^{-i\omega t}. \quad (3.87)$$

We thus find that the time-independent portion of the velocity potential satisfies the Helmholtz boundary value problem

$$[\nabla^2 + (\omega/c_s)^2] \Phi = 0 \quad z > 0 \quad (3.88a)$$

$$\hat{\mathbf{z}} \cdot \nabla \Phi = 0 \quad z = 0, r > a \quad (3.88b)$$

$$\hat{\mathbf{z}} \cdot \nabla \Phi = i\epsilon a \omega \quad z = 0, r \leq a. \quad (3.88c)$$

### 3.8.2 Solution in terms of the Green's function

Following our study in VOLUME 1 of elliptic partial differential equations, we make use of a Green's function to determine the velocity potential resulting from the oscillating piston. In particular, introduce the Green's function,  $G(\mathbf{x}|\mathbf{x}_0)$ , that satisfies the Helmholtz equation with a Dirac source at  $\mathbf{x}_0$  and a homogeneous Neumann boundary condition

$$[\nabla_{\mathbf{x}}^2 + (\omega/c_s)^2] G(\mathbf{x}|\mathbf{x}_0) = -\delta(\mathbf{x} - \mathbf{x}_0) \quad z > 0 \quad (3.89a)$$

$$\hat{\mathbf{z}} \cdot \nabla_{\mathbf{x}} G(\mathbf{x}|\mathbf{x}_0) = 0 \quad z = 0. \quad (3.89b)$$

Multiplying the Green's function equation (3.89a) by  $\Phi(\mathbf{x})$  and integrating over  $z \geq 0$  leads to

$$-\Phi(\mathbf{x}_0) = \int_{\mathcal{R}} \Phi(\mathbf{x}) [\nabla_{\mathbf{x}}^2 + (\omega/c_s)^2] G(\mathbf{x}|\mathbf{x}_0) dV_{\mathbf{x}}, \quad (3.90)$$

with integration by parts yielding

$$-\Phi(\mathbf{x}_0) = \int_{\mathcal{R}} [\nabla_{\mathbf{x}} \cdot (\Phi \nabla_{\mathbf{x}} G - G \nabla_{\mathbf{x}} \Phi) + G (\nabla^2 + (\omega/c)^2) \Phi] dV_{\mathbf{x}}. \quad (3.91)$$

Setting  $\nabla^2 + (\omega/c)^2) \Phi = 0$  as per the Helmholtz equation (3.88a), and use of the divergence theorem, renders

$$-\Phi(\mathbf{x}_0) = \int_{\partial\mathcal{R}} \hat{\mathbf{n}} \cdot (\Phi \nabla_{\mathbf{x}} G - G \nabla_{\mathbf{x}} \Phi) d\mathcal{S}. \quad (3.92)$$

There are two boundaries to consider. The first is at  $z = \infty$ , at which we assume the fields vanishes so we can ignore this boundary. The second boundary is at  $z = 0$ , where  $\hat{\mathbf{n}} = -\hat{\mathbf{z}}$  and the Neumann boundary conditions for  $G$  and  $\Phi$  bring about the expression

$$-\Phi(\mathbf{x}_0) = \int_0^a \int_0^{2\pi} G \hat{\mathbf{z}} \cdot \nabla_{\mathbf{x}} \Phi r dr d\vartheta, \quad (3.93)$$

where  $\vartheta$  is the polar angle. This is a very tidy result that says we merely need to determine the Green's function over the region of the piston,  $r \leq a$  at  $z = 0$ , in order to determine the velocity potential and hence the acoustic wave field for the  $z > 0$  half-space.

### 3.8.3 Method of images for the Green's function

We can make use of a special trick, known as the *method of images*, to determine the Green's function on the  $z \geq 0$  half-space. For this purpose, recall the free space Green's function for the Helmholtz equation from Section ??, which satisfies

$$[\nabla^2 + (\omega/c)^2] \mathcal{G}(\mathbf{x}|\mathbf{x}_0) = -\delta(\mathbf{x} - \mathbf{x}_0), \quad (3.94)$$

and takes the form

$$\mathcal{G}(\mathbf{x}|\mathbf{x}_0) = \frac{e^{i|\mathbf{x}-\mathbf{x}_0|\omega/c}}{4\pi |\mathbf{x} - \mathbf{x}_0|}. \quad (3.95)$$

This Green's function does not satisfy the Neumann boundary condition at  $z = 0$ . However, to generate a Green's function that does, introduce another Dirac delta point source,  $-\delta(\mathbf{x} - \bar{\mathbf{x}}_0)$ , where the source is positioned at

$$\bar{\mathbf{x}}_0 = x_0 \hat{\mathbf{x}} + y_0 \hat{\mathbf{y}} - z_0 \hat{\mathbf{z}}. \quad (3.96)$$

This Dirac source is at a position in the  $z < 0$  half-space reflected across the  $z = 0$  plane from the original source at  $\mathbf{x}_0$ . Notably, the new Dirac source never fires when the field point,  $\mathbf{x}$ , is in the upper,  $z > 0$ , half-space, just in the same manner that the original Dirac source,  $\delta(\mathbf{x} - \mathbf{x}_0)$ , never fires when the field point is in the  $z < 0$  half space. Consequently, the Green's function resulting from these two Dirac sources satisfies the following equation set

$$[\nabla^2 + (\omega/c)^2] G(\mathbf{x}|\mathbf{x}_0) = -\delta(\mathbf{x} - \mathbf{x}_0) \quad z > 0 \quad (3.97a)$$

$$[\nabla^2 + (\omega/c)^2] G(\mathbf{x}|\mathbf{x}_0) = -\delta(\mathbf{x} - \bar{\mathbf{x}}_0) \quad z < 0 \quad (3.97b)$$

$$\hat{\mathbf{z}} \cdot \nabla G = 0 \quad z = 0, \quad (3.97c)$$

and it is given by the sum of the two free space Green's functions,

$$G(\mathbf{x}|\mathbf{x}_0) = \mathcal{G}(\mathbf{x}|\mathbf{x}_0) + \mathcal{G}(\mathbf{x}|\bar{\mathbf{x}}_0) = \frac{e^{i|\mathbf{x}-\mathbf{x}_0|\omega/c}}{4\pi|\mathbf{x}-\mathbf{x}_0|} + \frac{e^{i|\mathbf{x}-\bar{\mathbf{x}}_0|\omega/c}}{4\pi|\mathbf{x}-\bar{\mathbf{x}}_0|}. \quad (3.98)$$

The vanishing Neumann boundary condition (3.97c) results from taking the sum of the original free space Green's function and its image across the  $z = 0$  plane, so that the vertical derivative of the two Green's functions cancel at  $z = 0$ . To verify this result, write

$$\partial_z[\mathcal{G}(\mathbf{x}|\mathbf{x}_0) + \mathcal{G}(\mathbf{x}|\bar{\mathbf{x}}_0)] = \partial_z[F(|\mathbf{x}-\mathbf{x}_0|) + F(|\mathbf{x}-\bar{\mathbf{x}}_0|)] \quad (3.99a)$$

$$= F'(|\mathbf{x}-\mathbf{x}_0|) \frac{z-z_0}{|\mathbf{x}-\mathbf{x}_0|} + F'(|\mathbf{x}-\bar{\mathbf{x}}_0|) \frac{z-\bar{z}_0}{|\mathbf{x}-\bar{\mathbf{x}}_0|} \quad (3.99b)$$

$$= F'(|\mathbf{x}-\mathbf{x}_0|) \frac{z-z_0}{|\mathbf{x}-\mathbf{x}_0|} + F'(|\mathbf{x}-\bar{\mathbf{x}}_0|) \frac{z+z_0}{|\mathbf{x}-\bar{\mathbf{x}}_0|}, \quad (3.99c)$$

where  $F'$  is the derivative of the function. At  $z = 0$ , the two distances are equal so that

$$|\mathbf{x}-\mathbf{x}_0| = |\mathbf{x}-\bar{\mathbf{x}}_0| \quad \text{and} \quad F'(|\mathbf{x}-\mathbf{x}_0|) = F'(|\mathbf{x}-\bar{\mathbf{x}}_0|) \quad \text{at } z = 0, \quad (3.100)$$

which then yields the desired result

$$\partial_z G(\mathbf{x}|\mathbf{x}_0) = 0 \quad z = 0. \quad (3.101)$$

Note that we are only interested in the boundary condition at  $z = 0$  and the behavior of the Green's function in the half-space  $z > 0$ . We have no concern for what the Green's function does in the region  $z < 0$ . So placement of an image source in the lower half-space is merely a trick to enable proper behavior in the  $z \geq 0$  region of interest. This "method of images" constructed Green's function is indeed somewhat magical, as it produces precisely what we need yet without solving any new Green's function equation. Hence, it is a very useful method to construct the Green's function for certain highly symmetric configurations such as given here. However, it is not a general method, so we can only make use of it for certain very special cases. Even so, we accept such gifts when they are available, here providing the Green's function needed to determine the velocity potential,  $\Phi$ , as per equation (3.93).

### 3.8.4 Velocity potential

Making use of the Green's function (3.98) within the expression (3.93) leads to the velocity potential

$$-\Phi(\mathbf{x}_0) = \frac{i\epsilon a \omega}{2\pi} \int_0^a \int_0^{2\pi} \frac{e^{i|\mathbf{x}-\mathbf{x}_0|\omega/c}}{|\mathbf{x}-\mathbf{x}_0|} r dr d\vartheta, \quad (3.102)$$

where the integral is computed at  $z = 0$  within the region  $r \leq a$  of the piston, and where we used the Neumann boundary condition (3.88c) for the velocity potential at  $z = 0$ . Making use of the expression (3.87) for the velocity potential,  $\Psi$ , renders

$$\Psi(\mathbf{x}_0, t) = \Phi(\mathbf{x}_0) e^{-i\omega t} = -\frac{i\epsilon a \omega}{2\pi} \int_0^a \int_0^{2\pi} \frac{e^{i(|\mathbf{x}-\mathbf{x}_0|\omega/c - \omega t)}}{|\mathbf{x}-\mathbf{x}_0|} r dr d\vartheta. \quad (3.103)$$

The velocity potential is constructed by integrating outgoing spherical waves over the area of the piston, with these waves having a strength that is proportional to the piston velocity. As shown in section 51 of [Fetter and Walecka \(2003\)](#), further evaluation of this integral reveals that the acoustic energy and power intensity display characteristic lobes as a function of the angle,  $\beta$ , relative to the vertical axis (see Figure 3.3).

### 3.8.5 Further study

This section is based on section 51 of [Fetter and Walecka \(2003\)](#). Further study of acoustic waves from a planar source can be found in section I.12 of [Lighthill \(1978\)](#).

## 3.9 Acoustic waves in a gently varying background

In Section 3.4.3 we developed the acoustic wave equations for linear waves propagating through a static yet inhomogeneous background state with both  $\rho_e$  and  $c_s$  functions of space. However, subsequent analysis of acoustic waves was restricted to homogeneous background state whereby  $\rho_e$  and  $c_s$  are taken to be space and time constants. In this section we extend the geometric optics formalism from Section 2.3 to render equations for the propagation of acoustic wave energy through a spatially inhomogeneous yet static background.<sup>8</sup> We make use of the *WKBJ asymptotic method*, which is a powerful approach for a variety of purposes.<sup>9</sup> We also present Whitham's variational principle, which, as we show, offers a far more streamlined means to deriving the leading order equations. Both the WKBJ and variational methods make use of the *eikonal wave ansatz* (equation (3.124)) rather than a Fourier plane wave ansatz, with the more general ansatz needed since Fourier methods are not suited to inhomogeneous backgrounds.

### 3.9.1 Scalar potential for density-weighted velocity

Return to the derivation of the acoustic wave equation in Sections 3.4.3 and 3.4.4, now generalizing the velocity potential to include the spatially variable background density. Again, the linearized Euler equation (3.23), its curl, and the linearized continuity equation (3.24) are given by

$$\partial_t(\rho_e \mathbf{v}') = -\nabla p' \quad \text{and} \quad \partial_t[\nabla \times (\rho_e \mathbf{v}')] = 0 \quad \text{and} \quad \partial_t p' + \rho_e c_s^2 \nabla \cdot \mathbf{v}' = 0. \quad (3.104)$$

Assuming a vanishing initial curl,  $\nabla \times (\rho_e \mathbf{v}') = 0$ , as appropriate for a static initial condition, allows us to focus on fluctuations that satisfy  $\nabla \times (\rho_e \mathbf{v}') = 0$  for all time. We are thus led to introduce a scalar potential for the density-weighted velocity

$$\rho_e \mathbf{v}' = -\nabla \chi, \quad (3.105)$$

with  $\chi$  having dimensions of density times squared length per time (compare to the velocity potential  $\psi$  in equation (3.30)). Use of  $\chi$  in the linearized Euler equation (first of equation (3.104)) leads to

$$\nabla(\partial_t \chi - p') = 0. \quad (3.106)$$

<sup>8</sup>Restricting to static background simplifies the maths. Also, it is sufficient for many applications since acoustic waves are quite fast relative to the slower movement of fluid particles. Finally, as discussed by [Pierce \(1990\)](#), the acoustic wave equation picks up extra terms when moving through a time evolving background state, and we are only interested here in the familiar form of the acoustic wave equation.

<sup>9</sup>Chapter 10 of [Bender and Orszag \(1978\)](#) presents a thorough discussion of the WKBJ method from an applied mathematical perspective.

Following the procedure from Section 3.4.4, we choose a gauge so that

$$p' = \partial_t \chi. \quad (3.107)$$

Using this expression for the pressure in the linearized continuity equation (third of equation (3.104)) renders the wave equation for the velocity potential

$$\frac{1}{\rho_e c_s^2} \frac{\partial^2 \chi}{\partial t^2} - \nabla \cdot (\rho_e^{-1} \nabla \chi) = 0. \quad (3.108)$$

Once we have determined  $\chi$ , then the pressure field is determined by taking the time derivative in equation (3.107), and the velocity is determined by taking the gradient according to equation (3.105).

### 3.9.2 Energetics

We studied the energetics of acoustic waves in Section 3.6. Here we provide a terse version of that discussion starting from the wave equation (3.108) and deriving the corresponding energy equation. Multiplying equation (3.108) by  $\partial_t \chi$ , and recalling that  $\rho_e$  and  $c_s$  are time independent, brings the first term to

$$[\rho_e c_s^2]^{-1} \partial_t \chi \partial_{tt} \chi = [2 \rho_e c_s^2]^{-1} \partial_t (\partial_t \chi \partial_t \chi), \quad (3.109)$$

whereas the second term in equation (3.108) gives

$$-\partial_t \chi \nabla \cdot (\rho_e^{-1} \nabla \chi) = -\nabla \cdot (\rho_e^{-1} \partial_t \chi \nabla \chi) + \partial_t (\rho_e^{-1} \nabla \chi \cdot \nabla \chi)/2. \quad (3.110)$$

We are thus led to the energy conservation law

$$(1/2) \partial_t [(\rho_e c_s^2)^{-1} (\partial_t \chi)^2 + \rho_e^{-1} (\nabla \chi)^2] = \nabla \cdot (\rho_e^{-1} \partial_t \chi \nabla \chi), \quad (3.111)$$

which can be equivalently written

$$(1/2) \partial_t [(\rho_e c_s^2)^{-1} (p')^2 + \rho_e \mathbf{v}' \cdot \mathbf{v}'] = -\nabla \cdot (p' \mathbf{v}'), \quad (3.112)$$

which compares directly to the energy equation (3.77a) derived for fluctuations around a homogeneous background state.

### 3.9.3 Stress-energy-momentum tensor

As an aside, we here display the stress-energy-momentum tensor for acoustic waves, following the field theory formalism from Section 2.4.3. The Lagrangian density for acoustic waves can be written as

$$\mathcal{L} = (2 \rho_e)^{-1} [c_s^{-2} (\partial_t \chi)^2 - (\nabla \chi)^2], \quad (3.113)$$

whose Euler-Lagrange field equation is the acoustic wave equation (3.108). The corresponding generalized momentum, Hamiltonian density, and energy flux (equation (??)) are given by

$$\mathcal{P} = \frac{\partial \mathcal{L}}{\partial (\partial_t \chi)} = (\rho_e c_s^2)^{-1} \partial_t \chi = p' / (\rho_e c_s^2) \quad (3.114a)$$

$$\mathcal{H} = \mathcal{P} \partial_t \chi - \mathcal{L} = (2 \rho_e)^{-1} [c_s^{-2} (\partial_t \chi)^2 + (\nabla \chi)^2] = (1/2) [(\rho_e c_s^2)^{-1} (p')^2 + \rho_e \mathbf{v}' \cdot \mathbf{v}'] \quad (3.114b)$$

$$F^a = \frac{\partial \mathcal{L}}{\partial(\partial_{x^a}\chi)} \frac{\partial\chi}{\partial t} = -\rho_e^{-1} \nabla\chi \partial_t\chi = \mathbf{v}' p'. \quad (3.114c)$$

Since the Lagrangian has no explicit time dependence, the Hamiltonian satisfies a source-free continuity equation, which is equivalent to the wave energy equation (3.111)

$$\partial_t \mathcal{H} + \nabla \cdot \mathbf{F} = 0. \quad (3.115)$$

Likewise, for a background state that has no dependence on  $x = x^1$ , the corresponding wave momentum is conserved so that

$$\partial_t T^0_1 + \partial_a T^a_1 = 0. \quad (3.116)$$

Elements to the stress-energy-momentum tensor are generally given by equation (2.48) and take on the following form for acoustic waves

$$T^0_1 = (\rho_e c_s^2)^{-1} \partial_t \chi \partial_x \chi \quad (3.117a)$$

$$T^1_1 = -(2\rho_e)^{-1} [c_s^{-2} (\partial_t \chi)^2 - (\partial_x \chi)^2 + (\partial_y \chi)^2 + (\partial_z \chi)^2] \quad (3.117b)$$

$$T^2_1 = -(\rho_e)^{-1} \partial_y \chi \partial_x \chi \quad (3.117c)$$

$$T^3_1 = -(\rho_e)^{-1} \partial_z \chi \partial_x \chi, \quad (3.117d)$$

which can be written in the equivalent form

$$T^0_1 = -p' u'/c_s^2 \quad (3.118a)$$

$$T^1_1 = -[(\rho_e c_s^2)^{-1} (p')^2 - \rho_e (u')^2 + \rho_e (v')^2 + \rho_e (w')^2]/2 \quad (3.118b)$$

$$T^2_1 = -\rho_e u' v' \quad (3.118c)$$

$$T^3_1 = -\rho_e u' w'. \quad (3.118d)$$

In Exercise 3.3 we verify that the budget equation (3.116) holds for the acoustic waves using elements of the stress-energy-momentum tensor given by equations (3.118a)–(3.118d).

### 3.9.4 Alternative expression for the action

Following the observation in Sections 2.6 and 2.6, we write the action for acoustic waves in the form

$$\mathcal{S} = \frac{1}{2} \int_R \chi \hat{D} \chi \, dV \, dt, \quad (3.119)$$

where the acoustic linear operator,  $\hat{D}$ , is given by

$$-\hat{D} = (\rho_e c_s^2)^{-1} \partial_{tt} - \nabla(\rho_e^{-1}) \cdot \nabla - \rho_e^{-1} \nabla^2. \quad (3.120)$$

To verify the expression (3.119), perform an integration by parts to find

$$2\mathcal{S} = - \int_R \partial_t [(\rho_e c_s^2)^{-1} \chi \partial_t \chi] \, dV \, dt + \int_R \nabla \cdot (\rho_e^{-1} \chi \nabla \chi) \, dV \, dt + 2 \int_R \mathcal{L} \, dV \, dt, \quad (3.121)$$

where we introduced the Lagrangian density

$$\mathcal{L} = (2\rho_e)^{-1} [c_s^{-2} (\partial_t \chi)^2 - (\nabla \chi)^2], \quad (3.122)$$

from equation (3.113). Dropping boundary terms, which renders  $\hat{D}$  a self-adjoint wave operator, leads to the traditional form of the action

$$\mathcal{S} = \int_R \mathcal{L} dV dt. \quad (3.123)$$

### 3.9.5 Space and time scale separation

We are motivated by the hypothesis that waves moving through a gently varying inhomogeneous background are locally close to the plane wave form realized for homogeneous background. Asymptotic methods arising from this hypothesis were developed in Chapter 2, and here we apply it to acoustic waves. Rather than the traveling plane wave ansatz (1.18) used for homogeneous media, we here consider the *eikonal wave ansatz*

$$\chi(\mathbf{x}, t) = \operatorname{Re}[A(\mathbf{x}, t) e^{i\mathcal{P}(\mathbf{x}, t)}] = A(\mathbf{x}, t) \cos[\mathcal{P}(\mathbf{x}, t)], \quad (3.124)$$

where  $A > 0$  is the real amplitude and  $\mathcal{P}$  is the phase introduced in Section 2.2. We assume that spatial variations of the wave amplitude scale according to the length scale,  $L$ , introduced by equation (2.3), so that

$$|\nabla A|/A \sim L^{-1} \sim |\nabla \rho_e|/\rho_e. \quad (3.125)$$

The complement assumption is that the phase function varies over a length scale that is much smaller than  $L$ , so that the local wavenumber satisfies equation (2.3), here written as

$$|\nabla \mathcal{P}| \gg L^{-1} \implies \frac{1/|\nabla \mathcal{P}|}{L} \ll 1. \quad (3.126)$$

Combined with the definition (3.125), the WKBJ approximation considers a wave amplitude and wave phase that satisfy

$$|\nabla A|/A \ll |\nabla \mathcal{P}| = |\mathbf{k}|. \quad (3.127)$$

The same arguments hold for the time scale of the waves, so that

$$|\partial_t A|/A \ll |\partial_t \mathcal{P}| = \omega. \quad (3.128)$$

In developing the WKBJ asymptotic equations from the eikonal wave ansatz, it can be useful to scale the phase by a small non-dimensional parameter,  $\epsilon$ , in which case

$$\mathcal{P} = \varphi/\epsilon \quad \text{where} \quad |\nabla A|/A \sim |\nabla \varphi|, \quad (3.129)$$

and equation (3.127) then takes on the form

$$|\nabla A|/A \ll |\nabla \varphi|/\epsilon. \quad (3.130)$$

An equivalent means to organize similarly scaled terms is to write the wave function (3.124) in the form

$$\chi = A e^{i\mathcal{P}/\sigma}, \quad (3.131)$$

where  $\sigma = 1$  is used to organize the terms, with this approach used in Section 3.9.6.

### 3.9.6 The WKBJ asymptotic solution

We now plug in the eikonal wave ansatz (3.131) to the wave equation (3.133) to develop the WKBJ asymptotic equations that determine how the amplitude and phase evolve. The manipulations are straightforward but somewhat tedious and uninspired. We expose sufficient details to facilitate checking the maths.

#### Notation used for the WKBJ asymptotic expansion

In this subsection we find it useful to introduce the following notation

$$W = 1/(\rho_e c_s^2) \quad (3.132)$$

so that the wave equation (3.108) is written

$$W \frac{\partial^2 \chi}{\partial t^2} - \nabla \cdot (W c_s^2 \nabla \chi) = 0. \quad (3.133)$$

Likewise, the energy equation (3.111) takes on the form

$$(1/2) \partial_t [W (\partial_t \chi)^2 + W c_s^2 (\nabla \chi)^2] = \nabla \cdot (W c_s^2 \partial_t \chi \nabla \chi), \quad (3.134)$$

so that the Hamiltonian density (i.e., the energy) and the energy flux are

$$\mathcal{H} = W [(\partial_t \chi)^2 + c_s^2 (\nabla \chi)^2]/2 \quad \text{and} \quad \mathbf{F} = -W c_s^2 \partial_t \chi \nabla \chi. \quad (3.135)$$

$W$  does not appear in the dispersion relation derived below, thus motivating its introduction.

#### The derivatives

The derivatives are given by the following expressions, organized according to powers of  $\sigma$ :

$$e^{-i\mathcal{P}/\sigma} \nabla \chi = (i/\sigma) A \nabla \mathcal{P} + \nabla A \quad (3.136a)$$

$$e^{-i\mathcal{P}/\sigma} \nabla^2 \chi = -\sigma^{-2} A (\nabla \mathcal{P})^2 + (i/\sigma) (2 \nabla A \cdot \nabla \mathcal{P} + A \nabla^2 \mathcal{P}) + \nabla^2 A \quad (3.136b)$$

$$e^{-i\mathcal{P}/\sigma} \partial_{tt} \chi = -\sigma^{-2} A (\partial_t \mathcal{P})^2 + (i/\sigma) (2 \partial_t A \partial_t \mathcal{P} + A \partial_{tt} \mathcal{P}) + \partial_{tt} A, \quad (3.136c)$$

along with

$$\begin{aligned} e^{-i\mathcal{P}/\sigma} \nabla \cdot (W c_s^2 \nabla \chi) &= W c_s^2 [-\sigma^{-2} A (\nabla \mathcal{P})^2 + (i/\sigma) (2 \nabla A \cdot \nabla \mathcal{P} + A \nabla^2 \mathcal{P}) + \nabla^2 A] \\ &\quad + \nabla (W c_s^2) \cdot [(i/\sigma) A \nabla \mathcal{P} + \nabla A]. \end{aligned} \quad (3.137)$$

#### $\mathcal{O}(\sigma^{-2})$ terms

As for deriving the quasi-geostrophic equations in Part ?? of this book, when deriving asymptotic equations we balance terms according to powers of the expansion parameter, with  $\sigma$  the parameter for the present analysis. The  $\mathcal{O}(\sigma^{-2})$  terms render

$$(\partial_t \mathcal{P})^2 = c_s^2 \nabla \mathcal{P} \cdot \nabla \mathcal{P} \implies \omega^2 = c_s^2 |\mathbf{k}|^2. \quad (3.138)$$

This local dispersion relation is the same as the dispersion relation (3.42) holding for acoustic waves moving in a homogeneous media. Now, however, the angular frequency, wave speed, and

wavenumber are each functions of the spatial position. This result accords with our original hypothesis that waves move through the inhomogeneous media with a local plane wave behavior. It also accords with the assumptions built into the ray theory studied in Sections 2.2 and 2.3.

### $\mathcal{O}(\sigma^{-1})$ terms

The  $\mathcal{O}(\sigma^{-1})$  balance yields

$$W(2\partial_t A \partial_t \mathcal{P} + A \partial_{tt} \mathcal{P}) = W c_s^2 (2\nabla A \cdot \nabla \mathcal{P} + A \nabla^2 \mathcal{P}) + A \nabla \mathcal{P} \cdot \nabla(W c_s^2), \quad (3.139)$$

which can be rearranged to the form

$$(\partial_t + \mathbf{c}_p \cdot \nabla) A = -\frac{A}{2W\omega} [W \partial_t \omega + \nabla \cdot (\mathbf{k} W c_s^2)]. \quad (3.140)$$

The left hand side is the time derivative of the amplitude computed by following the phase velocity,

$$\mathbf{c}_p = c_s \hat{\mathbf{k}} = (\omega/|\mathbf{k}|) \hat{\mathbf{k}}. \quad (3.141)$$

The right hand side of equation (3.140) is a source term that contributes to the amplitude change following the phase velocity. Furthermore, since the acoustic waves are non-dispersive, in which  $\mathbf{c}_p = \mathbf{c}_g$ , equation (3.140) can be written using the time derivative following a ray

$$-\frac{2\omega}{A} \frac{D_r A}{Dt} = \partial_t \omega + \nabla \cdot (\mathbf{k} W c_s^2)/W \quad (3.142a)$$

$$= \partial_t \omega + c_s^2 \nabla \cdot \mathbf{k} + \mathbf{k} \cdot \nabla c_s^2 + c_s^2 \mathbf{k} \cdot (\nabla W)/W. \quad (3.142b)$$

### Massaging the source terms

Terms on the right hand side of the amplitude equation (3.142b) provide sources for the wave amplitude along a ray. We find it useful to rearrange these terms by using the following expression for the wavevector in an acoustic wave

$$\mathbf{k} = \mathbf{c}_p |\mathbf{k}|/c_s = \mathbf{c}_p \omega/c_s^2 = \mathbf{c}_g \omega/c_s^2, \quad (3.143)$$

so that

$$\nabla \cdot \mathbf{k} = (\omega/c_s^2) \nabla \cdot \mathbf{c}_g + \mathbf{c}_g/c_s^2 \cdot \nabla \omega + \omega \mathbf{c}_g \cdot \nabla c_s^{-2}, \quad (3.144)$$

thus leading to

$$\partial_t \omega + c_s^2 \nabla \cdot \mathbf{k} + \mathbf{k} \cdot \nabla c_s^2 = (\partial_t + \mathbf{c}_g \cdot \nabla) \omega + \omega \nabla \cdot \mathbf{c}_g + \omega c_s^2 \mathbf{c}_g \cdot \nabla c_s^{-2} + \mathbf{k} \cdot \nabla c_s^2 \quad (3.145a)$$

$$= \frac{D_r \omega}{Dt} + \omega \nabla \cdot \mathbf{c}_g, \quad (3.145b)$$

where we used

$$\omega c_s^2 \mathbf{c}_g \cdot \nabla c_s^{-2} = -2(\omega/c_s) \mathbf{c}_g \cdot \nabla c_s = -\omega c_s^{-2} (c_s^2 \mathbf{k}/\omega) \cdot \nabla c_s^2 = -\mathbf{k} \cdot \nabla c_s^2. \quad (3.146)$$

These results then lead to the amplitude equation

$$\frac{1}{A} \frac{D_r A}{Dt} = -\frac{1}{2\omega} \frac{D_r \omega}{Dt} - \nabla \cdot (\mathbf{c}_g W)/(2W). \quad (3.147)$$

The dispersion relation has no explicit time dependence given that the background state is time independent. As a result, equation (2.34) from geometric optics means that the angular frequency remains constant along a ray, so that  $D_r\omega/Dt = 0$ . Even so, we retain this term since it hints at the more general case holding for time dependent background states.

For the second term on the right hand side of equation (3.147) we set  $W^{-1} = \rho_e c_s^2$  from equation (3.132) to have

$$-\nabla \cdot (\mathbf{c}_g W)/(2W) = -(\rho_e c_s^2/2) \nabla \cdot (\mathbf{c}_g / \rho_e c_s^2) = -\nabla \cdot \mathbf{c}_g/2 + \mathbf{c}_g \cdot \nabla(\rho_e c_s^2)/(2\rho_e c_s^2), \quad (3.148)$$

so that equation (3.147) takes the form

$$\frac{2}{A} \frac{D_r A}{Dt} + \frac{1}{\omega} \frac{D_r \omega}{Dt} = -\nabla \cdot \mathbf{c}_g + \mathbf{c}_g \cdot \nabla(\rho_e c_s^2)/(\rho_e c_s^2). \quad (3.149)$$

Since  $\rho_e c_s^2$  is assumed to be time independent, we can write this equation as

$$\frac{2}{A} \frac{D_r A}{Dt} + \frac{1}{\omega} \frac{D_r \omega}{Dt} - \frac{D_r(\rho_e c_s^2)}{Dt} = -\nabla \cdot \mathbf{c}_g, \quad (3.150)$$

which combines to read

$$\frac{D_r[(A^2 \omega)/(\rho_e c_s^2)]}{Dt} = -(A^2 \omega)/(\rho_e c_s^2) \nabla \cdot \mathbf{c}_g, \quad (3.151)$$

which can be written as the flux-form conservation equation

$$\partial_t[A^2 \omega/(\rho_e c_s^2)] + \nabla \cdot [\mathbf{c}_g A^2 \omega/(\rho_e c_s^2)] = 0. \quad (3.152)$$

In Section 3.9.7 we interpret this equation in terms of phase averaged energy and wave action.

### 3.9.7 Phase averaged energy

We here compute the phase averaged Hamiltonian (energy) consistent with the assumed scaling from Section 3.9.5. Working with the real expression (3.124) leads to

$$\partial_t \chi = A [A^{-1} \partial_t A \cos \mathcal{P} - \partial_t \mathcal{P} \sin \mathcal{P}] \approx A \omega \sin \mathcal{P}, \quad (3.153)$$

where  $A^{-1} \partial_t A \ll \omega$  as per the assumed scaling (3.128), and  $\omega = -\partial_t \mathcal{P}$  as per equation (2.2). Similarly, we compute the gradient as

$$\nabla \chi = A [A^{-1} \nabla A \cos \mathcal{P} - \nabla \mathcal{P} \sin \mathcal{P}] \approx -A \mathbf{k} \sin \mathcal{P}, \quad (3.154)$$

which follows from the scaling (3.127) and  $\mathbf{k} = \nabla \mathcal{P}$  from equation (2.2). We thus have the expression for the Hamiltonian

$$\mathcal{H} = \frac{A^2 \sin^2 \mathcal{P} (\omega^2 + c_s^2 |\mathbf{k}|^2)}{2 \rho_e c_s^2} = \frac{\omega^2 A^2 \sin^2 \mathcal{P}}{\rho_e c_s^2}, \quad (3.155)$$

where we inserted the dispersion relation (3.138). Taking the phase average leads to

$$\langle \mathcal{H} \rangle = \frac{\omega^2 A^2}{2 \rho_e c_s^2}. \quad (3.156)$$

Making use of the phase averaged energy (3.156) in the amplitude equation (3.152) leads to

$$\partial_t(\langle \mathcal{H} \rangle / \omega) + \nabla \cdot (\mathbf{c}_g \langle \mathcal{H} \rangle / \omega) = 0. \quad (3.157)$$

The quantity

$$\mathcal{A} = \langle \mathcal{H} \rangle / \omega \quad (3.158)$$

is the *wave action* that we studied in Section 2.5.

### 3.9.8 Whitham's variational principle

We studied Whitham's variational principle in Section 2.5, where we claimed that it offers a more streamlined means to derive the leading order phase averaged equations than the WKBJ method used in Sections 3.9.6 and 3.9.7. Taking the results from Section 2.5, we write the phase averaged acoustic Lagrangian (3.113)

$$\langle \mathcal{L} \rangle = (4 \rho_e)^{-1} A^2 (c_s^{-2} \omega^2 - |\mathbf{k}|^2). \quad (3.159)$$

The dispersion relation arises from the Euler-Lagrange equation resulting from a variation of the wave amplitude,

$$\partial \langle \mathcal{L} \rangle / \partial A = 0 \implies \omega^2 = c_s^2 |\mathbf{k}|^2. \quad (3.160)$$

Likewise, the Euler-Lagrange equation corresponding to varying the phase function yields the wave action conservation equation

$$\partial_t(\mathcal{A}) + \nabla \cdot (\mathbf{c}_g \mathcal{A}) = 0, \quad (3.161)$$

where  $\mathbf{c}_g = c_s \hat{\mathbf{k}}$  is the group velocity and the wave action is

$$\mathcal{A} = \partial \langle \mathcal{L} \rangle / \partial \omega = \frac{\omega A^2}{2 \rho_e c_s^2} = \langle \mathcal{H} \rangle / \omega. \quad (3.162)$$

Evidently, we only needed the phase averaged Lagrangian to leading order when using Whitham's principle to derive both the dispersion relation and the wave action conservation equation. In contrast, the WKBJ method required one order higher in the asymptotics when working with the wave equation to derive wave action conservation, which required a relatively large amount of manipulation.

The relative simplicity of Whitham's variational principle, compared to the WKBJ method, is reminiscent of analytic mechanics versus Newtonian mechanics. As shown through a variety of case studies in Chapter ??, analytical mechanics is far more powerful for the study of systems that are generally quite tedious, if not unavailable, using Newtonian mechanics. The price to pay is that analytical mechanics requires some relatively nontrivial theoretical work up front to establish the framework. But once established, exposure of the underlying symmetries and conservation laws is far more streamlined, if not trivial. Likewise, Whitham's variational principle requires some theoretical work up front to derive the framework (Section 2.5). The payoff is that the variational framework is optimized for capturing the leading order physics of phase averaged wave mechanics.

### 3.9.9 Further study

The WKBJ treatment here followed that given in Section 7.3 of [Thorne and Blandford \(2017\)](#). However, we limited attention to the case of a static background state since we only developed the physics for acoustic waves moving in a static inhomogenous background (Section 3.4.2). [Pierce \(1990\)](#) considers the more general case with a space and time dependent background, in which the acoustic wave equation contains further terms.



## 3.10 Exercises

### EXERCISE 3.1: ACOUSTIC MODES IN RECTANGULAR CAVITY (PROBLEM 9.1 OF [Fetter and Walecka \(2003\)](#))

A rectangular cavity with dimensions  $x \in [0, L_x]$ ,  $y \in [0, L_y]$ , and  $z \in [0, L_z]$  is bounded by rigid material walls on all sides. The fluid is homogeneous within the cavity. Determine the eigenfrequencies and eigenfunctions for the acoustic normal modes in this cavity.

Hint: this exercise requires solving the acoustic wave equation in a closed domain with associated kinematic boundary conditions. The resulting acoustic modes are standing wave modes rather than traveling waves, and the wavenumbers are quantized rather than continuous. Note that standing wave modes can be thought of as the superposition of two oppositely traveling waves with identical structure. For example, the sum of a right and left moving wave with equal amplitude, wavenumber, and frequency is given by the standing wave pattern

$$A \cos(k x - \omega t) + A \cos(k x + \omega t) = 2 A \cos(\omega t) \cos(k x). \quad (3.163)$$

### EXERCISE 3.2: PRESSURE FLUCTUATIONS RELATIVE TO A UNIFORM FLOW (PROBLEM 9.4 OF [Fetter and Walecka \(2003\)](#))

Consider a homogeneous and compressible fluid with uniform flow,  $\mathbf{v}$ . Show that the pressure fluctuations relative to this fluid flow state satisfy

$$(\partial_t + \mathbf{v} \cdot \nabla)^2 p - c_s^2 \nabla^2 p = 0. \quad (3.164)$$

Hint: this solution is a one-liner that results from linearizing equation (3.19) with a nonzero background flow.

### EXERCISE 3.3: MOMENTUM BUDGET FOR ACOUSTIC WAVES

Verify that the budget equation (3.116) holds for the acoustic waves, with elements of the stress-energy-momentum tensor given by equations (3.118a)–(3.118d).





## Chapter 4

# INTERFACIAL WAVES ON POTENTIAL FLOW

In this chapter we study *interfacial waves*, which are waves occurring at the interface between two homogeneous (constant density) fluid layers. We limit the study to the case where the upper layer has zero density, with this idealization motivated by the study of waves on the surface of the ocean under a massless atmosphere, referred to here as *surface waves*. We are here concerned with two restoring forces that affect small amplitude motion relative to a static equilibrium base state. One force arises from gravitation in the presence of a density jump between the two fluid layers; i.e., the buoyancy studied in Chapter ???. The other force arises from surface tension due to molecular forces at the interface (see Section ??).

To help understand the basics of surface waves along the ocean's free surface, consider a water parcel that rises above its equilibrium level into an environment where it is heavier than the surrounding atmosphere, in which case the water parcel experiences a downward buoyancy force that returns it to its equilibrium level. However, this motion generally overshoots the equilibrium level, at which point the parcel feels an upward buoyancy force. This up and down motion results in the exchange between kinetic energy and potential energy for the parcel, with the spatio-temporal organization of the oscillations constituting a *surface gravity wave*. A similar picture holds for *capillary waves* that arise from surface tension acting as the restoring force. By ignoring planetary rotation we tacitly focus on ocean surface waves whose lateral extent is too short to be affected by the planetary Coriolis acceleration. That is, we are concerned with surface waves and capillary waves that can be visually observed.

The interfacial waves studied in this chapter do not carry vorticity within the fluid interior, thus enabling use of irrotational fluid mechanics. In this case, the fluid velocity can be written as the gradient of a scalar potential, thus leading to the term *potential flow*. We also pursued the methods of potential flow in Chapter 3 when studying acoustic waves. Yet here the surface waves appear in a fluid with a strictly constant density, thus removing any acoustic modes and rendering non-divergent flow. A flow that is both irrotational and non-divergent is described by a scalar potential that satisfies Laplace's equation (Section ??); i.e., the potential is a harmonic function. Waves arise solely through the role of the boundary condition placed at the interface. Furthermore, the waves propagate in the horizontal direction along the interface and exponentially decay in the vertical, with a vertical decay scale directly related to the horizontal wavenumber. Mathematically, this coupling of the horizontal and vertical length scales is a direct result of the scalar potential satisfying Laplace's equation.

Gravity waves along an interface are transverse in the horizontal direction, meaning fluid particles move perpendicular to the wave. Furthermore, the waves do not propagate vertically and they induce vertical particle motion whose amplitude exponentially vanishes with depth in the ocean layer. Although the surface gravity waves we study in this chapter are linear, the depth decay in their amplitude leads to a net drift of fluid particles and hence to the transport

of matter. This matter transport is known as *Stokes drift*, which plays an important role in the transport of matter at the ocean surface. Stokes drift provides the canonical example of how averaging at a fixed point in space (Eulerian average) yields distinct behaviors from averaging on a fixed fluid particle (Lagrangian average).<sup>1</sup>

#### READER'S GUIDE TO THIS CHAPTER

We make use of dynamical ideas from Chapter ??, elements of the filtered equations from Chapter ??, and salient features of wave kinematics from Chapter 1. The study of capillary waves requires an understanding of surface tension in Section ?? . We also use ideas from partial differential equations introduced in Chapter ?? . The mathematical description of Stokes drift requires an understanding of Eulerian and Lagrangian kinematic descriptions from Chapter ?? . Generalizations of Stokes drift appear in Chapter ?? in our study of wave-mean flow interactions, isopycnal averaging, and the corresponding eddy-induced tracer transport. In this chapter we only consider waves at the upper surface of a single massive fluid layer, whereas in Chapter 12 we extend this analysis to the case of two massive fluid layers as part of a study of the Rayleigh-Taylor and Kelvin-Helmholtz instabilities. Finally, note that Chapter 10 in *Fetter and Walecka* (2003) and chapter 13 *Whitham* (1974) work through a number of examples, and in so doing they provide great practice in the mathematical physics of surface waves.

We use Cartesian coordinates throughout this chapter.

<b>4.1</b>	<b>Loose threads</b>	<b>97</b>
<b>4.2</b>	<b>Potential flow in a homogeneous fluid layer</b>	<b>98</b>
4.2.1	Motivating the irrotational assumption	98
4.2.2	Harmonic scalar potential	99
4.2.3	Equation of motion and Bernoulli's principle	100
4.2.4	Concerning the pressure field	101
4.2.5	Bernoulli equation of motion and boundary conditions	102
4.2.6	Local energetic balances	104
4.2.7	Expressions for the globally integrated kinetic energy	105
4.2.8	Kelvin's minimum kinetic energy theorem	107
4.2.9	Hamilton's principle and Luke's variational principle	109
<b>4.3</b>	<b>Linearized dynamics</b>	<b>112</b>
4.3.1	Linear relations between the velocity potential and pressure	112
4.3.2	Dynamic boundary condition at the free surface	112
4.3.3	Kinematic boundary conditions	113
4.3.4	Summary of the linear equations	113
<b>4.4</b>	<b>Energetics for the linearized equations</b>	<b>114</b>
4.4.1	Domain integrated kinetic energy	114
4.4.2	Domain integrated potential and available potential energies	115
4.4.3	Equipartition for the phase averaged domain integrated energies	115
4.4.4	Energetics for the depth integrated linear flow	116
<b>4.5</b>	<b>Traveling gravity waves in a flat domain</b>	<b>117</b>
4.5.1	Horizontally traveling plane wave	117
4.5.2	Domain integrated mechanical energy of a traveling wave	119
4.5.3	Dispersion relation	119

<sup>1</sup>Certain treatments of linear waves suggests that they affect a zero drift of matter. However, the Stokes drift by surface gravity waves provides an example where linear waves can transport matter. As noted in Section 4.11.5, the Lagrangian kinematics of particle trajectories introduces nonlinearities that lead to Stokes drift.

4.5.4	Alternative forms for the velocity potential and velocity . . . . .	120
4.5.5	Phase speed, group velocity, and angular frequency . . . . .	120
4.5.6	Particle trajectories ignoring Stokes drift . . . . .	122
4.5.7	Depth integrated mechanical energy of a traveling plane wave .	123
4.5.8	Further study . . . . .	125
4.6	<b>Qualitative features of deep water waves</b> . . . . .	<b>125</b>
4.7	<b>Shallow water waves approaching a shore</b> . . . . .	<b>126</b>
4.7.1	Wavenumber changes . . . . .	126
4.7.2	Wave energy and wave action . . . . .	126
4.7.3	Wave refraction . . . . .	127
4.7.4	Further study . . . . .	128
4.8	<b>Standing gravity waves in a closed basin</b> . . . . .	<b>128</b>
4.8.1	Solution for the standing waves . . . . .	129
4.8.2	Gravest seiche mode as an example . . . . .	129
4.8.3	Further study . . . . .	130
4.9	<b>Wave packets of surface gravity waves</b> . . . . .	<b>130</b>
4.9.1	Initializing the packet and dispensing with conjugate symmetry .	130
4.9.2	Expressions for the amplitude function . . . . .	131
4.9.3	Wave packet in terms of a propagator function . . . . .	132
4.9.4	Further study . . . . .	132
4.10	<b>Capillary-gravity waves</b> . . . . .	<b>132</b>
4.10.1	Pressure jump across the air-sea surface . . . . .	132
4.10.2	Dynamic boundary condition with surface tension . . . . .	133
4.10.3	Dispersion relation for capillary-gravity waves . . . . .	133
4.10.4	Deep water capillary-gravity waves . . . . .	134
4.10.5	Comments and further study . . . . .	135
4.11	<b>Particle trajectories and Stokes drift</b> . . . . .	<b>135</b>
4.11.1	Formulation of Stokes drift . . . . .	136
4.11.2	Particle trajectories in a homogeneous wave . . . . .	138
4.11.3	Stokes drift from an inhomogeneous wave . . . . .	139
4.11.4	Stokes drift for surface gravity waves . . . . .	141
4.11.5	Comments and further study . . . . .	143
4.12	<b>Exercises</b> . . . . .	<b>144</b>

---

## 4.1 Loose threads

- Waves using Luke's variational principle as per [Miles \(1977\)](#) and [Milder \(1977\)](#).
- When can we naively use Eulerian coordinates for Hamilton's principle versus the introduction of auxiliary fields? I believe potential flow is quite forgiving on these matters.
- Build on the following ideas. Much more to think about here.

Potential flow provides an example of a *scalar field theory*, which is distinctively less complex than the vector field theory encountered when the flow is less constrained. In particular, one need not worry about the Lagrangian approach, and can thus use Eulerian coordinates naively, much as done in other areas of field theory.

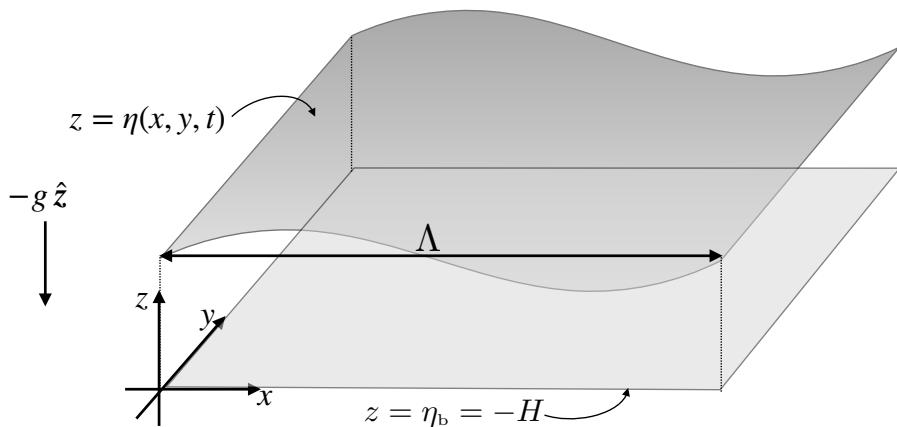


FIGURE 4.1: A depiction of an “ocean” comprised of a homogeneous (i.e., constant density) fluid with an upper free surface that is a moving material interface. The atmosphere applies a pressure to the ocean due to its mass; however, that mass is assumed to be uniform and static so that it does not affect surface motion. Linear fluctuations of the free surface exhibit gravity wave motion due to the restoring effects from a uniform gravitational field,  $\mathbf{g} = -g \hat{\mathbf{z}}$ , as well as capillary waves due to surface tension. We here depict a single wave with wavelength,  $\Lambda$ , with the wavevector parallel to the  $\hat{\mathbf{x}}$  direction. We assume the fluid layer retains a fixed volume so that the domain integral of the free surface is constant,  $\int \eta dx dy = \text{constant}$ . This property follows from the assumption that the free surface is material.

## 4.2 Potential flow in a homogeneous fluid layer

Throughout this chapter we assume the fluid is inviscid with constant density and with a flow that has zero vorticity.<sup>2</sup> These assumptions greatly simplify the expression for the velocity field, which is both non-divergent and irrotational. In this section we establish some general results for a perfect homogeneous fluid with a velocity that is non-divergent and irrotational; i.e., for *potential flow*.

We are concerned with a single layer of homogeneous fluid bounded below by a solid material surface and above by a free material surface, and refer to this layer as an “ocean”. We depict the flat bottom case in Figure 4.1, showing the material free upper ocean surface at  $z = \eta(x, y, t)$  separating a homogeneous ocean from a homogeneous atmosphere. The mass of the atmosphere is assumed to be horizontally uniform and static, even as its bottom boundary (the layer free surface) undulates. This approximation treatment means that atmospheric pressure does not contribute to motion of the ocean fluid layer. The ocean boundaries are material so that the mass of the ocean (equal to the ocean volume times the constant density,  $\rho$ ) is fixed. Consequently, the domain integral of the free surface is constant,

$$\int \eta dx dy = \int \eta dA = \text{constant}. \quad (4.1)$$

### 4.2.1 Motivating the irrotational assumption

We are familiar with the non-divergent flow assumption, which is part of the Boussinesq ocean studied in Chapter ???. However, we find it necessary to justify the assumption of irrotational flow, which also appeared in Chapter 3 when studying acoustic waves, and yet is not the case

<sup>2</sup>A more general approach can be considered in which the flow is decomposed into a potential flow (irrotational) and a vortical flow. However, for linear fluctuations these two flows are uncoupled. As we are unconcerned with vortical flow in this chapter, we set the vorticity to zero and thus focus on the potential flow.

for most other geophysical waves encountered in this book. For this purpose, consider the equation of motion for a perfect non-rotating homogeneous fluid in a gravity field

$$\rho D\mathbf{v}/Dt = -\nabla p - \rho \nabla \Phi, \quad (4.2)$$

where we assume the simple form of the geopotential (Section ??),

$$\Phi = g z, \quad (4.3)$$

with  $g$  the constant and uniform gravitational acceleration. As in our discussion of the Boussinesq ocean in Section ??, we remove the static background pressure by writing

$$p = -g \rho z + p_d = -\rho \Phi + p_d, \quad (4.4)$$

in which case the velocity equation (4.2) becomes

$$\rho D\mathbf{v}/Dt = -\nabla p_d. \quad (4.5)$$

When studying linear waves later in this chapter, this equation is linearized by dropping self-advection so that

$$\rho \partial_t \mathbf{v} = -\nabla p_d. \quad (4.6)$$

The pressure gradient cannot impart any vorticity to the velocity time tendency since  $\nabla \times \nabla p_d = 0$ . Evidently, if the velocity is initialized with zero vorticity, then the linearized equations of motion retain that zero vorticity. We thus examine linear fluctuations around a zero vorticity rest state, just as we did for the study of acoustic waves in Chapter 3.

This result offers motivation for studying properties of fluid flow with an irrotational velocity. In the remainder of this section we establish some general results for a perfect homogeneous fluid in non-rotating reference frame feeling a uniform gravity field and with an irrotational velocity. These results hold for the full nonlinear equations of motion and will later be specialized to the linear equations.

### 4.2.2 Harmonic scalar potential

A velocity field that has zero vorticity

$$\boldsymbol{\omega} = \nabla \times \mathbf{v} = 0, \quad (4.7)$$

can be written as the gradient of a scalar potential (see Section ??)<sup>3</sup>

$$\mathbf{v} = -\nabla \psi. \quad (4.8)$$

The scalar potential is unspecified up to an arbitrary function of time. The reason for this arbitrariness is that  $\psi$  and  $\psi + F(t)$  yield the same velocity field, where  $F(t)$  is any spatially constant function of time. We make use of this *gauge* degree of freedom in Section 4.2.3, just like we did for acoustic waves in Section 3.4.4.

Since the fluid is assumed to have uniform density, mass conservation in the form of the continuity equation (??) implies that the velocity field is non-divergent. Consequently, the

---

<sup>3</sup>The minus sign in equation (4.8) is conventional.

scalar potential satisfies Laplace's equation (Section ??)

$$\nabla \cdot \mathbf{v} = -\nabla \cdot \nabla \psi = -\nabla^2 \psi = 0, \quad (4.9)$$

in which we say that  $\psi$  is a *harmonic function*. We develop salient mathematical properties for harmonic functions in Section ??.

Harmonic functions do not support spatial oscillations in all three directions since the sum of the curvature in each direction (i.e., second partial derivatives) must vanish. Correspondingly, we will find that the velocity potential supports traveling waves in the horizontal and exponentially decaying in the vertical, with decay when moving away from the surface boundary. It is remarkable that the wavenumber of the horizontally traveling waves also determines the vertical decay scale. That is, the structure of the horizontal waves directly determines the vertical structure.

### 4.2.3 Equation of motion and Bernoulli's principle

To fully specify the scalar potential requires boundary conditions, which enter the development via the equation of motion. The vector-invariant equation of motion (??) for a non-rotating, irrotational, inviscid, uniform density fluid is given by

$$\partial_t \mathbf{v} = -\nabla(\Phi + \mathcal{K} + p/\rho), \quad (4.10)$$

with

$$\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2 \quad (4.11)$$

the kinetic energy per mass of a fluid element. Inserting the scalar potential,  $\mathbf{v} = -\nabla\psi$ , brings the equation of motion (4.10) to the form

$$\nabla(\Phi + \mathcal{K} + p/\rho - \partial_t \psi) = 0. \quad (4.12)$$

This equation means that everywhere in the fluid the dynamical fields satisfy

$$\Phi + \mathcal{K} + p/\rho - \partial_t \psi = C(t), \quad (4.13)$$

for some arbitrary time dependent function,  $C(t)$ . This equation is a particular expression of Bernoulli's theorem studied in Section ??.

We ascribe no physical significance to the arbitrary function,  $C(t)$ , appearing in equation (4.13). In fact, it can be completely removed by exploiting the gauge degree of freedom in the scalar potential as noted following equation (4.8). We do so by introducing a modified scalar potential

$$\Psi(\mathbf{x}, t) = \psi(\mathbf{x}, t) + \int_0^t C(t') dt'. \quad (4.14)$$

Both  $\psi$  and  $\Psi$  lead to the same velocity vector

$$\mathbf{v} = -\nabla\psi = -\nabla\Psi, \quad (4.15)$$

and as such the two scalar potentials are physically indistinguishable. However,  $\Psi$  is more convenient for our use since it absorbs the arbitrary time dependent function,  $C(t)$ , thus

rendering the simpler expression for the equation of motion

$$\partial_t \Psi = \Phi + \mathcal{K} + p/\rho. \quad (4.16)$$

In this manner we have dispensed with the need to compute  $C(t)$  since it is sufficient to work with  $\Psi$ . In the following, we refer to equation (4.16) as the *Bernoulli equation of motion*.

#### 4.2.4 Concerning the pressure field

We here explore facets of the pressure field as decomposed into either its hydrostatic and non-hydrostatic components, or its dynamically active and inactive components.

##### Two methods for decomposing pressure

In Section 4.2.1 we decomposed the pressure into its dynamically inactive component,  $-\rho g z$ , and dynamically active component,  $p_d$ . We can consider the alternative decomposition into hydrostatic and non-hydrostatic pressure components. As we will see, the dynamically active pressure is partly hydrostatic, due to motion of the free surface, and partly non-hydrostatic. We also considered these two decompositions of pressure when discussing effective buoyancy in Section ??, yet the discussion here is far simpler since the fluid has constant density.

##### Hydrostatic and non-hydrostatic pressures

The Bernoulli equation of motion (4.16) yields the vertical derivative of pressure

$$\partial_z p = -\rho \partial_z \Phi + \rho \partial_z (\partial_t \Psi - \mathcal{K}) = \partial_z p_h + \partial_z p_{nh}. \quad (4.17)$$

We here identified two contributions to the pressure. The hydrostatic pressure satisfies

$$\partial_z p_h = -\rho \partial_z \Phi = -\rho g \implies p_h = -\rho g (z - \eta) = -\rho \Phi + \rho g \eta. \quad (4.18)$$

This expression reveals that part of the hydrostatic pressure is the dynamically inactive pressure,  $-\rho \Phi = -\rho g z$ , as discussed in Section 4.2.1, plus a dynamically active portion due to undulations of the free surface. The non-hydrostatic pressure has a vertical derivative given by

$$\partial_z p_{nh} = \rho \partial_z (\partial_t \Psi - \mathcal{K}). \quad (4.19)$$

Note that plugging in the hydrostatic pressure (4.18) into the Bernoulli equation of motion (4.16) leads to

$$\partial_t \Psi = \mathcal{K} + g \eta + p_{nh}/\rho, \quad (4.20)$$

whose vertical derivative yields equation (4.19) (recall  $\eta = \eta(x, y, t)$  is depth independent).

Equation (4.19) indicates that depth variations in the kinetic energy and depth-time variations in the velocity potential lead the pressure to deviate from locally hydrostatic. We do not generally expect the flow to be hydrostatically balanced for two reasons: (i) the fluid layer has a uniform density so there is no stratification to suppress vertical accelerations that contribute to non-hydrostatic pressures; (ii) the fluid is nonrotating and so there is no vertical stiffening via the Taylor-Proudman result (Section ??), with vertical stiffening acting to suppress vertical accelerations that cause deviations from hydrostatic balance. Casual observations of surface ocean waves also supports the nontrivial vertical accelerations present in surface gravity waves, thus suggesting a key role for the non-hydrostatic pressure.

**Comments on a hydrostatic shallow water layer**

For the shallow water model we also consider a homogenous density layer. However, as emphasized in Chapter ??, the hydrostatic balance is fundamental to shallow water dynamics. Indeed, in Section ?? we see that the hydrostatic balance over a single homogeneous layer leads to horizontal motion that is depth independent throughout the layer. Hence,  $\partial_z(\partial_t \Psi) = 0$  and the kinetic energy contained in the horizontal flow is depth independent,  $\partial_z(u^2 + v^2) = 0$ . Furthermore, the vertical motion has a linear depth dependence across the shallow water layer (Section ??) and its magnitude is far smaller than horizontal motions. Therefore, we can drop all contributions to  $\partial_z \mathcal{K}$  for the shallow water layer, in which case equation (4.17) reduces to the hydrostatic limit

$$\partial_z p = -\rho \partial_z \Phi = -\rho g \quad \text{hydrostatic (shallow water) limit.} \quad (4.21)$$

This limit is relevant when the horizontal scales of motion are far larger than the vertical:  $L \gg H$ , in which case the flow is well approximated as hydrostatic.

So in summary, a homogeneous layer of fluid can have a depth dependence to its horizontal flow, and that depth dependence is driven through non-hydrostatic pressure forces. In contrast, hydrostatic pressure has a depth-independent horizontal gradient within a homogeneous layer

$$\partial_z(\nabla_h p_h) = -g \nabla_h \rho = 0 \quad \text{homogeneous density layer.} \quad (4.22)$$

Consequently, a hydrostatic pressure cannot drive depth dependence to the horizontal velocity field in a homogeneous fluid.

### 4.2.5 Bernoulli equation of motion and boundary conditions

Decompositions of the pressure discussed in Sections 4.2.1 and 4.2.4 render the following equivalent expressions for the Bernoulli equation of motion

$$\partial_t \Psi = \mathcal{K} + g z + p/\rho \quad \text{full pressure form} \quad (4.23a)$$

$$\partial_t \Psi = \mathcal{K} + p_d/\rho \quad \text{dynamically active/inactive pressure split} \quad (4.23b)$$

$$\partial_t \Psi = \mathcal{K} + g \eta + p_{nh}/\rho \quad \text{hydrostatic/non-hydrostatic pressure split.} \quad (4.23c)$$

With the velocity given by  $\mathbf{v} = -\nabla \Psi$ , we take the gradient of equations (4.23a)-(4.23c) to render the velocity equations

$$\partial_t \mathbf{v} = -\nabla(\mathcal{K} + g z + p/\rho) \quad \text{full pressure form} \quad (4.24a)$$

$$\partial_t \mathbf{v} = -\nabla(\mathcal{K} + p_d/\rho) \quad \text{dynamically active/inactive pressure split} \quad (4.24b)$$

$$\partial_t \mathbf{v} = -\nabla(\mathcal{K} + g \eta + p_{nh}/\rho) \quad \text{hydrostatic/non-hydrostatic pressure split.} \quad (4.24c)$$

A complete specification of the mathematical physics for this system requires the kinematic and dynamic boundary conditions given by the following.

**Kinematic free surface boundary condition**

The free surface is a moving material interface, so that its kinematics are described by the surface kinematic boundary condition (??) holding for a non-divergent flow at a material

boundary

$$(\partial_t + \mathbf{u} \cdot \nabla) \eta = w \quad \text{kinematic b.c. at } z = \eta(x, y, t). \quad (4.25)$$

When  $\hat{\mathbf{n}}$  has a nonzero vertical component (always assumed to hold in this chapter), the surface kinematic boundary condition can be written in the equivalent form

$$\mathbf{v} \cdot \hat{\mathbf{n}} = \frac{\partial_t \eta}{|\nabla(z - \eta)|}, \quad (4.26)$$

where the outward normal is

$$\hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} = \frac{\hat{\mathbf{z}} - \nabla\eta}{\sqrt{1 + \nabla\eta \cdot \nabla\eta}}. \quad (4.27)$$

It is furthermore convenient to make use of the relation between the area elements on the free surface given by equation (??)

$$d\mathcal{S} = |\nabla(z - \eta)| dA \quad \text{with} \quad dA = dx dy, \quad (4.28)$$

where  $dA$  is the horizontal projection of  $d\mathcal{S}$ . Consequently, we are led to the expression of the kinematic boundary condition

$$\mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \mathbf{v} \cdot \nabla(z - \eta) dA = \partial_t \eta dA. \quad (4.29)$$

### **Dynamic free surface boundary condition**

At the ocean free surface we assume zero atmospheric pressure,<sup>4</sup> in which case the ocean pressure must also vanish there. We are thus led to the following pressure boundary conditions

$$p = 0 \quad \text{and} \quad p_d = g \eta \rho \quad \text{and} \quad p_{nh} = 0 \quad \text{at } z = \eta, \quad (4.30)$$

each of which leads to the surface boundary condition for the velocity potential

$$\partial_t \Psi = \mathcal{K} + g \eta \quad \text{at } z = \eta. \quad (4.31)$$

### **Kinematic bottom boundary condition**

The bottom is a rigid material surface so that the bottom kinematic boundary condition (Section ??) is the no-normal flow condition

$$\hat{\mathbf{n}} \cdot \mathbf{v} = -\nabla\Psi \cdot \hat{\mathbf{n}} = 0 \quad \text{at } z = \eta_b(x, y). \quad (4.32)$$

When studying surface waves later in this chapter, we only consider the flat bottom ( $\hat{\mathbf{n}} = -\hat{\mathbf{z}}$ ), in which case  $w = 0$  at the bottom

$$w = -\partial_z \Psi = 0 \quad \text{at } z = \eta_b = -H. \quad (4.33)$$

<sup>4</sup>In Section 4.3.2 we show that a uniform and constant atmospheric pressure plays no role in the physics of concern here since the fluid density is itself a uniform constant.

### Dynamic bottom boundary condition

To develop the pressure boundary condition at the bottom, recall that a static material bottom has a static outward normal,  $\hat{\mathbf{n}}$ , so that

$$\hat{\mathbf{n}} \cdot \partial_t \mathbf{v} = \partial_t(\hat{\mathbf{n}} \cdot \mathbf{v}) = 0, \quad (4.34)$$

which then brings equations (4.24a)-(4.24c) to the form

$$\hat{\mathbf{n}} \cdot \nabla \mathcal{K} = -\hat{\mathbf{n}} \cdot \nabla(g z + p/\rho) \quad (4.35a)$$

$$\hat{\mathbf{n}} \cdot \nabla \mathcal{K} = -\hat{\mathbf{n}} \cdot \nabla p_d/\rho \quad (4.35b)$$

$$\hat{\mathbf{n}} \cdot \nabla \mathcal{K} = -\hat{\mathbf{n}} \cdot \nabla(g \eta + p_{nh}/\rho). \quad (4.35c)$$

Specializing to the flat bottom with  $\hat{\mathbf{n}} = -\hat{\mathbf{z}}$  yields the bottom boundary conditions

$$\partial_z \mathcal{K} = -g - \rho^{-1} \partial_z p \quad (4.36a)$$

$$\partial_z \mathcal{K} = -\rho^{-1} \partial_z p_d \quad (4.36b)$$

$$\partial_z \mathcal{K} = -\rho^{-1} \partial_z p_{nh}, \quad (4.36c)$$

where  $\partial_z \eta = 0$ . In the linearized theory,  $\partial_z \mathcal{K}$  is far smaller than  $g$  or  $\rho^{-1} \partial_z p$ , which means that the balance in equation (4.36a) must be hydrostatic,  $\partial_z p = -\rho g$ . This result then means that the non-hydrostatic boundary condition is  $\partial_z p_{nh} = 0$ . Likewise, with  $p = -\rho g z + p_d$ , and  $\partial_z p = -\rho g$  at the bottom, we must have  $\partial_z p_{nh} = 0$ . These results are thus summarized as

$$\partial_z p = -g \rho \quad \text{and} \quad \partial_z p_d = 0 \quad \text{and} \quad \partial_z p_{nh} = 0 \quad \text{at } z = \eta_b = -H. \quad (4.37)$$

### Comments on the need for two boundary conditions

Mathematically, the solution to Laplace's equation requires only a single boundary condition when posed in a domain with specified boundaries (e.g., see Section ??). For the current setup, however, the free surface is a moving boundary, thus offering one more dynamical degree of freedom that necessitates an extra boundary condition. Physically, there are two dynamical fields that describe the fluid layer, the velocity potential,  $\Psi$ , and the free surface height,  $\eta$ . Consequently, there are two boundary conditions that arise when specifying these fields: one from kinematics (the boundary interfaces are material) and one from dynamics (forces on both sides of the boundary interfaces must balance as per Newton's third law).

### 4.2.6 Local energetic balances

Since the fluid has a constant density, the only energy arises from mechanical energy due to motion (kinetic energy) plus the gravity field (gravitational potential energy)

$$\mathcal{E} = \mathcal{K} + \Phi = \mathbf{v} \cdot \mathbf{v}/2 + g z. \quad (4.38)$$

The internal energy is a constant and so plays no role in the energetic analysis. Also, we ignore dissipation and heat transfer so that energy is modified only through reversible processes. Local energetic budget equations are readily computed by taking the scalar product of the velocity with the velocity tendency. We consider here the three forms of the velocity equation (4.24a)-(4.23c) and their corresponding energy equations.

**Velocity equation with unsplit pressure**

With  $\partial_t \mathbf{v}$  written in the form of equation (4.24a) we have

$$\partial_t \mathcal{K} = -\mathbf{v} \cdot \nabla (\mathcal{K} + \Phi + p/\rho) = -\nabla \cdot [\mathbf{v} (\mathcal{K} + \Phi + p/\rho)], \quad (4.39)$$

where we used  $\nabla \cdot \mathbf{v} = 0$  for the second equality. Since the geopotential has a zero Eulerian time derivative, the kinetic energy equation readily leads to the total mechanical energy equation

$$\partial_t \mathcal{E} = -\nabla \cdot [\mathbf{v} (\mathcal{E} + p/\rho)] \implies \rho D\mathcal{E}/Dt = -\nabla \cdot (\mathbf{v} p). \quad (4.40)$$

Evidently, convergence of the pressure flux,  $\mathbf{v} p$ , leads to a material time change in the total energy.

**Dynamically active/inactive pressure split**

An equivalent form of the energy equation can be found by making use of the pressure split into dynamically active and inactive components according to the velocity equation (4.24b), which yields

$$\partial_t \mathcal{K} = -\nabla \cdot [\mathbf{v} (\mathcal{K} + p_d/\rho)] \implies \rho D\mathcal{K}/Dt = -\nabla \cdot (\mathbf{v} p_d). \quad (4.41)$$

We here find that convergence of the dynamic pressure flux,  $\mathbf{v} p_d$ , leads to a material time change in the kinetic energy.

**Hydrostatic/non-hydrostatic pressure split**

We now develop the energetics with the hydrostatic/non-hydrostatic pressure decomposition (4.23c) to render

$$\partial_t \mathcal{K} = -\nabla \cdot [\mathbf{v} (\mathcal{K} + g\eta + p_{nh}/\rho)]. \quad (4.42)$$

This form only lends itself to a total energy budget written in the form

$$\partial_t \mathcal{E} = -\nabla \cdot [\mathbf{v} (\mathcal{E} + p/\rho)] \implies \rho D\mathcal{E}/Dt = -\nabla \cdot [\mathbf{v} (p_h + p_{nh})]. \quad (4.43)$$

### 4.2.7 Expressions for the globally integrated kinetic energy

The kinetic energy per mass can be written in terms of the scalar potential

$$2\mathcal{K} = \mathbf{v} \cdot \mathbf{v} = \nabla \Psi \cdot \nabla \Psi = \nabla \cdot (\Psi \nabla \Psi), \quad (4.44)$$

where we used  $\nabla^2 \Psi = 0$  for the final equality. This divergence form of the kinetic energy means that its global integral is fully determined by properties at the boundaries.

**Domain integrated kinetic energy is just due to surface properties**

Integration of the kinetic energy in the form of equation (4.44) over the full ocean domain,  $\mathcal{R}$ , leads to the total kinetic energy

$$E_{KE} = \int_{\mathcal{R}} \rho \mathcal{K} dV = \frac{\rho}{2} \int_{\mathcal{R}} \nabla \cdot (\Psi \nabla \Psi) dV = \frac{\rho}{2} \int_{\partial\mathcal{R}} \Psi \nabla \Psi \cdot \hat{\mathbf{n}} d\mathcal{S}, \quad (4.45)$$

where we made use of the divergence theorem. The ocean bottom is material so that the flow satisfies the no-normal flow condition at each point along the bottom (Section ??)

$$\mathbf{v} \cdot \hat{\mathbf{n}} = -\nabla\Psi \cdot \hat{\mathbf{n}} = 0, \quad (4.46)$$

where  $\hat{\mathbf{n}}$  is the outward normal on the rigid material boundaries. Lateral boundaries are either periodic or rigid material walls. We thus find the remarkable result that the domain integrated kinetic energy arises solely from properties integrated over the free surface

$$E_{\text{KE}} = \frac{\rho}{2} \int_{z=\eta} \Psi \nabla\Psi \cdot \hat{\mathbf{n}} d\mathcal{S} = -\frac{\rho}{2} \int_{z=\eta} \Psi \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S}. \quad (4.47)$$

Evidently, contributions from interior motion play no role in the domain integrated kinetic energy. Furthermore, the kinetic energy is a non-negative number, so that the right hand side of equation (4.47) is non-negative although it is not obvious without noting that  $\Psi$  is a harmonic function in the domain interior. That is, the nature of the kinetic energy is fundamentally related to the harmonic nature of the velocity potential that allows the kinetic energy to be written as the total divergence in equation (4.44). We encountered a similar feature of harmonic functions in Section ?? when studying their mean-value property.

### Gauge invariance of the domain integrated kinetic energy

The kinetic energy per mass,  $\mathcal{K} = \nabla\Psi \cdot \nabla\Psi / 2$ , is manifestly gauge invariant since it remains unchanged if  $\Psi$  is shifted by a spatial constant. This property also holds for the domain integrated kinetic energy, (4.47), as follows from use of the divergence theorem and then the non-divergence of the velocity field

$$\int_{z=\eta} \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\partial\mathcal{R}} \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \int_{\mathcal{R}} \nabla \cdot \mathbf{v} dV = 0. \quad (4.48)$$

Evidently, if the scalar potential is shifted by a function of time that is spatially independent, then the globally integrated kinetic energy (4.47) remains unchanged and so is gauge invariant. An alternative derivation makes use of the surface kinematic boundary condition (4.29) that leads to the identity

$$\int_{z=\eta} \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \int \partial_t \eta dA = - \int \nabla \cdot \mathbf{U} dA = 0, \quad (4.49)$$

where we set  $\partial_t \eta = -\nabla \cdot \mathbf{U}$  as per the free surface equation (??) holding for a volume conserving fluid, with  $\mathbf{U} = \int_{-H}^{\eta} \mathbf{u} dz$  the depth integrated horizontal velocity.<sup>5</sup>

### Kinetic energy in terms of time tendencies

We write yet another form for the domain integrated kinetic energy, which proves of use when studying Hamilton's principle in Section 4.2.9. For this purpose, make use of the free surface kinematic boundary condition (4.29) to bring the kinetic energy equation (4.47) into the form

$$E_{\text{KE}} = -\frac{\rho}{2} \int_{z=\eta} \Psi \mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = -\frac{\rho}{2} \int_{z=\eta} \Psi \partial_t \eta dA. \quad (4.50)$$

---

<sup>5</sup>See also Exercise ??.

Now reintroduce the vertical integral via Leibniz's rule (Section ??) to write

$$\int_{z=\eta} \Psi \partial_t \eta \, dA = \int \left[ \frac{\partial}{\partial t} \int_{\eta_b}^{\eta} \Psi \, dz \right] \, dA - \int \left[ \int_{\eta_b}^{\eta} \partial_t \Psi \, dz \right] \, dA. \quad (4.51)$$

If the horizontal extent of the domain has a static extent, such as when the horizontal domain is periodic or it is bounded by rigid vertical walls (e.g., see Figure ??), then the time derivative commutes with the area integral so that equation (4.51) becomes

$$\int_{z=\eta} \Psi \partial_t \eta \, dA = \frac{d}{dt} \left[ \int \int_{\eta_b}^{\eta} \Psi \, dz \, dA \right] - \int \int_{\eta_b}^{\eta} \partial_t \Psi \, dz \, dA = \frac{d}{dt} \int \Psi \, dV - \int \partial_t \Psi \, dV. \quad (4.52)$$

We make use of this identity in the discussion of Hamilton's principle in Section 4.2.9.

#### 4.2.8 Kelvin's minimum kinetic energy theorem

The domain integrated kinetic energy equation (4.47) points to the central role of boundary conditions. We here expose a property of the kinetic energy associated with *Kelvin's minimum kinetic energy theorem*.

##### Basic formulation

Following Kelvin (as detailed in Section 45 of [Lamb \(1993\)](#)), consider a non-divergent flow,  $\mathbf{v}^{\text{tot}}$ , built from the sum of an irrotational and non-divergent velocity field (i.e., potential flow),  $\mathbf{v}$ , plus a non-divergent and rotational flow,  $\mathbf{v}_r$

$$\mathbf{v}_{\text{tot}} = \mathbf{v} + \mathbf{v}_r \quad \text{with} \quad \nabla \cdot \mathbf{v} = \nabla \cdot \mathbf{v}_r = 0 \quad \text{and} \quad \nabla \times \mathbf{v}_{\text{tot}} = \nabla \times \mathbf{v}_r. \quad (4.53)$$

The domain integrated kinetic energy of this flow,

$$E_{\text{KE}}^{\text{tot}} = \frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v}_{\text{tot}} \cdot \mathbf{v}_{\text{tot}} \, dV, \quad (4.54)$$

has three terms

$$E_{\text{KE}}^{\text{tot}} = \frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v} \cdot \mathbf{v} \, dV + \frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v}_r \cdot \mathbf{v}_r \, dV + \rho \int_{\mathcal{R}} \mathbf{v} \cdot \mathbf{v}_r \, dV. \quad (4.55)$$

The domain integrated kinetic energy for the irrotational flow equals to the boundary integral from equation (4.47)

$$\frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v} \cdot \mathbf{v} \, dV = -\frac{\rho}{2} \int_{\partial\mathcal{R}} \Psi \mathbf{v} \cdot \hat{\mathbf{n}} \, dS. \quad (4.56)$$

The cross-term also takes the form of a boundary integral

$$\int_{\mathcal{R}} \mathbf{v} \cdot \mathbf{v}_r \, dV = - \int_{\mathcal{R}} \nabla \Psi \cdot \mathbf{v}_r \, dV = - \int_{\mathcal{R}} \nabla \cdot (\Psi \mathbf{v}_r) \, dV = - \int_{\partial\mathcal{R}} \Psi \mathbf{v}_r \cdot \hat{\mathbf{n}} \, dS, \quad (4.57)$$

so that the domain integrated kinetic energy is

$$E_{\text{KE}}^{\text{tot}} = \frac{\rho}{2} \int_{\mathcal{R}} (\mathbf{v}_r \cdot \mathbf{v}_r + \mathbf{v} \cdot \mathbf{v}) \, dV - \rho \int_{\partial\mathcal{R}} \Psi \mathbf{v}_r \cdot \hat{\mathbf{n}} \, dS \quad (4.58a)$$

$$= \frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v}_r \cdot \mathbf{v}_r \, dV - \frac{\rho}{2} \int_{\partial\mathcal{R}} \Psi (\mathbf{v} + 2\mathbf{v}_r) \cdot \hat{\mathbf{n}} \, dS. \quad (4.58b)$$

**Kelvin's minimum kinetic energy theorem**

Consider the case where the flow is irrotational and so it satisfies the boundary conditions appropriate for potential flow in a homogeneous fluid layer with a free surface

$$\mathbf{v} \cdot \hat{\mathbf{n}} = 0 \quad \text{at solid bottom boundary, } z = \eta_b(x, y) \quad (4.59a)$$

$$\mathbf{v} \cdot \hat{\mathbf{n}} d\mathcal{S} = \partial_t \eta dA \quad \text{at free surface boundary, } z = \eta(x, y, t). \quad (4.59b)$$

The free surface condition (4.59b) arises from equation (4.26) as well as the area relation in equation (4.28). Now add a rotational flow, and assume the rotational flow leaves the potential flow's boundary conditions (4.59a) and (4.59b) unaffected, which can be ensured if the rotational flow satisfies the no-normal flow condition at all boundaries

$$\mathbf{v}_r \cdot \hat{\mathbf{n}} = 0 \quad \mathbf{x} \in \partial\mathcal{R}. \quad (4.60)$$

Kelvin's minimum kinetic energy theorem follows from equation (4.58a) with  $\mathbf{v}_r \cdot \hat{\mathbf{n}} = 0$ , in which case

$$E_{KE}^{\text{tot}} = \frac{\rho}{2} \int_{\mathcal{R}} (\mathbf{v} \cdot \mathbf{v} + \mathbf{v}_r \cdot \mathbf{v}_r) dV, \quad (4.61)$$

so that kinetic energy of the pure irrotational flow is less than that for the flow based on the same irrotational flow plus a rotational perturbation. Evidently, the irrotational flow minimizes the kinetic energy for the simply connected material domain.

In formulating the theorem (4.61), we considered the rotational component to the flow as a perturbation to the original irrotational flow, with the perturbation not altering the boundary conditions satisfied by the irrotational flow. In this manner, the rotational flow is akin to a variation added to the irrotational flow in the sense used for Hamilton's principle (e.g., Sections ?? and 4.2.9), in which variations do not touch boundary (or initial) conditions.

**Material boundary condition for  $\mathbf{v}_r + \mathbf{v}$** 

Now consider the material boundary conditions at the bottom and surface to be satisfied by the full flow

$$(\mathbf{v} + \mathbf{v}_r) \cdot \hat{\mathbf{n}} = 0 \quad \text{at solid bottom boundary, } z = \eta_b(x, y) \quad (4.62a)$$

$$(\mathbf{v} + \mathbf{v}_r) \cdot \hat{\mathbf{n}} d\mathcal{S} = \partial_t \eta dA \quad \text{at free surface boundary, } z = \eta(x, y, t), \quad (4.62b)$$

so that these boundary conditions couple the rotational and irrotational components. The kinetic energy equation (4.58b) now takes the form

$$E_{KE}^{\text{tot}} = \frac{\rho}{2} \int_{\mathcal{R}} \mathbf{v}_r \cdot \mathbf{v}_r dV - \frac{\rho}{2} \int_{z=\eta} \Psi \partial_t \eta dA - \frac{\rho}{2} \int_{\partial\mathcal{R}} \Psi \mathbf{v}_r \cdot \hat{\mathbf{n}} d\mathcal{S} \quad (4.63a)$$

$$= \frac{\rho}{2} \int_{\mathcal{R}} (\mathbf{v}_r \cdot \mathbf{v}_r - \mathbf{v} \cdot \mathbf{v}) dV - \rho \int_{z=\eta} \Psi \partial_t \eta dA \quad (4.63b)$$

For a rigid material upper boundary, so that  $\partial_t \eta = 0$ , the domain integrated kinetic energy in the irrotational flow is less than that in the rotational flow. Any further general statements are unavailable for the case with a time dependent free surface. The absence of a general theorem can be traced to coupling between the rotational and irrotational flow as realized via the boundary conditions. This situation contrasts to the case of Kelvin's minimum kinetic energy theorem, where there we purposely disabled any coupling.

### 4.2.9 Hamilton's principle and Luke's variational principle

Following the discussion of classical field theory in Chapter ??, we here consider the Lagrangian density functional and the corresponding equations of motion that follow from *Hamilton's principle*. The special nature of the dynamical free surface boundary condition prompts a modified version of the action that allows for a treatment of both the interior potential flow and the nonlinear surface boundary conditions. This modification is due to [Luke \(1967\)](#), prompting the name *Luke's variational principle*.<sup>6</sup>

#### Conventional form of the Lagrangian density

The Lagrangian density (dimensions of energy per unit volume) is the kinetic energy per volume minus the gravitational potential energy per volume, which takes the following form for potential flow of a homogeneous fluid layer

$$\mathcal{L} = \rho (\mathbf{v} \cdot \mathbf{v}/2 - g z) = \rho (\nabla \Psi \cdot \nabla \Psi - g z)/2. \quad (4.64)$$

The corresponding action is

$$\mathcal{S} = \int_R \mathcal{L} dV dt, \quad (4.65)$$

where  $R$  is the space-time domain. Hamilton's principle says that variation of the action vanishes for the physically realized fields

$$\delta \mathcal{S} = \int_R \delta \mathcal{L} dV dt = 0, \quad (4.66)$$

where the variation operator,  $\delta$ , does not touch space or time and so it commutes with integrals and derivatives (see Sections ?? and ??).

The puzzle we have, however, is that the free surface does not explicitly appear in the Lagrangian (4.64). Rather, it only appears as a boundary in the action integral (4.66). That is not a problem *per se*, but the problem is that varying the action using the Lagrangian (4.64) fails to produce the dynamical boundary condition at the free surface. To resolve this puzzle we follow [Luke \(1967\)](#) by transforming the action into a modified form that differs from the original action by a time derivative which, as shown in Section ??, does not alter the mechanics. Furthermore, the appropriate action turns out to be, quite remarkably, the space-time integral of the pressure.

#### Surface Lagrangian density

Before considering the approach of [Luke \(1967\)](#), observe that the Lagrangian density (4.64) can be written

$$\rho^{-1} \int_{\mathcal{R}} \mathcal{L} dV = \int_{\mathcal{R}} (\mathbf{v} \cdot \mathbf{v}/2 - g z) dV = -\frac{1}{2} \int [\Psi_s \partial_t \eta + g (\eta^2 - \eta_b^2)] dA, \quad (4.67)$$

where we made use of the kinematic boundary conditions (4.59a) and (4.59b), in which case the action is

$$\mathcal{S} = -\frac{\rho}{2} \int [\Psi_s \partial_t \eta + g (\eta^2 - \eta_b^2)] dA dt \equiv \int \mathcal{L}_s dA dt. \quad (4.68)$$

<sup>6</sup>Elements of this section follow from [Luke \(1967\)](#), Section 13.2 of [Whitham \(1974\)](#), and Exercise 10.14 of [Fetter and Walecka \(2003\)](#). See also [Miles \(1977\)](#) and [Milder \(1977\)](#).

In this equation we introduced the velocity potential evaluated at the free surface

$$\Psi_s = \Psi(x, y, z = \eta, t), \quad (4.69)$$

and the surface Lagrangian density (dimensions of energy per unit area)

$$\mathcal{L}_s = -\rho[\Psi_s \partial_t \eta + g(\eta^2 - \eta_b^2)]/2. \quad (4.70)$$

This Lagrangian forms the starting point for [Milder \(1977\)](#).

### Transforming the action into the integral of pressure

We transform the action through use of equation (4.52) for the kinetic energy, in which

$$\int \left[ \int \nabla \Psi \cdot \nabla \Psi dV \right] dt = - \int \left[ \int [\Psi \partial_t \eta]_{z=\eta} dA \right] dt \quad (4.71a)$$

$$= - \int \left[ \frac{\partial}{\partial t} \int \Psi dV \right] dt + \int \partial_t \Psi dV dt. \quad (4.71b)$$

As noted in deriving equation (4.52), we assumed here that the horizontal bounds for the domain are static. For the more general case of sloping side boundaries (e.g., see Figure ??), then we need to introduce yet another dynamical field, namely the moving horizontal bounds. That added dynamical degree of freedom is not the focus here, so that we assume the domain has static horizontal boundaries, thus allowing the time derivative to commute with the area integral.

Use of equation (4.71b) brings the action to the form

$$S = -\frac{\rho}{2} \int \left[ \frac{\partial}{\partial t} \int \Psi dV \right] dt + \rho \int \left[ \frac{1}{2} \frac{\partial \Psi}{\partial t} - gz \right] dV dt. \quad (4.72)$$

Following the discussion in Section ??, the first term on the right hand side is mechanically irrelevant since it evaluates to the time bounds, during which the velocity potential has zero variation. We can thus drop the first right hand side term in equation (4.72) to focus on the modified action

$$S^{\text{mod}} = \rho \int \left[ \frac{1}{2} \frac{\partial \Psi}{\partial t} - gz \right] dV dt. \quad (4.73)$$

Making use again of equation (4.52) along with the mechanical equivalence (4.72) gives

$$S^{\text{Luke}} = \rho \int \left[ \frac{\partial \Psi}{\partial t} - \frac{1}{2} \frac{\partial \Psi}{\partial t} - gz \right] dV dt = \rho \int \left[ \frac{\partial \Psi}{\partial t} - \frac{1}{2} \nabla \Psi \cdot \nabla \Psi - gz \right] dV dt. \quad (4.74)$$

### Variation of the action with respect to $\eta$

Exposing the vertical integration limits on the action (4.74) yields

$$S^{\text{Luke}} = \rho \int \int \int_{\eta_b}^{\eta} \left[ \frac{\partial \Psi}{\partial t} - \frac{1}{2} \nabla \Psi \cdot \nabla \Psi - gz \right] dz dA dt, \quad (4.75)$$

so that the variation arising from  $\delta\eta$  is straightforward to compute. In so doing we find

$$\frac{\delta S^{\text{Luke}}}{\delta \eta} = 0 \implies \partial_t \Psi = \frac{1}{2} \nabla \Psi \cdot \nabla \Psi + g\eta \quad \text{at } z = \eta, \quad (4.76)$$

which is the dynamic boundary condition (4.31).

### Variation of the action with respect to $\Psi$

Variation of the action with respect to the velocity potential leads to

$$\delta_\Psi \mathcal{S}^{\text{Luke}} = \frac{\rho}{2} \int \partial_t(\delta\Psi) dV dt - \rho \int \delta\Psi \nabla\Psi \cdot \hat{\mathbf{n}} d\mathcal{S} dt + \rho \int \nabla^2\Psi \delta\Psi dV dt, \quad (4.77)$$

where we noted that  $\delta(\partial_t\Psi) = \partial_t(\delta\Psi)$ , and integrated by parts to get the second and third terms on the right hand side. For the time derivative term we write

$$\int \partial_t(\delta\Psi) dV = \int \left[ \int_{\eta_b}^\eta \partial_t(\delta\Psi) dz \right] dA = \int \left[ \frac{\partial}{\partial t} \int_{\eta_b}^\eta \delta\Psi dz - \partial_t \eta \delta\Psi_s \right] dA, \quad (4.78)$$

whereas the kinematic boundary conditions (equations (4.59a) and (4.59b)) bring the spatial boundary term to

$$- \int \delta\Psi \nabla\Psi \cdot \hat{\mathbf{n}} d\mathcal{S} = \int \partial_t \eta \delta\Psi_s dA, \quad (4.79)$$

thus leading to the action variation

$$\delta_\Psi \mathcal{S}^{\text{Luke}} = \int \nabla^2\Psi \delta\Psi dV dt. \quad (4.80)$$

The action variation vanishes so long as  $\nabla^2\Psi = 0$  within the fluid domain. Bringing everything together leads to the boundary value problem for the velocity potential and the free surface

$$\nabla^2\Psi = 0 \quad \mathbf{x} \in \mathcal{R} \quad (4.81a)$$

$$\hat{\mathbf{n}} \cdot \nabla\Psi = 0 \quad \mathbf{x} \in \partial\mathcal{R} \text{ kinematic rigid condition} \quad (4.81b)$$

$$(\nabla\eta - \hat{\mathbf{z}}) \cdot \nabla\Psi = \partial_t\eta \quad z = \eta \text{ kinematic free surface condition} \quad (4.81c)$$

$$\partial_t\Psi = \frac{1}{2} \nabla\Psi \cdot \nabla\Psi + g\eta \quad z = \eta \text{ dynamic free surface condition.} \quad (4.81d)$$

### Comments

Remarkably, the Bernoulli equation of motion (4.16) means that the Lagrangian functional appearing in the action (4.74) equals to the pressure

$$\mathcal{S}^{\text{Luke}} = \rho \int \left[ \frac{\partial\Psi}{\partial t} - \frac{1}{2} \nabla\Psi \cdot \nabla\Psi - gz \right] dV dt = \int p dV dt. \quad (4.82)$$

Starting from this form of the action, Hamilton's principle extremizes the space-time integral of the pressure. *Seliger and Whitham* (1968) provide some discussion of pressure as the action within the context of the Boussinesq approximation, where the same result holds.

A free surface certainly adds subtleties to the use of Hamilton's principle. The presentation given in this subsection proceeds in the opposite order to that given by *Luke* (1967) as well as section 13.2 of *Whitham* (1974). Here, we started with the conventional form of the Lagrangian (4.64) written as the difference between kinetic energy and potential energy, and then showed how to transform the action into the pressure action integral (4.82) that captures the dynamical boundary condition. In that transformation we made use of the kinematic boundary conditions at both the static bottom boundary (equation (4.59a)) and the free surface (equation (4.59b)).

The alternative approach taken by [Luke \(1967\)](#) starts from the inspired guess of a pressure based action integral (4.82), and shows that it indeed produces the proper dynamical equations so long as the kinematic boundary conditions are satisfied. The present approach is thus offered as a complement to that from [Luke \(1967\)](#) and [Whitham \(1974\)](#)

## 4.3 Linearized dynamics

We here develop the boundary value problem describing linear surface gravity wave motions of the free surface, and characterize physical aspects of the waves. Surface tension is ignored so that pressure is continuous across the free surface. In Section 4.10 we remove this assumption by considering the pressure jump at the ocean surface due to surface tension, with this pressure jump leading to capillary waves. The fundamental parameter measuring nonlinearity concerns the ratio of the free surface undulation to the length of a wave disturbance. A small value for this non-dimensional ratio allows us to confidently make use of the linear equations.

Note that we focus on the velocity potential in our analysis of linear surface gravity waves. A directly analogous approach focuses on the dynamic pressure,  $p_d$ , which is a harmonic function for the linearized system (take the divergence of the linearized velocity equation (4.6)). Section 7.1 of [Vallis \(2017\)](#) takes the pressure approach. The two methods are equivalent, with the velocity potential and pressure closely connected.

### 4.3.1 Linear relations between the velocity potential and pressure

The Bernoulli equations of motion (4.23a)-(4.23c) provide an expression for the time tendency of the velocity potential. When linearizing the equations we drop the contribution from the kinetic energy since it is second order in the velocity field.<sup>7</sup> We are thus led to the linear relations

$$\partial_t \Psi = g z + p/\rho = p_d/\rho = g \eta + p_{nh}/\rho. \quad (4.83)$$

This equation provides the linearized relation between the velocity potential and the various pressure fields.

### 4.3.2 Dynamic boundary condition at the free surface

The equation of motion (4.16) applies to any point within the fluid and at any time. In particular, it applies at the free surface,  $z = \eta(x, y, t)$ , where pressure equals to the atmospheric pressure. As stated earlier, we assume that the atmospheric pressure is constant in space and time so that

$$g \eta + \mathcal{K} - \partial_t \Psi = -p_a/\rho = \text{constant}. \quad (4.84)$$

Without loss of generality we can set this constant to zero,<sup>8</sup> thus leaving the boundary condition

$$g \eta + \mathcal{K} - \partial_t \Psi = 0 \quad \text{linearized dynamic b.c. at } z = \eta. \quad (4.85)$$

We now linearize relative to a state of rest with  $\eta = 0$ ,  $\mathbf{v} = 0$ , and  $\partial_t \Psi = 0$ . Linear fluctuations about this rest state have small velocities. Consequently, the kinetic energy, which is second order in velocity, is small relative to the remaining terms and so we arrive at the

<sup>7</sup>We do *not* drop kinetic energy when studying the energetics of linear waves. But here we are studying their momentum, in which we linearize by dropping all second and higher order terms.

<sup>8</sup>Alternatively, can use a gauge transformation  $\Psi' = \Psi - t(p_a/\rho)$  to eliminate the constant.

linearized dynamic boundary condition

$$g\eta = \partial_t \Psi \quad \text{linearized dynamic b.c. at } z = \eta. \quad (4.86)$$

This boundary condition directly connects the free surface to time tendencies of the velocity potential. The free surface fluctuates upward when the velocity potential has a positive tendency, and vice versa. Also note that we arrive at this boundary condition by making use of the pressure boundary conditions (4.30) within equation (4.83).

### 4.3.3 Kinematic boundary conditions

The free surface is assumed to be a material interface, meaning that there is no matter transported across this surface. Consequently, following the discussion of kinematic boundary conditions in Section ??, we have

$$(\partial_t + \mathbf{u} \cdot \nabla)\eta = w \quad \text{kinematic b.c. at } z = \eta. \quad (4.87)$$

Linearizing this boundary condition about the state of rest, and introducing the scalar potential, leads to

$$\partial_t \eta = w = -\partial_z \Psi \quad \text{linearized kinematic b.c. at } z = \eta. \quad (4.88)$$

This is yet another constraint that links the free surface to the velocity potential.

### 4.3.4 Summary of the linear equations

The boundary value problem for the velocity potential and free surface is given by

$$\mathbf{v} = -\nabla \Psi \quad \text{velocity potential} \quad (4.89a)$$

$$\nabla^2 \Psi = 0 \quad \text{irrotational and non-divergent velocity for } \mathbf{x} \in \text{ocean} \quad (4.89b)$$

$$\partial_t \Psi = g\eta \quad \text{linearized dynamic b.c. at } z = \eta \quad (4.89c)$$

$$\partial_z \Psi = -\partial_t \eta \quad \text{linearized kinematic b.c. at } z = \eta \quad (4.89d)$$

$$\hat{\mathbf{n}} \cdot \nabla \Psi = 0 \quad \text{no-normal flow kinematic b.c. on rigid boundaries.} \quad (4.89e)$$

Equations (4.89a) and (4.89b) hold throughout the fluid whereas the remaining equations hold only at the boundaries.

Although the equations (4.89b)-(4.89e) were derived through linearization, there is one remaining nonlinearity that needs to be removed to enable a fully linear analytical treatment. Namely, when combining the boundary conditions into a single equation we compute the time derivative of equation (4.89c) according to

$$g \frac{\partial \eta}{\partial t} = \left[ \frac{\partial}{\partial t} + \frac{\partial \eta}{\partial t} \frac{\partial}{\partial z} \right] \frac{\partial \Psi}{\partial t}, \quad (4.90)$$

which follows since  $\Psi = \Psi(x, y, z = \eta(x, y, t), t)$  at the surface boundary. Combining with equation (4.89d) renders

$$\left[ \frac{\partial}{\partial t} + \frac{\partial \eta}{\partial t} \frac{\partial}{\partial z} \right] \frac{\partial \Psi}{\partial t} = -g \frac{\partial \Psi}{\partial z}. \quad (4.91)$$

With  $w \approx \partial \eta / \partial t$  at the free surface, we identify  $(\partial \eta / \partial t) \partial_z$  as a vertical advection operator. The corresponding term  $(\partial \eta / \partial t) \partial_{zt} \Psi$  is nonlinear and second order in fluctuating fields. Hence,

we drop this term as part of the linearization process. An equivalent means to realize this linearization is to evaluate the free surface boundary condition at  $z = 0$  rather than at  $z = \eta(x, y, t)$ . For this approximation to be self-consistent requires the amplitude of free surface undulations to be much smaller than the typical wavelengths ( $\Lambda$ ) of the fluctuations, in which case

$$|\eta|/\Lambda \ll 1 \iff |\eta| |\mathbf{k}| \ll 1, \quad (4.92)$$

where  $|\mathbf{k}| = 2\pi/\Lambda$  is the wavenumber for the waves. The condition (4.92) holds for the waves considered here.

In summary, the fully linearized equation set takes the form

$$\mathbf{v} = -\nabla\Psi \quad \text{velocity potential} \quad (4.93a)$$

$$\nabla^2\Psi = 0 \quad \text{irrotational and non-divergent velocity for } \mathbf{x} \in \text{ocean} \quad (4.93b)$$

$$\partial_t\Psi = g\eta \quad \text{linearized dynamic b.c. at } z = 0 \quad (4.93c)$$

$$\partial_z\Psi = -\partial_t\eta \quad \text{linearized kinematic b.c. at } z = 0 \quad (4.93d)$$

$$\hat{\mathbf{n}} \cdot \nabla\Psi = 0 \quad \text{no-normal flow kinematic b.c. on rigid boundaries.} \quad (4.93e)$$

Observe that these equations for surface gravity waves involve a harmonic scalar potential,  $\Psi$ , defined throughout the full fluid domain (equation (4.93b)), whereas the time tendencies are determined by the kinematic boundary condition (equation (4.93d)) and dynamic boundary condition (equation (4.93c)). Mathematically, the free surface,  $\eta$ , lives on a two-dimensional manifold of the surface interface, whereas the velocity potential,  $\Psi$ , lives on a three-dimensional manifold defined by the ocean domain. The velocity potential and free surface are coupled by the kinematic and dynamic boundary conditions, thus requiring the velocity potential to be determined throughout the three-dimensional ocean domain even if we might only care about fluctuations of the free surface. These features of surface gravity waves make them inherently more complex, and rich, than surface fluctuations of a membrane<sup>9</sup>, or the volume fluctuations of a compressible fluid leading to acoustic waves (Chapter 3). The surface gravity wave system provides a canonical example of a surface boundary dynamical system, with surface quasi-geostrophy another prominent geophysical example.<sup>10</sup>

## 4.4 Energetics for the linearized equations

We here specialize the energetic analysis from Section 4.2.6 to study the energetics of the linearized equations in a manner analogous to the acoustic wave energetics considered in Section 3.6.

### 4.4.1 Domain integrated kinetic energy

The surface integral (4.50) for the domain integrated kinetic energy is computed at  $z = \eta(x, y, t)$ . For the linearized system, this integral is approximated at  $z = 0$ , so that the domain integrated

<sup>9</sup>A drumhead provides the canonical example of a vibrating membrane. For determining waves moving on a drumhead, we focus exclusively on the drumhead without also considering dynamics of the surrounding fluid. See Chapter 8 of [Fetter and Walecka \(2003\)](#) for details.

<sup>10</sup>See [Held et al. \(1995\)](#) for a classic treatment of surface quasi-geostrophy. [Yassin \(2021\)](#) and [Yassin and Griffies \(2022\)](#) provide further studies of surface quasi-geostrophy in the context of normal mode theory.

kinetic energy is

$$E_{\text{KE}} = -\frac{\rho}{2} \int [\Psi \partial_t \eta]_{z=0} dA = \frac{\rho}{2} \int [\Psi \partial_z \Psi]_{z=0} dA = -\frac{\rho}{2} \int [\Psi w]_{z=0} dA, \quad (4.94)$$

where we made use of the kinematic boundary condition (4.93d).

#### 4.4.2 Domain integrated potential and available potential energies

Measuring the zero of gravitational potential energy at  $z = -H$  (see Figure 4.1), yields the domain integrated gravitational potential energy

$$g \rho \int_{z=0} \left[ \int_{-H}^{\eta} z dz \right] dA = \frac{g \rho}{2} \int (\eta^2 - H^2) dA. \quad (4.95)$$

The available potential energy (see Sections ?? and ??) is the difference between the gravitational potential energy and that contained in an ocean at rest with  $\eta = 0$ , so that

$$E_{\text{APE}} = g \rho \int \left[ \int_{-H}^{\eta} z dz \right] dA - g \rho \int \left[ \int_{-H}^0 z dz \right] dA = \frac{g \rho}{2} \int \eta^2 dA = \frac{\rho}{2g} \int (\partial_t \Psi)^2 dA, \quad (4.96)$$

where the final step made use of the dynamic boundary condition (4.93c). Hence, a non-negative available potential energy is associated with any undulation of the free surface, whether the undulation is positive or negative.

In Section ?? we computed the available potential energy for a single layer of shallow water fluid. In that discussion we chose to set  $z = 0$  at the resting free surface (Figure ??), whereas in the current discussion we chose  $z = 0$  at the flat bottom (Figure 4.1). Even so, the available potential energy (4.96) is identical to the shallow water case, which we see by noting that the area average of  $\eta$  vanishes for the current choice in Figure 4.1 (due to volume conservation)

$$\bar{\eta} = \frac{1}{A_{\text{ocn}}} \int \eta dA = 0, \quad (4.97)$$

where

$$A_{\text{ocn}} = \int dA \quad (4.98)$$

is the surface area of the ocean. We thus find

$$E_{\text{APE}} = \frac{g \rho}{2} \int \eta^2 dA = \frac{g \rho}{2} \int (\eta - \bar{\eta})^2 dA = \frac{g \rho}{2} \int (\eta')^2 dA, \quad (4.99)$$

which agrees with equation (??) derived for the shallow water layer. We expect the two energies to agree since the gravitational energy depends only on the density and the position within the gravity field; it has no concern for dynamical assumptions such as whether the fluid motion is approximately hydrostatic (as for the shallow water model) or general (as considered here).

#### 4.4.3 Equipartition for the phase averaged domain integrated energies

The expression (4.96) for the available potential energy can be written

$$E_{\text{APE}} = \frac{\rho}{2} \int \eta \partial_t \Psi dA = \frac{\rho}{2} \int [-\Psi \partial_t \eta + \partial_t(\eta \Psi)] dA = E_{\text{KE}} + \frac{\rho}{2} \int \partial_t(\eta \Psi) dA, \quad (4.100)$$

where we made use of equation (4.94) for the total kinetic energy. If the fields exhibit periodicity in time, such as for a surface gravity wave field, then integration over an integer multiple of the wave period results in an equipartition between the phase averaged available potential energy and phase averaged kinetic energy

$$\langle E_{\text{APE}} \rangle = \langle E_{\text{KE}} \rangle \quad \text{with} \quad \langle E_{\text{APE}} \rangle = \frac{\omega}{2\pi} \int_0^{2\pi/\omega} E_{\text{APE}} dt, \quad (4.101)$$

where  $2\pi/\omega$  is the wave period. We found a similar equipartition of energy in Section 3.6.2 for acoustic waves.

#### 4.4.4 Energetics for the depth integrated linear flow

We here study the mechanical energy contained in the fluid layer within the linear theory. Notably, when computing the kinetic energy in the depth integrated flow, we only integrate to  $z = 0$  since going to  $z = \eta$  involves third order terms that are neglected in the linear theory. Hence, the time derivative of the depth integrated kinetic energy per mass is

$$\frac{\partial}{\partial t} \int_{-H}^0 \mathcal{K} dz = \frac{1}{2} \frac{\partial}{\partial t} \int_{-H}^0 \nabla \Psi \cdot \nabla \Psi dz = \int_{-H}^0 \nabla (\partial_t \Psi) \cdot \nabla \Psi dz, \quad (4.102)$$

which follows since the integral bounds are static so that the time derivative commutes with the integral. Note that we used the partial time derivative operator,  $\partial/\partial t$ , as it is computed holding the horizontal position fixed. We next make use of the harmonic nature of the velocity potential ( $\nabla^2 \Psi = 0$ ) to write

$$\frac{\partial}{\partial t} \int_{-H}^0 \mathcal{K} dz = \int_{-H}^0 \nabla \cdot (\partial_t \Psi \nabla \Psi) dz = [\partial_t \Psi \partial_z \Psi]_{z=0} + \nabla_h \cdot \int_{-H}^0 \partial_t \Psi \nabla_h \Psi dz, \quad (4.103)$$

where we used the bottom kinematic boundary condition,  $\partial_z \Psi = 0$  at  $z = -H$ , and noted that  $H$  is a constant (flat bottom) so that the horizontal derivative commutes with the integral. For the boundary term we use the linearized dynamic boundary condition (4.93c) and linearized kinematic boundary condition (4.93d) to yield

$$\frac{\partial}{\partial t} \int_{-H}^0 \mathcal{K} dz = -g \eta \partial_t \eta + \nabla_h \cdot \int_{-H}^0 \partial_t \Psi \nabla_h \Psi dz. \quad (4.104)$$

Observe that the boundary term is the time tendency of the depth integrated potential energy per mass

$$\int_{-H}^{\eta} \Phi dz = \int_{-H}^{\eta} g z dz = (g/2) (\eta^2 - H^2) \implies \frac{\partial}{\partial t} \int_{-H}^{\eta} g z dz = g \eta \partial_t \eta. \quad (4.105)$$

The potential energy in the column is computed by integrating all the way to the free surface, even for the linear theory. That integration limit is needed since changes in the potential energy over a column of constant density fluid arise solely through changes in the free surface. Bringing the pieces together allows us to write the time derivative of the depth integrated mechanical energy per mass in the linear theory

$$\frac{\partial}{\partial t} \left[ \frac{1}{2} \int_{-H}^0 \mathbf{v} \cdot \mathbf{v} dz + \int_{-H}^{\eta} g z dz \right] = \nabla_h \cdot \int_{-H}^0 \partial_t \Psi \nabla_h \Psi dz. \quad (4.106)$$

Finally, set  $\mathbf{u} = -\nabla_h \Psi$ , and use the linear theory identity,  $\partial_t \Psi = p_d / \rho$  from equation (4.83), thus leading to

$$\frac{\partial}{\partial t} \left[ \int_{-H}^0 \rho \mathcal{K} dz + \int_{-H}^\eta \rho \Phi dz \right] = -\nabla_h \cdot \int_{-H}^0 p_d \mathbf{u} dz. \quad (4.107)$$

Evidently, it is the horizontal convergence of the layer integrated flux of dynamical pressure,  $p_d \mathbf{u}$ , that affects a time change to the layer integrated mechanical energy for the linear theory. That is, the layer integrated mechanical energy has a time tendency due to work by the dynamical pressure. Recall we saw the importance of pressure work in Section 4.2.6 when considering the local energy balances for the fully nonlinear system. We also saw the importance of pressure work for acoustic wave energetics in Section 3.6.3. Other wave systems we encounter in this part of the book also have pressure work central to their energetics.

## 4.5 Traveling gravity waves in a flat domain

We now study a traveling plane wave solution to the equations (4.93b)-(4.93e) as posed in a flat bottom domain such as illustrated in Figure 4.1. The waves are assumed to travel horizontally, with the example in Figure 4.1 showing waves in the  $\pm \hat{x}$  direction. There are no lateral boundaries. The wave amplitude, as we show, exponentially decays when moving from the surface into the interior.

Besides providing an explicit realization of surface gravity waves, our analysis offers experience with the *separation of variables* method for solving certain partial differential equations. In our analysis, we are not interested in the most general wave solution. Instead, we aim to determine a particular solution of sufficient generality to expose the underlying physics of the linear wave fluctuations, and in particular to expose the exponential decay of the wave amplitude with depth. Furthermore, given linearity, the *superposition principle* holds whereby the sum or integral of particular solutions are also solutions.

### 4.5.1 Horizontally traveling plane wave

We seek a traveling plane wave solution with angular frequency,  $\omega > 0$ , and horizontal wavevector and wave direction

$$\mathbf{k} = k_x \hat{x} + k_y \hat{y} \quad \text{and} \quad \hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|. \quad (4.108)$$

For this purpose we assume the waves appear in the velocity potential in the shape of a cosine modulated by a vertical structure function<sup>11</sup>

$$\Psi(x, y, z, t) = \Psi_0 \Gamma(z) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (4.109)$$

Plugging this ansatz into Laplace's equation,  $\nabla^2 \Psi = 0$ , leads to the ordinary differential equation satisfied by the non-dimensional vertical structure function

$$\frac{d^2 \Gamma}{dz^2} = |\mathbf{k}|^2 \Gamma \quad -H \leq z \leq 0 \quad (4.110a)$$

$$\frac{d\Gamma}{dz} = 0 \quad \text{at } z = -H, \quad (4.110b)$$

---

<sup>11</sup>We could use complex exponentials for the traveling wave, as discussed in Chapter 1. We here work with the real trigonometric functions to exemplify their use.

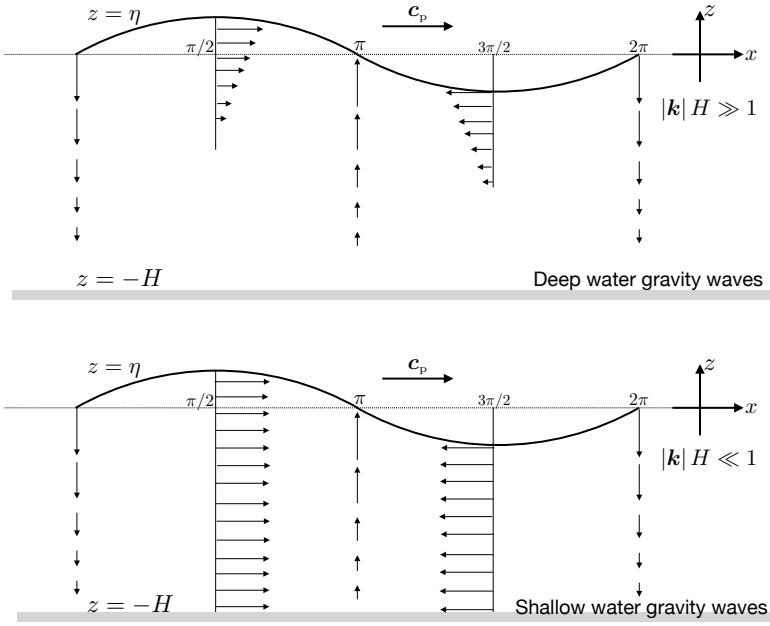


FIGURE 4.2: Snapshot of surface gravity waves over deep water ( $|\mathbf{k}|H \gg 1$ , top panel) and shallow water ( $|\mathbf{k}|H \ll 1$ , lower panel), with velocities computed according to equations (4.113a) and (4.113b). Shown here are the horizontal and vertical components to the velocity at a particular time instance, with the horizontal velocity in phase with the free surface undulations, and the vertical velocity  $\pi/2$  out of phase. The horizontal axis along the top designates values for  $\mathbf{k} \cdot \mathbf{x}$ , with values of  $0, \pi/2, \pi, 3\pi/2, 2\pi$  providing samples along the wave.

where the bottom boundary condition is required to satisfy the no-normal flow condition (4.93e). We write the solution in the form

$$\Psi = \Psi_0 \cosh[|\mathbf{k}|(z + H)] \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (4.111a)$$

$$\Psi_0 = \frac{g \eta_0 / \omega}{\cosh(|\mathbf{k}|H)}, \quad (4.111b)$$

so that the dynamic boundary condition (4.93c) renders the free surface height

$$\eta(\mathbf{x}, t) = \eta_0 \sin(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (4.112)$$

The corresponding fluid velocity field,  $\mathbf{v} = -\nabla\Psi$ , is given by

$$\mathbf{u} = \frac{g \eta_0 \hat{\mathbf{k}}}{C_p} \frac{\cosh[|\mathbf{k}|(z + H)] \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}{\cosh(|\mathbf{k}|H)} \quad (4.113a)$$

$$w = -\frac{g \eta_0}{C_p} \frac{\sinh[|\mathbf{k}|(z + H)] \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}{\cosh(|\mathbf{k}|H)}, \quad (4.113b)$$

where the wave phase speed is given by

$$C_p = \omega / |\mathbf{k}| > 0. \quad (4.114)$$

Figure 4.2 depicts the horizontal and vertical velocity in a snapshot of a deep water wave ( $|\mathbf{k}|H \gg 1$ ) and shallow water wave ( $|\mathbf{k}|H \ll 1$ ). We emphasize the following properties of these waves.

- The horizontal fluid particle velocity,  $\mathbf{u}$ , is parallel to the horizontal wavevector,  $\mathbf{k}$ . Hence,

the gravity waves are horizontally longitudinal.

- The horizontal velocity is in phase with the free surface,  $\eta$ , whereas they are both  $\pi/2$  out of phase with the vertical velocity.
- The vertical velocity vanishes at the bottom,  $z = -H$ , as needed to satisfy the no-normal flow boundary condition.
- The horizontal wave number,  $|\mathbf{k}|$ , both determines the horizontal wavelength,  $2\pi/|\mathbf{k}|$ , as well as the vertical decay scale,  $|\mathbf{k}|$ . This coupling of the horizontal to the vertical is a notable property of surface gravity waves.
- Letting time progress at a fixed space position reveals a clockwise progression of a fluid particle, as can be imagined by letting Figure 4.2 evolve in time. Further details of fluid particle trajectories are developed in Section 4.11 when studying Stokes drift.
- In the shallow water limit,  $|\mathbf{k}| H \ll 1$ , the horizontal velocity is depth independent. The vertical velocity is a linear function of depth and it is a factor of  $|\mathbf{k}| H$  times smaller in magnitude than the horizontal velocity. We provide a focused study of such shallow water gravity waves in Section 7.5.

### 4.5.2 Domain integrated mechanical energy of a traveling wave

The domain integrated kinetic energy (4.94) contained in a traveling surface gravity wave is

$$E_{\text{KE}} = -\frac{\rho}{2} \int_{z=0} \Psi \partial_t \eta \, dA \quad (4.115a)$$

$$= (\rho/2) \Psi_0 \eta_0 \omega \cosh(|\mathbf{k}| H) \int \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \, dA \quad (4.115b)$$

$$= (\rho/2) g \eta_0^2 \int \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \, dA, \quad (4.115c)$$

and likewise the domain integrated available potential energy (4.96) is

$$E_{\text{APE}} = (\rho/2) g \eta_0^2 \int \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \, dA, \quad (4.116)$$

so that their sum is a space and time constant

$$E_{\text{KE}} + E_{\text{APE}} = A_{\text{ocn}} \rho g \eta_0^2 / 2, \quad (4.117)$$

where  $A_{\text{ocn}}$  is the total ocean area.<sup>12</sup> We also see that the phase average of the domain integrated kinetic energy and available potential energy manifest the equipartition property (4.101)

$$\langle E_{\text{APE}} \rangle = \langle E_{\text{KE}} \rangle = A_{\text{ocn}} \rho g \eta_0^2 / 4. \quad (4.118)$$

### 4.5.3 Dispersion relation

Combining the two  $z = 0$  boundary conditions (4.93c) and (4.93d) yields

$$(\partial_{tt} + g \partial_z) \Psi = 0 \quad \text{at } z = 0. \quad (4.119)$$

---

<sup>12</sup>For the single traveling surface gravity wave to be a valid wave solution requires no lateral boundaries, in which case the ocean area is formally infinite.

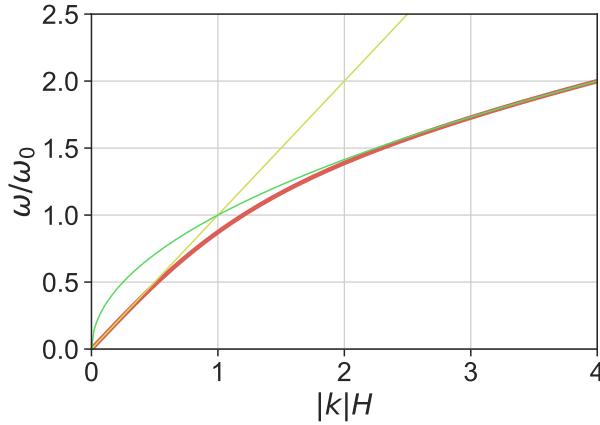


FIGURE 4.3: The dispersion relation (4.120) for surface gravity waves, plotted here as  $\omega/\omega_0 = \sqrt{|\mathbf{k}|H} \tanh(|\mathbf{k}|H)$  (dark thick curved line), where  $\omega_0^2 = g/H$  is the square of a fundamental frequency for surface gravity waves. We also show the dispersion relation for non-dispersive shallow water waves,  $\omega/\omega_0 = |\mathbf{k}|H$  (straight line), which holds for  $|\mathbf{k}|H \ll 1$ , and the dispersive deep-water gravity waves,  $\omega/\omega_0 = \sqrt{|\mathbf{k}|H}$ , which holds for  $|\mathbf{k}|H \gg 1$  (see Section 4.5.5).

Substituting the traveling plane wave (4.111a) into this relation leads to the *dispersion relation*

$$\omega^2 = g |\mathbf{k}| \tanh(|\mathbf{k}|H) \implies \omega = \sqrt{g |\mathbf{k}| \tanh(|\mathbf{k}|H)}. \quad (4.120)$$

The dispersion relation constrains those values available for the angular frequency,  $\omega$ , and wavenumber,  $|\mathbf{k}|$ . That is, the surface gravity waves only exist if their frequency and wavenumber are related according to the dispersion relation (4.120). We depict the dispersion relation in Figure 4.3.

#### 4.5.4 Alternative forms for the velocity potential and velocity

The dispersion relation (4.120) allows for a slight rewrite of the velocity potential (4.111a) and velocity field (4.113a) and (4.113b), with these variety of forms appearing in the literature

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t \quad (4.121a)$$

$$\Psi_0 = \frac{g \eta_0}{\omega \cosh(|\mathbf{k}|H)} = \frac{\eta_0 C_p}{\sinh(|\mathbf{k}|H)} \quad (4.121b)$$

$$\eta = \eta_0 \sin \mathcal{P} \quad (4.121c)$$

$$\Psi = \frac{g \eta_0 \cosh[|\mathbf{k}|(z+H)] \cos \mathcal{P}}{\omega \cosh(|\mathbf{k}|H)} = \frac{\eta_0 C_p \cosh[|\mathbf{k}|(z+H)] \cos \mathcal{P}}{\sinh(|\mathbf{k}|H)} \quad (4.121d)$$

$$\mathbf{u} = \frac{g \eta_0 \hat{\mathbf{k}}}{C_p} \frac{\cosh[|\mathbf{k}|(z+H)] \sin \mathcal{P}}{\cosh(|\mathbf{k}|H)} = \frac{\eta_0 \omega \hat{\mathbf{k}} \cosh[|\mathbf{k}|(z+H)] \sin \mathcal{P}}{\sinh(|\mathbf{k}|H)} \quad (4.121e)$$

$$w = -\frac{g \eta_0}{C_p} \frac{\sinh[|\mathbf{k}|(z+H)] \cos \mathcal{P}}{\cosh(|\mathbf{k}|H)} = -\frac{\eta_0 \omega \sinh[|\mathbf{k}|(z+H)] \cos \mathcal{P}}{\sinh(|\mathbf{k}|H)}. \quad (4.121f)$$

#### 4.5.5 Phase speed, group velocity, and angular frequency

We here characterize the phase speed, group velocity, and angular frequency for the surface gravity waves, and introduce the limits for deep water and shallow water waves.

### Phase speed

The phase speed for the surface gravity wave is given by

$$C_p = \omega / |\mathbf{k}| = \sqrt{(g/|\mathbf{k}|) \tanh(|\mathbf{k}| H)}. \quad (4.122)$$

We consider the two limits:  $|\mathbf{k}| H \gg 1$  are known as shortwaves or deep water waves, and  $|\mathbf{k}| H \ll 1$  are known as longwaves or shallow water waves. In these two limits the phase speed is given by

$$\omega \approx \sqrt{g |\mathbf{k}|} \quad C_p \approx \sqrt{g/|\mathbf{k}|} \quad |\mathbf{k}| H \gg 1 \quad \text{shortwave/deep water} \quad (4.123a)$$

$$\omega \approx |\mathbf{k}| \sqrt{g H} \quad C_p \approx \sqrt{g H} \quad |\mathbf{k}| H \ll 1 \quad \text{longwave/shallow water.} \quad (4.123b)$$

Observe that the shallow water gravity waves are non-dispersive since the phase speed is the same for all wave numbers,  $C_p \approx \sqrt{g H}$ . In contrast, the deep water gravity waves are dispersive, with the shorter (and higher frequency) waves having a slower phase speed than longer waves. To further emphasize the point about deep water waves, write the squared phase speed for the deep water waves as

$$C_p^2 = \frac{g}{|\mathbf{k}|} = \frac{g^2}{\omega^2} \implies C_p = g/\omega. \quad (4.124)$$

Hence, so long as the deep water limit is maintained, higher frequency and shorter waves have smaller phase speed than lower frequency and longer waves.

### Group velocity

As shown in Section 1.6, the group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}} \omega$ , measures the direction and speed of a group of waves (e.g., a wave train or wave packet), with the group velocity for surface gravity waves given by

$$\mathbf{c}_g = \frac{g \hat{\mathbf{k}}}{2 \omega} \left[ \frac{|\mathbf{k}| H + \cosh(|\mathbf{k}| H) \sinh(|\mathbf{k}| H)}{\cosh^2(|\mathbf{k}| H)} \right]. \quad (4.125)$$

The ratio of the group speed to the phase speed is given by

$$\frac{\hat{\mathbf{k}} \cdot \mathbf{c}_g}{C_p} = \frac{g |\mathbf{k}|^2 [|\mathbf{k}| H + \cosh(|\mathbf{k}| H) \sinh(|\mathbf{k}| H)]}{2 \omega^2 \cosh^2(|\mathbf{k}| H)} = \frac{1}{2} \left[ 1 + \frac{2 |\mathbf{k}| H}{\sinh(2 |\mathbf{k}| H)} \right]. \quad (4.126)$$

The shallow water limit, with  $|\mathbf{k}| H \ll 1$ , has phase speed and group speed equal, whereas the deep water waves have group speed one-half the phase speed. More precisely, we find the following limiting behaviors.

$$\omega \approx \sqrt{g |\mathbf{k}|} \quad C_p \approx \sqrt{g/|\mathbf{k}|} \quad \mathbf{c}_g \approx \mathbf{c}_p/2 \quad |\mathbf{k}| H \gg 1 \quad \text{shortwave/deep water} \quad (4.127a)$$

$$\omega \approx |\mathbf{k}| \sqrt{g H} \quad C_p \approx \sqrt{g H} \quad \mathbf{c}_g \approx \mathbf{c}_p \quad |\mathbf{k}| H \ll 1 \quad \text{longwave/shallow water.} \quad (4.127b)$$

For shallow water gravity waves, the group velocity equals to the phase velocity. In contrast, for deep water gravity waves, the group velocity magnitude is one-half the phase speed. When watching a packet of deep water gravity waves, we see the phase of the carrier waves appear at the back of the packet and move forward at twice the speed of the packet, only to then disappear at the front of the packet.

### Angular frequency

The deep water limit,  $|\mathbf{k}|H \gg 1$ , and shallow water limit,  $|\mathbf{k}|H \ll 1$ , are set according to the wavenumber. These limits lead to a distinct frequency for the two waves. To compute the ratio of the corresponding angular frequencies, introduce two non-dimensional numbers according to

$$\Gamma_{\text{sw}} \equiv H |\mathbf{k}|_{\text{sw}} \ll 1 \quad \text{and} \quad \Gamma_{\text{dw}} \equiv H |\mathbf{k}|_{\text{dw}} \gg 1, \quad (4.128)$$

in which case

$$\frac{\omega_{\text{dw}}^2}{\omega_{\text{sw}}^2} = \frac{g |\mathbf{k}|_{\text{dw}}}{|\mathbf{k}|_{\text{sw}}^2 c_{\text{grav}}^2} = \frac{g H^{-1} \Gamma_{\text{dw}}}{\Gamma_{\text{sw}}^2 H^{-2} g H} = \frac{\Gamma_{\text{dw}}}{\Gamma_{\text{sw}}^2} \gg 1. \quad (4.129)$$

We conclude that frequency of those short gravity waves whose extent is largely confined to the upper ocean (deep water waves) is much higher than the long gravity waves that extend throughout the water column (shallow water gravity waves). So although the shallow water waves and deep water waves feel the same gravitational acceleration, the huge scale for the shallow water waves leads to far lower frequency than for the deep water gravity waves.

#### 4.5.6 Particle trajectories ignoring Stokes drift

Following the discussion in Section ?? of fluid particles and the motion of points within a continuum, we here determine the trajectory of a fluid particle labeled by the material coordinate,  $\mathbf{a}$ , and do so by time integrating the ordinary differential equation (Section ??)

$$\frac{\partial \mathbf{X}(\mathbf{a}, T)}{\partial T} = \mathbf{v}[\mathbf{X}(\mathbf{a}, T), T], \quad (4.130)$$

where  $T = t$  is the material time label. In Section 4.11 we study the fluid particle motion within a surface gravity wave, noting that there is a net motion in the direction of the phase. This *Stokes drift* arises from the distinction between averaging at a fixed point in space versus averaging that follows a fluid particle. Here, we provide a step towards the Stokes drift discussion by considering the leading order motion of fluid particles found by ignoring the distinction between Eulerian and Lagrangian averaging. In effect, we time integrate equation (4.130) by fixing the particle trajectory on the right hand side to the value it had at an arbitrary initial time,  $T = t_0$ . Using the horizontal velocity (4.121e) and vertical velocity (4.121f) we have (now dropping the  $\mathbf{a}$  label for brevity)

$$\frac{d\mathbf{X}_h}{dT} = \frac{\eta_0 \omega \hat{\mathbf{k}} \cosh[|\mathbf{k}|(z_0 + H)] \sin(\mathbf{k} \cdot \mathbf{x}_0 - \omega T)}{\sinh(|\mathbf{k}| H)} \quad (4.131a)$$

$$\frac{dZ}{dT} = -\frac{\eta_0 \omega \sinh[|\mathbf{k}|(z_0 + H)] \cos(\mathbf{k} \cdot \mathbf{x}_0 - \omega T)}{\sinh(|\mathbf{k}| H)}, \quad (4.131b)$$

which integrates to the horizontal and vertical trajectories

$$\mathbf{X}_h = \mathbf{x}_0 + \frac{\eta_0 \hat{\mathbf{k}} \cosh[|\mathbf{k}|(z_0 + H)] \cos(\mathbf{k} \cdot \mathbf{x}_0 - \omega T)}{\sinh(|\mathbf{k}| H)} \quad (4.132a)$$

$$Z = z_0 + \frac{\eta_0 \sinh[|\mathbf{k}|(z_0 + H)] \sin(\mathbf{k} \cdot \mathbf{x}_0 - \omega T)}{\sinh(|\mathbf{k}| H)}. \quad (4.132b)$$

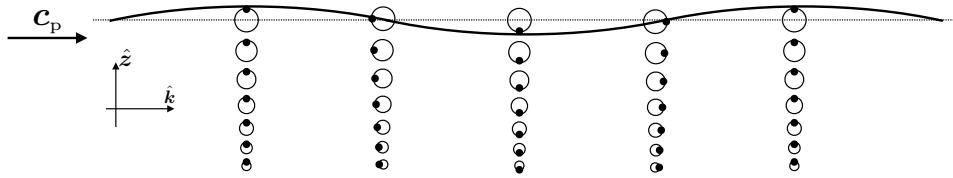


FIGURE 4.4: Motion of fluid particles in deep water waves. Shown here is a snapshot of the wave, with fluid particles denoted by black dots positioned at a point along their nearly circular orbits. The wavevector is directed to the right so that particles exhibit a clockwise orbital motion as the wave moves. Correspondingly, by moving ones eye to the left in this snapshot, then the series of fluid particles exhibit a clockwise rotation. As the horizontal motion is larger on the top of the orbit than the bottom, there is a net particle drift in the direction of the wave; this is *Stokes drift* discussed in Section 4.11. For these deep water waves, the particle motion becomes exponentially small when moving away from the interface. In contrast, for shallow water gravity waves, the particle orbits are elliptical and the horizontal major axis is depth independent whereas the vertical excursion vanishes at the bottom. This figure is taken after Figure 50 of Lighthill (1978).

Eliminating the time dependence leads to the equation for an ellipse in the horizontal-vertical plane

$$\frac{(\mathbf{X}_h - \mathbf{x}_0)^2}{\cosh^2[|\mathbf{k}|(z_0 + H)]} + \frac{(Z - z_0)^2}{\sinh^2[|\mathbf{k}|(z_0 + H)]} = \frac{\eta_0^2}{\sinh^2(|\mathbf{k}|H)}. \quad (4.133)$$

In the deep water limit, with  $|\mathbf{k}|H \gg 1$ , the ellipse becomes circular and the radius exponentially decreases with depth, with short waves (large  $|\mathbf{k}|$ ) decaying over a shorter depth range. As a result, deep water waves have nearly circular particle trajectories such as those shown in Figure 4.4. In the shallow water limit, with  $|\mathbf{k}|H \ll 1$ , the vertical motion reduces to zero so that particles move horizontally. Furthermore, their horizontal excursions are independent of  $z$ , so that particles throughout the full column move coherently by the same amount.

### 4.5.7 Depth integrated mechanical energy of a traveling plane wave

In Section 4.4.4 we derived general expressions for the layer integrated mechanical energy budget. Here we specialize those results to the case of traveling plane surface gravity waves, and limit our focus to the phase averaged quantities.

#### Available potential energy

The layer integrated potential energy (per horizontal area) is given by

$$g\rho \int_{-H}^{\eta} z dz = (\rho g \eta_0^2/2) \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t) - \rho g H^2/2. \quad (4.134)$$

We are concerned with the potential energy relative to the constant term,  $-\rho g H^2/2$ . That is, we focus on the available potential energy of the waves rather than the potential energy (see Section 4.5.2), in which we compute

$$g\rho \int_{-H}^{\eta} z dz - g\rho \int_{-H}^0 z dz = g\rho \int_0^{\eta} z dz = (\rho g \eta_0^2/2) \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (4.135)$$

The phase average of the available potential energy (per horizontal area) is thus given by

$$\left\langle g\rho \int_0^{\eta} z dz \right\rangle = \rho g \eta_0^2/4. \quad (4.136)$$

Evidently, the phase averaged and layer integrated available potential energy is proportional to the square of the amplitude of the free surface undulations,  $\eta_0^2$ .

### Kinetic energy

Making use of the expression (4.113a) for the horizontal wave velocity leads to the layer integrated kinetic energy (per horizontal area) in the horizontal flow

$$\frac{\rho}{2} \int_{-H}^0 \mathbf{u}^2 dz = \frac{\rho \Psi_0^2 |\mathbf{k}|^2}{2} \int_{-H}^0 \cosh^2[|\mathbf{k}|(z + H)] \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t) dz, \quad (4.137)$$

which has the phase average

$$\left\langle \frac{\rho}{2} \int_{-H}^0 \mathbf{u}^2 dz \right\rangle = \frac{\rho \Psi_0^2 |\mathbf{k}|^2}{4} \int_{-H}^0 \cosh^2[|\mathbf{k}|(z + H)] dz. \quad (4.138)$$

Likewise, the phase averaged kinetic energy of the layer integrated vertical wave motion (4.113b) is given by

$$\left\langle \frac{\rho}{2} \int_{-H}^0 w^2 dz \right\rangle = \frac{\rho \Psi_0^2 |\mathbf{k}|^2}{4} \int_{-H}^0 \sinh^2[|\mathbf{k}|(z + H)] dz. \quad (4.139)$$

With  $\cosh^2 y + \sinh^2 y = \cosh(2y)$ , the phase average of the layer integrated kinetic energy per area of the wave is

$$\left\langle \frac{\rho}{2} \int_{-H}^0 \mathbf{v}^2 dz \right\rangle = \frac{\rho \Psi_0^2 |\mathbf{k}|^2}{4} \int_{-H}^0 \cosh[2|\mathbf{k}|(z + H)] dz = (\Psi_0^2 |\mathbf{k}|/8) \sinh(2|\mathbf{k}|H). \quad (4.140)$$

Making use of the expression (4.111b) for  $\Psi_0$ , as well as the dispersion relation (4.120) yields

$$\left\langle \frac{\rho}{2} \int_{-H}^0 \mathbf{v}^2 dz \right\rangle = \frac{\rho g^2 \eta_0^2 |\mathbf{k}|}{8 \omega^2} \frac{\sinh(2|\mathbf{k}|H)}{\cosh^2(|\mathbf{k}|H)} = \rho g \eta_0^2 / 4, \quad (4.141)$$

which is identical to the phase averaged layer integrated available potential energy (4.136). We thus find, again, the equipartition of the phase averaged kinetic and available potential energies, here computed per horizontal area.

### Energy flux vector

Now we determine the phase averaged horizontal flux of layer integrated mechanical energy. The general expression is given by equations (4.106) and (4.107), and is here specialized to the plane wave

$$\int_{-H}^0 p_d \mathbf{u} dz = -\rho \int_{-H}^0 \partial_t \Psi \nabla_h \Psi dz \quad (4.142a)$$

$$= (\rho \Psi_0^2 \mathbf{k} \omega) \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \int_{-H}^0 \cosh^2[|\mathbf{k}|(z + H)] dz, \quad (4.142b)$$

which has a phase average

$$\left\langle \int_{-H}^0 p_d \mathbf{u} dz \right\rangle = \frac{\rho \Psi_0^2 \mathbf{k} \omega}{2} \int_{-H}^0 \cosh^2[k(z + H)] dz \quad (4.143a)$$

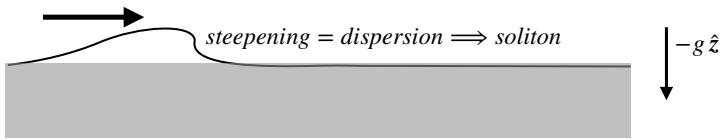


FIGURE 4.5: A soliton in the deep water limit results when the wave dispersion (long waves travel faster than short waves) balances the nonlinear steepening. The result is a soliton, which has an exact analytic expression following from the KdV equation (see [Drazin and Johnson \(1989\)](#)).

$$= \frac{\rho \Psi_0^2 \omega \hat{k}}{4} [|\mathbf{k}| H + \sinh(|\mathbf{k}| H) \cosh(|\mathbf{k}| H)] \quad (4.143b)$$

$$= (\rho g \eta_0^2 / 2) \mathbf{c}_g, \quad (4.143c)$$

where we used equation (4.125) for the group velocity and equation (4.111b) for  $\Psi_0$ . We thus see that the mechanical energy fluxed by the waves is, in the phase average and depth integral, equal to the mechanical energy times the group velocity

$$\left\langle \int_{-H}^0 p_d \mathbf{u} dz \right\rangle = \rho \mathbf{c}_g \left\langle \int_{-H}^0 \mathcal{K} dz + \int_0^\eta \Phi dz \right\rangle. \quad (4.144)$$

This connection between wave energy flux and the group velocity provides yet another reason why the group velocity is more relevant to wave mechanics than the phase velocity.

### Comments on group velocity and a single wave

The analysis in this subsection focused on energetics for a single wave with wavevector,  $\mathbf{k}$ . However, the group velocity is the wavevector space gradient of the dispersion relation,  $c_g = \nabla_{\mathbf{k}} \omega$ , which tacitly considers more than a single wavevector. From our study of wavepackets in Section 1.6, we consider the packet to be localized in  $\mathbf{k}$ -space around a single wavevector, so that we can develop wave energetics by focusing on that single wave. When this wave forms the central carrier wave for a packet, then the group velocity serves as the velocity of the packet and it appears in the energy flux.

#### 4.5.8 Further study

Further discussions of surface gravity waves, following that provided here, can be found in Section 54 of [Fetter and Walecka \(2003\)](#), Lectures 3 and 4 of [Pedlosky \(2003\)](#), and Section 7.1 of [Vallis \(2017\)](#).

## 4.6 Qualitative features of deep water waves

The shortwave/deep water waves are notable for having shorter waves travel slower than longer waves. In the event of a perturbation to the fluid, such as from a stone dropped into a pond or a storm on a lake or the ocean, deep water waves are energized. The dispersion relation (4.123a) means that longer waves spread away from the source faster than the shorter waves, leading to a self-organization of the wavelengths and corresponding wave packets.

Now imagine a deep water wave packet that somehow steepens and takes on a nonlinear form. Fourier decomposing this nonlinear wave into linear deep water modes requires more shortwave modes in the steep region, whereas the less steep portion of the wave requires longer deep water modes, which travel faster. If the nonlinear steepening on the wave face is exactly

balanced by the faster dispersion of the long waves near the wave base and backside, then the wave pattern remains stable; it does not break. This balance of steepening and dispersion describes the fundamental features of a *soliton* as depicted in Figure 4.5, with a soliton a *nonlinear wave*.

## 4.7 Shallow water waves approaching a shore

The shallow water limit is notable for the absence of wave dispersion; i.e., shallow water gravity waves of all wavelengths travel at speed  $\sqrt{gH}$ . Tsunamis are the prototypical shallow water waves. The shallow water dispersion relation also means that shallow water gravity waves slow down when the ocean depth shoals, as when approaching a shoreline. Consequently, as waves reach the shoreline there is a tendency to accumulate wave energy as the deeper waves pile up behind the shallower waves. Furthermore, the steeper part of the wave, being part of a thicker region of the fluid and thus a larger effective  $H$ , travels slightly faster than the wave trough. As such, the steeper part of the wave overtakes the trough and, at some point, the assumptions of linearity breakdown and the shallow water waves break on the beach as depicted in Figure 4.6. In the remainder of this section, we provide some quantitative context for these remarks.

### 4.7.1 Wavenumber changes

Consider the ray equations from Section 2.3 for a one-dimensional shallow water gravity wave approaching a beach with  $\mathbf{k} = k \hat{\mathbf{x}}$  where  $k > 0$ . Let the resting depth be a function,  $H(x)$ , that decreases in the  $+\hat{\mathbf{x}}$  direction,  $\partial_x H < 0$ . Since the background state is time-independent (i.e., the resting depth is static), the angular frequency remains fixed while following along a ray, so that<sup>13</sup>

$$\frac{D_r \varpi}{Dt} = 0. \quad (4.145)$$

With the local shallow water dispersion relation given by

$$\varpi(x) = k(x) \sqrt{g H(x)}, \quad (4.146)$$

we have

$$\frac{D_r \varpi}{Dt} = 0 \implies \frac{1}{k} \frac{D_r k}{Dt} = -\frac{1}{2H} \frac{D_r H}{Dt} > 0. \quad (4.147)$$

The inequality holds since the shoreline gets shallower as the ray approaches the beach. Consequently, the wavenumber increases and the wavelength decreases for waves approaching the shore.

### 4.7.2 Wave energy and wave action

What happens to the wave energy as a wave approaches the shore? Common experience suggests that the wave energy increases since the amplitude increases when the wave moves into shallower water. To see this effect analytically, consider the wave action equation introduced in Section 2.5. Recall that the wave action,  $\mathcal{A}$ , equals to the phase averaged wave energy,  $\langle \mathcal{H} \rangle$ , divided by the angular frequency,  $\omega$ . When following a ray, the wave action satisfies equation

---

<sup>13</sup>Standing on a pier looking at a periodic wave field approaching the beach should reveal that the angular frequency is a constant even as the waves move into shallower water.

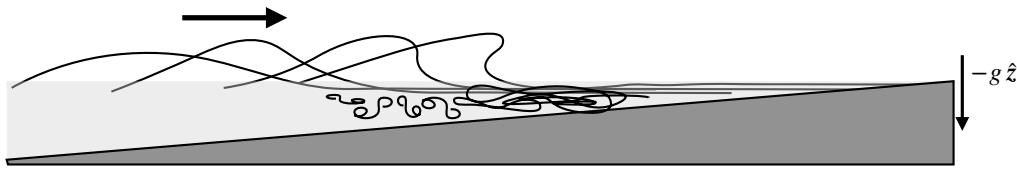


FIGURE 4.6: Shallow water waves approaching a shoreline steepen and eventually break. We can infer this behavior from the phase speed,  $C_p = \sqrt{g H}$ , whereby waves in deeper water move slightly faster than those in shallower water, so that the wave energy accumulates near the shore. Furthermore, water on the steeper part of the wave moves slightly faster than water in the trough, due to the difference in thickness of the water. This process causes water on the steeper portion of the wave to travel slightly faster than in the trough, leading to steepening of the waves. Nonlinearities eventually invalidate the assumptions made in deriving the linear waves. Even so, the qualitative characterization based on the linear analysis allows for a useful heuristic understanding of shallow water wave breaking on the beach. [This video](#) offers a pedagogical introduction to shallow water waves, with the 12-minute mark describing “compression” waves, in which waves steepen due to differences in the layer thickness, much like shallow water waves approaching a beach

(2.88), which here takes the form

$$\frac{1}{\mathcal{A}} \frac{D_r \mathcal{A}}{Dt} = -\nabla \cdot \mathbf{c}_g. \quad (4.148)$$

For the one-dimensional shallow water example considered here, the convergence of the group velocity is

$$-\nabla \cdot \mathbf{c}_g = -\partial_x c_g = -\partial_x(gH)^{1/2} = -(gH)^{1/2} \partial_x H / (2H) > 0. \quad (4.149)$$

Hence, the wave action increases along a ray moving towards a beach. Furthermore, as noted earlier, a static background environment means that the angular frequency remains constant along a ray. Consequently, the increase in wave action following a ray onto the beach occurs since the phase averaged wave energy increases, which means that the wave amplitude is increasing. Eventually the linear analysis fails when the nonlinear effects amplify with increasing wave amplitude, as illustrated in Figure 4.6 where the waves eventually break. Even so, it is satisfying that the linear theory provides a clear picture of its ultimate demise as nonlinear effects take over.

### 4.7.3 Wave refraction

When distant from the shore, waves typically approach at an oblique angle, but, observations readily verify, waves bend as they start to feel the bottom, so that their phase velocity is nearly straight onto the shoreline upon reaching the beach. This bending of the wave phase lines is known as *refraction*, and the amount of refraction is a function of the bottom slope and rules for refraction are embodied in *Snell's law* from optics.

To introduce the notion of shallow water gravity wave refraction, return to the ray equation (2.37c), now with the wavevector two-dimensional,

$$\mathbf{k} = \hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y. \quad (4.150)$$

If the bottom remains a function just of  $x$ , then the angular frequency remains constant along a ray, as does  $k_y$ . So writing the squared angular frequency as

$$\omega^2 = g H (k_x^2 + k_y^2), \quad (4.151)$$

it is only  $H$  and  $k_x$  that change when following along a ray. As the ray approaches the shore,  $H(x)$  gets smaller. So to keep the angular frequency constant along the ray requires  $k_x$  to

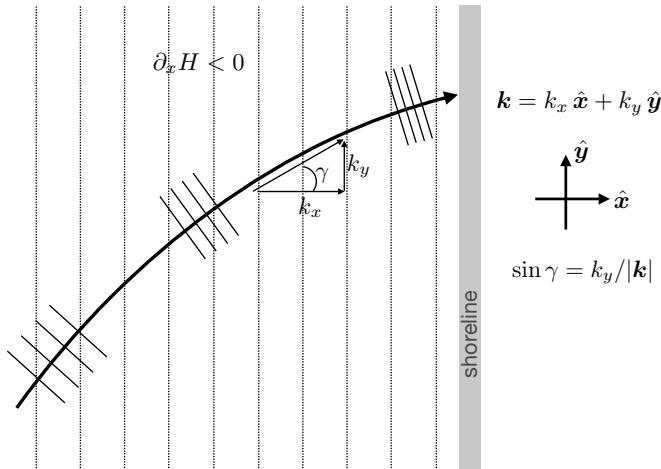


FIGURE 4.7: Depicting the refraction of shallow water gravity waves as a packet approaches the shoreline, thus causing the phase velocity to bend more directly into the shore and lines of constant phase to become more parallel to the shore. Refraction with the bottom shoaling ( $\partial_x H > 0$ ) occurs by increasing  $k_x$  (wave length on a ray gets smaller approaching the shore) while keeping the angular frequency and  $k_y$  constant along a ray. In this manner, the angle,  $\gamma$ , reduces toward zero as the bottom shoals, simply through increasing  $k_x$  as a wave packet moves toward the shoreline. The dotted straight lines are lines of constant bottom depth,  $H(x) = \text{constant}$ .

increase, which is the same result as in Section 4.7.1. However, now the increase in  $k_x$  means that the wave phase velocity turns into the coast so that phase lines become more parallel as the wave approaches the shore. Figure 4.7 illustrates the basic physics.

#### 4.7.4 Further study

This video from Prof. A. Hogg provides a pedagogical introduction to shallow water wave breaking along with deep water solitons, both as realized in a laboratory. Also, section 2.2 of Johnson (1997) provides a thorough mathematical analysis of shallow water waves approaching a shoreline.

## 4.8 Standing gravity waves in a closed basin

We now consider the surface gravity wave equations (4.93a)-(4.93e) for a closed rectangular basin of constant depth and with horizontal dimensions  $x \in [0, L_x]$  and  $y \in [0, L_y]$ . As we show, waves in a bounded domain are no longer traveling, but instead they are standing waves that oscillate in place rather than travel. We can think of such standing wave modes as a superposition of two oppositely traveling waves with the same frequency and wavenumber that are locked in-phase in a manner that satisfies the boundary conditions. For example, the sum of a right and left moving wave with equal amplitude, wavenumber, and frequency is given by the standing pattern

$$A \cos(kx - \omega t) + A \cos(-kx - \omega t) = 2A \cos(\omega t) \cos(kx). \quad (4.152)$$

We encountered a similar situation in Exercise 3.1 when considering acoustic waves in a closed rectangular cavity.

### 4.8.1 Solution for the standing waves

To satisfy the no-normal flow conditions at the four walls (equation (4.93e)) requires the horizontal wave numbers to be quantized. Correspondingly, the waves are not traveling plane waves since the flow is confined in a box. Rather, the solution is in the form of spatially standing wave that oscillates in time

$$\Psi(x, y, z, t) = \Psi_0 \Gamma(z) \cos(k_m x) \cos(l_n y) \cos(\omega t), \quad (4.153)$$

where the quantized wave numbers are

$$k_m = m \pi / L_x \quad \text{and} \quad l_n = n \pi / L_y \quad \text{with } m, n \text{ integers.} \quad (4.154)$$

The vertical structure function satisfies same differential equation as for the traveling waves in a channel from Section 4.5,

$$\frac{d^2\Gamma}{dz^2} = (k_m^2 + l_n^2) \Gamma \quad -H \leq z \leq 0 \quad (4.155a)$$

$$\frac{d\Gamma}{dz} = 0 \quad \text{at } z = -H, \quad (4.155b)$$

thus leading to the standing gravity wave solution

$$\Psi = \Psi_0 \cosh[K_{m,n}(z + H)] \cos(k_m x) \cos(l_n y) \sin(\omega t) \quad (4.156a)$$

$$\Psi_0 = \frac{g \eta_0 / \omega_{m,n}}{\cosh(K_{m,n} H)} \quad (4.156b)$$

$$\eta = \eta_0 \cos(k_m x) \cos(l_n y) \cos(\omega_{m,n} t) \quad (4.156c)$$

$$u = \frac{g \eta_0 k_m}{\omega_{m,n}} \frac{\cosh[K_{m,n}(z + H)]}{\cosh(K_{m,n} H)} \sin(k_m x) \cos(l_n y) \sin(\omega_{m,n} t) \quad (4.156d)$$

$$v = \frac{g \eta_0 l_n}{\omega_{m,n}} \frac{\cosh[K_{m,n}(z + H)]}{\cosh(K_{m,n} H)} \cos(k_m x) \sin(l_n y) \sin(\omega_{m,n} t) \quad (4.156e)$$

$$w = -\frac{g \eta_0 K_{m,n}}{\omega_{m,n}} \frac{\sinh[K_{m,n}(z + H)]}{\cosh(K_{m,n} H)} \cos(k_m x) \cos(l_n y) \sin(\omega_{m,n} t) \quad (4.156f)$$

$$K_{m,n}^2 = k_m^2 + l_n^2. \quad (4.156g)$$

As for the channel, we find the dispersion relation by substituting the standing wave solution into the boundary condition equation (4.119), thus leading to the quantized angular frequencies

$$\omega_{m,n}^2 = g K_{m,n} \tanh(K_{m,n} H). \quad (4.157)$$

### 4.8.2 Gravest seiche mode as an example

The standing wave solutions are commonly referred to as *seiches*, and they can be found in enclosed lakes and ocean basins after strong and persistent winds. Winds, generally with large scale variations, tend to force the lowest or *gravest* seiche mode. For example, assume the wind blows in the zonal direction so that it excites the lowest zonal standing wave with frequency

$$\omega_{1,0}^2 = g K_{1,0} \tanh(K_{1,0} H) = (g \pi / L_x) \tanh(\pi H / L_x), \quad (4.158)$$

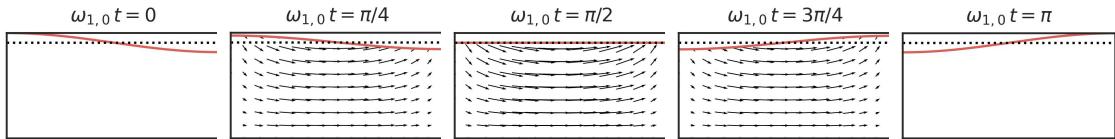


FIGURE 4.8: Flow in the  $(m, n) = (1, 0)$  seiche mode, with panels corresponding to an increment of  $\omega_{1,0} t = \pi/4$ . Note that there is no flow at times  $\omega_{1,0} t = 0$  and  $\omega_{1,0} t = \pi$ . Units are arbitrary.

which has the shallow water limit ( $H/L_x \rightarrow 0$ )

$$\omega_{1,0}^2 \approx g H (\pi/L_x)^2. \quad (4.159)$$

As the winds relax, the seiche mode oscillates according to the angular frequency  $\omega_{1,0}^2$ . We depict such oscillations in Figure 4.8 over one-half a period.

### 4.8.3 Further study

More discussion of seiche modes can be found in Section 12.5 of [Cushman-Roisin and Beckers \(2011\)](#), Section 1.6.4 of [Brown \(1999\)](#), and Chapter 10 of [Neumann and Pierson \(1966\)](#).

## 4.9 Wave packets of surface gravity waves

In this section we derive expressions for surface gravity wave packets moving over a flat bottom domain, specializing the general approach considered in Section 1.6. It is sufficient to focus on the free surface, though note that the free surface packets are associated with packets of the velocity potential and velocity field, with amplitudes related by equations (4.121b)-(4.121f). We apply the wave packet technology from Sections 1.6 and 1.7 to write a free surface wave packet as

$$\eta(\mathbf{x}, t) = \frac{1}{(2\pi)^2} \int \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} d\mathbf{k}. \quad (4.160)$$

In this equation,  $\mathbf{k}$  is the horizontal wavevector, the wavevector integral extends over all of the horizontal wavevector space, and the amplitude function,  $\mathcal{A}(\mathbf{k})$ , is generally complex and peaked around a wavevector,  $\mathbf{k}_0$ . The surface gravity wave dispersion relation (4.120) specifies the angular frequency,  $\omega = \varpi(\mathbf{k})$ , as a function of the wavevector. Since the surface gravity wave dispersion relation is a function only of the wavevector magnitude, it satisfies

$$\varpi(\mathbf{k}) = \varpi(-\mathbf{k}) = \varpi(|\mathbf{k}|). \quad (4.161)$$

This property affects some simplification relative to the general discussion of wave packets in Section 1.6.

### 4.9.1 Initializing the packet and dispensing with conjugate symmetry

To initialize the packet we consider the possibility of initializing the free surface,  $\eta(\mathbf{x}, t = 0)$ , and/or its time derivative,  $\partial_t \eta(\mathbf{x}, t = 0)$ . As shown here, by allowing for either possibility we must dispense with the conjugate symmetry property (??) otherwise assumed for the amplitudes of wave packets in Sections 1.6 and 1.7.

Given that the free surface height,  $\eta$ , is a real field, we can write its packet as one-half the

sum of the real and imaginary parts of the integral (4.160) so that

$$\eta(\mathbf{x}, t) = \frac{1}{2(2\pi)^2} \int \left[ \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} + \left( \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} \right)^* \right] d\mathbf{k} \quad (4.162a)$$

$$= \frac{1}{2(2\pi)^2} \int \left[ \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} + \mathcal{A}^*(\mathbf{k}) e^{-i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} \right] d\mathbf{k} \quad (4.162b)$$

$$= \frac{1}{2(2\pi)^2} \int \left[ \mathcal{A}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{x} - \varpi(\mathbf{k})t]} + \mathcal{A}^*(-\mathbf{k}) e^{-i[-\mathbf{k} \cdot \mathbf{x} - \varpi(-\mathbf{k})t]} \right] d\mathbf{k} \quad (4.162c)$$

$$= \frac{1}{2(2\pi)^2} \int e^{i\mathbf{k} \cdot \mathbf{x}} \left[ \mathcal{A}(\mathbf{k}) e^{-i\varpi(\mathbf{k})t} + \mathcal{A}^*(-\mathbf{k}) e^{i\varpi(\mathbf{k})t} \right] d\mathbf{k}, \quad (4.162d)$$

where the second equality made use of the associative property (??) of the complex conjugate operation, and the final equality used the property (4.161) of the surface gravity wave dispersion relation. We can compute the time derivative of the wave packet by differentiating equation (4.162d) to have

$$\partial_t \eta(\mathbf{x}, t) = \frac{i}{2(2\pi)^2} \int \varpi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} \left[ -\mathcal{A}(\mathbf{k}) e^{-i\varpi(\mathbf{k})t} + \mathcal{A}^*(-\mathbf{k}) e^{i\varpi(\mathbf{k})t} \right] d\mathbf{k}. \quad (4.163)$$

It is now clear why we must dispense with the conjugate symmetry property (??), which says that  $\mathcal{A}^*(-\mathbf{k}) = \mathcal{A}(\mathbf{k})$ . Namely, if conjugate symmetry holds, then  $\partial_t \eta(\mathbf{x}, t = 0) = 0$ . If we instead choose conjugate anti-symmetry, whereby  $\mathcal{A}^*(-\mathbf{k}) = -\mathcal{A}(\mathbf{k})$ , then  $\eta(\mathbf{x}, t = 0) = 0$ . So to enable a nonzero initial free surface and/or a nonzero free surface time derivative requires an amplitude function that does not satisfy conjugate symmetry. If one is physically motivated to choose just one of these initial conditions, then either conjugate symmetry or conjugate anti-symmetry is available. For the current analysis, we do not constrain the amplitudes. Even so,  $\eta$  remains a real function given that it is written in equation (4.162a) as the sum of a complex number plus its complex conjugate.

## 4.9.2 Expressions for the amplitude function

Setting  $t = 0$  in equation (4.160) for the free surface and equation (4.163) for its time derivative leads to

$$\eta(\mathbf{x}, 0) = \frac{1}{2(2\pi)^2} \int e^{i\mathbf{k} \cdot \mathbf{x}} [\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k})] d\mathbf{k} \quad (4.164a)$$

$$\partial_t \eta(\mathbf{x}, 0) = \frac{i}{2(2\pi)^2} \int \varpi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} [-\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k})] d\mathbf{k}. \quad (4.164b)$$

Inverting these Fourier integral yields

$$[\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k})]/2 = \int \eta(\mathbf{x}, 0) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} \quad (4.165a)$$

$$i\varpi(\mathbf{k}) [-\mathcal{A}(\mathbf{k}) + \mathcal{A}^*(-\mathbf{k})]/2 = \int \partial_t \eta(\mathbf{x}, 0) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}, \quad (4.165b)$$

which then provide expressions for the amplitude functions in terms of the prescribed initial conditions

$$\mathcal{A}(\mathbf{k}) = \int [\eta(\mathbf{x}, 0) + (i/\varpi) \partial_t \eta(\mathbf{x}, 0)] e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} \quad (4.166a)$$

$$\mathcal{A}^*(-\mathbf{k}) = \int [\eta(\mathbf{x}, 0) - (\mathrm{i}/\varpi) \partial_t \eta(\mathbf{x}, 0)] e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}} d\mathbf{x}. \quad (4.166b)$$

Note that the  $\mathbf{x}$ -integral extends over all of the horizontal  $\mathbf{x}$ -space. As discussed in Section 4.9.1, we find that the amplitude function indeed does not satisfy conjugate symmetry

$$\mathcal{A}(\mathbf{k}) \neq \mathcal{A}^*(-\mathbf{k}). \quad (4.167)$$

### 4.9.3 Wave packet in terms of a propagator function

Making use of the amplitude functions (4.166a) and (4.166b) in the integral expansion (4.162d) leads to the expression for the free surface height

$$\eta(\mathbf{x}, t) = \frac{1}{2(2\pi)^2} \int \int e^{i\mathbf{k}\cdot(\mathbf{x}-\boldsymbol{\xi})-i\varpi t} [\eta(\boldsymbol{\xi}, 0) + (\mathrm{i}/\varpi) \partial_t \eta(\boldsymbol{\xi}, 0)] d\mathbf{k} d\boldsymbol{\xi} \quad (4.168)$$

$$+ \frac{1}{2(2\pi)^2} \int \int e^{i\mathbf{k}\cdot(\mathbf{x}-\boldsymbol{\xi})+i\varpi t} [\eta(\boldsymbol{\xi}, 0) - (\mathrm{i}/\varpi) \partial_t \eta(\boldsymbol{\xi}, 0)] d\mathbf{k} d\boldsymbol{\xi}. \quad (4.169)$$

Rearrangement then renders the tidy expression

$$\eta(\mathbf{x}, t) = \int [\partial_t G(\mathbf{x} - \boldsymbol{\xi}, t) \eta(\boldsymbol{\xi}, 0) + G(\mathbf{x} - \boldsymbol{\xi}, t) \partial_t \eta(\boldsymbol{\xi}, 0)] d\boldsymbol{\xi}, \quad (4.170)$$

where we introduced the propagator function and its time derivative

$$G(\mathbf{x}, t) = \frac{1}{(2\pi)^2} \int e^{i\mathbf{k}\cdot\mathbf{x}} \frac{\sin((\varpi(|\mathbf{k}|) t))}{(\varpi(|\mathbf{k}|))} d\mathbf{k} \quad \text{and} \quad \partial_t G(\mathbf{x}, t) = \frac{1}{(2\pi)^2} \int e^{i\mathbf{k}\cdot\mathbf{x}} \cos(\varpi(|\mathbf{k}|) t) d\mathbf{k}. \quad (4.171)$$

As written, the function,  $G(\mathbf{x}, t)$ , and its time derivative encapsulate all the temporal behavior of the wave packet, acting to propagate the initial conditions,  $\eta(\boldsymbol{\xi}, 0)$  and  $\partial_t \eta(\boldsymbol{\xi}, 0)$ , forward in time. As noted in Section 1.6. equation (4.170) is quite elegant since  $G(\mathbf{x}, t)$  is independent of details of the initial conditions. In this manner it acts like a [Green's function](#).

### 4.9.4 Further study

Elements of this section follow from Section 55 of [Fetter and Walecka \(2003\)](#).

## 4.10 Capillary-gravity waves

As studied in Section ??, surface tension exists at the interface between two media, such as a liquid and solid or liquid and gas. In this section we consider the effects from surface tension on surface waves, with surface tension providing the means to realize *capillary waves*.

### 4.10.1 Pressure jump across the air-sea surface

As studied in Section ??, surface tension on an interface between two fluids leads to a pressure jump across the interface. That is, the pressure on one side of the interface is different from the pressure on the other side. The pressure jump is given by the Young-Laplace formula (??), which when applied to the ocean free surface renders

$$p_a - p_{ocn} = \gamma \nabla^2 \eta \implies p_{ocn} = p_a - \gamma \nabla^2 \eta. \quad (4.172)$$

In this equation,  $\gamma > 0$  is the surface tension (dimensions of force per length = M T<sup>-2</sup>),  $p_a$  is the pressure on the atmospheric side of the free surface, and  $p_{ocn}$  is the pressure on the ocean side of the free surface. To help remember signs for the pressure jump, note that the Young-Laplace formula (4.172) says that pressure on the concave side of the interface is higher than on the convex side. For example, if the free surface extends upward then  $p_{ocn} - p_a > 0$  since the ocean is on the concave side and so it has the higher pressure. This result also follows since  $\nabla^2\eta < 0$  for an upward extension, which leads to a local free surface maximum.<sup>14</sup>

#### 4.10.2 Dynamic boundary condition with surface tension

We can continue to apply Bernoulli's theorem even in the presence of surface tension to determine the dynamic boundary condition. Hence, we proceed as in Section 4.3.2 to evaluate the Bernoulli potential at the free surface. Now, however, it is important to specify which side of the free surface we evaluate the Bernoulli potential. Being interested in ocean waves, we evaluate the Bernoulli potential on the ocean side, in which case equation (4.84) takes on the form

$$g\eta + \mathcal{K} - \partial_t\Psi = -p_{ocn}/\rho = -(p_a - \gamma\nabla^2\eta)/\rho, \quad (4.173)$$

where the second equality follows from the Young-Laplace formula (4.172). We again assume the atmospheric pressure is a given constant that can be trivially absorbed by a gauge transformation (just like we did in Section 4.2.3 for surface gravity waves). We thus have the equation of motion

$$\partial_t\Psi = \rho^{-1} [g\rho - \gamma\nabla^2]\eta + \mathcal{K}. \quad (4.174)$$

The surface tension term,  $-(\gamma/\rho)\nabla^2\eta$ , is new relative to equation (4.84) holding for gravity waves.

#### 4.10.3 Dispersion relation for capillary-gravity waves

In Section 4.3.4 we detailed the steps needed to derive the linear equations (4.93b)-(4.93e) for surface gravity waves. Those steps also hold for capillary-gravity waves, with the only difference being a modification to the dynamic boundary condition arising from surface tension

$$\nabla^2\Psi = 0 \quad \text{irrotational and non-divergent velocity} \quad (4.175a)$$

$$\partial_t\Psi = \rho^{-1} [g\rho - \gamma\nabla^2]\eta \quad \text{linearized dynamic b.c. at } z=0 \quad (4.175b)$$

$$\partial_z\Psi = -\partial_t\eta \quad \text{linearized kinematic b.c. at } z=0 \quad (4.175c)$$

$$\hat{\mathbf{n}} \cdot \nabla\Psi = 0 \quad \text{no-normal flow kinematic b.c. on rigid boundaries.} \quad (4.175d)$$

The analysis in Section 4.5.1 concerns the interior harmonic scalar potential as well as the kinematic boundary condition. Both of these properties hold equivalently for surface gravity waves and for capillary-gravity waves. Hence, we can write the scalar potential for capillary-gravity waves as in equations (4.111a) and (4.111b), and the corresponding free surface as equation (4.112). The dispersion relation is derived by using these expressions for  $\Psi$  and  $\eta$  in the dynamic boundary condition (4.175b), with a few lines of algebra rendering

$$\omega^2 = g|\mathbf{k}|(1 + \Upsilon)\tanh(|\mathbf{k}|H) \quad \text{capillary-gravity waves,} \quad (4.176)$$

---

<sup>14</sup>The discussion in Section ?? considered bubbles, where the Young-Laplace formula shows that surface tension causes pressure inside of a bubble (concave side) to be larger than outside the bubble.

where we introduced the non-dimensional parameter

$$\Upsilon = |\mathbf{k}|^2 \gamma / (g \rho). \quad (4.177)$$

The capillary-gravity wave dispersion relation (4.176) generalizes the gravity wave dispersion (4.120). The  $\Upsilon$  parameter provides a regime boundary where capillary waves dominate ( $\Upsilon \gg 1$ ) versus where gravity waves dominate ( $\Upsilon \ll 1$ ).

#### 4.10.4 Deep water capillary-gravity waves

The hydrostatic limit is not relevant for capillary waves since capillary waves are generally very small (as seen below) and thus do not satisfy hydrostatic scaling. Hence, we find it physically most interesting to examine the limit of deep water capillary-gravity waves, in which case  $|\mathbf{k}| H \rightarrow \infty$  so that the dispersion relation (4.176) simplifies to

$$\omega^2 = g |\mathbf{k}| (1 + \Upsilon) \quad \text{deep water capillary-gravity waves.} \quad (4.178)$$

##### Phase speed

The phase speed for deep water capillary-gravity waves is given by

$$C_p = \omega / |\mathbf{k}| = \sqrt{(g / |\mathbf{k}|) (1 + \Upsilon)} = \sqrt{g / |\mathbf{k}| + \gamma |\mathbf{k}| / \rho}, \quad (4.179)$$

which has the longwave and shortwave limits

$$C_p \approx C_p^{\text{dwg}} \quad \text{for } \Upsilon \ll 1 \quad \text{and} \quad C_p \approx \sqrt{\gamma |\mathbf{k}| / \rho} \quad \text{for } \Upsilon \gg 1, \quad (4.180)$$

where we introduced the deep water gravity wave phase speed from equation (4.123a)

$$C_p^{\text{dwg}} = \sqrt{g / |\mathbf{k}|}. \quad (4.181)$$

At both extremes the phase speed is unbounded, with gravity waves dominant for longwaves and capillary waves dominant for shortwaves.

##### Group velocity

The transport of energy within a packet of capillary-gravity waves is determined by the group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}} \omega$ , which takes the form

$$\mathbf{c}_g = \frac{\hat{\mathbf{k}} C_p^{\text{dwg}}}{2} \frac{1 + 3 \Upsilon}{(1 + \Upsilon)^{1/2}} = \frac{\hat{\mathbf{k}}}{2} \left[ \frac{g \gamma}{\rho} \right]^{1/4} \frac{1 + 3 \Upsilon}{\sqrt{\Upsilon^{1/2} + \Upsilon^{3/2}}}. \quad (4.182)$$

The group velocity has the following shortwave and longwave limits

$$\mathbf{c}_g \approx \hat{\mathbf{k}} C_p^{\text{dwg}} / 2 \quad \text{for } \Upsilon \ll 1 \quad \text{and} \quad \mathbf{c}_g \approx \frac{3}{2} \hat{\mathbf{k}} \left[ \frac{|\mathbf{k}| \gamma}{\rho} \right]^{1/2} \quad \text{for } \Upsilon \gg 1. \quad (4.183)$$

As for the phase speed, we find that the group velocity is unbounded at both extremes, with gravity waves dominant for longwaves and capillary waves dominant at shortwaves.

### Wavelength of the minimum group velocity

As seen above, there is a continuum of wavelengths for capillary-gravity waves, with gravity waves dominating for longwaves and capillary waves dominating for shortwaves. To delineate between shortwaves and longwaves, we seek the wavenumber where the group velocity is a minimum. To simplify the algebra, assume  $\mathbf{k} = k \hat{\mathbf{x}}$  so that we reach an extrema of the group velocity when

$$\frac{dc_g}{dk} = 0 \implies 3 \Upsilon_{\min}^2 + 6 \Upsilon_{\min} - 1 = 0, \quad (4.184)$$

in which case

$$\Upsilon_{\min} = 2/\sqrt{3} - 1 \approx 0.1547 \quad (4.185a)$$

$$k_{\min} = (\Upsilon_{\min} g \rho / \gamma)^{1/2} \approx 0.393 (g \rho / \gamma)^{1/2} \quad (4.185b)$$

$$(c_g)_{\min} \approx 1.086 (g \gamma / \rho)^{1/4}. \quad (4.185c)$$

For an air-water interface we take  $\gamma = 0.072 \text{ N m}^{-1} = 0.072 \text{ kg s}^{-2}$ , along with the water density of  $\rho = 10^3 \text{ kg m}^{-3}$  and gravitational acceleration  $g = 9.8 \text{ m s}^{-2}$ , thus leading to

$$k_{\min} \approx 145 \text{ m}^{-1} \quad (4.186a)$$

$$\Lambda_{\min} \approx 4.3 \times 10^{-2} \text{ m} \quad (4.186b)$$

$$(c_g)_{\min} \approx 1.7 \times 10^{-1} \text{ m s}^{-1}. \quad (4.186c)$$

Evidently, for wavelengths smaller than roughly 4 cm the surface tension effects are important in their support of capillary waves. In contrast, larger wavelengths are dominated by gravity waves.

#### 4.10.5 Comments and further study

As noted in Section ??, we can ignore the pressure jump induced by surface tension across the air-sea interface if the radius of curvature of the air-sea interface is larger than a few centimeters. Here, we also see that the boundary between gravity waves and capillary waves occurs for wavelengths of a few centimeters. It is for these reasons that capillary waves, and surface tension more generally, can be ignored when considering geophysical fluid motions with scales larger than a few centimeters. Even so, we highlight the importance of surface tension and capillary waves for the study of fundamental processes affecting air-sea exchanges of matter, energy, and momentum.

Much of the discussion in this section follows that in Section 54 of [Fetter and Walecka \(2003\)](#). We again encounter surface tension effects in Section 12.2 when studying the Rayleigh-Taylor instability in the presence of gravity and surface tension.

## 4.11 Particle trajectories and Stokes drift

We here consider the trajectories of fluid particles moving as part of a wave field, going to the next order in asymptotics beyond the leading order results from Section 4.5.6. With a spatially constant wave amplitude, a fluid particle periodically returns to its original position. However, in the presence of wave inhomogeneities, such as the surface gravity waves considered in Sections 4.3 and 4.5, fluid particles oscillate between regions where the undulation in one direction does not match that in the other direction. In effect, a fluid particle spends a bit

more time in the forward moving part of the wave crest than the backward moving part of the wave trough. This asymmetry leads to a net drift of fluid particles in the direction of the wave. Figure 4.4 illustrates particle trajectories in a deep water wave for the case where the Stokes drift is ignored, in which case the particles exhibit periodic orbits. In this section we focus on the Stokes drift, in which case the particle orbits are nearly periodic, but not exactly, so that they do not return to their initial point.

To describe fluid particle drift induced by waves requires us to distinguish between an average computed at a fixed point in space (Eulerian mean), versus along a fixed fluid particle (Lagrangian mean). Their difference defines the *Stokes correction*

$$\text{Lagrangian mean} = \text{Eulerian mean} + \text{Stokes correction}. \quad (4.187)$$

When applied to the trajectory of fluid particles, the Stokes correction is referred to as the *Stokes drift*. In linear waves, Stokes drift arises when the wave field has spatial inhomogeneities that cause a particle to sample distinct portions of the wave that lead to a net, or rectified, transport. For the examples considered in this section, where there are no boundaries, then the Eulerian mean vanishes so that the Lagrangian mean particle position equals to the Stokes drift.

In this section we introduce the basic mathematics to support equation (4.187) when applied to the fluid particle position within a plane wave. These ideas form part of the rudiments for wave-mean flow interaction theory further studied in Chapter ?? and pursued in far more detail by [Böhler \(2014\)](#).

### 4.11.1 Formulation of Stokes drift

Consider a three-dimensional particle trajectory written in Cartesian coordinates,

$$\mathbf{X}(\mathbf{a}, t) = X(\mathbf{a}, t) \hat{\mathbf{x}} + Y(\mathbf{a}, t) \hat{\mathbf{y}} + Z(\mathbf{a}, t) \hat{\mathbf{z}}. \quad (4.188)$$

In the analysis of waves, it is common to assume the material coordinate,  $\mathbf{a}$ , is the initial position of a fluid particle, which we here assume. As discussed in Section ??, the particle trajectory is determined by time integrating the particle velocity

$$\frac{\partial \mathbf{X}(\mathbf{a}, t)}{\partial t} = \mathbf{v}[\mathbf{X}(\mathbf{a}, t), t] \quad (4.189)$$

so that

$$\mathbf{X}(\mathbf{a}, t) = \mathbf{X}(\mathbf{a}, 0) + \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt'. \quad (4.190)$$

This equation is a trivial result of time integrating the particle velocity. Nonetheless, it is useful to express the content of this equation in words. It says that the position at time,  $t$ , of a fluid particle labelled by the material coordinate,  $\mathbf{a}$ , is given by the initial position of the particle,  $\mathbf{X}(\mathbf{a}, 0)$ , plus the time integrated movement of the particle following the fluid flow.

Equation (4.190) provides the trajectory, but only by knowing the velocity following the trajectory. To produce a result that is more readily computed analytically, we develop a Taylor series computed relative to the initial position of the fluid particle. In this manner we make use of the approximate expression for the particle velocity at time  $t$

$$\mathbf{v}^n[\mathbf{X}(\mathbf{a}, t), t] \approx \mathbf{v}^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla \mathbf{v}^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot [\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)] \quad (4.191a)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \frac{d\mathbf{X}(\mathbf{a}, t')}{dt'} dt' \quad (4.191b)$$

$$= v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, t'), t'] dt', \quad (4.191c)$$

where the Taylor series was truncated after terms linear in the particle displacement  $\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)$ . We emphasize two points regarding equation (4.191c).

- The velocity,  $v^n[\mathbf{X}(\mathbf{a}, 0), t]$ , is the n'th component of the fluid velocity field evaluated at the initial point of the fluid particle trajectory,  $\mathbf{X}(\mathbf{a}, 0)$ , and at the time  $t$ . That is,  $v^n[\mathbf{X}(\mathbf{a}, 0), t]$  is the Eulerian velocity evaluated at the fixed Eulerian point,  $\mathbf{X}(\mathbf{a}, 0)$ .
- What determines the accuracy of the Taylor series? A suitable non-dimensional expansion coefficient for the Taylor expansion is the ratio of the particle displacement to the wave length,  $L$ ,

$$\epsilon = \frac{|\mathbf{X}(\mathbf{a}, t) - \mathbf{X}(\mathbf{a}, 0)|}{L}. \quad (4.192)$$

This ratio is small for the small amplitude waves considered here, whereby the particle displacements are far smaller than the wavelength.

The integrand on the right hand side of equation (4.191c) is the Lagrangian velocity integrated over the time interval. To within the same order of accuracy as maintained for writing equation (4.191c), we can use the Eulerian velocity evaluated at the initial position, thus rendering

$$v^n[\mathbf{X}(\mathbf{a}, t), t] \approx v^n[\mathbf{X}(\mathbf{a}, 0), t] + \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt', \quad (4.193)$$

with rearrangement leading to

$$v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t] \approx \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt'. \quad (4.194)$$

The left hand side is the difference between the velocity following a fluid particle (the Lagrangian velocity for the moving fluid particle) from the velocity at the initial particle point (the Eulerian velocity at the initial point of the trajectory). The right hand side terms are all evaluated at the initial position,  $\mathbf{X}(\mathbf{a}, 0)$ . Furthermore, the right hand side is non-zero where the velocity at the initial position has a nonzero gradient (i.e., it is spatially inhomogeneous), with its inhomogeneity projecting onto the time integrated velocity at that point. Equation (4.194) says that the velocity following a fluid particle is modified from the velocity at its initial position if the particle moves through an inhomogeneous velocity field.

Stokes drift of fluid particles is defined as the difference of the velocities in equation (4.194) when phased averaged, which we write as<sup>15</sup>

$$(v^{(S)})^n[\mathbf{X}(\mathbf{a}, 0), t] = \langle v^n[\mathbf{X}(\mathbf{a}, t), t] - v^n[\mathbf{X}(\mathbf{a}, 0), t] \rangle \quad (4.195a)$$

$$\approx \left\langle \nabla v^n[\mathbf{X}(\mathbf{a}, 0), t] \cdot \int_0^t \mathbf{v}[\mathbf{X}(\mathbf{a}, 0), t'] dt' \right\rangle. \quad (4.195b)$$

<sup>15</sup>We introduced the phase average in Section ???. It is realized by averaging over a  $2\pi$  extent of the wave phase. It can be performed via a time average over a period, a space average over a wavelength, or a combination.

This expression holds for any arbitrary initial point in the fluid, so that we can write it in a concise Eulerian form that dispenses with trajectories

$$(v^{(S)})^n(\mathbf{x}, t) \approx \left\langle \nabla v^n(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt' \right\rangle. \quad (4.196)$$

We can draw an analogy between Stokes drift and surfing. Namely, the more a fluid particle samples larger amplitude variations in the velocity field (the gradient term), the further it drifts (the integral term).

### 4.11.2 Particle trajectories in a homogeneous wave

The expression (4.196) for the Stokes drift is general and will be specialized in Section 4.11.4 for the case of surface gravity waves. Before doing so, in this section and in Section 4.11.3 we determine particle trajectories for the traveling plane wave

$$\frac{dX}{dt} = U \sin(kx - \omega t) \quad (4.197a)$$

$$\frac{dZ}{dt} = -U \cos(kx - \omega t). \quad (4.197b)$$

We here set the wavevector to  $\mathbf{k} = k \hat{\mathbf{x}}$ , so that it is purely zonal, let  $U > 0$  be the speed of the particle motion, and wrote  $X$  and  $Z$  for the Cartesian components of the particle trajectory. To simplify the mathematics we perform the analysis in a frame moving with the waves so that the phase  $kx - \omega t$  can be replaced by  $-\omega t$ , so that the particle trajectories satisfy

$$\frac{dX}{dt} = -U \sin(\omega t) \quad (4.198a)$$

$$\frac{dZ}{dt} = -U \cos(\omega t). \quad (4.198b)$$

Figure 4.4 shows a schematic of the particle trajectories appropriate for a deep water gravity wave, with the trajectories resulting from equations (4.198a) and (4.198b) directly analogous.

We start by examining particle motion in the case with a constant wave amplitude,  $U = U_0$ . Particle trajectories in this case are clockwise in the  $x$ - $z$  plane around a circle with radius  $U_0/\omega$

$$X(t) - X_0 = (U_0/\omega) [\cos(\omega t) - 1] \quad (4.199a)$$

$$Z(t) - Z_0 = -(U_0/\omega) \sin(\omega t), \quad (4.199b)$$

where the initial position at time  $t = 0$  is

$$\mathbf{X}(t = t_0) = \hat{\mathbf{x}} X_0 + \hat{\mathbf{z}} Z_0, \quad (4.200)$$

and the center of the circle is

$$\mathbf{X}_{\text{center}} = [X_0 - U_0/\omega] \hat{\mathbf{x}} + Z_0 \hat{\mathbf{z}}. \quad (4.201)$$

There is no Stokes drift since the particles return to their initial position each wave period.

### 4.11.3 Stokes drift from an inhomogeneous wave

Now consider the case of a vertically dependent wave amplitude,  $U = U(z)$ . The canonical example is where the wave amplitude decreases with depth, as for the surface gravity waves from Section 4.5. In turn, we expect there to be a fluid particle drift in the zonal direction introduced by the vertical wave inhomogeneity. This drift is a particular realization of Stokes drift.

To compute the leading order expression for the Stokes drift, expand  $U$  in a Taylor series about the initial position

$$U \approx U_0 + \sigma(Z - Z_0) \quad (4.202)$$

where the vertical shear,  $\sigma$ , has units of inverse time and is given by

$$\sigma = \left[ \frac{dU}{dZ} \right]_{Z=Z_0}. \quad (4.203)$$

The Taylor series (4.202) is valid so long as the vertical trajectories maintain the inequality

$$|\sigma| |Z - Z_0| \ll U_0, \quad (4.204)$$

which says that the vertical shear is small

$$|\sigma| \ll \frac{U_0}{|Z - Z_0|}. \quad (4.205)$$

We use the Taylor series expansion (4.202) to solve for the vertical trajectory as determined by

$$\frac{d(Z - Z_0)}{dt} = -[U_0 + \sigma(Z - Z_0)] \cos(\omega t). \quad (4.206)$$

Rearrangement leads to

$$\int_{Z_0}^Z \frac{d(Z - Z_0)}{U_0 + \sigma(Z - Z_0)} = - \int_0^t \cos(\omega t) dt. \quad (4.207)$$

The left hand side integral can be computed by changing variables

$$\Sigma = U_0 + \sigma(Z - Z_0) \implies d\Sigma = \sigma d(Z - Z_0), \quad (4.208)$$

so that equation (4.207) becomes

$$\int_{U_0}^{\Sigma} \frac{d\Sigma}{\Sigma} = -\sigma \int_0^t \cos(\omega t) dt. \quad (4.209)$$

Performing the integrals and evaluating the end points renders

$$\ln \left[ 1 + \frac{\sigma}{U_0}(Z - Z_0) \right] = -\frac{\sigma \sin(\omega t)}{\omega}, \quad (4.210)$$

which yields the exponential solution

$$1 + \frac{\sigma}{U_0}(Z - Z_0) = e^{-(\sigma/\omega) \sin(\omega t)} \implies Z - Z_0 = \frac{U_0}{\sigma} \left[ -1 + e^{-(\sigma/\omega) \sin(\omega t)} \right]. \quad (4.211)$$

The vertical particle position is seen to oscillate around its initial position  $Z_0$ .

We next consider the zonal particle position, in which case

$$\frac{d(X - X_0)}{dt} = -U_0 \left[ 1 + \frac{\sigma}{U_0}(Z - Z_0) \right] \sin(\omega t) = -U_0 e^{-(\sigma/\omega) \sin(\omega t)} \sin(\omega t), \quad (4.212)$$

where we used equation (4.211) for the vertical trajectory. To make progress, we expand the exponential assuming the ratio of inverse time scales,  $\sigma/\omega$ , is small

$$|\sigma/\omega| \ll 1. \quad (4.213)$$

In this limit, the vertical trajectory retains its unperturbed form (4.199b), and the zonal trajectory satisfies

$$\frac{d(X - X_0)}{dt} \approx -U_0 \sin(\omega t) \left[ 1 - \frac{\sigma}{\omega} \sin(\omega t) \right], \quad (4.214)$$

where we dropped terms of order  $(\sigma/\omega)^2$ . We can understand the scaling in equation (4.213) by noting that the period for the circular motion is given by

$$\tau_{\text{circle}} = 2\pi/\omega. \quad (4.215)$$

The inverse time,  $\sigma$ , introduces a time scale for the drift, defined according to

$$\tau_{\text{drift}} = 2\pi/|\sigma|. \quad (4.216)$$

A small ratio  $|\sigma/\omega|$  thus implies

$$|\sigma/\omega| = \tau_{\text{circle}}/\tau_{\text{drift}} \ll 1. \quad (4.217)$$

Hence, we are solving for the zonal trajectory in the limit where the time scale for the circular motion is small (i.e., fast oscillations around the circle) relative to the time scale for the drift (i.e., slow drift).

Returning now to the approximate zonal trajectory equation (4.214) yields

$$\frac{d(X - X_0)}{dt} = -U_0 \sin(\omega t) \left[ 1 - \frac{\sigma}{\omega} \sin(\omega t) \right] \quad (4.218a)$$

$$= -U_0 \sin(\omega t) + \frac{U_0 \sigma}{2\omega} [1 - \cos(2\omega t)], \quad (4.218b)$$

which integrates to

$$X - X_0 = \left( \frac{U_0}{\omega} \right) \left[ \cos(\omega t) - 1 - \frac{\sigma \sin(2\omega t)}{4\omega} + \frac{\sigma t}{2} \right] \quad (4.219a)$$

$$= \underbrace{\left( \frac{U_0}{\omega} \right) [\cos(\omega t) - 1]}_{\text{homogeneous}} + \underbrace{\frac{U_0 \sigma t}{2\omega}}_{\text{Stokes drift}} - \underbrace{\frac{U_0 \sigma \sin(2\omega t)}{4\omega^2}}_{\text{higher harmonic}} + \mathcal{O}(\sigma/\omega)^2. \quad (4.219b)$$

The leading order term is the homogeneous motion given by equation (4.199a). The next term is the Stokes drift, followed by a higher order harmonic and then further terms on the order of  $(\sigma/\omega)^2$ . There is no vertical Stokes drift to this order in  $(\sigma/\omega)$ , so that the Stokes drift velocity

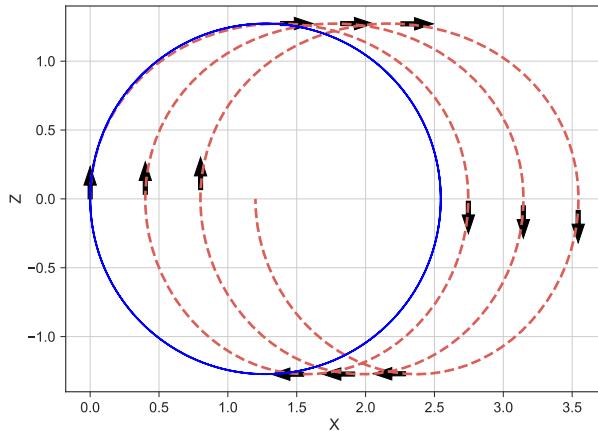


FIGURE 4.9: Example trajectories of fluid particles undergoing Stokes drift. Particle motion is clockwise in the  $x$ - $z$  plane. For homogeneous waves, there is zero Stokes drift with circular trajectories given by equations (4.199a) and (4.199b), as depicted here by the blue trajectory. There is a Stokes drift in the presence of vertical derivative in the wave amplitude and thus in the particle velocity, with trajectories for this example given by equation (4.211) for the vertical component and equation (4.219b) for the horizontal component. We set the parameters as follows:  $T = 2\pi/\omega = 60$  s,  $U_0 = 0.1$  m s $^{-1}$ , and  $\sigma = \omega/10$  and exhibit trajectories over four minutes.

is given by

$$\left[ \frac{\mathbf{X} - \mathbf{X}_0}{t} \right]^{\text{drift}} = \frac{\sigma U_0}{2\omega} \hat{\mathbf{x}}. \quad (4.220)$$

The circular motion of the parcels is therefore deformed by the zonal Stokes drift. The drift increases with larger wave amplitude ( $U_0$  large); with larger vertical shear ( $\sigma$  large); and with longer period waves ( $\omega$  small). Each of these wave properties affects the time that a fluid particle samples the wave as it moves forward versus backward, thus determining the magnitude of the Stokes drift. See Figure 4.9 for an illustration based on a particular choice for the dimensional parameters.

#### 4.11.4 Stokes drift for surface gravity waves

The velocity field for surface gravity waves, given by equations (4.113a) and (4.113b), is far more complicated than the prototypical wave considered in Sections 4.11.2 and 4.11.3. For that reason we make use of the general expression (4.196) to determine the Stokes drift velocity for surface gravity waves in a flat channel, rather than directly integrate to determine the trajectories. Hence, we need to compute terms in the following equations for surface gravity waves

$$u^{(S)}(\mathbf{x}) \approx \left\langle \nabla u(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt' \right\rangle \quad (4.221a)$$

$$v^{(S)}(\mathbf{x}) \approx \left\langle \nabla v(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt' \right\rangle \quad (4.221b)$$

$$w^{(S)}(\mathbf{x}) \approx \left\langle \nabla w(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt' \right\rangle. \quad (4.221c)$$

The components to the velocity field for surface gravity waves are given by equations (4.113a) and (4.113b), repeated here to be self-contained

$$\mathbf{u} = \frac{g \eta_0 \hat{\mathbf{k}}}{C_p} \frac{\cosh[|\mathbf{k}|(z + H)] \sin(\mathcal{P})}{\cosh(|\mathbf{k}|H)} \quad (4.222a)$$

$$w = -\frac{g \eta_0}{C_p} \frac{\sinh[|\mathbf{k}|(z + H)] \cos(\mathcal{P})}{\cosh(|\mathbf{k}|H)}, \quad (4.222b)$$

and their spatial gradients are given by

$$\nabla u = \frac{g \eta_0 k_x [\hat{\mathbf{k}} \cos(\mathcal{P}) \cosh[|\mathbf{k}|(z + H)] + \hat{\mathbf{z}} \sin(\mathcal{P}) \sinh[|\mathbf{k}|(z + H)]]}{C_p \cosh(|\mathbf{k}|H)} \quad (4.223a)$$

$$\nabla v = \frac{g \eta_0 k_y [\hat{\mathbf{k}} \cos(\mathcal{P}) \cosh[|\mathbf{k}|(z + H)] + \hat{\mathbf{z}} \sin(\mathcal{P}) \sinh[|\mathbf{k}|(z + H)]]}{C_p \cosh(|\mathbf{k}|H)} \quad (4.223b)$$

$$\nabla w = \frac{g \eta_0 |\mathbf{k}| [\hat{\mathbf{k}} \sin(\mathcal{P}) \sinh[|\mathbf{k}|(z + H)] - \hat{\mathbf{z}} \cos(\mathcal{P}) \cosh[|\mathbf{k}|(z + H)]]}{C_p \cosh(|\mathbf{k}|H)}, \quad (4.223c)$$

where we introduced the phase

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t \quad (4.224)$$

to produce more tidy expressions, and recall that

$$\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}| = (\hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y)/|\mathbf{k}| \quad (4.225)$$

is the wave direction in the horizontal plane. We also require the following time integrals

$$\int_0^t \sin(\mathbf{k} \cdot \mathbf{x} - \omega t') dt' = \omega^{-1} [\cos(\mathbf{k} \cdot \mathbf{x} - \omega t) - \cos(\mathbf{k} \cdot \mathbf{x})] \quad (4.226a)$$

$$\int_0^t \cos(\mathbf{k} \cdot \mathbf{x} - \omega t') dt' = -\omega^{-1} [\sin(\mathbf{k} \cdot \mathbf{x} - \omega t) - \sin(\mathbf{k} \cdot \mathbf{x})]. \quad (4.226b)$$

Use of these results in equations (4.221a)-(4.221c) leads to the horizontal Stokes drift velocity for surface gravity waves in a flat bottom domain, here written in three equivalent manners through use of the dispersion relation and the phase speed

$$\mathbf{v}^{(S)} = \frac{\mathbf{k} (g \eta_0)^2 \cosh[2|\mathbf{k}|(z + H)]}{2 \omega C_p^2 \cosh^2(|\mathbf{k}|H)} \quad (4.227a)$$

$$= \frac{\mathbf{k} \omega \eta_0^2 \cosh[2|\mathbf{k}|(z + H)]}{2 \sinh^2(|\mathbf{k}|H)} \quad (4.227b)$$

$$= \frac{\hat{\mathbf{k}} C_p (|\mathbf{k}| \eta_0)^2 \cosh[2|\mathbf{k}|(z + H)]}{2 \sinh^2(|\mathbf{k}|H)}. \quad (4.227c)$$

The purely horizontal nature of the Stokes drift agrees with that found for the prototypical wave in Section 4.11.3.

The ratio of the Stokes speed at  $z = 0$  to that at  $z = -H$  is given by

$$\frac{\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)}(z = 0)}{\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)}(z = -H)} = \cosh(2|\mathbf{k}|H). \quad (4.228)$$

In the deep water limit,  $k H \gg 1$ , this ratio is much greater than unity, indicating a nontrivial vertical shear in the zonal Stokes velocity. In contrast, the ratio becomes unity in the shallow water limit,  $|k| H \ll 1$ , so that there is no vertical shear to the Stokes velocity. In this case the Stokes speed takes on the depth-independent shallow water form

$$\mathbf{v}^{(S)} \cdot \hat{\mathbf{k}} \approx (C_p/2) (\eta_0/H)^2 \quad \text{for } |k| H \ll 1, \quad (4.229)$$

which has no dependence on the wavelength (other than  $|k| H \ll 1$ ). This speed is much smaller than the deep water Stokes speed at  $z = 0$ , which we can see by computing

$$\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)}(z = 0) \approx \eta_0^2 g^{1/2} |k|^{3/2}/4 \quad \text{for } |k| H \gg 1 \quad (4.230a)$$

$$\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)}(z = 0) \approx (\eta_0/H)^2 \sqrt{g H} \quad \text{for } |k| H \ll 1, \quad (4.230b)$$

so that the ratio of the deep water to shallow water Stokes speed at  $z = 0$  is

$$\frac{\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)dw}(z = 0)}{\hat{\mathbf{k}} \cdot \mathbf{v}^{(S)sw}(z = 0)} \approx |k H|^{3/2}/4 \gg 1. \quad (4.231)$$

Since the phase average vanishes at a fixed point in space, the Eulerian mean vanishes so that the Lagrangian mean in equation (4.187) is given by the Stokes drift. Hence, the presence of Stokes drift indicates that there is a net movement of matter in the direction of the wave. In particular, for the shallow water case there is a steady column of fluid moving in the direction of the wave according to the Stokes drift expression (4.229). It is important to appreciate that this example is on an unbounded domain. If the domain is bounded so that the phase averaged center of mass for the fluid is fixed in space, then the Eulerian mean exactly compensates the Stokes drift to render a zero Lagrangian mean motion.

#### 4.11.5 Comments and further study

Is Stokes drift a nonlinear phenomena? In answering this question we note that Stokes drift occurs with particle motion in linear waves, but the waves must be inhomogeneous such as the surface gravity waves studied in this chapter. Nonlinearity appears in the form of the particle-following (Lagrangian) average, as can be seen by the expression of Stokes drift given by equation (4.196)

$$(v^{(S)})^n(\mathbf{x}, t) \approx \left\langle \nabla v^n(\mathbf{x}, t) \cdot \int_0^t \mathbf{v}(\mathbf{x}, t') dt' \right\rangle. \quad (4.232)$$

The dot product of the velocity gradient with the time integrated velocity (to give the time integrated position) is nonlinear. So although the waves are linear, the Lagrangian kinematics of particle trajectories introduces nonlinearities.

Stokes drift occurs in many guises when studying the motion of fluid particles within wave fields. We revisit elements of Stokes drift in Chapter ?? when studying the rudiments of eddy-induced tracer transport. This video from Prof. Hogg at Australian National University provides an overview of the discussion in this section along with some laboratory experiments to illustrate Stokes drift. Section 10.1.1 of [Bühler \(2014\)](#) discusses Stokes corrections in the context of generalized Lagrangian mean.



## 4.12 Exercises

### EXERCISE 4.1: UNIQUENESS THEOREM FOR POTENTIAL FLOW

Assume there are two potential flows,  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , that satisfy the same boundary conditions on  $\partial\mathcal{R}$ . Show that  $\mathbf{v}_1 = \mathbf{v}_2$  throughout the domain. That is, show that there is a unique potential flow that satisfies any given boundary condition. Hint: consider the domain integrated kinetic energy contained in the difference field,  $\mathbf{V} = \mathbf{v}_1 - \mathbf{v}_2$ , and make use of the results from Section 4.2.7.

### EXERCISE 4.2: SURFACE GRAVITY WAVES ON THE VIDEO CHANNEL *Veritasium*

Veritasium is a science channel that has the following [educational video on surface ocean waves](#). However, the host makes a few minor errors in this video. Discuss the errors.

### EXERCISE 4.3: 2004 BOXING DAY TSUNAMI IN THE INDIAN OCEAN

On 26 December 2004, an earthquake near Sumatra generated a tsunami that travelled westward across the Indian Ocean that hit the Indian coastline. Assume the amplitude of the free surface  $\eta_0 = 2 \times 10^{-2}$  m, depth of the ocean  $H = 4 \times 10^3$  m, and wavelength  $\Lambda = 2 \times 10^5$  m.

- Are these approximately shallow water or deep water waves? Justify your answer with a non-dimensional number.
- What is the period of these waves?
- What is the phase averaged power (in Watts) transmitted by the waves to a coastline of length  $L = 10^3$  m? Hint: make use of the depth integrated energy flux in Section 4.5.7. Depending on whether these are deep water or shallow water waves, use the relevant approximate expressions from Section 4.5.5.

### EXERCISE 4.4: SCALE OF THE ADVECTION TERM IN SURFACE GRAVITY WAVES

In linearizing the Boussinesq equations in Section 4.3, we dropped the nonlinear advection of velocity. In this exercise we show that this assumption is self-consistent with the velocity field found within a surface gravity wave. To be specific, make use of the second form of the surface gravity wave velocity within equations (4.121e) and (4.121f).

- Derive the expression for  $(\mathbf{v} \cdot \nabla)\mathbf{u}$ .
- Derive the expression for  $(\mathbf{v} \cdot \nabla)w$ .
- Are these terms indeed smaller than the linear terms retained? Hint: in the previous parts you should find that

$$|(\mathbf{v} \cdot \nabla)w| \ll g \quad \text{and} \quad |(\mathbf{v} \cdot \nabla)\mathbf{u}| \ll g. \quad (4.233)$$

### EXERCISE 4.5: SURFACE KINEMATIC BOUNDARY CONDITION

Show that for an irrotational and non-divergent flow, the surface kinematic boundary condition (4.25) can be written

$$\partial_t \eta = -\nabla \cdot [(z - \eta) \nabla \Psi] \quad \text{at } z = \eta. \quad (4.234)$$

Hint: read Section ??.

### EXERCISE 4.6: GAUGE INVARIANCE OF KINETIC ENERGY

In Section 4.2.6 we argued that the gauge invariance of the globally integrated kinetic energy follows from equation (4.49):  $\int \mathbf{v} \cdot \hat{\mathbf{n}} dS = 0$ , which follows for the case where the free surface is a material interface so that  $\partial_t \eta = -\nabla \cdot \mathbf{U}$ . Discuss whether the kinetic energy remains gauge invariant if the upper free surface is not material, so that  $\partial_t \eta = -\nabla \cdot \mathbf{U} + Q_m/\rho$ , where  $Q_m \neq 0$  is a mass flux crossing the free surface and with this mass flux having density equal to the

constant layer density,  $\rho$ .

Hint: to write  $\mathbf{v} = -\nabla\Psi$  requires  $\nabla \times \mathbf{v} = 0$ . Do we expect the flow to remain irrotational in the presence of nonzero  $Q_m$  when the density of  $Q_m$  equals to that of the homogeneous domain?

#### EXERCISE 4.7: ENERGETICS OF DEPTH INTEGRATED FLOW IN FULL NONLINEAR THEORY

In Section 4.4.4 we computed the energy budget for the depth integrated flow within the linearized theory. Derive the mechanical energy equation for the full nonlinear theory integrated over  $-H \leq z \leq \eta$ , thus providing the nonlinear analog to the energy equation (4.107) derived for the linear theory. Do not assume the bottom is flat for the nonlinear theory. Confirm that all terms missing from the energy budget (4.107) are third order or higher. Hint: make use of results from Section 4.2.6.

#### EXERCISE 4.8: ENERGETICS OF CAPILLARY-GRAVITY WAVES

In Section 4.4.4 we derived equations for the energetics of depth integrated linearized flow in absence of surface tension. Extend that discussion to include surface tension as discussed in Section 4.10. Specifically, derive the equation for the time tendency of the depth integrated sum of the kinetic plus gravitational potential energy. Interpret the extra term arising from surface tension. Hint: start at a point in the derivation from Section 4.4.4 that is valid whether surface tension is present or not. Thereafter, make use of the dynamic boundary condition that includes surface tension.

#### EXERCISE 4.9: STOKES DRIFT FOR ONE-DIMENSIONAL MONOCHROMATIC WAVE

Consider a one-dimensional monochromatic longitudinal wave with velocity

$$u = u_0 \sin(kx - \omega t), \quad (4.235)$$

where  $u_0$  is the wave amplitude,  $k = 2\pi/\Lambda > 0$  the wave number,  $\Lambda$  the wavelength,  $\omega = 2\pi/T > 0$  the angular frequency,  $T$  the wave period, and  $C_p = \omega/k = \Lambda/T$  is the phase speed. A longitudinal wave is one whose particle motions are parallel to the wave vector, which in this exercise are both in the  $\hat{x}$  direction. Determine the wave period averaged Stokes velocity to first order accuracy in the small parameter

$$\epsilon = u_0/C_p = u_0 k/\omega = u_0 T/\Lambda, \quad (4.236)$$

with this parameter the ratio of the wave amplitude to wave speed, or equivalently the ratio of the length scale of particle displacements to the wavelength. Hint: make use of the general result given by equation (4.196).





## Chapter 5

# INERTIAL WAVES ON THE $f$ -PLANE

Inertial waves are dispersive waves that arise from the Coriolis acceleration. Fluid particles within inertial waves exhibit the anti-cyclonic *inertial oscillations* studied in Section ???. We study inertial waves in an inviscid, homogeneous (constant density),  $f$ -plane fluid, and examine traveling plane inertial waves in an unbounded domain as well as forced inertial waves in both unbounded domains and vertically bounded domains. Inertial waves provide the mechanism for vertical stiffness exhibited by the *Taylor-Proudman effect* (Section ??), and we discuss that role in this chapter.

### READER'S GUIDE TO THIS CHAPTER

We assume a working knowledge of fluid mechanics in a rotational reference frame as studied in Chapter ??. Our study of inertial waves was inspired by Section 9.2 of [Davidson \(2015\)](#), who emphasizes the role of inertial waves in forming vertically stiff structures in rapidly rotating turbulent flows, as well as Section 2.2 of [Stern \(1975\)](#), the Epilogue of [Lighthill \(1978\)](#), and the concise review by [Mory \(1992\)](#). For a visualization of inertial waves, refer to the 18 minute mark from the [rotating tank experiments of Prof. Fultz](#), which illustrates inertial oscillations within a bounded rotating homogeneous fluid. We again encounter inertial waves in Section 9.8 when studying free inertia-gravity waves, in which case the limit of zero buoyancy frequency reduces to the inertial waves of this chapter.<sup>a</sup>

<sup>a</sup>Some authors refer to Rossby waves as a type of inertial wave, since Rossby waves also owe their existence to the Coriolis acceleration. However, we consider *inertial waves* to be those waves occurring on an  $f$ -plane, where Rossby waves do not occur. We study Rossby waves in Chapters 6, 7, and 14.

<b>5.1</b>	<b>Equations for a uniformly rotating homogeneous fluid</b>	<b>148</b>
5.1.1	Velocity equation	148
5.1.2	Vorticity equation	149
5.1.3	Energy equations	149
<b>5.2</b>	<b>Linearized equations</b>	<b>150</b>
5.2.1	Linearized velocity equation	150
5.2.2	Forced oscillator equation for the velocity	150
5.2.3	Forced oscillator equation for $\nabla_h \cdot \mathbf{u} = -\partial_z w$	151
5.2.4	Wave equation for the vertical velocity	151
5.2.5	Linearized vorticity equation	151
5.2.6	The inertial wave equation for general $\Omega$	152
<b>5.3</b>	<b>Plane inertial waves</b>	<b>152</b>
5.3.1	Dispersion relation for inertial waves	152
5.3.2	Interpreting the dispersion relation	154

5.3.3	Group velocity . . . . .	155
5.3.4	Polarization relations for the velocity components . . . . .	156
5.3.5	Phase averaged kinetic energy . . . . .	156
5.3.6	High frequency inertial waves with $\omega \approx  f $ . . . . .	157
5.3.7	Low frequency inertial waves with $\omega \approx 0$ . . . . .	158
5.4	<b>Radially symmetric high frequency inertial waves</b> . . . . .	158
5.4.1	Qualitative presentation . . . . .	159
5.4.2	Radially symmetric inertial oscillations . . . . .	159
5.5	<b>Low frequency inertial waves and vertical stiffening</b> . . . . .	160
5.5.1	Slowly oscillating disk . . . . .	160
5.5.2	Inertial waves from a moving sinusoidal boundary . . . . .	161
5.5.3	Stationary wave solution . . . . .	162
5.5.4	Vertically coherent motion . . . . .	163

---

## 5.1 Equations for a uniformly rotating homogeneous fluid

We study inertial motions in a homogeneous and unbounded inviscid fluid on an  $f$ -plane. To start that study, we here formulate the nonlinear equations for the fluid motion and follow up in Section 5.2 with the linearized equations. Recall that we also studied motion of surface waves in a homogeneous fluid in Chapter 4. One key distinction is that here we are concerned with motion away from any boundary, including a free surface. Additionally, we now work on the  $f$ -plane.

### 5.1.1 Velocity equation

A homogeneous inviscid fluid on the  $f$ -plane is governed by the momentum equation

$$[\partial_t + (\mathbf{v} \cdot \nabla)] \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\nabla p/\rho - g \hat{\mathbf{z}}, \quad (5.1)$$

where  $g$  is the effective gravitational acceleration that arises from central gravity and planetary centrifugal (Section ??), and

$$\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}} = (f/2) \hat{\mathbf{z}} \quad (5.2)$$

is the constant angular frequency of the  $f$ -plane. Recall from Section ?? that the  $f$ -plane is based on assuming a locally flat geopotential, in which case the centrifugal acceleration from the rotating reference frame is incorporated into the effective gravitational acceleration. In this manner, each point on the  $f$ -plane experiences the same effects from rotation of the reference frame.

As when formulating the Boussinesq approximation in Section ??, we find it useful to decompose pressure according to a static background hydrostatic pressure, plus a dynamical pressure

$$p = p_0 + \rho \varphi \quad \text{where} \quad dp_0/dz = -\rho g, \quad (5.3)$$

with  $\rho$  the constant density and  $p_0(z)$  the static background hydrostatic pressure that exactly balances the fluid weight. This decomposition brings the momentum equation 5.1 to the form

$$[\partial_t + (\mathbf{v} \cdot \nabla)] \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\nabla \varphi \quad (5.4)$$

There is no buoyancy since density is uniform (Chapter ??).

Furthermore, again since the density is uniform everywhere, and since we are ignoring all boundaries, there can be no hydrostatic pressure induced either by gradients of the density nor by gradients of the boundary<sup>1</sup> As a result, the dynamical pressure is fully non-hydrostatic for the system we are studying here.

### 5.1.2 Vorticity equation

In addition to the velocity equation, we make use of the vorticity equation, whose nonlinear form is derived from the vector invariant velocity equation (see Section ??)

$$\partial_t \mathbf{v} + (2\boldsymbol{\Omega} + \boldsymbol{\omega}) \times \mathbf{v} = -\nabla(\varphi + \mathcal{K}), \quad (5.5)$$

where we introduced the kinetic energy per mass and the relative vorticity

$$\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2 \quad \text{and} \quad \boldsymbol{\omega} = \nabla \times \mathbf{v}. \quad (5.6)$$

The vorticity equation (??) takes on the following form in this constant density inviscid fluid

$$[\partial_t + (\mathbf{v} \cdot \nabla)] \boldsymbol{\omega}_a = (\boldsymbol{\omega}_a \cdot \nabla) \mathbf{v} \quad \text{with} \quad \boldsymbol{\omega}_a = 2\boldsymbol{\Omega} + \boldsymbol{\omega}. \quad (5.7)$$

Noting that  $\boldsymbol{\Omega}$  is independent of space and time allows us to write the relative vorticity equation

$$\partial_t \boldsymbol{\omega} = [(2\boldsymbol{\Omega} + \boldsymbol{\omega}) \cdot \nabla] \mathbf{v}. \quad (5.8)$$

### 5.1.3 Energy equations

Taking the scalar product of  $\mathbf{v}$  with the velocity equation (5.1) yields the equation for the kinetic energy per mass,  $\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2$ , and gravitational potential energy per mass,  $\Phi = gz$  (i.e., the geopotential), and their sum (the mechanical energy per mass)

$$\frac{D\mathcal{K}}{Dt} = -\mathbf{v} \cdot \nabla p/\rho - w g \quad (5.9a)$$

$$\frac{D\Phi}{Dt} = w g \quad (5.9b)$$

$$\frac{D(\mathcal{K} + \Phi)}{Dt} = -\mathbf{v} \cdot \nabla p/\rho. \quad (5.9c)$$

Since the fluid density is uniform and constant, the gravitational potential energy decouples from the kinetic energy through use of the pressure decomposition (5.3), thus leading to

$$\frac{D\mathcal{K}}{Dt} = -\mathbf{v} \cdot \nabla \varphi = -\nabla \cdot (\mathbf{v} \varphi). \quad (5.10)$$

Hence, the kinetic energy is materially modified for flows where the velocity is misaligned with constant dynamic pressure surfaces (e.g., kinetic energy increases when flow is down the pressure gradients), or equivalently where there is a convergence of the dynamical pressure flux.

---

<sup>1</sup>Recall that for surface waves in a homogeneous fluid layer, a hydrostatic pressure is generated through undulations of the free ocean surface as given by equation (4.18).

## 5.2 Linearized equations

We here formulate the linear equations for an inviscid homogeneous fluid moving on an  $f$ -plane, with these equations providing the basis for the study of inertial waves.

### 5.2.1 Linearized velocity equation

We are interested in small amplitude fluctuations relative to a state of zero motion. Linearization of the velocity equation (5.4) occurs by dropping the self-advection term,  $(\mathbf{v} \cdot \nabla) \mathbf{v}$ , thus leading to

$$\partial_t \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v} = -\nabla \varphi. \quad (5.11)$$

With  $\boldsymbol{\Omega}$  independent of time, we see that the velocity projected onto  $\boldsymbol{\Omega}$  satisfies

$$\partial_t (\mathbf{v} \cdot \boldsymbol{\Omega}) = -\boldsymbol{\Omega} \cdot \nabla \varphi. \quad (5.12)$$

Hence, if there is no pressure gradient along the direction of the rotation axis, then the velocity of the linear flow in that direction remains constant in time. For rotation around the vertical axis,  $\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}$ , and with no vertical dynamic pressure gradient,  $\partial_z \varphi = 0$ , then the vertical velocity is static,  $\partial_t w = 0$ .

### 5.2.2 Forced oscillator equation for the velocity

Taking a time derivative on the velocity equation (5.11) and back-substituting the velocity equation leads to

$$\partial_{tt} \mathbf{v} + 4\Omega^2 \mathbf{v} - 4\boldsymbol{\Omega}(\mathbf{v} \cdot \boldsymbol{\Omega}) = 2\boldsymbol{\Omega} \times \nabla \varphi - \partial_t \nabla \varphi. \quad (5.13)$$

Now decompose the velocity according to the orientation of the rotation vector

$$\hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega}/|\boldsymbol{\Omega}| \quad (5.14a)$$

$$\mathbf{v} = \mathbf{v}_\perp + \mathbf{v}_\parallel \quad (5.14b)$$

$$\mathbf{v}_\perp = \mathbf{v} - \hat{\boldsymbol{\Omega}} (\hat{\boldsymbol{\Omega}} \cdot \mathbf{v}) \quad (5.14c)$$

$$\mathbf{v}_\parallel = \mathbf{v} - \mathbf{v}_\perp = \hat{\boldsymbol{\Omega}} (\hat{\boldsymbol{\Omega}} \cdot \mathbf{v}), \quad (5.14d)$$

which allows us to decompose equation (5.13) into an equation for the parallel velocity and one for the perpendicular velocity

$$\partial_t [\hat{\boldsymbol{\Omega}} \cdot (\partial_t \mathbf{v} + \nabla \varphi)] = 0 \quad (5.15a)$$

$$(\partial_{tt} + 4\Omega^2) \mathbf{v}_\perp = 2\hat{\boldsymbol{\Omega}} \times \nabla \varphi - \partial_t [\nabla \varphi - \hat{\boldsymbol{\Omega}} (\hat{\boldsymbol{\Omega}} \cdot \nabla \varphi)]. \quad (5.15b)$$

Equation (5.15a) says that the projection of  $\partial_t \mathbf{v} + \nabla \varphi$  onto the rotation axis remains constant in time. With  $\hat{\boldsymbol{\Omega}} = \hat{\mathbf{z}}$  then

$$\partial_t (\partial_t w + \partial_z \varphi) = 0. \quad (5.16)$$

Equation (5.15b) is a forced simple harmonic oscillator equation for motion in the plane perpendicular to the rotation axis. The natural angular frequency is  $2\Omega$  and the forcing arises from pressure gradients in the perpendicular plane. Again, with  $\hat{\boldsymbol{\Omega}} = \hat{\mathbf{z}}$ , equation (5.15b) becomes

$$(\partial_{tt} + 4\Omega^2) \mathbf{u} = 2\Omega \hat{\mathbf{z}} \times \nabla \varphi - \partial_t \nabla_h \varphi. \quad (5.17)$$

The appearance of an oscillator equation anticipates the simple harmonic oscillations of fluid particles within a linear wave.

### 5.2.3 Forced oscillator equation for $\nabla_h \cdot \mathbf{u} = -\partial_z w$

We find it useful to determine expressions for the horizontal velocity divergence in the special case of  $f = 2\Omega = 2\hat{z}\Omega$ , in which case equation (5.17) has the two components

$$(\partial_{tt} + f^2) u = -f \partial_y \varphi - \partial_{xt} \varphi \quad (5.18a)$$

$$(\partial_{tt} + f^2) v = f \partial_x \varphi - \partial_{yt} \varphi. \quad (5.18b)$$

Taking  $\partial_x$  on the first equation and  $\partial_y$  on the second, and then adding leads to

$$(\partial_{tt} + f^2) \nabla_h \cdot \mathbf{u} = -\nabla_h (\partial_t \varphi) \iff (\partial_{tt} + f^2) \partial_z w = \nabla_h (\partial_t \varphi). \quad (5.19)$$

Evidently, if the vertical velocity is a linear function of  $z$  (as in the long wave limit), then the pressure is independent of  $z$ , which in turn from equations (5.18a) and (5.18b) mean that  $\partial_z \mathbf{u} = 0$ . We make use of this property in Section 5.5.2 when discussing inertial waves forced from a moving sinusoidal boundary.

### 5.2.4 Wave equation for the vertical velocity

Taking the vertical derivative of equation (5.19) and then using equation (5.16) renders the wave equation for the vertical velocity in the case that  $\Omega = \hat{z}\Omega$

$$(\partial_{tt} \nabla^2 + f^2 \partial_{zz}) w = 0. \quad (5.20)$$

This equation provides the starting point for studies of inertial oscillations on an  $f$ -plane. In Section 5.2.6 we derive a slightly more general form of this equation for an arbitrary oriented rotation vector.

### 5.2.5 Linearized vorticity equation

Taking the curl of the linear velocity equation (5.11) eliminates the pressure gradient and yields the linear vorticity equation

$$\partial_t \boldsymbol{\omega} = 2(\boldsymbol{\Omega} \cdot \nabla) \mathbf{v}. \quad (5.21)$$

To reach this result requires the identity (??) for the curl of a cross product. We also assumed  $\boldsymbol{\Omega}$  has no spatial dependence so that all of its derivatives vanish.

The linear vorticity equation (5.21) can also be derived by linearizing the nonlinear vorticity equation (5.8) by dropping the contributions from nonlinear stretching and tilting, which are processes we studied in Section ???. So the only source for vorticity in the linear theory arises from stretching and tilting along the rotational axis. That is, from the derivative of velocity in a direction aligned with the rotational axis. With  $\boldsymbol{\Omega} = \hat{z}\Omega$ , the vorticity evolves according to

$$\partial_t \boldsymbol{\omega} = 2\Omega \partial_z \mathbf{v}, \quad (5.22)$$

in which case the horizontal vorticity components evolve according to vertical tilting whereas

the vertical component evolves according to vertical stretching<sup>2</sup>

$$\partial_t(\hat{\mathbf{x}} \cdot \boldsymbol{\omega}) = 2\Omega \partial_z u \quad \text{and} \quad \partial_t(\hat{\mathbf{y}} \cdot \boldsymbol{\omega}) = 2\Omega \partial_z v \quad \text{and} \quad \partial_t(\hat{\mathbf{z}} \cdot \boldsymbol{\omega}) = 2\Omega \partial_z w. \quad (5.23)$$

Evidently, inertial waves carry a non-zero vorticity, and that vorticity is directly generated by the rotation vector,  $\boldsymbol{\Omega}$ , in the presence of velocity gradients.

### 5.2.6 The inertial wave equation for general $\boldsymbol{\Omega}$

To develop a wave equation, we take a time derivative of the linear vorticity equation (5.21), and make use of the linearized momentum equation (5.11), thus leading to

$$\partial_{tt}\boldsymbol{\omega} = -2(\boldsymbol{\Omega} \cdot \nabla)(2\boldsymbol{\Omega} \times \mathbf{v} + \nabla\varphi). \quad (5.24)$$

To eliminate the pressure gradient we take another curl and again make use of the identity (??) to write

$$\nabla \times \boldsymbol{\omega} = \nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v} = -\nabla^2 \mathbf{v}, \quad (5.25)$$

where  $\nabla \cdot \mathbf{v} = 0$  since the fluid density is constant. We also make use of Cartesian tensor calculus from Chapter ??, remembering that  $\boldsymbol{\Omega}$  and the permutation symbol are constants, to derive the identity

$$(\nabla \times [(\boldsymbol{\Omega} \cdot \nabla)(\boldsymbol{\Omega} \times \mathbf{v})])^m = \epsilon^{mnp} \partial_n [\Omega^s \partial_s (\boldsymbol{\Omega} \times \mathbf{v})_p] \quad (5.26a)$$

$$= \epsilon^{mnp} \partial_n [\Omega^s \partial_s \epsilon_{pqr} \Omega^q v^r] \quad (5.26b)$$

$$= \epsilon^{mnp} \epsilon_{qrp} \Omega^s \Omega^q \partial_n \partial_s v^r \quad (5.26c)$$

$$= (\delta^m{}_q \delta^n{}_r - \delta^m{}_r \delta^n{}_q) \Omega^s \Omega^q \partial_n \partial_s v^r \quad (5.26d)$$

$$= \Omega^s (\Omega^m \partial_r \partial_s v^r - \Omega^n \partial_n \partial_s v^m) \quad (5.26e)$$

$$= -(\boldsymbol{\Omega} \cdot \nabla)^2 v^m, \quad (5.26f)$$

where we made use of the  $\epsilon$ -tensor identity (??) as well as  $\nabla \cdot \mathbf{v} = 0$ . We are thus led to the wave equation for inertial waves, which is satisfied separately for each of the Cartesian components of the velocity field

$$[\partial_{tt} \nabla^2 + (2\boldsymbol{\Omega} \cdot \nabla)^2] \mathbf{v} = 0 \iff \partial_{tt}(\nabla \times \boldsymbol{\omega}) = (2\boldsymbol{\Omega} \cdot \nabla)^2 \mathbf{v}. \quad (5.27)$$

The vertical component of the first equation reduces to equation (5.20) when  $\boldsymbol{\Omega} = \hat{\mathbf{z}}\Omega$ . The second expression makes use of equation (5.25) that relates the curl of the vorticity to the Laplacian of the velocity.

## 5.3 Plane inertial waves

In this section we study the physics of plane inertial waves moving in an unbounded domain.

### 5.3.1 Dispersion relation for inertial waves

The inertial wave equation (5.27) provides the starting point for developing mechanical properties of inertial waves. To develop the dispersion relation we consider a traveling plane wave solution

---

<sup>2</sup>We study vortex stretching and tilting in Section ??.

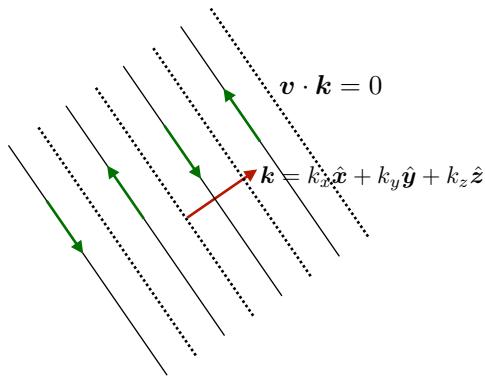


FIGURE 5.1: Illustrating the transverse nature of plane waves appearing in a homogeneous fluid with constant density, whereby  $\mathbf{v} \cdot \mathbf{k} = 0$ . The alternating solid-dotted lines depict lines of constant phase that differ by  $\pi/2$  radians so that the velocity field switches sign between every  $\pi$  radians.

of the form

$$\mathbf{v} = \tilde{\mathbf{v}} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \quad \text{and} \quad \varphi = \tilde{\varphi} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad (5.28)$$

with three-dimensional wavevector and wave direction unit vector

$$\mathbf{k} = \hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y + \hat{\mathbf{z}} k_z \quad \text{and} \quad \hat{\mathbf{k}} = (\hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y + \hat{\mathbf{z}} k_z) / |\mathbf{k}|. \quad (5.29)$$

We introduced complex amplitudes,  $\tilde{\mathbf{v}}$  and  $\tilde{\varphi}$ , for the velocity and pressure. The angular frequency,  $\omega \geq 0$ , is determined as a function of the wavevector according to the dispersion relation (5.31) derived below.<sup>3</sup> As we are considering free space waves (no boundaries), there is no preferred length scale for the inertial waves.<sup>4</sup> Furthermore, since the flow is non-divergent and the waves have a three dimensional wavevector, the constraint  $\nabla \cdot \mathbf{v} = 0$  means that the velocity of fluid particles is perpendicular to the wavevector

$$\nabla \cdot \mathbf{v} = 0 \implies \mathbf{v} \cdot \mathbf{k} = 0. \quad (5.30)$$

This orientation of wavevector and fluid velocity characterizes *transverse waves*, in which lines of constant wave phase (e.g., wave crests and troughs) are everywhere perpendicular to  $\mathbf{k}$ . We illustrate this property of transverse waves in Figure 5.1.

Plugging the wave ansatz (5.28) into the inertial wave equation (5.27) leads to the *dispersion relation*

$$\omega^2 = \varpi^2(\mathbf{k}) = (2\Omega \cdot \mathbf{k})^2 / |\mathbf{k}|^2 = (2\Omega \cdot \hat{\mathbf{k}})^2. \quad (5.31)$$

As per our previous encounter with linear waves, we recognize that the dispersion relation specifies the inertial wave's angular frequency,  $\omega$ , once the wavevector is chosen.<sup>5</sup> We illustrate the basics of plane inertial waves in Figure 5.2.

The dispersion relation (5.31) means that the angular frequency of inertial waves is directly proportional to the orientation of the wavevector relative to the rotation vector. Furthermore, it is independent of the magnitude of the wave vector. One way to display these geometric

<sup>3</sup>Be careful to distinguish the angular frequency,  $\omega$ , from the vorticity vector,  $\boldsymbol{\omega}$ .

<sup>4</sup>Inertial waves in a bounded domain will generally have quantized wavenumbers, as required by the boundary conditions.

<sup>5</sup>We write  $\omega = \varpi(\mathbf{k})$  when aiming to distinguish the angular frequency,  $\omega$ , from the function,  $\varpi$ , determining the angular frequency.

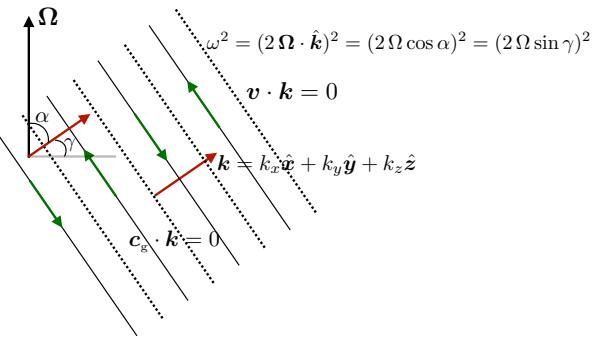


FIGURE 5.2: Illustrating the transverse nature of inertial waves, with their dispersion relation given by equation (5.31), whereby  $\omega^2 = (2\Omega \cdot \hat{k})^2$ . We show lines of constant phase separated by  $\pi/2$  radians so that the fluid particle velocity switches sign every alternative line. Evidently, the frequency of inertial waves is maximized when the wavevector,  $\mathbf{k}$ , is aligned with the rotation vector,  $\boldsymbol{\Omega}$ , in which the angles  $\alpha = 0$  and  $\gamma = \pi/2$ . In contrast, the angular frequency vanishes when the wavevector is orthogonal to  $\boldsymbol{\Omega}$ . The waves are transverse,  $\mathbf{v} \cdot \mathbf{k} = 0$ , so that fluid particle motion occurs within surfaces of constant phase. Maximum frequency waves with  $\alpha = 0$  correspond to fluid particle motion on planes orthogonal to the rotation vector, whereby the particles exhibit pure inertial oscillations. Zero frequency waves (standing inertial waves) occur with  $\alpha = \pi/2$  and  $\gamma = 0$ , whereby fluid particle motion is vertical. Vertical particle motion does not feel a Coriolis acceleration so that the wave frequency vanishes in this case. The group velocity is parallel to the fluid particle velocity since, according to equation (5.39), it satisfies  $\mathbf{c}_g \cdot \mathbf{k} = 0$ . Note that we make use of the angle  $\alpha$  in this chapter, though switch to  $\gamma = \pi/2 - \alpha$  when studying internal waves in Chapter 9.

properties is to write

$$\boldsymbol{\Omega} \cdot \hat{\mathbf{k}} = |\boldsymbol{\Omega}| \cos \alpha = |\boldsymbol{\Omega}| \sin \gamma \quad \text{with } 0 \leq \alpha \leq \pi \quad \text{and} \quad -\pi/2 \leq \gamma \leq \pi/2, \quad (5.32)$$

where  $\alpha$  is the angle between  $\boldsymbol{\Omega}$  and  $\mathbf{k}$ , whereas  $\gamma$  is the complement angle. We mostly use  $\alpha$  in this chapter, though switch to  $\gamma$  when discussing internal gravity waves and inertia-gravity waves in Chapter 9. The dispersion relation (5.31) thus takes on the particularly compact form

$$\omega^2 = (2\Omega \cos \alpha)^2 = (2\Omega \sin \gamma)^2. \quad (5.33)$$

Sweeping through the possible orientation angles,  $\alpha$ , reveals that the angular frequency (which is a non-negative number) for free space inertial waves spans the continuum range

$$0 \leq \omega \leq 2|\boldsymbol{\Omega}|. \quad (5.34)$$

Since the magnitude of the angular frequency is bounded above by  $2|\boldsymbol{\Omega}| = |f|$ , such waves are referred to as *sub-inertial*.<sup>6</sup>

### 5.3.2 Interpreting the dispersion relation

Since inertial waves are transverse,  $\mathbf{v} \cdot \mathbf{k} = 0$  (Figure 5.2), fluid particle motion associated with inertial waves is parallel to constant phase surfaces, and there is no particle motion in the direction of the wave vector.<sup>7</sup> Dynamical fields have the same geometric structure within a plane wave, so that all fields are spatially constant along a phase surface. It follows that there is

<sup>6</sup>In Section 7.8.5 we find that shallow water waves in the presence of gravity and rotation have their angular frequency bounded below by  $|f|$ , in which case these shallow water *inertia-gravity* waves are *super-inertial*.

<sup>7</sup>It is the oscillatory motion of the fluid particles in the transverse direction that constitutes motion of the phase, thus constituting the traveling wave. A fluid particle moves in the direction of a phase surface. However, the particle does not remain on a fixed phase surface since the phase travels in the  $\hat{\mathbf{k}}$  direction and yet particles have  $\mathbf{v} \cdot \hat{\mathbf{k}} = 0$  since the waves are transverse.

no spatial pressure gradient force along a constant phase surface at any particular time instance. Hence, fluid particles moving parallel to the phase surfaces feel the Coriolis acceleration arising from the projection,  $\mathbf{k} \cdot \boldsymbol{\Omega}$ , of the rotational vector along the wave vector.<sup>8</sup>

Recall that a particle moving in a rotating reference frame exhibits inertial oscillations (Section ??) when the particle does not feel any pressure forces. Evidently, fluid particle motion induced by inertial waves exhibits inertial oscillations in the phase plane orthogonal to the wavevector. If the wavevector is aligned parallel to the rotation vector, then a fluid particle feels the full extent of the Coriolis acceleration, whereby inertial oscillations have a squared angular frequency  $(2\boldsymbol{\Omega} \cdot \hat{\mathbf{k}})^2$ . This situation corresponds to a particle at either of the planetary poles.<sup>9</sup> When the wavevector is misaligned from the rotation vector, then only that portion of the rotation vector projected onto  $\hat{\mathbf{k}}$  acts to produce inertial oscillations. Finally, if the wavevector is perpendicular to the rotation vector, then the particle feels no Coriolis acceleration, just like a particle on the equator feels no planetary Coriolis acceleration.

### 5.3.3 Group velocity

As seen from our discussion of wave packets in Section 1.6, the group velocity determines the speed and direction of wave energy propagation within a packet, with the group velocity given by the wavevector gradient of the dispersion relation

$$\mathbf{c}_g = \nabla_{\mathbf{k}} \varpi = \hat{\mathbf{x}} \frac{\partial \varpi}{\partial k_x} + \hat{\mathbf{y}} \frac{\partial \varpi}{\partial k_y} + \hat{\mathbf{z}} \frac{\partial \varpi}{\partial k_z}. \quad (5.35)$$

Making use of the dispersion relation (5.31) renders

$$\mathbf{c}_g = \frac{4(\boldsymbol{\Omega} \cdot \hat{\mathbf{k}})[\boldsymbol{\Omega} - \hat{\mathbf{k}}(\boldsymbol{\Omega} \cdot \hat{\mathbf{k}})]}{\omega |\mathbf{k}|} = \frac{(\boldsymbol{\Omega} \cdot \hat{\mathbf{k}})}{|\boldsymbol{\Omega} \cdot \hat{\mathbf{k}}|} \frac{\hat{\mathbf{k}} \times (2\boldsymbol{\Omega} \times \hat{\mathbf{k}})}{|\mathbf{k}|} \quad (5.36)$$

where we made use of the vector identity (??). Since

$$\boldsymbol{\Omega} \cdot \hat{\mathbf{k}} = \pm |\boldsymbol{\Omega} \cdot \hat{\mathbf{k}}| \quad (5.37)$$

we can write the group velocity as

$$\mathbf{c}_g = \pm \frac{\hat{\mathbf{k}} \times (2\boldsymbol{\Omega} \times \hat{\mathbf{k}})}{|\mathbf{k}|}. \quad (5.38)$$

This expression leads to a particularly remarkable property of the group velocity for inertial waves

$$\mathbf{k} \cdot \mathbf{c}_g = 0. \quad (5.39)$$

Since the waves are transverse,  $\mathbf{k} \cdot \mathbf{v} = 0$ , so that the group velocity is aligned with the fluid particle velocity. The Coriolis acceleration acts perpendicular to the direction of a moving fluid particle. That orientation manifests for inertial waves via  $\mathbf{k} \cdot \mathbf{c}_g = 0$ , so that inertial waves

---

<sup>8</sup>The argument here holds either in an Eulerian reference frame, in which case we ignore the nonlinear self-advection contribution as per the linearized velocity equation (5.11), or in a Lagrangian reference frame, in which we are following a material fluid particle. In either case, motion of a fluid particle in an inertial wave, moving in a direction that parallels the constant phase surfaces, only feels the Coriolis acceleration; pressure forces are zero.

<sup>9</sup>From Figure 5.2 we note that the angle  $\alpha$  is the co-latitude, so to connect to the planetary Coriolis parameter we set  $\phi = \pi/2 - \alpha$  where  $\phi$  is the latitude.

carry energy (via the group velocity) in a direction parallel to wave crests (perpendicular to  $\mathbf{k}$ ), which is aligned with fluid particle motion. A second property of the group velocity is found by projecting it onto the direction of the rotational axis

$$\mathbf{c}_g \cdot \boldsymbol{\Omega} = \pm 2 [\Omega^2 \mathbf{k} \cdot \mathbf{k} - (\boldsymbol{\Omega} \cdot \mathbf{k})^2] / |\mathbf{k}|^3 = \pm 2 [\Omega^2 - (\omega/2)^2] / |\mathbf{k}| = \pm 2 (\Omega^2 / |\mathbf{k}|) \sin^2 \alpha, \quad (5.40)$$

where we made use of the expressions (5.31) and (5.33) for the dispersion relation.

### 5.3.4 Polarization relations for the velocity components

The velocity amplitude,  $\tilde{\mathbf{v}}$ , is generally a complex number, which allows for there to be a variety of phase shifts between the velocity components. We here determine some general relations between these amplitudes by returning to the linear velocity equation (5.11) and inserting the wave ansatz (5.28) to render

$$-i\omega \tilde{u} - f \tilde{v} = -i k_x \tilde{\varphi} \quad (5.41a)$$

$$-i\omega \tilde{v} + f \tilde{u} = -i k_y \tilde{\varphi} \quad (5.41b)$$

$$-i\omega \tilde{w} = -i k_z \tilde{\varphi}, \quad (5.41c)$$

where we set  $2\boldsymbol{\Omega} = f\hat{\mathbf{z}}$ , so that the vertical and horizontal wavenumbers are related by

$$k_z^2 = |\mathbf{k}|^2 \cos^2 \alpha \implies k_z^2 \sin^2 \alpha = (k_x^2 + k_y^2) \cos^2 \alpha. \quad (5.42)$$

Equations (5.41a), (5.41b), and (5.41c) define *polarization* relations that specify the relative phases for the velocity components and pressure. We have the freedom to choose how to reference the phases. In the following discussion of kinetic energy, we choose to measure phases relative to the pressure amplitude, which means  $\tilde{\varphi}$  is a real amplitude.

### 5.3.5 Phase averaged kinetic energy

The phase averaged kinetic energy for a plane inertial wave is given by

$$\langle \mathcal{K} \rangle = \langle \text{Re}(\mathbf{v}) \cdot \text{Re}(\mathbf{v}) \rangle / 2 = (|\tilde{u}|^2 + |\tilde{v}|^2 + |\tilde{w}|^2) / 4, \quad (5.43)$$

where we used equation (??) for the phase average of the square of a periodic function. To derive the squared amplitudes, multiply the zonal and meridional velocity equations (5.41a) and (5.41b) by  $i\omega$ , and then back-substitute

$$\tilde{u} (f^2 - \omega^2) = -\tilde{\varphi} (k_x \omega + ik_y f) \quad (5.44a)$$

$$\tilde{v} (f^2 - \omega^2) = \tilde{\varphi} (-k_y \omega + ik_x f), \quad (5.44b)$$

so that

$$|\tilde{u}|^2 = \tilde{\varphi}^2 \frac{(k_x \omega)^2 + (k_y f)^2}{(\omega^2 - f^2)^2} = \tilde{\varphi}^2 \frac{(k_x \cos \alpha)^2 + k_y^2}{f^2 \sin^4 \alpha} \quad (5.45a)$$

$$|\tilde{v}|^2 = \tilde{\varphi}^2 \frac{(k_y \omega)^2 + (k_x f)^2}{(\omega^2 - f^2)^2} = \tilde{\varphi}^2 \frac{(k_y \cos \alpha)^2 + k_x^2}{f^2 \sin^4 \alpha}, \quad (5.45b)$$

where the second equalities made use of the dispersion relation

$$\omega^2 = f^2 \cos^2 \alpha. \quad (5.46)$$

Likewise, equation (5.41c) renders the squared magnitude of the vertical velocity amplitude

$$|\tilde{w}|^2 = \tilde{\varphi}^2 k_z^2 / \omega^2. \quad (5.47)$$

Adding equations (5.45a), (5.45b), and (5.47) then leads to

$$\frac{\langle \mathcal{K} \rangle}{\tilde{\varphi}^2} = \frac{\omega^2 (k_x^2 + k_y^2)(1 + \cos^2 \alpha) + k_z^2 f^2 \sin^4 \alpha}{4 f^2 \omega^2 \sin^4 \alpha} \quad (5.48a)$$

$$= \frac{f^2 \cos^2 \alpha (k_x^2 + k_y^2)(1 + \cos^2 \alpha) + k_z^2 f^2 \sin^4 \alpha}{4 f^4 \cos^2 \alpha \sin^4 \alpha}. \quad (5.48b)$$

Use of equation (5.42) relating the horizontal and vertical wave numbers leads to

$$\langle \mathcal{K} \rangle = \frac{\tilde{\varphi}^2 k_z^2}{2 f^2 \cos^2 \alpha \sin^2 \alpha} = \frac{\tilde{\varphi}^2 (k_x^2 + k_y^2)}{2 f^2 \sin^4 \alpha}. \quad (5.49)$$

Consider the two limiting cases of  $\alpha = 0$  and  $\alpha = \pi/2$ , which we show are non-singular. A vertical inertial wave, with  $\mathbf{k} = \hat{\mathbf{z}} k_z$ ,  $\alpha = 0$ , and  $\omega^2 = f^2$ , has zero vertical particle motion,  $\tilde{w} = 0$ , so that the pressure fluctuation vanishes,  $\tilde{\varphi} = 0$ , according to equation (5.41c). As discussed in Section 5.3.6, this case corresponds to free inertial oscillations in the horizontal plane. The complement case of a horizontal inertial wave, with  $k_z = 0$ ,  $\alpha = \pi/2$ , and  $\omega = 0$ , has vertical particle motion and is discussed in Section 5.3.7.

### 5.3.6 High frequency inertial waves with $\omega \approx |f|$

The dispersion relation (5.33) leads to a maximum angular frequency magnitude for a wavevector aligned parallel or anti-parallel to the rotation axis,

$$\mathbf{k} \times \boldsymbol{\Omega} = 0 \implies \omega = 2 |\boldsymbol{\Omega}| = |f|. \quad (5.50)$$

Recall that inertial waves are transverse so that  $\mathbf{v} \cdot \mathbf{k} = 0$ . This property, in combination with  $\mathbf{k} \times \boldsymbol{\Omega} = 0$ , means that high frequency inertial waves have fluid particle motion in a plane perpendicular to the rotation axis:  $\mathbf{v} \cdot \boldsymbol{\Omega} = 0$ . With a vertical rotation vector, then the wavevector only has a vertical component,  $\mathbf{k} = k_z \hat{\mathbf{z}}$ . In this case, the plane waves propagate vertically ( $k_x = k_y = 0$ ) while fluid particle motion is restricted to the horizontal plane.

For inertial waves with  $\omega = |f| > 0$  and  $\mathbf{k} = k_z \hat{\mathbf{z}}$ , the fluid velocity amplitude relations (5.41a) and (5.41b) indicate that

$$\tilde{u} = i \tilde{v}, \quad (5.51)$$

so that the horizontal velocity of the wave is given by

$$\mathbf{u}/\tilde{u} = \hat{\mathbf{x}} e^{i(k_z z - |f| t)} + \hat{\mathbf{y}} e^{i(k_z z - |f| t - \pi/2)}, \quad (5.52)$$

which is a vertically propagating plane wave. Taking the real part renders

$$\mathbf{u}/\tilde{u} = \hat{\mathbf{x}} \cos(k_z z - |f| t) + \hat{\mathbf{y}} \sin(k_z z - |f| t). \quad (5.53)$$

With  $k_x = k_y = 0$  in the velocity equations (5.41a) and (5.41b), we see that there is no coupling between the horizontal velocity components and pressure. Hence, the motion of fluid particles reduces to inertial oscillations in the horizontal plane just as we studied for point particle motion in Section ??, whereby fluid particle motion occurs with a balance between Coriolis acceleration and centrifugal acceleration (see Figure ??). Furthermore, since there is no dependence on the horizontal position (since  $k_x = k_y = 0$ ), fluid particles move together in a coherent oscillation within each horizontal plane while the wave propagates vertically. As a check, we see that for a fixed vertical position, say  $z = 0$ , the velocity relation (5.53) for inertial waves is identical to the velocity relation (??) for particles undergoing inertial oscillations in a circle with a constant radius

$$\mathbf{u}/\tilde{u} = \hat{\mathbf{x}} \cos(|f|t) - \hat{\mathbf{y}} \sin(|f|t) \quad \text{for } z = 0 \text{ and } \omega = |f|. \quad (5.54)$$

In ocean physics, inertial waves with  $\omega \approx |f|$  are referred to as **nearinertialwaves**, which refers to their angular frequency being close to the Coriolis frequency. Since near inertial waves have their wavevector oriented close to the rotation axis, equation (5.38) indicates that they also have a vanishingly small group velocity.

### 5.3.7 Low frequency inertial waves with $\omega \approx 0$

Low frequency inertial waves occur when the wavevector is nearly perpendicular to the rotation axis

$$\boldsymbol{\Omega} \cdot \mathbf{k} \approx 0 \implies \omega/\Omega \approx 0. \quad (5.55)$$

Hence, the wave number parallel to the rotation axis is vanishingly small. For example, if the rotation axis is vertical, then low frequency inertial waves have a vanishingly small vertical wave number,

$$k_z^2 \ll k_x^2 + k_y^2. \quad (5.56)$$

Correspondingly, for the velocity vector in the form given by equation (5.28),  $\boldsymbol{\Omega} \cdot \mathbf{k} \approx 0$  means that

$$(\boldsymbol{\Omega} \cdot \nabla) \mathbf{v} = i(\boldsymbol{\Omega} \cdot \mathbf{k}) \mathbf{v} \approx 0. \quad (5.57)$$

That is, the velocity vector for low frequency inertial waves is coherent in the direction aligned with the rotation axis. Furthermore, when  $\boldsymbol{\Omega} \cdot \mathbf{k} \approx 0$ , the group velocity (5.40) has a magnitude

$$|\mathbf{c}_g| \approx 2|\Omega|/|\mathbf{k}|. \quad (5.58)$$

Evidently, low frequency inertial waves with long wavelength (small wavenumber  $|\mathbf{k}|$ ) quickly transmit their energy along the direction of the rotation axis. We return to this property in Section 5.5 when considering the Taylor-Proudman effect.

## 5.4 Radially symmetric high frequency inertial waves

To help further our understanding of inertial oscillations, consider a particularly simple case of a coherent axially symmetric fluid ring within a horizontal  $f$ -plane as depicted in Figure 5.3. Slightly perturb the ring by expanding its radius outward and then let the ring move freely. What happens? We study the motion using angular momentum arguments as well as Coriolis arguments, and assume zero pressure gradients throughout the discussion. As we show, the ring oscillates at frequency  $\omega = f$  and thus displays a canonical form of inertial oscillations.

### 5.4.1 Qualitative presentation

The outward radial perturbation gives the ring a larger moment of inertia computed relative to the rotational axis. In Section ??, we studied a fluid ring looped around the planet. As in that case, the constraint imposed by angular momentum conservation (computed relative to the rotational axis) requires the ring to rotate anti-cyclonically (clockwise if  $\Omega > 0$ ) in response to a perturbation that increases its radius.<sup>10</sup> Equivalently, radially outward motion induces a Coriolis acceleration that causes the ring to rotate clockwise (anti-cyclonic). In turn, as the ring rotates anti-cyclonically, each fluid particle within the ring experiences a radial Coriolis acceleration pointing towards the center of the ring (to the right of the particle motion). This Coriolis acceleration halts the outward perturbation and returns the ring towards a smaller radius, with the inward motion leading to a further Coriolis acceleration that causes the ring to rotate cyclonically (again, to the right of the inward particle motion). The whole process oscillates between radially outward and anti-cyclonic rotation, and radially inward and cyclonic rotation. The oscillations of a fluid ring exhibit the basic mechanism of inertial waves propagating along the rotational axis ( $\mathbf{k} \times \boldsymbol{\Omega} = 0$ ) with frequency  $f$  (Section 5.3.6).

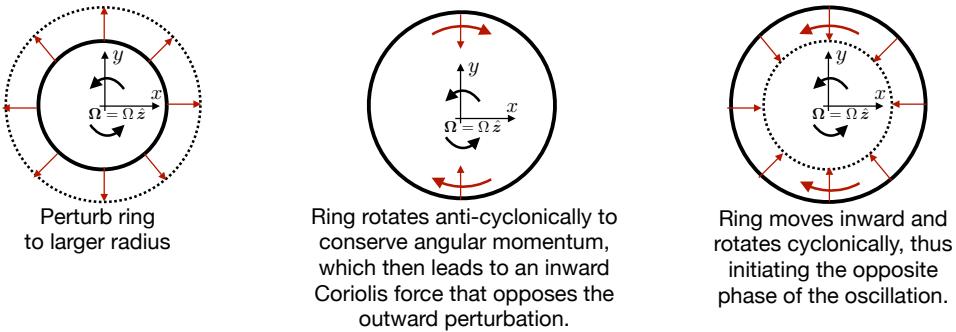


FIGURE 5.3: Schematic of inertial oscillations of an axially symmetric ring of fluid in the horizontal plane in the presence of rotation,  $\boldsymbol{\Omega} = \Omega \hat{z}$  with  $\Omega > 0$ , where the rotation axis extends through the ring center and out of the page. The left panel shows the ring perturbed outward, with this perturbation increasing the ring's moment of inertia about the vertical axis running through the center of the ring. To conserve angular momentum the ring must turn opposite to the sense of the rotating reference frame; that is, it rotates anti-cyclonically, as shown in the middle panel. As it turns anti-cyclonically the ring generates a Coriolis acceleration that points inward (to the right of the motion), thus causing the ring to oscillate back to a smaller radius (right panel), where the oscillation turns around. The physics depicted in this figure, representing a three-way balance between linear acceleration, Coriolis acceleration, and pressure gradient acceleration, is summarized by the linear velocity equation (5.11).

### 5.4.2 Radially symmetric inertial oscillations

The thought experiment in Figure 5.3 can be mathematically described by writing the linearized equation of motion using polar-coordinates from Section ???. We assume all fields are axially symmetric around any point, and all motion is two-dimensional on a horizontal plane. We also assume there is no horizontal dynamical pressure gradient acting on the fluid, so that motion is purely inertial. Decomposing the velocity equation (5.4) into radial and angular directions leads to

$$\frac{Dv^r}{Dt} - (f + \dot{\vartheta}) v^\vartheta = 0 \quad \text{and} \quad \frac{Dv^\vartheta}{Dt} + (f + \dot{\vartheta}) v^r = 0, \quad (5.59)$$

<sup>10</sup>In Figure ?? we studied the angular momentum of an axially symmetric ring of fluid around the planet. Axial symmetry means there are no zonal pressure gradients so that axial angular momentum is materially constant for the earth spanning fluid ring. Here we are making use of the same angular momentum constraint for an axially symmetric ring of fluid in a rotating  $f$ -plane.

where we introduced the polar components to the velocity

$$\mathbf{v} = (\mathbf{v} \cdot \hat{\mathbf{r}}) \hat{\mathbf{r}} + (\mathbf{v} \cdot \hat{\boldsymbol{\vartheta}}) \hat{\boldsymbol{\vartheta}} + 0 \hat{\mathbf{z}}, \quad (5.60)$$

and the radial and azimuthal unit vectors

$$\hat{\mathbf{r}} = \hat{\mathbf{x}} \cos \vartheta + \hat{\mathbf{y}} \sin \vartheta \quad \text{and} \quad \hat{\boldsymbol{\vartheta}} = -\hat{\mathbf{x}} \sin \vartheta + \hat{\mathbf{y}} \cos \vartheta. \quad (5.61)$$

The presence of  $\dot{\vartheta}$  along with the Coriolis parameter accounts for the centrifugal acceleration due to the fluid motion (as distinct from the planetary centrifugal acceleration). This is a nonlinear effect that is dropped in the linear analysis.

Linearizing the velocity equation (5.59) leads to

$$\partial_t v^r - f v^\vartheta = 0 \quad \text{and} \quad \partial_t v^\vartheta + f v^r = 0, \quad (5.62)$$

which then renders a linear oscillator equation satisfied by each velocity component

$$(\partial_{tt} + f^2) v^r = 0 \quad \text{and} \quad (\partial_{tt} + f^2) v^\vartheta = 0. \quad (5.63)$$

Assuming a monochromatic time dependence

$$v^r = \tilde{v}^r e^{-i\omega t} \quad \text{and} \quad v^\vartheta = \tilde{v}^\vartheta e^{-i\omega t} \quad (5.64)$$

leads to the dispersion relation

$$\omega^2 = f^2. \quad (5.65)$$

The linear velocity equation (5.62) ensures that the velocity components are  $\pi/2$  out of phase with

$$\tilde{v}^r = i \tilde{v}^\vartheta = \tilde{v}^\vartheta e^{i\pi/2}. \quad (5.66)$$

The motion is thus a coherent oscillation of the fluid consisting of oscillations between radial and angular motion whereby the Coriolis acceleration acts to turn the motion to the right (assuming  $f > 0$ ). There is no preferred length scale in the horizontal plane. Indeed, since the flow is horizontally non-divergent (there is no vertical fluid particle motion so  $\tilde{w} = 0$ ), and due to the assumed symmetry in the angular direction, there can be no radial dependence to the motion. That is, the radial wavenumber is zero. However, there can be vertical propagation of the waves, just as discussed in Section 5.3.6.

## 5.5 Low frequency inertial waves and vertical stiffening

We consider two more thought experiments focused on low frequency inertial waves and their connection to the vertical stiffening that arises either from a small aspect ratio flow (as in shallow water theory) or in flow in a rapidly rotating reference frame (as in the Taylor-Proudman effect). The first experiment is treated heuristically whereas we include some mathematical analysis for the second one.

### 5.5.1 Slowly oscillating disk

Imagine a slowly oscillating disk that moves in a direction aligned with the axis of rotation, such as depicted in Figure 5.4. If the oscillation frequency is much slower than the rotation

frequency,  $\omega_{\text{disk}} \ll |\Omega|$ , then the disk generates low frequency inertial waves at the frequency of the oscillating disk,  $\omega = \omega_{\text{disk}}$ . Following from our discussion in Section 5.3.7, we know that the low frequency inertial waves have a wavevector oriented perpendicular to the rotation axis,  $\boldsymbol{\Omega} \cdot \mathbf{k} = 0$ , as depicted in Figure 5.4. Since inertial waves have a group velocity that is itself perpendicular to the wavevector, the low frequency inertial waves have a group velocity parallel to the rotation axis:  $\boldsymbol{\Omega} \times \mathbf{c}_g = 0$ .

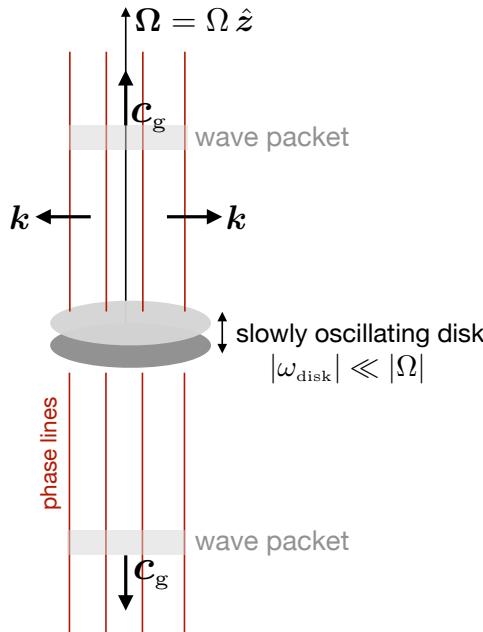


FIGURE 5.4: Schematic of inertial waves generated by a slowly oscillating disk in a rotating homogeneous fluid. The disk moves along the rotation axis (vertical axis) in small amplitude oscillations whose angular frequency is much smaller than the rotational angular frequency,  $\omega_{\text{disk}} \ll |\Omega|$ . The oscillating disk preferentially generates inertial waves whose frequency is close to  $\omega_{\text{disk}}$ ; i.e., low frequency inertial waves as discussed in Section 5.3.7. The phase lines and group velocity for these waves are parallel to the rotation axis, and the wavevector is perpendicular to the rotation axis. We depict two wave packets that send energy vertically away from the disk, with the long wave and low frequency waves having the highest magnitude for the group velocity. Since  $\boldsymbol{\Omega} \cdot \mathbf{k} = 0$ , the fluid particle velocity associated with the inertial waves is constant in the direction along the rotation axis:  $(\boldsymbol{\Omega} \cdot \nabla) \mathbf{v} = \Omega \partial_z \mathbf{v} = 0$ .

As seen by the equation (5.58) for the group velocity magnitude, information (i.e., energy) concerning the oscillating disk is most rapidly transmitted by long wavelength low frequency inertial waves. Such low frequency and long wavelength inertial waves generate fluid particle motion that is independent of the position along the rotational axis since  $(\boldsymbol{\Omega} \cdot \nabla) \mathbf{v} = 0$ . Hence, the particle motion is parallel to the rotational axis, and this motion is coherent. We conclude that inertial waves transmit information about stiffening along the rotation axis, thus providing a mechanism for the Taylor-Proudman effect from Section ?? found when the flow is strictly geostrophic.

### 5.5.2 Inertial waves from a moving sinusoidal boundary<sup>11</sup>

We here provide a bit more substance to the previous discussion by exploring the steady linear waves generated by a moving sinusoidal lower boundary, as depicted in Figure 5.5. In particular, consider a horizontally unbounded region of homogeneous fluid on an *f*-plane that is bounded

<sup>11</sup>This thought experiment comes from Section 2.2 of *Stern* (1975).

above by a rigid lid at  $z = H$  and bounded below by a moving sinusoidal boundary with vertical position

$$z = \eta_b(x, t) = h \sin(q x - U q t) = h \sin(q x - \omega_b t), \quad (5.67)$$

where we introduced the frequency set by the moving boundary

$$\omega_b = U q > 0. \quad (5.68)$$

The amplitude,  $h$ , of the lower boundary is assumed to be small compared to the wavenumber of the boundary

$$h q \ll 1, \quad (5.69)$$

which ensures that movement of the boundary generates linear waves. Because the fluid is homogeneous and on an  $f$ -plane, the undulating bottom boundary forces inertial waves. The excited inertial waves have a horizontal wavenumber given by that of the topography,  $q$ , and their vertical wavenumber is set according to the inertial wave equation (5.27). We also require the top and bottom boundary conditions to derive an expression for the vertical velocity

$$w(z = 0) = \partial_t \eta_b = -h \omega_b \cos(q x - \omega_b t) \quad \text{and} \quad w(z = H) = 0. \quad (5.70)$$

The bottom boundary condition results from linearizing the kinematic boundary condition,

$$\partial_t \eta_b + \mathbf{u} \cdot \nabla \eta_b = w \implies w(z = 0) \approx \partial_t \eta_b. \quad (5.71)$$

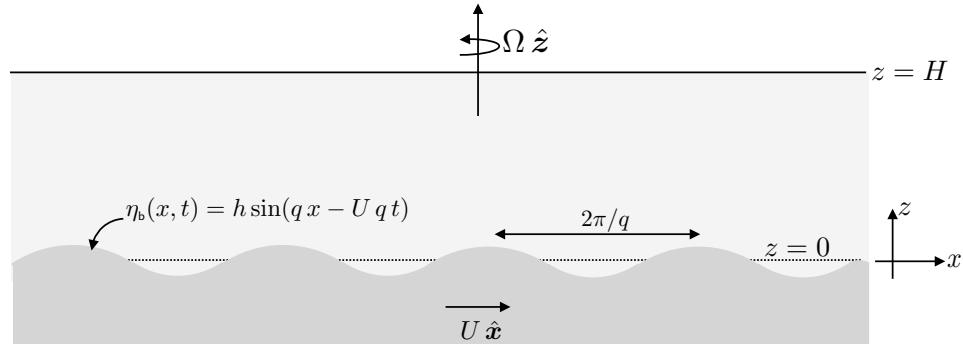


FIGURE 5.5: Schematic for the study of inertial waves generated by a moving sinusoidal lower boundary at  $z(x, t) = h_0 \sin(q x - \omega_0 t)$ , with the frequency generated by the moving boundary given by  $\omega_0 = U q$ . The region is filled with a homogenous fluid of density  $\rho$  and bounded above by a rigid lid at  $z = H$ .

### 5.5.3 Stationary wave solution

We examine waves after they have reached a steady state, and as such they are *stationary*. Given the meridional symmetry, we are only concerned with inertial waves that have a zero meridional wavenumber,  $k_y = 0$ , so that

$$\mathbf{k} = k_x \hat{x} + k_z \hat{z} = q \hat{x} + k_z \hat{z}, \quad (5.72)$$

where we set  $k_x = q$  given the forcing from the lower boundary.

Since the horizontal domain is unbounded, we consider the horizontal fluid velocity arising from a traveling plane wave

$$\mathbf{u}(\mathbf{x}, t) = \tilde{\mathbf{u}} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \quad (5.73)$$

Assuming the waves are indeed inertial with frequency,  $\omega_b$ , and horizontal wavenumber,  $q$ , the dispersion relation (5.31) renders the corresponding vertical wavenumber

$$\omega_b^2 = \frac{f^2 k_z^2}{q^2 + k_z^2} \implies k_z^2 = \frac{\omega_b^2 q^2}{f^2 - \omega_b^2}. \quad (5.74)$$

Evidently, there are two distinct regimes for the excited fluctuations: one that leads to exponential decay away from the lower boundary, with these *evanescent* waves trapped next to the bottom. The other excitation appears as inertial waves

$$\omega_b^2 > f^2 \quad \text{exponential decay away from lower boundary} \quad (5.75a)$$

$$\omega_b^2 < f^2 \quad \text{inertial waves are excited.} \quad (5.75b)$$

We only consider the case of sub-inertial forcing so that  $\omega_b^2 < f^2$ .

The structure of the vertical velocity associated this forced motion is determined by the following boundary value problem

$$[\partial_{tt} \nabla^2 + (2\boldsymbol{\Omega} \cdot \nabla)^2] w = 0 \quad (5.76a)$$

$$w(z=0) = -h \omega_b \cos(qx - \omega_b t) \quad (5.76b)$$

$$w(z=H) = 0. \quad (5.76c)$$

To find a particular solution we take the ansatz

$$w(x, t) = w(0) \Gamma(z), \quad (5.77)$$

with the non-dimensional structure function satisfying

$$\frac{d^2 \Gamma}{dz^2} + k_z^2 \Gamma = 0 \quad \text{for } 0 < z < H \quad \text{with} \quad \Gamma(0) = 1 \quad \text{and} \quad \Gamma(H) = 0, \quad (5.78)$$

which readily leads to the vertical fluid velocity

$$w(x, z, t) = w(0) \frac{\sin[k_z(H-z)]}{\sin(k_z H)} = -h \omega_b \cos(qx - \omega_b t) \frac{\sin[k_z(H-z)]}{\sin(k_z H)}. \quad (5.79)$$

#### 5.5.4 Vertically coherent motion

Consider the longwave limit in which

$$\Gamma(z) = \lim_{k_z H \rightarrow 0} \frac{\sin[k_z(H-z)]}{\sin(k_z H)} = (H-z)/H, \quad (5.80)$$

so that the vertical velocity is a linear function of  $z$

$$w(x, z, t) \approx -h \omega_b \cos(qx - \omega_b t) (H-z)/H. \quad (5.81)$$

As seen in equation (5.19), a linear vertical dependence to  $w$  means that the dynamic pressure is independent of  $z$ . Correspondingly, equations (5.18a) and (5.18b) reveal that  $\partial_z \mathbf{u} = 0$ , which means that horizontal fluid motion is vertically coherent. This motion corresponds to the shallow water fluid from Part ?? of this book. It also corresponds to the Taylor-Proudman effect discussed in Section 5.5.

To determine the non-dimensional scaling that leads to vertically coherent motion, return to equation (5.78) for the vertical structure function and introduce the non-dimensional vertical coordinate

$$\hat{z} = z/H, \quad (5.82)$$

in which equation (5.78) becomes

$$\frac{d^2\Gamma}{d\hat{z}^2} + (k_z H)^2 \Gamma = 0. \quad (5.83)$$

The vertical structure of the vertical velocity is a linear function of  $\hat{z}$  in the limit that  $(k_z H)^2 \rightarrow 0$ , which means

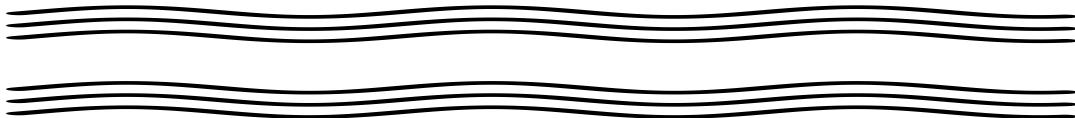
$$(k_z H)^2 \rightarrow 0 \implies \frac{(q H)^2}{(f/U q)^2 - 1} \rightarrow 0. \quad (5.84)$$

This limit can be realized if

$$q H \ll 1 \quad \text{small aspect ratio (shallow water limit)} \quad (5.85a)$$

$$U q/f \ll 1 \quad \text{small Rossby number (Taylor-Proudman limit).} \quad (5.85b)$$

Evidently, a small aspect ratio, with a Rossby number bounded away from unity, leads to vertically coherent motion even if the fluid is not in a rotating reference frame. This motion corresponds to that of shallow water fluid studied in Part ?? of this book. As emphasized in Section ??, the small aspect ratio shallow water fluid displays vertically coherent motion due to the fluid being homogeneous and hydrostatic. Alternatively, we realize vertically coherent columnar motion with a small Rossby number. This motion corresponds to the Taylor-Proudman effect for rotating fluids. It is notable that vertically coherent motion becomes more restricted (i.e., needs a smaller Rossby number) when the upper boundary moves far away,  $H \rightarrow \infty$ , thus moving to the deep water limit.



## Chapter 6

# BAROTROPIC VORTICITY WAVES

In this chapter we study *vorticity waves*, also called *Rossby waves* or *vortical modes*, which rely on the presence of a gradient in the base state potential vorticity field. We examine a particularly simple realization of Rossby waves as found in the inviscid two-dimensional non-divergent barotropic model on an unbounded  $\beta$ -plane. This model supports two general kinds of vorticity waves. One arises from the gradient of planetary vorticity (i.e.,  $\beta$ -effect), which gives rise to *planetary Rossby waves*. The second arises from gradients in the vorticity of the base flow. *Edge waves* are a particularly simple kind of vorticity waves that arise from assuming a jump in the background vorticity field.

### READER'S GUIDE TO THIS CHAPTER

We make extensive use of the horizontally non-divergent barotropic model from Chapter ??, as well as the wave kinematics from Chapter 1. Rossby waves are encountered again when discussing shallow water waves in Chapter 7, and edge waves are encountered when studying shear instability in Chapter 13 and baroclinic instability in Chapter 14.

<b>6.1</b>	<b>Loose threads</b>	<b>166</b>
<b>6.2</b>	<b>A single plane wave in the non-divergent barotropic fluid</b>	<b>166</b>
6.2.1	Transverse plane waves	167
6.2.2	Absence of inertial waves and gravity waves	167
6.2.3	Zero advection for a single plane wave	168
6.2.4	Pressure equation for a single plane wave	168
6.2.5	Velocity equation for a single plane wave	169
6.2.6	Structure of a single plane wave	169
6.2.7	Kinetic energy of a single plane wave	170
<b>6.3</b>	<b>Barotropic and non-divergent Rossby waves</b>	<b>171</b>
6.3.1	The vorticity mechanism for planetary Rossby waves	171
6.3.2	Flow relative to a zonal base state	173
6.3.3	Equations for the fluctuating vorticity and streamfunction	173
6.3.4	Rossby wave dispersion relation	174
6.3.5	Extrinsic and intrinsic angular frequency	175
6.3.6	Concerning the westward phase velocity	176
6.3.7	Stationary Rossby waves	177
6.3.8	Group velocity	178
<b>6.4</b>	<b>Geometry of planetary Rossby waves</b>	<b>178</b>
6.4.1	Group and phase velocities for planetary Rossby waves	179
6.4.2	Dispersion circle for planetary Rossby waves	179
6.4.3	Reflection of planetary Rossby waves	181

---

6.4.4	Further study . . . . .	184
<b>6.5</b>	<b>Edge waves . . . . .</b>	<b>184</b>
6.5.1	Base state and the meridionally modulated wave ansatz . . . . .	184
6.5.2	Rayleigh-Kuo equation . . . . .	185
6.5.3	The point jet . . . . .	186
6.5.4	Kinematic boundary condition at the interface . . . . .	187
6.5.5	Dynamic boundary condition at the interface . . . . .	187
6.5.6	Edge wave dispersion relation . . . . .	188
6.5.7	Further study . . . . .	189
<b>6.6</b>	<b>Exercises . . . . .</b>	<b>190</b>

---

## 6.1 Loose threads

- Reflection and tunneling as per Section  $J_I$  of [Cohen-Tannoudji et al. \(1977\)](#). Follow Isaac's notes for the problem setup.
- Hamilton/Whitham principle
- energetics
- Rossby wave particle velocity and trajectories
- Plot the dispersion relation for Rossby waves
- Rectification by Rossby waves that are radiated by a source; Section 5.4 of [McWilliams \(2006\)](#).
- Work through the Green's function problem as in [Haidvogel and Rhines \(1983\)](#) as well as Bill Young's Les Houches lectures.
- Stokes drift
- Rossby waves on a rotating planet as per continuum Lagrangian field theory
- Rossby wave packets in a non-constant background flow
- Rossby waves using spherical coordinates and spherical harmonics. See Isaac's notes.

## 6.2 A single plane wave in the non-divergent barotropic fluid

In this section we study properties of a single plane wave in an inviscid barotropic and horizontally non-divergent fluid on an unbounded beta plane. We then follow in Section 6.3 with a study of Rossby waves in this model.<sup>1</sup> Recall that the barotropic and horizontally non-divergent model was studied in Chapter ??, with the flow fully described by the vorticity equation

$$\frac{D(\zeta + f)}{Dt} = (\partial_t + \mathbf{u} \cdot \nabla)(\zeta + f) = 0, \quad (6.1)$$

where the horizontal velocity is non-divergent and so can be written in terms of a streamfunction

$$\nabla \cdot \mathbf{u} = 0 \implies \mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi \quad \text{and} \quad \zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \nabla^2 \psi. \quad (6.2)$$

---

<sup>1</sup>The edge waves studied in Section 6.5 are not plane waves, and so they do not satisfy the special properties described in this section.

As we show in this section, a single plane wave in this model exactly satisfies linear velocity and vorticity equations, with such equations summarized in Table 6.1. Since there are no nonlinear terms affecting the single plane wave, there is no need to linearize the equations of motion when studying properties of this wave. We emphasize, however, that this distinct property holds only for a single plane wave. In particular, the advection operator that vanishes for a single wave is nonzero when the flow has more than a single plane wave. Indeed, the nonlinear interactions between distinct wave modes provide the mechanism for the inverse turbulent cascade in this model (e.g., see Chapter 11 of [Vallis \(2017\)](#)).

NAME	GENERAL RELATION	SINGLE PLANE WAVE
streamfunction	$\psi$	$\psi = A \cos \mathcal{P}$
velocity	$\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi$	$\mathbf{u} = -(\hat{\mathbf{z}} \times \mathbf{k}) A \sin \mathcal{P}$
velocity tendency	$\partial_t \mathbf{u} = \hat{\mathbf{z}} \times \nabla (\partial_t \psi)$	$\partial_t \mathbf{u} = \omega (\hat{\mathbf{z}} \times \mathbf{k}) \psi$
non-divergence	$\nabla \cdot \mathbf{u} = 0$	$\mathbf{k} \cdot \mathbf{u} = 0$
relative vorticity	$\zeta = \nabla^2 \psi$	$\zeta = - \mathbf{k} ^2 \psi$
$\beta$ -plane vorticity equation	$(\partial_t + \mathbf{u} \cdot \nabla) \zeta = -v \beta$	$\partial_t \zeta = -v \beta$
pressure equation	$-\nabla \cdot (\nabla \varphi - f \nabla \psi) = \mathbf{S}^2 - \mathbf{R}^2$	$-\nabla \cdot (\nabla \varphi - f \nabla \psi) = 0$
Coriolis acceleration	$-f \hat{\mathbf{z}} \times \mathbf{u} = f \nabla \psi$	$-f A \mathbf{k} \sin \mathcal{P}$
velocity equation	$\partial_t \mathbf{u} + (f + \zeta) \hat{\mathbf{z}} \times \mathbf{u} = -\nabla (\varphi + \mathcal{K})$	$\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla \varphi$
kinetic energy equation	$D\mathcal{K}/Dt = -\nabla \cdot (\mathbf{u} \varphi)$	$\partial_t \mathcal{K} = -\nabla \cdot (\mathbf{u} \varphi)$

TABLE 6.1: Properties of the inviscid horizontally non-divergent barotropic model. The left column holds for a general flow whereas the right column holds for a single plane wave, with  $\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t$  the phase function.  $\mathbf{S}$  is the strain rate tensor, with  $\mathbf{S}^2 = S_{mn} S^{mn}$ . Likewise,  $\mathbf{R}$  is the rotation tensor, with  $\mathbf{R}^2 = R_{mn} R^{mn}$ . Both of these tensors are introduced in Section ???. This table highlights the remarkable properties of the single plane wave in the inviscid horizontally non-divergent barotropic model, in which the nonlinear terms in the velocity and vorticity equation vanish identically.

### 6.2.1 Transverse plane waves

Consider a plane wave ansatz for the streamfunction

$$\psi(\mathbf{x}, t) = A \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad \text{with} \quad \mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} \quad \text{and} \quad \mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t, \quad (6.3)$$

with  $A$  a constant amplitude. The velocity of fluid particles in the plane wave is thus given by

$$\mathbf{u} = -A (\hat{\mathbf{z}} \times \mathbf{k}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (6.4)$$

which then leads to

$$\mathbf{k} \cdot \mathbf{u} = 0. \quad (6.5)$$

The horizontally propagating plane waves are transverse, as illustrated in Figure 6.1, which arises since the flow is horizontally non-divergent. This property compares to the three dimensional transverse inertial plane waves discussed in Section 5.3.1 and depicted in Figure 5.1.

### 6.2.2 Absence of inertial waves and gravity waves

Since there is no vertical motion in the two-dimensional non-divergent barotropic model, any wavevector must be horizontal and thus perpendicular to the vertically oriented rotation vector. Hence, the inertial wave dispersion relation (5.31) only admits a zero frequency mode, which is geostrophic balance. Furthermore, gravity waves are absent from this model since gravity

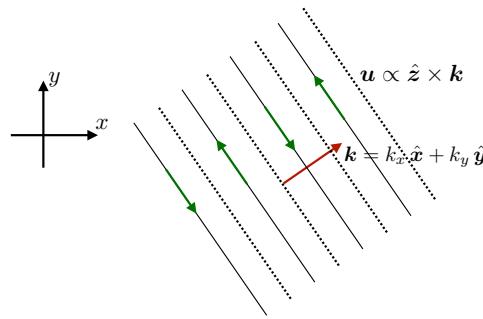


FIGURE 6.1: Illustrating the transverse nature of plane waves appearing in a horizontally non-divergent barotropic model, whereby  $\mathbf{u} \cdot \mathbf{k} = 0$ . The alternating solid-dotted lines depict lines of constant phase that differ by  $\pi/2$  radians so that the velocity field switches sign between every  $\pi$  radians. Compare this figure to Figure 5.1, which illustrates transverse plane waves in three dimensions.

waves vanish with a flow that has zero horizontal divergence (see the discussion of shallow water gravity waves in Section 7.5). We thus conclude that the two-dimensional non-divergent barotropic model has neither inertial waves nor gravity waves.

As we will see, the only wave supported by this model occurs in the presence of a background or base state vorticity gradient, such as from planetary vorticity or the vorticity of a mean flow. In the absence of vorticity gradients, there are no linear waves in the two-dimensional non-divergent barotropic model. Such is the case, for example, in non-rotating and homogeneous two-dimensional flows, which are commonly used to study two-dimensional turbulence.

### 6.2.3 Zero advection for a single plane wave

Linearizing the equations of motion is a basic step in the development of a dispersion relation. For the vorticity equation appropriate for Rossby waves, linearization means neglecting the advection of relative vorticity. However, quite remarkably, there is no advection of relative vorticity for a plane wave in the horizontally non-divergent barotropic model. That is, the advection operator,  $\mathbf{u} \cdot \nabla \zeta$ , vanishes identically when  $\mathbf{u}$  and  $\zeta$  are built from a single plane wave. To see this property, consider a traveling plane wave streamfunction

$$\psi(x, y, t) = A \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \implies \zeta = -|\mathbf{k}|^2 \psi \quad \text{and} \quad \nabla \zeta = -|\mathbf{k}|^2 \nabla \psi, \quad (6.6)$$

in which case we readily find

$$\mathbf{u} \cdot \nabla \zeta = (\hat{\mathbf{z}} \times \nabla \psi) \cdot (-|\mathbf{k}|^2 \nabla \psi) = 0. \quad (6.7)$$

Evidently, for this model the velocity of fluid particles in a plane wave are aligned parallel to surfaces of constant relative vorticity of the wave. Equivalently, this model supports no nonlinear self-interactions for a single plane wave.

### 6.2.4 Pressure equation for a single plane wave

Recall the discussion in Section ?? where we showed that pressure in the horizontally non-divergent barotropic model satisfies the Poisson equation ??

$$-\nabla \cdot (\nabla \varphi - f \nabla \psi) = 2 [(\partial_{xy} \psi)^2 - \partial_{xx} \psi \partial_{yy} \psi]. \quad (6.8)$$

Making use of the traveling plane wave (6.6) readily reveals that the nonlinear source term vanishes identically

$$(\partial_{xy}\psi)^2 - \partial_{xx}\psi \partial_{yy}\psi = 0. \quad (6.9)$$

Although the plane wave supports a nonzero strain rate tensor and a nonzero rotation tensor, their respective squares cancel identically. As a result, the pressure source from a plane wave is due only to the Coriolis acceleration

$$-\nabla \cdot (\nabla\varphi - f \nabla\psi) = -\nabla \cdot (\nabla\varphi + f \hat{z} \times \mathbf{u}) = 0. \quad (6.10)$$

Consequently,

$$\nabla\varphi + f \hat{z} \times \mathbf{u} = -\hat{z} \times \nabla\lambda, \quad (6.11)$$

where  $\lambda(x, y, t)$  is a gauge function. In Section (6.2.5) we show that  $\nabla\lambda = \nabla(\partial_t\psi)$ .

### 6.2.5 Velocity equation for a single plane wave

So what velocity equation does a single plane wave satisfy for the inviscid horizontally non-divergent barotropic model? To answer that question we make use of the vector-invariant velocity equation (??) written in the form

$$\partial_t\mathbf{u} + f \hat{z} \times \mathbf{u} + \nabla\varphi = -(\nabla\mathcal{K} + \hat{z} \times \zeta \mathbf{u}) \quad \text{with } \mathcal{K} = \mathbf{u} \cdot \mathbf{u}/2. \quad (6.12)$$

Introducing the streamfunction brings the right hand side terms into the form

$$\nabla\mathcal{K} + \hat{z} \times \zeta \mathbf{u} = \hat{x}(\partial_y\psi \partial_{xy}\psi - \partial_x\psi \partial_{yy}\psi) + \hat{y}(\partial_x\psi \partial_{xy}\psi - \partial_y\psi \partial_{xx}\psi) \quad (6.13a)$$

$$= \hat{x}[\hat{z} \cdot (\partial_y\nabla\psi \times \nabla\psi)] + \hat{y}[\hat{z} \cdot (\partial_x\nabla\psi \times \nabla\psi)]. \quad (6.13b)$$

We readily find that each of these nonlinear terms vanishes when the streamfunction is given by the single plane wave function (6.3)

$$\partial_y\nabla\psi \times \nabla\psi = A^2 k_y(\mathbf{k} \times \mathbf{k}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) = 0 \quad (6.14a)$$

$$\partial_x\nabla\psi \times \nabla\psi = A^2 k_x(\mathbf{k} \times \mathbf{k}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) = 0. \quad (6.14b)$$

Evidently, a single plane wave in the horizontally non-divergent barotropic model exactly satisfies the linear velocity equation

$$\partial_t\mathbf{u} + f \hat{z} \times \mathbf{u} = -\nabla\varphi. \quad (6.15)$$

Comparing to the result from the pressure equation (6.11) reveals that  $\nabla\lambda = \nabla(\partial_t\psi)$ . Furthermore, recall that we made no assumptions about  $f$  in arriving at equation (6.15), so that  $f$  can be a function of latitude as per the  $\beta$ -plane. Indeed,  $f$  can be an arbitrary function of space,  $f(x, y)$ . Finally, equation (6.15) indicates that a stationary plane wave ( $\omega = 0$ ) is in exact geostrophic balance. We further discuss this result in Section 6.2.6 when studying the structure of a single plane wave.

### 6.2.6 Structure of a single plane wave

In Table 6.1 we summarize the properties satisfied by a single plane wave in the horizontally non-divergent barotropic model. In particular, note that the pressure gradient has been decomposed into linearly independent directions parallel to the wave,  $\hat{\mathbf{k}}$ , and perpendicular to the wave,

TERM	$\mathcal{P} = 0$	$\mathcal{P} = \pi/2$	$\mathcal{P} = \pi$	$\mathcal{P} = 3\pi/2$	$\mathcal{P} = 2\pi$
$\mathbf{u}$	0	$-A(\hat{\mathbf{z}} \times \mathbf{k})$	0	$A(\hat{\mathbf{z}} \times \mathbf{k})$	0
$-f\hat{\mathbf{z}} \times \mathbf{u}$	0	$-Af\mathbf{k}$	0	$Af\mathbf{k}$	0
$-\nabla\varphi$	$A\omega(\hat{\mathbf{z}} \times \mathbf{k})$	$Af\mathbf{k}$	$-A\omega(\hat{\mathbf{z}} \times \mathbf{k})$	$-Af\mathbf{k}$	$A\omega(\hat{\mathbf{z}} \times \mathbf{k})$
$\partial_t \mathbf{u}$	$A\omega(\hat{\mathbf{z}} \times \mathbf{k})$	0	$-A\omega(\hat{\mathbf{z}} \times \mathbf{k})$	0	$A\omega(\hat{\mathbf{z}} \times \mathbf{k})$

TABLE 6.2: Values for plane wave terms in the velocity equation as the phase,  $\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t$ , moves from 0 to  $2\pi$ . The time tendency,  $\partial_t \mathbf{u}$ , is always perpendicular to the wavevector, which accords with the transverse nature of the wave. Also note that the time tendency is  $\pi/2$  out of phase with the velocity itself. When the time tendency vanishes, then the pressure gradient and Coriolis accelerations are in geostrophic balance, with this balance occurring every  $\pi$  radians.

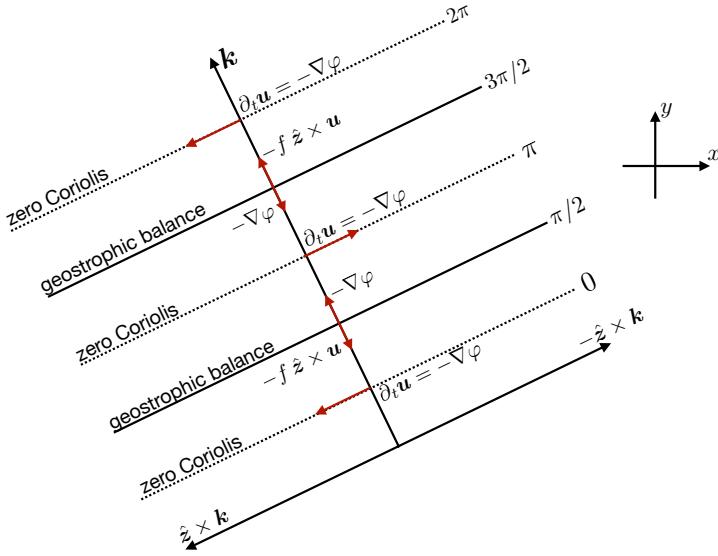


FIGURE 6.2: Schematic of the terms appearing in a horizontally propagating plane wave in the horizontally non-divergent barotropic model. We display terms as a function of the phase,  $\mathcal{P} = (\mathbf{k} \cdot \mathbf{x} - \omega t)$ , whose value is shown for  $\mathcal{P} = 0, \pi/2, \pi, 3\pi/2, 2\pi$ . The terms correspond to those given in Table 6.2. Note the oscillation between geostrophic balance, with zero time tendency, with a phase where there is a downgradient acceleration (i.e. ageostrophic motion). Also note that the pressure gradient rotates in a clockwise direction when the phase increases.

$$\hat{\mathbf{z}} \times \hat{\mathbf{k}}$$

$$-\nabla\varphi = A [\omega(\hat{\mathbf{z}} \times \mathbf{k}) \cos \mathcal{P} + f \mathbf{k} \sin \mathcal{P}]. \quad (6.16)$$

Table 6.2 considers the values for each term in the velocity equation as the phase moves around the unit circle, and Figure 6.2 provides a schematic. Evidently, the transverse plane waves oscillate between geostrophic balance, with  $\partial_t \mathbf{u} = 0$ , and pressure driven tendency, where  $\partial_t \mathbf{u} = -\nabla\varphi$ .

### 6.2.7 Kinetic energy of a single plane wave

Taking the scalar product of the velocity with the velocity equation (6.12) leads to the kinetic energy equation (??)

$$\frac{D\mathcal{K}}{Dt} = -\nabla \cdot (\mathbf{u} \varphi). \quad (6.17)$$

Likewise, taking the scalar product with the linear velocity equation (6.15) satisfied by the single plane wave renders

$$\partial_t \mathcal{K} = -\nabla \cdot (\mathbf{u} \varphi). \quad (6.18)$$

As for the velocity, we see that the kinetic energy of the plane wave experiences no advection. Furthermore, making use of the wave ansatz (6.3) leads to the kinetic energy per mass within a single plane wave

$$\mathcal{K} = (A^2 |\mathbf{k}|^2 / 2) \sin^2 \mathcal{P}, \quad (6.19)$$

thus indicating more energy in higher wavenumber waves. In contrast, the gravitational potential energy is constant given that the fluid has a uniform density and rigid lid. Hence, as the kinetic energy fluctuates within a wave, there is no exchange with potential energy. Instead, there is an exchange with the external dynamical system that affects a rigid lid on the fluid.

The phase average of the kinetic energy (6.19) is

$$\langle \mathcal{K} \rangle = A^2 |\mathbf{k}|^2 / 4. \quad (6.20)$$

Now consider the energy flux convergence, again as rendered by the single plane wave

$$-\nabla \cdot (\mathbf{u} \varphi) = -\mathbf{u} \cdot \nabla \varphi = \omega A^2 |\mathbf{k}|^2 \sin \mathcal{P} \cos \mathcal{P}, \quad (6.21)$$

where we made use of equation (6.16) for the pressure gradient. The phase average of this flux convergence vanishes

$$\langle \nabla \cdot (\mathbf{u} \varphi) \rangle = 0. \quad (6.22)$$

Evidently, the zero kinetic energy flux convergence arises since the plane wave is present throughout space, so that there is no means to converge phase averaged wave energy to any particular region. Only when there is a symmetry breaking, such as by modulating the plane wave into a localized packet, will there be a nonzero phase average flux convergence.

## 6.3 Barotropic and non-divergent Rossby waves

In this section we build on the general properties of the plane wave as developed in Section 6.2. The key new ingredient considered here is the dispersion relation that couples the wavevector to the wave angular frequency. As we see, the fluctuations are organized into *Rossby waves*, which are waves that carry a nonzero vorticity and are reliant on gradients in the background potential vorticity field.

### 6.3.1 The vorticity mechanism for planetary Rossby waves

Before developing the detailed properties of Rossby waves, we discuss the underlying mechanism for planetary Rossby waves. This discussion serves as both a motivation and guide for the mathematics to follow. The foundational principle is that fluctuations constrained by material conservation of potential vorticity organize into Rossby waves when they are presented with a background potential vorticity gradient. The background potential vorticity gradient can arise from the meridional gradient of the planetary vorticity (giving rise to planetary waves), gradients in the vorticity of the base flow (e.g., edge waves), and, in more general models, buoyancy gradients and topography gradients (topographic Rossby waves). Note that the seeds for these arguments were planted in Section ?? when studying vorticity constraints on the flow for the non-divergent barotropic model.

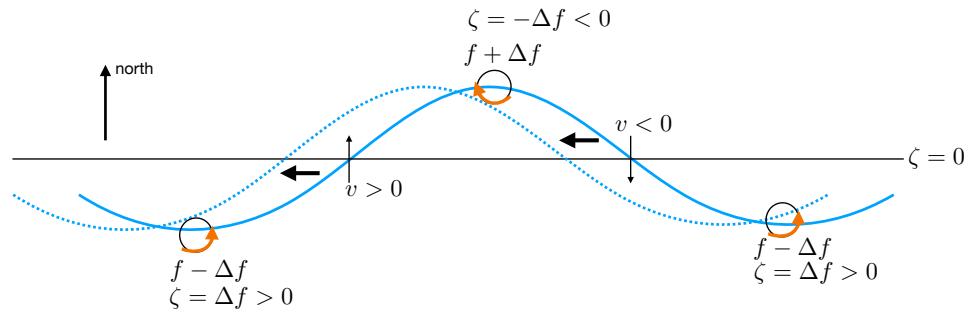


FIGURE 6.3: The westward phase propagation of a planetary Rossby wave arises from the presence of  $\beta > 0$  and the constraint that absolute vorticity ( $\zeta_a = f + \zeta$ ) is materially constant for a two-dimensional non-divergent and inviscid barotropic flow. In this figure we consider a material curve within the fluid in either the northern or southern hemispheres. The background flow is assumed to be static to allow us to focus on the role of the planetary vorticity gradient as measured by  $\beta$ . The relative vorticity is assumed to vanish for points on the straight constant latitude line, so that  $\zeta_a = f$  must be maintained by any latitudinal perturbation. For a northward perturbation relative to the latitude line, a fluid parcel finds itself at a latitude with Coriolis parameter more positive than its original value ( $f \rightarrow f + \Delta f > f$ ), with absolute vorticity conservation requiring the parcel to pick up a negative relative vorticity anomaly,  $\zeta = -\Delta f < 0$ . The opposite occurs for a southward perturbation. The counter-rotating secondary flow induced by the relative vorticity anomaly acts to move the wave pattern westward, so that the solid wave pattern is, at a future time, moved to the dashed wave pattern. We depict that the meridional motion induced by the wave as the pattern crosses the constant latitude line. In the absence of  $\beta$ , the meridional movement of parcels does not render a change in the planetary vorticity since in this case  $f$  is a constant. So for the  $f$ -plane there is no induced relative vorticity anomaly so there is no coherent movement of the wave pattern. We thus see the central role of  $\beta \neq 0$  for planetary Rossby waves. We also see that the sign of the Coriolis parameter is not relevant; it is only  $\beta > 0$  that determines the westward wave motion in both northern and southern hemispheres. A generic way to orient the phase is to note that when looking in the phase direction, higher planetary vorticity is to the right, with this rule holding for Rossby waves arising from potential vorticity gradients other than planetary  $\beta$ .

### Westward phase propagation

In Figure 6.3 we display the essential physics of planetary Rossby waves as realized in the horizontally non-divergent barotropic model. As described in the figure caption, the constraint of absolute vorticity conservation for fluctuating fluid parcels, in the presence of  $\beta > 0$ , gives rise to the westward phase propagation of the planetary Rossby wave. As we see in Section 6.5.6, a pseudo-westward propagation arises for edge waves generated by vorticity jumps. This preferred direction for propagation is a canonical property of vorticity waves, and it distinguishes these waves from other waves whose phase propagation has no directional preference.

### Vorticity and momentum arguments

The argument offered in Figure 6.3 does not consider forces. Rather, we make use of the constraint imposed by material conservation of absolute vorticity and infer the motion of fluid parcels by noting how the relative vorticity anomaly induces flow of a particular orientation. Rossby waves carry vorticity, so vorticity arguments offer the natural means to understand their mechanism. Even so, a complementary approach to understanding Rossby waves considers the forces acting in the wave, and as such is referred to as a *momentum argument*. This argument is concerned with the nature of pressure fluctuations within the wave. We note in Section 6.2.5 that plane waves in an inviscid horizontally non-divergent barotropic model oscillate between a state with exact geostrophic balance and a state with zero Coriolis acceleration so that acceleration is down the pressure gradient.

### 6.3.2 Flow relative to a zonal base state

We here examine wave fluctuations relative to a static base flow whose vorticity satisfies

$$\mathbf{u}_b \cdot \nabla(f + \zeta_b) = 0, \quad (6.23)$$

along with a corresponding velocity potential so that

$$\mathbf{u}_b = \hat{\mathbf{z}} \times \nabla\psi_b \quad \text{and} \quad \zeta_b = \nabla^2\psi_b. \quad (6.24)$$

This base flow maintains a materially constant absolute vorticity, so that it is an exact static solution to the equations of motion for the inviscid two-dimensional non-divergent barotropic model, and thus serves as a suitable base flow to study wave fluctuations.<sup>2</sup>

We focus in this chapter on a static background zonal flow written in the form<sup>3</sup>

$$\mathbf{u}_b = \hat{\mathbf{x}} U(y) \quad \text{with} \quad \psi_b = - \int^y u_b(y') dy' \quad \text{and} \quad \zeta_b = -\partial_y u_b = \partial_{yy}\psi_b. \quad (6.25)$$

As we show, the base flow with constant  $u_b = U$  (and thus with  $\zeta_b = 0$ ) supports plane waves. The study of wave fluctuations on general base states requires more general methods, such as the asymptotics from Chapter 2. We encounter another case not admitting plane waves when studying edge waves in Section 6.5. Although not admitting plane waves, the edge waves do support interfacial waves similar to the surface waves in Chapter 4.

### 6.3.3 Equations for the fluctuating vorticity and streamfunction

In the presence of a base state flow, we write the vorticity, velocity, and streamfunction in the form<sup>4</sup>

$$\zeta_{full} = \zeta + \zeta_b \quad \text{and} \quad \mathbf{u}_{full} = \mathbf{u} + \mathbf{u}_b \quad \text{and} \quad \psi_{full} = \psi + \psi_b, \quad (6.26)$$

with the vorticity equation

$$[\partial_t + (\mathbf{u} + \mathbf{u}_b) \cdot \nabla](\zeta + \zeta_b + f) = 0. \quad (6.27)$$

Again, from equation (6.23) we assume the static base state satisfies  $\mathbf{u}_b \cdot \nabla(f + \zeta_b) = 0$ , which means the vorticity equation (6.27) reduces to the equation for the fluctuating vorticity

$$\partial_t \zeta + (\mathbf{u} + \mathbf{u}_b) \cdot \nabla \zeta + \mathbf{u} \cdot \nabla(f + \zeta_b) = 0. \quad (6.28)$$

Introducing a streamfunction for the fluctuating flow,  $\zeta = \nabla^2\psi$ , brings the vorticity equation (6.28) into the form

$$\partial_t(\nabla^2\psi) + \hat{\mathbf{z}} \cdot [\nabla\psi \times \nabla(\nabla^2\psi + f + \nabla^2\psi_b)] + \hat{\mathbf{z}} \cdot [\nabla\psi_b \times \nabla(\nabla^2\psi)] = 0. \quad (6.29)$$

As we show below, this equation supports traveling plane vorticity waves. It is notable that it has only one time derivative, which contrasts to all the other wave equations we have encountered

<sup>2</sup>Although we cannot generally determine an analytic expression for the base flow, we know that a pressure can be found that accords with the flow configuration and the non-divergent nature of the flow (see Section ??).

<sup>3</sup>Recall that the streamfunction is arbitrary up to a constant (Section ??). It is for this reason that we have no concern for the lower integration limit in equation (6.25).

<sup>4</sup>An alternative notation is to write  $\zeta = \zeta' + \zeta_b$ , where  $\zeta'$  is the fluctuating vorticity. We choose the notation in equation (6.26) to reduce the abundance of primes appearing in the equations.

in this book (e.g., equation (3.34) for acoustic waves, equation (4.119) for surface gravity waves, and equation (5.27) for inertial waves). It leads to an asymmetric phase propagation of the linear wave fluctuations.

### 6.3.4 Rossby wave dispersion relation

To develop a dispersion relation we substitute the plane wave ansatz (6.3) into the vorticity equation (6.29). For the plane wave ansatz to lead to a self-consistent dispersion relation requires the background vorticity field to be extremely simple. In particular, the background velocity and background vorticity gradient must both be independent of space

$$\nabla f = \text{constant} \quad \text{and} \quad \mathbf{u}_b = U \hat{\mathbf{x}} \implies \zeta_b = 0, \quad (6.30)$$

with the first assumption holding for the  $\beta$  plane. In general, these assumptions ensure that the angular frequency for the plane wave is independent of space. In Section 6.5 we consider a slightly less trivial background state that supports edge waves rather than plane waves.

Given our focus on plane waves, assume a domain without boundaries and consider a horizontal traveling plane wave ansatz in the form of equation (6.6). Plugging this ansatz into the vorticity equation (6.29), and recalling that  $\mathbf{u} \cdot \nabla \zeta = 0$  for a plane wave as discussed in Section 6.2.3, leads to

$$A \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) [-\omega |\mathbf{k}|^2 + (\mathbf{k} \cdot \mathbf{u}_b) |\mathbf{k}|^2 - (\hat{\mathbf{z}} \times \mathbf{k}) \cdot \nabla f] = 0. \quad (6.31)$$

This equation is generally satisfied only when the bracketed term vanishes, which gives the dispersion relation that expresses the angular frequency as a function of the wavevector, the base flow, and geophysical parameters

$$\varpi = \underbrace{\mathbf{k} \cdot \mathbf{u}_b}_{\text{Doppler}} + \underbrace{(\mathbf{k} \times \hat{\mathbf{z}}) \cdot \nabla f / |\mathbf{k}|^2}_{\text{planetary vorticity gradient}} = k_x (U - \beta / |\mathbf{k}|^2). \quad (6.32)$$

As noted at the end of Section 6.3.2, the Rossby wave dispersion relation (6.32) results from a wave equation with only a single time derivative. We commented in Section 1.3.3 on such wave equations as being notable for possessing a preferred direction for wave propagation, which we see in Section 6.3.6 results in Rossby waves having a phase that moves westward.

#### Doppler shift from the base flow

The term

$$\varpi_{\text{Doppler}} \equiv \mathbf{k} \cdot \mathbf{u}_b = k_x U \quad (6.33)$$

provides a shift in the angular frequency relative to the case with zero base flow. If the wave direction is aligned with the base flow, then  $\varpi_{\text{Doppler}} > 0$ , so that the angular frequency of the wave is increased. In contrast, the wave frequency is decreased when the wave is directed anti-parallel to the base flow. This frequency shift is referred to as a *Doppler shift*. A Doppler shift is familiar from acoustic waves when, for example, the frequency of a train whistle received by a stationary listener is higher when the train approaches and lower when it is moving away.

### Rossby waves supported by a gradient in the planetary vorticity

The planetary beta effect leads to the term

$$\varpi_\beta \equiv -(\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \cdot \nabla f / |\mathbf{k}| = -\beta k_x / |\mathbf{k}|^2, \quad (6.34)$$

which gives rise to the *planetary wave* or *planetary Rossby wave*. We see that planetary Rossby waves rely on planetary curvature, in which case  $\beta \neq 0$ . Equivalently, planetary Rossby waves rely on a nonzero gradient in the planetary vorticity,  $\nabla f = \beta \hat{\mathbf{y}}$ , so that planetary Rossby waves do not exist on an  $f$ -plane.

The dispersion relation (6.34) reveals that long planetary waves (small wavenumber) have higher angular frequency than short planetary waves. The maximum angular frequency is given by the Rossby wave with zero meridional wavenumber

$$\omega_{\beta\text{-max}} = \beta / |k_x|. \quad (6.35)$$

This frequency corresponds to a purely zonal Rossby wave with no meridional structure. Correspondingly, the transverse nature of the waves means that fluid particles are moving meridionally in the presence of a zonal Rossby wave. Evidently, the frequency of the waves is directly related to the degree to which fluid particles move through the background potential vorticity field, with highest frequency for particles moving meridionally and zero frequency for particles moving zonally.

We compute the ratio of the maximum angular frequency for a Rossby wave to the central value of the Coriolis parameter,  $f_0$ , used for the beta-plane approximation (Section ??)

$$\omega_{\beta\text{-max}} / f_0 = \beta / |k_x f_0| = \beta L_x / |f_0| \ll 1. \quad (6.36)$$

In this equation we set the zonal wavenumber equal to the inverse of a zonal length scale of the flow,  $k_x = 1/L_x$ . The ratio  $\beta L_x / |f_0|$  is much less than unity so long as the  $\beta$ -plane approximation is accurate.<sup>5</sup> We thus see that the maximum angular frequency of the Rossby waves is much smaller than the Coriolis frequency, thus making the planetary Rossby wave a *sub-inertial wave*.

#### 6.3.5 Extrinsic and intrinsic angular frequency

We here introduce some terminology sometimes applied to waves in the presence of a background mean flow. For this purpose, consider again the dispersion relation,  $\varpi$ , given by equation (6.32). This relation renders the angular frequency as measured by an observer stationary with respect to the moving frame. As such, it is sometimes referred to as the *ground-based frequency* or the *extrinsic frequency*

$$\text{extrinsic frequency} = \varpi = \mathbf{k} \cdot \mathbf{u}_b + (\mathbf{k} \times \hat{\mathbf{z}}) \cdot \nabla f / |\mathbf{k}|^2 = k_x (U - \beta / |\mathbf{k}|^2). \quad (6.37)$$

The frequency measured by an observer moving with the background flow does not have a Doppler contribution, motivating the name *intrinsic frequency*

$$\text{intrinsic frequency} = \varpi - \mathbf{k} \cdot \mathbf{u}_b = -\beta / |\mathbf{k}|^2. \quad (6.38)$$

---

<sup>5</sup>See Section ?? for more on the  $\beta$ -plane approximation.

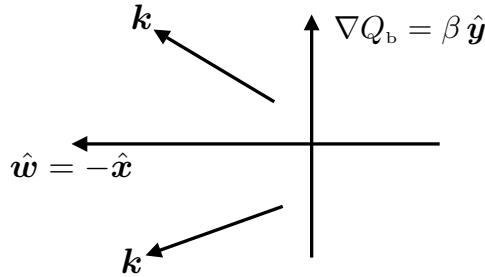


FIGURE 6.4: Illustrating the westward phase propagation of a planetary Rossby wave. In this case the westward unit vector is  $\hat{w} = -\hat{x}$ , and the phase velocity is projected in the westward direction,  $[\mathbf{c}_p \cdot \hat{w}]_\beta > 0$ . We show two possible wavevectors,  $\mathbf{k} = k_x \hat{x} + k_y \hat{y}$ , with  $k_x < 0$  so the wavevector has a westward component, but the meridional component can be either positive or negative.

The intrinsic frequency is also sometimes referred to as the *Doppler-shifted frequency*, but that name should perhaps more clearly be the “frequency with the Doppler shift removed”.

The above definitions for extrinsic and intrinsic angular frequencies accord with the conventional definitions. However, it is notable than Section 6.2 of [Sutherland \(2010\)](#) offers the exact opposite definitions. One should thus be mindful of these different naming conventions.

### 6.3.6 Concerning the westward phase velocity

The phase velocity,  $\mathbf{c}_p = (\omega/|\mathbf{k}|) \hat{\mathbf{k}}$  (see equation (1.26)), takes on the following form for a barotropic Rossby wave<sup>6</sup>

$$\mathbf{c}_p = \hat{\mathbf{k}} [\hat{\mathbf{k}} \cdot \mathbf{u}_b - (\hat{z} \times \hat{\mathbf{k}}) \cdot \nabla f / |\mathbf{k}|^2] = \hat{\mathbf{k}} [\hat{\mathbf{k}} \cdot \mathbf{u}_b + \hat{\mathbf{k}} \cdot (\hat{z} \times \nabla f) / |\mathbf{k}|^2] = \hat{\mathbf{k}} (\hat{\mathbf{k}} \cdot \hat{x}) (U - \beta / |\mathbf{k}|^2). \quad (6.39)$$

The phase velocity arising from planetary beta

$$[\mathbf{c}_p]_\beta = -\beta \hat{\mathbf{k}} (\hat{x} \cdot \hat{\mathbf{k}}) / |\mathbf{k}|^2 \quad (6.40)$$

has a sign-definite zonal component

$$[\mathbf{c}_p \cdot \hat{x}]_\beta = -(\hat{x} \cdot \hat{\mathbf{k}})^2 / |\mathbf{k}|^2 = -k_x^2 \beta / |\mathbf{k}|^4 < 0. \quad (6.41)$$

We depict this property of the Rossby wave phase velocity in Figure 6.4. This westward phase propagation holds for both hemispheres since  $\beta \geq 0$  over the globe. Furthermore, the westward propagation is larger in magnitude at lower latitudes where  $\beta$  is larger, with a ratio given by

$$\frac{[\mathbf{c}_p(\phi_1) \cdot \hat{x}]_\beta}{[\mathbf{c}_p(\phi_2) \cdot \hat{x}]_\beta} = \frac{\cos \phi_1}{\cos \phi_2}. \quad (6.42)$$

For example,  $\mathbf{c}_p(\phi_1) \cdot \hat{x}$  at  $60^\circ$  latitude is one-half that at the equator. Both the westward phase propagation and the faster phase speed at lower latitudes are canonical features of planetary Rossby waves. These properties are readily seen in large-scale flow patterns in both the atmosphere and ocean.

<sup>6</sup>In this book, we eschew the notion of components to the phase speed since the phase speed is not a vector and so it has no components. Rather, as discussed in Section 1.5.2, the phase speed is the magnitude of the phase velocity,  $C_p = \mathbf{c}_p \cdot \hat{\mathbf{k}} = \omega/|\mathbf{k}| \geq 0$ , with the phase velocity  $\mathbf{c}_p = C_p \hat{\mathbf{k}}$ .

### Westward phase velocity implied by positive angular frequency

Another indication that Rossby waves have a westward phase propagation is to recall that the angular frequency of a wave is positive (Section 1.2.3). This convention is maintained by noting that the direction of the phase propagation is carried by the wavevector

$$\mathbf{k} = \hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y, \quad (6.43)$$

rather than allowing the angular frequency to be negative.<sup>7</sup> Planetary Rossby waves in two-dimensional non-divergetnt barotropic flow have the dispersion relation (6.34)

$$\varpi_\beta = -\beta k_x / |\mathbf{k}|^2. \quad (6.44)$$

The resulting angular frequency is positive if  $k_x < 0$ , meaning that propagating planetary waves have a westward component to the phase velocity.

The result (6.44) generalizes by considering the dispersion relation arising from a more general background potential vorticity<sup>8</sup>

$$\varpi_{\text{base}+\beta} \equiv -(\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \cdot \nabla Q^b / |\mathbf{k}|, \quad (6.45)$$

with non-negative values assured only for wavevectors oriented in the pseudo-westward direction so that

$$(\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \cdot \nabla Q^b = -(\hat{\mathbf{z}} \times \nabla Q^b) \cdot \hat{\mathbf{k}} = -(\hat{\mathbf{k}} \cdot \hat{\mathbf{w}}) |\hat{\mathbf{z}} \times \nabla Q^b| < 0, \quad (6.46)$$

where we introduced the *pseudo-westward* unit vector

$$\hat{\mathbf{w}} \equiv \frac{\hat{\mathbf{z}} \times \nabla Q^b}{|\hat{\mathbf{z}} \times \nabla Q^b|}. \quad (6.47)$$

### Emphasizing the special nature of the westward phase velocity

The westward phase velocity is a very distinct feature of Rossby waves. As seen in Section 6.3.1, it results from the constraint of material conservation of potential vorticity in the presence of a background potential vorticity gradient. Other waves that we have studied, such as acoustic waves (Chapter 3), surface waves (Chapter 4), and inertial waves (Chapter 5), support an arbitrary orientation for their phase velocity. As a result, a source for these sorts of waves will generate waves whose phases are oriented in directions constrained by details of the source rather than by any intrinsic property of the waves. In contrast, an arbitrary source for Rossby waves can only produce Rossby waves with a westward oriented phase. There are no Rossby waves with an eastward phase propagation. In this manner, Rossby waves are *anisotropic waves*.

#### 6.3.7 Stationary Rossby waves

*Stationary Rossby waves* have zero phase velocity, which occurs if the base flow satisfies

$$c_p = 0 \implies \hat{\mathbf{k}} \cdot \hat{\mathbf{x}} (U - \beta / |\mathbf{k}|^2) = 0. \quad (6.48)$$

---

<sup>7</sup>Our use of a positive angular frequency is not universally maintained in the literature. For example, [Pedlosky \(2003\)](#) considers  $\omega < 0$  for Rossby waves.

<sup>8</sup>For the angular frequency to be independent of spatial position (assumed for plane waves) requires  $\nabla Q^b$  to be spatially independent. See discussion in Section 6.3.2 for more on the restrictions of the background flow enabling plane waves.

That is, the zonal portion of the Doppler shift exactly cancels the westward phase propagation from planetary  $\beta$

$$U = \beta/|\mathbf{k}|^2. \quad (6.49)$$

The corresponding wavelength,  $\Lambda = 2\pi/|\mathbf{k}|$ , is given by

$$\Lambda_{\text{stationary}} = 2\pi \sqrt{U/\beta}. \quad (6.50)$$

For example, assuming an eastward base flow speed of  $U = 1 \text{ m s}^{-1}$  (as in portions of the Antarctic Circumpolar Current) at  $\phi = 60^\circ\text{S}$ , where  $\beta = (2\Omega/R)\cos\phi \approx 1.14 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$ , renders a stationary barotropic Rossby wavelength of  $\Lambda_{\text{stationary}} \approx 1860 \text{ km}$ , whereas for the atmosphere with  $\mathbf{u}_b \cdot \hat{\mathbf{x}} = 25 \text{ m s}^{-1}$  we find  $\Lambda_{\text{stationary}} \approx 9300 \text{ km}$ . Evidently, at these large scales the Rossby waves feel the Coriolis acceleration and thus properly earn the name *planetary wave*.

### 6.3.8 Group velocity

If Rossby waves were non-dispersive, then the westward phase velocity would introduce a puzzle: how can all the wave energy propagate only in the westward direction? Since wave energy follows the group velocity (Section 1.6), would there be an unbounded accumulation of Rossby wave energy in the western side of a domain? In fact, this puzzle does not arise since Rossby waves are dispersive, with their group velocity not constrained to be westward. We here introduce the group velocity and then follow up in Sections 6.4.1 and 6.4.2 by focusing on the group and phase velocities for planetary Rossby waves.

The Rossby wave group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}}\varpi$ , is given by

$$\mathbf{c}_g = \mathbf{u}_b + \frac{\hat{\mathbf{z}} \times \nabla f - 2\hat{\mathbf{k}}[\hat{\mathbf{k}} \cdot (\hat{\mathbf{z}} \times \nabla f)]}{|\mathbf{k}|^2} = U\hat{\mathbf{x}} - \frac{\beta[\hat{\mathbf{x}} - 2\hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \hat{\mathbf{x}})]}{|\mathbf{k}|^2}. \quad (6.51)$$

The presence of  $U \neq 0$  signals the bulk transport of a Rossby wave packet by the base flow. The remaining terms can be related to the phase velocity (6.39) by projecting onto the wavevector direction

$$(\mathbf{c}_g - \mathbf{c}_p) \cdot \hat{\mathbf{k}} = 2\beta\hat{\mathbf{x}} \cdot \hat{\mathbf{k}}/|\mathbf{k}|^2. \quad (6.52)$$

Since wave dispersion is signaled by a difference between the phase velocity and group velocity, equation (6.52) indicates that Rossby wave dispersion arises from a nonzero gradient in the planetary vorticity,  $\nabla Q^b = \nabla f = \beta\hat{\mathbf{y}}$ , along with a nonzero projection of the wavevector onto the westward direction.

## 6.4 Geometry of planetary Rossby waves

We here focus on some special properties of planetary Rossby waves revealed by studying the geometry of the group velocity using a diagrammatic method developed by [Longuet-Higgins \(1964\)](#).

### 6.4.1 Group and phase velocities for planetary Rossby waves

We start by focusing on the relation between group and phase velocities for planetary Rossby waves. The  $\beta$  contribution to the group velocity (6.51) is given by

$$[\mathbf{c}_g]_\beta = \frac{\beta [(k_x^2 - k_y^2) \hat{\mathbf{x}} + 2 k_x k_y \hat{\mathbf{y}}]}{|\mathbf{k}|^4} \implies [\mathbf{c}_g \cdot \mathbf{c}_g]_\beta = \beta^2 / |\mathbf{k}|^4. \quad (6.53)$$

Recall the westward component of the phase velocity arising from planetary beta as discussed in Section 6.3.6. In contrast, the zonal component to the group velocity

$$[\mathbf{c}_g \cdot \hat{\mathbf{x}}]_\beta = \frac{\beta (k_x^2 - k_y^2)}{|\mathbf{k}|^4} = -[\mathbf{c}_p \cdot \hat{\mathbf{x}}]_\beta - \frac{\beta k_y^2}{|\mathbf{k}|^4}, \quad (6.54)$$

can be directed in either direction. Note that to reach this equality we used equation (6.41) for  $[\mathbf{c}_p \cdot \hat{\mathbf{x}}]_\beta$ .

Evidently, the group velocity (6.53) for planetary Rossby waves depends on the shape of the wave as characterized by  $(k_x^2 - k_y^2) \hat{\mathbf{x}} + 2 k_x k_y \hat{\mathbf{y}}$ . Consequently, the group velocity has the following properties for its zonal component

$$[\mathbf{c}_g \cdot \hat{\mathbf{x}}]_\beta > 0 \quad \text{if } k_x^2 > k_y^2 \quad \Rightarrow \text{eastward } \mathbf{c}_g \text{ for short zonal planetary waves} \quad (6.55a)$$

$$[\mathbf{c}_g \cdot \hat{\mathbf{x}}]_\beta < 0 \quad \text{if } k_x^2 < k_y^2 \quad \Rightarrow \text{westward } \mathbf{c}_g \text{ for long zonal planetary waves} \quad (6.55b)$$

$$[(\mathbf{c}_g + \mathbf{c}_p) \cdot \hat{\mathbf{x}}]_\beta = 0 \quad \text{if } k_y = 0 \quad \Rightarrow \text{eastward } \mathbf{c}_g \text{ for } k_y = 0 \text{ planetary waves.} \quad (6.55c)$$

Wave energy moves eastward in packets of zonally short ( $k_x^2 > k_y^2$ ) planetary Rossby waves, whereas wave energy is westward in zonally long ( $k_x^2 < k_y^2$ ) planetary Rossby waves. That is, zonally elongated Rossby waves have westward group velocity whereas zonally compressed Rossby waves have eastward group velocity, where ‘‘elongated’’ and ‘‘compressed’’ are relative to the meridional structure. Indeed, if  $k_y = 0$ , in which case there is no meridional structure to the wave, then Rossby waves of any zonal wavenumber have eastward group velocity, even while the phase velocity is westward

$$[\mathbf{c}_g \cdot \hat{\mathbf{x}}]_\beta = -[\mathbf{c}_p \cdot \hat{\mathbf{x}}]_\beta > 0 \quad \text{if } k_y = 0. \quad (6.56)$$

This property is consistent with the ratio

$$\frac{[\mathbf{c}_p \cdot \mathbf{c}_p]_\beta}{[\mathbf{c}_g \cdot \mathbf{c}_g]_\beta} = \frac{k_x^2}{|\mathbf{k}|^2} = (\hat{\mathbf{k}} \cdot \hat{\mathbf{x}})^2, \quad (6.57)$$

which then allows us to write the dispersion relation in the form

$$[\omega^2]_\beta = |\mathbf{k}|^2 [\mathbf{c}_p \cdot \mathbf{c}_p]_\beta = k_x^2 [\mathbf{c}_g \cdot \mathbf{c}_g]_\beta. \quad (6.58)$$

### 6.4.2 Dispersion circle for planetary Rossby waves

Figure 6.5 illustrates the geometry of the phase and group velocity as realized for planetary Rossby waves. We refer to this diagram as the *dispersion circle*, and it arises from noting that the dispersion relation,  $\omega = -\beta k_x / |\mathbf{k}|^2$ , can be written as an equation for a circle in the  $(k_x, k_y)$  plane

$$[k_x + \beta/(2\omega)]^2 + k_y^2 = [\beta/(2\omega)]^2. \quad (6.59)$$

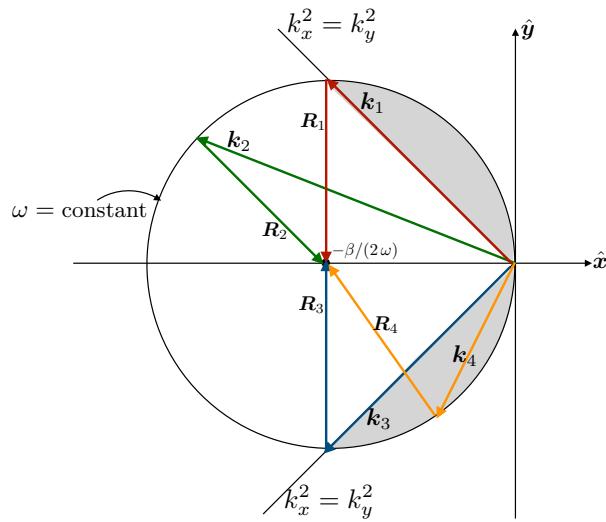


FIGURE 6.5: Dispersion circle for planetary Rossby waves, with this diagram orienting the group velocity and phase velocity in wavevector space,  $(k_x, k_y)$ . The angular frequency,  $\omega$ , determines a particular dispersion circle. We depict four example wavevectors,  $\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}$ , that orient the phase velocity,  $\mathbf{c}_p = \hat{\mathbf{k}} \omega / |\mathbf{k}|$ . Each wavevector extends from the origin to a point on the dispersion circle perimeter,  $[k_x + \beta/(2\omega)]^2 + k_y^2 = [\beta/(2\omega)]^2$ , with the circle having center at  $\mathbf{k}_{\text{center}} = -\beta/(2\omega) \hat{\mathbf{x}}$  and radius  $\beta/(2\omega)$ . Each wavevector has an associated group velocity orientation vector,  $\mathbf{R} = -\mathbf{k} - \beta/(2\omega) \hat{\mathbf{x}}$ , that points from the circle perimeter to the circle center. The group velocity is westward for those wavevectors that intersect the circle perimeter within the gray-shaded region. Such wavevectors characterize Rossby waves with zonal wavenumbers that are smaller than their meridional wavenumbers; i.e., relatively long zonal Rossby waves. The group velocity has an eastward component for wavevectors outside the gray region, with the lines  $k_x^2 = k_y^2$  separating these regions where the group velocity is eastward or westward. Such wavevectors characterize Rossby waves with zonal wavenumbers that are larger than their meridional wavenumbers; i.e., relatively short zonal Rossby waves. The group velocity for wavevector  $\mathbf{k}_1$  is exactly southward; for  $\mathbf{k}_2$  it is southeastward; for  $\mathbf{k}_3$  it is exactly northward, and for  $\mathbf{k}_4$  it is northwestward. This figure is taken after [Longuet-Higgins \(1964\)](#).

The center of the circle is at the wavevector

$$\mathbf{k}_{\text{center}} = -\beta/(2\omega) \hat{\mathbf{x}}, \quad (6.60)$$

and with a radius equal to  $\beta/(2\omega)$ . The circle has angular frequency,  $\omega$ , acting as a parameter, with lower frequency Rossby waves yielding larger circles. We further reveal the geometry of the group velocity (6.53) by writing it in the form

$$[\mathbf{c}_g]_\beta = \frac{\beta [(k_x^2 - k_y^2) \hat{\mathbf{x}} + 2 k_x k_y \hat{\mathbf{y}}]}{|\mathbf{k}|^4} = -\frac{2\omega [[k_x + \beta/(2\omega)] \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}]}{|\mathbf{k}|^2}. \quad (6.61)$$

Furthermore, we introduce the group velocity orientation vector

$$\mathbf{R} = -\mathbf{k} - \beta/(2\omega) \hat{\mathbf{x}} = -[k_x + \beta/(2\omega)] \hat{\mathbf{x}} - k_y \hat{\mathbf{y}} \quad \text{with} \quad |\mathbf{R}| = \beta/(2\omega), \quad (6.62)$$

so that the group velocity can be written in the rather tidy form

$$[\mathbf{c}_g]_\beta = 2\omega \mathbf{R}/|\mathbf{k}|^2. \quad (6.63)$$

Notice that  $\mathbf{R}$  has magnitude equal to the radius of the circle. Furthermore, this vector points from the circle perimeter to the circle center, as seen since  $\mathbf{R} + \mathbf{k} = -\beta/(2\omega) \hat{\mathbf{x}} = \mathbf{k}_{\text{center}}$ .

The geometry depicted in Figure 6.5 partitions the group velocity according to the wavevector. Again, the phase velocity always has a westward component, yet the group velocity can be westward or eastward. Additionally, for each angular frequency there is one wave whose group velocity is precisely northward and another that is southward. The squared magnitude of the group velocity is given by equation (6.53)

$$[\mathbf{c}_g \cdot \mathbf{c}_g]_\beta = \beta^2/|\mathbf{k}|^4, \quad (6.64)$$

so that longer waves (smaller wavenumber) have higher group velocity magnitude.

### 6.4.3 Reflection of planetary Rossby waves

The diagrammatic method developed in Section 6.4.2 provides a basis to characterize the reflection of Rossby wave packets from a smooth solid boundary. In the left panel of Figure 6.6 we depict a straight and smooth wall sloped with angle  $\gamma$  in the counter-clockwise direction from the positive  $x$ -axis. An incident southwestward group velocity carries a Rossby wave packet to the wall at an angle,  $\theta_i$ , relative to the wall's normal direction,  $\hat{\mathbf{n}}$ . We assume the group velocity represents a packet whose central carrier Rossby wave has a wavevector  $\mathbf{k}_i$ , and seek information about the reflected wave packet's carrier wavevector,  $\mathbf{k}_r$ . Geometric optics from Section 2.3 provides the foundation for the approach taken in the following, in which we describe a diagrammatic approach for understanding how the wave packet reflects. Details are summarized in the right panel of Figure 6.6. We assume that the waves reflect from the boundary without dissipation (i.e., perfectly smooth and straight wall), so that reflection only involves kinematic boundary conditions.

#### Kinematic boundary condition at the wall

The kinematic boundary condition along the wall requires the velocity of fluid particles in the wave to have zero normal component at the wall, so that the streamfunction must be

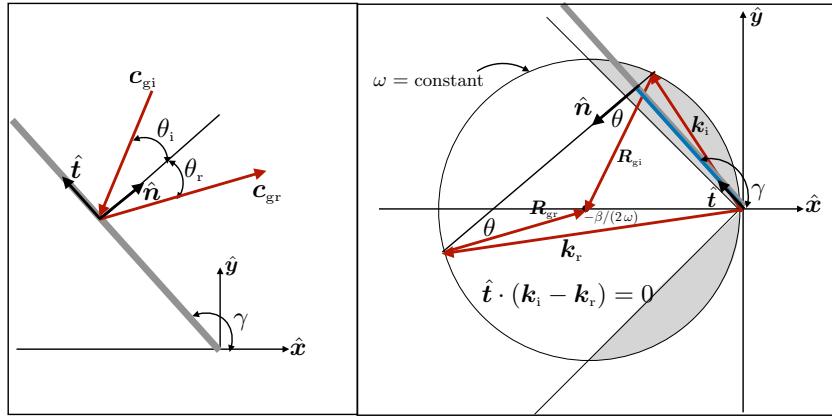


FIGURE 6.6: Depicting the reflection of a Rossby wave packet from a solid wall. The left panel shows the wall sloped at an angle,  $\gamma$ , relative to the positive  $x$ -axis. We orient the wall using both its tangential unit vector,  $\hat{t} = \hat{z} \times \hat{n} = \hat{x} \cos \gamma + \hat{y} \sin \gamma$ , as well as its unit normal,  $\hat{n} = \hat{x} \sin \gamma - \hat{y} \cos \gamma$ . The incident wave packet has group velocity  $\mathbf{c}_{gi}$  directed to the southwest and reflected group velocity  $\mathbf{c}_{gr}$  directed to the northeast. The angle of incidence equals to the angle of reflectance,  $\theta_i = \theta_r = \theta$ , as required by the kinematic boundary condition (6.67). The right panel shows the wave packet reflection using the dispersion diagram from Figure 6.5. Since the incident and reflected wave have the same angular frequency, we can use the same circle for deriving the wavevectors. The incident and reflected group velocities are oriented by the vectors  $\mathbf{R}_{gi}$  (southwest) and  $\mathbf{R}_{gr}$  (northeast) that point from the perimeter to the center of the  $\omega$ -circle. They both make an angle of  $\theta$  with respect to the wall's normal direction,  $\hat{n}$ . The incident and reflected wavevectors satisfy equation (6.72), which says  $(\mathbf{k}_i - \mathbf{k}_r) \cdot \hat{t} = 0$ . This relation provides the means to determine the reflected wavevector,  $\mathbf{k}_r$  using the diagrammatic method illustrated here. Note that the right panel depicts the orientation of the western wall relative to the wavevectors and group velocities in wavevector space. One should not interpret this depiction as somehow making the western wall boundary into an eastern wall.

constant along the wall

$$\mathbf{u} \cdot \hat{n} = (\hat{z} \times \nabla \psi) \cdot \hat{n} = (\hat{n} \times \hat{z}) \cdot \nabla \psi = -\hat{t} \cdot \nabla \psi = 0, \quad (6.65)$$

where we introduced the wall's unit tangent vector

$$\hat{t} = \hat{z} \times \hat{n} = \hat{x} \cos \gamma + \hat{y} \sin \gamma. \quad (6.66)$$

The streamfunction can be any constant on the wall, which we take to be zero without loss of generality

$$\psi(\mathbf{x} = \mathbf{x}_{\text{wall}}, t) = 0, \quad (6.67)$$

with  $\mathbf{x}_{\text{wall}}$  the coordinate for a point on the wall.

### Relating incident and reflected wave properties

Write the velocity streamfunctions for the incident and reflected carrier waves as

$$\psi_i = A_i \cos(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t) \quad \text{and} \quad \psi_r = A_r \cos(\mathbf{k}_r \cdot \mathbf{x} - \omega_r t), \quad (6.68)$$

where  $A_i$  and  $A_r$  are the real wave amplitudes,  $\mathbf{k}_i$  and  $\mathbf{k}_r$  are the wavevectors, and  $\omega_i$  and  $\omega_r$  are the angular frequencies. At any point in the fluid, the streamfunction is the sum of the streamfunctions for the incident and reflected waves

$$\psi(\mathbf{x}, t) = \psi_i(\mathbf{x}, t) + \psi_r(\mathbf{x}, t). \quad (6.69)$$

Satisfaction of the boundary condition (6.67) requires

$$A_i \cos(\mathbf{k}_i \cdot \mathbf{x}_{\text{wall}} - \omega_i t) + A_r \cos(\mathbf{k}_r \cdot \mathbf{x}_{\text{wall}} - \omega_r t) = 0. \quad (6.70)$$

For this equality to hold at each point along the wall and for all time requires

$$\omega_i = \omega_r \quad \text{equal incident and reflected angular frequency} \quad (6.71a)$$

$$A_i = -A_r \quad \text{equal wave amplitudes but opposite sign} \quad (6.71b)$$

$$(\mathbf{k}_i - \mathbf{k}_r) \cdot \hat{\mathbf{t}} = 0 \quad \text{equal projection of wavevectors onto } \hat{\mathbf{t}}. \quad (6.71c)$$

The final equality means that there is an equal projection of the incident and reflected wavevectors onto the wall's tangential direction. It arises by writing the position for a point on the wall as

$$\mathbf{x}_{\text{wall}} = |\mathbf{x}_{\text{wall}}| \hat{\mathbf{t}} \implies \mathbf{k}_i \cdot \hat{\mathbf{t}} = \mathbf{k}_r \cdot \hat{\mathbf{t}}. \quad (6.72)$$

Considering the wall to be an extreme case of a static inhomogeneous media, we can connect these relations to the ray equations in Section 2.3.6. In that discussion we found that a wave packet moving through a static but inhomogeneous media maintains a constant angular frequency along a ray, whereas the wavevector changes.

### Geometry of the reflection in wavevector space

Our considerations thus far have been generic, holding for any wave packet described by a wave function such as the velocity streamfunction. To determine further details of the reflected Rossby wavevector, return to the Rossby wave dispersion diagram in Figure 6.5, depicting the reflection process in wavevector space. We only need one dispersion circle since both the incident and reflected waves have the same angular frequency, as required by the kinematic boundary condition (6.70).

To construct the dispersion diagram we start with the known incident group velocity,  $\mathbf{c}_{gi}$ , which is assumed to be directed towards the southwest. Draw the corresponding group velocity orientation vector,  $\mathbf{R}_{gi}$ , from the circle perimeter to the center, also oriented in the same southwesterly direction. From knowledge of  $\mathbf{R}_{gi}$ , draw the incident carrier wavevector,  $\mathbf{k}_i$ , extending from the origin to where  $\mathbf{R}_{gi}$  meets the perimeter. Next make use of the kinematic condition (6.72) that allows us to compute the unique reflected carrier wavevector,  $\mathbf{k}_r$ , constructed by setting  $\mathbf{k}_r \cdot \hat{\mathbf{t}} = \mathbf{k}_i \cdot \hat{\mathbf{t}}$ . Finally, we can now compute the reflected group velocity orientation vector,  $\mathbf{R}_{gr}$ , which points to the center of the circle from the point where  $\mathbf{k}_r$  hits the circle perimeter. It is through this construction that we find the incident and reflected group velocities make the same angle with the wall normal:

$$\theta_i = \theta_r = \theta. \quad (6.73)$$

Reflections that satisfy this property are known as *specular*.

### Features of the incident and reflected waves

Both the incident and reflected carrier wavevectors have a westward component, as required for Rossby waves. However, the southwestward orientation of the incident group velocity is reflected at the wall into a northeastward group velocity. The reflected wave packet, moving eastward, has a larger zonal wavenumber than the incident wave packet:

$$|\mathbf{k}_r| > |\mathbf{k}_i|. \quad (6.74)$$

This increase in wavenumber arises from an increase in zonal wavevector component, so that the zonal wavelength of the waves within the reflected eastward wave packet are shorter than those in the incident wave packet. The larger wavenumber decreases the group velocity, so that the northeastward reflected packet is slower than the southwestward incident packet. In a fluid with dissipation, such as through viscosity (see Section ??), we expect smaller scale features to be dissipated more readily than larger scale features. Hence, the northeastward reflected wave packet is expected to be dissipated more readily than the southwestward incident packet.

The reflection of westerly moving Rossby wave packets off a western boundary hold in their converse for the reflection at easterly packets hitting an eastern wall. Namely, a slowly moving easterly wave packet, which is comprised of short wavelength Rossby waves, is reflected off the eastern wall as a faster moving westerly packet of longer wavelength Rossby waves. These results have particular relevance to the ocean, such as through middle latitude western boundary current intensification and the El Niño / Southern Oscillation phenomena in the tropics (see [Vallis \(2017\)](#) for further discussion).

#### 6.4.4 Further study

In addition to working through the geometry of planetary waves on a  $\beta$ -plane, [Longuet-Higgins \(1964\)](#) studies planetary waves on a sphere, thus making use of spherical harmonics.

### 6.5 Edge waves

*Edge waves* are vorticity fluctuations that live on the interface separating two regions of different background vorticity.<sup>9</sup> As vorticity waves, the edge waves of this section share many features with planetary Rossby waves studied in Section 6.3. Additionally, edge waves share features with surface waves studied in Chapter 4, in that they travel along the interface while exponentially decaying in the direction away from the interface.

#### 6.5.1 Base state and the meridionally modulated wave ansatz

Following the decomposition (6.26) we write the relative vorticity as

$$\zeta_{\text{full}} = \zeta + \zeta_b \quad (6.75)$$

where  $\zeta_b$  is a static base state vorticity and  $\zeta$  the vorticity fluctuating around the base state. The absolute vorticity equation thus takes the form

$$D(\zeta + \zeta_b + f)/Dt = 0, \quad (6.76)$$

in which we see that the base state vorticity,  $\zeta_b$ , plays a role directly analogous to planetary vorticity,  $f$ . This analog allows us to transfer concepts of planetary Rossby waves from Section 6.3 directly over to the edge waves of this section.

We are concerned with a base state comprised of a meridionally dependent zonal flow and corresponding vorticity

$$\mathbf{u}_b = \hat{x} u_b(y) \quad \text{and} \quad \zeta_b = \zeta_b(y) = -\partial_y u_b. \quad (6.77)$$

---

<sup>9</sup> [Sutherland \(2010\)](#) in his section 2.6.2 refers to edge waves as *Rayleigh waves*.

In particular, to generate edge waves we assume in Section 6.5.3 that the background vorticity has a jump at  $y = y_0$ . The vorticity equation (6.28) describing fluctuations relative to the background flow (6.77) is given by

$$\partial_t \zeta + (\mathbf{u} + \hat{\mathbf{x}} u_b) \cdot \nabla \zeta + v (\beta + \partial_y \zeta_b) = 0, \quad (6.78)$$

which, when introducing the streamfunction  $\zeta = \nabla^2 \psi$ , leads to

$$\partial_t (\nabla^2 \psi) + \hat{\mathbf{z}} \cdot [\nabla \psi \times \nabla (\nabla^2 \psi)] + u_b \partial_x (\nabla^2 \psi) + \partial_x \psi (\beta + \partial_y \zeta_b) = 0. \quad (6.79)$$

If we substitute the plane wave ansatz (6.3) into the vorticity equation (6.79), assuming the angular frequency and wavevector are spatially independent, then we are led to

$$\omega = k_x u_b - \frac{k_x (\beta + \partial_y u_b)}{|\mathbf{k}|^2} \quad \text{inconsistent equation.} \quad (6.80)$$

This equation looks like a straightforward generalization of the Rossby wave dispersion relation (6.32). However, since the background flow,  $u_b(y)$ , has a meridional dependence, the plane wave assumption that  $\omega$  and  $\mathbf{k}$  are spatial constants is, in fact, flawed. We conclude that the vorticity equation (6.79) does not admit a plane wave solution when the background flow is not a constant.

For the case of a background  $u_b(y)$  that is gently varying, we can make use of the asymptotic methods from Chapter 2. However, for the edge waves considered in this section we can make progress with a somewhat simpler approach, in which we assume the streamfunction is a zonal traveling wave modulated by a meridionally dependant amplitude

$$\psi(x, y, t) = \tilde{\psi}(y) \cos(k x - \omega t). \quad (6.81)$$

We pursue this ansatz in the following.

### 6.5.2 Rayleigh-Kuo equation

The plane wave properties studied in Section 6.2 do not hold for the meridionally modulated wave (6.81). Consequently, to derive a dispersion relation requires us to linearize the vorticity equation (6.79), which takes the form

$$\partial_t (\nabla^2 \psi) + u_b \partial_x (\nabla^2 \psi) + \partial_x \psi (\beta + \partial_y \zeta_b) = 0. \quad (6.82)$$

Substituting the modulated wave ansatz (6.81) into the linearized vorticity equation (6.82) gives

$$\partial_t (\nabla^2 \psi) = \omega (-k^2 + \partial_{yy}) \tilde{\psi} \sin(k x - \omega t) \quad (6.83a)$$

$$u_b \partial_x (\nabla^2 \psi) = -u_b k (-k^2 + \partial_{yy}) \tilde{\psi} \sin(k x - \omega t) \quad (6.83b)$$

$$\partial_x \psi (\beta + \partial_y \zeta_b) = -k (\beta + \partial_y \zeta_b) \tilde{\psi} \sin(k x - \omega t), \quad (6.83c)$$

which then leads to the *Rayleigh-Kuo* equation for the meridional wave amplitude function<sup>10</sup>

$$(u_b - c) (\partial_{yy} - k^2) \tilde{\psi} + (\beta - \partial_{yy} u_b) \tilde{\psi} = 0 \quad \text{with} \quad c = \omega/k, \quad (6.84)$$

---

<sup>10</sup>Kuo extended the original Rayleigh equation to account for the  $\beta$  effect. In the absence of  $\beta$  we refer to just the *Rayleigh equation*.

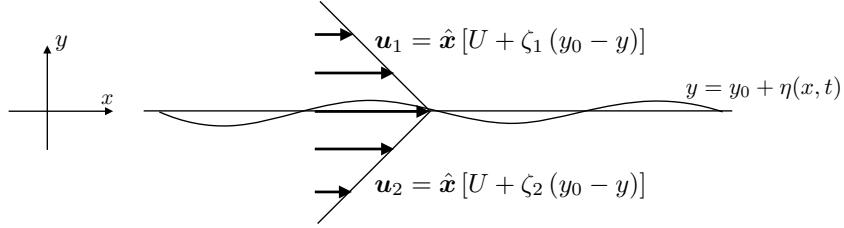


FIGURE 6.7: A zonal point jet in the horizontal  $x$ - $y$  plane with the jet maxima at  $y = y_0$ . The velocity is linear on both sides of the maxima and there is a finite jump in the vorticity at the maxima. Vorticity in the upper region is given by  $\zeta_1 = -\partial_y u_1$  and in the lower region  $\zeta_2 = -\partial_y u_2$ , with  $\zeta_1 > 0$  and  $\zeta_2 < 0$  depicted in this figure. Oscillations of the material interface,  $y = y_0 + \eta(x, t)$ , are known as *edge waves*, which exist due to the jump in the vorticity. The top half-plane is referred to as region 1 and the lower half-plane is region 2. Any surface with a finite vorticity jump supports edge waves, with the symmetric point jet shown in this figure a prototypical case.

which can also be written in the form

$$\partial_y[(u_b - c) \partial_y \tilde{\psi} - \tilde{\psi} \partial_y u_b] + [\beta - k^2 (u_b - c)] \tilde{\psi} = 0. \quad (6.85)$$

The Rayleigh-Kuo equation is fundamental to our study of barotropic shear instability in Chapter 13. For current purposes we use it to derive the dispersion relation for edge waves.

### 6.5.3 The point jet

Figure 6.7 illustrates the zonal flow characterizing a *point jet*, in which there is a linear velocity profile on both sides of a maxima

$$\mathbf{u}_b = u_b(y) \hat{x} \quad \text{with} \quad u_b(y) = U + (y_0 - y) \begin{cases} \zeta_1 & \text{for } y > y_0 \\ \zeta_2 & \text{for } y < y_0, \end{cases} \quad (6.86)$$

where  $\zeta_1$  and  $\zeta_2$  are constants that measure the vorticity in the two half-planes

$$\zeta_b = -\partial_y u_b = \begin{cases} \zeta_1 & \text{for } y > y_0 \\ \zeta_2 & \text{for } y < y_0. \end{cases} \quad (6.87)$$

Because of the finite vorticity jump at  $y = y_0$ , the meridional derivative of the vorticity is proportional to the Dirac delta<sup>11</sup>

$$\partial_y \zeta_b(y = y_0) = \lim_{\epsilon \rightarrow 0} \frac{\zeta_b(y = y_0 + \epsilon/2) - \zeta_b(y = y_0 - \epsilon/2)}{\epsilon} = (\zeta_1 - \zeta_2) \delta(y - y_0). \quad (6.88)$$

As a check on this expression, we take an integral across the interface to find

$$\int_{y_0 - \epsilon/2}^{y_0 + \epsilon/2} \partial_y \zeta_b dy = \zeta_b(y_0 + \epsilon/2) - \zeta_b(y_0 - \epsilon/2) = (\zeta_1 - \zeta_2) \int_{y_0 - \epsilon/2}^{y_0 + \epsilon/2} \delta(y - y_0) dy. \quad (6.89)$$

Finally, note that the point yet is zonally symmetric so that the base flow only has a nonzero meridional pressure gradient

$$\nabla \varphi_b = \hat{y} \partial_y \varphi_b. \quad (6.90)$$

We study waves traveling on the material vorticity interface at  $y = y_0 + \eta(x, t)$ , and written

<sup>11</sup>See Chapter ?? for more on the Dirac delta. In particular, note that the Dirac delta,  $\delta(y)$ , has dimensions of inverse length.

in the form

$$\eta(x, t) = \tilde{\eta} \cos(k x - \omega t). \quad (6.91)$$

To develop properties of the flow on each side of the interface requires us to develop the kinematic and dynamic boundary conditions at the interface. Furthermore, the amplitude,  $\tilde{\eta}$ , is assumed to be much smaller than the wavelength, thus allowing us to linearize the boundary conditions and to evaluate the conditions at  $y = y_0$ .<sup>12</sup> Note that we assume the interface undulation has the same phase as the streamfunction (6.81), and this assumption will be seen to be self-consistent.

#### 6.5.4 Kinematic boundary condition at the interface

The material nature of the interface means that the meridional velocity of a fluid particle at the interface is given by<sup>13</sup>

$$v = D\eta/Dt = (\partial_t + u \partial_x)\eta \approx (\partial_t + u_b \partial_x)\eta, \quad (6.92)$$

where the approximation follows from linearization. This kinematic boundary condition holds on both sides of the interface so that

$$(\partial_t + u_1 \partial_x)\eta = \partial_x \psi_1 \quad \text{and} \quad (\partial_t + u_2 \partial_x)\eta = \partial_x \psi_2. \quad (6.93)$$

Applying this relation to the wave ansatz (6.91) for the interface, and the wave ansatz (6.81) for the streamfunction as evaluated at  $y = y_0$ , leads to

$$(u_1 - c)\tilde{\eta} = \tilde{\psi}_1 \quad \text{and} \quad (u_2 - c)\tilde{\eta} = \tilde{\psi}_2. \quad (6.94)$$

Since  $u_1(y_0) = u_2(y_0) = U$ , the kinematic boundary condition says that the streamfunctions match at the interface

$$\tilde{\psi}_1 = \tilde{\psi}_2 \quad \text{at } y = y_0. \quad (6.95)$$

#### 6.5.5 Dynamic boundary condition at the interface

For the dynamic boundary condition we assume there is no surface tension on the interface, in which case pressure matches

$$\varphi_1 = \varphi_2 \quad \text{at } y = y_0. \quad (6.96)$$

To make use of this boundary condition, consider the zonal velocity equation (??) with the zonal flow decomposed as in equation (6.26)

$$\partial_t(u + u_b) + (\mathbf{u} + \mathbf{u}_b) \cdot \nabla(u + u_b) - f v = -\partial_x \varphi, \quad (6.97)$$

which linearizes to

$$(\partial_t + u_b \partial_x)u + v(\partial_y u_b - f) = -\partial_x \varphi \implies (\partial_t + u_b \partial_x)\partial_y \psi - \partial_x \psi(\partial_y u_b - f) = \partial_x \varphi, \quad (6.98)$$

---

<sup>12</sup>We took the same approach for the surface gravity waves in Section 4.3.4.

<sup>13</sup>Equation (6.92) is the direct analog of the kinematic boundary condition developed in Section ?? for vertical position of a material surface,  $z = \eta(x, y, t)$ .

where we set  $\partial_x \varphi_b = 0$  as per the assumed zonal symmetry of the background state in equation (6.90). Consider an ansatz for the pressure in the same form as for the streamfunction

$$\varphi(x, y, t) = \tilde{\varphi}(y) \cos(k x - \omega t), \quad (6.99)$$

which, along with the streamfunction ansatz (6.81), brings the velocity equation (6.98) to

$$[(u_b - c) \partial_y + (f - \partial_y u_b)] \tilde{\psi} = \tilde{\varphi} \quad \text{at } y = y_0. \quad (6.100)$$

Invoking continuity of pressure at the interface yields

$$[(U - c) \partial_y + (f - \partial_y u_1)] \tilde{\psi}_1 = [(U - c) \partial_y + (f - \partial_y u_2)] \tilde{\psi}_2 \quad \text{at } y = y_0. \quad (6.101)$$

This dynamic boundary condition also follows from integrating the Rayleigh-Kuo equation (6.85) across the interface and noting that  $[\beta - k^2 (u_b - c)] \tilde{\psi}$  is continuous at the interface. We can further simplify the dynamic boundary condition (6.101) by making use of the kinematic boundary condition (6.95),  $\tilde{\psi}_1 = \tilde{\psi}_2$  at  $y = y_0$ . Doing so eliminates the Coriolis term and yields

$$(U - c) (\partial_y \tilde{\psi}_1 - \partial_y \tilde{\psi}_2) = [\partial_y u_1 - \partial_y u_2] \tilde{\psi} \quad \text{at } y = y_0. \quad (6.102)$$

### 6.5.6 Edge wave dispersion relation

Set  $\beta = 0$  to focus on wave solutions arising just from the point jet, in which the Rayleigh-Kuo equation (6.84) simplifies to

$$(u_b - c) (\partial_{yy} - k^2) \tilde{\psi} - \partial_{yy} u_b \tilde{\psi} = 0. \quad (6.103)$$

Evaluating this equation on the two sides of the interface, and assuming the base flow is distinct from the phase velocity so that  $u_b \neq c$ , leads to

$$\tilde{\psi} = \psi_0 \begin{cases} e^{-|k|(y-y_0)} & \text{for } y > y_0 \\ e^{+|k|(y-y_0)} & \text{for } y < y_0. \end{cases} \quad (6.104)$$

As anticipated, the edge wave exponentially decays away from the interface. Just like the surface waves in Chapter 4, the horizontal wave number,  $|k|$ , determines the exponential decay scale.

Using equation (6.104) in the dynamic boundary condition (6.102) yields the phase velocity<sup>14</sup>

$$c = \omega/k = U + \frac{\partial_y u_1 - \partial_y u_2}{2|k|} = \underbrace{U}_{\text{Doppler}} - \underbrace{\frac{(\zeta_1 - \zeta_2)}{2|k|}}_{\text{intrinsic}}. \quad (6.105)$$

The background zonal flow,  $U$ , provides a Doppler shift to the phase velocity. The second piece arises from the vorticity jump across the interface, and we refer to it as the *intrinsic* portion to the phase velocity. If the vorticity increases northward, so that  $\zeta_1 - \zeta_2 > 0$ , then  $c - U < 0$ , thus signaling a westward phase velocity relative to the background flow. This behavior is precisely that found for planetary Rossby waves, thus prompting some to refer to edge waves as prototypical Rossby waves. In Figure 6.8, we summarize the vorticity mechanism for the

<sup>14</sup>The phase of the edge wave only moves in the zonal direction, with  $c > 0$  for a phase velocity in the  $+\hat{x}$  direction and  $c < 0$  for a phase velocity in the  $-\hat{x}$  direction.

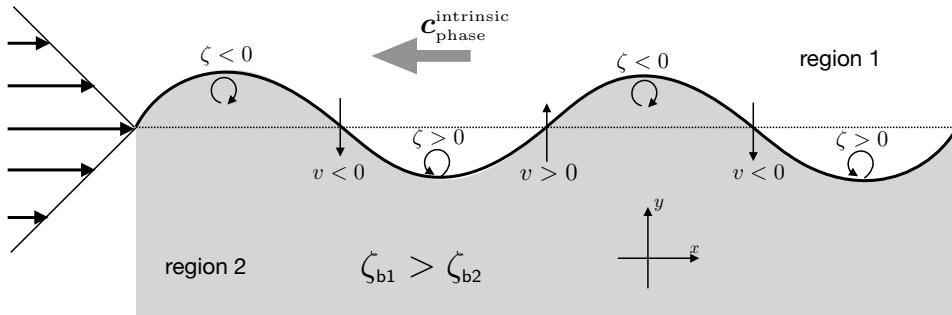


FIGURE 6.8: Depicting the *pseudo-westward* intrinsic phase propagation of an edge wave along a material surface separating two regions of vorticity. Here the larger background relative vorticity is oriented to the north,  $\zeta_{b1} > \zeta_{b2}$ , so that the intrinsic phase velocity for the edge wave propagates to the west. We also show the point jet velocity profile, whose relative vorticity jump supports the westward phase velocity of the edge wave. The variables,  $v$  and  $\zeta$ , are the anomalies relative to the unperturbed reference state, as per the decomposition equation (6.26). The vorticity mechanism for the phase propagation is identical to that for planetary Rossby waves depicted in Figure 6.3, where the background vorticity jump from the point jet serves the same role as the planetary vorticity gradient. Consider a material line initially along  $y = y_0$  and whose relative vorticity equals to that of the background point jet. A southward extension of this line leads to a local positive relative vorticity anomaly,  $\zeta' > 0$ , since the particle moves into a region with lower background relative vorticity. Conversely, a northward extension leads to a local negative vorticity anomaly and associated clockwise secondary flow. The anomalous meridional motion is depicted every  $\pi$  radians, with maximal anomalies at wave nodes. The action of the coherent counter-rotating secondary vortices leads to westward propagation of the phase. In general, the phase moves with the higher vorticity to the right, which means that it moves from the convex side of the velocity profile towards the concave side.

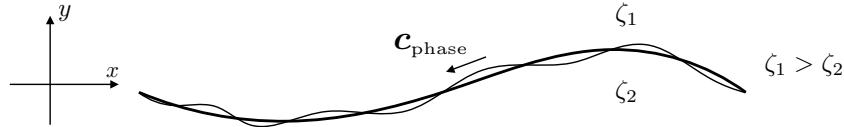


FIGURE 6.9: Edge waves moving along a jump in the background vorticity field, with the background vorticity equal to  $\zeta_1$  north of the jump and  $\zeta_2$  to the south. The phase velocity is oriented with the higher relative vorticity to the right when facing in the direction of the phase velocity.

*pseudo-westward* phase velocity, with the mechanism identical to the planetary Rossby wave in Figure 6.3. In general, the phase velocity for edge waves is directed so that the region of higher vorticity is to the right when facing in the direction of the phase velocity, with this orientation defining “pseudo-westward”. This terminology is taken from the truly westward phase velocity corresponding to planetary waves, in which the larger planetary vorticity is to the north. Figure 6.9 provides a schematic for a curved vorticity jump line, whereby the edge waves propagate along the jump, again with the phase in the pseudo-westward direction.

### 6.5.7 Further study

A similar treatment of edge waves can be found in Section 3.12 of [Smyth and Carpenter \(2019\)](#) and Section 9.2.3 of [Vallis \(2017\)](#). As studied in Chapter 13, interactions between vorticity waves can lead to shear instabilities, with such instabilities forming a fundamental feature of unstable and turbulent geophysical flows.



## 6.6 Exercises

### EXERCISE 6.1: ROSSBY WAVES IN A ZONAL CHANNEL

In Section 6.3 we examined the physics of Rossby waves on an unbounded  $\beta$ -plane. Here we consider a zonal channel on a  $\beta$ -plane, with the channel unbounded in the zonal direction but bounded meridionally. The meridional domain is given by  $y_s \leq y \leq y_n$  where  $y_n - y_s = L$ , and with rigid material walls at the meridional boundaries. Everything else is just as in Section 6.3, with a zonal and constant base flow,  $\mathbf{u}_b = U \hat{\mathbf{x}}$ .

- (a) Write a single plane wave mode that travels in the zonal direction but is standing in the meridional direction. Hint: the plane wave ansatz (6.6) should be modified so that  $\hat{\mathbf{n}} \cdot (\hat{\mathbf{z}} \times \nabla \psi) = 0$  at the northern and southern walls. The boundary condition imposes a constraint on the meridional wavenumber. What is the constraint?
- (b) What is the dispersion relation for the Rossby waves in this channel?

Hint: recall our study of standing gravity wave modes in Section 4.8, with such modes occurring when waves live in a bounded domain and thus can no longer travel freely. Instead, the wave modes must fit inside of the domain in a manner that satisfies the boundary conditions. We can think of such standing wave modes as a superposition of two oppositely traveling waves with the same frequency and wavenumber that are locked in-phase in a manner that satisfies the boundary conditions. For example, the sum of a right and left moving wave with equal amplitude, wavenumber, and frequency is given by the standing pattern

$$A \cos(k x - \omega t) + A \cos(-k x - \omega t) = 2 A \cos(\omega t) \cos(k x). \quad (6.106)$$

### EXERCISE 6.2: VORTEX MECHANISM FOR EDGE WAVES

Provide a sketch like Figure 6.8 but for edge waves in the case with  $\zeta_1 < \zeta_2$ . Hint: be sure to orient the wave so that it moves from the convex side of the velocity profile towards the concave side.



## Chapter 7

# SHALLOW WATER WAVES

In this chapter we develop the basic theory for waves appearing in the shallow water fluid. In the literature these waves are sometimes referred to as *long waves* since the shallow water approximation applies when the horizontal length scale is much larger than the vertical, in which case the horizontal wavelength is larger than the fluid depth. Our presentation focuses on the case of a single shallow water layer, but we offer brief discussions of two-layer cases to illustrate the generalizations to multiple layers. Notably, the algebraic tedium increases greatly when adding layers. Hence, for theoretical analysis concerned with the role of stratification, it is typically more fruitful to move to the continuous vertical stratification in Chapter 9, rather than study waves in more than two stacked shallow water layers.

We first consider general features of both the nonlinear and then the linear shallow water equations. For the linearized system, we develop a unified wave equation for the free surface. The derivation of this equation is rather detailed and the resulting linear equation somewhat complicated. Even so, it serves to unify across the full suite of linear shallow water waves on a rotating  $\beta$ -plane, and offers insights into both the methods used to derive wave equations and the scalings used to extract their core physical features. In unpacking this wave equation we pursue a focused study of gravity waves, inertia-gravity waves, and Rossby waves. We also consider the case of a Kelvin wave, which relies on a boundary in the presence of the Coriolis parameter.

### READER'S GUIDE TO THIS CHAPTER

This chapter builds from the shallow water studies in Chapters ?? and ??, as well as other waves chapters in this part of the book. In Chapter 8 we extend the study of shallow water waves by considering a variety of case studies. [The first half of this video](#) offers a pedagogical introduction to non-rotating shallow water waves.

<b>7.1</b>	<b>Loose threads</b>	<b>192</b>
<b>7.2</b>	<b>Shallow water equations</b>	<b>193</b>
7.2.1	Nonlinear wave equation	193
7.2.2	Another form of the nonlinear wave equation	193
7.2.3	Comments	194
<b>7.3</b>	<b>Linearized shallow water equations</b>	<b>194</b>
7.3.1	Linearized thickness and velocity equations	194
7.3.2	Energetics	196
7.3.3	Potential vorticity	196
7.3.4	Traveling plane wave ansatz and $\nabla \cdot \mathbf{v} = 0$	197
<b>7.4</b>	<b>Unified shallow water wave equation</b>	<b>197</b>

7.4.1	Use of the divergence equation . . . . .	198
7.4.2	Use of the vorticity equation . . . . .	199
7.4.3	The unified wave equation just with $\eta'$ . . . . .	200
7.4.4	Scaling for the $\beta$ -plane with small topography . . . . .	201
7.4.5	Super-inertial wave equation . . . . .	201
7.4.6	Sub-inertial wave equation . . . . .	202
7.4.7	Comments on bottom topography . . . . .	205
7.5	<b>Shallow water gravity waves</b> . . . . .	<b>205</b>
7.5.1	Flat bottom gravity waves . . . . .	205
7.5.2	Structure of the gravity wave . . . . .	206
7.5.3	Dispersion relation . . . . .	207
7.5.4	Steady one-dimensional flow over an obstacle . . . . .	208
7.6	<b>Gravity waves in two layers</b> . . . . .	<b>209</b>
7.6.1	Linearized two-layer equations . . . . .	209
7.6.2	Gravity wave equations . . . . .	210
7.6.3	Dispersion relation . . . . .	211
7.6.4	Structure of a plane gravity wave . . . . .	211
7.6.5	Energetic scaling for the waves . . . . .	213
7.6.6	The depth averaged velocity and the velocity difference . . . . .	215
7.6.7	Comments . . . . .	216
7.7	<b>Kelvin waves</b> . . . . .	<b>216</b>
7.7.1	Wave solutions with a southern boundary . . . . .	216
7.7.2	Kelvin wave solutions . . . . .	217
7.8	<b>Inertia-gravity waves</b> . . . . .	<b>218</b>
7.8.1	Forced oscillator equation for horizontal velocity . . . . .	219
7.8.2	Free wave equation and potential vorticity . . . . .	219
7.8.3	Dispersion relation . . . . .	220
7.8.4	Zero frequency geostrophic mode . . . . .	220
7.8.5	Inertia-gravity wave modes . . . . .	220
7.8.6	Group velocity . . . . .	221
7.8.7	Shortwave limit for inertia-gravity waves . . . . .	222
7.8.8	Longwave limit for inertia-gravity waves . . . . .	223
7.8.9	Polarization relations for a plane inertia-gravity wave . . . . .	223
7.8.10	Energetics . . . . .	224
7.9	<b>Rossby waves</b> . . . . .	<b>226</b>
7.9.1	Dispersion relation . . . . .	226
7.9.2	Connecting to quasi-geostrophic potential vorticity . . . . .	227
7.9.3	Vorticity mechanism . . . . .	227
7.9.4	Dispersion circle for planetary Rossby waves . . . . .	227
7.9.5	Comments . . . . .	229
7.10	<b>Exercises</b> . . . . .	<b>229</b>

---

## 7.1 Loose threads

- Equatorial waves
- Energetics of Rossby waves as per Section 6.6 of [Vallis \(2017\)](#).

## 7.2 Shallow water equations

The shallow water equations of motion are given by the velocity equation (??) and thickness equation (??), written here in their Eulerian form and with zero atmospheric pressure

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta \quad (7.1a)$$

$$\partial_t h + \nabla \cdot (h \mathbf{u}) = 0. \quad (7.1b)$$

In Section 7.3 we linearize these shallow water equations and study their properties in anticipation of later study of their wave fluctuations. Before doing so, however, we formulate a wave-like equation that holds for the nonlinear shallow water model. Doing so anticipates certain features of the linear equations while offering yet another view of the shallow water system beyond that provided in Chapters ?? and ???. We pursued the analogous manipulations for acoustic waves in Section 3.4.1.

### 7.2.1 Nonlinear wave equation

Consider the shallow water equations written using the material time derivative

$$D\mathbf{u}/Dt + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad Dh/Dt = -h \nabla \cdot \mathbf{u}. \quad (7.2)$$

Taking the material time derivative of the thickness equation, and then using the velocity equation, yields

$$\frac{D}{Dt} \left[ \frac{1}{h} \frac{Dh}{Dt} \right] = - \frac{D(\nabla \cdot \mathbf{u})}{Dt} \quad (7.3a)$$

$$= -\nabla \cdot \frac{D\mathbf{u}}{Dt} + \partial_m u^n \partial_n u^m \quad (7.3b)$$

$$= -\nabla \cdot (-f \hat{\mathbf{z}} \times \mathbf{u} - g \nabla \eta) + \mathbf{S}^m{}_n \mathbf{S}^n{}_m - \mathbf{R}^m{}_n \mathbf{R}^n{}_m, \quad (7.3c)$$

where  $\mathbf{S}$  is the strain rate tensor and  $\mathbf{R}$  is the rotation tensor (both are discussed in Section ??). Rearrangement then leads to

$$\frac{D}{Dt} \left[ \frac{1}{h} \frac{Dh}{Dt} \right] - g \nabla^2 \eta = \nabla \cdot (f \hat{\mathbf{z}} \times \mathbf{u}) + \mathbf{S}^m{}_n \mathbf{S}^n{}_m - \mathbf{R}^m{}_n \mathbf{R}^n{}_m, \quad (7.4)$$

which corresponds to equation (3.19) holding for the compressible fluid. We interpret this equation as a generalized wave equation holding for movement following a shallow water fluid column. The second order material time operator acts on the thickness to propagate signals relative to the moving fluid. The linear terms,  $g \nabla^2 \eta$  and  $\nabla \cdot (f \hat{\mathbf{z}} \times \mathbf{u})$ , also appear in the linearized equations and will be discussed later, whereas  $\mathbf{S}^m{}_n \mathbf{S}^n{}_m - \mathbf{R}^m{}_n \mathbf{R}^n{}_m$  is a nonlinear source arising from gradients in the fluid velocity. We drop the nonlinear source when working with the linearized equations, and likewise the material time derivative becomes a local time derivative.

### 7.2.2 Another form of the nonlinear wave equation

To align more closely with the linear wave equation derived in Section (7.3), we consider again the velocity and thickness equations in the form of equation (7.2) yet first determine the evolution equation for the divergence of the thickness flux. For this purpose, multiply the

thickness equation by the velocity and the velocity equation by the thickness, and then add to find

$$\mathrm{D}(h \mathbf{u})/\mathrm{Dt} = -f h \hat{\mathbf{z}} \times \mathbf{u} - g h \nabla \eta - (h \mathbf{u}) \nabla \cdot \mathbf{u}. \quad (7.5)$$

Divergence of the left hand side of this equation is given by

$$\nabla \cdot [\mathrm{D}(h \mathbf{u})/\mathrm{Dt}] = \mathrm{D}[\nabla \cdot (h \mathbf{u})]/\mathrm{Dt} + (\partial_m \mathbf{u}) \cdot \nabla(h u^m) = -\mathrm{D}(\partial_t h)/\mathrm{Dt} + (\partial_m \mathbf{u}) \cdot \nabla(h u^m), \quad (7.6)$$

where we used the thickness equation for the final equality. Combining with the divergence of the right hand side of equation (7.5) yields

$$-\mathrm{D}(\partial_t h)/\mathrm{Dt} + \partial_m \mathbf{u} \cdot \nabla(h u^m) = -\nabla \cdot (f h \hat{\mathbf{z}} \times \mathbf{u}) - \nabla \cdot (g h \nabla \eta) - \nabla \cdot (h \mathbf{u} \nabla \cdot \mathbf{u}), \quad (7.7)$$

with rearrangement leading to the nonlinear wave equation

$$\frac{\mathrm{D}}{\mathrm{Dt}} \frac{\partial \eta}{\partial t} - \nabla \cdot (g h \nabla \eta) = \nabla \cdot (f h \hat{\mathbf{z}} \times \mathbf{u}) + \partial_m \mathbf{u} \cdot \nabla(h u^m) + \nabla \cdot (h \mathbf{u} \nabla \cdot \mathbf{u}), \quad (7.8)$$

where we noted that  $\partial_t h = \partial_t \eta$ . This nonlinear equation corresponds directly to the linear wave equation (7.31) derived below. It is this alternative form that readily offers a direct decomposition of the various waves appearing in the shallow water fluid. Namely, in the linear equations we drop the nonlinear sources on the right hand side, and convert the material time derivative to the local Eulerian time derivative.

### 7.2.3 Comments

This section stopped short of pursuing an analysis of nonlinear wave solutions. However, as in the analogous discussion of the compressible fluid in Section 3.4.1, these manipulations signal that wave-like equations are not restricted to linearized equations. Rather, they are basic to the nonlinear equations, though with a more complex mathematical structure that generally requires numerical or asymptotic methods to penetrate.

For the remainder of this chapter we focus on the linearized set of equations, thus enabling the use of linear analysis methods. Even though much simpler than the nonlinear equations, we uncover a rich and complex variety of linear wave phenomena within the shallow water model.

## 7.3 Linearized shallow water equations

In this section we linearize the shallow water velocity equation (7.1a) and thickness equation (7.1b) in a manner that supports subsequent analysis of small amplitude fluctuations. In the process we derive a linear wave equation for free surface undulations, with a number of source terms that support the variety of waves studied in later sections. This wave equation is the linearized version of the nonlinear shallow water equation (7.8).

### 7.3.1 Linearized thickness and velocity equations

We linearize the equations around a base state that is at rest and thus with a flat free surface. Referring to Figure ?? leads to the expressions for the layer thickness

$$h(x, y, t) = [\bar{\eta} + \eta'(x, y, t)] - [\bar{\eta}_b + \eta'_b(x, y)] = H + \eta'(x, y, t) - \eta'_b(x, y), \quad (7.9)$$

along with the corresponding velocity

$$\mathbf{u}(x, y, t) = 0 + \mathbf{u}'(x, y, t). \quad (7.10)$$

The terms  $\bar{\eta}$  and  $\bar{\eta}_b$  are area means for the free surface and bottom topography, whereas  $\eta'$  and  $\eta'_b$  are deviations from these means. We assume there are no boundary sources of volume so that  $\bar{\eta}$  is constant in time. The prime on the velocity in equation (7.10) acts as a reminder that it is driven by gradients in the fluctuating free surface,  $\eta'$ .

We make the following assumptions about the terms comprising the layer thickness.

- The layer thickness is everywhere positive,  $h > 0$ , thus ensuring there is water at each point in the layer at each time. We make this assumption since a vanishing layer thickness represents a nontrivial change in the domain that requires the full nonlinear equations.
- A nonzero  $\eta'_b$  allows for local variations in the bottom topography. We assume these variations are bounded by the mean thickness,

$$H_r \equiv H - \eta'_b > 0, \quad (7.11)$$

thus ensuring the resting thickness,  $H_r$ , is nonzero everywhere. This assumption does not place strong constraints on  $\eta'_b$ . In particular, we do not generally assume  $\eta'_b$  is small beyond assuming that  $H_r > 0$ . However, we do assume that  $\eta'_b$  is small when studying the quasi-geostrophic motion associated with Rossby waves in Sections 7.9.

- We assume the free surface undulations are much smaller than the resting thickness

$$|\eta'| \ll H_r. \quad (7.12)$$

This is the key assumption in the linearization process. It follows from our interest in studying linear wave fluctuations of the free surface.

The above assumptions lead to the linearized version of the thickness equation (7.1b)

$$\partial_t \eta' + \nabla \cdot (H_r \mathbf{u}') = 0. \quad (7.13)$$

This equation says that a time tendency of the free surface is driven by the convergence of a thickness flux, where the thickness is approximated by its resting value

$$h = H + \eta' - \eta'_b \approx H - \eta'_b = H_r. \quad (7.14)$$

We conclude from equation (7.13) that free surface transients (i.e., waves) require a nonzero horizontal convergence,  $-\nabla \cdot (H_r \mathbf{u}') \neq 0$ . It follows that waves in a flat bottom domain arise only when there is a nonzero convergence in the horizontal flow,  $-\nabla \cdot \mathbf{u}' \neq 0$ . Such horizontal convergences are fundamental to the shallow water fluid, and provide the key distinction from waves appearing in the horizontally non-divergent barotropic fluid from Chapter 6.

The linearized version of the shallow water velocity equation (7.1a) is

$$\partial_t \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -g \nabla \eta', \quad (7.15)$$

which is reached by dropping the nonlinear advection,  $(\mathbf{u}' \cdot \nabla) \mathbf{u}'$ . Advection is indeed smaller than the other terms since velocity tendencies are driven by the assumed small amplitude fluctuations in the free surface, thus making  $(\mathbf{u}' \cdot \nabla) \mathbf{u}'$  second order in small terms.

### 7.3.2 Energetics

We studied energetics of a single layer of shallow water fluid in Section ???. We here summarize how the energetics appear for the linearized thickness equation (7.13) and velocity equation (7.15). As per the norm with energetics of linearized systems (e.g., see the study of acoustic energy in Section 3.6 and surface waves in Section 4.4), we work to second order accuracy in fluctuating fields.

The gravitational potential energy per horizontal area is given by

$$\mathcal{P}^{\text{sw}} = g \rho \int_{\eta_b}^{\eta} z \, dz = (g \rho / 2) (\eta^2 - \eta_b^2), \quad (7.16)$$

and its time tendency is

$$\partial_t \mathcal{P}^{\text{sw}} = g \rho (\bar{\eta} + \eta') \partial_t \eta' \approx -g \rho (\bar{\eta} + \eta') \nabla \cdot (H_r \mathbf{u}'). \quad (7.17)$$

Likewise, the kinetic energy per area is

$$\mathcal{K}^{\text{sw}} = \frac{\rho}{2} \int_{\eta_b}^{\eta} \mathbf{u} \cdot \mathbf{u} \, dz = (\rho / 2) h \mathbf{u}' \cdot \mathbf{u}' \approx (\rho / 2) H_r \mathbf{u}' \cdot \mathbf{u}', \quad (7.18)$$

and its time tendency is

$$\partial_t \mathcal{K}^{\text{sw}} = \rho H_r \mathbf{u}' \cdot \partial_t \mathbf{u}' = -g \rho H_r \mathbf{u}' \cdot \nabla \eta' = -g \rho H_r \mathbf{u}' \cdot \nabla (\bar{\eta} + \eta'). \quad (7.19)$$

The time tendency for the mechanical energy is thus given by

$$\partial_t (\mathcal{K}^{\text{sw}} + \mathcal{P}^{\text{sw}}) = -g \rho \nabla \cdot [H_r (\bar{\eta} + \eta') \mathbf{u}']. \quad (7.20)$$

This equation is a linearized version of the mechanical energy equation (??). Here, we only have the transfer of energy due to the advection of pressure, whereas the nonlinear equations also have the advection of kinetic energy and potential energy.

### 7.3.3 Potential vorticity

The shallow water potential vorticity (Section ??) is given by

$$Q = \frac{f + \zeta}{h} = \frac{f + \zeta}{H - \eta'_b + \eta'} \approx Q' \quad (7.21)$$

where, to first order in primed quantities,

$$Q' = \frac{f + \zeta'}{H_r} - \frac{f \eta'}{H_r^2} \approx \frac{f}{H} + \frac{H \zeta' + f (\eta'_b - \eta')}{H^2}, \quad (7.22)$$

with the approximation following from assuming  $|\eta'| \ll H_r = H - \eta'_b \approx H$ . That is, we assume both small amplitude free surface fluctuations, and small amplitude bottom topography variations. Note that for the  $f$ -plane, the  $f/H$  term can be dropped since it is a constant. Use of the linearized thickness equation (7.13) and linearized velocity equation (7.15), yields the time tendency for the linearized potential vorticity (see Exercise 7.1)

$$\partial_t Q' + \mathbf{u}' \cdot \nabla (f / H_r) = 0. \quad (7.23)$$

For a flat bottom  $f$ -plane domain, the linearized potential vorticity remains static at each point in the fluid,  $\partial_t Q' = 0$ . In contrast, nonzero gradients in  $f$  or  $H_r$  cause the linearized potential vorticity to be modified through advection of  $f/H_r$ , with  $f/H_r$  the potential vorticity of the background rest state. Hence, if  $f$  and/or  $H_r$  are spatially varying, solutions to the linear equations, including waves, have an evolving potential vorticity.

### 7.3.4 Traveling plane wave ansatz and $\nabla \cdot \mathbf{v} = 0$

To derive the dispersion relation for the various shallow water waves, we consider the traveling plane wave ansatz from Section 1.5, here written in the form

$$(u', v', \eta') = (\tilde{u}, \tilde{v}, \tilde{\eta}) e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad (7.24)$$

where the real part of the right hand side is assumed. The amplitudes  $(\tilde{u}, \tilde{v}, \tilde{\eta})$  are generally complex numbers that are independent of space and time. Recall from Section ?? that the horizontal velocity has no depth dependence in the shallow water layer, which follows from making the hydrostatic approximation in the homogeneous fluid layer. Hence, a plane wave solution to the shallow water wave equations must have a zero vertical component to the wavevector,  $k_z = 0$ , so that

$$\mathbf{k} = \hat{\mathbf{x}} k_x + \hat{\mathbf{y}} k_y. \quad (7.25)$$

Evidently, the phase of shallow water plane waves (7.24) travels horizontally within a layer.

As a layer with a constant density, the fluid in the shallow water layer is incompressible so that  $\nabla \cdot \mathbf{v} = 0$ . For plane waves moving in three-dimensions, this non-divergence constraint means that the velocity of the plane waves satisfies,  $\mathbf{k} \cdot \mathbf{u} + k_z w = 0$ , (e.g., see our study of inertial waves in Chapter 5). However, for the shallow water system we saw above that  $k_z = 0$ , which would then seem to imply that  $\mathbf{k} \cdot \mathbf{u} = 0$ . Yet shallow water waves do *not* satisfy  $\mathbf{k} \cdot \mathbf{u} = 0$  since they are horizontally divergent.

The resolution of this quandary concerns the vertical velocity in a shallow water layer. Namely, the vertical velocity component is diagnosed via the horizontal convergence,  $\partial_z w = -\nabla \cdot \mathbf{u}$ . The depth independence of the horizontal velocity leads to a linear depth-dependence of the vertical velocity component (equation ??))

$$\nabla \cdot \mathbf{u} + \partial_z w = 0 \implies w(z) = w(\eta_b) - (z - \eta_b) \nabla \cdot \mathbf{u}, \quad (7.26)$$

which for the plane wave (7.24) leads to

$$w(z) = w(\eta_b) - i(z - \eta_b) \mathbf{k} \cdot \tilde{\mathbf{u}} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \quad (7.27)$$

The factor of  $i = \sqrt{-1}$  out front means that the vertical velocity component is  $\pi/2$  out of phase with the horizontal velocity. Also note that  $|z - \eta_b| |\mathbf{k}| \leq H |\mathbf{k}|$ . Since  $H |\mathbf{k}| \ll 1$  for shallow water flow, we see that the vertical velocity component has a much smaller magnitude than the horizontal velocity.

## 7.4 Unified shallow water wave equation

In this section we combine the linearized thickness equation (7.13) and linearized velocity equation (7.15) to render a unified wave equation for the free surface.<sup>1</sup> We start by closely

---

<sup>1</sup>Many of the manipulations in this section follow Lecture 14 of [Pedlosky \(2003\)](#).

following the steps taken for the nonlinear wave equation in Section 7.2.2. Yet we go much further for the purpose of obtaining a linear equation with the free surface as the only prognostic field. Although the unified wave equation is rather tedious, it does serve to unify the variety of waves supported by the rotating shallow water system. Once deriving this wave equation (equation (7.42)), we study its scaling for high frequency ( $\omega^2 > f^2$ ; super-inertial) and low frequency ( $\omega^2 < f^2$ ; sub-inertial) wave motion. That discussion anticipates analysis pursued in more depth in the remaining sections of this chapter.

### 7.4.1 Use of the divergence equation

We start by forming an equation for the evolution of the divergence of  $H_r \mathbf{u}'$ , which is found by multiplying the linearized velocity equation (7.15) by  $H_r$  and then taking the divergence

$$\partial_t [\nabla \cdot (H_r \mathbf{u}')] + \nabla \cdot (\hat{\mathbf{z}} \times f H_r \mathbf{u}') = -\nabla \cdot (g H_r \nabla \eta'), \quad (7.28)$$

where the divergence and time derivative commute. The thickness equation (7.13) can be used to eliminate the divergence, in which case we are led to

$$\partial_{tt} \eta' - \nabla \cdot (g H_r \nabla \eta') = -\hat{\mathbf{z}} \cdot [\nabla \times (f H_r \mathbf{u}')]. \quad (7.29)$$

where we made use of the identity

$$\nabla \cdot (\hat{\mathbf{z}} \times f H_r \mathbf{u}') = -\hat{\mathbf{z}} \cdot [\nabla \times (f H_r \mathbf{u}')]. \quad (7.30)$$

The wave equation (7.29) is the linear analog of the nonlinear equation (7.8). We can expand the cross product on the right hand side to identify the variety of physical processes supporting shallow water waves

$$\partial_{tt} \eta' = \underbrace{\nabla \cdot (g H_r \nabla \eta')}_{\text{gravity waves}} - \underbrace{\hat{\mathbf{z}} \cdot (f H_r \nabla \times \mathbf{u}')}_{\text{inertial waves}} - \underbrace{\hat{\mathbf{z}} \cdot (H_r \nabla f \times \mathbf{u}')}_{\text{planetary Rossby waves}} - \underbrace{\hat{\mathbf{z}} \cdot (f \nabla H_r \times \mathbf{u}')}_{\text{topographic Rossby waves}}. \quad (7.31)$$

Although suggestive, it is useful to further manipulate this equation into one that involves a single prognostic field, here chosen to be the free surface. In so doing we provide a unified discussion of the dispersion relation for the various waves admitted by the shallow water model, and allow for a seamless decomposition of these motions in terms of space and time scales. We pursue the somewhat tedious manipulations in Section 7.4.2, but only after considering a heuristic discussion of the physical processes in equation (7.31).

#### Non-rotating gravity waves

Without rotation ( $f = 0$ ) yet with gravity ( $g \neq 0$ ), equation (7.31) reduces to the shallow water gravity wave equation

$$\partial_{tt} \eta' - \nabla \cdot (g H_r \nabla \eta') = 0. \quad (7.32)$$

With a flat bottom, these waves are non-dispersive plane gravity waves, whereas the case with a gently sloping bottom requires the WKBJ method developed in Section 3.9. These waves are sometimes referred to as *long gravity waves*, in contrast to the shorter waves that occur in deep water (see Section 4.6). We study shallow water gravity waves in Section 7.5.

## Inertia-gravity waves

The wave equation

$$\partial_{tt}\eta' - \nabla \cdot (g H_r \nabla \eta') = -f H_r \zeta' \quad (7.33)$$

describes dispersive inertia-gravity waves in a shallow water layer. The inertial portion of the waves rely on both rotation ( $f \neq 0$ ) and vorticity ( $\zeta' \neq 0$ ). In Chapter 5 we studied inertial waves in a homogeneous fluid. The key difference here is that the shallow water fluid is hydrostatic, which, as we will see, alters the dispersion relation relative to that derived for the non-hydrostatic homogeneous fluid in Chapter 5. It furthermore motivates us to examine inertial waves along with gravity waves since a hydrostatic fluid only arises in a gravity field. Such *inertia-gravity* or *Poincaré waves*, as they appear in the shallow water fluid, are the topic of Section 7.8.

## Planetary and topographic Rossby waves

The remaining two terms on the right hand side of equation (7.31) lead to waves supported by a background vorticity gradient. The *planetary Rossby waves* rely on a nonzero planetary vorticity gradient (i.e., the  $\beta$  effect), whereas *topographic Rossby waves* rely on a nonzero gradient in the bottom topography. Both of these waves are dispersive, and both have frequencies lower than inertia-gravity waves. As shown in Section 7.4.6, and as already detailed for the horizontally non-divergent barotropic model in Section 6.3, the wave equation for Rossby waves has just a single time derivative acting on the free surface. This situation contrasts to the two time derivatives found for inertia-gravity waves and currently appearing in equation (7.31). Although shallow water Rossby waves rely on a nonzero  $\nabla(H_r f) \times \mathbf{u}'$ , we postpone writing their wave equation until working through some more technical details.

In Chapter 6 we studied many properties of Rossby waves in the horizontally non-divergent barotropic fluid, with these waves also relying on a gradient in the background (potential) vorticity. It is thus not surprising that shallow water Rossby waves share key physical properties with Rossby waves in the non-divergent barotropic flow. We require more technical steps to extract the physics within the shallow water model, and to expose their key distinction from Rossby waves in the horizontally non-divergent barotropic model. Note also that the shallow water model supports topographic Rossby waves. In contrast, no such waves appear in the horizontally non-divergent barotropic model, since that flow is restricted to strictly move along constant depth contours, with no cross isobath motion.

### 7.4.2 Use of the vorticity equation

We now return to the wave equation (7.31) and perform some further manipulations in a quest to derive an equation that only has the free surface along with geophysical parameters. Start by writing the divergence equation (7.29) in its form that extracts the  $\beta$  term

$$\partial_{tt}\eta' - \nabla \cdot (g H_r \nabla \eta') = -f \hat{\mathbf{z}} \cdot [\nabla \times (H_r \mathbf{u}')] + \beta H_r u. \quad (7.34)$$

Next work with  $\hat{\mathbf{z}} \cdot [\nabla \times (H_r \mathbf{u}')]$ , which is the vorticity of the thickness (resting thickness) weighted velocity. To compute its time evolution, multiply the velocity equation (7.15) by  $H_r$  and take the curl

$$\nabla \times \partial_t(H_r \mathbf{u}') = -\nabla \times (\hat{\mathbf{z}} \times f H_r \mathbf{u}') - g \nabla H_r \times \nabla \eta' \quad (7.35a)$$

$$= -\hat{\mathbf{z}} [\nabla \cdot (f H_r \mathbf{u})] - g \nabla H_r \times \nabla \eta', \quad (7.35b)$$

where the second equality made use of the identity

$$\nabla \times (\hat{\mathbf{z}} \times f H_r \mathbf{u}') = \hat{\mathbf{z}} [\nabla \cdot (f H_r \mathbf{u})]. \quad (7.36)$$

Now take the time derivative of the divergence equation (7.34) and use the vorticity equation (7.35b) to find

$$\partial_t [\partial_{tt} \eta' - \nabla \cdot (g H_r \nabla \eta')] = -f \hat{\mathbf{z}} \cdot [\nabla \times \partial_t (H_r \mathbf{u}')] + \beta H_r \partial_t u' \quad (7.37a)$$

$$= f \nabla \cdot (f H_r \mathbf{u}) + f g \hat{\mathbf{z}} \cdot (\nabla H_r \times \nabla \eta') + \beta H_r \partial_t u' \quad (7.37b)$$

$$= -f^2 \partial_t \eta' + f g \hat{\mathbf{z}} \cdot (\nabla H_r \times \nabla \eta') + \beta H_r (f v' + \partial_t u'). \quad (7.37c)$$

Rearrangement leads to

$$\partial_t [\mathcal{L}(\eta') - \nabla \cdot (g H_r \nabla \eta')] = f g \hat{\mathbf{z}} \cdot (\nabla H_r \times \nabla \eta') + \beta H_r (f v' + \partial_t u'), \quad (7.38)$$

where we introduced the linear time operator as a shorthand

$$\mathcal{L} = \partial_{tt} + f^2. \quad (7.39)$$

This operator commutes with time derivatives but commutes with space derivatives only for the  $f$ -plane. For the case of  $\beta = 0$ , equation (7.38) only has the free surface height as a prognostic field. Hence, this equation is suited to developing the dispersion relation for gravity waves and inertial waves. Yet, as we show in the following, we need more work for sub-inertial Rossby waves.

### 7.4.3 The unified wave equation just with $\eta'$

The final step is to eliminate the term  $f v' + \partial_t u'$  from equation (7.38) in favor of terms proportional to the free surface. The key step is to make use of the following identities derived in Exercise 7.2

$$\mathcal{L}(u') = -g (\partial_{xt} \eta' + f \partial_y \eta') \quad \text{and} \quad \mathcal{L}(v') = -g (\partial_{yt} \eta' - f \partial_x \eta'). \quad (7.40)$$

Acting with  $\mathcal{L}$  on equation (7.38), and using the identities (7.40), yield

$$\partial_t [\mathcal{L}[\mathcal{L}(\eta') - \nabla \cdot (g H_r \nabla \eta')]] = f g \hat{\mathbf{z}} \cdot [\nabla H_r \times \nabla \mathcal{L}(\eta')] + \beta H_r f \mathcal{L}(v') + \beta H_r \partial_t \mathcal{L}(u'). \quad (7.41)$$

Expanding the linear operator and rearranging leads to the desired unified equation

$$\partial_t [\mathcal{L}[\mathcal{L}(\eta') - \nabla \cdot (g H_r \nabla \eta')]] = f g \hat{\mathbf{z}} \cdot [\nabla H_r \times \nabla \mathcal{L}(\eta')] + g \beta H_r (f^2 \partial_x \eta' - 2 f \partial_{yt} \eta' - \partial_{xtt} \eta'). \quad (7.42)$$

As shown in the following sections, this linear partial differential equation encapsulates the full suite of linear wave processes active in a shallow water layer on a  $\beta$ -plane with topography. However, it is not so simple to parse in its current form, thus motivating the examination of limiting forms of this equation. In particular, we consider super-inertial waves ( $\omega^2 > f^2$ ) versus sub-inertial waves ( $\omega^2 < f^2$ ). In both cases we assume horizontal length scales according to the  $\beta$ -plane, along with small fluctuations of the topography.

#### 7.4.4 Scaling for the $\beta$ -plane with small topography

To help analyze the super-inertial and sub-inertial wave motions contained in equation (7.42), introduce the following length and time scales of the motions

$$L = \text{horizontal scale} \quad \text{and} \quad \omega = \text{angular frequency}. \quad (7.43)$$

We work under the assumptions of a  $\beta$ -plane approximation (see equation (??)) so that the horizontal scales of motion satisfy

$$\beta L \ll |f_0|, \quad (7.44)$$

where we wrote the Coriolis parameter as  $f = f_0 + \beta y$ . We also assume the topography undulation,  $\eta'_b$ , is much smaller than the mean layer thickness,

$$\eta'_b \ll H. \quad (7.45)$$

Both approximations (7.44) and (7.45) are familiar from the development of quasi-geostrophy theory in Section ???. Finally, assume  $f > 0$  as per the northern hemisphere, with occurrences of  $f$  changed to  $|f|$  for southern hemisphere results.

Introducing the above scales into the wave equation (7.42) leads to

$$\partial_t \mathcal{L}^2(\eta') \sim \omega (\omega^2 + f^2)^2 \eta' \quad (7.46a)$$

$$\partial_t \mathcal{L}[\nabla \cdot (g H_r \nabla \eta')] \sim \omega (\omega^2 + f^2) (g H_r / L^2) \eta' \quad (7.46b)$$

$$f g \nabla H_r \times \nabla \mathcal{L}(\eta') \sim (f g \eta'_b / L^2) (\omega^2 + f^2) \eta' \quad (7.46c)$$

$$g \beta H_r f^2 \partial_x \eta' \sim (g \beta H_r f^2 / L) \eta' \quad (7.46d)$$

$$-2 \beta g H_r f \partial_{yt} \eta' \sim (\beta g H_r f \omega / L) \eta' \quad (7.46e)$$

$$\beta g H_r \partial_{xtt} \eta' \sim \beta g H_r (\omega^2 / L) \eta'. \quad (7.46f)$$

Next, we separately examine the super-inertial regime,  $\omega^2 > f^2$ , and the sub-inertial regime,  $\omega^2 < f^2$ .

#### 7.4.5 Super-inertial wave equation

In this subsection we derive an approximate wave equation relevant for super-inertial motions ( $\omega^2 > f^2$ ).

##### Super-inertial scaling for the LHS of equation (7.42)

From equations (7.46a) and (7.46b), the two terms on the left hand side of the wave equation (7.42) scale as

$$\text{LHS} = \partial_t \mathcal{L}^2(\eta') - \partial_t \mathcal{L}[\nabla \cdot (g H_r \nabla \eta')] \sim \omega^3 [\omega^2 + g H_r / L^2] \eta'. \quad (7.47)$$

Assuming the horizontal length scale is given by the inverse wavenumber,  $L = |\mathbf{k}|^{-1}$ , yields

$$g H_r / L^2 \approx c_{\text{grav}}^2 |\mathbf{k}|^2 \approx \omega^2 \implies L^2 \sim g H_r / \omega^2. \quad (7.48)$$

The first approximation follows from assuming  $\eta'_b \ll H$  so that  $g H_r \approx g H = c_{\text{grav}}^2$ , and the second approximation follows from assuming a non-rotating gravity wave dispersion relation.<sup>2</sup> In essence, we are assuming the two terms in equation (7.47) scale the same, which holds so long as the length scale of the super-inertial waves satisfies  $L^2 \sim g H_r / \omega^2$ .

### Ratio of RHS terms to the LHS

Now consider the magnitude for each term on the right hand side of the wave equation (7.42) relative to the left hand side terms just found in equation (7.47). The ratio for terms in equation (7.46c) satisfy

$$\frac{f g \nabla H_r \times \nabla \mathcal{L}(\eta')}{\text{LHS}} \sim \frac{\omega^2 f g \eta'_b / L^2}{\omega^3 g H_r / L^2} = \frac{f \eta'_b}{\omega H_r} \ll 1, \quad (7.49)$$

where the inequality follows from assuming that the bottom topography undulation is small as per equation (7.45). The ratio for the remaining three terms (equations (7.46d), (7.46e), and (7.46f)) are also much less than unity, with these inequalities following from the  $\beta$ -plane approximation (7.44)

$$\frac{g \beta H_r f^2 \partial_x \eta'}{\text{LHS}} \sim \frac{g \beta H_r f^2 / L}{\omega^3 g H_r / L^2} = (\beta L / \omega) (f / \omega)^2 < (\beta L / f) (f / \omega)^2 \ll 1 \quad (7.50a)$$

$$\frac{-2 f \partial_{yt} \eta'}{\text{LHS}} \sim \frac{\beta g H_r f \omega / L}{\omega^3 g H_r / L^2} = (\beta L / \omega) (f / \omega) < (\beta L / f) (f / \omega) \ll 1 \quad (7.50b)$$

$$\frac{\beta g H_r \partial_{xtt} \eta'}{\text{LHS}} \sim \frac{\beta g H_r (\omega^2 / L)}{\omega^3 g H_r / L^2} = (\beta L / \omega) < (\beta L / f) \ll 1. \quad (7.50c)$$

### Inertia-gravity wave equation for super-inertial waves

We thus conclude that for super-inertial waves on a  $\beta$ -plane and with small amplitude topography, then the two terms on the left hand side of equation (7.42) balance so that

$$\partial_t \mathcal{L}[\mathcal{L}(\eta') - \nabla \cdot (g H_r \nabla \eta')] \approx 0. \quad (7.51)$$

This relation can be maintained by setting

$$(\partial_{tt} + f_o^2) \eta' - \nabla \cdot (g H_r \nabla \eta') = 0, \quad (7.52)$$

which is the equation for inertia-gravity waves on an  $f$ -plane with small amplitude topography. Note that we put  $f = f_o$  in equation (7.52) since the  $\beta$  effect is most crucial for sub-inertial wave motions as discussed in Section 7.4.6, rather than super-inertial motions. We study non-rotating gravity waves in Section 7.5 and inertia-gravity waves in Section 7.8.

### 7.4.6 Sub-inertial wave equation

We here derive an equation relevant for sub-inertial wave motions ( $\omega^2 < f^2$ ). In fact, we further restrict the motion to those that are very low frequency so that

$$\omega / f \sim \beta L / f_o \sim \eta'_b / H \ll 1. \quad (7.53)$$

---

<sup>2</sup>We derive the shallow water gravity wave dispersion relation in equation (7.80). Also, see the discussion of long surface gravity waves in Section 4.5.3.

We also assume that the horizontal length scales are on the order of

$$L \sim L_d = c_{\text{grav}}/f = \sqrt{g H}/f, \quad (7.54)$$

which is known as the *deformation radius* for the shallow water layer, first introduced in Section ?? . This length scale appears throughout our discussion of rotating shallow water waves.

The assumed scalings (7.53) along with (7.54) are precisely those assumed for shallow water quasi-geostrophy as detailed in Section ?? . The corresponding sub-inertial waves are referred to as *planetary Rossby waves* and *topographic Rossby waves*. We here derive the wave equation for these waves, with further details of their dispersion relation provided in Section 7.9.

### Sub-inertial scaling for the LHS of equation (7.42)

For the sub-inertial scalings (7.53), the linear operator is order  $\mathcal{L} \sim \mathcal{O}(f^2)$ , so that the left hand side of the wave equation (7.42) scales as

$$\text{LHS} = \partial_t \mathcal{L}(\eta') + \partial_t \mathcal{L}[\nabla \cdot (g H_r \nabla \eta')] \sim \omega f^2 [f^2 + g H_r/L^2] \eta'. \quad (7.55)$$

As for the super-inertial waves in Section 7.4.5, we assume the two terms in the right expression scale the same. To ensure that scaling holds, we focus on horizontal length scales satisfying

$$L^2 \sim g H_r/f^2 \sim g H/f^2 = c_{\text{grav}}^2/f^2 = L_d^2. \quad (7.56)$$

This assumption of deformation scale motions for sub-inertial waves can be compared to the super-inertial length scale in equation (7.48). We find that super-inertial motions are generally smaller than the deformation radius whereas sub-inertial motions are on the order or larger than the deformation radius.

With the above scaling (7.56) for the horizontal length, the left hand side of the wave equation (7.42) scales as

$$\text{LHS} \sim (\omega f g H_r/L^2) \eta'. \quad (7.57)$$

We will later replace  $f$  with  $f_0$  in equation (7.57), with that replacement warranted by the  $\beta$ -plane approximation,  $\beta L \ll f_0$ . For now, we reduce clutter by just writing  $f$ .

### Ratio of RHS terms to the LHS

Now consider the magnitude for each term on the right hand side of the wave equation (7.42) relative to the left hand side terms just found in equation (7.57). The ratio for terms in equation (7.46c) is given by

$$\frac{f g \nabla H_r \times \nabla \mathcal{L}(\eta')}{\text{LHS}} \sim \frac{f^3 g \eta'_b/L^2}{\omega f^2 g H_r/L^2} = \frac{f \eta'_b}{\omega H_r} = \mathcal{O}(1), \quad (7.58)$$

with the final equality following from the assumed scales in equation (7.53). We also find that the ratio for terms in equation (7.46d) is given by

$$\frac{g \beta H_r f^2 \partial_x \eta'}{\text{LHS}} \sim \frac{g \beta H_r f^2/L}{\omega f^2 g H_r/L^2} = \beta L/\omega = \mathcal{O}(1), \quad (7.59)$$

where we again made use of equation (7.53) for the final equality. In contrast to these order unity ratios, the ratios involving terms in equations (7.46e) and (7.46f) are much less than unity

$$\frac{-2f\partial_{yt}\eta'}{\text{LHS}} \sim \frac{\beta g H_r f \omega / L}{\omega f^2 g H_r / L^2} = \beta L / f \ll 1 \quad (7.60\text{a})$$

$$\frac{\beta g H_r \partial_{xtt}\eta'}{\text{LHS}} \sim \frac{\beta g H_r (\omega^2 / L)}{\omega f^2 g H_r / L^2} = (\beta L / f) (\omega / f) \ll 1. \quad (7.60\text{b})$$

The first inequality follows from the  $\beta$ -plane approximation (7.44), and the second follows also from the  $\beta$ -plane approximation as well as the low frequency assumption,  $\omega^2 \ll f^2$ .

### Shallow water planetary and topographic Rossby wave equation

The above scalings lead to the approximate wave equation for sub-inertial waves moving on a  $\beta$ -plane with small amplitude bottom topography

$$\partial_t [f^2 \eta' - \nabla \cdot (g H \nabla \eta')] = f g \hat{z} \cdot (\nabla H_r \times \nabla \eta') + g H \beta \partial_x \eta'. \quad (7.61)$$

Note that we set  $H_r = H$  everywhere except where it is differentiated, with this replacement consistent with the assumed  $\eta'_b \ll H$  (equation (7.53)). A slightly more tidy version arises by introducing the squared gravity wave speed  $c_{\text{grav}}^2 = g H$  and the deformation radius,  $L_d = c_{\text{grav}} / f$ , so that

$$\partial_t [(L_d^{-2} - \nabla^2) \eta'] = (f/H) \hat{z} \cdot (\nabla H_r \times \nabla \eta') + \beta \partial_x \eta'. \quad (7.62)$$

The first term on the right hand side gives rise to topographic Rossby waves, whereas the second term gives rise to planetary Rossby waves. It is notable that this equation has only a single time derivative, which contrasts with the super-inertial wave equation (7.52).

### Rossby waves are supported by potential vorticity gradients

We can write the right hand side of equation (7.62) in the following equivalent form

$$\partial_t [(L_d^{-2} - \nabla^2) \eta'] = H \hat{z} \cdot (\nabla \eta' \times \nabla Q_r). \quad (7.63)$$

In this manner we see that both planetary Rossby waves and topographic Rossby waves are supported by gradients in the resting fluid's potential vorticity<sup>3</sup>

$$Q_r = f / H_r. \quad (7.64)$$

Equation (7.63) reveals that shallow water Rossby waves arise from a misalignment between contours of constant resting state potential vorticity and contours of constant surface height.

### The effective $\beta$ vector

We offer one further expression for the Rossby wave equation

$$\partial_t [(L_d^{-2} - \nabla^2) \eta'] = \hat{z} \cdot (\nabla \eta' \times \boldsymbol{\beta}_{\text{eff}}), \quad (7.65)$$

---

<sup>3</sup>When studying waves in the horizontally non-divergent barotropic fluid (Section 6.3), we also noted the need for a background potential vorticity gradient to support Rossby wave modes.

where we introduced the effective beta vector that combines the planetary vorticity gradient with the topographic gradient

$$\boldsymbol{\beta}_{\text{eff}} = \nabla f - Q_r \nabla H_r = \beta \hat{\mathbf{y}} - Q_r \nabla H_r. \quad (7.66)$$

Again, this form exhibits the parallel role of both planetary beta and topographic slopes in supporting Rossby waves. Equation (7.65) says that shallow water Rossby waves arise from a misalignment between the effective beta vector and free surface height gradients.

### 7.4.7 Comments on bottom topography

Throughout this section we allowed for the bottom topography, as defined by  $H_r = H - \eta'_b$  (equation (7.11)), to have arbitrary  $(x, y)$  spatial dependence. In the following sections we make simplifications to the topography, thus facilitating the study of plane waves, planetary waves, and topographically trapped waves. The study of waves with more general bottom topographies requires the asymptotic methods from Chapter 2.

## 7.5 Shallow water gravity waves

In this section we study shallow water gravity waves in the absence of planetary rotation ( $f = 0$ ). We already encountered features of these waves in Chapter 4 when studying surface gravity waves. Even so, it is useful to start from the shallow water equations to directly derive the properties of the non-dispersive gravity waves. Additionally, we here do not introduce a velocity potential as done for the surface gravity waves, even though the non-rotating shallow water gravity waves are irrotational. We do not introduce the velocity potential since the other waves studied in this chapter carry vorticity, thus making the velocity potential an incomplete description for those cases.

### 7.5.1 Flat bottom gravity waves

Setting  $\eta'_b = 0$  in equation (7.32) yields the shallow water gravity wave equation holding for a flat bottom domain

$$(\partial_{tt} - c_{\text{grav}}^2 \nabla^2) \eta' = 0, \quad (7.67)$$

where the gravity wave speed is determined by the two geophysical properties of the system

$$c_{\text{grav}} = \sqrt{g H}. \quad (7.68)$$

To help emphasize a few key properties, we return to the linear thickness equation (7.13) and linear velocity equation (7.15), now written with  $f = 0$  and  $\eta'_b = 0$

$$\partial_t \eta' = -H \nabla \cdot \mathbf{u}' \quad \text{and} \quad \partial_t \mathbf{u}' = -g \nabla \eta'. \quad (7.69)$$

Taking the time derivative of the thickness equation and then using the velocity equation readily yields the wave equation (7.67). Conversely, taking the horizontal divergence of the velocity equation yields

$$\partial_t (\nabla \cdot \mathbf{u}') = -g \nabla^2 \eta', \quad (7.70)$$

so that time changes in the horizontal divergence are driven by curvature in the free surface. Taking the time derivative of this equation, and using the thickness equation, reveals that the

horizontal divergence also satisfies the wave equation

$$(\partial_{tt} - c_{\text{grav}}^2 \nabla^2) (\nabla \cdot \mathbf{u}') = 0. \quad (7.71)$$

Hence, both the free surface and the horizontal divergence travel as non-dispersive gravity waves in the flat bottom non-rotating shallow water layer.

Finally, observe that the linearized potential vorticity equation (7.23), in the presence of  $f = 0$  and  $\eta'_b = 0$ , means that the relative vorticity is static at each point in the fluid

$$\partial_t \zeta' = 0. \quad (7.72)$$

Hence, these gravity wave fluctuations do not alter the flow vorticity. In particular, if the flow starts with zero vorticity then shallow water gravity wave fluctuations retain the zero vorticity.

### 7.5.2 Structure of the gravity wave

Substituting the plane wave ansatz (7.24) into the linearized thickness and velocity equation (7.69) allows us to connect the wave amplitudes for the free surface and velocity

$$\omega \tilde{\mathbf{u}} = g \mathbf{k} \tilde{\eta} \implies \mathbf{k} \times \tilde{\mathbf{u}} = 0 \quad \text{and} \quad \omega \tilde{\eta} = H \mathbf{k} \cdot \tilde{\mathbf{u}}. \quad (7.73)$$

The identity  $\mathbf{k} \times \tilde{\mathbf{u}} = 0$  for the plane wave means that the fluid particle's horizontal velocity is oriented parallel to the horizontal wave vector, so that

$$\tilde{\mathbf{u}} = \hat{\mathbf{k}} (\hat{\mathbf{k}} \cdot \tilde{\mathbf{u}}) \quad \text{with} \quad \hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|. \quad (7.74)$$

This alignment of the horizontal velocity with the wavevector is a direct result of the irrotational nature of the gravity waves, which was already noted when discussing the linearized vorticity equation (7.72). Additionally, with a real wavevector and real angular frequency, the amplitude relation  $\omega \tilde{\mathbf{u}} = g \mathbf{k} \tilde{\eta}$  also holds for the traveling plane wave velocity and free surface so that

$$\omega \mathbf{u} = g \mathbf{k} \eta \implies \mathbf{k} \cdot \mathbf{u} = g |\mathbf{k}|^2 \eta / \omega \quad \text{and} \quad \mathbf{u} = \hat{\mathbf{k}} (g/c_{\text{grav}}) \tilde{\eta} \cos \mathcal{P}, \quad (7.75)$$

where

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t \quad (7.76)$$

is the phase.

Alignment of the horizontal fluid particle velocity with the wavevector (equation (7.75)) indicates that motion in the shallow water gravity wave is horizontally longitudinal. However, as noted in Section 7.3.4, a shallow water fluid has a nonzero vertical velocity component that is a linear function of depth, as given by equations (7.26) and (7.27). The presence of vertical motion within the wave means that the waves are not longitudinal in three dimensions. Rather, a fluid particle moves horizontally in the direction of the wavevector but the particle also moves vertically, thus tracing out an elliptical path in the vertical-horizontal plane.

To describe the vertical particle motion in the wave, make use of equation (7.27) for the vertical velocity and equation (7.75) for the horizontal velocity, thus leading to

$$w = -i(z - \eta_b) \mathbf{k} \cdot \tilde{\mathbf{u}} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} = (z - \eta_b) (g |\mathbf{k}|^2 \tilde{\eta} / \omega) (-i \cos \mathcal{P} + \sin \mathcal{P}), \quad (7.77)$$

where we set  $w(\eta_b) = 0$  for a flat bottom. Taking the real part renders the vertical velocity

within the shallow water gravity wave

$$w = (z - \eta_b) (g |\mathbf{k}|^2 / \omega) \tilde{\eta} \sin \mathcal{P} = (z - \eta_b) |\mathbf{k}| (g/c_{\text{grav}}) \tilde{\eta} \sin \mathcal{P} \quad (7.78)$$

The vertical velocity is  $\pi/2$  out of phase with both the free surface and the horizontal velocity. Furthermore, the ratio of the maximum magnitude for the vertical and horizontal velocities is given by

$$\frac{|w|_{\max}}{|\mathbf{u}|_{\max}} = (z - \eta_b) |\mathbf{k}|. \quad (7.79)$$

Shallow water gravity waves are characterized by wavelengths that are long relative to the fluid depth. Hence, even at the free upper surface, the ratio (7.79) is much less than unity. As a result, the elliptical particle paths are longer in the horizontal direction than vertical direction. Furthermore, since the horizontal motion is depth-independent, the elliptical particle trajectories have the same excursion along the major axis (the horizontal axis) throughout the layer depth. This behavior contrasts to the deepwater gravity waves studied in Chapter 4, whose amplitude in all directions decreases with depth. In Figure 7.1, we depict the motion of fluid particles with a shallow water gravity wave. In Section 4.11.4 we study the Stokes drift resulting from the phase averaged particle motion in these waves.

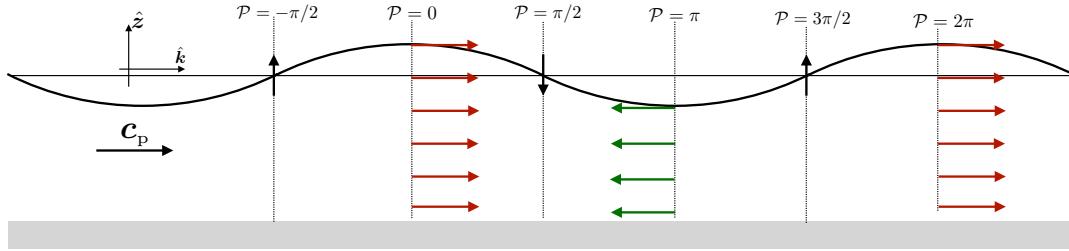


FIGURE 7.1: Illustrating motion within a shallow water gravity wave over a flat bottom, here depicted at a fixed point in time with the phase velocity to the right. The horizontal arrows are the horizontal velocity as per equation (7.75), and the vertical arrows are for the vertical velocity as per equation (7.78). The horizontal and vertical velocity components are  $\pi/2$  out of phase, so that the vertical velocity vanishes when the horizontal velocity has maximum magnitude, and vice versa. This figure is a slightly simplified version of the lower panel in Figure 4.2.

### 7.5.3 Dispersion relation

Substitution of the plane wave ansatz (7.24) into the wave equation (7.67) leads to

$$\omega^2 = g H |\mathbf{k}|^2 = c_{\text{grav}}^2 |\mathbf{k}|^2. \quad (7.80)$$

Taking the positive square root then leads to the shallow water gravity wave dispersion relation<sup>4</sup>

$$\omega = |\mathbf{k}| c_{\text{grav}} \implies C_p = \omega / |\mathbf{k}| = c_{\text{grav}}, \quad (7.81)$$

where we identified the phase speed as the gravity wave speed

$$C_p = c_{\text{grav}} = \sqrt{g H}. \quad (7.82)$$

<sup>4</sup>Recall that we always consider the angular frequency of waves to be non-negative (Section 1.2), hence we only take the positive root for the dispersion relation (7.81).

Each wavevector propagates at the same speed since the phase speed is only dependent on geophysical parameters (gravitational acceleration and the resting layer thickness). Equivalently, higher wavenumber waves have correspondingly higher frequency, and the relation between the two is linear. We conclude that non-rotating shallow water gravity waves are *non-dispersive*.

#### 7.5.4 Steady one-dimensional flow over an obstacle

Consider a steady one-dimensional inviscid flow of a single shallow water layer of fluid in a channel with a varying bottom,  $z = \eta_b(x)$ , such as depicted in Figure 7.2. This flow approximates that in a straight canal as water moves over a *weir*. The thicker water upstream of the weir creates a pressure gradient that speeds up the flow as it moves over the weir. Throughout the analysis we assume the fluid has a positive layer thickness,  $h > 0$ , so that there is no region where the fluid layer vanishes.

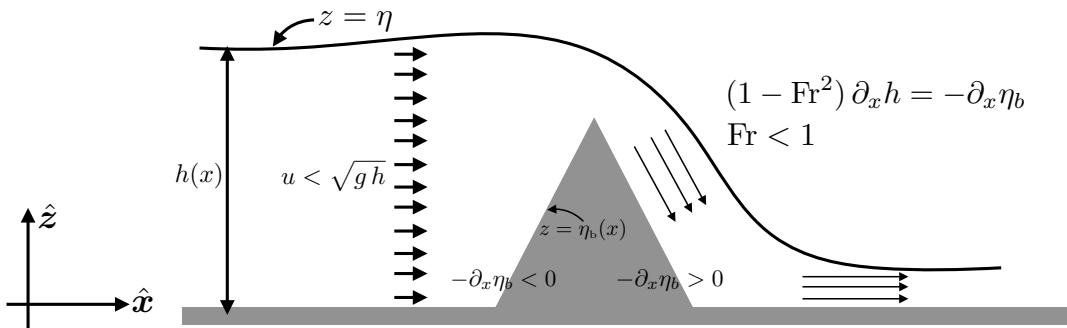


FIGURE 7.2: One-dimensional flow of a shallow water fluid layer in a channel with an obstacle. This flow approximates that for water flowing over a weir in a straight canal. If a fluid particle moves faster than the locally defined gravity wave speed,  $u > \sqrt{gh}$  or  $\text{Fr} > 1$ , then the flow is said to undergo a *hydraulic jump*, which typically occurs first at the peak of the weir with  $u = \sqrt{gh}$  or  $\text{Fr} = 1$ . Flow that is faster than the gravity wave speed is said to be *super-critical*, whereas slower flow is *sub-critical*. We here depict sub-critical flow where the local Froude number is everywhere less than unity,  $\text{Fr} < 1$ . The fluid is thicker upstream than downstream of the weir as per equation (7.89), thus enabling a higher pressure upstream to create a pressure drop across the weir.

The steady, one-dimensional, non-rotating shallow water equations take the form

$$u \partial_x u = -g \partial_x (\eta_b + h) \quad (7.83a)$$

$$u \partial_x h + h \partial_x u = 0, \quad (7.83b)$$

where we wrote the free surface height,  $\eta$ , as the sum of the bottom topography position plus the layer thickness (see Figure ??)

$$\eta = \eta_b + h. \quad (7.84)$$

The momentum equation can be written in the form of a mechanical energy equation

$$\partial_x (u^2/2 + gh) = -g \partial_x \eta_b. \quad (7.85)$$

We see that the mechanical energy,  $u^2/2 + gh$ , has spatial variations according to the variations in the bottom topography. Otherwise, the mechanical energy is a spatial constant.

Writing the steady state thickness equation (7.83b) in the form

$$\partial_x u = -(u/h) \partial_x h, \quad (7.86)$$

allows us to write the momentum equation (7.83a) as

$$(g - u^2/h) \partial_x h = -g \partial_x \eta_b. \quad (7.87)$$

Introducing the local *Froude number*,

$$\text{Fr} = u / \sqrt{gh}, \quad (7.88)$$

then leads to

$$(1 - \text{Fr}^2) \partial_x h = -\partial_x \eta_b. \quad (7.89)$$

The bottom is flat in regions away from the weir, so that  $\partial_x \eta_b = 0$ , in which case the left hand side of equation (7.89) must vanish. The left hand side vanishes if  $\partial_x h = 0$ , in which the free surface is flat, such as depicted for the region away from the weir in Figure 7.2. The left hand side also vanishes if the Froude number is unity,  $\text{Fr} = 1$ . A unit Froude number means that the particle velocity equals to the local gravity wave speed

$$u^2 = gh. \quad (7.90)$$

Flow moving at speeds less than the gravity wave speed is called *sub-critical*, such as flow upstream of the weir. Flow whose speed equals the gravity wave speed is said to be under *hydraulic control*, which commonly occurs near the peak of the weir in Figure 7.2. Flow that is faster than the local gravity wave speed is said to be *super-critical*, with the steady assumptions going into equation (7.89) breaking down for super-critical flow. Super-critical flow is found to exhibit a *hydraulic jump*, which is an instability with flow seemingly overtaking itself since it cannot “see” where it is going. The situation is akin to the sonic boom occurring in a compressible fluid moving at speeds greater than the acoustic wave speed.

The discussion here is rather descriptive, with more analysis required to deductively support the presentation, particularly for super-critical flow. Chapter 1 of [Pratt and Whitehead \(2008\)](#) considers this topic in more detail. The remaining chapters in [Pratt and Whitehead \(2008\)](#) consider planetary rotation, thus developing the subject of *rotating hydraulics*, which is particularly important for describing geophysical flows moving over topographic slopes.

## 7.6 Gravity waves in two layers

When adding more shallow water layers, how much can we use from the single layer results? To help answer that question, consider the linearized equations for two shallow water layers. The equations can be generalized to multiple layers, though with increased algebraic complexity. Indeed, even the two-layer case with rotation proves to be algebraically tedious, thus motivating us to examine just the case of gravity waves without rotation. For more general depth dependence, we find it simpler analytically to study the continuously stratified fluid in Chapter 9.

### 7.6.1 Linearized two-layer equations

We derived the equations for a stacked shallow water model in Section ??, and focused on two layers in Section ?? . With reference to Figure ?? for notation, write the layer thicknesses in the

form

$$h_1 = H_1 + h'_1 = \eta_{1/2} - \eta_{3/2} = (\overline{\eta_{1/2}} - \overline{\eta_{3/2}}) + (\eta'_{1/2} - \eta'_{3/2}) \quad (7.91\text{a})$$

$$h_2 = H_2 + h'_2 = \eta_{3/2} - \eta_b = (\overline{\eta_{3/2}} - \overline{\eta_b}) + (\eta'_{3/2} - \eta'_b), \quad (7.91\text{b})$$

where we introduced the resting layer thicknesses

$$H_1 = \overline{\eta_{1/2}} - \overline{\eta_{3/2}} \quad \text{and} \quad H_2 = \overline{\eta_{3/2}} - \overline{\eta_b} \quad \text{and} \quad H = H_1 + H_2, \quad (7.92)$$

with the overbar representing the area mean, and with the half-integer labels representing interface fields. This notation leads to the linearized version of the thickness equations (??) and (??)

$$\partial_t h'_1 + H_1 \nabla \cdot \mathbf{u}'_1 = 0 \quad (7.93\text{a})$$

$$\partial_t h'_2 + H_2 \nabla \cdot \mathbf{u}'_2 = 0, \quad (7.93\text{b})$$

where  $\mathbf{u}'_1$  and  $\mathbf{u}'_2$  are the horizontal velocities in the two layers. Likewise, assuming a zero atmospheric pressure leads to the linearized version of velocity equations (??) and (??)

$$\partial_t \mathbf{u}'_1 + f \hat{\mathbf{z}} \times \mathbf{u}'_1 = -g \nabla \eta'_{1/2} \quad (7.94\text{a})$$

$$\partial_t \mathbf{u}'_2 + f \hat{\mathbf{z}} \times \mathbf{u}'_2 = -g \nabla \eta'_{1/2} - g^r_{3/2} \nabla \eta'_{3/2}. \quad (7.94\text{b})$$

For the linearized equations, the only layer coupling occurs through undulations of the free surface impacting on the pressure felt in the lower layer.

## 7.6.2 Gravity wave equations

We restrict further analysis to the case with zero Coriolis acceleration so that the layer equations take the form

$$\partial_t (\eta'_{1/2} - \eta'_{3/2}) + H_1 \nabla \cdot \mathbf{u}'_1 = 0 \quad \text{and} \quad \partial_t \mathbf{u}'_1 = -g \nabla \eta'_{1/2} \quad (7.95\text{a})$$

$$\partial_t \eta'_{3/2} + H_2 \nabla \cdot \mathbf{u}'_2 = 0 \quad \text{and} \quad \partial_t \mathbf{u}'_2 = -g \nabla \eta'_{1/2} - g^r \nabla \eta'_{3/2}, \quad (7.95\text{b})$$

where we introduced the reduced gravity for the interior layer interface

$$g^r = g^r_{3/2} = g(\rho_2 - \rho_1)/\rho_1. \quad (7.96)$$

Taking time derivatives of the thickness equations and substituting the divergence of the velocity equations leads to the matrix-vector equation

$$\begin{bmatrix} (\partial_{tt} - g H_1 \nabla^2) & -\partial_{tt} \\ -g H_2 \nabla^2 & (\partial_{tt} - g^r H_2 \nabla^2) \end{bmatrix} \begin{bmatrix} \eta'_{1/2} \\ \eta'_{3/2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (7.97)$$

This equation has non-trivial solutions if the determinant of the matrix vanishes, and we use that property to derive the dispersion relation.

### 7.6.3 Dispersion relation

To derive the dispersion relation, consider the plane wave ansatz for the free surface undulation in each layer

$$\eta'_{1/2} = \tilde{\eta}_{1/2} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \quad \text{and} \quad \eta'_{3/2} = \tilde{\eta}_{3/2} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \quad (7.98)$$

Note that we assume each layer has the same wavevector and same angular frequency. However, the amplitudes,  $\tilde{\eta}_{1/2}$  and  $\tilde{\eta}_{3/2}$ , are generally complex so that the phase of the layers can differ. The corresponding velocity for the two layers is given by

$$\mathbf{u}'_1 = \tilde{\mathbf{u}}_1 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \quad \text{and} \quad \mathbf{u}'_2 = \tilde{\mathbf{u}}_2 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \quad (7.99)$$

Plugging the wave ansatz (7.98) into equation (7.97) converts the differential operators to algebraic expressions

$$\begin{bmatrix} (-\omega^2 + g H_1 |\mathbf{k}|^2) & \omega^2 \\ g H_2 |\mathbf{k}|^2 & (-\omega^2 + g^r H_2 |\mathbf{k}|^2) \end{bmatrix} \begin{bmatrix} \tilde{\eta}_{1/2} \\ \tilde{\eta}_{3/2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (7.100)$$

Setting the determinant to zero leads to the quadratic equation in  $\omega^2$

$$(\omega^2 - g H_1 |\mathbf{k}|^2)(\omega^2 - g^r H_2 |\mathbf{k}|^2) - g H_2 |\mathbf{k}|^2 \omega^2 = 0. \quad (7.101)$$

Solving for the squared phase speed ( $C_p^2 = \omega^2 / |\mathbf{k}|^2$ ) leads to

$$2\omega^2 / |\mathbf{k}|^2 = (g H + g^r H_2) \pm \sqrt{(g H + g^r H_2)^2 - 4 g g^r H_1 H_2}. \quad (7.102)$$

Introducing the small non-dimensional parameter,

$$\epsilon = g^r H_2 / (g H) = g^r H_2 / c_{\text{grav}}^2 \ll 1, \quad (7.103)$$

and expanding the dispersion relation (7.102) to expose the leading order behavior, yields the squared phase speeds

$$C_p^2 \approx c_{\text{grav}}^2 = g H \quad \text{squared barotropic phase speed} \quad (7.104a)$$

$$C_p^2 \approx c_{\text{grav}}^2 g^r H_1 H_2 / (g H^2) = g^r H_1 H_2 / H \quad \text{squared baroclinic phase speed.} \quad (7.104b)$$

We motivate the names *barotropic mode* and *baroclinic mode* in the following, based on our study of vorticity in Section ?? for a Boussinesq fluid. Note that the phase speed for the barotropic mode is roughly given by the gravity wave speed for a single layer of fluid with resting thickness  $H = H_1 + H_2$ . This speed is much greater than that for the baroclinic mode since  $g^r \ll g$ .

### 7.6.4 Structure of a plane gravity wave

In this subsection we examine the structure of a plane gravity wave by relating the wave amplitudes. To start, we relate interface height amplitudes by making use of equation (7.100) to write

$$\frac{\tilde{\eta}_{1/2}}{\tilde{\eta}_{3/2}} = \frac{\omega^2}{\omega^2 - g H_1 |\mathbf{k}|^2} = \frac{\omega^2 - g^r H_2 |\mathbf{k}|^2}{g H_2 |\mathbf{k}|^2}. \quad (7.105)$$

Since the right hand side is real, so too is the ratio of the amplitudes. Hence, the interfaces undulate either in phase or  $\pi$  radians out of phase. Correspondingly, we can, without loss of generality, take the interface height amplitudes to be real.

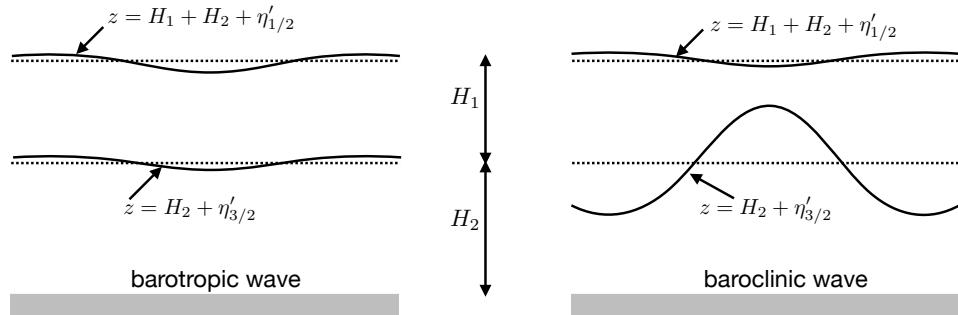


FIGURE 7.3: Illustrating the interface displacements for a two-layer shallow water gravity wave as discussed in Section 7.6.4. The left panel depicts a barotropic wave, where the undulations are both relatively small, in-phase, and on the same order of magnitude. The right panel shows a baroclinic wave, where the undulations of the free surface are small whereas those of the interior interface are much larger (order  $g/g'$ ) and the two undulations are  $\pi$  radians out of phase. For the baroclinic mode, recall that the same structure occurs for the reduced gravity model shown in Figure ??.

We relate the velocity amplitude to the interface amplitude by evaluating the linearized velocity equations (7.95a) and (7.95b) for plane waves, in which

$$\omega \tilde{\mathbf{u}}_1 = \mathbf{k} g \tilde{\eta}_{1/2} \quad \text{and} \quad \omega \tilde{\mathbf{u}}_2 = \mathbf{k} (g^r \tilde{\eta}_{1/2} + g \tilde{\eta}_{3/2}). \quad (7.106)$$

Since the interface height amplitudes are real, so too are the velocity amplitudes. Evidently, the fluid particle velocity in each layer is aligned with the phase velocity, so that the gravity waves are horizontally longitudinal. Correspondingly, the ratio of the velocity amplitudes is found by taking the scalar product of equation (7.106) with the wavevector to find

$$\omega \tilde{\mathbf{u}}_1 \cdot \mathbf{k} = |\mathbf{k}|^2 g \tilde{\eta}_{1/2} \quad \text{and} \quad \omega \tilde{\mathbf{u}}_2 \cdot \mathbf{k} = |\mathbf{k}|^2 (g^r + g H_2/H) \tilde{\eta}_{1/2}, \quad (7.107)$$

whose ratio is

$$\frac{\tilde{\mathbf{u}}_1 \cdot \mathbf{k}}{\tilde{\mathbf{u}}_2 \cdot \mathbf{k}} = \frac{g \tilde{\eta}_{1/2}}{g^r \tilde{\eta}_{1/2} + g \tilde{\eta}_{3/2}}. \quad (7.108)$$

### Structure of the barotropic mode

For the barotropic mode we set  $\omega^2 \approx g H |\mathbf{k}|^2$  within equation (7.105) to render the amplitude ratio

$$\tilde{\eta}_{1/2}/\tilde{\eta}_{3/2} \approx H/H_2 > 1. \quad (7.109)$$

This interface height ratio then leads, through equation (7.108), to the velocity amplitude

$$\frac{\tilde{\mathbf{u}}_1 \cdot \mathbf{k}}{\tilde{\mathbf{u}}_2 \cdot \mathbf{k}} \approx \tilde{\eta}_{1/2}/\tilde{\eta}_{3/2} \approx H/H_2. \quad (7.110)$$

Evidently, the two interfaces undulate in-phase and with the surface interface fluctuating more than the interior interface. Likewise, the horizontal velocities move in-phase with the interface heights, and with an amplitude ratio that concurs with the interface height ratio. We depict this wave motion in the left panel of Figure 7.3.

### Structure of the baroclinic wave mode

For the baroclinic mode, setting  $\omega^2 \approx g^r H_1 H_2 |\mathbf{k}|^2 / H$  within equation (7.100) leads to the amplitude ratio

$$\tilde{\eta}_{1/2} / \tilde{\eta}_{3/2} \approx -g^r H_2 / (g H). \quad (7.111)$$

Hence, the two interfaces undulate oppositely (i.e.,  $\pi$  radians out of phase) and with the amplitude of the upper interface far less than the interior interface

$$|\tilde{\eta}_{1/2} / \tilde{\eta}_{3/2}| \ll 1. \quad (7.112)$$

That is, the free surface is nearly rigid relative to the interior interface, as depicted in the right panel of Figure 7.3. For the velocity amplitude, we make use of equation (7.108) along with the interface height amplitude ratio (7.111), so that

$$\frac{\tilde{\mathbf{u}}_1 \cdot \mathbf{k}}{\tilde{\mathbf{u}}_2 \cdot \mathbf{k}} \approx -\frac{g^r H_2}{g H}, \quad (7.113)$$

so that the velocities are  $\pi$  radians out of phase and their ratio is identical to that for the interface heights.

Evidently, the baroclinic mode consists of layer velocities and interface heights that oscillate  $\pi$  radians out of phase, and with the velocity and surface height in the upper layer undulating much less than in the interior layer. Indeed, to leading order the upper interface is rigid and the corresponding flow stagnant. In contrast, the lower interface fluctuates as a gravity wave with the squared phase speed,  $g^r H_1 H_2 / H$ . This wave speed is far less than the  $\approx \sqrt{g H}$  speed of the barotropic mode.

#### 7.6.5 Energetic scaling for the waves

We here examine the relative amounts of phase averaged kinetic and available potential energies contained in the waves. From Section ??, we write the linearized layer integrated kinetic energy per area

$$\mathcal{K} = (\rho_0/2) (H_1 \mathbf{u}'_1 \cdot \mathbf{u}'_1 + H_2 \mathbf{u}'_2 \cdot \mathbf{u}'_2), \quad (7.114)$$

whose phase average is

$$\langle \mathcal{K} \rangle = (\rho_0/4) (H_1 |\tilde{\mathbf{u}}_1|^2 + H_2 |\tilde{\mathbf{u}}_2|^2). \quad (7.115)$$

Within the gravity wave, the velocity amplitudes are related to the interface amplitudes via equation (7.106), so that the phase averaged kinetic energy per horizontal area is

$$\langle \mathcal{K} \rangle = (\rho_0/4) (|\mathbf{k}|/\omega)^2 [H_1 (g \tilde{\eta}_{1/2})^2 + H_2 (g^r \tilde{\eta}_{1/2} + g \tilde{\eta}_{3/2})^2]. \quad (7.116)$$

For the available potential energy, we make use of the phase averaged version of equation (??) to write the phase averaged available potential energy per area as

$$\langle \mathcal{A} \rangle = (\rho_0/4) (g \tilde{\eta}_{1/2}^2 + g^r \tilde{\eta}_{3/2}^2). \quad (7.117)$$

#### Scaling within a barotropic gravity wave

For the barotropic gravity wave, the interface height undulations are related by equation (7.109). Hence, the phase averaged kinetic energy and available potential energy are given by

the approximate expressions

$$\langle \mathcal{K} \rangle \approx \frac{\rho_0 \tilde{\eta}_{1/2}^2}{4gH} [H_1 g^2 + H_2 (g^r + g H_2/H)^2] \quad (7.118a)$$

$$\langle \mathcal{A} \rangle \approx (\rho_0/4) \tilde{\eta}_{1/2}^2 (g + g^r H_2^2/H^2), \quad (7.118b)$$

and their ratio is

$$\frac{\langle \mathcal{K} \rangle}{\langle \mathcal{A} \rangle} = \frac{1}{gH} \frac{H_1 H^2 g^2 + H_2 (g^r H + g H_2)^2}{g H^2 + g^r H_2^2}. \quad (7.119)$$

This ratio is on the order of unity, with specific values determined by the layer thickness and reduced gravity. Hence, for the barotropic wave there is roughly the same amount of energy contained in the kinetic energy as in the available potential energy.

### Scaling within a baroclinic gravity wave

For the baroclinic gravity wave, the interface height undulations are related by equation (7.111). Hence, the phase averaged kinetic energy and available potential energy are given by the approximate expressions

$$\langle \mathcal{K} \rangle \approx \frac{\rho_0 \tilde{\eta}_{1/2}^2}{4 C_p^2} (H_1 g^2 + H_2 [g^r - (g^2 H)/(g^r H_2)]^2) \approx \frac{\rho_0 \tilde{\eta}_{1/2}^2}{4 C_p^2} \frac{g^2 H^2}{(g^r)^2 H_2} \quad (7.120)$$

$$\langle \mathcal{A} \rangle \approx (\rho_0/4) \tilde{\eta}_{1/2}^2 \frac{g^2 H^2}{g^r H_2^2}. \quad (7.121)$$

Use of equation (7.104b) for the squared phase speed in the baroclinic mode leads to the energy ratio

$$\frac{\langle \mathcal{K} \rangle}{\langle \mathcal{A} \rangle} \approx \frac{1}{C_p^2} \frac{g^2 H_2}{g^r} = \frac{g^2 H}{(g^r)^2 H_1} \gg 1. \quad (7.122)$$

This ratio is generally much larger than unity, indicating that the baroclinic gravity wave carries far more kinetic energy than available potential energy.

### Comparing the available potential energies between the waves

Making use of the above results leads to the ratio of the available potential energies contained in the barotropic gravity wave and baroclinic gravity wave

$$\frac{\langle \mathcal{A}^{bt} \rangle}{\langle \mathcal{A}^{bc} \rangle} \approx \frac{g^r H_2^2}{g H^2} \frac{(\tilde{\eta}_{1/2}^2)^{bt}}{(\tilde{\eta}_{1/2}^2)^{bc}} \approx \frac{g H^2}{g^r H_2^2} \frac{(\tilde{\eta}_{3/2}^2)^{bt}}{(\tilde{\eta}_{3/2}^2)^{bc}}. \quad (7.123)$$

The ratio thus depends on the assumed ratio of the undulations found in the two waves. In general we expect that

$$\frac{(\tilde{\eta}_{1/2}^2)^{bt}}{(\tilde{\eta}_{1/2}^2)^{bc}} \gg 1 \quad \text{and} \quad \frac{(\tilde{\eta}_{3/2}^2)^{bt}}{(\tilde{\eta}_{3/2}^2)^{bc}} \ll 1. \quad (7.124)$$

Even so, we cannot make any general statements about the ratio of available potential energies without further information. That is, we cannot *a priori* state that the barotropic wave requires more or less available potential energy than the baroclinic wave. Whereas the baroclinic wave involves large undulations of the interior interface, these undulations are coupled to the relatively small reduced gravity ( $g^r \ll g$ ), thus ameliorating the available potential energy cost.

In contrast, the barotropic wave involves a relatively small undulation of the interior interface and somewhat larger free surface undulation (larger than for the baroclinic mode). The free surface motion is coupled to the relatively large buoyancy through  $g \gg g^r$ , thus enhancing the potential energy cost for the free surface undulation, making the barotropic wave available potential energy comparable to that of the baroclinic wave.

### 7.6.6 The depth averaged velocity and the velocity difference

There are occasions in which it is useful to combine the layer velocity equations in a manner that directly approximates the barotropic and baroclinic motions. For this purpose we introduce the depth averaged velocity<sup>5</sup>

$$H \bar{\mathbf{u}} = H_1 \mathbf{u}'_1 + H_2 \mathbf{u}'_2, \quad (7.125)$$

with  $H = H_1 + H_2$ , along with the layer deviations from the depth average

$$\mathbf{u}_{1b} = \mathbf{u}'_1 - \bar{\mathbf{u}} \quad \text{and} \quad \mathbf{u}_{2b} = \mathbf{u}'_2 - \bar{\mathbf{u}}. \quad (7.126)$$

Making use of the linearized equations (7.94a) and (7.94b) (here returning to the case with rotation) leads to the equations of motion (see Exercise 7.3)

$$\partial_t \bar{\mathbf{u}} + f \hat{\mathbf{z}} \times \bar{\mathbf{u}} = -g \nabla \eta'_{1/2} - (H_2 g^r / H) \nabla \eta'_{3/2} \quad (7.127a)$$

$$\partial_t \mathbf{u}_{1b} + f \hat{\mathbf{z}} \times \mathbf{u}_{1b} = (g^r H_2 / H) \nabla \eta'_{3/2} \quad (7.127b)$$

$$\partial_t \mathbf{u}_{2b} + f \hat{\mathbf{z}} \times \mathbf{u}_{2b} = -(g^r H_1 / H) \nabla \eta'_{3/2}. \quad (7.127c)$$

Notice that the two deviation velocities,  $\mathbf{u}_{1b}$  and  $\mathbf{u}_{2b}$ , are independent of the free surface fluctuations. Furthermore, a layer integration of these two velocities vanishes

$$H_1 \mathbf{u}_{1b} + H_2 \mathbf{u}_{2b} = 0, \quad (7.128)$$

which is consistent with a vanishing integral for their equations of motion

$$H_1 (\partial_t \mathbf{u}_{1b} + f \hat{\mathbf{z}} \times \mathbf{u}_{1b}) + H_2 (\partial_t \mathbf{u}_{2b} + f \hat{\mathbf{z}} \times \mathbf{u}_{2b}) = 0. \quad (7.129)$$

Finally, introduce the vertical shear velocity

$$\mathbf{u}_s = \mathbf{u}_1 - \mathbf{u}_2 = \mathbf{u}_{1b} - \mathbf{u}_{2b} \quad (7.130)$$

whose equation of motion is

$$\partial_t \mathbf{u}_s + f \hat{\mathbf{z}} \times \mathbf{u}_s = g^r \nabla \eta'_{3/2}. \quad (7.131)$$

For the barotropic wave, whereby both layer interfaces undulate in phase and with a relatively small amplitude (Figure 7.3), then the shear velocity is nearly zero since the pressure gradient,  $g^r \nabla \eta'_{3/2}$  in equation (7.131) is small for the barotropic wave. In contrast, the depth averaged velocity described by equation (7.127a) is dominated by the larger (in magnitude) pressure gradient arising from free surface height undulations,  $-g \nabla \eta'_{1/2}$ . For the baroclinic wave, the depth averaged velocity is far smaller in magnitude than found in the barotropic wave, and the pressure gradient is dominated by interior interface undulations via  $g^r \nabla \eta'_{3/2}$ . These behaviors motivate the oceanographic colloquial terminology whereby the depth averaged

---

<sup>5</sup>We study vorticity of the depth averaged velocity for a continuous fluid in Section ??.

velocity is referred to as the *barotropic velocity* and the shear velocity is referred to as the *baroclinic velocity*.

### 7.6.7 Comments

As noted at the start of this section, the addition of further layers greatly increases the algebraic complexity of the analysis, thus motivating the use of numerical models for studies with  $N > 2$  layers. One generally finds that each layer adds another wave mode, with  $N$  layers realizing  $N$  modes (one barotropic mode and  $N - 1$  baroclinic modes). We further the study of gravity waves in Chapter 9 by studying internal gravity waves, with such waves corresponding here to a continuum of baroclinic modes.

## 7.7 Kelvin waves

The Kelvin wave is a non-dispersive gravity wave. It arises from the combined presence of a boundary and the Coriolis acceleration. The boundary considered here is a solid vertical wall. Additionally, Kelvin waves occur along the equator, with the equator acting as a boundary due to the change in sign of the Coriolis parameter,  $f$ .

### 7.7.1 Wave solutions with a southern boundary

To expose the key points about the shallow water Kelvin wave, it is sufficient to orient the  $f$ -plane with a boundary at  $y = y_0$  and to consider flow in the region  $y > y_0$ . The meridional velocity component must vanish at  $y = y_0$  to satisfy the no-normal flow condition. We are thus motivated to seek nontrivial solutions with  $v' = 0$  everywhere, in which case the linearized  $f$ -plane thickness and velocity equations are

$$\partial_t \eta' = -H \partial_x u' \quad (7.132a)$$

$$\partial_t u' = -g \partial_x \eta' \quad (7.132b)$$

$$f u' = -g \partial_y \eta'. \quad (7.132c)$$

It is notable that the meridional velocity equation (7.132c) expresses geostrophic balance between the Coriolis acceleration,  $f u'$ , and meridional pressure gradient,  $-g \partial_y \eta'$ .

Taking the time derivative of the zonal velocity equation (7.132b) and making use of the free surface equation (7.132a) leads to the one-dimensional wave equation for the zonal velocity

$$(\partial_{tt} - c_{\text{grav}}^2 \partial_{xx}) u' = 0, \quad (7.133)$$

where

$$c_{\text{grav}} = \sqrt{g H} \quad (7.134)$$

is the shallow water gravity wave speed. Likewise, a time derivative of the free surface equation (7.132a) and substitution of zonal velocity equation (7.132b) recovers the same wave equation satisfied by the free surface

$$(\partial_{tt} - c_{\text{grav}}^2 \partial_{xx}) \eta' = 0. \quad (7.135)$$

## 7.7.2 Kelvin wave solutions

In Section ?? we studied how to solve the wave equations (7.133) and (7.135), in which we write the general solutions in the form

$$u'(x, y, t) = F_1(x^L(t), y) + F_2(x^R(t), y), \quad (7.136a)$$

$$\eta'(x, y, t) = E_1(x^L(t), y) + E_2(x^R(t), y). \quad (7.136b)$$

In these expressions,  $F_1, E_1, F_2, E_2$  are functions of space that are specified by the initial conditions, and

$$x^L(t) = x + c_{\text{grav}} t \quad \text{and} \quad x^R(t) = x - c_{\text{grav}} t \quad (7.137)$$

are points along the  $x$ -axis that, as time increases, move to the left and right, respectively, at the gravity wave speed. Evidently, the wave signal transmits, without distortion, the  $F_1$  and  $E_1$  patterns in the negative  $\hat{x}$ -direction, and the  $F_2$  and  $E_2$  patterns propagate in the positive  $\hat{x}$ -direction.

The free surface height and zonal velocity are coupled via the equations of motion (7.132a)–(7.132c), so that the functions  $F_1, F_2$  are related to  $E_1, E_2$ . To determine this relation we use the velocity equation (7.132b) along with the chain rule

$$\frac{\partial F_1}{\partial t} = \frac{\partial F_1(x^L, y)}{\partial x^L} \frac{dx^L}{dt} = \frac{\partial F_1(x^L, y)}{\partial x^L} c_{\text{grav}} \quad (7.138a)$$

$$\frac{\partial F_2}{\partial t} = \frac{\partial F_2(x^R, y)}{\partial x^R} \frac{dx^R}{dt} = -\frac{\partial F_2(x^R, y)}{\partial x^R} c_{\text{grav}}, \quad (7.138b)$$

in which case

$$\partial_t u' = c_{\text{grav}} (\partial F_1 / \partial x^L - \partial F_2 / \partial x^R) = -g \partial_x \eta' = -g (\partial E_1 / \partial x^L - \partial E_2 / \partial x^R), \quad (7.139)$$

so that

$$\eta'(x, y, t) = (c_{\text{grav}}/g) [-F_1(x + c_{\text{grav}} t, y) + F_2(x - c_{\text{grav}} t, y)]. \quad (7.140)$$

To determine the meridional dependence to the wave solutions requires the geostrophic balance (7.132c), with the free surface equation (7.140) and zonal velocity equation (7.136a) leading to

$$\partial_y F_1 = f F_1 / c_{\text{grav}} \quad \text{and} \quad \partial_y F_2 = -f F_2 / c_{\text{grav}}, \quad (7.141)$$

whose solutions take the form

$$F_1 = F(x + c_{\text{grav}} t) e^{(y-y_0)/L_d} \quad \text{and} \quad F_2 = G(x - c_{\text{grav}} t) e^{-(y-y_0)/L_d}, \quad (7.142)$$

where

$$L_d = c_{\text{grav}} / f \quad (7.143)$$

is the shallow water deformation radius defined in equation (??) (and used in Section 7.4.6). We also introduced the functions  $F$  and  $G$ , which are functions of a single space coordinate. To ensure boundedness in the region  $y > y_0$ , where the fluid is assumed to exist, we drop the  $F_1$  solution, thus leaving the free surface

$$\eta' = (H/g)^{1/2} e^{-(y-y_0)/L_d} G(x - c_{\text{grav}} t), \quad (7.144)$$

and the horizontal velocity components

$$u' = e^{-(y-y_0)/L_d} G(x - c_{\text{grav}} t) \quad \text{and} \quad v' = 0. \quad (7.145)$$

For the vertical velocity component we return to equation (7.26) to write

$$w'(z) = w'(z = -H) - (z + H) \partial_x u' \quad (7.146a)$$

$$= w'(z = -H) + H^{-1} (z + H) \partial_t \eta' \quad (7.146b)$$

$$= w'(z = -H) - (z + H) e^{-(y-y_0)/L_d} \frac{dG(x^R)}{dx^R}, \quad (7.146c)$$

where the second equality used the linearized thickness equation (7.132a), and the third equality used the chain rule.

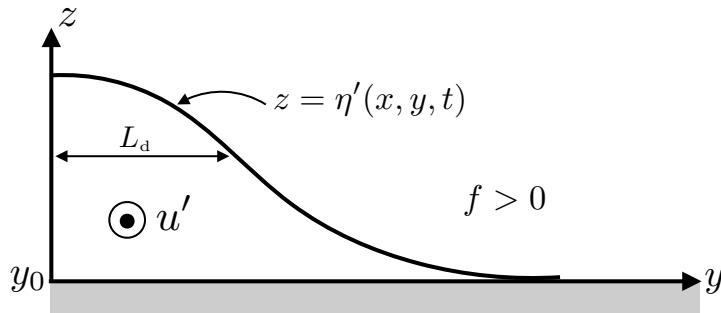


FIGURE 7.4: Illustrating free surface and zonal velocity for a northern hemisphere Kelvin wave with a solid boundary on the south. The free surface exponentially decays away from the boundary with a decay scale set by the deformation radius,  $L_d$ . With  $f > 0$  the Kelvin wave propagates with the boundary on the right, so that for this orientation the wave phase velocity is coming out of the page.

These non-dispersive gravity wave signals propagate in the positive  $\hat{x}$  direction, in which case the boundary  $y = y_0$  is on the right. This orientation holds for any boundary orientation in the northern hemisphere, whereby Kelvin waves propagate with the solid boundary on the right when looking in the direction of wave phase velocity. For the southern hemisphere, Kelvin waves propagate with the boundary to the left of the wave motion. Hence, Kelvin waves propagate cyclonically around a closed basin. In Figure 7.4 we illustrate the free surface and velocity for the Kelvin wave solution just derived.

## 7.8 Inertia-gravity waves

In this section we set the Coriolis parameter,  $f$ , to a nonzero constant while retaining a flat bottom and removing all side boundaries. We also retain a nonzero gravity, which is consistent with the shallow water being in hydrostatic balance. The linearized thickness equation (7.13), velocity equation (7.15), and potential vorticity equation (7.23) take the form

$$\partial_t \eta' = -H \nabla \cdot \mathbf{u}' \quad (7.147a)$$

$$\partial_t \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -g \nabla \eta' \quad (7.147b)$$

$$\partial_t Q' = 0, \quad (7.147c)$$

where the linear flat bottom  $f$ -plane shallow water potential vorticity is given from equation (7.22)

$$Q' = \zeta'/H - f \eta'/H^2, \quad (7.148)$$

where we dropped the constant  $f$  since it plays no dynamical role.

### 7.8.1 Forced oscillator equation for horizontal velocity

Taking a time derivative of the horizontal velocity equation (7.147b), and then back-substituting the velocity equation, leads to

$$(\partial_{tt} + f^2) \mathbf{u}' = -g (\partial_t \nabla \eta' - f \hat{\mathbf{z}} \times \nabla \eta'). \quad (7.149)$$

This equation for the horizontal velocity is in the form of a forced oscillator, with forcing from gradients in the free surface. We make use of this equation when developing the mathematical expressions for plane inertia-gravity waves in Section 7.8.9.

### 7.8.2 Free wave equation and potential vorticity

For the case of a flat bottom on an  $f$ -plane, we can write the wave equation (7.31) as

$$(\partial_{tt} - c_{\text{grav}}^2 \nabla^2) \eta' = -H f \zeta'. \quad (7.150)$$

This equation describes a forced shallow water gravity wave with forcing from relative vorticity coupled to the Coriolis parameter. Recall that this coupling between Coriolis and relative vorticity is fundamental to the inertial waves studied in Chapter 5. We can now use the vorticity equation as in Section 7.4.2 to eliminate  $\zeta'$  to reveal a free wave equation. Equivalently, in equation (7.38) we set  $f$  to a constant and the bottom to be flat, thus leading to

$$\partial_t [\partial_{tt} \eta' + f^2 \eta' - c_{\text{grav}}^2 \nabla^2 \eta'] = 0. \quad (7.151)$$

The linear fluctuations described by this equation are known as shallow water *inertia-gravity* or *Poincaré waves*. The name “inertia-gravity” is due to the presence of both the Coriolis frequency,  $f$ , and gravitational acceleration,  $g$ , with both playing a role as restoring forces to support the waves.

The wave equation (7.151) is in the form of a local conservation law where the term in square brackets is static. We already know about another static field, namely the potential vorticity,  $Q'$ , given by equation (7.148). We here show that the wave equation (7.151) is indeed identical to the potential vorticity equation (7.147c). For this purpose, substitute the expression (7.148) for the potential vorticity into the forced wave equation (7.150), which readily yields

$$f H^2 Q' = -(\partial_{tt} + f^2 - c_{\text{grav}}^2 \nabla^2) \eta'. \quad (7.152)$$

We thus find that

$$\partial_t Q' = 0 \implies \partial_t [\partial_{tt} \eta' + f^2 \eta' - c_{\text{grav}}^2 \nabla^2 \eta'] = 0, \quad (7.153)$$

so that the wave equation (7.151) for inertia-gravity waves is identical to the conservation equation for linear shallow water potential vorticity. We saw a similar connection between potential vorticity conservation and waves in our study of Rossby waves in Section 6.3 for the horizontally non-divergent barotropic fluid, and will see the connection yet again for shallow water Rossby waves in Section 7.9.

### 7.8.3 Dispersion relation

Substituting the wave ansatz (7.24) into equations (7.147a)-(7.147b) renders the homogeneous matrix-vector equation

$$\begin{bmatrix} -i\omega & -f & igk_x \\ f & -i\omega & igk_y \\ iHk_x & iHk_y & -i\omega \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{\eta} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (7.154)$$

This equation has a nontrivial solution only when the determinant of the matrix vanishes. The real part of the determinant cancels exactly, thus leaving just the imaginary part. Setting the imaginary part to zero yields the dispersion relation

$$\omega [\omega^2 - f^2 - gH|\mathbf{k}|^2] = 0. \quad (7.155)$$

We can derive the same dispersion relation by substituting the wave ansatz into the wave equation (7.153). There are three solutions to this cubic equation described in the following subsections.

### 7.8.4 Zero frequency geostrophic mode

The zero frequency solution to the dispersion relation (7.155) corresponds to  $f$ -plane geostrophic motion. Such motion is static so that the linearized continuity equation (7.147a) means that the flow is horizontally non-divergent,  $\nabla \cdot \mathbf{u}' = 0$ . Furthermore, the geostrophic solution corresponds to a static yet non-zero potential vorticity

$$f \hat{z} \times \mathbf{u}' = -g \nabla \eta' \implies (f/g) Q' = (\nabla^2 - L_d^{-2}) \eta' \neq 0, \quad (7.156)$$

where we introduced the deformation radius,  $L_d = c_{\text{grav}}/f$ , from equation (7.54). Turning equation (7.156) around, we see that if the potential vorticity is known, then the geostrophically balanced free surface can be found by inverting the elliptic operator,  $\nabla^2 - L_d^{-2}$ .

The static geostrophic mode with nonzero potential vorticity is decoupled, in the linear theory, from the ageostrophic inertia-gravity wave whose potential vorticity is identically zero and yet whose relative vorticity and free surface are time dependent. For this reason we can separately study the two linear modes without concern for interactions.

### 7.8.5 Inertia-gravity wave modes

The  $\omega \neq 0$  solution to the dispersion relation (7.155) satisfies the dispersion relation

$$\omega^2 = f^2 (1 + L_d^2 |\mathbf{k}|^2). \quad (7.157)$$

Recall from Section 1.2 that we consider the angular frequency of waves to be non-negative. Hence, we only take the positive root

$$\omega = |f| \sqrt{1 + L_d^2 |\mathbf{k}|^2}, \quad (7.158)$$

with Figure 7.5 depicting this relation.

The  $\omega > 0$  modes satisfying the dispersion relation (7.158) are inertia-gravity waves. These

waves have an angular frequency greater than or equal to the inertial frequency

$$\omega \geq |f|, \quad (7.159)$$

and are thus said to be *super-inertial waves*.<sup>6</sup> Furthermore, they carry zero potential vorticity (equation (7.152))

$$f H^2 Q' = -(\partial_{tt} + f^2 - c_{\text{grav}}^2 \nabla^2) \eta' = 0, \quad (7.160)$$

where we verify this property holds for plane inertia-gravity waves in Exercise 7.5. With  $Q' = 0$ , equation (7.148) for the potential vorticity shows that the inertia-gravity waves carry a relative vorticity given by

$$\zeta' = f \eta'/H. \quad (7.161)$$

Since  $|\eta'|/H \ll 1$ , we see that the relative vorticity carried by shallow water inertia-gravity waves is small relative to the planetary vorticity

$$|\zeta'| \ll |f|. \quad (7.162)$$

## 7.8.6 Group velocity

Taking the  $\mathbf{k}$ -space gradient of the dispersion relation (7.158) leads to the group velocity for the shallow water inertia-gravity waves

$$\mathbf{c}_g = \nabla_{\mathbf{k}} \varpi(\mathbf{k}) = \frac{c_{\text{grav}}^2 \mathbf{k}}{\omega} = \frac{\mathbf{c}_p c_{\text{grav}}^2}{|\mathbf{c}_p|^2} = \frac{\mathbf{c}_p}{1 + (L_d |\mathbf{k}|)^{-2}}, \quad (7.163)$$

where the phase velocity is  $\mathbf{c}_p = (\omega/|\mathbf{k}|) \hat{\mathbf{k}}$ . Evidently, the group velocity is parallel to the phase velocity and the ratio of their magnitudes is given by

$$\frac{|\mathbf{c}_g|}{|\mathbf{c}_p|} = \frac{c_{\text{grav}}^2}{|\mathbf{c}_p|^2} = \frac{1}{1 + (L_d |\mathbf{k}|)^{-2}}. \quad (7.164)$$

In the non-rotating case, where  $f = 0$  so that  $L_d^{-2} = (c_{\text{grav}}/f)^{-2} = 0$ , the group and phase velocities are equal, which we expect since the non-rotating shallow water gravity waves from Section 7.5 are non-dispersive. The inertia-gravity waves approach the non-dispersive limit for wavelengths much smaller than the deformation radius, in which the waves are too small to feel the effects from the Coriolis acceleration (we further discuss the shortwave limit in Section 7.8.7). But for the general case with dispersion, the Coriolis acceleration causes the group velocity to be smaller in magnitude than the phase velocity. Hence, the wave energy, which is carried by the group velocity (see Section 7.8.10 on energetics), is more slowly transmitted than the phase.

---

<sup>6</sup>The inertial waves considered in Chapter 5, which we studied in a homogeneous fluid, have their angular frequency bounded by  $\omega^2 < f^2$ . These are *sub-inertial* waves. They are again encountered in Section 9.8.1 as a special case of a rotating internal gravity wave in the limit where the reference fluid state is homogeneous. They are distinct from the shallow water inertial waves since the shallow water fluid is hydrostatic whereas the inertial waves in Chapter 5 rely on non-hydrostatic pressure.

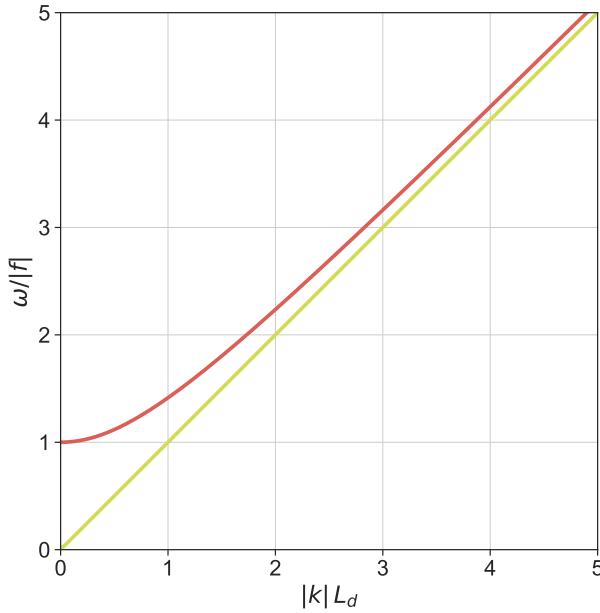


FIGURE 7.5: Illustrating the dispersion relation for shallow water inertia-gravity waves. The red hyperbolic curve is the dispersion relation (7.157), which asymptotes at high wavenumber (small wavelength) to the straight line for the non-rotating case. The angular frequency,  $\omega$ , is scaled by the magnitude of the Coriolis frequency,  $|f|$ , and the horizontal wavenumber,  $|\mathbf{k}|$ , is multiplied by the deformation radius  $L_d = \sqrt{gH}/f = c_{\text{grav}}/f$ . For small wave numbers ( $|\mathbf{k}| L_d \ll 1$  or  $\Lambda \gg 2\pi L_d$ ), the inertial-gravity wave frequency approaches the inertia frequency,  $f$ , with this behavior seen at the minimum of the dispersion curve. We expect this result since waves large relative to the deformation radius feel the Coriolis acceleration. At the opposite extreme of high wave numbers ( $\Lambda \ll 2\pi L_d$ ), the wave frequency approaches the non-rotating gravity wave frequency, shown here by the linear dispersion relation  $\omega = |\mathbf{k}| \sqrt{gH} = |\mathbf{k}| c_{\text{grav}}$ . Waves small relative to the deformation radius do not feel the Coriolis acceleration and thus converge to non-rotating gravity waves. Since all shallow water inertia-gravity waves satisfy  $\omega^2 \geq f^2$ , they are said to be *super-inertial* waves; i.e., waves whose frequency is larger in magnitude than the inertial frequency.

### 7.8.7 Shortwave limit for inertia-gravity waves

The shortwave limit is in the regime where

$$|\mathbf{k}|^2 L_d^2 \gg 1, \quad (7.165)$$

so that the shortwave limit occurs when the wavelength is much shorter than the deformation radius, in which case the waves do not feel the effects from the Coriolis acceleration. For example, consider a middle latitude shallow water gravity wave in a layer with  $H = 10^3$  m and  $f = 10^{-4}$  s $^{-1}$ , in which case the shallow water deformation radius is  $L_d = \sqrt{gH}/f = 10^3$  km. Maintaining the constraint (7.165) means that the wavelength of the gravity wave,  $\Lambda = 2\pi/|\mathbf{k}|$ , must satisfy

$$\Lambda \ll 2\pi L_d. \quad (7.166)$$

Within the shortwave limit, however, the wavelength cannot be too small since the flow must retain the hydrostatic balance as per a shallow water model. As discussed in Section ??, maintaining the hydrostatic balance means that the flow retains a small vertical to horizontal aspect ratio. In terms of the wavenumber for gravity waves, the hydrostatic balance implies

$$|\mathbf{k}| H \ll 1 \iff \text{hydrostatic balance}. \quad (7.167)$$

We conclude that the shortwave limit for shallow water gravity waves is given by the regime

$$L_d^{-1} \ll |\mathbf{k}| \ll H^{-1} \iff 2\pi H \ll \Lambda \ll 2\pi L_d. \quad (7.168)$$

Although considered short from the perspective of the shallow water model, these waves are hydrostatic and thus considered long gravity waves from the perspective of the surface gravity waves studied in Chapter 4.

Finally, for the shortwave limit, the dispersion relation (7.157) reduces to the dispersion relation (7.81)

$$\omega \approx |\mathbf{k}| c_{\text{grav}}. \quad (7.169)$$

Such waves are only weakly affected by the Coriolis acceleration so that their dispersion relation reduces to linear and non-dispersive gravity waves of Section 7.5.

### 7.8.8 Longwave limit for inertia-gravity waves

The longwave limit occurs when

$$|\mathbf{k}|^2 L_d^2 \ll 1, \quad (7.170)$$

so the waves are much longer than the deformation radius. In this limit the wave is strongly affected by the Coriolis acceleration. Indeed, the wave dispersion relation becomes

$$\omega^2 = f^2, \quad (7.171)$$

in which fluid particles exhibit inertial oscillations (Sections 7.8.9 and ??).

### 7.8.9 Polarization relations for a plane inertia-gravity wave

As for the gravity waves in Section 7.5, we study the behavior of the traveling plane wave ansatz (7.24). In contrast to the case with  $f = 0$ , here we require the complex nature of the wave amplitudes  $(\tilde{u}, \tilde{v}, \tilde{\eta})$  in order to realize nontrivial inertia-gravity wave solutions. Such complex amplitudes mean there are differences in phase between the velocity and free surface, with phase differences arising from the Coriolis parameter.

#### Polarization relations for the wave

Substituting the traveling plane wave ansatz (7.24) into the forced oscillator equation (7.149) leads to the amplitude relation

$$(-\omega^2 + f^2) \tilde{\mathbf{u}} = -g (\omega \mathbf{k} - i f \hat{\mathbf{z}} \times \mathbf{k}) \tilde{\eta}. \quad (7.172)$$

Without loss of generality, we assume the free surface amplitude,  $\tilde{\eta}$ , is real, which then leads to the free surface height

$$\eta' = \tilde{\eta} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (7.173)$$

and the fluid velocity within a plane wave

$$\mathbf{u}' = \frac{g |\mathbf{k}| \tilde{\eta}}{\omega^2 - f^2} \left[ \underbrace{\omega \hat{\mathbf{k}} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}_{\text{horizontally longitudinal}} + \underbrace{f (\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}_{\text{horizontally transverse}} \right] \quad (7.174a)$$

$$= \frac{\tilde{\eta}}{|\mathbf{k}| H} \left[ \underbrace{\omega \hat{\mathbf{k}} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}_{\text{horizontally longitudinal}} + \underbrace{f (\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}_{\text{horizontally transverse}} \right], \quad (7.174b)$$

where the unit vector

$$\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}| \quad (7.175)$$

points in the direction of the wave, and equation (7.174b) follows from use of the dispersion relation (7.157) in equation (7.174a). The vanishing potential vorticity in these waves means that they carry a nonzero relative vorticity as given by equation (7.161), so that

$$\zeta' = f \eta'/H = (f \tilde{\eta}/H) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (7.176)$$

Hence, in the northern hemisphere the relative vorticity is in-phase with the free surface, whereas in the southern hemisphere it is  $\pi$  radians out of phase.

These mathematical expressions for the fields within a wave are sometimes referred to as *polarization relations*. Observe that the horizontally transverse component of the velocity vector is  $\pi/2$  out of phase from the horizontally longitudinal component. Furthermore, the transverse component corresponds to fluid particle motion that is perpendicular to the wavevector, in which case we say that the transverse component is *polarized* perpendicular to the wavevector.

### Drawing the polarization relations

To draw the free surface (7.173), vorticity (7.176), and velocity (7.174b) we assume that  $f > 0$  for the northern hemisphere, and recall that  $\omega^2 \geq f^2$  since shallow water inertia-gravity waves are super-inertial. We also find it convenient to write the velocity as

$$\frac{\mathbf{u}' |\mathbf{k}| H}{\tilde{\eta} f} = (\omega/f) \hat{\mathbf{k}} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) + (\hat{\mathbf{z}} \times \hat{\mathbf{k}}) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (7.177)$$

Consider a point fixed in space and let time progress so that the phase,

$$\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t, \quad (7.178)$$

decreases. Consequently, the velocity vector rotates in a clockwise direction, forming an ellipse with the major axis along the longitudinal direction,  $\hat{\mathbf{k}}$ , and minor axis along the transverse direction,  $(\hat{\mathbf{z}} \times \hat{\mathbf{k}})$ . This motion corresponds to the inertial oscillations studied in Section ?? (where  $\omega^2 = f^2$ ). Now consider a fixed time and sample the velocity field in the direction of the wave. In this case the phase increases as we move in the wave direction, so that the sampled velocity progresses counter-clockwise around the ellipse. Figure 7.6 offers three depictions of the wave field.

### 7.8.10 Energetics

In Section 7.3.2 we considered the general form of energy balances for the linearized shallow water equations integrated over the shallow water layer. For the special case of a flat bottom domain (with  $\eta_b = 0$  for simplicity), these equations take the form

$$\partial_t (\mathcal{K}^{sw} + \mathcal{P}^{sw}) = -g \rho H \nabla \cdot [(\bar{\eta} + \eta') \mathbf{u}'] \quad (7.179a)$$

$$\mathcal{P}^{sw} = (\rho/2) g \eta'^2 \quad (7.179b)$$

$$\mathcal{K}^{sw} = (\rho/2) H \mathbf{u}' \cdot \mathbf{u}'. \quad (7.179c)$$

Here we consider the energy carried by a plane inertia-gravity wave, and focus on the phase averaged energetics. Note that since the waves are present throughout space (as per the plane

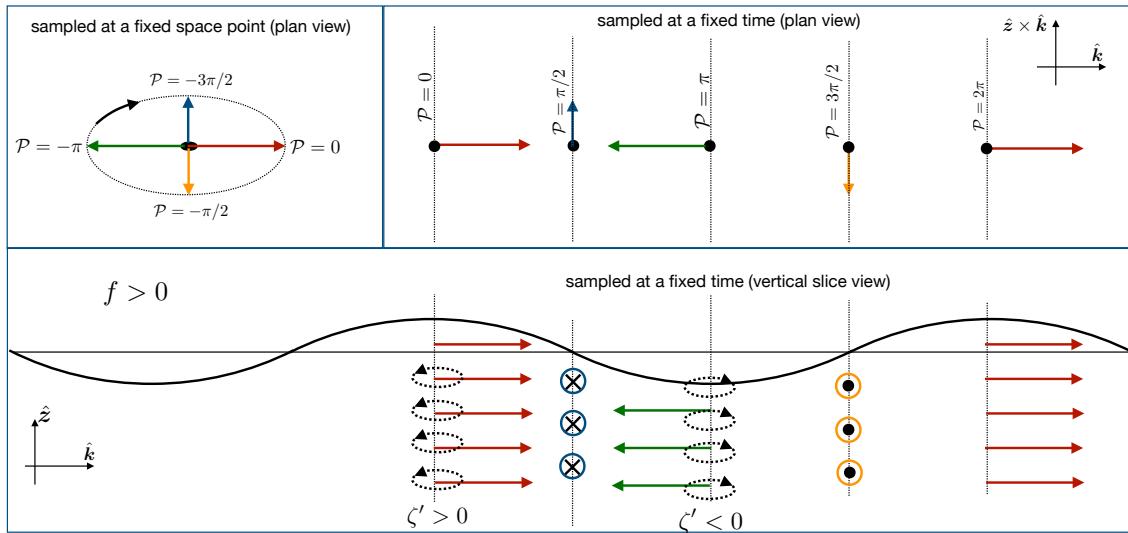


FIGURE 7.6: Illustrating the shallow water inertia gravity wave from three perspectives, with  $\mathcal{P} = \mathbf{k} \cdot \mathbf{x} - \omega t$  the phase and  $f > 0$ . The upper left panel shows the horizontal velocity vector at a fixed point in space as time increases and so the phase decreases (becomes more negative). The resulting velocity vector rotates clockwise and exhibits inertial oscillations. The upper right panel shows the horizontal velocity vector sampled along the wave direction,  $\hat{k}$ , at a fixed time from a plan (horizontal) view, in which case the phase increases moving in the  $\hat{k}$  direction. The lower panel shows the horizontal velocity, relative vorticity, and free surface as viewed from a vertical slice aligned with the wavevector direction. Note that the relative vorticity is the vertical component,  $\zeta'$ , so that the arrowed ellipses on the lower panel are in the horizontal plane.

wave assumption), we do not expect to have energy accumulate in any particular region when phase averaged. Instead, we expect the phase averaged energy to remain constant. So our question then concerns how that energy is partitioned according to kinetic energy and potential energy, and how the energy moves or is fluxed.

### Kinetic and potential energies contained in a plane wave

Writing the free surface as in equation (7.173) and the horizontal velocity as in equations (7.174a) and (7.174b) leads to the potential energy and kinetic energy carried by the wave field

$$\mathcal{P}^{\text{sw}} = \frac{\rho g \tilde{\eta}^2}{2} \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (7.180\text{a})$$

$$\mathcal{K}^{\text{sw}} = \frac{\rho \tilde{\eta}^2}{|\mathbf{k}|^2 H^2} [\omega^2 \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) + f^2 \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t)], \quad (7.180\text{b})$$

with a corresponding phase average given by

$$\langle \mathcal{P}^{\text{sw}} \rangle = \frac{\rho g \tilde{\eta}^2}{4} \quad (7.181\text{a})$$

$$\langle \mathcal{K}^{\text{sw}} \rangle = \frac{\rho \tilde{\eta}^2 (\omega^2 + f^2)}{2 |\mathbf{k}|^2 H^2} = \frac{\rho g \tilde{\eta}^2}{4} \frac{\omega^2 + f^2}{\omega^2 - f^2}. \quad (7.181\text{b})$$

Evidently, for non-rotating shallow water gravity waves,  $f = 0$ , there is an equipartition between phase averaged potential and kinetic energy. However, for the general case with rotation, the

phase averaged kinetic energy is larger than the potential energy

$$\frac{\langle \mathcal{K}^{\text{sw}} \rangle}{\langle \mathcal{P}^{\text{sw}} \rangle} = \frac{\omega^2 + f^2}{\omega^2 - f^2} = 1 + 2(L_d |\mathbf{k}|)^{-2}, \quad (7.182)$$

with equipartition approached only for wave lengths smaller than the deformation radius (high wave numbers).

### Mechanical energy and energy flux contained in a plane wave

The phase averaged mechanical energy contained in a plane wave is given by

$$\langle \mathcal{P}^{\text{sw}} \rangle + \langle \mathcal{K}^{\text{sw}} \rangle = \langle \mathcal{P}^{\text{sw}} \rangle \left[ 1 + \frac{\omega^2}{\omega^2 - f^2} \right] = 2 \langle \mathcal{P}^{\text{sw}} \rangle \frac{|\mathbf{c}_p|^2}{c_{\text{grav}}^2} = \frac{\rho g \tilde{\eta}^2}{2} \frac{|\mathbf{c}_p|}{|\mathbf{c}_g|}, \quad (7.183)$$

where we made use of equation (7.164) for the ratio of the group velocity magnitude to the phase velocity magnitude. For the flux of energy contained in the wave, we return to the energy equation (7.179a) and only consider the phase averaged flux, which takes the form

$$g \rho H \langle \eta' \mathbf{u}' \rangle = \frac{g \rho \tilde{\eta}^2 \omega \mathbf{k}}{2 |\mathbf{k}|^2} = \frac{g \rho \tilde{\eta}^2 \mathbf{c}_p}{2} = \frac{g \rho \tilde{\eta}^2 \mathbf{c}_g}{2} \frac{|\mathbf{c}_p|}{|\mathbf{c}_g|} = \mathbf{c}_g (\langle \mathcal{P}^{\text{sw}} \rangle + \langle \mathcal{K}^{\text{sw}} \rangle). \quad (7.184)$$

Hence, the phase averaged mechanical energy flux contained in the plane wave is given by the group velocity times the phase averaged mechanical energy. This is a standard result that we have seen before in the study of other linear waves.

## 7.9 Rossby waves

We now focus on the sub-inertial wave equation (7.62) derived in Section (7.4.6)

$$\partial_t [(L_d^{-2} - \nabla^2) \eta'] = H \hat{\mathbf{z}} \cdot (\nabla \eta' \times \nabla Q_r), \quad (7.185)$$

where

$$Q_r = f/H_r \quad (7.186)$$

is the potential vorticity in the resting fluid. Equation (7.185) describes shallow water *Rossby waves*.

### 7.9.1 Dispersion relation

Making use of the plane wave ansatz (7.24) in the wave equation (7.185) readily leads to the shallow water Rossby wave dispersion relation

$$\omega = \frac{H (\mathbf{k} \times \hat{\mathbf{z}}) \cdot \nabla Q_r}{k_d^2 + |\mathbf{k}|^2} \quad (7.187)$$

For the angular frequency to be independent of space,  $\nabla Q_r$  must be spatially independent. We can ensure this property by assuming a  $\beta$ -plane along with linear and gently varying topography

$$\nabla H_r^{-1} = -H_r^{-2} \nabla H_r = H_r^{-2} \nabla \eta'_b \approx H_r^{-2} \nabla \eta'_b, \quad (7.188)$$

where  $\nabla\eta'_b$  is spatially independent with linear topography. More general (but gently varying) bottom topography requires the asymptotic methods from Chapter 2.

The dispersion relation (7.187) compares directly to equation (6.32) for Rossby waves in the horizontally non-divergent barotropic model. The sole difference concerns the presence of the deformation radius,  $L_d$  from equation (7.54), and its associated wavenumber

$$k_d = L_d^{-1} = f/c_{\text{grav}}, \quad (7.189)$$

which is present in the shallow water dispersion relation. This term is missing from the non-divergent barotropic model since  $c_{\text{grav}}$  is formally infinite (there are no gravity waves in that model), in which case  $L_d^{-2} = k_d^2 = 0$ .

### 7.9.2 Connecting to quasi-geostrophic potential vorticity

When deriving the sub-inertial equations in Section 7.4.6, we noted that the assumptions made in that derivation are identical to the assumptions made when deriving shallow water quasi-geostrophy in Section ???. Indeed, the Rossby wave equation (7.185) is the linearized equation for the material evolution of shallow water quasi-geostrophic potential vorticity. We see this equality by recalling the discussion in Section (??), where we derived equation (??)

$$(f_o/g) Dq/Dt = \partial_t [(L_d^{-2} - \nabla^2) \eta'] - H \hat{z} \cdot (\nabla \eta' \times \nabla Q_r) - (g/f_o) \hat{z} \cdot [\nabla \eta' \times \nabla(\nabla^2 \eta')], \quad (7.190)$$

where  $q$  is the quasi-geostrophic potential vorticity. For an inviscid fluid,  $Dq/Dt = 0$ , and for small amplitude fluctuations the nonlinear term in equation (7.190) (final term on the right hand side) is neglected. In this case we see that the linearized equation for material evolution of quasi-geostrophic potential vorticity is identical to the Rossby wave equation (7.185).

The nonlinear term in equation (7.190) arises from the geostrophic advection of geostrophic relative vorticity. Although it is small for small amplitude fluctuations, and thus commonly dropped when deriving the dispersion relation, we note that it vanishes identically for a plane wave. It does so in precisely the same way as it vanishes for the non-divergent barotropic vorticity equation in Section 6.2.3. Namely, this result follows since for a plane wave,  $\nabla(\nabla^2 \eta') = -|\mathbf{k}|^2 \nabla \eta'$ , so that it follows immediately that  $\nabla \eta' \times \nabla(\nabla^2 \eta') = 0$ . Hence, a plane Rossby wave is an exact solution to the shallow water quasi-geostrophic potential vorticity equation in an inviscid fluid.

### 7.9.3 Vorticity mechanism

In Section 6.3 we studied Rossby waves in the horizontally non-divergent barotropic model, with a vorticity mechanism for the waves presented in Figure 6.3. This mechanism follows from the material evolution of absolute vorticity in the horizontally non-divergent barotropic model. The identical mechanism holds for the shallow water fluid yet with quasi-geostrophic potential vorticity replacing absolute vorticity. Hence, all conceptual points from Figure 6.3 also hold for the shallow water fluid.

### 7.9.4 Dispersion circle for planetary Rossby waves

In Section 6.4.2 we described a geometric method to help interpret the dispersion relation for planetary Rossby waves in the horizontally non-divergent barotropic model. That method also proved useful in Section 6.4.3 for describing the reflection of Rossby waves from a smooth and flat wall. Here we extend the geometric method to shallow water planetary Rossby waves, in

which the dispersion relation is given by the  $\beta$  portion of the general dispersion relation (7.187)

$$\varpi_\beta = -\frac{\beta k_x}{k_d^2 + |\mathbf{k}|^2}. \quad (7.191)$$

This equation compares to the  $\beta$  portion of the dispersion relation (6.32) holding for the non-divergent barotropic model. Again, the sole distinction is that the shallow water Rossby wave dispersion has a nonzero deformation wavenumber,  $k_d \neq 0$ . Furthermore, as shown in Exercise 7.6, the shallow water planetary geostrophic equations support Rossby waves with  $|\mathbf{k}|^2 \rightarrow 0$ , which are referred to as long Rossby waves.

### Dispersion circle

Following the geometric approach from Section 6.4, we write the dispersion relation (7.191) as the equation of a circle in wavevector space

$$(k_x + \beta/2\omega)^2 + k_y^2 = (\beta/2\omega)^2 - k_d^2. \quad (7.192)$$

As written, the angular frequency,  $\omega$ , is a parameter for the circle whose center is

$$\mathbf{k}_{\text{center}} = -(\beta/2\omega) \hat{\mathbf{x}} \quad (7.193)$$

and squared radius is  $(\beta/2\omega)^2 - k_d^2$ . The deformation wavenumber appearing on the right hand side of the circle equation (7.192) places an upper bound on the angular frequency allowed for propagating Rossby waves

$$\omega \leq \omega_{\max} = \beta/(2 k_d) = \beta L_d/2. \quad (7.194)$$

The shallow water model supports gravity waves that render  $k_d \neq 0$ , thus imposing an upper bound on the shallow water Rossby wave angular frequency. In contrast, for the horizontally non-divergent barotropic model,  $k_d = 0$  since there are no gravity waves, in which case there is no maximum frequency for Rossby waves. Note that for the  $\beta$ -plane scaling used to derive the Rossby waves,

$$\omega_{\max}/f_o = \beta L_d/(2 f_o) \ll 1, \quad (7.195)$$

thus confirming that the shallow water Rossby waves are strictly sub-inertial.

### Group velocity

The group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}} \varpi$ , is given by

$$\mathbf{c}_g = \frac{\beta [(k_x^2 - k_y^2 - k_d^2) \hat{\mathbf{x}} + 2 k_x k_y \hat{\mathbf{y}}]}{(|\mathbf{k}|^2 + k_d^2)^2} = -\frac{2 \omega \mathbf{R}}{|\mathbf{k}|^2 + k_d^2}, \quad (7.196)$$

where we introduced the group velocity orientation vector

$$\mathbf{R} = -\mathbf{k} - \beta/(2\omega) \hat{\mathbf{x}} = -[k_x + \beta/(2\omega)] \hat{\mathbf{x}} - k_y \hat{\mathbf{y}} \quad \text{with} \quad |\mathbf{R}|^2 = (\beta/2\omega)^2 - k_d^2. \quad (7.197)$$

As for the orientation vector (6.62) in the horizontally non-divergent barotropic model, the vector  $\mathbf{R}$  points from the perimeter of the dispersion circle to the center (its magnitude equals to the radius of the circle). We thus conclude that the dispersion geometry for shallow water planetary waves directly carries over from the horizontally non-divergent barotropic model

detailed in Section 6.4.2. We illustrate the dispersion geometry for shallow water Rossby waves in Figure 7.7, with many details provided in the figure and its caption.

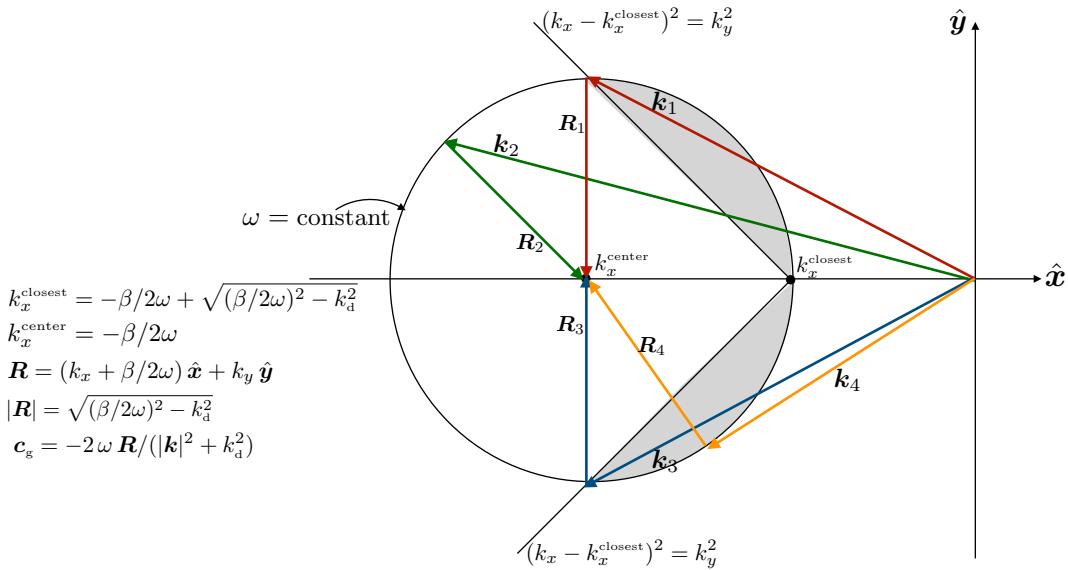


FIGURE 7.7: Dispersion circle for shallow water planetary Rossby waves as depicted by a circle in wavevector space,  $(k_x, k_y)$ , parameterized by the angular frequency,  $\omega$ . The center of the circle is at  $\mathbf{k}_{\text{center}} = -(\beta/2\omega)\hat{\mathbf{x}}$  and the squared radius is  $(\beta/2\omega)^2 - k_d^2$ . A positive radius requires the angular frequency to be less than the maximum,  $\omega_{\max} = \beta/(2k_d)$ . We depict four sample wavevectors,  $\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}$ , that orient the phase velocity,  $\mathbf{c}_p = \hat{\mathbf{k}} \omega / |\mathbf{k}|$ . Each wavevector extends from the origin to a point on the circle. Each wavevector has an associated group velocity orientation vector,  $\mathbf{R} = -\mathbf{k} - \beta/(2\omega)\hat{\mathbf{x}}$ , that points from the circle perimeter to the circle center. The group velocity is westward for those wavevectors that intersect the circle perimeter within the gray-shaded region, which generally includes Rossby waves with low zonal wavenumbers. The group velocity is eastward for wavevectors outside the gray region, with the lines  $(k_x - k_x^{\text{closest}})^2 = k_y^2$  separating these regions where the group velocity is westward or eastward. In particular, the group velocity for wavevector  $\mathbf{k}_1$  is southward; for  $\mathbf{k}_2$  it is southeastward; for  $\mathbf{k}_3$  it is northward, and for  $\mathbf{k}_4$  it is northwestward. This dispersion circle directly compares to Figure 6.5 constructed for the horizontally non-divergent barotropic model, with the key difference being  $k_d \neq 0$  for the shallow water so that the dispersion circle does not touch the origin at  $\mathbf{k} = 0$ .

### 7.9.5 Comments

A more streamlined approach to deriving the Rossby wave dispersion relation starts directly from the quasi-geostrophic potential vorticity equation. That approach makes use of the quasi-geostrophic theory derived in Section ???. Even so, the longer approach taken in the current section benefits by exposing the direct connection to other shallow water waves through the unified shallow water equation (7.42).

In Chapter 14 we study waves in a continuous quasi-geostrophic fluid, thus returning to some of the material in this section while extending it to continuous stratification.



## 7.10 Exercises

### EXERCISE 7.1: STEPS IN DERIVING THE LINEAR PV EQUATION

Fill in the mathematical details needed to derive the linearized potential vorticity equation

(7.23). Hint: form the linearized relative vorticity equation as an intermediate step.

**EXERCISE 7.2: STEPS IN DERIVING EQUATION (7.40)**

Derive equation (7.40) from the linearized velocity equation (7.15).

**EXERCISE 7.3: EQUATIONS FOR BAROTROPIC AND BAROCLINIC VELOCITIES**

Fill in the details for deriving the equations (7.127a)-(7.127c) for the depth averaged and depth-deviation velocities.

**EXERCISE 7.4: GRAVITY WAVES ON A CONSTANT REFERENCE FLOW**

Consider a one-dimensional shallow water layer moving with a constant zonal speed,  $U > 0$ , over a flat bottom and in a non-rotating reference frame. We here examine the linear gravity wave disturbances on this constant background flow, thus generalizing the case from Section 7.5 for gravity waves propagating on a stationary background.

- (a) Determine a general expression for the free surface wave fluctuation. Hint: perform a Galilean transformation (Section ??) to a reference frame moving with the constant background flow.
- (b) Determine a general expression for the zonal velocity wave fluctuation.
- (c) Making use of the ideas from hydraulic control in Section 7.5.4, discuss the cases where  $U < c_{\text{grav}}$  (subcritical flow),  $U > c_{\text{grav}}$  (supercritical flow), and  $U = c_{\text{grav}}$  (critical flow).

**EXERCISE 7.5: VANISHING POTENTIAL VORTICITY FOR INERTIA-GRAVITY WAVES**

Verify that the linearized potential vorticity,  $H Q' = \zeta' - f \eta'/H$ , vanishes for the plane shallow water inertia-gravity waves ( $f$ -plane) given by the polarization relations in Section 7.8.9.

**EXERCISE 7.6: WAVES IN PLANETARY GEOSTROPHY**

As part of our study of balanced models in VOLUME 3, we derive the shallow water planetary geostrophic equations

$$f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \frac{Dh}{Dt} = -h \nabla \cdot \mathbf{u}. \quad (7.198)$$

What linear waves are supported by these equations, assuming a flat bottom, zero mean flow, and  $\beta$ -plane?



## Chapter 8

# SHALLOW WATER WAVES: CASE STUDIES

In this chapter we examine some case studies in shallow waves, thus furthering our study of shallow water wave mechanics.

### READER'S GUIDE TO THIS CHAPTER

This chapter is a direct extension of the shallow water wave theory studied in Chapter 7.

<b>8.1</b>	<b>Loose threads</b>	<b>231</b>
<b>8.2</b>	<b>Waves excited by flow over topography</b>	<b>232</b>
8.2.1	Linearized governing equations	232
8.2.2	Galilean transformation to the frame of the reference flow	233
8.2.3	Forced wave equation and potential vorticity	233
8.2.4	Monochromatic topography	234
8.2.5	Stationary waves and causality	235
8.2.6	Free stationary inertia-gravity waves	237
8.2.7	Non-stationary gravity wave adjustment	238
8.2.8	Comments and further study	240
<b>8.3</b>	<b>Geostrophic adjustment</b>	<b>240</b>
8.3.1	Potential vorticity inversion	240
8.3.2	Posing the initial value problem	241
8.3.3	Adjustment with $f = 0$	242
8.3.4	Adjustment on the $f$ -plane	243
8.3.5	Concerning the deformation radius	244
8.3.6	Comments and further reading	246
<b>8.4</b>	<b>Exercises</b>	<b>246</b>

## 8.1 Loose threads

- WKBJ approximation for shallow water gravity waves so to compute the change in the amplitude of the wave. Need to reinterpret the acoustic amplitude equation 2.37c for shallow water.
- Salmon class notes Chapter 9 on shallow water waves induced by an earthquake: solving the initial value problem.
- String function discussion from [Tyler and Käse \(2000\)](#) and [Tyler and Käse \(2001\)](#).
- non-Doppler for long Rossby waves as on page 14 of Liu's notes.

## 8.2 Waves excited by flow over topography

Consider the  $f$ -plane flow of a single layer of shallow water fluid with a static and prescribed reference flow,  $\mathbf{u}_R$ , and let this flow pass over a non-flat bottom with vertical position

$$\eta_b(x, y) = \bar{\eta}_b + \eta'_b(x, y). \quad (8.1)$$

If the topographic amplitude is much smaller than the resting layer thickness,  $|\eta'_b| \ll H$ , then the motion consists of linear waves, and with the principle of superposition allowing Fourier analysis to construct the linear wave fields generated by arbitrary (small amplitude) topography. Additionally, the linearized form of shallow water potential vorticity conservation constrains the waves. Indeed, it provides the wave equation. This section works through many elements of the kinematics and dynamics encountered in a variety of forced wave problems, here with the special restrictions imposed by the vertically columnar motion of a shallow water layer.<sup>1</sup>

### 8.2.1 Linearized governing equations

Following the linearization process detailed in Section 7.3, here with a nonzero reference flow, leads to the thickness and velocity decomposition

$$h = H + \eta' - \eta'_b = H + h' \quad \text{and} \quad \mathbf{u} = \mathbf{u}_R + \mathbf{u}'. \quad (8.2)$$

The reference flow is assumed to be in geostrophic balance with a prescribed pressure gradient generated by a static free surface,

$$f_0 \hat{\mathbf{z}} \times \mathbf{u}_R = -g \nabla \eta_R. \quad (8.3)$$

Making use of equations (8.2) and (8.3) in the shallow water equations (7.1a) and (7.1b) leads to the linearized governing equations

$$[\partial_t + \gamma + (\mathbf{u}_R \cdot \nabla)] \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -g \nabla \eta' \quad (8.4a)$$

$$(\partial_t + \mathbf{u}_R \cdot \nabla) h' = -H \nabla \cdot \mathbf{u}'. \quad (8.4b)$$

The constant,  $\gamma \geq 0$ , is the inverse time scale for a Rayleigh drag (Section ??). We include Rayleigh drag since in many wave problems it is used to retain a finite solution in the presence of resonances. Even so, we drop it when examining the structure of the waves generated in the presence of flow over topography.

Taking the curl of the velocity equation (8.4a) and then making use of the thickness equation (8.4b) leads to the linearized version of the potential vorticity equation

$$(\partial_t + \mathbf{u}_R \cdot \nabla) Q' = -\gamma \zeta'/H, \quad (8.5)$$

where  $\zeta' = \partial_x v' - \partial_y u'$  is the relative vorticity of the perturbation, and

$$Q' = f/H + \zeta'/H - f h'/H^2 \quad (8.6)$$

is the linearized potential vorticity (equation (7.22)), with  $f/H$  a constant for the  $f$ -plane examined here. Equation (8.5) reveals that the linearized potential vorticity locally evolves

---

<sup>1</sup>The setup is analogous to that considered for inertial waves in Section 5.5.2, which considered inertial waves generated in a resting flow forced by moving topography.

according to dissipation from Rayleigh drag acting on the relative vorticity.

### 8.2.2 Galilean transformation to the frame of the reference flow

The reference flow is constant in space and time, so that a Galilean transformation to the frame moving with this flow acts to remove advection from the equations of motion (8.4a)-(8.4b). For this purpose, consider the Galilean transformation

$$\bar{t} = t \quad \text{and} \quad \bar{\mathbf{x}} = \mathbf{x} - \mathbf{u}_R t. \quad (8.7)$$

Following our discussion of Galilean transformations in Section ??, we know that the derivative operators transform according to

$$\partial_{\bar{t}} = \partial_t + \mathbf{u}_R \cdot \nabla \quad \text{and} \quad \partial_{\bar{x}} = \partial_x \quad \text{and} \quad \partial_{\bar{y}} = \partial_y, \quad (8.8)$$

which then brings the linear equations (8.4a), (8.4b), and (8.5) to

$$(\partial_{\bar{t}} + \gamma) \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -g \nabla \eta' \quad (8.9a)$$

$$\partial_{\bar{t}} h' = -H \nabla \cdot \mathbf{u}' \quad (8.9b)$$

$$\partial_{\bar{t}} Q' = -\gamma \zeta'/H. \quad (8.9c)$$

Whereas the reference flow moves with the velocity  $\mathbf{u}_R$  relative to the topography, the topography moves with a velocity  $-\mathbf{u}_R$  relative to the reference flow. Correspondingly, by moving to the boosted frame of the reference flow, the topography, which is independent of the rest frame time,  $t$ , is a function of time as measured in the boosted frame,  $\bar{t}$ .

### 8.2.3 Forced wave equation and potential vorticity

To derive a wave equation, take  $\partial_{\bar{t}}$  of the thickness equation (8.9b) to render

$$\partial_{\bar{t}\bar{t}} h' = -H \nabla \cdot \partial_{\bar{t}} \mathbf{u}'. \quad (8.10)$$

Replacing  $\partial_{\bar{t}} \mathbf{u}'$  using the linearized velocity equation (8.9a) leads to the forced wave equation

$$\partial_{\bar{t}} [(\partial_{\bar{t}} + \gamma) h'] - c_{\text{grav}}^2 \nabla^2 \eta' = -f H \zeta', \quad (8.11)$$

where we introduced the squared shallow water gravity wave speed

$$c_{\text{grav}}^2 = g H. \quad (8.12)$$

Following the approach in Section 7.8.2, make use of the linearized potential vorticity,  $Q'$  (equation (8.6)) to replace  $\zeta'$ , in which

$$f H^2 Q' = -(\partial_{\bar{t}\bar{t}} + \gamma \partial_{\bar{t}} + f^2) h' + c_{\text{grav}}^2 \nabla^2 \eta', \quad (8.13)$$

which agrees with equation (7.152) for the case with  $\mathbf{u}_R = 0$ ,  $\gamma = 0$ , and  $\eta'_b = 0$ . Equation (8.13) brings the potential vorticity evolution equation (8.9c) to the form

$$\partial_{\bar{t}} [(\partial_{\bar{t}\bar{t}} + \gamma \partial_{\bar{t}} + f^2) h' - c_{\text{grav}}^2 \nabla^2 \eta'] = f H \gamma \zeta', \quad (8.14)$$

which can be rewritten to isolate a forced wave equation for  $\eta'$

$$\partial_{\bar{t}}[(\partial_{\bar{t}\bar{t}} + f^2)\eta' - c_{\text{grav}}^2 \nabla^2 \eta'] = \gamma [f H \zeta' - \partial_{\bar{t}\bar{t}}(\eta' - \eta'_b)] + \partial_{\bar{t}}[(\partial_{\bar{t}\bar{t}} + f^2)\eta'_b]. \quad (8.15)$$

In the absence of dissipation ( $\gamma = 0$ ) and for a flat bottom ( $\eta'_b = 0$ ), this equation has solutions given by the free inertia-gravity wave modes discussed in Section 7.8.

For the remainder of this section we set  $\gamma = 0$  to focus on responses of a non-dissipative flow moving over small amplitude topography. In this case equation (8.15) simplifies to

$$\partial_{\bar{t}}[(\partial_{\bar{t}\bar{t}} + f^2)\eta' - c_{\text{grav}}^2 \nabla^2 \eta'] = \partial_{\bar{t}}[(\partial_{\bar{t}\bar{t}} + f^2)\eta'_b]. \quad (8.16)$$

Furthermore, consistent with our treatment of free inertia-gravity waves in Section 7.8, we focus on wave solutions with zero potential vorticity, which leads to the forced linear wave equation

$$(\partial_{\bar{t}\bar{t}} + f^2)\eta' - c_{\text{grav}}^2 \nabla^2 \eta' = (\partial_{\bar{t}\bar{t}} + f^2)\eta'_b. \quad (8.17)$$

Finally, since the bottom topography is independent of time,  $t$ , the source on the right hand side takes the form

$$(\partial_{\bar{t}\bar{t}} + f^2)\eta' - c_{\text{grav}}^2 \nabla^2 \eta' = [(\mathbf{u}_R \cdot \nabla)^2 + f^2]\eta'_b. \quad (8.18)$$

This is a forced linear wave equation for a shallow water layer moving in a uniform background flow over topography.

### 8.2.4 Monochromatic topography

We consider bottom topography in the form of a monochromatic wave

$$\eta'_b = \eta_o e^{ik_b \cdot x} = \eta_o e^{ik_b \cdot (\bar{x} + u_R \bar{t})} = \eta_o e^{i(k_b \cdot \bar{x} - \omega_R \bar{t})}, \quad (8.19)$$

where  $\eta_o$  is a constant real amplitude,  $\mathbf{k}_b$  is the topography wavevector, and

$$\omega_R = -\mathbf{k}_b \cdot \mathbf{u}_R > 0 \quad (8.20)$$

is the angular frequency implied by the reference flow that moves over the topography. Since the wave response is assumed to be linear, the response to more complex topography can be built using Fourier analysis.<sup>2</sup>

The topography (8.19) is stationary in the rest frame, but it moves in the direction opposite to the reference flow when viewed in the boosted frame. This direction swap is reflected in the choice  $\mathbf{k}_b \cdot \mathbf{u}_R < 0$ , which orients the topography wavevector,  $\mathbf{k}_b$ , according to the reference flow,  $\mathbf{u}_R$ . For example, a zonal referential flow,  $\mathbf{u}_R = U \hat{x}$ , with  $U > 0$ , has a topography wavevector  $\mathbf{k}_b = -|\mathbf{k}_b| \hat{x}$ , in which case the topography (8.19) takes the form

$$\eta'_b = \eta_o e^{-i|\mathbf{k}_b| x} = \eta_o e^{-i|\mathbf{k}_b| (\bar{x} + U \bar{t})}. \quad (8.21)$$

Evidently, a boosted frame observer rides along with the reference flow while the topography moves as a plane wave in the  $-\hat{x}$  direction. Conversely, a rest frame observer sees static topography with the reference flow moving in the  $+\hat{x}$  direction.

---

<sup>2</sup>We illustrate the Fourier analysis approach in Section 10.3 for stationary internal gravity waves generated by flow over a single mountain.

### 8.2.5 Stationary waves and causality

Before considering the non-stationary (or non-steady) wave response in Section 8.2.7, we determine the stationary response from a uniform flow moving over the monochromatic topography (8.21). Notably, stationary flow (also steady flow) refers to stationary in the rest frame of the topography, so that

$$\partial_t = 0 \implies \partial_{\bar{t}} = \mathbf{u}_R \cdot \nabla. \quad (8.22)$$

In effect, we assume all transient (traveling) waves have propagated far away from the region of interest, leaving just the stationary wave response to the forcing. For flow over topography, the forcing angular frequency is  $\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b > 0$ , whereas this frequency might have a different origin in other systems, such as the frequency of a piston in a laboratory experiment.

#### Causality and the non-negative forcing frequency

Stationary waves are generally simpler to mathematically determine than their transient cousins. Furthermore, stationary waves are often of primary physical interest when concerned with long term wave responses. However, there are subtleties related to *causality* that must be kept in mind when studying stationary solutions. Namely, the stationary equations have time symmetry merely because all fields are time independent (in the rest frame). However, stationary waves result from equilibration of transients that arise from a particular forcing. For example, with flow over topography the stationary wave solution knows about the direction of the reference flow, thus respecting causality (e.g., waves are swept downstream not upstream). So although there are no time derivatives in the stationary equations, we build in causality through specifying the forcing frequency,  $\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b > 0$  (equation (8.20)). Our choice to insist on a positive forcing angular frequency provides a robust means to maintain causality. It also accords with our choice in this book to only consider non-negative angular frequencies (see Section 1.2.3).

#### Free surface and velocity for the stationary flow

We insert the topography Fourier mode (8.19) into the potential vorticity equation (8.16), and take a plane wave ansatz for the free surface

$$\eta' = \tilde{\eta} e^{i(\mathbf{k} \cdot \bar{x} - \bar{\omega} \bar{t})}, \quad (8.23)$$

where  $\bar{\omega}$  is the angular frequency seen in the boosted reference frame. This substitution yields

$$\bar{\omega}(-\bar{\omega}^2 + f^2 + c_{\text{grav}}^2 |\mathbf{k}|^2) \tilde{\eta} e^{i(\mathbf{k} \cdot \bar{x} - \bar{\omega} \bar{t})} = \omega_R (-\omega_R^2 + f^2) \eta_o e^{i(\mathbf{k}_b \cdot \bar{x} - \omega_R \bar{t})}. \quad (8.24)$$

Since the fluctuations are assumed to be linear, we expect the reference flow over the topography to excite linear waves. Furthermore, in the steady state we expect the wavevector and angular frequency to be set by the flow and topography

$$\bar{\omega}^2 = \omega_R^2 = (\mathbf{k}_b \cdot \mathbf{u}_R)^2 \quad \text{and} \quad |\mathbf{k}|^2 = |\mathbf{k}_b|^2, \quad (8.25)$$

in which case equation (8.24) yields the amplitude ratio

$$\frac{\tilde{\eta}}{\eta_o} = \frac{\omega_R^2 - f^2}{\omega_R^2 - f^2 - c_{\text{grav}}^2 |\mathbf{k}_b|^2}, \quad (8.26)$$

and the corresponding free surface wave form

$$\eta' = \tilde{\eta} e^{ik_b \cdot x} = \frac{(\omega_R^2 - f^2) \eta_o e^{ik_b \cdot x}}{\omega_R^2 - f^2 - c_{\text{grav}}^2 |k_b|^2} = \left[ \frac{\omega_R^2 - f^2}{\omega_R^2 - f^2 - c_{\text{grav}}^2 |k_b|^2} \right] \eta'_b, \quad (8.27)$$

with an example given by Figure 8.1. Evidently, the stationary solution has a free surface and bottom topography that are either in-phase (same sign) or  $\pi$  radians out of phase (opposite sign), depending on properties of the reference flow, the topography, the gravity wave speed, and the Coriolis parameter. The stationary velocity field can be found just like for the free inertia-gravity waves in Section 7.8.9 through use of the linear velocity equation (8.9a) (with the Rayleigh drag,  $\gamma = 0$ ), in which

$$\mathbf{u}' = \tilde{\mathbf{u}} e^{ik_b \cdot x} \quad (8.28)$$

with the complex amplitude

$$\tilde{\mathbf{u}} = \frac{g \tilde{\eta} (-k_b \omega_R - i f \hat{z} \times k_b)}{\omega_R^2 - f^2} = \frac{g \eta_o (-k_b \omega_R - i f \hat{z} \times k_b)}{\omega_R^2 - f^2 - c_{\text{grav}}^2 |k_b|^2}. \quad (8.29)$$

These waves are not the free inertia-gravity waves studied in Section 7.8 since  $\omega_R^2 \neq f^2 (1 + L_d^2 |k_b|^2)$ . Rather, they are stationary waves forced by the motion of the shallow water fluid over the sinusoidal topography. Even so, they share many properties with inertia-gravity waves, including the polarization relation exhibited by equation (8.29), which directly compares to the free inertia-gravity wave in equation (7.174b).

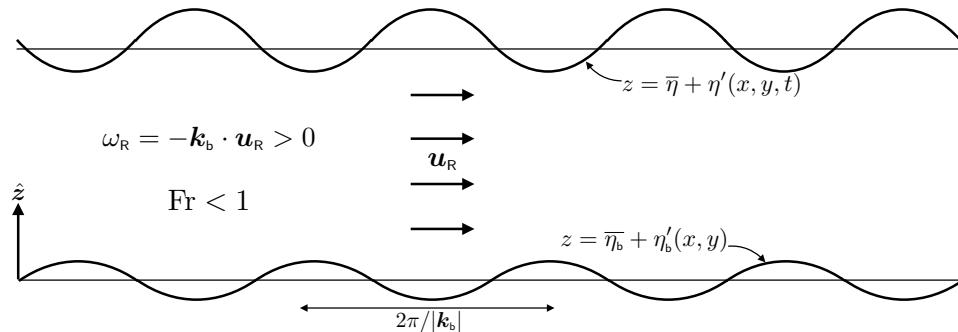


FIGURE 8.1: A shallow water fluid layer with reference flow that moves over a sinusoidal bottom topography for the sub-critical case ( $\text{Fr} < 1$ ) from equation (8.33) with  $f = 0$ . The stationary gravity waves have amplitude,  $\tilde{\eta} = -\eta_o \text{Fr}^2 / (1 - \text{Fr}^2)$ , that is opposite to that of the topography. Note that the specification,  $\mathbf{k}_b \cdot \mathbf{u}_R < 0$ , orients the topography wavevector,  $\mathbf{k}_b$ , according to the reference flow,  $\mathbf{u}_R$ . This orientation ensures that the topography observed in the boosted frame moves in the opposite direction to the reference flow as viewed in the rest frame. It also builds in causality as discussed in Section 8.2.5.

### Flow regimes and free surface deflections

To help understand the amplitude ratio (8.26), consider the special case of zonal reference flow with  $\mathbf{u}_R = U \hat{x}$ , so that the squared frequency is

$$\omega_R^2 = U^2 |k_b|^2. \quad (8.30)$$

We also find it useful to introduce the Froude number, shallow water deformation radius, and period of an inertial oscillation

$$\text{Fr} = U/c_{\text{grav}} \quad \text{and} \quad L_d = c_{\text{grav}}/f \quad \text{and} \quad T_{\text{inertial}} = 2\pi/f. \quad (8.31)$$

We can thus write the amplitude ratio (8.26) in the equivalent forms

$$\frac{\tilde{\eta}}{\eta_o} = \frac{\omega_R^2 - f^2}{\omega_R^2 - f^2 - c_{\text{grav}}^2 |\mathbf{k}_b|^2} = \frac{U^2 - f^2 |\mathbf{k}_b|^{-2}}{U^2 - f^2 |\mathbf{k}_b|^{-2} - c_{\text{grav}}^2} = \frac{\text{Fr}^2 - (L_d |\mathbf{k}_b|)^{-2}}{\text{Fr}^2 - (L_d |\mathbf{k}_b|)^{-2} - 1}, \quad (8.32)$$

with the first a ratio of angular frequencies, the second a ratio of speeds, and the third a ratio of non-dimensional numbers.

There are three regimes for the amplitude as determined by

$$\frac{\tilde{\eta}}{\eta_o} = \begin{cases} > 0 & \text{if } \omega_R^2 > f^2 + c_{\text{grav}}^2 |\mathbf{k}_b|^2 \\ < 0 & \text{if } f^2 < \omega_R^2 < f^2 + c_{\text{grav}}^2 |\mathbf{k}_b|^2 \\ > 0 & \text{if } \omega_R^2 < f^2, \end{cases} \quad (8.33)$$

as well as their equivalent forms in terms of speeds and non-dimensional numbers. Actually, it is simplest to start by considering the  $f = 0$  case, in which there are just two regimes as determined by

$$\left[ \frac{\tilde{\eta}}{\eta_o} \right]_{\text{non-rotating}} = \begin{cases} > 0 & \text{if } \text{Fr} > 1 \\ < 0 & \text{if } \text{Fr} < 1. \end{cases} \quad (8.34)$$

With  $\text{Fr} > 1$ , the free surface is in-phase with the topography, so that the free surface rises when the bottom rises, and vice versa. As discussed in Section 7.5.4, the  $\text{Fr} > 1$  flow is a state of hydraulic control and is generally unstable (the fluid particle speed is greater than the gravity wave speed, and the flow generally breaks down into turbulence). For the sub-critical flow with  $\text{Fr} < 1$ , the free surface falls when the topography rises, and vice versa. Equation (8.33) shows that the addition of a non-zero Coriolis parameter, and thus a finite deformation radius and finite inertial oscillation period, introduces a third regime for in-phase free surface and topography, while it modifies the sub-critical and super-critical regimes.

### 8.2.6 Free stationary inertia-gravity waves

Consider the special case of

$$\omega_R^2 = (\mathbf{k}_b \cdot \mathbf{u}_R)^2 = f^2, \quad (8.35)$$

in which case the topographic forcing vanishes on the right hand side of the potential vorticity equation (8.24). This case allows for any free inertia-gravity wave to fit within the domain so that

$$\omega^2 = f^2 (1 + L_d^2 |\mathbf{k}|^2). \quad (8.36)$$

That is, the free inertia-gravity wave can exist just as in the case of the flat bottom domain examined in Section 7.8. Transforming the free surface back to the rest frame yields

$$\eta' = \tilde{\eta} e^{i(\mathbf{k} \cdot \bar{\mathbf{x}} - \bar{\omega} \bar{t})} = \tilde{\eta} e^{i[\mathbf{k} \cdot \mathbf{x} - (\mathbf{k} \cdot \mathbf{u}_R + \bar{\omega}) t]} = \tilde{\eta} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad (8.37)$$

where we related the boosted frame frequency,  $\bar{\omega}$ , according to the *Doppler shift* relative to the rest frame frequency,  $\omega$ ,

$$\bar{\omega} = \omega - \mathbf{u}_R \cdot \mathbf{k}. \quad (8.38)$$

Furthermore note that the free surface amplitude is unconstrained since these waves are invisible to the topography. We have thus identified a free wave solution, which arises for the special case of topography and reference flow satisfying  $\omega_R^2 = f^2 > 0$  so that  $\omega^2/f^2 = 1 + (|\mathbf{k}| L_d)^2$ , which is the flat bottom frequency for free traveling inertia-gravity waves found in Section 7.8.

### 8.2.7 Non-stationary gravity wave adjustment

So far we have focused on the stationary waves that result after allowing for transient non-stationary waves to propagate away from the area of interest (formally, to propagate out to infinity). We here consider the case of non-stationary gravity waves ( $f = 0$ ) generated by flow over topography. We assume the free surface fluctuation,  $\eta'$ , is initially equal to the bottom topography and to have a zero initial tendency. With a zonal reference flow ( $\mathbf{u}_R = U \hat{\mathbf{x}}$ ) and zonally dependent bottom topography,  $\eta'_b = \eta'_b(x)$ , we are led to the initial value problem from equation (8.18)

$$(\partial_{\bar{t}\bar{t}} - c_{\text{grav}}^2 \partial_{\bar{x}\bar{x}}) \eta' = U^2 \partial_{\bar{x}\bar{x}} \eta'_b \quad (8.39a)$$

$$\eta'(\bar{x}, \bar{t} = 0) = \eta'_b(\bar{x}) \quad (8.39b)$$

$$\partial_t \eta'(\bar{x}, \bar{t} = 0) = 0. \quad (8.39c)$$

This is an initial value problem formulated using the boosted reference frame coordinates,  $(\bar{x}, \bar{t}) = (x - Ut, t)$  from Section 8.2.2. Additionally, as shallow water gravity waves are non-dispersive, we can naively allow the bottom topography to be arbitrarily shaped.<sup>3</sup>

We do not expect the initial condition to remain fixed for all time. Rather, we expect the flow (in the limit of linear behavior) to adjust through gravity waves that propagate in both directions. Additionally, we expect this propagation to occur in the presence of the stationary solution from Section 8.2.5. This expectation motivates the ansatz

$$\eta'(\bar{x}, \bar{t}) = \eta'_{\text{stationary}}(\bar{x}) + \eta'_{\text{transient}}(\bar{x}, \bar{t}) = \frac{\text{Fr}^2}{\text{Fr}^2 - 1} \eta'_b(x) + \eta'_{\text{transient}}(\bar{x}, \bar{t}), \quad (8.40)$$

where we wrote the stationary solution in terms of the rest frame coordinates,  $(x, t)$ , which is the frame where  $\eta'_{\text{stationary}}$  is stationary. The transient solution embodies linear gravity waves propagating in both directions, and it satisfies equation (8.39a) with zero forcing and with initial conditions set according to equations (8.39b) and (8.39c)

$$(\partial_{\bar{t}\bar{t}} - c_{\text{grav}}^2 \partial_{\bar{x}\bar{x}}) \eta'_{\text{transient}} = 0 \quad (8.41a)$$

$$\eta'_{\text{transient}}(\bar{x}, \bar{t} = 0) = -\eta'_{\text{transient}}(\bar{x}, \bar{t} = 0) \quad (8.41b)$$

$$\partial_t \eta'_{\text{transient}}(\bar{x}, \bar{t} = 0) = 0. \quad (8.41c)$$

The solution to the non-dispersive wave equation (8.41a) is given by the D'Alembert formula from Section ??, which takes the form

$$\eta'_{\text{transient}} = A \eta'_b(\bar{x} - c_{\text{grav}} \bar{t}) + B \eta'_b(\bar{x} + c_{\text{grav}} \bar{t}) \quad (8.42a)$$

$$= A \eta'_b[x - t(U + c_{\text{grav}})] + B \eta'_b[x - t(U - c_{\text{grav}})] \quad (8.42b)$$

$$= A \eta'_b(x - c_{(+)} t) + B \eta'_b(x - c_{(-)} t), \quad (8.42c)$$

---

<sup>3</sup>We consider dispersive internal gravity waves in Section 10.3, which requires Fourier analysis methods to account for wave dispersion.

where

$$c_{(\pm)} = U \pm c_{\text{grav}} \quad (8.43)$$

are the gravity wave speeds relative to the reference flow. The constants  $A$  and  $B$  are set according to the initial conditions (8.41b) and (8.41c), which yields

$$\eta'(x, t) = \frac{\text{Fr}^2}{\text{Fr}^2 - 1} \eta'_b(x) + \frac{1}{2} \left[ \frac{\eta'_b(x - c_{(-)} t)}{1 - \text{Fr}} + \frac{\eta'_b(x - c_{(+)} t)}{1 + \text{Fr}} \right]. \quad (8.44)$$

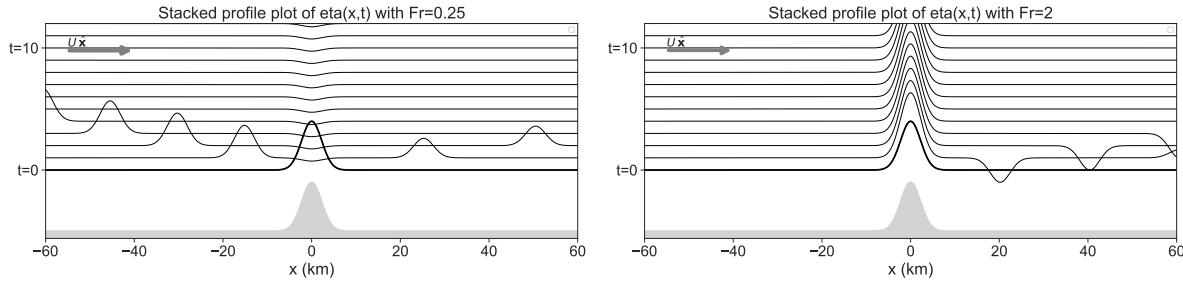


FIGURE 8.2: Stacked profiles of the free surface height that depicts the transient linear gravity wave adjustment of a shallow water layer flowing over a mountain, with the reference flow moving from left to right,  $\mathbf{u}_R = U \hat{x}$  with  $U > 0$ . Each line represents the free surface incremented in time (moving upward) by the amount  $\Delta t = 20 \times 10^3 \text{ m} / \sqrt{1000 \text{ m } 9.8 \text{ m s}^{-2}} \approx 200 \text{ s}$ . The left panel shows results for subcritical flow with  $\text{Fr} = 1/4$ , and the right panel shows supercritical flow with  $\text{Fr} = 2$ , both according to the solution (8.44). The initial condition for the free surface,  $\eta'(\bar{x}, \bar{t} = 0) = \eta'_b(\bar{x})$ , is shown as a thick black line. For the subcritical flow, the stationary solution has a relatively small and negative amplitude, whereas for the supercritical flow the stationary solution has a large a positive amplitude. The subcritical case reveals leftward and rightward moving gravity waves that propagate away from the mountain, with the leftward wave having a relatively large amplitude and slow speed. The supercritical case has both waves moving to the right, with the relatively small amplitude positive signal moving much faster than the larger amplitude negative signal.

For subcritical flow, with  $U < c_{\text{grav}}$ , we have

$$\text{subcritical} \implies c_{(+)} = U + c_{\text{grav}} > 0 \quad \text{and} \quad c_{(-)} = U - c_{\text{grav}} < 0. \quad (8.45)$$

Since  $|c_{(-)}| < |c_{(+)}|$ , we find the transient solution consists of a relatively fast rightward moving signal,  $\eta'_b(x - c_{(+)} t)/[2(1 + \text{Fr})]$ , plus a relatively slow leftward moving signal,  $\eta'_b(x - c_{(-)} t)/[2(1 - \text{Fr})]$ . Since  $\text{Fr} < 1$ , the amplitude of the right moving signal is smaller than the left moving signal according to the ratio  $0 < (1 - \text{Fr})/(1 + \text{Fr}) < 1$ . The stationary solution has a negative amplitude of relatively absolute value. We depict an example in the left panel of Figure 8.2 for the case of a single Gaussian mountain with  $\text{Fr} = 1/4$ . We can make use of a Gaussian shaped mountain rather than a single Fourier mode since the linear responses are non-dispersive gravity waves that travel at the same speed. Hence, we can sum any number of non-dispersive gravity wave modes to render the Gaussian shaped wave response that remains coherent.

For supercritical flow, with  $U > c_{\text{grav}}$ , have

$$\text{supercritical} \implies c_{(+)} = U + c_{\text{grav}} > 0 \quad \text{and} \quad c_{(-)} = U - c_{\text{grav}} > 0. \quad (8.46)$$

Since  $c_{(\pm)} > 0$ , both signals move to the right. The slower signal,  $\eta'_b(x - c_{(-)} t)/[2(1 - \text{Fr})]$ , has a negative amplitude since  $\text{Fr} > 1$ , with this amplitude larger in absolute sense than the faster signal,  $\eta'_b(x - c_{(+)} t)/[2(1 + \text{Fr})]$ , whose amplitude is positive. Furthermore, the stationary

solution has a positive amplitude. We depict an example supercritical response in the right panel of Figure 8.2 for the case of a single Gaussian mountain with  $\text{Fr} = 2$ .

It is notable that the potential energy of the stationary state for subcritical flow is less than that of the initial condition, given the depressed free surface height, whereas the potential energy of the supercritical flow's stationary state is greater than the initial condition. We cannot perform a closed energy budget since the reference flow is assumed to be fixed. Even so, we understand the ability of the supercritical reference flow to lift the full column of shallow water fluid up and over the mountain, given its relatively large source of kinetic energy. In contrast, the subcritical reference flow insufficient kinetic energy to lift the free surface over the mountain.

### 8.2.8 Comments and further study

There is no vertical wave propagation in a shallow water layer (see Section 7.3.4). Hence, the layer responds to movement of the reference flow over variable bottom topography by conforming to the constraints from shallow water potential vorticity conservation. It is for this reason that we focused on the linearized potential vorticity equation (8.9c). A further focus on fluctuations with zero potential vorticity,  $Q' = 0$ , leads to the wave equation (8.18), just as for the free inertia-gravity waves in Section 7.8. In Section 10.2 we study the generation of internal inertia-gravity waves by flow over topography. That study extends the work done here for the shallow water fluid, and it reveals a richer phenomenology that is supported by continuous stratification and non-hydrostatic pressure. Chapter 5 of *Sutherland* (2010) works through a variety of examples for shallow water fluid layers moving over topography, including a study of the initial value problem that we considered in Section 8.2.7.

## 8.3 Geostrophic adjustment

The geostrophic balance presented in Sections ?? and ?? is very well maintained by the observed large-scale atmosphere and ocean. Hence, geostrophy (and the associated thermal wind) is a powerful diagnostic. In this section, we examine how a flow state that is initially not in geostrophic balance evolves towards geostrophy. We thus study the dynamical processes associated with the *geostrophic adjustment* problem. As we see, the adjustment occurs through the propagation of linear inertia-gravity waves.

A single shallow water layer on a flat  $f$ -plane is sufficient to introduce the main physical ideas of geostrophic adjustment. Furthermore, we focus on linear perturbations so that the governing equations are those derived in Section 7.8 when studying inertia-gravity waves. The adjustment consists of linear inertia-gravity waves that maintain a locally static potential vorticity (equation (7.23)). For brevity in notation, we here drop all primes on the linear fluctuating terms.

### 8.3.1 Potential vorticity inversion

Before studying the geostrophic adjustment problem, we offer a few comments about *potential vorticity inversion*, which generally refers to the process of determining the flow field given information about the potential vorticity. In a shallow water layer, the potential vorticity is given by

$$Q = h^{-1} (f + \zeta) = h^{-1} (f + \partial_x v - \partial_y u). \quad (8.47)$$

If we further assume the flow to be in geostrophic balance (Section ??), then

$$Q = \frac{f}{h} + \frac{1}{h} \left[ \frac{\partial}{\partial x} \left( \frac{g}{f} \frac{\partial \eta}{\partial x} \right) + \frac{\partial}{\partial y} \left( \frac{g}{f} \frac{\partial \eta}{\partial y} \right) \right]. \quad (8.48)$$

Assuming we know  $Q$  throughout the domain; assuming  $f$  and  $Q$  are uniformly of the same sign within the domain; and assuming we know boundary conditions for  $\eta$ , then equation (8.48) is a nonlinear elliptic boundary value problem (Section ??) for  $\eta$ . Nonlinearities come from the  $h^{-1} = (\eta - \eta_b)^{-1}$  pre-factor, as well as the boundary conditions discussed below. Linearizing by setting  $h^{-1} \approx H^{-1}$  and simplifying the boundary conditions (see below) allows equation (8.48) to be solved for  $\eta$ . This solution process is referred to as *inverting* the elliptic operator, so that this particular inversion process is referred to as *potential vorticity inversion*.

General boundary conditions for  $\eta$  can be rather complex to handle mathematically. Namely, in a domain with a sloping bottom, such as in Figure ??, the free surface deviation equals to the bottom deviation,  $\eta = \eta_b$ , along the domain boundaries since the layer thickness vanishes there. Furthermore, the horizontal position of the domain boundary is a function of time since the layer moves up and down the sloping bottom. Vanishing layers and the associated moving boundaries are intrinsically nonlinear; i.e., there is no way to linearize the process without removing it altogether. Instead, to facilitate the use of linear physics requires us to assume the layer thickness remains nonzero throughout the domain. Furthermore, we assume the layer thickness deviates only a small amount from the layer averaged thickness:  $h/H \approx 1$ . These assumptions are made in the following discussion of geostrophic adjustment.

### 8.3.2 Posing the initial value problem

We solve for the  $t > 0$  evolution of surface height and velocity by making use of the linearized equations from Section 7.3<sup>4</sup>

$$\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = -g \nabla \eta \quad (8.49a)$$

$$\partial_t \eta + H \nabla \cdot \mathbf{u} = 0 \quad (8.49b)$$

$$\zeta - f \eta / H = H Q(x, y). \quad (8.49c)$$

$Q(x, y)$  is the linearized potential vorticity that is static for the flat bottom  $f$ -plane and so it is fully determined by the initial conditions (Section 7.3.3). To illustrate the geostrophic adjustment in an analytically tractable manner, consider the following step initial conditions for the surface height

$$\eta(x, t=0) = \begin{cases} +\eta_o & x < 0 \\ -\eta_o & x > 0, \end{cases} \quad (8.50)$$

which can be written

$$\eta(x, t=0) = \eta_o [1 - 2 \mathcal{H}(x)] = -\eta_o \operatorname{sgn}(x), \quad (8.51)$$

where the sign-function is given by

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0, \end{cases} \quad (8.52)$$

---

<sup>4</sup>Recall that to reduce notational clutter, we drop primes on the fluctuation variables in this section.

which can also be written in terms of the Heaviside step function

$$\operatorname{sgn}(x) = 2 \mathcal{H}(x) - 1, \quad (8.53)$$

where<sup>5</sup>

$$\mathcal{H}(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1/2 & \text{if } x = 0 \\ 1 & \text{if } x > 0. \end{cases} \quad (8.54)$$

The velocity is assumed to be zero initially

$$\mathbf{u}(x, y, t = 0) = 0. \quad (8.55)$$

Correspondingly, the initial relative vorticity vanishes so that the linearized potential vorticity (equation (8.49c)) is

$$Q(x, y) = \frac{f \eta_0}{H^2} \operatorname{sgn}(x). \quad (8.56)$$

Since  $\partial Q / \partial t = 0$  (equation (7.23)), the potential vorticity (8.56) is maintained at each point in space throughout the adjustment process. The velocity and surface height adjustment is thus constrained to keep potential vorticity static. This rather basic point is key to determining evolution of the velocity and surface height, and thus in determining the final (equilibrium) state for these fields.

### 8.3.3 Adjustment with $f = 0$

In the absence of planetary rotation ( $f = 0$ ), relative vorticity is constant at each grid point. With a zero initial velocity, relative vorticity remains zero throughout the adjustment. The adjustment is thus quite simple. Namely, it consists of linear gravity waves, which carry zero relative vorticity (equation (7.72)). These gravity waves propagate away from the initial step, converting the potential energy of the step into kinetic energy of waves that propagate to infinity. As the linear gravity waves are non-dispersive, they carry the initial pulse out to infinity without distortion in the wave form

$$\eta(x, t) = -\frac{\eta_0}{2} [\operatorname{sgn}(x + c_{\text{grav}} t) + \operatorname{sgn}(x - c_{\text{grav}} t)]. \quad (8.57)$$

The meridional velocity remains zero, whereas the zonal velocity equation

$$\partial_t u = -g \partial_x \eta, \quad (8.58)$$

leads to

$$u(x, t) = \frac{g \eta_0}{2 c_{\text{grav}}} [\operatorname{sgn}(x + c_{\text{grav}} t) - \operatorname{sgn}(x - c_{\text{grav}} t)]. \quad (8.59)$$

After the transient waves have passed, the steady solution is a flat surface height with zero velocity. This steady solution is familiar from the case of a rock dropped into a still pond. After dropping the rock into the pond, the surface gravity waves radiate outward from the rock and are eventually damped upon reaching the shore. In equilibrium, the pond returns to a state of rest with a flat surface height.<sup>6</sup>

---

<sup>5</sup>We discuss the Heaviside step function in Section ??.

<sup>6</sup>For most ponds, waves are better studied using deep water equations rather than shallow water equations; see Section 4.3. Even so, the key physical points in this example are maintained.

### 8.3.4 Adjustment on the $f$ -plane

On a rotating  $f$ -plane, the transient solution consists of the inertia-gravity waves studied in Section 7.8, with these waves transmitting information about the initial surface height perturbation out to infinity. After the transient waves have passed, we might guess that the steady solution is either the trivial solution with flat surface height (as for the  $f = 0$  case), or a nontrivial solution that is in geostrophic balance

$$f\hat{z} \times \mathbf{u} = -g \nabla \eta \quad \text{and} \quad \nabla \cdot \mathbf{u} = 0 \quad \text{and} \quad Q = H^{-2} f \eta_o \operatorname{sgn}(x). \quad (8.60)$$

Conservation of potential vorticity chooses the geostrophic solution, so that a steady solution of no-motion is not allowed by potential vorticity conservation. This is a profound distinction from the adjustment with  $f = 0$ .

#### Computing the steady state

As the steady flow is geostrophic on an  $f$ -plane, we make use of the geostrophic streamfunction

$$\psi = g \eta / f. \quad (8.61)$$

The steady state is written in terms of the streamfunction according to

$$u = -\partial_y \psi \quad \text{and} \quad v = \partial_x \psi \quad \text{and} \quad \zeta = \nabla^2 \psi. \quad (8.62)$$

Making use of these expressions for the linearized potential vorticity (8.49c) leads to the elliptic partial differential equation for the streamfunction

$$[\nabla^2 - L_d^{-2}] \psi = H Q(x, y), \quad (8.63)$$

where we introduced the shallow water deformation radius,  $L_d = c_{\text{grav}}/f$ , from equation (7.54).

The initial condition (8.51) has no  $y$ -dependence. Furthermore, there is nothing in the adjustment process that breaks meridional symmetry. Hence, the steady state is a function only of  $x$ , in which case the streamfunction satisfies the ordinary differential equation

$$\frac{d^2 \psi}{dx^2} - L_d^{-2} \psi = \frac{f \eta_o}{H} \operatorname{sgn}(x). \quad (8.64)$$

We solve this equation separately for  $x > 0$  and  $x < 0$  and match the function and its first derivative at  $x = 0$ , and furthermore constrain the streamfunction to vanish at  $\pm\infty$ . The  $x > 0$  streamfunction satisfies

$$\frac{d^2 \psi}{dx^2} - L_d^{-2} \psi = \frac{f \eta_o}{H}. \quad (8.65)$$

The particular solution is

$$\psi_p = -L_d^2 H Q = -L_d^2 f \eta_o / H = -g \eta_o / f, \quad (8.66)$$

and the homogeneous solution is

$$\psi_h = (g \eta_o / f) e^{-x/L_d}, \quad (8.67)$$

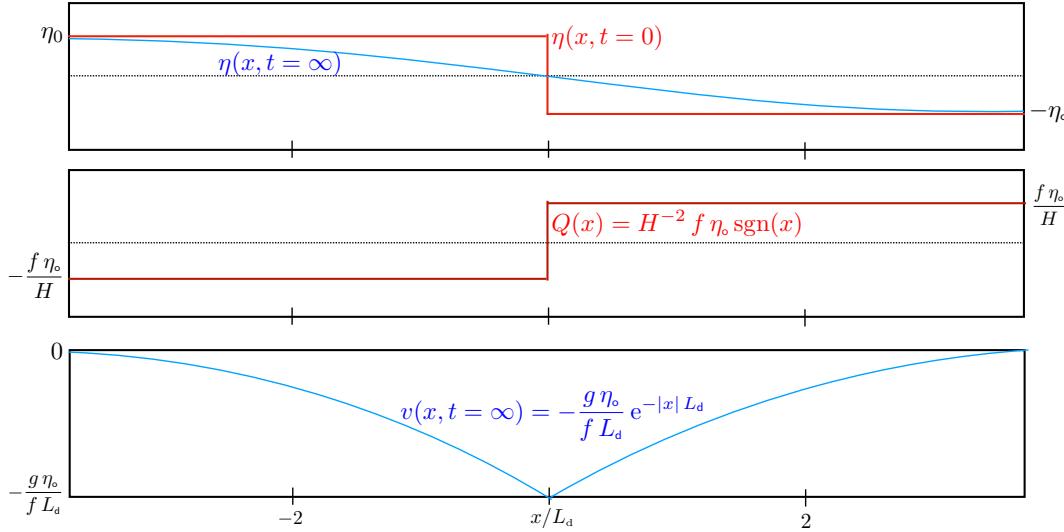


FIGURE 8.3: Depicting solutions to the linear geostrophic adjustment of a rotating shallow water layer on an  $f$ -plane. The top panel shows the initial (step-function) surface height (8.51) and the steady state (exponential) surface height (8.70). The second panel shows the static (step-function) potential vorticity (8.60). The third panel shows the steady state (exponential) meridional velocity (8.71) comprised of a jet centered at  $x = 0$ . The horizontal axis is scaled according to the shallow water deformation radius,  $L_d = f^{-1} \sqrt{gH}$ . This figure is adapted from figure 3.10 of [Vallis \(2017\)](#).

so that

$$\psi = -\frac{g\eta_0}{f} \left[ 1 - e^{-x/L_d} \right]. \quad (8.68)$$

The  $x < 0$  solution is found similarly, so that the full solution is

$$\psi = \frac{g\eta_0}{f} \begin{cases} -(1 - e^{-x/L_d}) & x > 0 \\ (1 - e^{x/L_d}) & x < 0, \end{cases} \quad (8.69)$$

which means that the steady state surface height is

$$\eta = \eta_0 \begin{cases} -(1 - e^{-x/L_d}) & x > 0 \\ (1 - e^{x/L_d}) & x < 0. \end{cases} \quad (8.70)$$

Note that the streamfunction vanishes at  $x = 0$  and has a first derivative of  $-\eta_0 \sqrt{gH}/H$ . Since the streamfunction only has a zonal dependence, the steady state velocity is purely meridional

$$u = 0 \quad \text{and} \quad v = -\frac{g\eta_0}{fL_d} e^{-|x|/L_d}. \quad (8.71)$$

The steady state velocity thus consists of a jet that is perpendicular to the surface height front.

### 8.3.5 Concerning the deformation radius

As illustrated in Figure 8.3, the steady state profiles for the surface height and velocity both have an exponential decay, with decay length scale given by the deformation radius. The deformation radius is this length scale over which a signal can propagate before being affected or “deformed” by the Coriolis acceleration, thus motivating the name *deformation radius*. More precisely, the deformation radius measures the horizontal length scale over which a wave can propagate within the time  $|f|^{-1}$  (the *inertial period* is  $2\pi/|f|$ ), before feeling the effects of the

Coriolis acceleration, thus making

$$L_d |f| = \sqrt{g H}. \quad (8.72)$$

In the  $f = 0$  limit, the deformation radius is infinity and the steady solution returns to the case considered in Section 8.3.3, whereby the steady state free surface is flat and there is no flow. A key feature of the  $f \neq 0$  case is that some of the potential energy contained within the initial perturbed free surface remains part of the steady state geostrophic flow. Conservation of potential vorticity constrains the flow so that all the initial potential energy cannot be converted to kinetic energy. Rather, the adjustment occurs only within a deformation radius distance from the initial perturbation.

We can extend the ideas introduced in this single-layer adjustment to a two-layer system as depicted in Figure 8.4, with the figure caption summarizing the physics. Again, adjustment leads to geostrophic flow when the lateral extent of the flow reaches the deformation scale. In this case, the deformation scale is much smaller than the single-layer fluid given that it is the reduced gravity that determines the velocity scale rather than the gravity.<sup>7</sup> The slower internal wave speeds propagate the internal wave signal a shorter distance before feeling the effects from Coriolis

$$L_d |f| = \sqrt{g^r H} \ll \sqrt{g H}. \quad (8.73)$$

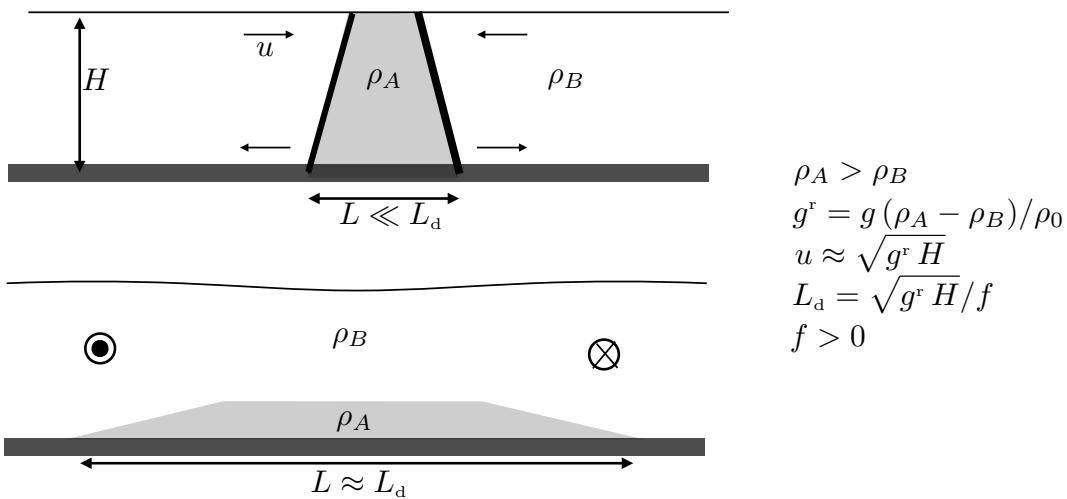


FIGURE 8.4: Two regions of uniform density fluid used to illustrate the adjustment under gravity acceleration and Coriolis acceleration in the absence of mixing. Top panel: two fluid layers each have uniform density with  $\rho_A > \rho_B$  and with the heavy fluid sitting between two regions of the lighter fluid, with the fluids separated by a barrier. Lower panel: upon releasing the barrier, gravity causes the heavy fluid to slump under the lighter fluid. This process halts when the lateral scale of the heavier fluid reaches the deformation scale,  $L_d = \sqrt{g^r H}/f$ , where  $\sqrt{g^r H} \approx N H$  is the speed of the internal gravity wave signal appearing in a two-layer fluid. The adjusted state reaches two-layer geostrophic balance as discussed in Section ??, whereby the difference in geostrophic flow in the two layers is proportional to the slope of the interface between the layers. By extension, if the initial region of heavy fluid had an initial lateral extent on the order of the deformation scale, then there will be minimal change in the lateral extent after the barrier is removed since the Coriolis acceleration will balance the pressure gradients to render a geostrophic flow. See Figure 15.4 of [Cushman-Roisin and Beckers \(2011\)](#) for more examples of geostrophic adjustment.

<sup>7</sup>Recall from Section 7.6 that we study gravity waves in two shallow water layers, where we indeed find the reduced gravity determines the baroclinic phase speed.

### 8.3.6 Comments and further reading

Our study of shallow water geostrophic adjustment shares some features with that of the gravity wave adjustment to flow over topography as studied in Section 8.2.7. For the topographic case we considered  $f = 0$ , whose linear gravity waves carry zero relative vorticity, whereas for the geostrophic adjustment we allow for  $f \neq 0$ , whose linear inertia-gravity waves carry zero potential vorticity. Transients for both cases consist of linear waves carrying information out to “infinity”, leaving behind a stationary flow. For the topographic case the stationary flow is set according to the topography and the Froude number, whereas for the geostrophic case the stationary flow is set according to the Coriolis acceleration and initial conditions.

Section 3.9 of [Vallis \(2017\)](#) presents a far more thorough discussion of this linear geostrophic adjustment problem, including an elegant variational approach. Chapter 3 of [Pratt and Whitehead \(2008\)](#) provide a thorough discussion for both linear and nonlinear geostrophic adjustment.



## 8.4 Exercises

### EXERCISE 8.1: DEFORMATION RADIUS

The deformation radius appears in many contexts within rotating fluid dynamics. Here, we compute this length scale for selective geophysical flow regimes at  $30^\circ N$  latitude, where  $f = 7.3 \times 10^{-5} \text{ s}^{-1}$ .

- Compute the shallow water deformation radius for an ocean continental shelf of depth 500 m.
- Compute the shallow water deformation radius for the deep ocean with depth 5000 m.
- The deformation radius defined in this chapter is sometimes called the *external deformation radius* as it makes use of the full depth of the fluid and the gravitational acceleration. In contrast, the deformation radius defined in terms of internal layer thickness and reduced gravity,  $g^r$ , leads to the internal deformation radius. The internal deformation radius,  $L_d^{\text{int}} = \sqrt{g^r h}/f$  is the appropriate rotational length scale for density layers in the interior of the ocean or isentropic layers in the interior of the atmosphere. Compute the deformation radius for a density layer of thickness  $h = 200$  m and reduced gravity of  $g^r = g/1000$ .

### EXERCISE 8.2: GEOSTROPHIC ADJUSTMENT (BASED ON EXERCISE 4.6 OF [Vallis \(2019\)](#))

Consider the linear geostrophic adjustment problem on an  $f$ -plane with a single layer of shallow water fluid over a flat bottom. Rather than assume an initial free surface profile, as we did in Section 8.3, here we assume an initial meridional velocity profile given by

$$v(x, t=0) = v_0 \operatorname{sgn}(x) = v_0 (2 \mathcal{H}(x) - 1), \quad (8.74)$$

where  $v_0 > 0$  is a constant,  $\operatorname{sgn}$  is the sign-function (equation (8.52)), and  $\mathcal{H}$  is the Heaviside step function (equation (8.54)). The free surface is assumed to be initially flat.

- Show that the linearized potential vorticity,  $H Q' = \zeta' - f \eta'/H$ , is given by

$$H Q'(x) = 2 v_0 \delta(x), \quad (8.75)$$

where  $\delta(x)$  is the Dirac delta with dimensions  $L^{-1}$ .

- (b) As we did in Section 8.3, solve for the geostrophic streamfunction  $\psi = g \eta / f$ .
- (c) Discuss the geostrophically adjusted streamfunction and velocity, and draw a sketch of  $\psi$  and  $v$ .

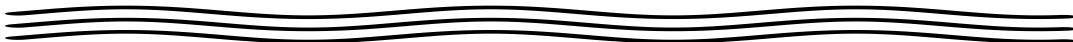
Here are some hints.

- We discuss properties of the Dirac delta in Chapter ???. However, to answer the first part of this question it is sufficient to know that the derivative of the Heaviside step function (a dimensionless step function) is the Dirac delta

$$\delta(x) = \frac{d\mathcal{H}(x)}{dx}, \quad (8.76)$$

so that the Dirac delta has dimensions of inverse length.

- For the second part, note that the streamfunction is exponentially decaying on either side of the  $x = 0$  according to  $\psi = \psi_0 e^{-|x|/L_d}$ , which then leads to a jump in the derivative approaching the origin from each side. Carefully use equations (8.63) and (8.76) to determine  $\psi_0$ .





## Chapter 9

# INTERNAL INERTIA-GRAVITY WAVES

In this chapter we continue the study of how gravity affects wave motion in geophysical fluids, here focusing on gravity waves realized in a continuously stratified fluid. As part of anticipating this discussion, recall Section 7.6 where we studied gravity waves in two shallow water layers and decomposed the motion into two modes: the *barotropic mode* and *baroclinic mode*. When moving to a continuously stratified fluid we encounter an infinity of baroclinic modes. These modes are associated with motion of interior stratification surfaces, and have very little projection onto motion of the upper free surface. Consequently, these waves are generally referred to as *internal gravity waves*.

Although the shallow water system anticipated some features of gravity waves in a continuous system, there are many properties revealed only when moving to continuous stratification. In particular, the dispersion relation for internal gravity waves presents the peculiar feature of having a group velocity that is perpendicular to the phase velocity. In fact, this property was already encountered in Chapter 5 when studying inertial waves in a homogeneous fluid. We here focus most attention on the case of internal gravity waves in a non-rotating reference frame. Extending to the case of an  $f$ -plane is relatively straightforward and builds on the inertial wave material from Chapter 5 as well as the shallow water inertia-gravity waves studied in Section 7.8.

### READER'S GUIDE TO THIS CHAPTER

We build on the prior wave mechanics chapters and assume familiarity with the equations for a Boussinesq ocean in Chapter ?? and the concept of Archimedean buoyancy from Chapter ???. Chapter 10 continues our study of internal inertia-gravity waves by focusing on a variety of geophysical mechanisms for the forcing of such waves.

Further resources for material in this chapter can be found in [Lighthill \(1978\)](#), [Gill \(1982\)](#), [Pedlosky \(2003\)](#), [Sutherland \(2010\)](#), [Cushman-Roisin and Beckers \(2011\)](#), [Kundu et al. \(2016\)](#), and [Vallis \(2017\)](#). The second half of this video offers some pedagogical visualizations of stratified flow phenomena, and this video provides more visualizations from simulations and laboratory tank experiments.

9.1	Loose threads . . . . .	250
9.2	Boussinesq ocean and its linearization . . . . .	251
9.2.1	Boussinesq ocean equations . . . . .	251
9.2.2	The prescribed reference state . . . . .	251
9.2.3	Buoyancy frequency compared to surface gravity waves . . . . .	252
9.2.4	Linearization around the background state of rest . . . . .	253

9.2.5	Energetics . . . . .	253
<b>9.3</b>	<b>Buoyancy oscillations</b> . . . . .	<b>255</b>
9.3.1	Unforced buoyancy oscillations . . . . .	255
9.3.2	The lack of oscillations with horizontal homogeneity . . . . .	256
<b>9.4</b>	<b>The linear Boussinesq ocean with <math>f = 0</math></b> . . . . .	<b>257</b>
9.4.1	Relative vorticity . . . . .	257
9.4.2	Wave equation for the vertical velocity . . . . .	257
9.4.3	Linearized boundary conditions . . . . .	258
<b>9.5</b>	<b>Free internal gravity waves with constant stratification</b> . . . . .	<b>260</b>
9.5.1	Plane wave ansatz . . . . .	260
9.5.2	Transverse nature of internal gravity waves . . . . .	261
9.5.3	Relative vorticity . . . . .	261
9.5.4	Amplitude of pressure fluctuations . . . . .	261
9.5.5	Vertical velocity component . . . . .	261
9.5.6	Horizontal velocity . . . . .	262
9.5.7	Dispersion relation . . . . .	262
9.5.8	Concerning the upper limit on the frequency . . . . .	263
9.5.9	Unpacking the dispersion relation . . . . .	263
9.5.10	Group velocity . . . . .	263
9.5.11	Force balance within an internal gravity wave . . . . .	266
9.5.12	Forced internal gravity wave packets . . . . .	268
9.5.13	Polarization relations and structure of a plane wave . . . . .	269
9.5.14	Energetics of a plane internal gravity wave . . . . .	271
<b>9.6</b>	<b>Reflection of gravity waves</b> . . . . .	<b>273</b>
9.6.1	Reflection conditions . . . . .	273
9.6.2	Specializing to internal gravity waves . . . . .	274
9.6.3	Comments and further reading . . . . .	276
<b>9.7</b>	<b>Linear Boussinesq ocean on an <math>f</math>-plane</b> . . . . .	<b>276</b>
9.7.1	Forced oscillator equation for horizontal velocity . . . . .	277
9.7.2	Vertical component to the relative vorticity . . . . .	277
9.7.3	Forced oscillator and free wave equations for vertical velocity . . . . .	277
9.7.4	Forced oscillator equation for buoyancy . . . . .	278
9.7.5	An equation for pressure . . . . .	279
9.7.6	Potential vorticity . . . . .	279
<b>9.8</b>	<b>Free inertia-gravity waves</b> . . . . .	<b>281</b>
9.8.1	Dispersion relation . . . . .	281
9.8.2	Near-inertial waves . . . . .	281
9.8.3	Force balance in an inertia-gravity wave . . . . .	282
9.8.4	Group velocity . . . . .	284
9.8.5	Polarization relations for a plane wave . . . . .	284
9.8.6	Energetics of a plane inertia-gravity wave . . . . .	285
9.8.7	Comments . . . . .	287
<b>9.9</b>	<b>Exercises</b> . . . . .	<b>287</b>

---

## 9.1 Loose threads

- energetic balances for reflected waves.
- Is there a general means to determine  $\mathbf{c}_g \cdot \mathbf{k} = 0$  without going through the process of computing the group velocity? Why do some waves have this property but others do not?

## 9.2 Boussinesq ocean and its linearization

Throughout this chapter we work with the Boussinesq ocean equations in their inviscid/adiabatic limit, and linearize the governing equations around a rest state in exact hydrostatic balance. In this section we review the governing equations from Chapter ?? and perform the linearization.

### 9.2.1 Boussinesq ocean equations

We make use of the governing equations for an adiabatic and inviscid Boussinesq ocean written in the form (see Section ??)

$$\rho_0 [\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + 2\boldsymbol{\Omega} \times \mathbf{v}] = -\nabla p - \rho g \hat{\mathbf{z}} \quad \text{inviscid velocity equation} \quad (9.1a)$$

$$\nabla \cdot \mathbf{v} = 0 \quad \text{non-divergent flow} \quad (9.1b)$$

$$(\partial_t + \mathbf{v} \cdot \nabla) \rho = 0 \quad \text{adiabatic density equation} \quad (9.1c)$$

$$\rho = -\rho_0 \alpha_\Theta \Theta \quad \text{linear equation of state.} \quad (9.1d)$$

The density field contains a nontrivial spatial structure

$$\rho = \rho(\mathbf{x}, t), \quad (9.2)$$

with the Boussinesq reference density,  $\rho_0$ , a constant, and we furthermore assume the thermal expansion coefficient,  $\alpha_\Theta$ , is a constant. Hence, the density and Conservative Temperature,  $\Theta$ , are linearly proportional, with the adiabatic density equation (9.1c) meaning that Conservative Temperature is materially constant. The velocity field resolved by the Boussinesq ocean is non-divergent so that it does not support acoustic waves.<sup>1</sup> The inhomogeneous density field couples to the gravity field to render a nonzero buoyancy that appears in the velocity equation (9.1a). Buoyancy is the essential ingredient for the internal gravity waves studied in this chapter. Finally, we make use of the Traditional Approximation (Section ??) and tangent plane approximation (Section ??) with the planetary rotation approximated by

$$2\boldsymbol{\Omega} \approx f_0 \hat{\mathbf{z}}, \quad (9.3)$$

where  $f_0$  is the planetary vorticity set either to zero or a nonzero constant.

### 9.2.2 The prescribed reference state

We consider a three-component decomposition of density into a constant Boussinesq reference density,  $\rho_0 > 0$ , plus a prescribed static reference density,  $\rho_R(z) > 0$ , and a perturbation density,  $\rho'(\mathbf{x}, t)$ ,

$$\rho(\mathbf{x}, t) = \rho_0 + \rho_R(z) + \rho'(\mathbf{x}, t). \quad (9.4)$$

We assume the following inequalities hold

$$\rho_R \ll \rho_0 \quad \text{and} \quad |\rho'| \ll \rho_R, \quad (9.5)$$

with the first inequality following from the Boussinesq ocean approximation, and the second inequality supporting linearization in Section 9.2.4. The buoyancy is decomposed according to

---

<sup>1</sup>See Section 3.2 for more on how we interpret the quasi-compressible properties of the Boussinesq ocean.

the density decomposition (9.4), so that

$$b = -g(\rho - \rho_0)/\rho_0 = -g(\rho_R + \rho')/\rho_0 = b_R + b'. \quad (9.6)$$

Likewise, we decompose pressure so that

$$p = p_o(z) + p_R(z) + p'(\mathbf{x}, t), \quad (9.7)$$

where

$$dp_o(z)/dz = -\rho_0 g \quad \text{and} \quad dp_R(z)/dz = -\rho_R(z) g, \quad (9.8)$$

so that  $p_o$  and  $p_R$  are both hydrostatic base state pressures given by

$$p_o(z) = -\rho_0 g z \quad \text{and} \quad p_R(z) = g \int_z^0 \rho_R(z') dz'. \quad (9.9)$$

The corresponding pressure and gravity contributions to the velocity equation (9.1a) take the form

$$-\nabla p - \rho g \hat{\mathbf{z}} = -\rho_o \nabla \varphi' - \rho' g \hat{\mathbf{z}} = -\rho_o (\nabla \varphi' - b' \hat{\mathbf{z}}), \quad (9.10)$$

where we introduced the normalized dynamic pressure (with dimensions of squared velocity)

$$\varphi' = p'/\rho. \quad (9.11)$$

Introducing the reference state into the velocity equation (9.1a) and density equation (9.1c) leads to

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + f_o \hat{\mathbf{z}} \times \mathbf{v} = -\nabla \varphi' + b' \hat{\mathbf{z}} \quad \text{decomposed velocity equation} \quad (9.12a)$$

$$(\partial_t + \mathbf{v} \cdot \nabla) b' = -w N_R^2 \quad \text{decomposed buoyancy equation.} \quad (9.12b)$$

In the buoyancy equation we introduced the reference state's squared buoyancy frequency

$$N_R^2 = -\frac{g}{\rho_0} \frac{d\rho_R}{dz}. \quad (9.13)$$

It is notable that the velocity equation (9.12a) is mathematically identical to the original velocity equation (9.1a). However, the fluctuating buoyancy equation (9.12b) is distinct from the original buoyancy equation (9.1c). Namely, the reference state squared buoyancy frequency couples to the vertical velocity to provide a source,  $-w N_R^2$ , for the material evolution of the perturbation buoyancy.

### 9.2.3 Buoyancy frequency compared to surface gravity waves

As revealed in this chapter, the buoyancy frequency of the background state,  $N_R$ , sets the time scale for the internal gravity wave motions. To garner a sense for its scale, we compare this frequency to that found for the deep water surface gravity waves studied in Section 4.5.5. Making use of equation (4.123a) for deep water gravity waves leads to the ratio

$$\frac{N_R^2}{\omega_{\text{dww}}^2} = \frac{N_R^2}{g |\mathbf{k}|_{\text{dww}}} = \left| \frac{g}{\rho_0} \frac{d\rho_R}{dz} \frac{\Lambda_{\text{dww}}}{2\pi g} \right| = \left| \frac{\Delta\rho_R}{2\pi\rho_0} \right|. \quad (9.14)$$

In the final step we took the deep water wave length,  $\Lambda_{\text{dww}}$ , as the vertical scale over which to measure the vertical density difference,  $\Delta\rho_{\text{R}}$ . Now the deep water wavelength is just a few meters, which means that  $|\Delta\rho_{\text{R}}| \ll \rho_{\text{o}}$ . We conclude that the buoyancy frequency is much less than the frequency of deep water surface gravity waves. We can understand this result by noting that the gravitational restoring force acting on deep water surface waves is far stronger than the reduced gravity acting on internal waves, so we expect the deep water waves to oscillate much faster.

Now consider the same calculation for shallow water surface gravity waves by making use of the dispersion relation (4.123b) to yield

$$\frac{N_{\text{R}}^2}{\omega_{\text{sww}}^2} = \frac{N_{\text{R}}^2}{|\mathbf{k}|_{\text{sww}}^2 c_{\text{grav}}^2} = \left| \frac{g}{\rho_{\text{o}}} \frac{d\rho_{\text{R}}}{dz} \frac{\Lambda_{\text{sww}}}{2\pi |\mathbf{k}|_{\text{sww}} g H} \right| = \left| \frac{\Delta\rho_{\text{R}}}{\rho_{\text{o}}} \frac{1}{2\pi |\mathbf{k}|_{\text{sww}} H} \right|. \quad (9.15)$$

In this case the density difference,  $\Delta\rho_{\text{R}}$ , is computed over a vertical length scale equal to the wavelength,  $\Lambda_{\text{sww}}$ . But since  $|\mathbf{k}|_{\text{sww}} H \ll 1$  for shallow water gravity waves, we take  $\Delta\rho_{\text{R}}$  over the full depth of the ocean. Even so, the density ratio remains small,  $\Delta\rho_{\text{R}}/\rho_{\text{o}} \ll 1$ . Yet now this ratio is divided by the small number  $2\pi |\mathbf{k}|_{\text{sww}} H = (2\pi)^2 H/\Lambda_{\text{sww}} \ll 1$ . We thus find that the internal gravity wave frequency can be on the order of the shallow water gravity wave frequency. So although the shallow water waves feel the full gravitational acceleration, just like deep water gravity waves, the huge scale for shallow water waves leads to far slower wave motion than the deep water gravity waves. As a result, the frequency for shallow water waves is on the order of that for internal waves.

#### 9.2.4 Linearization around the background state of rest

The second of the density inequalities (9.5) implies the buoyancy inequality

$$|b'| \ll b_{\text{R}}. \quad (9.16)$$

Evidently, fluctuations in the buoyancy field, which can be positive or negative, are far smaller in magnitude than the reference state buoyancy. These small fluctuations in buoyancy lead to correspondingly small fluctuations in dynamic pressure,  $\varphi'$ , and small fluctuations in the velocity,  $\mathbf{v}'$ . The linearized velocity equation (9.12a) and buoyancy equation (9.12b) are obtained by dropping nonlinear product of fluctuating fields

$$\partial_t \mathbf{u}' + f_{\text{o}} \hat{\mathbf{z}} \times \mathbf{u}' = -\nabla_{\text{h}} \varphi' \quad \text{linearized horizontal velocity equation} \quad (9.17a)$$

$$\partial_t w' = -\partial_z \varphi' + b' \quad \text{linearized vertical-velocity equation} \quad (9.17b)$$

$$\partial_t b' = -w' N_{\text{R}}^2 \quad \text{linearized buoyancy equation} \quad (9.17c)$$

$$\nabla \cdot \mathbf{v}' = 0 \quad \text{continuity for velocity fluctuations.} \quad (9.17d)$$

The final equation expresses the non-divergent nature of the fluctuating velocity field.<sup>2</sup> These three linear equations are the linearized governing equations used in this chapter.

#### 9.2.5 Energetics

We studied the energetics of the nonlinear equation set (9.1a)-(9.1d) in Section ???. We here revisit that discussion and then derive energetics for the linearized equations (9.17a)-(9.17d).

---

<sup>2</sup>The velocity has been decomposed into its background state flow,  $\mathbf{v}_{\text{R}}$ , which is assumed here to be zero, plus a fluctuation around the background flow,  $\mathbf{v}'$ . Since  $\nabla \cdot \mathbf{v} = 0$  and  $\mathbf{v}_{\text{R}} = 0$ , we see that  $\nabla \cdot \mathbf{v}' = 0$ .

### Adiabatic inviscid Boussinesq ocean

The kinetic energy per volume and potential energy per volume in a Boussinesq ocean are given by

$$\rho_0 \mathcal{K} = \rho_0 \mathbf{v} \cdot \mathbf{v}/2 \quad \text{and} \quad \rho \Phi = \rho g z, \quad (9.18)$$

where we introduced the kinetic energy per mass and potential energy per mass (i.e., the geopotential)

$$\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2 \quad \text{and} \quad \Phi = g z. \quad (9.19)$$

Material evolution of the kinetic energy is derived by making use of the velocity equation (9.1a), in which

$$\rho_0 D\mathcal{K}/Dt = -\nabla \cdot (\mathbf{v} p) - \rho g w. \quad (9.20)$$

Material evolution of the potential energy per volume for an adiabatic Boussinesq ocean is given by

$$D(\rho \Phi)/Dt = g \rho w. \quad (9.21)$$

We are thus led to the mechanical energy equation for the adiabatic and inviscid Boussinesq ocean

$$\rho_0 D\mathcal{K}/Dt + \rho D\Phi/Dt = -\nabla \cdot (\mathbf{v} p), \quad (9.22)$$

so that mechanical energy following a fluid element is modified by convergence of the pressure flux. We thus identify  $\mathbf{v} p$  as the flux of mechanical energy in the Boussinesq ocean.

### Linearized adiabatic inviscid Boussinesq ocean

In the process of deriving the linearized equation set (9.17a)-(9.17d), we gave no attention to the energetic balances. Hence, it is not *a priori* clear that a sensible energetic balance exists for these linear equations. However, as we now show, there is indeed an energetic balance that manifests a physically relevant exchange of mechanical energy between its kinetic and potential forms. For the kinetic energy per mass,  $\mathcal{K}' = \mathbf{v}' \cdot \mathbf{v}'/2$ , use of the velocity equation (9.17a) readily yields

$$\partial_t \mathcal{K}' = -\nabla \cdot (\mathbf{v}' \varphi') + w' b'. \quad (9.23)$$

The potential energy is a bit less straightforward. Namely, multiplying the linear buoyancy equation (9.17c) by  $b'$  readily finds

$$\partial_t \mathcal{A}' = -w' b' \quad \text{with} \quad \mathcal{A}' = (b'/N_R)^2/2. \quad (9.24)$$

Recall the discussion in Section ??, where  $\mathcal{A}$  is identified as a measure of the available potential energy contained in small amplitude fluctuations of buoyancy surfaces.<sup>3</sup> Evidently, the linearized Boussinesq ocean has a mechanical energy budget given by

$$\partial_t (\mathcal{K}' + \mathcal{A}') = -\nabla \cdot (\mathbf{v}' \varphi'). \quad (9.25)$$

Hence, a time tendency for the sum of the kinetic energy plus the available potential energy is driven by convergences in the dynamic pressure flux,  $\mathbf{v}' \varphi'$ .

---

<sup>3</sup>In Section 9.5.14 we offer a further interpretation of the available potential energy as it appears for an internal gravity wave.

## 9.3 Buoyancy oscillations

As a starting point to studying small amplitude flow features emerging from the linearized Boussinesq ocean equations (9.17a)-(9.17d), consider  $\Omega = 0$  (flow in a non-rotating reference frame), in which the buoyancy equation and vertical velocity equation are

$$\partial_t b' + w' N_R^2 = 0 \quad \text{and} \quad \partial_t w' = -\partial_z \varphi' + b'. \quad (9.26)$$

The time derivative of the vertical velocity equation and use of the buoyancy equation, along with the complement operations, yield

$$(\partial_{tt} + N_R^2) w' = -\partial_{tz} \varphi' \quad \text{and} \quad (\partial_{tt} + N_R^2) b' = N_R^2 \partial_z \varphi'. \quad (9.27)$$

Each of these equations describes a forced oscillator with angular frequency given by the buoyancy frequency,  $N_R$ . We determine the forcing by determining the pressure, which satisfies the Poisson equation

$$-\nabla^2 \varphi' = -\partial_z b', \quad (9.28)$$

along with boundary conditions specific to the domain. As studied in Sections ?? and ??, this pressure equation results from taking the divergence of the velocity equation (9.17a) (with  $\Omega = 0$ ) and using the continuity equation,  $\nabla \cdot \mathbf{v}' = 0$ . In the absence of any vertical structure to the buoyancy (i.e.,  $\partial_z b' = 0$ ), the pressure satisfies Laplace's equation,  $\nabla^2 \varphi' = 0$ , which renders a zero pressure fluctuation in the absence of boundary effects.

### 9.3.1 Unforced buoyancy oscillations

Ignoring the pressure forcing in the forced oscillator equations (9.27) leads to free oscillations for the buoyancy and vertical velocity

$$(\partial_{tt} + N_R^2) w'_{\text{free}} = 0 \quad \text{and} \quad (\partial_{tt} + N_R^2) b'_{\text{free}} = 0. \quad (9.29)$$

Evidently, both the vertical velocity and the buoyancy exhibit simple harmonic oscillations with angular frequency,  $N_R$ . Such free oscillations occur when the pressure field has zero vertical gradient,  $\partial_z \varphi' = 0$ , in which case there is no resistance to free oscillations in the vertical.

By ignoring the pressure fluctuations, we are in effect ignoring the impact of the buoyancy fluctuation on pressure. This approach is not dynamically self-consistent. However, it is a common approach, sometimes referred to as the *parcel method*, whereby we consider the motion of a test fluid element that is assumed to have no impact on the surrounding fluid.<sup>4</sup> We considered the pros and cons of this approach when studying Archimedean buoyancy in Section ?? and effective buoyancy in Section ??.

As seen in Section 9.5.11, fluid particles exhibit free oscillations when they move in directions that parallel the surfaces of constant phase for internal gravity waves *if* the phase surfaces are not horizontal. On constant phase surfaces the pressure is spatially constant so that the only force acting on the fluid particle arises from gravity in a vertically stratified fluid; i.e., buoyancy. We return to this conceptual picture in Section 9.5.11 as it is fundamental to the forces acting within an internal gravity wave.

---

<sup>4</sup>See Sections ?? and ?? for more on test fluid elements.

### 9.3.2 The lack of oscillations with horizontal homogeneity

When discussing buoyancy oscillations one sometimes finds it convenient to consider a horizontally homogenous fluid and conceive of the oscillations moving horizontal stratification surfaces up and down. However, there are caveats to this conceptual picture that offer further hints at the physics of gravity waves.

Horizontal homogeneity means that the horizontal pressure gradient vanishes, and thus the horizontal velocity vanishes (remember we are assuming a non-rotating reference frame so that  $\Omega = 0$ ). To maintain continuity for the Boussinesq ocean requires  $\partial_z w' = 0$ , which means, with zero horizontal motion, that  $w' = 0$  throughout the domain. It follows that for a Boussinesq ocean there can be no buoyancy fluctuation with  $\mathbf{v}' = 0$ . Although this conclusion might be clear enough, we step through the details to further support an understanding of the linear equations.

A vanishing vertical velocity in the forced oscillator equation (9.27) means that the vertical pressure gradient is time independent,  $\partial_t(\partial_z \varphi') = 0$ . Furthermore, horizontal homogeneity for the pressure equation (9.28) means that

$$\partial_z(\partial_z \varphi' - b') = 0, \quad (9.30)$$

so that  $\partial_z \varphi' - b' = B(t)$ , where  $B$  is an arbitrary function of time. Taking a time derivative, and using  $\partial_t(\partial_z \varphi') = 0$ , leads to  $\partial_t b' = -\partial_t B$ , which means that the buoyancy fluctuation is vertically independent. A vertically independent buoyancy in the forced buoyancy oscillator equation (9.27) means that  $N_R^2 \partial_z \varphi'$  is also vertically independent, which can be satisfied if  $\partial_z \varphi' = 0$ . But then  $w' = 0$  and  $\partial_z \varphi' = 0$  in the vertical velocity equation (9.29) then means that  $b' = 0$ . We are thus led to conclude that the horizontally homogeneous linear Boussinesq system is dynamically consistent only with the trivial solution:  $\mathbf{v}' = 0$  and  $b' = 0$ . That is, horizontal homogeneity means exact hydrostatic balance in which, for a Boussinesq fluid, there is zero motion and zero buoyancy oscillations.<sup>5</sup>

As we see in the following sections, the study of internal gravity waves leads to fluctuations that have both horizontal and vertical spatial variations, thus allowing for nontrivial wave motions. Now a surface of constant phase for a plane wave has homogeneous flow properties, in which case fluid particles exhibit free buoyant oscillations along these surfaces.<sup>6</sup> Yet as seen in the following, such buoyancy oscillations exist only if the phase surfaces are sloped relative to the horizontal plane, thus allowing the background buoyancy to produce buoyancy oscillations along the phase surfaces. No internal gravity waves occur if the phase surface is strictly horizontal, with this result consistent with the above arguments. We further detail this physical picture in Section 9.5.11.

---

<sup>5</sup>A non-Boussinesq fluid can exhibit vertical motion even if the fluid has its density modified in a manner that does not introduce horizontal density gradients. For example, a horizontally homogeneous heating in an ocean with constant salinity will reduce the density and expand the water column, without introducing any horizontal inhomogeneities. This expansion results from the divergent nature of the flow field in a non-Boussinesq fluid. However, this expansion is not represented by a Boussinesq ocean. Namely, the Boussinesq ocean has a prognostic flow field that is non-divergent, so that a uniform heating is not directly felt by the flow. In Section ?? we discuss implications of this limitation of the Boussinesq ocean for the study of global mean sea level.

<sup>6</sup>This conclusion follows since internal gravity waves in a Boussinesq ocean are transverse waves, so that fluid particles move along constant phase surfaces.

## 9.4 The linear Boussinesq ocean with $f = 0$

We here study properties of the linearized Boussinesq equations (9.17a)-(9.17d) in a non-rotating reference frame

$$\partial_t \mathbf{v}' = -\nabla \varphi' + b' \hat{\mathbf{z}} \quad \text{linearized } \boldsymbol{\Omega} = 0 \text{ velocity equation} \quad (9.31a)$$

$$\partial_t b' = -w' N_R^2 \quad \text{linearized buoyancy equation} \quad (9.31b)$$

$$\nabla \cdot \mathbf{v}' = 0 \quad \text{continuity for velocity fluctuations.} \quad (9.31c)$$

### 9.4.1 Relative vorticity

Taking the curl of the velocity equation (9.31a) leads to the evolution equation for the relative vorticity<sup>7</sup>

$$\partial_t (\nabla \times \mathbf{v}') = \nabla \times b' \hat{\mathbf{z}} \implies \partial_t \boldsymbol{\omega}' = -\hat{\mathbf{z}} \times \nabla b'. \quad (9.32)$$

The right hand side is the baroclinicity vector (??) for a Boussinesq fluid. Hence, in the linearized Boussinesq equations for a fluid in a non-rotating reference frame, relative vorticity has a local time tendency driven by baroclinicity. In the absence of a buoyancy gradient, the vorticity vector in the linear theory is static and remains zero if initialized to zero.

As noted in Section ??, baroclinicity in a Boussinesq fluid only affects vorticity in the horizontal directions, here seen for the linearized vorticity equation (9.32) since

$$\hat{\mathbf{z}} \cdot (\hat{\mathbf{z}} \times \nabla b') = 0. \quad (9.33)$$

One interpretation of internal gravity waves is that they arise from oscillations of the baroclinicity vector that leads to an oscillation of the horizontal components of the relative vorticity.<sup>8</sup> Since the vertical component to the vorticity is unaffected by baroclinicity, in the linear theory the vertical vorticity is locally static

$$\partial_t \zeta' = 0 \quad \text{with} \quad \zeta' = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{v}'). \quad (9.34)$$

So if the flow starts with a zero vertical vorticity, then the linearized Boussinesq ocean maintains  $\zeta' = 0$ . In Section 7.5.1 we also found a static vertical component of relative vorticity holds for the shallow water gravity waves in a non-rotating reference frame (see equation (7.72)).

### 9.4.2 Wave equation for the vertical velocity

In Section 9.3 we derived the forced harmonic oscillator equation (9.27) for the vertical velocity

$$(\partial_{tt} + N_R^2) w' = -\partial_{tz} \varphi'. \quad (9.35)$$

We find it useful to derive a wave equation solely in terms of  $w'$ , which is then used to derive the dispersion relation for internal gravity waves. To eliminate pressure, start by taking the horizontal divergence of the horizontal components to the velocity equation (9.31a)

$$\partial_t (\nabla_h \cdot \mathbf{u}') = -\nabla_h^2 \varphi', \quad (9.36)$$

---

<sup>7</sup>Be careful to distinguish the relative vorticity vector,  $\boldsymbol{\omega}' = \nabla \times \mathbf{v}'$ , from the angular frequency,  $\omega$ , of a plane wave, and from the vertical component of the velocity vector,  $\mathbf{v} = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v + \hat{\mathbf{z}} w$ .

<sup>8</sup>For example, see the 16 minute mark of [this video from Prof. Long](#).

where we introduced the horizontal gradient operator and horizontal Laplacian operator

$$\nabla_h = \hat{x} \partial_x + \hat{y} \partial_y \quad \text{and} \quad \nabla_h^2 = \nabla_h \cdot \nabla_h = \partial_{xx} + \partial_{yy}. \quad (9.37)$$

Equation (9.36) says that time evolution of the horizontal velocity divergence is driven by the negative horizontal Laplacian acting on the dynamic pressure. Since the full velocity is non-divergent,  $\nabla \cdot \mathbf{v}' = \nabla_h \cdot \mathbf{u}' + \partial_z w' = 0$ , equation (9.36) becomes an equation for the vertical divergence of the vertical velocity

$$\partial_t(\partial_z w') = \nabla_h^2 \varphi'. \quad (9.38)$$

Taking the horizontal Laplacian of the forced oscillator equation (9.35) (recall  $N_R = N_R(z)$ ) yields

$$(\partial_{tt} + N_R^2) \nabla_h^2 w' = -\partial_{tz} \nabla_h^2 \varphi', \quad (9.39)$$

and then using equation (9.38) leads to the wave equation for the vertical velocity

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) w' = 0. \quad (9.40)$$

There is anisotropy in this equation due to the gravity force that distinguishes the vertical direction from the horizontal. The anisotropy manifests by the fully three dimensional Laplacian,  $\nabla^2$ , in the first term in equation (9.40), whereas just the horizontal Laplacian,  $\nabla_h^2$ , appears in the second term. We furthermore observe that the wave equation (9.40) shares features with the wave equation (5.27) satisfied by the inertial waves in a homogeneous fluid in a constantly rotating reference frame. Gravity, coupled to nonzero density gradients, breaks symmetry in the Boussinesq ocean by introducing the buoyancy acceleration that then supports internal gravity waves. Analogously, a rotating reference frame breaks symmetry of the homogeneous fluid by introducing the Coriolis acceleration that then supports inertial waves.

### 9.4.3 Linearized boundary conditions

For much of this chapter, we study freely traveling internal inertia-gravity waves in the absence of boundaries. However, we do have occasions to study waves within a bounded or semi-bounded domain, in which case we need to apply boundary conditions. Hence, we here formulate the linearized boundary conditions for these cases. To be specific, we orient the domain with  $z = 0$  the vertical position of a resting upper free surface boundary, such as for a resting ocean domain. However, when studying waves generated by bottom topography in Section 10.2, we place  $z = 0$  at the position of a flat bottom domain.

#### Kinematic boundary conditions

Since the reference flow is assumed to vanish, the kinematic no-normal flow boundary condition holds for the fluctuating flow

$$\mathbf{v}' \cdot \hat{\mathbf{n}} = 0 \quad \text{at solid boundaries.} \quad (9.41)$$

At a moving material boundary (Section ??), such as the ocean free surface at  $z = \eta(x, y, t)$ , the kinematic boundary condition (??)

$$(\partial_t + \mathbf{u} \cdot \nabla) \eta = w' \quad \text{at } z = \eta \quad (9.42)$$

is linearized to

$$\partial_t \eta = w' \quad \text{at } z = \eta. \quad (9.43)$$

In fact, we must go one step further to fully linearize this boundary condition. The reason is that  $w(\eta)$  is nonlinear, as seen by a Taylor expansion about the resting  $z = 0$  free surface position

$$w'(z = \eta) \approx w'(z = 0) + \eta \partial_z w'. \quad (9.44)$$

The term,  $\eta \partial_z w'$ , and all higher order terms, are nonlinear and so are dropped for the linear theory. We are thus led to the linearized surface kinematic boundary condition

$$\partial_t \eta = w' \quad \text{at } z = 0. \quad (9.45)$$

We made use of this same boundary condition linearization in Section 4.3.4 when studying linear surface gravity waves and surface capillary waves, in which we also evaluate boundary terms at  $z = 0$  rather than  $z = \eta$ .

### Dynamic boundary condition for pressure

Following our treatment of stresses acting at an interface in Section ??, we make use of Newton's third law to set the pressure boundary condition at the free surface. Namely, pressure, in the absence of surface tension (which we ignore here), is continuous across the free surface. We find it useful to determine a boundary condition for the pressure written as the decomposition (9.7)

$$p(\mathbf{x}, t) = p_o(z) + p_R(z) + p'(\mathbf{x}, t) = -\rho_o g z + g \int_z^0 \rho_R(z') dz' + p'(\mathbf{x}, t). \quad (9.46)$$

Assuming the atmospheric pressure applied on the free surface is a constant, which we take to be zero without loss of generality, leads to

$$0 = -\rho_o g \eta + g \int_\eta^0 \rho_R(z') dz' + p'(x, y, z = \eta, t). \quad (9.47)$$

Since the free surface height is small for linear waves, we Taylor expand the reference density around  $z' = 0$  so that the integral becomes

$$\int_\eta^0 \rho_R(z') dz' \approx \int_\eta^0 [\rho_R(0) + (\partial \rho_R / \partial z)_{z=0} z'] dz' = -\eta \rho_R(0) - (\eta^2 / 2) (\partial \rho_R / \partial z)_{z=0}. \quad (9.48)$$

Dropping the term  $\eta^2$  since it is nonlinear then leads to the boundary condition for the dynamic pressure evaluated at the free surface

$$\varphi'(x, y, z = \eta, t) = g \eta [\rho_o + \rho_R] / \rho_o \quad \text{at } z = \eta. \quad (9.49)$$

We fully linearize by evaluating the pressure at  $z = 0$  and note that  $\rho_R \ll \rho_o$  so that the boundary condition becomes

$$\varphi' = g \eta \quad \text{at } z = 0. \quad (9.50)$$

Evidently, the dynamic pressure at the ocean surface is determined by the free surface height. This boundary condition is identical to that used for the surface gravity waves as derived in Sections 4.3.2 and 4.3.3. Here, we needed to work a bit harder than for surface gravity waves. The reason is that here the fluid is stratified whereas we studied surface gravity waves at the

surface of a homogeneous ocean that resulted in irrotational motion, and thus a rather simple expression of Bernoulli's theorem (Section 4.3).

### Vertical velocity boundary condition

We further combine the pressure boundary condition (9.50) with the linearized surface kinematic boundary condition (9.45) to eliminate the free surface in favor of the vertical velocity

$$\partial_t \varphi' = g w' \quad \text{at } z = 0. \quad (9.51)$$

When studying vertical normal modes in an ocean domain in Section 10.5, we find it more convenient to have a boundary condition just for the vertical velocity. To eliminate the pressure we take the divergence of the horizontal velocity equation,  $\partial_t \mathbf{u}' = -\nabla_h \varphi'$ , and use the non-divergence condition on the velocity to yield

$$\partial_{tz} w' = \nabla_h^2 \varphi'. \quad (9.52)$$

Taking another time derivative and evaluating the expression at  $z = 0$  yields

$$\partial_{ttz} w' = \nabla_h^2 \partial_t \varphi' \implies \partial_{ttz} w' = g \nabla_h^2 w' \quad \text{at } z = 0, \quad (9.53)$$

where we used the boundary condition (9.51) to eliminate pressure in favor of the vertical velocity.

## 9.5 Free internal gravity waves with constant stratification

Thus far we have assumed that the reference state buoyancy frequency is a function of the vertical direction. We now specialize to the case of a constant frequency,  $N_R$ , in which case we study free plane internal gravity waves. We return in Section 10.4 to the more realistic case of vertically varying buoyancy frequency,  $N_R(z)$ , which requires the WKBJ approximation.

### 9.5.1 Plane wave ansatz

To examine the physics of internal gravity waves in a fluid with constant buoyancy frequency, we introduce a traveling plane wave ansatz for the velocity, dynamic pressure, and buoyancy

$$\mathbf{v}' = \tilde{\mathbf{v}} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \quad \text{and} \quad \varphi' = \tilde{\varphi} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad \text{and} \quad b' = \tilde{b} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}, \quad (9.54)$$

where the amplitudes,  $\tilde{\mathbf{v}}$ ,  $\tilde{\varphi}$ , and  $\tilde{b}$ , are generally complex numbers, and where we write the three dimensional and horizontal wavevectors as

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}} \quad \text{and} \quad \mathbf{k}_h = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}. \quad (9.55)$$

Substitution of the ansatz (9.54) into the linearized Boussinesq equations (9.31a)-(9.31c) allows us to determine relations between the velocity, pressure, and buoyancy for a traveling plane wave. These relations are referred to as *polarization* relations. Furthermore, as part of developing polarization relations we derive the dispersion relation (Section 9.5.7) that connects the wavevector,  $\mathbf{k}$ , to the wave angular frequency,  $\omega$ .

### 9.5.2 Transverse nature of internal gravity waves

The non-divergence condition holding for the Boussinesq ocean flow,  $\nabla \cdot \mathbf{v}' = 0$ , means that internal gravity waves are transverse, so that the velocity amplitude satisfies

$$\mathbf{k} \cdot \tilde{\mathbf{v}} = \mathbf{k}_h \cdot \tilde{\mathbf{u}} + k_z \tilde{w} = 0. \quad (9.56)$$

Consequently, fluid particles within a plane internal gravity wave move within the planes defined by constant phase, and these planes are perpendicular to the wavevector. Recall from Chapter 5 that plane inertial waves are also transverse.

### 9.5.3 Relative vorticity

The relative vorticity vector satisfies equation (9.32), which for a plane wave takes on the form

$$\omega \mathbf{k} \times \tilde{\mathbf{v}} = -i \tilde{b} (\hat{z} \times \mathbf{k}) \implies \mathbf{k} \times (\omega \tilde{\mathbf{v}} - i \hat{z} \tilde{b}) = 0. \quad (9.57)$$

Evidently, the vector,  $\omega \tilde{\mathbf{v}} - i \hat{z} \tilde{b}$ , is parallel to the wavevector. Taking the dot product of the  $\mathbf{k}$ -space relative vorticity equation (9.57) with  $\hat{z}$  leads to

$$\hat{z} \cdot (\mathbf{k} \times \tilde{\mathbf{v}}) = 0, \quad (9.58)$$

which is the  $\mathbf{k}$ -space expression for a zero vertical component to the relative vorticity,  $\zeta' = 0$ . This constraint means that the plane waves maintain  $\zeta' = 0$  since the horizontal velocity amplitudes are related by

$$k_x \tilde{v} = k_y \tilde{u}. \quad (9.59)$$

### 9.5.4 Amplitude of pressure fluctuations

Substitution of the plane wave ansatz into the pressure equation (9.28) renders

$$\tilde{\varphi}/\tilde{b} = -i k_z / |\mathbf{k}|^2. \quad (9.60)$$

Hence, the pressure and buoyancy in the plane wave are  $\pi/2$  radians out of phase. Observe that the pressure amplitude vanishes for a horizontal wavevector, in which  $k_z = 0$ . We expect this result for a free wave since vertical structure in the fluctuating buoyancy (i.e.,  $k_z \neq 0$ ) provides the source for pressure fluctuations as per the pressure Poisson equation (9.28).

### 9.5.5 Vertical velocity component

For the vertical velocity component, make use of the forced oscillator equation (9.35) to render a relation between the vertical velocity amplitude and amplitude of the dynamic pressure

$$\tilde{w}/\tilde{\varphi} = -\frac{k_z \omega}{N_R^2 - \omega^2}. \quad (9.61)$$

We can also relate the vertical velocity amplitude to the buoyancy amplitude through use of the pressure equation (9.60)

$$\tilde{w}/\tilde{b} = \frac{i k_z^2 \omega}{|\mathbf{k}|^2 (N_R^2 - \omega^2)}. \quad (9.62)$$

Evidently, the vertical velocity is either in phase or  $\pi$  radians out of phase with the pressure (depending on the sign of  $k_z$ ), whereas the vertical velocity is  $\pi/2$  out of phase with buoyancy.

### 9.5.6 Horizontal velocity

The horizontal portion of the velocity equation (9.31a) leads to the relation satisfied by the plane wave

$$\omega \mathbf{k}_h \cdot \tilde{\mathbf{u}} = \tilde{\varphi} |\mathbf{k}_h|^2 = -i \tilde{b} k_z |\mathbf{k}_h|^2 / |\mathbf{k}|^2, \quad (9.63)$$

where we used the pressure equation (9.60) for the second equality. Equivalently, we can use the transverse nature of the wave as per equation (9.56) to write

$$\omega \mathbf{k}_h \cdot \tilde{\mathbf{u}} = -\omega k_z \tilde{w} = -\frac{i k_z^3 \omega^2 \tilde{b}}{|\mathbf{k}|^2 (N_R^2 - \omega^2)}, \quad (9.64)$$

where we used equation (9.62) for the vertical velocity.

### 9.5.7 Dispersion relation

Equating equations (9.63) and (9.64) yields the dispersion relation for internal gravity waves

$$\omega^2 = |\mathbf{k}_h|^2 N_R^2 / |\mathbf{k}|^2 = N_R^2 \cos^2 \gamma \implies 0 \leq \omega \leq N_R. \quad (9.65)$$

In the second equality we introduced the angle between the wavevector and the horizontal plane

$$\cos \gamma = |\mathbf{k}_h| / |\mathbf{k}|, \quad (9.66)$$

with the geometry illustrated in Figure 9.1. An alternative means to write the dispersion relation (9.65) is given by

$$\frac{N_R^2 - \omega^2}{\omega^2} = \frac{k_z^2}{|\mathbf{k}_h|^2} = \tan^2 \gamma. \quad (9.67)$$

This expression makes it clear that internal gravity waves with maximum frequency,  $\omega = N_R$ , correspond to a horizontal phase velocity,  $k_z = 0$  and  $\gamma = 0$ . Conversely, the angular frequency goes to zero when  $\gamma = \pi/2$  whereby the waves have a purely vertical phase velocity,  $\mathbf{k}_h = 0$ .

The dispersion relation (9.65) can be directly derived from the wave equation (9.40) satisfied by the vertical velocity

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) w' = 0 \implies \omega^2 |\mathbf{k}|^2 - N_R^2 |\mathbf{k}_h|^2 = 0. \quad (9.68)$$

Furthermore, note that if we only extract the horizontal and time harmonic portion of  $w'$ , so that

$$w' = W(z) e^{i(\mathbf{k}_h \cdot \mathbf{x} - \omega t)}, \quad (9.69)$$

then the vertical structure function,  $W(z)$ , satisfies the ordinary differential equation

$$\left[ \frac{d^2}{dz^2} + \frac{|\mathbf{k}_h|^2 (N_R^2 - \omega^2)}{\omega^2} \right] W = 0 \implies \left[ \frac{d^2}{dz^2} + k_z^2 \right] W = 0. \quad (9.70)$$

This equation is satisfied if  $W = e^{i k_z z}$ , which is what we use for the constant  $N_R$  case now being studied. However, in Section 10.4 we find that equation (9.70) is generalized for the case with the background buoyancy frequency a function of the vertical,  $N_R = N_R(z)$ , in which case

the vertically propagating plane wave is no longer suited,  $W \neq e^{ik_z z}$ .

### 9.5.8 Concerning the upper limit on the frequency

The internal gravity wave dispersion relation (9.65) says that no internal gravity waves exist with angular frequency greater than the buoyancy frequency of the background reference state,  $\omega \leq N_R$ . What happens if a stratified fluid is agitated at a frequency  $\omega_{\text{source}} > N_R$ ? In this case, fluid particles do not have time to exhibit buoyancy oscillations. Instead, the fluid particles follow whatever displacements are forced on them by the external forcing, and with fluid particles in phase with the forcing. No propagating internal gravity waves are formed, and any energy imparted to the fluid stays local to the agitation. Indeed, with enough agitation the fluid becomes a forced turbulent soup.<sup>9</sup> For a mechanical analog, consider a forced linear pendulum where the forcing frequency is greater than the pendulum's natural frequency. In this case, the forcing fails to generate a natural oscillation, but instead it causes an incoherent and non-periodic back and forth motion of the pendulum.

### 9.5.9 Unpacking the dispersion relation

The angular frequency for an internal gravity wave only depends on the buoyancy frequency and the cosine of the angle the wavevector makes with the horizontal plane, whereas it is independent of the wavenumber,  $|\mathbf{k}|$ , and thus of the wavelength,  $2\pi/|\mathbf{k}|$ . Furthermore, the angular frequency possesses rotational (azimuthal) symmetry around the vertical direction. As such, the angular frequency is the same along the surface of a cone for wavevectors extending out from the origin along the  $\hat{\mathbf{z}}$  axis in either direction. Figure 9.1 provides an illustration for the two cones associated with a particular  $\gamma$ . The upper cone has  $k_z > 0$  and so waves on this cone have an upward phase velocity, whereas waves on the lower cone, with  $k_z < 0$ , have a downward phase velocity.

### 9.5.10 Group velocity

The dispersion relation (9.65) leads to the group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}}\varpi(\mathbf{k})$ , which for internal gravity waves is given by

$$\mathbf{c}_g = \frac{N_R k_z}{|\mathbf{k}|^3 |\mathbf{k}_h|} (k_z \mathbf{k}_h - |\mathbf{k}_h|^2 \hat{\mathbf{z}}). \quad (9.71)$$

In this section we identify a number of properties satisfied by the group velocity and summarize these properties in Figure 9.2.

#### **Group velocity is orthogonal to the phase velocity**

The internal wave group velocity (9.71) is orthogonal to the phase velocity

$$\mathbf{c}_g \cdot \mathbf{c}_p = 0, \quad (9.72)$$

where the phase velocity is

$$\mathbf{c}_p = (\omega/|\mathbf{k}|) \hat{\mathbf{k}}. \quad (9.73)$$

Hence, the group velocity is aligned parallel to constant phase surfaces, as with motion of fluid particles within an internal gravity wave. Recall that these properties also hold for inertial waves discussed in Section 5.3.3.

---

<sup>9</sup>See Section 13.2 of *Cushman-Roisin and Beckers* (2011) for further discussion on this point.

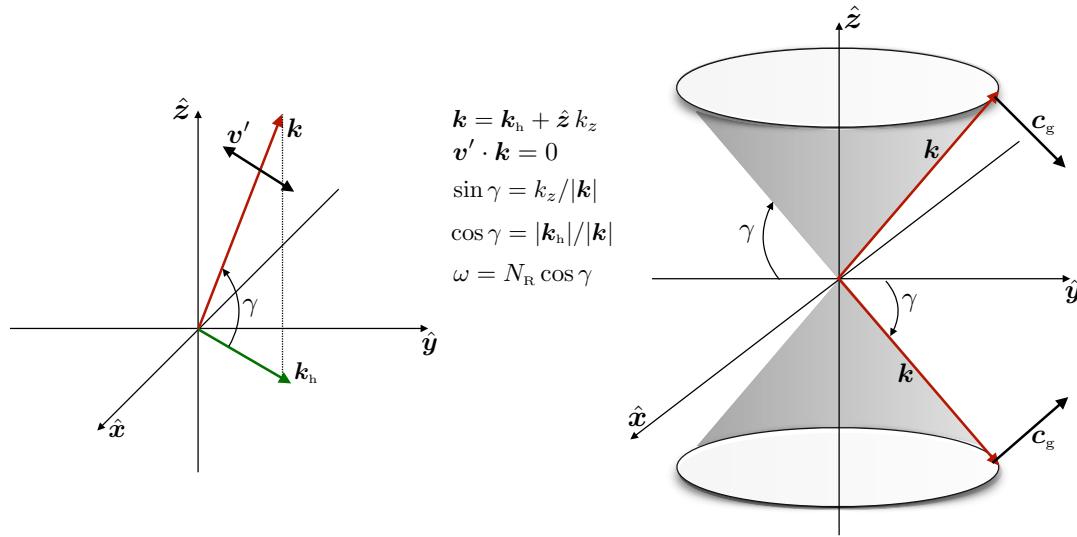


FIGURE 9.1: Left panel: A sample wavevector,  $\mathbf{k} = \mathbf{k}_h + \hat{z} k_z$ , for a plane internal gravity wave. The wave is transverse so that the velocity of fluid particles is orthogonal to the wavevector,  $\mathbf{k} \cdot \mathbf{v}' = 0$ . The dispersion relation,  $\omega = N_R \cos \gamma$ , is independent of the wavenumber,  $|\mathbf{k}|$ , but instead only depends on the cosine of the angle made with the horizontal plane,  $\cos \gamma = |\mathbf{k}_h|/|\mathbf{k}|$ . Right panel: another depiction of the dispersion relation. The angular frequency is the same for all wavevectors on the surface of an  $\omega$ -cone extending from the origin along the  $\hat{z}$  axis and with arbitrary magnitude. Correspondingly, with the phase velocity given by  $\mathbf{c}_p = \hat{\mathbf{k}}(\omega/|\mathbf{k}|)$ , higher wavenumber waves along a particular  $\omega$ -cone have slower phase speeds than lower wavenumber waves. Cones with  $k_z > 0$  correspond to internal gravity waves with an upward phase velocity, and those with  $k_z < 0$  are downward. The group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}}\varpi(\mathbf{k})$ , points in the  $\mathbf{k}$ -space direction of steepest ascent for the angular frequency, and so it is orthogonal to surfaces of constant  $\omega$ . For internal gravity waves, the group velocity is orthogonal to the phase velocity ( $\mathbf{c}_g \cdot \hat{\mathbf{k}} = 0$ ), and the group velocity points towards the horizontal plane (i.e., smaller  $|\gamma|$ ) since that is the direction of increasing angular frequency. Hence, waves with an upward phase velocity have a downward group velocity, and downward phase velocity waves have an upward group velocity. Evidently, internal wave energy is directed away from the inside of the  $\omega$ -cone and thus toward the horizontal plane.

We geometrically understand the orientation of the phase and group velocities as follows. The phase velocity is aligned with the wavevector,  $\mathbf{k}$ . All wavevectors emanating from the origin that make an angle,  $\gamma$ , have the same angular frequency and so form points along a particular  $\omega$ -cone (Figure 9.1). Consequently, surfaces of constant angular frequency are parallel to all these wavevectors. Now the group velocity,  $\mathbf{c}_g = \nabla_{\mathbf{k}}\varpi(\mathbf{k})$ , points in the  $\mathbf{k}$ -space direction of steepest ascent for the angular frequency, so that the group velocity is normal to surfaces of constant angular frequency. Hence, the group velocity points in a direction normal to the  $\omega$ -cone, and as such it is orthogonal to the phase velocity,  $\mathbf{c}_p \cdot \mathbf{c}_g = 0$ .

### The phase velocity and group velocity have opposing vertical components

The vertical component to the group velocity has the opposite sign of the vertical component to the phase velocity, which is seen by

$$(\mathbf{c}_p \cdot \hat{z})(\mathbf{c}_g \cdot \hat{z}) = -\frac{\omega k_z}{|\mathbf{k}|^2} \frac{N_R k_z |\mathbf{k}_h|}{|\mathbf{k}|^3} = -\frac{\omega^2 \sin^2 \gamma}{|\mathbf{k}|^2} < 0. \quad (9.74)$$

The phase velocity is directed away from the  $\gamma = 0$  horizontal plane, whereas the group velocity is directed toward the horizontal plane. This property is depicted in Figures 9.1 and 9.4.

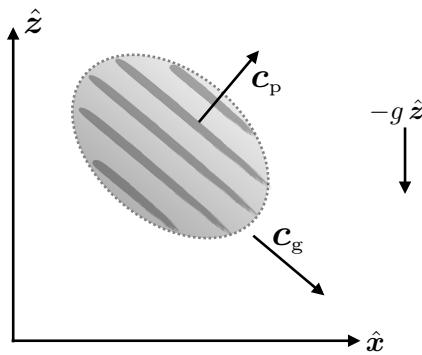


FIGURE 9.2: An example elliptical packet of internal gravity waves, with the packet moving down to the right according to the group velocity,  $\mathbf{c}_g$ , whereas phase lines move upward to the right. The packet moves parallel to the phase lines, with the phase velocity,  $\mathbf{c}_p$ , directed up and to the right.

Evidently, the internal gravity wave energy, which is fluxed in the direction of the group velocity, is directed away from the inside of the  $\omega$ -cone and towards the horizontal plane.

### The group velocity magnitude

The squared magnitude of the group velocity is given by

$$\mathbf{c}_g \cdot \mathbf{c}_g = \frac{N_R^2 k_z^2}{|\mathbf{k}|^4} = \frac{N_R^2 \sin^2 \gamma}{|\mathbf{k}|^2} \quad \text{where} \quad \sin \gamma = k_z/|\mathbf{k}|. \quad (9.75)$$

To help understand this expression it is useful to consider two extreme cases.

The group velocity vanishes when the wavevector is horizontal ( $k_z = 0$  and  $\gamma = 0$ ). As a result, no wave energy is propagated when the wavevector is horizontal.<sup>10</sup> When the wavevector is horizontal, all fluid particle motion in the wave is vertical, with fluid particles exhibiting vertical buoyancy oscillations at the maximum allowable gravity wave frequency,  $\omega = N_R$ . So even though the wave has energy when it has a horizontal phase velocity, the wave energy is not propagated since  $\mathbf{c}_g = 0$ . Rather, the wave energy is stationary as fluid particles exhibit vertical buoyancy oscillations.

At the opposite extreme, when the phase velocity is vertical, so that  $\mathbf{k}_h = 0$  and  $\gamma = \pi/2$ , the fluid particles move along a horizontal plane in an arbitrary horizontal direction. The group velocity is also horizontal and takes the form

$$\mathbf{c}_g = (N_R/|k_z|) \hat{\mathbf{k}}_h \quad \text{with} \quad \mathbf{k}_h = 0, \quad (9.76)$$

where  $\hat{\mathbf{k}}_h$  is any horizontal direction. The angular frequency (9.65) vanishes ( $\omega = 0$ ) so that the wave is stationary; i.e., phase lines do not move when  $\gamma = \pi/2$ .

### Relating the dispersion relation for internal gravity waves and inertial waves

The internal gravity wave dispersion relation (9.65) shares features with that for inertial waves, whose dispersion relation is given by equation (5.33), so that we have

$$\omega_{igw}^2 = N_R^2 \cos^2 \gamma \quad \text{and} \quad \omega_{inertial}^2 = f^2 \sin^2 \gamma. \quad (9.77)$$

<sup>10</sup>In Section 9.5.14 we see that the mechanical energy flux of an internal gravity wave is proportional to the group velocity, which is the case for any linear waves (see Chapter 2).

Both dispersion relations are independent of the wavenumber, depending only on the wave's orientation. For the gravity wave, the vertical is a special direction since it is the direction of the gravitational acceleration, whereas for the inertial wave, the horizontal direction is special since is the direction of the Coriolis acceleration. Indeed, the switch from  $\cos^2 \gamma$  to  $\sin^2 \gamma$  arises since gravitation acts in the vertical whereas the Coriolis acts in the horizontal.<sup>11</sup> Both angular frequencies have an upper bound, with inertial waves having angular frequencies no larger than the Coriolis frequency,  $|f|$ , and internal gravity waves having angular frequencies no larger than the buoyancy frequency,  $N_R$ .

### 9.5.11 Force balance within an internal gravity wave

To help further understand the internal gravity wave dispersion relation (9.65), consider a test fluid element that moves with the wave in the transverse direction; i.e., parallel to surfaces of constant phase. As discussed in Section ??, a test fluid element is assumed to have zero impact on the surrounding fluid, in particular it has no impact on the pressure. This assumption is broken when considering real fluid elements, since all fluid elements affect the surrounding fluid. However, the pressure is spatially constant along a constant phase direction. Hence, for a particle displaced in this direction it only feels the change in buoyancy. So our critique of the test fluid elements in Section 9.3.1 is tempered when considering motion along directions that parallel phase surfaces. In the following, we refer to the test fluid element as a fluid particle both for brevity and to accord with the literature.

#### Fluid particle displacement vector

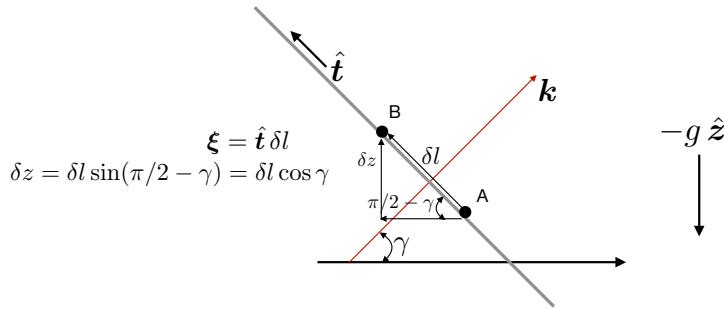


FIGURE 9.3: Depicting the displacement of a fluid particle along a direction transverse to a constant phase surface for a plane internal gravity wave whose wavevector makes an angle  $\gamma$  with the horizontal. We write the displacement vector as  $\xi(\mathbf{x}, t) = \delta l \hat{\mathbf{t}}$ , with the local tangent vector,  $\hat{\mathbf{t}}$ , pointing in a transverse direction orthogonal to the wavevector,  $\mathbf{k}$ . The displacement of the fluid particle from point  $A$  to point  $B$  occurs over a distance,  $\delta l$ , with a corresponding change in vertical position given by  $\delta z = \delta l \cos \gamma$ .

As depicted in Figure 9.3, we measure a fluid particle's position relative to its equilibrium position through a displacement field,

$$\xi(\mathbf{x}, t) = \hat{\mathbf{x}} \delta x + \hat{\mathbf{y}} \delta y + \hat{\mathbf{z}} \delta z. \quad (9.78)$$

The displacement field vanishes when the fluid particle is at its equilibrium position, and it has

<sup>11</sup>We discuss the force balances for an internal gravity wave in Section 9.5.11 and then for an inertia-gravity wave in Section 9.8.3.

a time tendency that equals to the velocity of a fluid particle<sup>12</sup>

$$\partial_t \boldsymbol{\xi}(\mathbf{x}, t) = \mathbf{v}'(\mathbf{x}, t). \quad (9.79)$$

Evidently, since the plane wave is transverse,

$$\mathbf{v}' \cdot \mathbf{k} = (\partial_t \boldsymbol{\xi}) \cdot \mathbf{k} = 0. \quad (9.80)$$

If the fluid particle moves in the vertical, it moves through the background reference density field,  $\rho_R(z)$ . Referring to Figure 9.3, assume the fluid particle starts at point *A* with local buoyancy equal to the reference buoyancy. As it rises an infinitesimal amount to point *B*, its buoyancy referenced to the local environment is negative since it started from a denser level. As such, the fluid particle wants to return to its deeper level. We can express the locally referenced buoyancy anomaly felt at point *B* in terms of the particle displacement

$$b' = -\boldsymbol{\xi} \cdot \nabla b_R = -\delta z (db_R/dz) = -\delta z N_R^2. \quad (9.81)$$

As a check on this expression, we find that if  $\delta z > 0$ , so that the fluid particle rises, then the buoyancy anomaly is negative,  $b' < 0$ , since it started from a deeper level where the density is larger. Hence, there is a buoyant acceleration back to the deeper level. Also observe that the time derivative of equation (9.81) leads to

$$\partial_t b' = -w' N_R^2, \quad (9.82)$$

which is the linearized buoyancy equation (9.17c).

### Equation of motion

We now make use of the result (9.81) in the equation of motion (9.31a). More precisely, we project that equation onto the transverse direction,  $\hat{\mathbf{t}}$ . This direction is fixed in time for a plane wave so that we have

$$\partial_t (\hat{\mathbf{t}} \cdot \mathbf{v}') = \hat{\mathbf{t}} \cdot (-\nabla \varphi' + b' \hat{\mathbf{z}}). \quad (9.83)$$

As noted above, there is no spatial gradient of flow properties along a constant phase surface. Hence, the tangential component of the pressure gradient vanishes

$$\hat{\mathbf{t}} \cdot \nabla \varphi' = 0, \quad (9.84)$$

which means pressure has no affect on fluid particles moving in the direction transverse to the wavevector. We write  $\delta l$  for the displacement along the phase surface, so that

$$\hat{\mathbf{t}} \cdot \mathbf{v}' = \partial_t (\delta l), \quad (9.85)$$

which brings the equation of motion (9.83) to the form

$$\partial_{tt}(\delta l) - \hat{\mathbf{z}} \cdot \hat{\mathbf{t}} b' = 0. \quad (9.86)$$

---

<sup>12</sup>We make use of a one dimensional displacement field in Section 3.3 when considering acoustic waves using a Lagrangian perspective. We also make use of fluid particle displacement vectors when studying the generalized Lagrangian mean in Section ??, as well as for tracer kinematics in Section ???. The present discussion is heuristic and skims over details explored in these other sections. We can afford some degree of informality since all perturbations here are small.

Noting that  $\hat{\mathbf{z}} \cdot \hat{\mathbf{t}} = \cos \gamma$  and making use of equation (9.81) for the buoyancy anomaly leads to

$$\partial_{tt}(\delta l) + \cos \gamma N_R^2 \delta z = 0. \quad (9.87)$$

Finally, we set  $\delta z = \delta l \cos \gamma$  to render the simple harmonic oscillator equation for displacements along the constant phase line

$$(\partial_{tt} + N_R^2 \cos^2 \gamma) \delta l = 0. \quad (9.88)$$

The angular frequency of the oscillations,  $N_R \cos \gamma$ , is precisely that arising from the internal gravity wave dispersion relation (9.65).

### Summary of the physical picture for internal gravity waves

The above analysis suggests the following physical picture for internal gravity waves. Namely, the waves consist of fluid particle motion in directions that parallel the constant phase surfaces, which accords with the property of all transverse waves. As the particles move along the transverse direction,  $\hat{\mathbf{t}}$ , they sample the background buoyancy field and thus experience buoyancy accelerations that act to return the particle to its neutral buoyancy position. The buoyancy force acting on the particles only depends on the angle of the phase lines relative to the horizontal. Oscillatory motion arises from over-shooting the neutral buoyancy position, with the oscillations having an angular frequency,  $N_R \cos \gamma$ . The oscillations have their maximum angular frequency when the angle,  $\gamma = 0$ , which arises since the particles are moving vertically and thus fully sampling the background buoyancy field. In contrast, there are no oscillations when the particles move along a horizontal phase surface ( $\gamma = \pi/2$  so that  $\hat{\mathbf{k}} = \hat{\mathbf{z}}$ ) since horizontal surfaces sample the same background buoyancy. The vanishing of oscillations with  $\gamma = \pi/2$  accords with our discussion in Section 9.3.2. The force balance dependence on the angle,  $\gamma$ , explains the geometry of the  $\omega$ -cones in Figure 9.1, which, in turn, explains why the group velocity is perpendicular to the phase velocity.

We emphasize that although the fluid particles are moving in a direction that parallels the constant phase surfaces, they are *not* moving with those surfaces. Rather, the fluid particles are oscillating in the transverse direction, and it is their oscillation that gives rise to the wave itself and thus to the movement of the phase surfaces through the fluid.

#### 9.5.12 Forced internal gravity wave packets

A further remark about the dispersion relation written as in equation (9.67) concerns the case where we know the frequency of the wave and the buoyancy frequency, in which case the angle  $\gamma$  is specified. For example, consider a local source (say an oscillating disk) with angular frequency,  $\omega_{\text{source}}$ , moving with a small amplitude in a stratified fluid with  $\omega_{\text{source}} < N_R$ . This source preferentially forces internal waves of angular frequency  $\omega = \omega_{\text{source}}$ . Consequently, the left hand side of the dispersion relation (9.67) is specified.

Gravity wave packets extend outward from the oscillating source and define the group velocity direction (not the phase velocity direction). Consequently, the group velocity extends outward from the source at an angle  $\pi/2 - \gamma$  from the horizontal. The reason the source defines the group velocity is that energy is input to the wave field by the oscillating source, and the wave energy propagates along the group velocity direction with the wave packets (we discuss internal gravity wave energy in Section 9.5.14). The phase velocity is directed perpendicular to the group velocity and it makes an angle  $\gamma$  with the horizontal. Hence, as the source angular frequency approaches the buoyancy frequency,  $\gamma$  approaches zero so that the group

velocity “cross” pattern steepens toward the vertical axis. Figure 9.4 provides a schematic of the experiment along with links to videos illustrating the remarkable wave patterns in a laboratory setting as well as a numerical simulation.<sup>13</sup>

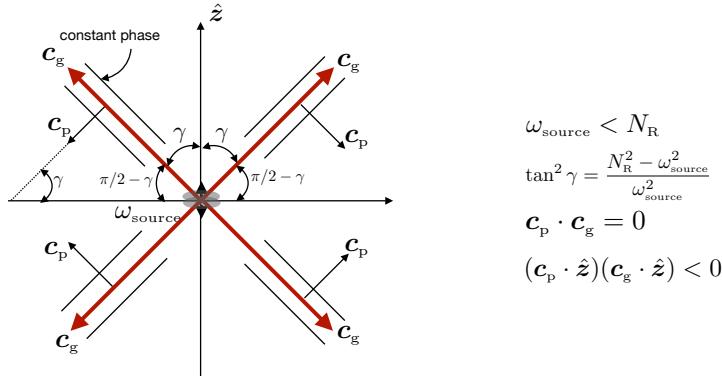


FIGURE 9.4: A small amplitude oscillating source in a stratified fluid with angular frequency,  $\omega_{\text{source}} < N_R$ . If the source has small amplitude then it generates linear internal gravity waves with  $\omega = \omega_{\text{source}}$ . These waves radiate wave packets with group velocity,  $\mathbf{c}_g$ , away from the source and with an angle,  $\pi/2 - \gamma$ , from the horizontal. The angle,  $\gamma$ , is determined by the dispersion relation through  $\tan^2 \gamma = (N_R^2 - \omega_{\text{source}}^2)/\omega_{\text{source}}^2$ . The group velocity is parallel to lines of constant phase and perpendicular to the phase velocity so that  $\mathbf{c}_p \cdot \mathbf{c}_g = 0$ . Furthermore, the phase and group velocities are oriented so that  $(\mathbf{c}_g \cdot \hat{\mathbf{z}})(\mathbf{c}_p \cdot \hat{\mathbf{z}}) < 0$ , so that if the group velocity is directed upward then the phase velocity is directed downward, and vice versa. As the source angular frequency approaches the buoyancy frequency,  $\gamma$  approaches zero so that the group velocity “cross” pattern steepens toward the vertical axis. The pattern is sometimes referred to as *St. Andrew’s cross*. [This video](#) and [this video](#), both from the geophysical fluids laboratory at The Australian National University, as well as [this video from Prof. Rhines’ lab at the University of Washington](#), illustrate the phenomena in a laboratory settings, whereas [this animation from Prof. Durran’s website at the University of Washington](#) illustrates the phenomena in a numerical simulation. [Sutherland \(2010\)](#) in his Chapter 5 works through the boundary value problems arising from both an oscillating cylinder and a sphere, thus illustrating the analytical methods available to solve for the generated internal gravity wave fields.

Internal gravity waves have an angular frequency that is bounded above by the buoyancy frequency:  $0 \leq \omega \leq N_R$ . There are no internal gravity waves with frequency larger than the buoyancy frequency of the background reference state. This limit exists since internal gravity waves are coherent patterns of buoyancy oscillations, with the angular frequency of that oscillation taken from the background stratification (see Section 9.5.11 for more on this perspective).

### 9.5.13 Polarization relations and structure of a plane wave

We now piece together the polarization relations to provide expressions for the velocity, pressure, and buoyancy within a wave. We first express all fields in terms of the buoyancy amplitude,  $\tilde{b}$ , and then in terms of the pressure amplitude,  $\tilde{\varphi}$ . These amplitudes are related via the pressure equation from Section 9.5.4, whereby

$$\tilde{\varphi}/\tilde{b} = -i k_z / |\mathbf{k}|^2 = -i \sin \gamma / |\mathbf{k}|. \quad (9.89)$$

<sup>13</sup>It is notable that [this video from Prof. Rhines’ lab](#) shows both a St. Andrew’s cross, which is well described by linear theory, as well as horizontal phase lines indicative of  $\gamma = \pi/2$  vertical phase propagation and horizontal group propagation. Paraphrasing from an email discussion with Prof. Rhines, the horizontal phase lines are in part the result of mixing (a nonlinear process) at the oscillating circular cylinder, which affects the background buoyancy frequency enough to alter the linear waves a bit. This mixing causes a diffusive drift away from the source that excites low frequency internal waves, which move rapidly ahead of the turbulently mixed fluid.

Neither is more fundamental, and yet one amplitude might be more readily available than the other so it is useful to provide both expressions.

### Vertical velocity component

The dispersion relation in the form of equation (9.67) brings the vertical velocity expressions (9.61) and (9.62) into the tidy forms

$$\tilde{w}/\tilde{\varphi} = -\frac{k_z \omega}{N_R^2 - \omega^2} = -\frac{|\mathbf{k}_h|^2}{\omega k_z} \quad (9.90a)$$

$$\tilde{w}/\tilde{b} = \frac{i k_z^2 \omega}{|\mathbf{k}|^2 (N_R^2 - \omega^2)} = \frac{i |\mathbf{k}_h|^2}{\omega |\mathbf{k}|^2} = \frac{i \cos \gamma}{N_R}. \quad (9.90b)$$

Notice how the vertical velocity and pressure are in phase (or  $\pi$  out of phase), whereas the vertical velocity and buoyancy are  $\pi/2$  out of phase.

### Horizontal velocity

For the horizontal velocity we use the transverse nature of the wave (Section 9.5.2) to relate the horizontal components to the vertical

$$\mathbf{k}_h \cdot \tilde{\mathbf{u}} = -k_z \tilde{w}, \quad (9.91)$$

as well as the zero vertical component of the relative vorticity (Section 9.5.3) so that

$$k_x \tilde{v} = k_y \tilde{u}. \quad (9.92)$$

These two relations, along with the vertical velocity amplitude equations (9.90a) and (9.90b), lead to

$$\tilde{\mathbf{u}} = -\frac{i \tilde{b} \mathbf{k}_h \sin \gamma}{N_R |\mathbf{k}_h|}. \quad (9.93)$$

As for the vertical velocity, we here find that the horizontal velocity is  $\pi/2$  out of phase with the buoyancy.

### Wave solutions in terms of the buoyancy amplitude

Bringing the previous results together allows us to express the solutions for the velocity, pressure, and buoyancy within a freely propagating internal gravity wave. The solutions in terms of a real buoyancy amplitude,  $\tilde{b}$ , are given by

$$b' = \tilde{b} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.94a)$$

$$\varphi' = \frac{\tilde{b} \sin \gamma}{|\mathbf{k}|} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.94b)$$

$$\mathbf{u}' = \frac{\tilde{b} \mathbf{k}_h \sin \gamma}{N_R |\mathbf{k}_h|} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.94c)$$

$$w' = -\frac{\tilde{b} \cos \gamma}{N_R} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.94d)$$

$$\mathbf{v}' = \frac{c_g \tilde{b} |\mathbf{k}|^2}{N_R^2 k_z} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (9.94e)$$

Equation (9.94e) for  $\mathbf{v}'$  made use of equation (9.71) for the group velocity, thus exposing the parallel nature of the particle velocity and group velocity for plane internal gravity waves. Note that the buoyancy is  $\pi/2$  out of phase with the pressure along with the three components of the velocity.

### Wave solutions in terms of the pressure amplitude

An alternative suite of wave solutions arises when assuming a real pressure amplitude,  $\tilde{\varphi}$ , which leads to

$$\varphi' = \tilde{\varphi} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.95a)$$

$$b' = -\frac{\tilde{\varphi} |\mathbf{k}|^2}{k_z} \sin(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.95b)$$

$$\mathbf{u}' = \frac{\tilde{\varphi} \mathbf{k}_h}{\omega} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.95c)$$

$$w' = -\frac{\tilde{\varphi} |\mathbf{k}_h|^2}{\omega k_z} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.95d)$$

$$\mathbf{v}' = \frac{c_g \tilde{\varphi} |\mathbf{k}|^4}{N_R^2 k_z^2} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (9.95e)$$

### 9.5.14 Energetics of a plane internal gravity wave

In Section 9.2.5 we derived the energetics for the linearized Boussinesq ocean equations and were led to the energy equation (9.25)

$$\partial_t (\mathcal{K}' + \mathcal{A}') = -\nabla \cdot (\mathbf{v}' \varphi') \quad \text{with} \quad \mathcal{K}' = \mathbf{v}' \cdot \mathbf{v}' / 2 \quad \text{and} \quad \mathcal{A}' = (b'/N_R)^2 / 2, \quad (9.96)$$

where  $\mathcal{K}'$  is the kinetic energy per mass of the linear fluctuation, and  $\mathcal{A}'$  is the corresponding available potential energy per mass. We only consider energetics of a plane internal wave. This idealized physical system has rather trivial energetics since for a plane wave there is no spatial convergence of the phase averaged energy flux, which means that the phase averaged energy is fixed in time. Even so, it is useful to work through the maths to gain practice in developing the phase averaged energy relations.

### Interpreting the available potential energy

We saw in Section 9.2.5 that  $\mathcal{A}' = (b'/N_R)^2 / 2$  measures the available potential energy per mass arising from the small amplitude buoyancy fluctuation. When those fluctuations are part of a plane wave we can introduce the fluid particle displacement as in equation (9.81), in which

$$\mathcal{A}' = (b'/N_R)^2 / 2 = (\delta z N_R)^2 / 2. \quad (9.97)$$

This expression is identical to the potential energy (??) that we encountered for a point mass connected to a spring, where here the spring constant per mass equals to  $N_R^2$ .

### Mechanical energy in the wave field

Assuming the small amplitude fluctuations are given by the wave relations (9.95a)-(9.95d) leads to

$$\mathcal{K}' = \frac{\tilde{\varphi}^2 |\mathbf{k}|^4}{2 N_R^2 k_z^2} \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (9.98a)$$

$$\mathcal{A}' = \frac{\tilde{\varphi}^2 |\mathbf{k}|^4}{2 N_R^2 k_z^2} \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (9.98b)$$

so that

$$\mathcal{K}' + \mathcal{A}' = \frac{\tilde{\varphi}^2 |\mathbf{k}|^4}{2 N_R^2 k_z^2} = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 |\mathbf{k}|^2}{2 \omega^2 k_z^2}. \quad (9.99)$$

It is notable that the wave's mechanical energy,  $\mathcal{K}' + \mathcal{A}'$ , is independent of the space and time position. Furthermore, the phase averaged kinetic energy equals to that of the available potential energy

$$\langle \mathcal{K}' \rangle = \langle \mathcal{A}' \rangle, \quad (9.100)$$

which manifests equipartition for the wave field.<sup>14</sup> We can relate the total mechanical energy to the phase average of the squared vertical velocity through use of equation (9.95d), which yields

$$\mathcal{K}' + \mathcal{A}' = \langle w' w' \rangle |\mathbf{k}|^2 / |\mathbf{k}_h|^2 = \langle w' w' \rangle / \cos^2 \gamma. \quad (9.101)$$

Note that the first expression in equation (9.99) indicates that for vertical waves, where  $\gamma = \pi/2$ , the sum  $\mathcal{K}' + \mathcal{A}'$  is non-singular since  $(w')^2$  vanishes. More precisely, we set  $\mathbf{k}_h = 0$  in equation (9.99) and find that a vertical phase velocity wave has mechanical energy

$$\mathcal{K}' + \mathcal{A}' = \frac{\tilde{\varphi}^2 k_z^2}{2 N_R^2} \quad \text{if } \mathbf{k}_h = 0. \quad (9.102)$$

### Energy flux and group velocity

The flux of mechanical energy can be written as

$$\mathbf{v}' \cdot \boldsymbol{\varphi}' = \frac{\tilde{\varphi}^2 \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t)}{k_z \omega} (k_z \mathbf{k}_h - \hat{\mathbf{z}} |\mathbf{k}_h|^2) = \frac{\mathbf{c}_g \tilde{\varphi}^2 |\mathbf{k}|^4 \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t)}{N_R^2 k_z^2}, \quad (9.103)$$

where we used equation (9.71) for the group velocity. Taking the phase average then leads to

$$\langle \mathbf{v}' \cdot \boldsymbol{\varphi}' \rangle = \frac{\mathbf{c}_g \tilde{\varphi}^2 |\mathbf{k}|^4}{2 N_R^2 k_z^2} = \mathbf{c}_g (\mathcal{K}' + \mathcal{A}'). \quad (9.104)$$

We thus confirm that the phase averaged flux of mechanical energy contained in an internal gravity wave equals to the group velocity times the total mechanical energy. This result accords with both the group velocity and the particle velocity being parallel to lines of constant phase and hence perpendicular to the wavevector

$$\mathbf{c}_g \cdot \mathbf{k} = \mathbf{v}' \cdot \mathbf{k} = 0. \quad (9.105)$$

---

<sup>14</sup>See Section ?? for the underlying reason for the equipartition of energy within linear waves, which is related to the virial theorem of classical mechanics.

### Energetics in terms of buoyancy amplitude

The above energetics made use of the polarization relations (9.95a)-(9.95e), written in terms of the pressure amplitude,  $\tilde{\varphi}$ . Here we briefly expose the results making use of equations (9.94a)-(9.94e), in which we write the fields in terms of the buoyancy amplitude,  $\tilde{b}$ . For this purpose we write for the kinetic energy

$$\mathcal{K}' = \frac{\mathbf{c}_g \cdot \mathbf{c}_g \tilde{b}^2 |\mathbf{k}|^4}{2 N_R^4 k_z^2} \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t) = \frac{\tilde{b}^2}{2 N_R^2} \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (9.106)$$

where we used equation (9.75) for  $\mathbf{c}_g \cdot \mathbf{c}_g$ . The corresponding available potential energy is

$$\mathcal{A}' = (b'/N_R)^2/2 = \frac{\tilde{b}^2}{2 N_R^2} \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (9.107)$$

so that the total mechanical energy is

$$\mathcal{K}' + \mathcal{A}' = \frac{\tilde{b}^2}{2 N_R^2}, \quad (9.108)$$

and the phase averaged energy flux is

$$\langle \mathbf{v}' \varphi' \rangle = \frac{\mathbf{c}_g \tilde{b}^2}{2 N_R^2} = \mathbf{c}_g (\mathcal{K}' + \mathcal{A}'). \quad (9.109)$$

## 9.6 Reflection of gravity waves

In Section 6.4.3 we studied the reflection of planetary Rossby wave packets from a smooth solid surface. That approach made use of some methods from geometric optics, in which we assume the waves reflect from the boundary without dissipation. Consequently, we only need to invoke the kinematic boundary condition to derive relations between incident and reflected waves. Here we pursue a similar study for internal gravity waves. Rossby waves (and many other waves such as acoustic waves and electromagnetic waves) exhibit **specular reflection**, whereby the angle the incident wave packet makes with the surface is preserved upon reflection. In contrast, the dispersion relation for internal gravity waves leads to a distinctly *non-specular* property. Namely, the angle the wave makes with the horizontal plane remains unchanged, rather than the angle the wave makes with the surface of reflection. Correspondingly, we encounter a particularly striking ability for the internal gravity wave, hitting the plane surface at a critical angle, to have an unbounded (infinite) wavenumber upon reflection.

### 9.6.1 Reflection conditions

Consider a packet of internal gravity waves with group velocity,  $\mathbf{c}_{gi}$ , that is incident on a plane solid boundary, with the boundary making an angle,  $\beta$ , with the horizontal (see Figure 9.5). Let the carrier wave in the wave packet have an angular frequency,  $\omega_i$ , and wavevector,  $\mathbf{k}_i$ , with  $\mathbf{c}_{gi} \cdot \mathbf{k}_i = 0$ . Since the velocity of fluid particles is parallel to the group velocity for internal waves, we write the velocity of fluid particles in the carrier wave as

$$\mathbf{v}'_i = \mathbf{c}_{gi} A_i \cos(\mathbf{k}_i \cdot \mathbf{x} - \omega_i t), \quad (9.110)$$

with the amplitude,

$$A_i = A_i(\mathbf{k}_i) \quad (9.111)$$

a shorthand for the amplitudes in either equation (9.94e) or (9.95e). As such,  $A_i$  is a function of the incident wavevector,  $\mathbf{k}_i$ , and the buoyancy frequency,  $N_R$ . The same considerations hold for the velocity of fluid particles in the reflected wave, so that

$$\mathbf{v}'_r = \mathbf{c}_{gr} A_r \cos(\mathbf{k}_r \cdot \mathbf{x} - \omega_r t), \quad (9.112)$$

with  $A_r = A_r(\mathbf{k}_r)$ ,  $\mathbf{k}_r$ , and  $\omega_r$ , the amplitude, wavevector, and angular frequency of the velocity of the reflected wave.

Now consider a steady state situation in which there are both incident and reflected waves, so that the fluid velocity at any given point in the fluid is given by the sum<sup>15</sup>

$$\mathbf{v}' = \mathbf{v}'_i + \mathbf{v}'_r. \quad (9.113)$$

The no-normal flow boundary condition at a solid boundary couples the incident and reflected wave properties, in which case

$$\mathbf{v}' \cdot \hat{\mathbf{n}} = (\mathbf{v}'_i + \mathbf{v}'_r) \cdot \hat{\mathbf{n}} = 0 \quad \text{at } \mathbf{x} = \mathbf{x}_w. \quad (9.114)$$

In these equations,  $\hat{\mathbf{n}}$  is the outward normal on the solid boundary and  $\mathbf{x}_w = |\mathbf{x}_w| \hat{\mathbf{t}}$  is a point on the boundary with  $\hat{\mathbf{t}}$  the unit vector pointing tangent to the boundary so that  $\hat{\mathbf{n}} \cdot \hat{\mathbf{t}} = 0$ .

As for Rossby wave reflection considered in Section 6.4.3, the boundary condition (9.114) leads to the reflection conditions for the angular frequency and wavevectors

$$\omega_i = \omega_r \quad \text{and} \quad (\mathbf{k}_i - \mathbf{k}_r) \cdot \hat{\mathbf{t}} = 0. \quad (9.115)$$

These conditions hold so long as the velocity amplitude and the group velocity satisfy

$$A_i \mathbf{c}_{gi} \cdot \hat{\mathbf{n}} = -A_r \mathbf{c}_{gr} \cdot \hat{\mathbf{n}}. \quad (9.116)$$

## 9.6.2 Specializing to internal gravity waves

The wavevector condition in equation (9.115) means that there is an equal projection onto the tangent direction of the incident and reflected wavevectors. This same condition holds for other waves, such as we found when studying Rossby wave reflection in Section 6.4.3 and as found for inertial waves in Exercise 9.4. Even so, the angular frequency condition,  $\omega_i = \omega_r$ , when coupled to the internal gravity wave dispersion relation (9.65), leads to non-specular wave reflection. The reason is that specification of the angular frequency and the background buoyancy frequency fixes the angle of the gravity wave relative to the horizontal plane. Hence, the angle that a reflected internal gravity wave makes relative to the solid boundary is generally different from the angle made by the incident wave. To prove this result requires trigonometry based on the boundary conditions (9.115), with details provided in Figure 9.5.

---

<sup>15</sup>Recall that steady state does not mean static. Here, we assume steady state in the sense that the incident and reflected waves are fully established, so that our concern is not with the initial value problem. Instead, we are concerned with the boundary value problem.

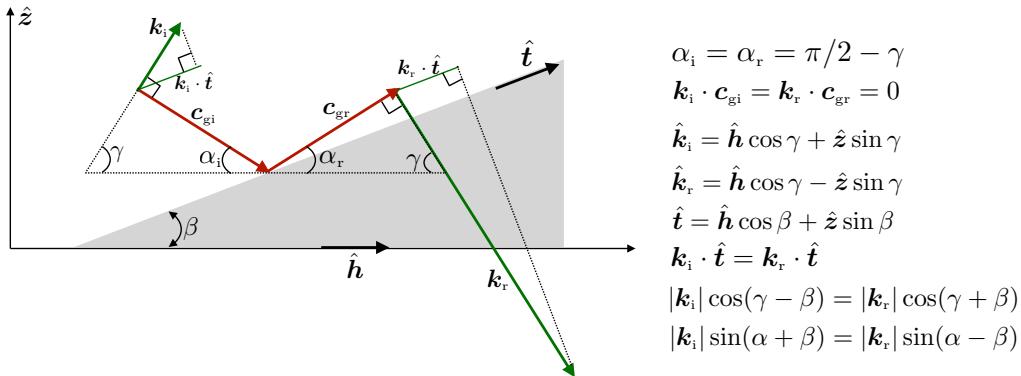


FIGURE 9.5: Depicting the reflection of an internal gravity wave packet as viewed within the geometrical optics approximation. Here the packet reflects from a planar inclined solid boundary that makes an angle,  $\beta$ , with the horizontal direction,  $\hat{h}$ . The downward incident wave packet has a group velocity,  $c_{gi}$ , and an upward carrier wavevector,  $\mathbf{k}_i$ , with  $\mathbf{c}_{gi} \cdot \mathbf{k}_i = 0$ . The incident group velocity makes an angle,  $\alpha_i = \pi/2 - \gamma$ , with the horizontal, whereas  $\mathbf{k}_i$  makes an angle,  $\gamma$ , with the horizontal. The upward reflected wave packet has group velocity,  $c_{gr}$ , and a downward carrier wavevector,  $\mathbf{k}_r$ , with  $\mathbf{c}_{gr} \cdot \mathbf{k}_r = 0$ . The angular frequency of the incident carrier wave is the same as the reflected wave,  $\omega_i = \omega_r = \omega = N_R \cos \gamma$ , which then means that  $\mathbf{k}_r$  makes an angle,  $\gamma$ , with the horizontal. Consequently, the reflected group velocity makes an angle  $\alpha_r = \pi/2 - \gamma$ , which is the same as the incident wave,  $\alpha_r = \alpha_i$ . The projection of the incident wavevector onto the surface tangent direction equals to that of the reflected wavevector,  $\mathbf{k}_i \cdot \hat{\mathbf{t}} = \mathbf{k}_r \cdot \hat{\mathbf{t}}$ , which means that  $|\mathbf{k}_i| \cos(\gamma - \beta) = |\mathbf{k}_r| \cos(\gamma + \beta)$ , or equivalently,  $|\mathbf{k}_i| \sin(\alpha + \beta) = |\mathbf{k}_r| \sin(\alpha - \beta)$ . For this example, the reflected wave has a larger wavenumber,  $|\mathbf{k}_r|$ , than the incident wave,  $|\mathbf{k}_i|$ , since  $\alpha - \beta$  is smaller than  $\alpha + \beta$ . Indeed, for a packet with incident group velocity angle equal to the boundary angle,  $\alpha = \beta$ , then the reflected wave has an infinite wavenumber, which is an indication that the linear theory breaks down. To the right of the schematic we list the basic relations between the incident and reflected waves.

### Non-specular nature of internal gravity wave reflection

For internal gravity waves, the frequency condition in equation (9.115) means that the angle the wavevector makes with the horizontal,  $\gamma$ , remains unchanged upon reflection

$$\omega_i = \omega_r = N_R \cos \gamma. \quad (9.117)$$

This relation holds no matter what angle,  $\beta$ , the solid wall makes with the horizontal. This identity is depicted in Figure 9.5, where the angle  $\alpha_i$ , which is the angle the incident packet makes with the horizontal, equals to the reflected angle,  $\alpha_r$ . It follows that the angle that the wave packet makes relative to the solid surface is different for the incident and reflected waves. This property is referred to as *non-specular*. For the example of Figure 9.5, the reflected wave is more nearly parallel to the solid boundary than the incident wave. Reversing the sense for the wave packet provides an example of a reflected packet that is less aligned with the boundary than the incident packet.

### Basic identities for internal gravity wave reflection

The frequency identity (9.117), the wavevector relation (9.115), along with some basic vector and trigonometric analysis, lead to the following identities

$$\hat{\mathbf{t}} = \hat{\mathbf{h}} \cos \beta + \hat{\mathbf{z}} \sin \beta \quad (9.118a)$$

$$\hat{\mathbf{k}}_i = \hat{\mathbf{h}} \cos \gamma + \hat{\mathbf{z}} \sin \gamma \quad (9.118b)$$

$$\hat{\mathbf{k}}_r = \hat{\mathbf{h}} \cos \gamma - \hat{\mathbf{z}} \sin \gamma \quad (9.118c)$$

$$\hat{\mathbf{k}}_i \cdot \hat{\mathbf{t}} = \cos \gamma \cos \beta + \sin \gamma \sin \beta = \cos(\gamma - \beta) = \sin(\alpha + \beta) \quad (9.118d)$$

$$\hat{\mathbf{k}}_r \cdot \hat{\mathbf{t}} = \cos \gamma \cos \beta - \sin \gamma \sin \beta = \cos(\gamma + \beta) = \sin(\alpha - \beta), \quad (9.118e)$$

where  $\hat{\mathbf{h}}$  is a unit vector in the horizontal. With these results we find that the wavevector boundary condition,  $\mathbf{k}_i \cdot \hat{\mathbf{t}} = \mathbf{k}_r \cdot \hat{\mathbf{t}}$ , (equation (9.115)) leads to the equivalent relations

$$|\mathbf{k}_i| \cos(\gamma - \beta) = |\mathbf{k}_r| \cos(\gamma + \beta) \quad (9.119a)$$

$$|\mathbf{k}_i| \sin(\alpha + \beta) = |\mathbf{k}_r| \sin(\alpha - \beta). \quad (9.119b)$$

### Critical reflection of internal gravity waves

The identity (9.119b) makes it clear that a most remarkable result holds when the incident packet hits the boundary at an angle that equals to the solid wall angle,  $\alpha = \beta$ . In this case the reflected wavenumber,  $|\mathbf{k}_r|$ , is unbounded, which means that the reflected waves have arbitrarily small wavelength. In a real fluid such small wavelength waves will eventually feel the impacts from viscous dissipation (e.g., Section ??), and/or they will break. In either case, such *critical reflection* of internal gravity waves provides an important mechanism for the transfer of energy from large to small scales, with the small scale features more prone to dissipative mixing.

### 9.6.3 Comments and further reading

Figure 9.5 provides an example where the reflected wave has a higher wavenumber than the incident wave. We can reverse all vectors to provide an example where a reflected wave has a smaller wavenumber than the incident wave. However, in a real fluid the reflections are not generally reversible since higher wavenumber waves are more subject to irreversible mixing, either through viscous dissipation or breaking. These, and other, wave-induced ocean mixing processes are reviewed in [MacKinnon et al. \(2013\)](#), [MacKinnon et al. \(2017\)](#), and [Buijsman et al. \(2019\)](#).

## 9.7 Linear Boussinesq ocean on an *f*-plane

We here establish some general properties of the linear Boussinesq ocean on an *f*-plane as described by equations (9.17a)-(9.17d)

$$\partial_t \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -\nabla_h \varphi' \quad \text{horizontal velocity equation} \quad (9.120a)$$

$$\partial_t w' = -\partial_z \varphi' + b' \quad \text{vertical velocity equation} \quad (9.120b)$$

$$\partial_t b' = -w' N_{\text{R}}^2 \quad \text{buoyancy equation} \quad (9.120c)$$

$$\nabla \cdot \mathbf{v}' = 0 \quad \text{continuity equation.} \quad (9.120d)$$

Since we are only concerned with *f*-plane motion, Rossby waves are not included in the physical system. As we see, the discussion reveals many forced oscillator equations satisfied by the linear fields, thus providing insight into the workings of the inertia-gravity waves discussed in Section 9.8. These oscillator equations also prove of use when developing the polarization relations for the plane inertia-gravity wave in Section 9.8.5.

### 9.7.1 Forced oscillator equation for horizontal velocity

Taking the time derivative of the horizontal velocity equation (9.120a) and back-substituting in the horizontal velocity tendency leads to

$$(\partial_{tt} + f^2) \mathbf{u}' = f \hat{\mathbf{z}} \times \nabla_h \varphi' - \partial_t (\nabla_h \varphi'). \quad (9.121)$$

This equation takes the form of a forced harmonic oscillator with natural angular frequency,  $f$ , and with forcing from the horizontal pressure gradient. Evidently, if we know the horizontal pressure gradient then that is sufficient to determine the horizontal velocity. Note that this is the same forced oscillator equation as (5.17) found for inertial oscillations in a homogeneous fluid.

### 9.7.2 Vertical component to the relative vorticity

Evolution of the vertical component to the relative vorticity is derived by taking the curl of the horizontal velocity equation (9.120a) and projecting onto the vertical, in which we find

$$\partial_t \zeta' = f \partial_z w'. \quad (9.122)$$

Hence, relative vorticity evolves when there is vertical stretching in the presence of planetary rotation. This property accords with the more general understanding of vortex mechanics studied in Section ???. It also reduces to the  $f = 0$  case, in which  $\zeta'$  is static since  $f = 0$  (Section 9.4.1). Finally, we can derive a forced oscillator equation for relative vorticity by taking the curl of the velocity equation (9.121) to find

$$(\partial_{tt} + f^2) \zeta' = f \nabla_h^2 \varphi'. \quad (9.123)$$

Evidently, the relative vorticity exhibits forced inertial oscillations, with the forcing proportional to the horizontal Laplacian of the pressure field as weighted by the Coriolis parameter.

### 9.7.3 Forced oscillator and free wave equations for vertical velocity

Taking the divergence of the horizontal velocity equation (9.120a), and then using the continuity equation (9.120d), leads to

$$\partial_t (\nabla_h \cdot \mathbf{u}') - f \zeta' = -\nabla_h^2 \varphi' \implies \partial_{tz} w' + f \zeta' = \nabla_h^2 \varphi'. \quad (9.124)$$

Now taking a time derivative and using the relative vorticity stretching equation (9.122) leads to the forced inertial oscillator equation for the vertical derivative of the vertical velocity

$$(\partial_{tt} + f^2) \partial_z w' = \partial_t \nabla_h^2 \varphi'. \quad (9.125)$$

This equation says that when the horizontal flow has a nonzero divergence in the presence of rotation, then it experiences a forced inertial oscillation. Notice that if the pressure is constant in the horizontal direction, then the inertial oscillations are unforced.

We can derive another forced oscillator equation, this one for  $w'$ . To do so, take the time derivative of the vertical velocity equation (9.120b) and then use the linearized buoyancy equation (9.120c), which yields the forced buoyancy oscillator equation

$$\partial_{tt} w' = -\partial_{tz} \varphi' + \partial_t b' \implies (\partial_{tt} + N_R^2) w' = -\partial_{tz} \varphi'. \quad (9.126)$$

We already encountered this equation in Section (9.3) when studying buoyancy oscillations. It says that vertical motion in the presence of a buoyancy field leads to forced buoyancy oscillations. Furthermore, if the pressure is constant in the vertical then the buoyancy oscillations are unforced.

Just like in the case of a non-rotating reference frame in Section 9.4.2, we can derive a free wave equation for  $w'$  by taking the vertical derivative of equation (9.125), the horizontal Laplacian of equation (9.126), and then adding to cancel the pressure contribution

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2 + f^2 \partial_{zz}) w' = 0. \quad (9.127)$$

This is the fundamental wave equation for internal inertia-gravity waves. The build-up to this equation offers signatures of the forces acting within these waves. Namely,  $\partial_z w'$  exhibits forced inertial oscillations with natural angular frequency  $f$  (equation (9.125)). Such oscillations are associated, through continuity, with horizontally diverging flow that feels the Coriolis acceleration. Likewise,  $w'$  exhibits forced buoyancy oscillations with natural angular frequency  $N_R$  (equation (9.126)). These oscillations are associated with vertical motion that feels the effects from the background vertical stratification and associated buoyancy. By eliminating the pressure forcing the two complementary modes of oscillation, we reveal a free wave equation for  $w'$ . The free wave equation exposes the roles for Coriolis and buoyancy accelerations, yet hides the intermediate role of pressure accelerations that force the oscillations.

#### 9.7.4 Forced oscillator equation for buoyancy

Since the flow is non-divergent at each time instant, we know that

$$\nabla \cdot \partial_t \mathbf{v}' = \partial_t \nabla \cdot \mathbf{v}' = 0. \quad (9.128)$$

Hence, taking the divergence of the horizontal velocity equation (9.120a) and adding to the vertical derivative of the vertical velocity equation (9.120b) leads to the pressure equation

$$-\nabla^2 \varphi' = -(f \zeta' + \partial_z b'). \quad (9.129)$$

Evidently, a source for the dynamic pressure arises from vertical buoyancy gradients along with relative vorticity. We characterized the physics of these sources in Section ?? when studying the pressure equation in a Boussinesq ocean.

Taking the time derivative of the pressure equation (9.129) and using the stretching relative vorticity equation (9.122) leads to

$$f^2 \partial_z w' = \partial_t (\nabla^2 \varphi' - \partial_z b'). \quad (9.130)$$

One more time derivative and use of the vertical velocity equation (9.120b) yields the forced inertial oscillator equation for the vertical derivative of buoyancy

$$(\partial_{tt} + f^2) \partial_z b' = (\partial_{tt} \nabla^2 + f^2 \partial_{zz}) \varphi'. \quad (9.131)$$

### 9.7.5 An equation for pressure

We now consider an equation for pressure that is based on taking the time derivative of the forced inertial oscillator equation (9.131) for  $\partial_z b'$ . Focusing on the left hand side we find

$$\partial_t[(\partial_{tt} + f^2) \partial_z b'] = \partial_z[(\partial_{tt} + f^2) \partial_t b'] \quad \text{swap } \partial_z \text{ and } \partial_t \quad (9.132a)$$

$$= -\partial_z[(\partial_{tt} + f^2) w' N_R^2] \quad \text{equation (9.120c)} \quad (9.132b)$$

$$= -\partial_z N_R^2 (\partial_{tt} + f^2) w' - N_R^2 (\partial_{tt} + f^2) \partial_z w' \quad \text{product rule} \quad (9.132c)$$

$$= -\partial_z N_R^2 (\partial_{tt} + f^2) w' - N_R^2 \partial_t \nabla_h^2 \varphi' \quad \text{equation (9.125)} \quad (9.132d)$$

Combining with the time derivative on the right hand side of equation (9.131) leads to the pressure equation

$$\partial_t[(\partial_{tt} \nabla^2 + f^2 \partial_{zz} + N_R^2 \nabla_h^2) \varphi'] = -\partial_z N_R^2 (\partial_{tt} + f^2) w'. \quad (9.133)$$

In the special case of a constant background stratification we find

$$\partial_t[(\partial_{tt} \nabla^2 + f^2 \partial_{zz} + N_R^2 \nabla_h^2) \varphi'] = 0 \quad \text{if } \partial_z N_R = 0, \quad (9.134)$$

which is nearly the same as the free wave equation (9.127) for  $w'$ .<sup>16</sup>

### 9.7.6 Potential vorticity

Although inertia-gravity waves modify relative vorticity, we here show they retain the linearized potential vorticity unchanged. For this purpose, use the relative vorticity equation (9.122) and the buoyancy equation (9.120c) to find<sup>17</sup>

$$\partial_t[\zeta' + f \partial_z(b'/N_R^2)] = 0 \implies \partial_t Q' = 0 \quad \text{with} \quad Q' = \zeta' + f \partial_z(b'/N_R^2). \quad (9.135)$$

Hence, an inertia-gravity wave does not alter the potential vorticity, at least to linear order. This means that any vortical portion of the flow (i.e., flow with  $Q' \neq 0$ ), such as a background static geostrophic flow, can exist without either the waves or the background flow affecting one another. Furthermore, we observe that the linearized potential vorticity (9.135) accords with the relative vorticity and buoyancy contribution to the quasi-geostrophic potential vorticity (??). In the present case, however, we are concerned with the  $f$ -plane rather than the  $\beta$ -plane considered for quasi-geostrophy. Another difference is that the potential vorticity (9.135) is locally static,  $\partial_t Q' = 0$ , whereas the quasi-geostrophic potential vorticity remains constant when following the geostrophic flow.

#### Potential vorticity and layer thickness

For yet another way to interpret the linear potential vorticity (9.135), write the buoyancy fluctuation in terms of the particle displacement as in equation (9.81)

$$\delta z = -b'/N_R^2, \quad (9.136)$$

<sup>16</sup>See Lecture 11 in [Pedlosky \(2003\)](#) for connection between the pressure equation (9.134) to the potential vorticity equation in the case of constant  $N_R$ .

<sup>17</sup>Note that in this chapter we define potential vorticity with dimensions of inverse time,  $T^{-1}$ , as for the continuously stratified quasi-geostrophic potential vorticity in Section ???. In contrast, in our study of shallow water waves in Chapter 7, we defined  $Q'$  in Section 7.3.3 to have dimensions of  $L^{-1} T^{-1}$ .

which brings the linearized potential vorticity to the form

$$Q' = \zeta' + f \partial_z(b'/N_R^2) = \zeta' - f \partial_z(\delta z). \quad (9.137)$$

Since  $Q'$  is a static field, the presence of  $\partial_z(\delta z)$  leads in a rotating reference frame to a compensating relative vorticity. In an internal gravity wave,  $\delta z$  measures the periodic compression and expansion of the vertical distance between constant buoyancy surfaces. This layer interpretation of potential vorticity accords with our understanding of potential vorticity from shallow water theory in Section ??, as well as potential vorticity in isopycnal/buoyancy coordinate models in Section ??.

### Connection to pressure and buoyancy

The various harmonic oscillator equations established thus far in this section are not satisfied by  $Q'$  since it is a static field. However, we can derive an expression for  $Q'$  in terms of pressure and buoyancy through the following manipulations

$$(\partial_{tt} + f^2) Q' = f^2 Q' \quad (9.138a)$$

$$= (\partial_{tt} + f^2)[\zeta' + f \partial_z(b'/N_R^2)] \quad (9.138b)$$

$$= f \nabla_h^2 \varphi' + N_R^{-2} (\partial_{tt} \nabla^2 + f^2 \partial_{zz}) \varphi' + \partial_z(N_R^{-2}) (\partial_{tt} + f^2) b', \quad (9.138c)$$

where equation (9.138a) follows since  $\partial_t Q' = 0$ , equation (9.138b) used the definition of  $Q'$  from equation (9.135), and equation (9.138c) used the forced harmonic oscillator equations (9.123) and (9.125). Multiplying by  $N_R^2$  leads to

$$f N_R^2 Q' = (\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2 + f^2 \partial_{zz}) \varphi' + N_R^2 \partial_z(N_R^{-2}) (\partial_{tt} + f^2) b'. \quad (9.139)$$

### Decomposing vortical and divergent motions for constant vertical stratification

In the special case of a constant  $N_R$ , we see that the potential vorticity can be written entirely in terms of the pressure

$$f N_R^2 Q' = (\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2 + f^2 \partial_{zz}) \varphi' \quad \text{if } \partial_z N_R = 0. \quad (9.140)$$

Evidently, for constant background stratification the static nature of the potential vorticity is equivalent to the wave equation (9.134)

$$\partial_t Q' = 0 \implies \partial_t[(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2 + f^2 \partial_{zz}) \varphi'] = 0 \quad \text{if } \partial_z N_R = 0. \quad (9.141)$$

We found the same connection between the potential vorticity equation and the wave equation in Section 7.8.2 when studying shallow water inertia-gravity waves (see equation (7.153)). As for the shallow water, we conclude that for the case of constant stratification the vortical motions associated with *f*-plane geostrophy (hence with zero vertical velocity) are decoupled from the divergent motions associated with inertia-gravity waves. Hence, we can describe the fluid motion by a static potential vorticity whose non-zero value is set by *f*-plane geostrophic (vortical) flow, plus a zero potential vorticity flow arising from inertia-gravity waves. In the linear theory, and for constant vertical stratification, there is no exchange of potential vorticity between the vortical flow and divergent waves. As a practical matter, one often has knowledge of the static potential vorticity arising from the geostrophic motion. We can obtain the associated geostrophic pressure field from inverting the elliptic operator in equation (9.140).

## 9.8 Free inertia-gravity waves

In this section we study the free *inertia-gravity* waves that arise under small amplitude fluctuations in a continuously stratified fluid on an  $f$ -plane. To enable plane wave solutions we assume the stratification is constant.

### 9.8.1 Dispersion relation

Returning to the wave equation (9.127) for  $w'$

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2 + f^2 \partial_{zz}) w' = 0, \quad (9.142)$$

we take the plane wave ansatz

$$w' = \tilde{w} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} \quad (9.143)$$

which readily leads to the dispersion relation

$$\varpi^2(\mathbf{k}) = \frac{N_R^2 |\mathbf{k}_h|^2 + f^2 k_z^2}{|\mathbf{k}|^2}. \quad (9.144)$$

Just like the case of internal gravity waves studied in Section 9.5, the dispersion relation is only a function of the wave direction. To manifest this property we introduce the angle,  $\gamma$ , between the wavevector and the horizontal plane

$$\varpi^2(\mathbf{k}) = N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma = (N_R^2 - f^2) \cos^2 \gamma + f^2 = N_R^2 - (N_R^2 - f^2) \sin^2 \gamma. \quad (9.145)$$

The dispersion relation reduces to that for internal gravity waves when  $f = 0$  given as by equation (9.65).

### 9.8.2 Near-inertial waves

In stably stratified geophysical fluids we typically find the squared buoyancy frequency is larger than the squared Coriolis frequency,  $N_R^2 > f^2$ . In this case the plane inertia-gravity wave angular frequency is bounded by

$$f^2 \leq \omega^2 \leq N_R^2 \quad \text{when } N_R^2 > f^2. \quad (9.146)$$

Indeed, in many cases we have  $f^2 \ll N_R^2$ . Even so, there are weakly stratified regions with  $N_R^2 < f^2$ , in which case the opposite frequency range holds whereby  $N_R^2 \leq \omega^2 \leq f^2$ . In the limit with  $N_R = 0$ , then the inertia-gravity waves reduce to the *inertial waves* studied in Chapter 5, which are waves appearing for motion on the  $f$ -plane within a fluid with homogeneous density ( $N_R = 0$ ).

Inertia-gravity waves typically have horizontal scales much larger than vertical, in which case  $|\mathbf{k}_h|^2 \ll k_z^2$ . As discussed in Exercise 9.6, such waves satisfy hydrostatic scaling in which

$$\frac{\text{vertical scales in the waves}}{\text{horizontal scales in the waves}} = \frac{|\mathbf{k}_h|}{|k_z|} \ll 1 \iff \text{hydrostatic motion.} \quad (9.147)$$

The reason for this anisotropy in the waves is related to the processes forcing the waves. In particular, for the ocean the forcing by winds occurs with large horizontal patterns characteristic of the atmosphere. With  $|\mathbf{k}_h|^2 \ll k_z^2$ , the phase velocity for the waves is nearly vertical so that

$\gamma \approx \pi/2$ . We thus find that the dominant forcing frequency for inertia-gravity waves is near that of the Coriolis frequency

$$\omega^2 = f^2 + (N_R^2 - f^2) \cos^2 \gamma \gtrsim f^2. \quad (9.148)$$

Such waves are referred to as *near-inertial waves*.

### 9.8.3 Force balance in an inertia-gravity wave

In Section 9.5.11 we studied the buoyancy forces acting on a fluid particle that moves transverse to the constant phase surfaces in a gravity wave. We pursue the same analysis here with the added feature of the Coriolis acceleration. This analysis offers a force balance interpretation of the dispersion relation (9.145).

For a phase surface that has a nonzero slope in the vertical, the displaced fluid particle moving parallel to the phase surface feels the effects from buoyancy forces. Likewise, when including the Coriolis acceleration on an  $f$ -plane, a horizontal displacement of a fluid particle induces a Coriolis acceleration. Given the nature of the Coriolis acceleration, the particle turns in the horizontal. Hence, a fluid particle moving in the direction of constant phase surfaces feels the effects of the buoyant acceleration when it moves vertically plus the Coriolis acceleration when it moves horizontally. For small amplitude motion, these forces give rise to simple harmonic oscillator motion, and they represent the essential features of an inertia-gravity wave. We here detail the mechanics of this motion.

#### Particle displacements and Coriolis acceleration

Following the approach in Section 9.5.11, we make use of the fluid particle displacement field,  $\xi(\mathbf{x}, t)$ , whose time derivative yields the velocity of the fluid particle in a small amplitude wave

$$\partial_t \xi(\mathbf{x}, t) = \mathbf{v}'(\mathbf{x}, t). \quad (9.149)$$

To understand the role of the Coriolis acceleration, consider an inertial particle that satisfies the equation of motion

$$\partial_t \mathbf{v}' + f \hat{\mathbf{z}} \times \mathbf{v}' = 0 \implies \partial_t (\partial_t \xi + f \hat{\mathbf{z}} \times \xi) = 0. \quad (9.150)$$

Hence, in terms of particle displacements, the inertial particle satisfies

$$\partial_t \xi + f \hat{\mathbf{z}} \times \xi = \text{constant}, \quad (9.151)$$

with the constant determined by initial conditions, which can be set to zero without loss of generality. Evidently, an inertial particle in a rotating reference frame, displaced in a direction perpendicular to the rotation axis (horizontal direction here), induces a time tendency for motion in the orthogonal horizontal direction

$$\partial_t \xi = -f \hat{\mathbf{z}} \times \xi. \quad (9.152)$$

#### Particle displacements with buoyancy and Coriolis

The linearized equation of motion that includes the effects from pressure, Coriolis, and buoyancy, is given by

$$\partial_t \mathbf{v}' + f \hat{\mathbf{z}} \times \mathbf{v}' = -\nabla \varphi' + b' \hat{\mathbf{z}}. \quad (9.153)$$

Focusing on a fluid particle in the presence of a transverse wave, we project the equation of motion onto the transverse direction,  $\hat{t}$ , with this direction static for a plane wave. Doing so eliminates the pressure gradient since it is constant along a surface of constant phase, so that fluid particle motion in the transverse direction satisfies

$$\partial_t(\hat{t} \cdot \mathbf{v}') + f \hat{t} \cdot (\hat{z} \times \mathbf{v}') = b' \hat{t} \cdot \hat{z}. \quad (9.154)$$

Introducing the particle displacement yields

$$\hat{t} \cdot \mathbf{v}' = \partial_t(\hat{t} \cdot \boldsymbol{\xi}) = \partial_t \delta l \quad (9.155a)$$

$$b' \hat{t} \cdot \hat{z} = -N_R^2 \delta z \cos \gamma = -N_R^2 \delta l \cos^2 \gamma, \quad (9.155b)$$

with the buoyancy expression in equation (9.155b) following from equation (9.81) derived when studying internal gravity waves ( $f = 0$ ). For the Coriolis acceleration we make use of equation (9.152) to write

$$f \hat{t} \cdot (\hat{z} \times \mathbf{v}') = f \mathbf{v}' \cdot (\hat{t} \times \hat{z}) \quad (9.156a)$$

$$= f \partial_t \boldsymbol{\xi} \cdot (\hat{t} \times \hat{z}) \quad (9.156b)$$

$$= -f^2 (\hat{z} \times \boldsymbol{\xi}) \cdot (\hat{t} \times \hat{z}) \quad (9.156c)$$

$$= f^2 \delta l \sin^2 \gamma, \quad (9.156d)$$

where we used the vector identity (??) and set  $\hat{z} \cdot \hat{t} = \sin \gamma$ .

Bringing the above results together leads to the equation of motion for particle displacements

$$(\partial_{tt} + N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma) \delta l = 0. \quad (9.157)$$

This is an equation for a simple harmonic oscillator with angular frequency

$$\omega^2 = N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma, \quad (9.158)$$

which is the dispersion relation derived in Section 9.8.1 through use of the plane wave ansatz.

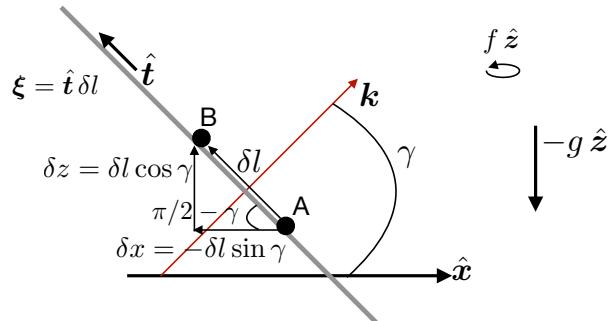


FIGURE 9.6: An extension of Figure 9.3 to now allow for a fluid particle to feel accelerations from both buoyancy and Coriolis as it moves in the direction transverse to the plane waves. The wavevector makes an angle,  $\gamma$ , with the horizontal. We show a view in the  $x$ - $z$  plane with the particle displacement vector,  $\boldsymbol{\xi}(\mathbf{x}, t) = \delta l \hat{t}$ . The local transverse unit vector,  $\hat{t}$ , points in a direction that is orthogonal to the wavevector,  $\mathbf{k}$ . The displacement of the fluid particle from point  $A$  and point  $B$  occurs over a distance,  $\delta l$ , with a corresponding change in vertical position given by  $\delta z = \delta l \cos \gamma$  and horizontal position changes by  $\delta x = -\delta l \sin \gamma$ . This displacement causes the fluid particle to experience accelerations from both buoyancy and Coriolis.

### 9.8.4 Group velocity

The wavevector gradient of the dispersion relation (9.158) yields the group velocity for inertia-gravity waves

$$\mathbf{c}_g = \frac{(N_R^2 - f^2) k_z}{|\mathbf{k}|^4 \omega} (k_z \mathbf{k}_h - |\mathbf{k}_h|^2 \hat{\mathbf{z}}). \quad (9.159)$$

As for internal gravity waves in equation (9.71), we find that the group velocity is perpendicular to the phase velocity ( $\mathbf{c}_p = \hat{\mathbf{k}} \omega / |\mathbf{k}|$ )

$$\mathbf{c}_g \cdot \mathbf{c}_p = 0. \quad (9.160)$$

We also find that

$$(\mathbf{c}_g \cdot \hat{\mathbf{z}}) (\mathbf{c}_p \cdot \hat{\mathbf{z}}) = -\frac{(N_R^2 - f^2) k_z^2 |\mathbf{k}_h|^2}{|\mathbf{k}|^4 \omega}, \quad (9.161)$$

so that

$$(\mathbf{c}_g \cdot \hat{\mathbf{z}}) (\mathbf{c}_p \cdot \hat{\mathbf{z}}) < 0 \quad \text{if } f^2 < N_R^2 \quad (9.162a)$$

$$(\mathbf{c}_g \cdot \hat{\mathbf{z}}) (\mathbf{c}_p \cdot \hat{\mathbf{z}}) > 0 \quad \text{if } f^2 > N_R^2. \quad (9.162b)$$

The usual case for the atmosphere and ocean finds  $f^2 < N_R^2$ , so that if the group velocity is upward then the phase velocity is downward, and vice versa. However, if  $f^2 > N_R^2$ , as for inertial waves where  $N_R = 0$  (Chapter 5, Exercise 9.4), or more generally for inertia-gravity waves in very weak vertical stratification, then the group and phase velocities have the same orientation with respect to the vertical.

### 9.8.5 Polarization relations for a plane wave

We now make use of the plane wave ansatz (9.54) to derive the polarization relations that determine the structure of a plane wave. For this purpose we make use of the various forced harmonic oscillator equations from Section 9.7 to express the velocity and buoyancy in terms of a real pressure amplitude,  $\tilde{\varphi}$ , where pressure is assumed to take the form

$$\varphi' = \tilde{\varphi} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t). \quad (9.163)$$

We also make use of the dispersion relation (9.145) to write the polarization relations in a variety of forms. In the limit with  $f = 0$ , each of the expressions below reduce to the gravity wave polarization relations (9.90a)-(9.90b) derived in Section 9.5.13.

#### Buoyancy

Buoyancy satisfies the forced oscillator equation (9.131), which for a plane wave takes the form

$$(-\omega^2 + f^2) i k_z \tilde{b} = (\omega^2 |\mathbf{k}|^2 - f^2 k_z^2) \tilde{\varphi}. \quad (9.164)$$

Various forms of the dispersion relation given in Section 9.8.1 lead to the identities

$$\omega^2 |\mathbf{k}|^2 - f^2 k_z^2 = N_R^2 |\mathbf{k}_h|^2 \quad (9.165a)$$

$$\omega^2 - f^2 = (N_R^2 - f^2) |\mathbf{k}_h|^2 / |\mathbf{k}|^2 = (N_R^2 - \omega^2) |\mathbf{k}_h|^2 / k_z^2 \quad (9.165b)$$

so that the ratio of amplitudes can be written in the equivalent manners

$$\tilde{b}/\tilde{\varphi} = \frac{i(\omega^2 |\mathbf{k}|^2 - f^2 k_z^2)}{k_z(\omega^2 - f^2)} = \frac{i N_R^2 |\mathbf{k}_h|^2}{k_z(\omega^2 - f^2)} = \frac{i N_R^2 |\mathbf{k}|^2}{k_z(N_R^2 - f^2)} = \frac{i N_R^2 k_z}{N_R^2 - \omega^2}, \quad (9.166)$$

which means that the real buoyancy wave is given by

$$b'/\tilde{\varphi} = -\frac{N_R^2 |\mathbf{k}_h|^2 \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}{k_z(\omega^2 - f^2)} = -\frac{N_R^2 k_z \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}{N_R^2 - \omega^2}. \quad (9.167)$$

### Vertical velocity component

From Section 9.7.3 we know that the vertical velocity satisfies the two forced oscillator equations (9.125) and (9.126)

$$(\partial_{tt} + f^2) \partial_z w' = \partial_t \nabla_h^2 \varphi' \quad \text{and} \quad (\partial_{tt} + N_R^2) w' = -\partial_{tz} \varphi'. \quad (9.168)$$

For a plane wave these equations lead to

$$\tilde{w}/\tilde{\varphi} = -\frac{\omega |\mathbf{k}_h|^2}{k_z(\omega^2 - f^2)} = -\frac{\omega k_z}{N_R^2 - \omega^2} = -\frac{\omega |\mathbf{k}|^2}{k_z(N_R^2 - f^2)}, \quad (9.169)$$

so that the real wave takes on the form

$$w'/\tilde{\varphi} = -\frac{\omega |\mathbf{k}_h|^2 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}{k_z(\omega^2 - f^2)} = -\frac{\omega k_z \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}{N_R^2 - \omega^2} = -\frac{\omega |\mathbf{k}|^2 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t)}{k_z(N_R^2 - f^2)}. \quad (9.170)$$

### Horizontal velocity

Equation (9.121) for the horizontal velocity leads to the relation for the plane wave amplitude

$$\tilde{\mathbf{u}}/\tilde{\varphi} = \frac{\omega \mathbf{k}_h - i f (\hat{\mathbf{z}} \times \mathbf{k}_h)}{\omega^2 - f^2}, \quad (9.171)$$

so that the real plane wave polarization relation is given by

$$\mathbf{u}/\tilde{\varphi} = \frac{\omega \mathbf{k}_h \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) + f (\hat{\mathbf{z}} \times \mathbf{k}_h) \sin(\mathbf{k} \cdot \mathbf{x} - \omega t)}{\omega^2 - f^2}, \quad (9.172)$$

which compares directly to the shallow water polarization relation (7.174a).

## 9.8.6 Energetics of a plane inertia-gravity wave

We here extend the energetic analysis of plane internal gravity waves in Section 9.5.14 to the case of plane inertia-gravity waves, making use of the polarization relations from Section 9.8.5. Each of the expressions found here reduce to the internal gravity waves case when taking  $f = 0$ .

### Instantaneous energetics

The plane inertia-gravity wave has squared velocity components

$$\mathbf{u}' \cdot \mathbf{u}'/\tilde{\varphi}^2 = \frac{\omega^2 |\mathbf{k}_h|^2 \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) + f^2 |\mathbf{k}_h|^2 \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t)}{(\omega^2 - f^2)^2} \quad (9.173a)$$

$$(w')^2/\tilde{\varphi}^2 = \frac{\omega^2 |\mathbf{k}_h|^4 \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t)}{k_z^2 (\omega^2 - f^2)^2}, \quad (9.173b)$$

so that the wave's kinetic energy is

$$\mathcal{K}' = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2}{2(\omega^2 - f^2)^2} [\omega^2 (1 + |\mathbf{k}_h|^2/k_z^2) \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t) + f^2 \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t)], \quad (9.174)$$

which can be written in the more tidy form through use of the dispersion relation (9.145)

$$\mathcal{K}' = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 [f^2 + (N_R^2 \cot^2 \gamma) \cos^2(\mathbf{k} \cdot \mathbf{x} - \omega t)]}{2(\omega^2 - f^2)^2}, \quad (9.175)$$

where we set

$$\cot \gamma = |\mathbf{k}_h|/k_z. \quad (9.176)$$

Likewise, the available potential energy is given by

$$\mathcal{A}' = \frac{\tilde{\varphi}^2 (|\mathbf{k}_h|^2 N_R^2 \cot^2 \gamma) \sin^2(\mathbf{k} \cdot \mathbf{x} - \omega t)}{2(\omega^2 - f^2)^2}. \quad (9.177)$$

Taking the sum leads to the total mechanical energy in a plane inertia-gravity wave

$$\mathcal{K}' + \mathcal{A}' = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 (f^2 k_z^2 + N_R^2 |\mathbf{k}_h|^2)}{2 k_z^2 (\omega^2 - f^2)^2} \quad (9.178)$$

Time independence of mechanical energy for the plane inertia-gravity wave accords with the result (9.99) for the plane internal gravity wave.

### Phase averaged energetics

Taking the phase averages on the kinetic energy and available potential energy leads to

$$\langle \mathcal{K}' \rangle = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 (\omega^2 |\mathbf{k}|^2 + f^2 k_z^2)}{4 k_z^2 (\omega^2 - f^2)^2} \quad \text{and} \quad \langle \mathcal{A}' \rangle = \frac{\tilde{\varphi}^2 N_R^2 |\mathbf{k}_h|^4}{4 k_z^2 (\omega^2 - f^2)^2}, \quad (9.179)$$

whose sum is

$$\langle \mathcal{K}' + \mathcal{A}' \rangle = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 (\omega^2 |\mathbf{k}|^2 + f^2 k_z^2 + N_R^2 |\mathbf{k}_h|^2)}{4 k_z^2 (\omega^2 - f^2)^2} = \frac{\tilde{\varphi}^2 |\mathbf{k}_h|^2 (f^2 k_z^2 + N_R^2 |\mathbf{k}_h|^2)}{2 k_z^2 (\omega^2 - f^2)^2} \quad (9.180)$$

and ratio is

$$\frac{\langle \mathcal{K}' \rangle}{\langle \mathcal{A}' \rangle} = 1 + 2(f/N_R)^2 \tan^2 \gamma = \frac{\omega^2 |\mathbf{k}|^2 + f^2 k_z^2}{N_R^2 |\mathbf{k}_h|^2}. \quad (9.181)$$

It is notable that the ratio is bounded below by unity, so that the kinetic energy is never less than the available potential energy. The ratio is unity when the phase velocity is horizontal ( $\gamma = 0$ ), in which fluid particles have vertical trajectories and thus exhibit purely vertical buoyancy oscillations. This equipartition of kinetic and available potential energies (again, holding with  $\gamma = 0$ ) was found for the internal gravity wave in Section 9.5.14. When the phase velocity is vertical, so that  $\gamma = \pi/2$ , the plane waves have no available potential energy since the fluid particles are exhibiting horizontal inertial oscillations and do not sample the background buoyancy field.

### Phase averaged mechanical energy flux

The phase averaged mechanical energy flux is given by

$$\frac{\langle \mathbf{v}' \varphi' \rangle}{\tilde{\varphi}^2} = \frac{\omega (k_z \mathbf{k}_h - \hat{\mathbf{z}} |\mathbf{k}_h|^2)}{2 k_z (\omega^2 - f^2)} \quad (9.182a)$$

$$= \frac{\omega^2 |\mathbf{k}|^4 c_g}{2 k_z^2 (N_R^2 - f^2) (\omega^2 - f^2)} \quad (9.182b)$$

$$= \frac{c_g |\mathbf{k}_h|^2 (f^2 k_z^2 + N_R^2 |\mathbf{k}_h|^2)}{2 k_z^2 (\omega^2 - f^2)^2}, \quad (9.182c)$$

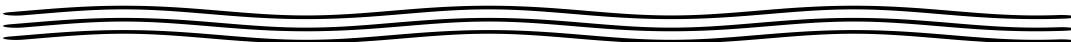
where we made use of the inertia-gravity wave group velocity (9.159). Now the phase averaged mechanical energy is given by equation (9.180), so that the plane inertia-gravity waves exhibit the group velocity property also found for internal gravity waves in equation (9.104)

$$\langle \mathbf{v}' \varphi' \rangle = c_g (\mathcal{K}' + \mathcal{A}'). \quad (9.183)$$

That is, a plane inertia-gravity wave has a phase averaged mechanical energy flux equal to the group velocity times the phase averaged mechanical energy.

#### 9.8.7 Comments

The inertial waves studied in Chapter 5 are the  $N_R = 0$  limit of inertia-gravity waves studied in the present section. It is notable that in the presence of any nonzero vertical stratification, the squared angular frequency of propagating inertia-waves is super-inertial,  $\omega^2 \geq f^2$ , whereas when  $N_R = 0$  the waves are sub-inertial,  $\omega^2 \leq f^2$ . If an inertial wave from an unstratified region, with  $\omega^2 \leq f^2$ , encounters stratification, then this low frequency wave cannot propagate into the stratified region.



## 9.9 Exercises

### EXERCISE 9.1: INTERNAL GRAVITY WAVE EQUATION FOR ALL PROGNOSTIC FIELDS

In Section 9.4.2 we derived the wave equation (9.40) for the vertical velocity

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) w' = 0. \quad (9.184)$$

- (a) The wave equation (9.184) holds for a background buoyancy frequency that is a function of depth,  $N_R = N_R(z)$ . Prove that the buoyancy satisfies the same equation

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) b' = 0. \quad (9.185)$$

- (b) Show that for the linear internal gravity wave fluctuations that pressure also satisfies the wave equation, only now for the case of  $N_R$  a constant

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) \varphi' = 0, \quad (9.186)$$

- (c) Again assume that  $N_R$  is a spatial constant, and show that for the linear internal gravity

wave fluctuations that the horizontal velocity satisfies the wave equation

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) \mathbf{u}' = 0. \quad (9.187)$$

**EXERCISE 9.2: TWO-DIMENSIONAL INTERNAL GRAVITY WAVES IN A BOUNDED DOMAIN**

In this exercise we consider internal gravity waves in a bounded domain. Rather than free waves, these waves must satisfy boundary conditions that leads to standing waves that oscillate in place. That is, the waves no longer propagate in bounded directions. We already studied similar standing wave patterns for acoustic waves in Exercise 3.1, and for surface gravity waves in Section 4.8, with the mathematics of the present problem very similar to those other systems.

- (a) Consider two-dimensional non-divergent flow in the  $y$ - $z$  plane confined to a rectangular box with  $0 \leq y \leq L$  the meridional extent of the box and  $0 \leq z \leq H$  the vertical extent. Assume a constant reference state buoyancy frequency,  $N_R$ . Find the discrete normal modes for the streamfunction,  $\psi$ , where  $\mathbf{v} = \hat{\mathbf{y}} v + \hat{\mathbf{z}} w = \hat{\mathbf{x}} \times \nabla \psi$ . Hint: make use of Exercise 9.1 to determine the wave equation satisfied by the streamfunction. Also recall that the streamfunction is a constant along a material boundary (see Section ??), which is here the rectangular boundary of the  $y$ - $z$  domain. We can set the streamfunction to zero around the rectangular boundary without loss of generality.
- (b) Assume the initial condition for the streamfunction,  $\psi(y, z, t = 0) = \Psi_0(y, z)$ . Use the methods of Fourier series from Section ?? to write the general expression for the streamfunction as an infinite Fourier series. Write the expression for the modal amplitude in terms of the initial condition as projected onto the discrete normal modes.
- (c) Consider the same two-dimensional non-divergent flow in the  $y$ - $z$  plane, but now consider an infinite zonal channel yet with zero zonal fluid flow. What extra wave degree of freedom is allowed by opening up the zonal direction? Write a mathematical expression that embodies this extra degree of freedom.

**EXERCISE 9.3: HORIZONTAL VELOCITY ROTATION WITH VERTICALLY PROPAGATING INERTIA–GRAVITY WAVES**

Consider a northern hemisphere internal gravity wave packet with  $N_R^2 > f^2$  and assume the wave packet has a downward group velocity,  $\mathbf{c}_g \cdot \hat{\mathbf{z}} < 0$ .

- (a) What is the sign of the vertical wavenumber,  $k_z$ ?
- (b) When looking down from above at a fixed point in space, what rotational sense (clockwise or counter-clockwise) does the horizontal velocity vector make as time progresses?
- (c) When looking down from above, what rotational sense (clockwise or counter-clockwise) does the horizontal velocity vector make at a fixed time as one moves upward, thereby increasing  $z$ ?

Hint: consider the case of shallow water inertia-gravity waves studied in Section 7.8.9.

**EXERCISE 9.4: REFLECTION CONDITIONS FOR INERTIAL WAVES**

Inertial waves studied in Chapter 5 are the  $N^2 = 0$  limit of inertia-gravity waves from Section 9.8. Emulating the analysis in Section 9.6.2 for internal gravity waves, determine the reflection conditions for packets of plane inertial waves. Figure 9.5 is drawn for internal gravity waves. Redraw this figure for inertial waves.

**EXERCISE 9.5: ENERGY FLUX IN TERMS OF VERTICAL VELOCITY AMPLITUDE**

In Section 9.5.14 we wrote the phase averaged mechanical energy flux for an internal gravity wave,  $\langle \mathbf{v}' \varphi' \rangle$ , in terms of the pressure amplitude,  $\tilde{\varphi}$ , as well as the buoyancy amplitude,  $\tilde{b}$ . Here we do the same but now in terms of the vertical velocity amplitude,  $\tilde{w}$ .

- (a) Write the energy flux for an internal gravity wave in terms of the vertical velocity amplitude,  $\tilde{w}$ .
- (b) Write the vertical component of the internal gravity wave energy flux,  $\langle \mathbf{v}' \varphi' \rangle \cdot \hat{\mathbf{z}}$ , in terms of the vertical velocity amplitude,  $\tilde{w}$ , the wavenumbers,  $\mathbf{k}_h$  and  $k_z$ , and the angular frequency,  $\omega$ .

## EXERCISE 9.6: HYDROSTATIC INERTIA-GRAVITY WAVES

Throughout this chapter we studied inertia-gravity waves as they appear in a non-hydrostatic flow that satisfies the linear Boussinesq equations (9.17a)-(9.17d). However, inertia-gravity waves also exist in flows maintaining an approximate hydrostatic balance, and in this exercise we derive their dispersion relation and discuss the relation to the non-hydrostatic waves. Hint: see Section 9.8.2 as well as page 280 of [Vallis \(2017\)](#).

- (a) Write the hydrostatic version of the linear Boussinesq equations (9.17a)-(9.17d).
- (b) Following the methods from Section 9.8, derive the free wave equation satisfied by the vertical velocity,  $w'$ , in a hydrostatic flow. That is, derive the hydrostatic version of equation (9.127). Show all steps.
- (c) Derive the dispersion relation for inertia-gravity waves in a hydrostatic flow with a constant reference buoyancy frequency,  $N_R$ .
- (d) Start from the dispersion relation for inertia-gravity waves in a non-hydrostatic flow. Discuss the maths and physics of the length scale limiting process that results in the dispersion relation for inertia-gravity waves in a hydrostatic flow.
- (e) What does the hydrostatic limit say about the angular frequency of inertia-gravity waves relative to non-hydrostatic inertia-gravity waves?





# Chapter 10

## INTERNAL GRAVITY WAVES: CASE STUDIES

In this chapter we consider some case studies that help further our understanding of internal gravity waves as they appear in geophysical fluids. Sections 10.2 and 10.3 consider the generation of internal gravity waves via flow over topography, thus giving rise to *mountain waves*. Mountain waves arise when stratified fluid flows over topography in both the atmosphere (hence their name) and the ocean. We limit the analysis to *stationary waves*, which are time independent in the rest frame of the mountains. Doing so allows us to avoid the mathematical questions of transient wave adjustment, while still exposing us to the Fourier analysis and synthesis methods central to wave mechanics. Section 10.4 examines gravity waves within a bounded domain that also has a gently varying stratification. This case study allows us to use the rudiments of ray theory whose general features were presented in Chapter 2.

### READER'S GUIDE TO THIS CHAPTER

This chapter is a natural extension of material studied in Chapter 9. Further resources for this chapter can be found in [Lighthill \(1978\)](#), [Pedlosky \(2003\)](#), [Sutherland \(2010\)](#), [Cushman-Roisin and Beckers \(2011\)](#), [Kundu et al. \(2016\)](#), [Vallis \(2017\)](#), and [Buijsman et al. \(2019\)](#). The second half of this video offers some pedagogical visualizations of stratified flow phenomena, and [this video provides more visualizations from simulations and laboratory tank experiments](#).

<b>10.1</b>	<b>Loose threads</b>	<b>292</b>
<b>10.2</b>	<b>Gravity waves from a sinusoidal mountain range</b>	<b>292</b>
10.2.1	Linearized equations with constant reference flow	293
10.2.2	Bottom topography and bottom boundary condition	294
10.2.3	Galilean transformation to the moving flow's reference frame	295
10.2.4	Free space dispersion relation with moving reference flow	295
10.2.5	Wavenumbers for stationary inertia-gravity waves	297
10.2.6	Stationary mountain inertia-gravity waves with $f^2 < \omega_R^2 < N_R^2$	298
10.2.7	Stationary mountain gravity waves with $0 < \omega_R^2 < N_R^2$	299
10.2.8	Trapped mountain waves with $\omega_R > N_R$	301
10.2.9	Comments on the phenomenology	303
<b>10.3</b>	<b>Gravity waves from an arbitrary mountain</b>	<b>303</b>
10.3.1	Mathematical formulation using Fourier analysis	304
10.3.2	Wave solution in $(k_x, z)$ -space	305
10.3.3	Wave solution in $(x, z)$ -space	305
10.3.4	Lorentzian topography	306
<b>10.4</b>	<b>Gravity waves in gently varying stratification</b>	<b>307</b>
10.4.1	The two length scale assumption	308

10.4.2	The wave equation with the WKB ansatz . . . . .	309
10.4.3	Wave amplitude and vertical wavenumber . . . . .	310
10.4.4	WKB solution for the vertical velocity . . . . .	311
10.4.5	Structure of an internal gravity wave . . . . .	311
10.4.6	Wave packets within a wave guide . . . . .	312
10.4.7	Stretched vertical coordinate for hydrostatic gravity waves . . . . .	313
10.4.8	Comments and further study . . . . .	314
<b>10.5</b>	<b>Vertically standing waves in a bounded domain</b> . . . . .	<b>315</b>
10.5.1	Formulation of the boundary value problem . . . . .	315
10.5.2	Rigid lid boundary condition . . . . .	315
10.5.3	Distinction between $\omega^2 < N_r^2$ and $\omega^2 > N_r^2$ . . . . .	316
10.5.4	Sturm-Liouville eigenvalue problem . . . . .	316
10.5.5	Orthogonality of the eigenmodes . . . . .	317
10.5.6	Non-uniform vertical stratification . . . . .	317
10.5.7	A complete set of eigenmodes with a free surface boundary . . . . .	318
10.5.8	Further study . . . . .	320
<b>10.6</b>	<b>Exercises</b> . . . . .	<b>320</b>

---

## 10.1 Loose threads

- Work through the examples Legg notes, as well as material from [Vallis \(2017\)](#).
- Vertical normal modes and WKB in Section 10.5, as per Lecture 9 of [Pedlosky \(2003\)](#) or Section 6.10 of [Gill \(1982\)](#).
- Expand the ray theory in Section 10.4. Follow elements from Chapter 6 of [Sutherland \(2010\)](#).
- Exercise on refraction and reflection as per [Gill \(1982\)](#) Section 6.9. Or make this into a section.
- Write the vertical eigenmodes from Section 10.5 for the case of  $f \neq 0$  and with  $f^2 < \omega^2 < N_r^2$ . See Section 8.12.1 of [Gill \(1982\)](#).

## 10.2 Gravity waves from a sinusoidal mountain range

Following the discussion for inertial waves in Section 5.5.2, and for shallow water inertia-gravity waves in Section 8.2, we here consider a prescribed and fixed constant horizontal reference (background) flow on a stratified  $f$ -plane that moves over small amplitude topography (e.g., mountains) in an otherwise unbounded domain. The flow over a non-flat bottom provides a forcing of vertical motion as imposed by the bottom kinematic boundary condition, with a vertically moving fluid parcel sampling the continuously stratified reference buoyancy. As such, flow over topography in a continuously stratified fluid serves as a forcing for internal inertia-gravity waves. Those waves that propagate do so both vertically and horizontally throughout the domain. The horizontal wavenumber of stationary waves is set by the horizontal wavenumber of the topography. Likewise, the angular frequency of stationary waves is fully determined by the Doppler shift from the moving reference flow, which is set by the wavevector of the topography and the flow speed. In a vertically unbounded domain, the vertical wavenumber of stationary waves is determined by the dispersion relation. As for surface gravity waves in Chapter 4, waves with high horizontal wavenumber are exponentially trapped near the mountains, whereas waves

with lower horizontal wavenumber propagate vertically. The transition between evanescence and propagation is set by the ratio of the reference stratification to the reference flow.

The *mountain waves* studied in this section represent a rich area of geophysical fluid mechanics of primary importance for waves and mixing in the ocean and atmosphere. For the ocean, barotropic (depth independent) tidal motion offers an important source for the reference flow, with barotropic tidal motions generating internal waves that are referred to as *internal tides*. Our goal is to introduce some of the richness of this geophysical system by mathematically formulating the generation of linear internal waves and studying their properties. For simplicity, we focus on the *stationary waves* that arise from a fully developed wave field.<sup>1</sup> Some of the formulation in this section is analogous to that of the shallow water in Section 8.2, yet with two key distinctions. First, the fluid here is continuously stratified, and second, the flow is non-hydrostatic. Both of these properties support a rich wave field that is not constrained by the columnar motion found in hydrostatic shallow water layers.

### 10.2.1 Linearized equations with constant reference flow

Figure 10.1 depicts the physical system studied in this section, in which we consider a reference flow whose velocity is given by

$$\mathbf{v} = \mathbf{u}_R + \mathbf{v}', \quad (10.1)$$

where  $\mathbf{u}_R$  is a prescribed space-time constant horizontal reference flow, and  $\mathbf{v}'$  is the space-time dependent flow relative to this reference. Furthermore, we assume that the steady reference flow is in  $f$ -plane geostrophic balance with a prescribed reference pressure field

$$f \hat{\mathbf{z}} \times \mathbf{u}_R = -\nabla \varphi_R. \quad (10.2)$$

Alternatively, for the case with  $f = 0$ , we simply assume the reference flow is down the prescribed gradient of the reference pressure. In either case we have no concern for how the reference pressure field is maintained, only that it supports the steady reference flow.

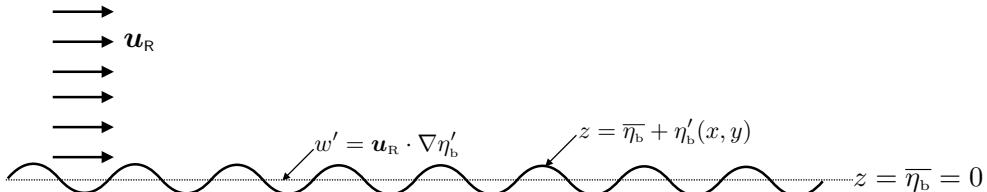


FIGURE 10.1: Schematic of a constant reference flow,  $\mathbf{u}_R$ , moving over small amplitude monochromatic bottom topography,  $z = \bar{\eta}_b + \eta'_b(x, y) = 0 + \eta_o e^{i\mathbf{k}_b \cdot \mathbf{x}}$ , where we assume  $\bar{\eta}_b = 0$ , and with  $\eta_o$  a real amplitude, and  $\mathbf{k}_b$  a horizontal wavevector. We choose to orient the topography wavevector so that  $\mathbf{u}_R \cdot \mathbf{k}_b < 0$ , which is based on noting that the reference flow,  $\mathbf{u}_R$ , when viewed in the rest frame, is equivalent to topography moving in the direction opposite to  $\mathbf{u}_R$  when viewed in the boosted frame. We assume the topography has a small amplitude in the sense that  $\eta_o |\mathbf{k}_b| \ll 1$ , thus ensuring that the generated waves are linear. There is no upper boundary nor side boundary, so waves generated by flow over the bottom are free to propagate horizontally and vertically. At the bottom, the full flow must satisfy the kinematic boundary condition,  $w = \mathbf{u} \cdot \nabla \eta_b$ . Linearizing this boundary condition brings about its evaluation at  $z = \bar{\eta}_b = 0$  rather than at  $z = \eta_b(x, y)$ . Additionally, with  $\eta'_b$  of small amplitude, the linearized bottom boundary condition is  $w' = \mathbf{u}_R \cdot \nabla \eta'_b = i(\mathbf{u}_R \cdot \mathbf{k}_b) \eta'_b$  at  $z = 0$ .

<sup>1</sup>Recall from Section 1.2 that stationary waves have no time dependence in the rest frame. We here encounter stationary waves as the steady solution to the propagating internal gravity wave equations. The initial value problem requires more mathematical machinery beyond our scope.

Following the steps in Section 9.2, we decompose the density and pressure according to

$$\rho(\mathbf{x}, t) = \rho_0 + \rho_R(\mathbf{x}) + \rho'(\mathbf{x}, t), \quad (10.3)$$

where the reference density,  $\rho_R$ , is a function of all three spatial coordinates, and where we consider two static pressures,  $p_0$  and  $p_R$ , that are in hydrostatic balance with their corresponding densities

$$\frac{dp_0}{dz} = -\rho_0 g \quad \text{and} \quad \frac{\partial p_R}{\partial z} = -\rho_R g. \quad (10.4)$$

Inserting this density and pressure decomposition in the Boussinesq equations (9.1a)-(9.1d), and then linearizing, leads to the linearized governing equations

$$(\partial_t + \mathbf{u}_R \cdot \nabla) \mathbf{u}' + f \hat{z} \times \mathbf{u}' = -\nabla_h \varphi' \quad \text{linear horizontal velocity equation} \quad (10.5a)$$

$$(\partial_t + \mathbf{u}_R \cdot \nabla) w' = -\partial_z \varphi' + b' \quad \text{linear vertical velocity equation} \quad (10.5b)$$

$$(\partial_t + \mathbf{u}_R \cdot \nabla) b' = -w' N_R^2 \quad \text{linear buoyancy equation} \quad (10.5c)$$

$$\nabla \cdot \mathbf{v}' = 0 \quad \text{continuity for velocity fluctuations,} \quad (10.5d)$$

which reduce to equations (9.17a)-(9.17d) when  $\mathbf{u}_R = 0$ .

### 10.2.2 Bottom topography and bottom boundary condition

As for the shallow water case studied in Section 8.2 (in particular, see Section 8.2.4), we consider the domain to be bounded from below with monochromatic bottom topography undulations given by

$$\eta'_b(x, y) = \eta_0 e^{i \mathbf{k}_b \cdot \mathbf{x}}, \quad (10.6)$$

such as depicted in Figure 10.1. In this expression,  $\eta_0$  is the real and constant amplitude of the topography and  $\mathbf{k}_b$  is a horizontal wave number that specifies the direction and wavelength of the topography. Linearity of the waves generated by the topography is ensured by assuming

$$\eta_0 |\mathbf{k}_b| \ll 1, \quad (10.7)$$

so that the amplitude of the topography is small on the length scale set by the topography's wavelength.

We orient the topography's wavevector,  $\mathbf{k}_b$ , to be opposite that of the reference flow, so that

$$\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b \geq 0, \quad (10.8)$$

where  $\omega_R$  is an angular frequency of the bottom boundary forcing induced by the flow over the topography. As discussed in Section 8.2.5 when studying shallow water waves generated by topography, the sign for the inequality (10.8) reflects the fact that topography is stationary in the rest frame but it moves in the direction opposite to the reference flow when viewed in the frame following the reference flow. This inequality can be considered a causality condition and it is central to the phase functions encountered in the following.<sup>2</sup>

The no-flow bottom kinematic boundary condition for the full nonlinear flow takes the form

---

<sup>2</sup>We are considering the steady state flow where one generally has time symmetry. However, keeping in mind the transient situation where time symmetry is broken, we here break time symmetry through acknowledging a directionality to the reference flow (left to right) and by insisting that the forcing angular frequency,  $\omega_R$ , is positive.

(Section ??)

$$w = \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \eta_b(x, y). \quad (10.9)$$

Inserting the velocity decomposition (10.1) and the bottom topography,

$$z = \bar{\eta}_b + \eta'_b(x, y), \quad (10.10)$$

leads to

$$w' = (\mathbf{u}' + \mathbf{u}_R) \cdot \nabla(\bar{\eta}_b + \eta'_b) = (\mathbf{u}' + \mathbf{u}_R) \cdot \nabla \eta'_b \approx \mathbf{u}_R \cdot \nabla \eta'_b \quad \text{at } z = 0, \quad (10.11)$$

where the approximation arises from dropping the relatively small term,  $\mathbf{u}' \cdot \nabla \eta'_b$ .

### 10.2.3 Galilean transformation to the moving flow's reference frame

Following the shallow water case from Section 8.2.2, we transform to the moving frame of the constant reference flow by introducing the boosted (moving reference frame) coordinates

$$\bar{t} = t \quad \text{and} \quad \bar{x} = x - \mathbf{u}_R t \quad \text{and} \quad \bar{v} = \mathbf{v} - \mathbf{u}_R. \quad (10.12)$$

Observers in the rest frame see the topography at rest and the reference flow moving with velocity,  $\mathbf{u}_R$ . In contrast, observers in the boosted reference frame see the topography moving with velocity  $-\mathbf{u}_R$  whereas the reference flow is at rest. Since the flow is constant in space and time, the two observers are inertial, so the transformation between the two reference frames is Galilean (Section ??). Following the discussion in Section ??, the Galilean transformation (10.12) leads to the transformed differential operators

$$\partial_{\bar{t}} = \partial_t + \mathbf{u}_R \cdot \nabla \quad \text{and} \quad \bar{\nabla} = \nabla, \quad (10.13)$$

and the corresponding transformation of the linearized governing equations (10.5a)-(10.5d)

$$\partial_{\bar{t}} \mathbf{u}' + f \hat{z} \times \mathbf{u}' = -\bar{\nabla}_h \varphi' \quad (10.14a)$$

$$\partial_{\bar{t}} w' = -\partial_{\bar{z}} \varphi' + b' \quad (10.14b)$$

$$\partial_{\bar{t}} b' = -w' N_R^2 \quad (10.14c)$$

$$\bar{\nabla} \cdot \mathbf{v}' = 0. \quad (10.14d)$$

Notice how transformation to the moving frame removed advection by the reference flow from the horizontal velocity equation. Indeed, that result is the main operational purpose of transforming to the moving reference frame.

### 10.2.4 Free space dispersion relation with moving reference flow

Before considering the case of a lower boundary, we here establish some basic results for waves moving in a uniform flow in free space. For that purpose, note that equations (10.14a)-(10.14d) are mathematically identical to the linear equations (9.17a)-(9.17d) that apply to the case of zero reference flow, only here with the time derivative computed in the boosted reference frame,  $\partial_{\bar{t}}$ , rather than the rest frame time derivative,  $\partial_t$ . Consequently, when assuming a constant reference state buoyancy frequency, the equation for the vertical velocity in the boosted reference frame is given by equation (9.142) with boosted coordinates

$$(\partial_{\bar{t}} \bar{\nabla}^2 + N_R^2 \bar{\nabla}_h^2 + f^2 \partial_{\bar{z}\bar{z}}) w' = 0. \quad (10.15)$$

In the absence of any boundaries, the phase function for linear waves is written<sup>3</sup>

$$\mathbf{k} \cdot \mathbf{x} - \omega t = \mathbf{k} \cdot (\bar{\mathbf{x}} + \mathbf{u}_R t) - \omega t = \mathbf{k} \cdot \bar{\mathbf{x}} - (\omega - \mathbf{u}_R \cdot \mathbf{k}) \bar{t} = \mathbf{k} \cdot \bar{\mathbf{x}} - \bar{\omega} \bar{t}, \quad (10.16)$$

where the final equality introduced<sup>4</sup>

$$\bar{\omega} = \omega - \mathbf{u}_R \cdot \mathbf{k}, \quad (10.17)$$

which relates the angular frequency in the boosted frame,  $\bar{\omega}$ , to that in the rest frame,  $\omega$ . The frequency shift,  $-\mathbf{u}_R \cdot \mathbf{k}$ , is known as the *Doppler shift*, which vanishes when the phase is directed orthogonal to the reference flow,  $\mathbf{u}_R \cdot \mathbf{k} = 0$ . Furthermore, the boosted frame's angular frequency,  $\bar{\omega}$ , vanishes for waves whose Doppler shift satisfies  $\omega = \mathbf{u}_R \cdot \mathbf{k}$ . For these waves, the boosted reference frame rides along a fixed wave and so there is no propagation within this reference frame.

### Critical levels

Plugging the free wave ansatz,

$$w' = \tilde{w} e^{i(\mathbf{k} \cdot \bar{\mathbf{x}} - \bar{\omega} \bar{t})}, \quad (10.18)$$

into the vertical velocity equation (10.15) leads to the dispersion relation

$$\bar{\omega}^2 = (\omega - \mathbf{u}_R \cdot \mathbf{k})^2 = N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma = [(N_R \mathbf{k}_h)^2 + (f k_z)^2] / |\mathbf{k}|^2. \quad (10.19)$$

This is a straightforward extension of the frequency from the rest frame value,  $\omega$ , to the moving frame.

Wave energy propagates according to the group velocity, which we compute from the wavevector gradient

$$\nabla_{\mathbf{k}} \bar{\omega}^2 = 2 \bar{\omega} \nabla_{\mathbf{k}} \bar{\omega} = 2 \bar{\omega} (\nabla_{\mathbf{k}} \omega - \mathbf{u}_R) = 2 \bar{\omega} (\mathbf{c}_g - \mathbf{u}_R), \quad (10.20)$$

where  $\mathbf{c}_g = \nabla_{\mathbf{k}} \omega$  is the group velocity (9.159) as measured in the rest frame, and  $\mathbf{c}_g - \mathbf{u}_R$  is that in the moving frame. Computing  $\nabla_{\mathbf{k}} \bar{\omega}^2$  yields

$$\mathbf{c}_g - \mathbf{u}_R = \frac{(N_R^2 - f^2) k_z}{\bar{\omega} |\mathbf{k}|^4} (k_z \mathbf{k}_h - |\mathbf{k}_h|^2 \hat{\mathbf{z}}). \quad (10.21)$$

It is particularly revealing to consider gravity waves with  $f = 0$ , in which we find the vertical component to the group velocity

$$\hat{\mathbf{z}} \cdot \mathbf{c}_g = -(\omega - \mathbf{u}_R \cdot \mathbf{k}) k_z / |\mathbf{k}|^2. \quad (10.22)$$

The group velocity slows down, and halts, as the wave frequency approaches  $\mathbf{u}_R \cdot \mathbf{k}$ . Although derived here for a constant reference flow, this equation holds to leading order for gently varying reference flows using WKB methods as in Section 10.4. In such cases we can find the vertical

---

<sup>3</sup>Throughout this book we are working with particle and wave speeds that are far smaller than the speed of light, thus enabling the use of Galilean space-time rather than Lorentz space-time. Hence, we do not encounter the special relativistic effects from length contraction, and so there are no changes to the wavelength when moving to the Galilean boosted reference frame. Consequently, movement to the Galilean boosted reference frame leaves the wavevector,  $\mathbf{k}$ , unchanged.

<sup>4</sup>We introduced the same angular frequency relation in equation (8.38) for shallow water waves generated by flow over topography.

component of the group velocity vanish at certain *critical levels*, which are levels where waves can deposit their energy to the mean flow.

### 10.2.5 Wavenumbers for stationary inertia-gravity waves

Now consider the case of a monochromatic bottom (mountains) as depicted in Figure 10.1, with bottom topography of the form

$$\eta'_b = \eta_o e^{i \mathbf{k}_b \cdot \mathbf{x}} = \eta_o e^{i (\mathbf{k}_b \cdot \bar{\mathbf{x}} - \omega_R \bar{t})}. \quad (10.23)$$

The topography is static in the rest frame, whereas it is a traveling plane wave in the boosted frame with angular frequency  $\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b \geq 0$  given from equation (10.8). In the presence of topography, the waves are not free in all three directions. Rather, they must satisfy the bottom boundary condition (10.11), which takes on the following form for a monochromatic mountain range

$$w'(z=0) = \mathbf{u}_R \cdot \nabla \eta'_b = i(\mathbf{u}_R \cdot \mathbf{k}_b) \eta'_b = -i \omega_R \eta'_b. \quad (10.24)$$

For waves that are stationary in the rest frame ( $\omega = 0$  so that  $\bar{\omega} = \omega_R$ ), we satisfy the boundary condition (10.24) by setting the vertical velocity to

$$w' = -i \omega_R \eta_o e^{i (\mathbf{k}_b \cdot \bar{\mathbf{x}} + k_z \bar{z} - \omega_R \bar{t})} = -i \omega_R \eta_o e^{i (\mathbf{k}_b \cdot \mathbf{x} + k_z z)}, \quad (10.25)$$

which has a real part given by

$$w' = \omega_R \eta_o \sin(\mathbf{k}_b \cdot \mathbf{x} + k_z z). \quad (10.26)$$

Notably, the horizontal wavevector is set by the topography,

$$\mathbf{k}_h = \mathbf{k}_b. \quad (10.27)$$

With constant  $N_R$ , we can use the stationary form of the dispersion relation (10.19),

$$\bar{\omega}^2 = \omega_R^2 = (\mathbf{u}_R \cdot \mathbf{k})^2 = N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma = [(N_R |\mathbf{k}_h|)^2 + (f k_z)^2] / |\mathbf{k}|^2, \quad (10.28)$$

to find the squared vertical wavenumber for the stationary inertia-gravity waves (again,  $\mathbf{k}_h = \mathbf{k}_b$  via equation (10.27))

$$k_z^2 = \frac{|\mathbf{k}_b|^2 (N_R^2 - \omega_R^2)}{\omega_R^2 - f^2}, \quad (10.29)$$

with the vertical wavenumber either real or imaginary according to

$$k_z^2 = \begin{cases} < 0 & \text{if } \omega_R^2 < f^2 \\ > 0 & \text{if } f^2 < \omega_R^2 < N_R^2 \\ < 0 & \text{if } \omega_R^2 > N_R^2. \end{cases} \quad (10.30)$$

The cases with  $k_z^2 < 0$  lead to stationary waves that are oscillating in the horizontal yet exponentially trapped in the vertical, much like the surface gravity waves studied in Chapter 4 or the edge waves studied in 6.5. In contrast, with  $k_z^2 > 0$  there are stationary waves extending throughout the vertical. The regime of vertically extended stationary waves is given by

$$2\pi/N_R < \Lambda_b/U < 2\pi/f, \quad (10.31)$$

where

$$\Lambda_b = 2\pi/|\mathbf{k}_b| \quad (10.32)$$

is the topography's wavelength. In the atmosphere, typical values for the reference state are  $N_R = 10^{-2} \text{ s}^{-1}$ ,  $U = 10 \text{ m s}^{-1}$ , and  $f = 10^{-4} \text{ s}^{-1}$ , in which case there are vertically extended inertia-gravity waves for mountains having wavelengths within the range

$$2\pi U/N_R < \Lambda_b < 2\pi U/f \implies 2\pi \times 10^3 \text{ m} < \Lambda_b < 2\pi \times 10^5 \text{ m}. \quad (10.33)$$

For the deep ocean we take  $N_R = 10^{-3} \text{ s}^{-1}$ ,  $U = 10^{-1} \text{ m s}^{-1}$  (barotropic tidal speeds), and  $f = 10^{-4} \text{ s}^{-1}$ , so that there are vertically extended stationary inertia-gravity waves for undersea mountains having wavelengths within the range

$$2\pi \times 10^2 \text{ m} < \Lambda_b < 2\pi \times 10^3 \text{ m}. \quad (10.34)$$

These numbers are only meant to give an impression of the approximate scales of mountain forcing that generate vertically extended internal inertia-gravity waves. One key point is the roughly one order smaller scales for the ocean relative to the atmosphere.

### 10.2.6 Stationary mountain inertia-gravity waves with $f^2 < \omega_R^2 < N_R^2$

For  $f^2 < \omega_R^2 < N_R^2$ , we have stationary inertia-gravity waves with  $k_z^2 > 0$  according to equation (10.29). As the reference flow interacts with the mountains at the lower boundary, internal inertia-gravity waves propagate energy upwards according to the group velocity. As seen in Section 9.5.10, an upward group velocity means a downward phase velocity, so that we take the negative root from equation (10.29)

$$k_z = -|\mathbf{k}_b| \sqrt{\frac{N_R^2 - \omega_R^2}{\omega_R^2 - f^2}} = -|k_z| < 0, \quad (10.35)$$

which renders the stationary vertical velocity in the wave<sup>5</sup>

$$w' = -i\omega_R \eta_o e^{i(\mathbf{k}_b \cdot \mathbf{x} - |k_z| z)}. \quad (10.36)$$

For the specific case of zonal reference flow,  $\mathbf{u}_R = U \hat{\mathbf{x}}$  with  $U > 0$ , we have  $\mathbf{k}_b = -|\mathbf{k}_b| \hat{\mathbf{x}}$ , so that

$$\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b = U |\mathbf{k}_b| \implies w' = -iU |\mathbf{k}_b| \eta_o e^{-i(|\mathbf{k}_b| x + |k_z| z)}, \quad (10.37)$$

which has real part

$$w' = \tilde{w} \sin(|\mathbf{k}_b| x + |k_z| z) = -U |\mathbf{k}_b| \eta_o \sin(|\mathbf{k}_b| x + |k_z| z), \quad (10.38)$$

where we defined the vertical velocity amplitude

$$\tilde{w} = -U |\mathbf{k}_b| \eta_o. \quad (10.39)$$

Evidently, the vertical velocity amplitude is directly related to the reference flow magnitude, the topography wavenumber, and the topography amplitude.

---

<sup>5</sup>In equation (10.35), and elsewhere in this chapter, we find it useful to introduce the absolute value around a wavenumber in order to clearly expose the sign of the wavenumber. Doing so can greatly help in capturing the proper phase relations for the waves.

To determine polarization relations for the other fields, we can return to those derived earlier for the free waves in Section 10.2.4, specializing to the case with  $\omega = \omega_R$ . Alternatively, we can set  $\partial_t = 0$  in the linear governing equations (10.5a)-(10.5d). This work is considered in Exercise 10.1 and follows a method similar to that taken for stationary mountain gravity waves in Section 10.2.7.

### 10.2.7 Stationary mountain gravity waves with $0 < \omega_R^2 < N_R^2$

We now specialize to  $f = 0$ , such as appropriate for mountain waves that are small enough laterally to allow us to ignore the Coriolis acceleration. Also, continue to assume a zonal reference flow,  $\mathbf{u}_R = U \hat{\mathbf{x}}$  with  $U > 0$  and  $\mathbf{k}_b = -|\mathbf{k}_b| \hat{\mathbf{x}}$ . The vertical velocity is still given by equation (10.38)

$$w' = -U |\mathbf{k}_b| \eta_o \sin(|\mathbf{k}_b| x + |k_z| z), \quad (10.40)$$

yet here with the angular frequency for stationary waves given by

$$\omega_R = N_R |\mathbf{k}_b| / |\mathbf{k}| = U |\mathbf{k}_b| \implies U = N_R / |\mathbf{k}|, \quad (10.41)$$

and the vertical wavenumber corresponding to an upward directed group velocity

$$k_z = -|\mathbf{k}_b| \sqrt{(N_R / \omega_R)^2 - 1} = -\sqrt{(N_R / U)^2 - |\mathbf{k}_b|^2} < 0. \quad (10.42)$$

As discussed in Exercise 10.2, the frequency ratio,

$$\text{Fr}(|\mathbf{k}_b|) = \frac{\omega_R}{N_R} = \frac{U |\mathbf{k}_b|}{N_R}, \quad (10.43)$$

can be considered a wavenumber dependent Froude number. Inserting this Froude number into equation (10.42) yields

$$k_z = -|\mathbf{k}_b| \sqrt{\text{Fr}^{-2} - 1}, \quad (10.44)$$

so that the Froude number distinguishes vertically propagating stationary gravity waves ( $\text{Fr} < 1$ ) from vertically evanescent waves ( $\text{Fr} > 1$ ).

#### Lines of constant phase

For the vertical propagating waves ( $\text{Fr} < 1$ ), the vertical velocity (10.40) is constant along phase lines that satisfy

$$|\mathbf{k}_b| x + |k_z| z = \text{constant}. \quad (10.45)$$

These phase lines have a slope

$$\frac{dz}{dx} = -|\mathbf{k}_b| / |k_z|, \quad (10.46)$$

which indicates that the phase lines slope up to the left such as depicted in Figure 10.2. Evidently, the phase lines tilt into the reference flow.

#### Polarization relations

The polarization relations for the stationary waves are determined by the steady linear governing equations (10.5a)-(10.5d) with  $f = 0$  and  $\partial_t = 0$

$$(\mathbf{u}_R \cdot \nabla) \mathbf{u}' = -\nabla_h \varphi' \quad \text{steady linearized horizontal velocity equation} \quad (10.47a)$$

$$(\mathbf{u}_R \cdot \nabla) w' = -\partial_z \varphi' + b' \quad \text{steady linearized vertical velocity equation} \quad (10.47b)$$

$$(\mathbf{u}_R \cdot \nabla) b' = -w' N_R^2 \quad \text{steady linearized buoyancy equation} \quad (10.47c)$$

$$\nabla \cdot \mathbf{v}' = 0 \quad \text{continuity for velocity fluctuations.} \quad (10.47d)$$

All motion is in the  $x$ - $z$  plane, so that the continuity equation (10.47d) gives the zonal velocity

$$u' = -(|k_z|/|\mathbf{k}_b|) w' \implies u' \frac{dz}{dx} = w' \quad \text{and} \quad u' = U |k_z| \eta_o \sin(|\mathbf{k}_b| x + |k_z| z), \quad (10.48)$$

where we used the slope relation (10.46). The steady buoyancy equation (10.47c) yields

$$U \partial_x b' = N_R^2 U |\mathbf{k}_b| \eta_o \sin(|\mathbf{k}_b| x + |k_z| z) \implies b' = -\eta_o N_R^2 \cos(|\mathbf{k}_b| x + |k_z| z). \quad (10.49)$$

Likewise, the steady horizontal velocity equation (10.47a) gives the wave pressure

$$\varphi' = -U u' = -\eta_o |k_z| U^2 \sin(|\mathbf{k}_b| x + |k_z| z). \quad (10.50)$$

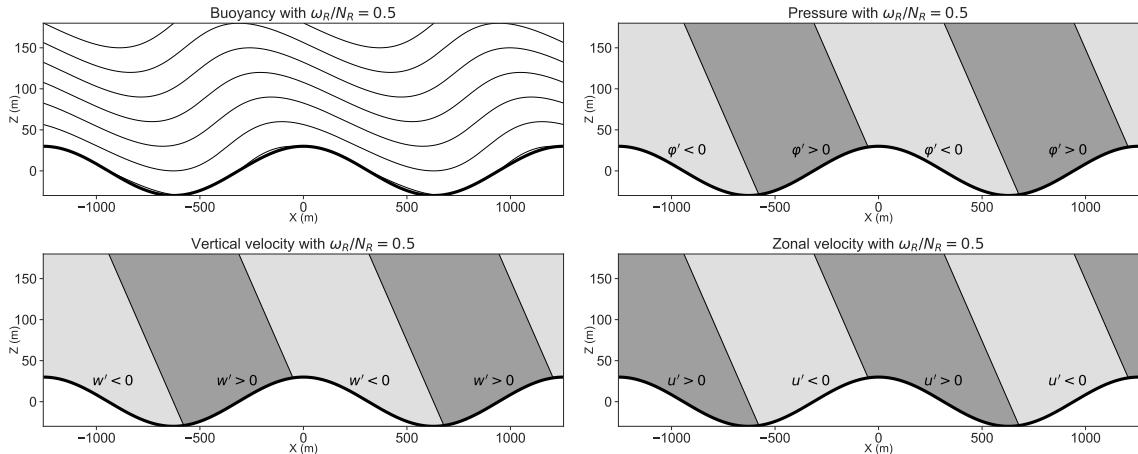


FIGURE 10.2: Stationary mountain gravity waves generated by an eastward reference flow,  $\mathbf{u}_R = U \hat{x}$ , over monochromatic topography. We set the parameters according to ocean values, with  $\eta_o = 30$  m,  $U = 0.1 \text{ m s}^{-1}$ ,  $N_R = 10^{-3} \text{ s}^{-1}$ , and  $\Lambda_b = 2\pi \times 200$  m, so that  $\omega_R/N_R = 0.5$ . Top left panel: contours of the buoyancy field,  $b = b_R + b'$ , with  $b_R = N_R^2 z$  and  $b' = -\eta_o N_R^2 \cos(|\mathbf{k}_b| x + |k_z| z)$  from equation (10.49). Top right panel: perturbation pressure field,  $\varphi' = -U u' = U w' |k_z| / |\mathbf{k}_b|$ , from equation (10.50). Lower left panel: the vertical velocity anomaly,  $w' = -U |\mathbf{k}_b| \eta_o \sin(|\mathbf{k}_b| x + |k_z| z)$  from equation (10.40). Lower right panel: the zonal velocity anomaly,  $u' = -|k_z| / |\mathbf{k}_b| w'$  from equation (10.48). The color scale is set to highlight just the sign of the field. The phase shift, whereby high pressures are on the windward side of the mountains and low pressures on the lee side, leads to a topographic pressure form stress acting on the fluid according to equation (10.51).

### Wave induced topographic form stress

In Figure 10.2 we illustrate the buoyancy, pressure, vertical velocity, and zonal velocity within a stationary gravity wave for a particular set of parameters. Note the phase lines that are tilted into the mean flow with the positive pressure perturbation on the upwind side of the mountains. This orientation is reflective of the general discussion surrounding Figure ?? when studying pressure form stress acting between the fluid and the solid lower boundary. From that discussion we know that a positive pressure anomaly on the upwind side of a mountain renders an eastward form stress from the fluid to the solid, and a corresponding (through Newton's third law) westward topographic form stress directed from the solid to the fluid.

From the discussion in Section ??, we write the topographic form stress acting on the fluid as  $-p \nabla \eta_b$ . The pressure is evaluated at the solid-fluid boundary for the nonlinear case, whereas for linear waves we evaluate the topographic form stress at  $z = 0$ , with a phase average yielding

$$\langle -\varphi' \partial_x \eta_b \rangle = -U^2 \eta_o^2 |\mathbf{k}_b| |k_z| / 2. \quad (10.51)$$

This westward topographic form stress acts from the solid onto the fluid, with a corresponding eastward form stress of the same magnitude from the fluid to the solid. The stress is directly related to the square of the reference flow speed and the square of the topography amplitude. It is also linearly related to the vertical wavenumber of the gravity wave and the vertical wavenumber of the topography. Since the stress is directed contrary to the reference flow direction, it is commonly referred to as a *mountain drag*.

### Wave energy flux

From Section 9.5.14 we know that the energy flux carried by a gravity wave is given by  $\mathbf{v}' \varphi'$ . For stationary mountain waves in the  $x$ - $z$  plane we have

$$\mathbf{v}' \varphi' = -(\hat{x} u' + \hat{z} w') U u' = \frac{U w' w' |k_z|}{|\mathbf{k}_b|^2} (-|k_z| \hat{x} + |\mathbf{k}_b| \hat{z}). \quad (10.52)$$

Now the group velocity for internal gravity waves is given by equation (9.71), which takes on the following form for the forced stationary waves

$$\mathbf{c}_g = \frac{N_R k_z}{|\mathbf{k}|^3 |\mathbf{k}_b|} (k_z \mathbf{k}_b - |\mathbf{k}_b|^2 \hat{z}) \quad (10.53a)$$

$$= \frac{N_R |k_z|}{|\mathbf{k}|^3} (-|k_z| \hat{x} + |\mathbf{k}_b| \hat{z}) \quad (10.53b)$$

$$= \frac{U |k_z|}{|\mathbf{k}|^2} (-|k_z| \hat{x} + |\mathbf{k}_b| \hat{z}), \quad (10.53c)$$

where the second equality set  $\mathbf{k}_b = -|\mathbf{k}_b| \hat{x}$  and  $k_z = -|k_z| < 0$ , and the third equality set  $U = N_R / |\mathbf{k}|$  from equation (10.41). Next, we make use of equation (9.101) to write the mechanical energy within the stationary internal gravity wave

$$\mathcal{K}' + \mathcal{A}' = \langle w' w' \rangle |\mathbf{k}|^2 / |\mathbf{k}_b|^2 = \frac{1}{2} \eta_o^2 U^2 |\mathbf{k}|^2 = \frac{1}{2} \eta_o^2 N_R^2, \quad (10.54)$$

which then brings the phase averaged energy flux to the standard form holding for a linear wave

$$\langle \mathbf{v}' \varphi' \rangle = \mathbf{c}_g (\mathcal{K}' + \mathcal{A}'). \quad (10.55)$$

### 10.2.8 Trapped mountain waves with $\omega_R > N_R$

For the case with mountain forcing at a frequency greater than buoyancy frequency,  $\omega_R > N_R$ , the vertical wavenumber becomes imaginary

$$k_z = i \sqrt{|\mathbf{k}_b|^2 - (N_R/U)^2} = i |k_z|, \quad (10.56)$$

which yields a wave solution that is exponentially decaying away from the mountains (see equation (10.36) and remember that  $\mathbf{k}_b = -|\mathbf{k}_b| \hat{\mathbf{x}}$  and  $\omega_R = U |\mathbf{k}_b| > 0$ )

$$w' = -i\omega_R \eta_o e^{i\mathbf{k}_b \cdot \mathbf{x} - |k_z| z} \xrightarrow{\text{Re}} -U |\mathbf{k}_b| \eta_o \sin(|\mathbf{k}_b| x) e^{-|k_z| z}. \quad (10.57)$$

We refer to these fluctuations as *vertically evanescent* or *vertically trapped* gravity waves. The corresponding zonal velocity, buoyancy, and pressure fields are

$$u' = U |k_z| \eta_o \cos(|\mathbf{k}_b| x) e^{-|k_z| z} \quad (10.58a)$$

$$b' = -\eta_o N_R^2 \cos(|\mathbf{k}_b| x) e^{-|k_z| z} \quad (10.58b)$$

$$\varphi' = -U^2 |k_z| \eta_o \cos(|\mathbf{k}_b| x) e^{-|k_z| z}, \quad (10.58c)$$

with Figure 10.3 providing an example.

### Zero energy propagation and zero topographic form stress

Since the vertical velocity (10.57) and pressure (10.58c) are  $\pi/2$  radians out of phase, vertically trapped waves generate no vertical propagation of phase averaged wave energy

$$\langle w' \varphi' \rangle = 0. \quad (10.59)$$

Likewise, there is no topographic form stress since there is no phase shift between the topography and the pressure

$$\langle \varphi' \partial_x \eta_b \rangle = 0. \quad (10.60)$$

### Limit where $\omega_R \gg N_R$

The case of  $\omega_R \gg N_R$  corresponds to large topographic frequency and/or small buoyancy frequency. In this limit the vertical wavenumber has a magnitude  $|k_z| \approx |\mathbf{k}_b|$  so that  $N_R$  drops out from the system. This case corresponds to the surface gravity waves studied in Chapter 4 where stratification was completely ignored. In that case, as well as here, the exponential trapping is determined by the horizontal wavenumber,  $|\mathbf{k}_b|$ . Furthermore, in the case with vertical length scales,  $H$ , such that  $|k_z| H \approx |\mathbf{k}_b| H \ll 1$ , then the vertical velocity becomes a linear function of  $z$ . This case corresponds to the shallow water inertia-gravity waves studied in Section 8.2

### Comments on exponential trapping

As discussed in Section 9.5.8, forcing a stratified fluid at a frequency greater than the buoyancy frequency does not lead to propagating internal gravity waves. For the case of mountain gravity waves, we see that exponential trapping arises when the topographic forcing has a frequency,  $\omega_R$ , greater than the maximum frequency of the buoyancy oscillations,  $N_R$ . In this case, motion induced by the forcing is too fast for fluid particles to exhibit buoyancy oscillations. The result is fluid particles largely following whatever displacements are imposed on them by the external forcing, with exponential damping when moving away from the forcing and with fluid particles in phase with the forcing. In this case, no propagating internal gravity waves are formed, and energy imparted to the fluid stays local to the topographic forcing.

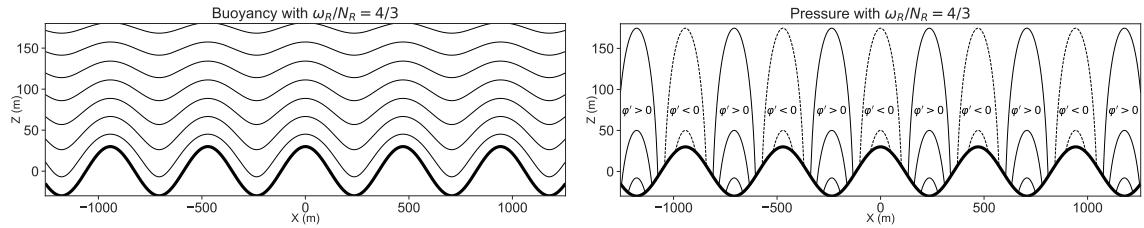


FIGURE 10.3: Stationary trapped mountain gravity waves generated by an eastward reference flow,  $\mathbf{u}_R = U \hat{\mathbf{x}}$ , over monochromatic topography. We set the parameters according to ocean values, with  $\eta_o = 30$  m,  $U = 0.1$  m s $^{-1}$ ,  $N_R = 10^{-3}$  s $^{-1}$ , and  $\Lambda_b = 2\pi \times 75$  m (we used the larger wavelength of  $\Lambda_b = 2\pi \times 200$  m in Figure 10.2), so that  $\omega_R/N_R = 4/3$ . These waves are exponentially trapped near the topography with e-folding scale  $|k_z|^{-1} = 1/\sqrt{(N_R/U)^2 - |\mathbf{k}_b|^2} \approx 110$  m. Left panel: contours of the buoyancy field,  $b = b_R + b'$ , with  $b_R = N_R^2 z$  and  $b' = -\eta_o N_R^2 \cos(|\mathbf{k}_b| x) e^{-|k_z|z}$  from equation (10.58c). Right panel: perturbation pressure field,  $\varphi' = -U^2 |\mathbf{k}_b| \eta_o \cos(|\mathbf{k}_b| x) e^{-|k_z|z}$  from equation (10.58c). Note how the pressure is anomalously high over the valleys (solid contours with  $\varphi' > 0$ ) and low over the crests (dashed contours with  $\varphi' < 0$ ). We chose the horizontal extent to be the same as in Figure 10.2, thus emphasizing how the smaller scale topography in this figure can lead to exponentially trapped waves, whereas the larger scale topography in Figure 10.2 allows for vertical propagation. Since there is no phase shift with height for the vertically trapped waves, there is no topographic form stress, which we confirm analytically in equation (10.60).

### 10.2.9 Comments on the phenomenology

Gravity waves are ubiquitous in the ocean and atmosphere, with flow over topography one of the key methods for generating such waves. As seen in this section, the reference flow speed,  $U > 0$ , and mountain wavenumber,  $|\mathbf{k}_b|$ , together define the angular frequency for the forcing,  $\omega_R = U |\mathbf{k}_b| > 0$ . This forcing generates propagating gravity waves if  $\omega_R < N_R$ , whereas the gravity waves are evanescent (trapped) if  $\omega_R > N_R$ . That is, fast reference flows and/or high wavenumber mountains produce gravity waves that are trapped next to the mountain, whereas slower reference flow or lower wavenumber mountains generate gravity waves that propagate vertically.

Consider the case of an atmosphere with  $N_R = 10^{-2}$  s $^{-1}$  and  $U = 10$  m s $^{-1}$ . Propagating gravity waves are generated for mountains with a horizontal wavelength ( $|\mathbf{k}_b| = 2\pi/\Lambda$ ) larger than  $\Lambda = 2\pi \times 10^3$  m. For the deep ocean, with  $N_R = 10^{-3}$  m and  $U = 10^{-2}$  m s $^{-1}$ , propagating gravity waves are generated for abyssal mountains with wavelengths larger than  $\Lambda = 2\pi \times 10^1$  m. These numbers suggest that all topographic features larger than a minimal scale will generate gravity waves. However, there is an upper bound on the scale that arises from the Coriolis parameter. Namely, for large enough scales the flow feels the Coriolis acceleration, thus requiring us to recall from Section 9.8.1 that propagating inertia-gravity waves have a dispersion relation (9.145)

$$\omega^2 = N_R^2 \cos^2 \gamma + f^2 \sin^2 \gamma \implies f^2 < \omega^2 < N_R^2, \quad (10.61)$$

where we here assume the Coriolis frequency is less than the buoyancy frequency. Hence, the condition for an upper bound on the forcing frequency, set by the buoyancy frequency, is coupled to a lower bound set by the Coriolis frequency. For the atmosphere at middle latitudes with  $|f| = 10^{-4}$  s $^{-1}$ , we find  $2\pi \times 10^3$  m  $< \Lambda < 2\pi \times 10^5$  m, whereas the ocean requires  $2\pi \times 10^1$  m  $< \Lambda < 2\pi \times 10^2$  m.

## 10.3 Gravity waves from an arbitrary mountain

In this section we examine stationary mountain waves resulting from constant zonal flow,  $\mathbf{u}_R = U \hat{\mathbf{x}}$ , moving over arbitrary ( $x$ -dependent) topography. This analysis extends our study of

waves emanating from flow over monochromatic topography considered in Section 10.2. To handle arbitrary topography, we make use of Fourier analysis from Chapter ?? by exploiting the zonal symmetry of the reference flow and geometry. For pedagogy, we find it especially useful to carefully step through the formulation, with a particular realization given by a mountain in the shape of a Lorentzian function.

### 10.3.1 Mathematical formulation using Fourier analysis

Following our discussion in Section ??, we here introduce the Fourier transform and inverse Fourier transform for the bottom topography, velocity, buoyancy, and pressure. Furthermore, from our discussion of causality in Section 8.2.5, and with the reference flow  $\mathbf{u}_R = U \hat{\mathbf{x}}$  with  $U > 0$ , we build in causality by only considering  $k_x \leq 0$ , so that the forcing frequency is non-negative

$$\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b = -U k_x \geq 0 \implies k_x \leq 0. \quad (10.62)$$

These considerations lead to the Fourier transform pairs

$$\widehat{\eta}_b(k_x) = \int_{-\infty}^{\infty} \eta'_b(x) e^{-ik_x x} dx \quad \eta'(x) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{\eta}_b(k_x) e^{ik_x x} dk_x \quad (10.63a)$$

$$\widehat{u}(k_x, z) = \int_{-\infty}^{\infty} u'(x, z) e^{-ik_x x} dx \quad u'(x, z) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{u}(k_x, z) e^{ik_x x} dk_x \quad (10.63b)$$

$$\widehat{w}(k_x, z) = \int_{-\infty}^{\infty} w'(x, z) e^{-ik_x x} dx \quad w'(x, z) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{w}(k_x, z) e^{ik_x x} dk_x \quad (10.63c)$$

$$\widehat{b}(k_x, z) = \int_{-\infty}^{\infty} b'(x, z) e^{-ik_x x} dx \quad b'(x, z) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{b}(k_x, z) e^{ik_x x} dk_x \quad (10.63d)$$

$$\widehat{\varphi}(k_x, z) = \int_{-\infty}^{\infty} \varphi'(x, z) e^{-ik_x x} dx \quad \varphi'(x, z) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{\varphi}(k_x, z) e^{ik_x x} dk_x, \quad (10.63e)$$

where we only perform the  $x$  direction Fourier transform since the domain has a lower boundary and so is not symmetric in  $z$ . The hatted Fourier transform fields have an extra length dimension relative to their  $\mathbf{x}$ -space partners.

Now write the stationary linear equations (10.47a)-(10.47d) using the reference flow,  $\mathbf{u}_R = U \hat{\mathbf{x}}$ ,

$$U \partial_x u' = -\partial_x \varphi' \quad (10.64a)$$

$$U \partial_x w' = -\partial_z \varphi' + b' \quad (10.64b)$$

$$U \partial_x b' = -w' N_R^2 \quad (10.64c)$$

$$\partial_x u' + \partial_z w' = 0, \quad (10.64d)$$

where we set  $v' = 0$  and  $\partial_y = 0$ . Operating on these equations with  $\int_{-\infty}^{\infty} e^{-ik_x x}$ , and assuming all boundary terms vanish, brings about the Fourier space equations

$$U \widehat{u} = -\widehat{\varphi} \quad (10.65a)$$

$$ik_x U \widehat{w} = -\partial_z \widehat{\varphi} + \widehat{b} \quad (10.65b)$$

$$ik_x U \widehat{b} = -\widehat{w} N_R^2 \quad (10.65c)$$

$$ik_x \widehat{u} + \partial_z \widehat{w} = 0, \quad (10.65d)$$

along with the Fourier space version of the linearized bottom boundary condition (10.11)

$$\widehat{w}(k_x, z = 0) = i k_x U \widehat{\eta}(k_x). \quad (10.66)$$

### 10.3.2 Wave solution in $(k_x, z)$ -space

To satisfy the bottom kinematic boundary condition (10.66), we follow the approach from Section 10.2.5 used for the monochromatic topography by writing

$$\widehat{w}(k_x, z) = i k_x U \widehat{\eta}_b e^{i k_z z}. \quad (10.67)$$

The vertical wavenumber is determined through use of the stationary gravity wave dispersion relation as in Section 10.2.7

$$k_z = -\sqrt{(N_R/U)^2 - k_x^2}. \quad (10.68)$$

The negative sign ensures that  $k_z < 0$  for long horizontal waves with  $k_x < N_R/U$ , thus ensuring that gravity wave energy, which follows the group velocity, propagates vertically upward (away from the mountains), thus satisfying causality. The Fourier space continuity equation (10.65d) leads to the zonal velocity Fourier transform

$$\widehat{u} = -i k_z U \widehat{\eta}_b e^{i k_z z}, \quad (10.69)$$

which, with the pressure equation (10.65a), leads to

$$\widehat{\varphi} = i k_z U^2 \widehat{\eta}_b e^{i k_z z}. \quad (10.70)$$

Finally, we make use of the buoyancy equation (10.65c) to find

$$\widehat{b} = \frac{i \widehat{w} N_R^2}{U k_x} = -N_R^2 \widehat{\eta}_b e^{i k_z z}. \quad (10.71)$$

### 10.3.3 Wave solution in $(x, z)$ -space

Making use of the inverse Fourier transforms from equations (10.63a)-(10.63e) renders the  $(x, z)$ -space expressions

$$\eta'(x) = \frac{1}{2\pi} \int_{-\infty}^0 \widehat{\eta}_b e^{i k_x x} dk_x \quad (10.72a)$$

$$u'(x, z) = \frac{-i U}{2\pi} \int_{-\infty}^0 k_z \widehat{\eta}_b e^{i(k_x x + k_z z)} dk_x \quad (10.72b)$$

$$w'(x, z) = \frac{i U}{2\pi} \int_{-\infty}^0 k_x \widehat{\eta}_b e^{i(k_x x + k_z z)} dk_x \quad (10.72c)$$

$$b'(x, z) = \frac{-N_R^2}{2\pi} \int_{-\infty}^0 \widehat{\eta}_b e^{i(k_x x + k_z z)} dk_x \quad (10.72d)$$

$$\varphi'(x, z) = \frac{i U^2}{2\pi} \int_{-\infty}^0 k_z \widehat{\eta}_b e^{i(k_x x + k_z z)} dk_x. \quad (10.72e)$$

Again, the vertical wavenumber is given by equation (10.68), in terms of the horizontal wavenumber and the prescribed background stratification and reference flow speed. Evidently, stationary gravity waves with relatively low horizontal wavenumber ( $k_x^2 < (N_R/U)^2$ ) are

vertically propagating ( $k_z^2 > 0$ ), whereas higher horizontal wavenumber waves are exponentially trapped next to the mountains ( $k_z^2 < 0$ ).

### 10.3.4 Lorentzian topography

Consider the case of topography given by the *Lorentzian* form

$$\eta'_b(x) = \frac{\eta_o \ell^2}{x^2 + \ell^2} = \frac{\eta_o}{1 + (x/\ell)^2}, \quad (10.73)$$

where  $\ell > 0$  is a length scale and  $\eta_o > 0$  is an amplitude. The Fourier transform of the Lorentzian topography is given by

$$\hat{\eta}_b(k_x) = \pi \eta_o \ell e^{-|k_x| \ell}. \quad (10.74)$$

When  $\ell \rightarrow 0$ , the topography sharpens in  $x$ -space whereas the  $k_x$ -space distribution broadens, with the high  $|k_x|$  waves exponentially trapped near the mountain ( $k_z^2 < 0$ ). Conversely, when  $\ell \rightarrow \infty$ , the topography flattens in  $x$ -space whereas the  $k_x$ -space distribution sharpens around  $k_x = 0$ , with such long horizontal waves vertically propagating ( $k_z^2 > 0$ ).

#### Pressure in vertically propagating stationary waves

Getting the signs correct for the inverse Fourier transform can be a bit tricky, so let us work through the case of pressure with some care. First consider the case of vertically propagating stationary waves so that

$$k_z^2 > 0 \quad \text{and} \quad k_z = -\sqrt{(N_R/U)^2 - k_x^2} < 0, \quad (10.75)$$

in which the pressure equation (10.72e) takes the form

$$\varphi'_{\text{prop}}(x, z) = \frac{iU^2}{2\pi} \int_{-\infty}^0 k_z(k_x) \hat{\eta}_b(k_x) e^{i(k_x x + k_z z)} dk_x \quad (10.76a)$$

$$= \frac{iU^2}{2\pi} \int_0^\infty k_z(-k_x) \hat{\eta}_b(-k_x) e^{i(-k_x x + k_z z)} dk_x \quad (10.76b)$$

$$= \frac{iU^2}{2\pi} \int_0^\infty k_z(k_x) \hat{\eta}_b(k_x) e^{i(-k_x x + k_z z)} dk_x, \quad (10.76c)$$

where the final step noted that  $k_z$  is an even function of  $k_x$ , as is  $\hat{\eta}_b(k_x)$  for the Lorentzian topography according to equation (10.74). Next make use of causality with  $k_z = -|k_z| < 0$ , so that

$$\varphi'_{\text{prop}}(x, z) = -\frac{iU^2}{2\pi} \int_0^\infty |k_z| \hat{\eta}_b e^{-i(k_x x + |k_z| z)} dk_x. \quad (10.77)$$

Assuming the Fourier transform of the topography is real (as for the Lorentzian in equation (10.74)), and taking the real part of the pressure field, renders the vertically propagating pressure

$$\varphi'_{\text{prop}}(x, z) = -\frac{U^2}{2\pi} \int_0^\infty |k_z| \hat{\eta}_b \sin(|k_x| x + |k_z| z) dk_x, \quad (10.78)$$

which accords with equation (10.50) found for a monochromatic topography. Finally, we insert the Fourier transform (10.74) to have

$$\varphi'_{\text{prop}}(x, z) = -\frac{U^2 \eta_o \ell}{2} \int_0^{N_R/U} |k_z| e^{-|k_x| \ell} \sin(|k_x| x + |k_z| z) dk_x, \quad (10.79)$$

where we cutoff the upper limit in recognition that the horizontal wavenumber satisfies,  $|k_x| < N_R/U$ , for vertically propagating waves.

### Pressure in vertically trapped waves

With an imaginary vertical wavenumber,

$$k_z^2 < 0 \quad \text{and} \quad k_z = i \sqrt{k_x^2 - (N_R/U)^2} = i |k_z|, \quad (10.80)$$

the pressure equation (10.72e) takes the form

$$\varphi'_{\text{trap}}(x, z) = -\frac{U^2}{2\pi} \int_{-\infty}^0 |k_z(k_x)| \hat{\eta}_b(k_x) e^{i k_x x - |k_z| z} dk_x \quad (10.81a)$$

$$= -\frac{U^2}{2\pi} \int_0^\infty |k_z(-k_x)| \hat{\eta}_b(-k_x) e^{-i k_x x - |k_z| z} dk_x \quad (10.81b)$$

$$= -\frac{U^2}{2\pi} \int_0^\infty |k_z(k_x)| \hat{\eta}_b(k_x) \cos(k_x x) e^{-|k_z| z} dk_x, \quad (10.81c)$$

where the final step took the real part and made use of  $|k_z(-k_x)| = |k_z(k_x)|$ , as well as  $\hat{\eta}_b(-k_x) = \hat{\eta}_b(k_x)$  and its reality. This expression for the trapped pressure accords with equation (10.58c) for the case of monochromatic topography. Making use of the Fourier transform (10.74) renders

$$\varphi'_{\text{trap}}(x, z) = -\frac{U^2 \eta_o \ell}{2} \int_{N_R/U}^\infty |k_z| e^{-|k_x| \ell} \cos(k_x x) e^{-|k_z| z} dk_x, \quad (10.82)$$

where we set the lower limit to  $N_R/U$  in recognition that the waves are trapped with  $|k_x| > N_R/U$ .

### Summary expression for the pressure field

Bringing the two pieces together yields the pressure field

$$\begin{aligned} \varphi'(x, z) = & -\frac{U^2 \eta_o \ell}{2} \int_0^{N_R/U} |k_z| e^{-|k_x| \ell} \sin(|k_x| x + |k_z| z) dk_x \\ & - \frac{U^2 \eta_o \ell}{2} \int_{N_R/U}^\infty |k_z| e^{-|k_x| \ell} \cos(k_x x) e^{-|k_z| z} dk_x. \end{aligned} \quad (10.83)$$

We provide two examples in Figure 10.4, one with a relatively sharp mountain and one that is broader. Note that the sharp mountain exhibits oscillatory features on the lee side of the mountain, which are referred to as *lee waves*.

## 10.4 Gravity waves in gently varying stratification

In much of our study of gravity waves, we assumed a constant background buoyancy frequency,  $N_R$ , thus enabling the use of plane internal gravity waves. However, for geophysical fluids it is

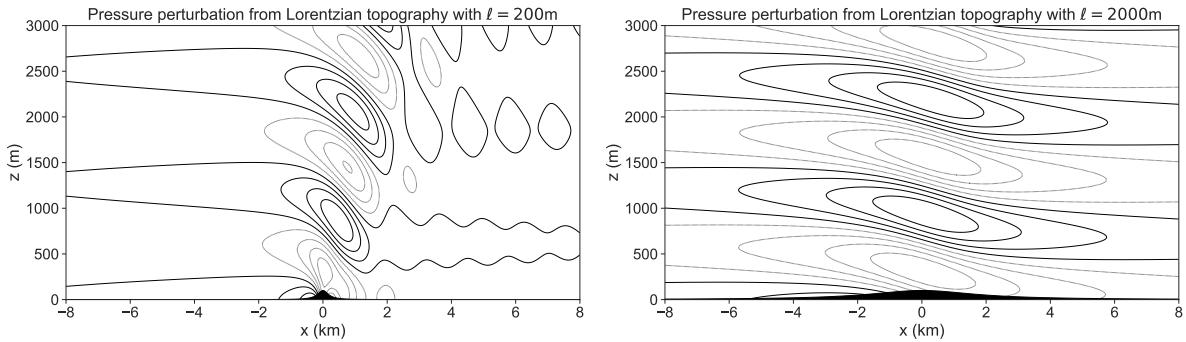


FIGURE 10.4: Stationary gravity waves over a Lorentzian mountain as seen in the pressure field, and as generated with buoyancy frequency,  $N = 0.005 \text{ s}^{-1}$ , reference flow speed  $U = 1 \text{ m s}^{-1}$  (flow from left to right), and topographic amplitude,  $\eta_0 = 100 \text{ m}$ . The left panel makes use of the Lorentzian mountain (10.73) with  $\ell = U/N = 200 \text{ m}$ , whereas the right panel uses  $\ell = 10U/N = 2000 \text{ m}$ . Dotted contours depict negative pressure anomalies (such as those adjacent to the mountain) and solid contours are positive anomalies. The pressure field resulting from the relatively sharp mountain (left panel) exhibits oscillatory features on the lee side of the mountain, which are referred to as *lee waves*. These solutions were generated using a trapezoidal numerical integration to compute the integrals in equation (10.83).

common for reference state stratification to be a function of the vertical,  $N_R(z)$ . Indeed, such vertical dependence was considered in Sections 9.2, 9.3, and 9.4 when studying properties of the linear Boussinesq equations before specializing to plane waves in Sections 9.5 and 9.6. In this section we make use of the WKB approximation also used for acoustic waves in Section 3.9 (see Chapter 2 for a more general discussion), here using it to study internal gravity waves in the presence of  $N_R(z)$ . Our fundamental assumption is that the background reference density, which is static, displays vertical variations over a vertical scale that is much larger than the vertical wavelength of the gravity waves. This scale separation is essential to use WKB methods.

We can anticipate the basic results of this analysis by noting that if a propagating internal gravity wave, with frequency  $\omega < N_R$ , moves into a region where  $\omega > N_R$ , then it can no longer propagate. Instead, it becomes an exponentially damped or *evanescent* wave. In this case, the propagating wave can reflect back to regions where propagation is available. Alternatively, if the region of evanescence is narrow, then the wave can tunnel through such regions.

### 10.4.1 The two length scale assumption

We introduced the WKB formalism in Section 3.9 for acoustic waves moving in a fluid with a spatially dependent equilibrium density. For internal gravity waves we consider the somewhat simpler case with vertical variations in the background density stratification, rather than the fully three dimensional variations considered for acoustic waves. We thus introduce the length scale for vertical variations in the stably stratified background density field

$$L \equiv |\partial_z \rho_R / \rho_R|^{-1} = (\rho_0 N_R^2 / (g \rho_R))^{-1}. \quad (10.84)$$

We assume that the internal gravity waves have a phase that has a vertical wavevector component that is a function of vertical position,  $k_z = k_z(z)$ , which satisfies the following scaling

$$L \gg |k_z|^{-1} \implies |k_z| \gg \rho_0 N_R^2 / (g \rho_R). \quad (10.85)$$

This assumption means that we are concerned with a vertical length scale of the waves, as measured by  $|k_z(z)|^{-1}$ , that is much smaller than the length scale,  $L$ , that is set by variations

in the vertical density stratification. Furthermore, we assume that the vertical variations of  $k_z^{-1}$  occur over the same length scale,  $L$ , so that

$$|dk_z^{-1}/dz| = |k_z L|^{-1} \ll 1. \quad (10.86)$$

In addition to having  $k_z$  now a function of vertical, we allow the wave amplitude to slowly vary with  $z$ . Writing  $A(z)$  for that amplitude we assume, as for  $k_z(z)$ , that

$$|dA/dz| = |A/L|, \quad (10.87)$$

which implies that

$$|A^{-1} d^2 A/dz^2| = L^{-2} \ll k_z^2. \quad (10.88)$$

### 10.4.2 The wave equation with the WKB ansatz

The wave equation (9.40) for the vertical velocity holds for  $N_R(z)$ , so we focus our WKB analysis on that equation, rewritten here in the form

$$(\partial_{tt} \nabla^2 + N_R^2 \nabla_h^2) w' = \partial_{tt} \partial_{zz} w' + (\partial_{tt} + N_R^2) \nabla_h^2 w' = 0. \quad (10.89)$$

Recall the plane wave ansatz (9.54) assumed for constant background density, which for the vertical velocity is given by

$$w'(\mathbf{x}, t) = \tilde{w} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} = (\tilde{w} e^{i k_z z}) e^{i(k_x x + k_y y - \omega t)} \quad \text{for constant } N_R. \quad (10.90)$$

For the case with  $N_R(z)$  we generalize this ansatz to the form

$$w'(\mathbf{x}, t) = A(z) e^{i(k_x x + k_y y + \sigma(z) - \omega t)} \quad \text{for vertically varying } N_R(z), \quad (10.91)$$

with the real amplitude,  $A(z)$ , and real phase,  $\sigma(z)$ , each of which are to be determined. For the ansatz (10.91), we took a horizontal wavevector,  $\mathbf{k}_h = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}}$ , and angular frequency,  $\omega$ , just as for the constant  $N_R$  waves. This form is motivated by the results from ray theory developed in Section 2.3. Namely, when following a ray, a wavevector component is constant if the background structure is constant in its corresponding direction, and the angular frequency is constant if the background state is time independent. Given that in the present case we are only concerned with vertical dependence of the background state, it is only the vertical wavenumber that has a vertical dependence.

Making use of the WKB wave ansatz (10.91) in the wave equation (10.89) leads to the following expressions

$$\partial_{zz} w'/w' = A^{-1} \partial_{zz} A - (d\sigma/dz)^2 + i [d^2\sigma/dz^2 + 2 A^{-1} (d\sigma/dz) (dA/dz)] \quad (10.92a)$$

$$\nabla_h^2 w'/w' = -|\mathbf{k}_h|^2 \quad (10.92b)$$

$$\partial_{tt} w'/w' = -\omega^2, \quad (10.92c)$$

which brings the wave equation (10.89) to the form

$$[\partial_{zz} + k_z^2(z)] w' = 0, \quad (10.93)$$

where we defined the vertically dependent vertical wave number

$$k_z^2(z) = \frac{|\mathbf{k}_h|^2 (N_R^2(z) - \omega^2)}{\omega^2}. \quad (10.94)$$

This expression for  $k_z(z)$  is identical to the case with a constant  $N_R$  (see equation (9.67)), only now it holds with  $N_R(z)$  a function of vertical position. Equation (10.93) generalizes equation (9.70) that holds when  $N_R$  is a spatial constant.

### 10.4.3 Wave amplitude and vertical wavenumber

Making use of equation (10.92a) for  $\partial_{zz}w'$  in equation (10.94), and setting the real and imaginary parts to zero, leads to

$$k_z^2 - (\mathrm{d}\sigma/\mathrm{d}z)^2 = A^{-1} \mathrm{d}^2A/\mathrm{d}z^2 \quad (10.95a)$$

$$\mathrm{d}^2\sigma/\mathrm{d}z^2 + 2A^{-1}(\mathrm{d}\sigma/\mathrm{d}z)(\mathrm{d}A/\mathrm{d}z) = 0. \quad (10.95b)$$

Recall the amplitude scaling (10.88), which says that  $A^{-1} \mathrm{d}^2A/\mathrm{d}z^2 \ll k_z^2$ . Equation (10.95a) thus leads to the vertical phase function

$$\mathrm{d}\sigma/\mathrm{d}z = k_z \implies \sigma(z) = \int_z^{z_0} k_z(z') \mathrm{d}z'. \quad (10.96)$$

The vertical position,  $z_0$ , results only in an overall constant and arbitrary shift in the phase, and so it is arbitrary. Equation (10.95b) can be written in the form of a total derivative

$$\frac{\mathrm{d}}{\mathrm{d}z} \left[ A \left| \frac{\mathrm{d}\sigma}{\mathrm{d}z} \right|^{1/2} \right] = \frac{\mathrm{d}(A |k_z|^{1/2})}{\mathrm{d}z} = 0, \quad (10.97)$$

which means that the wave amplitude is related to the vertical wavenumber via

$$A \propto |\mathrm{d}\sigma/\mathrm{d}z|^{-1/2} = |k_z|^{-1/2}. \quad (10.98)$$

To help interpret the relation (10.98), we make use of Exercise 9.5, in which we show that the vertical component to the energy flux within an internal gravity wave can be written as

$$\langle \mathbf{v}' \varphi' \rangle \cdot \hat{\mathbf{z}} = -\frac{1}{2} A^2 k_z \omega / |\mathbf{k}_h|^2. \quad (10.99)$$

As noted above, the waves are propagating through the depth dependent stratification with a constant angular frequency and constant horizontal wavenumber. Since the reference state is time independent, the mechanical energy remains a constant following a ray. In order to avoid energy converging to any particular vertical position (which would then lead to energy changing at that position), we must have

$$\partial_z [\langle \mathbf{v}' \varphi' \rangle \cdot \hat{\mathbf{z}}] = 0 \implies \frac{\mathrm{d}(A^2 k_z)}{\mathrm{d}z} = 0. \quad (10.100)$$

Hence, the relation (10.98) is a consequence of mechanical energy conservation for the internal gravity waves propagating through a reference state with depth dependent stratification.

#### 10.4.4 WKB solution for the vertical velocity

Bringing the pieces together leads to the WKB solution for the vertical velocity component

$$w'(\mathbf{x}, t) = C |k_z|^{-1/2} e^{i(k_x x + k_y y + \sigma(z) - \omega t)}, \quad (10.101)$$

where we wrote the amplitude as

$$A = C |k_z|^{-1/2}, \quad (10.102)$$

with  $C$  a constant real constant with dimensions  $L^{1/2} T^{-1}$ . We are thus led to the WKB solution

$$w'(\mathbf{x}, t) = C |k_z|^{-1/2} \cos(k_x x + k_y y + \sigma(z) - \omega t) \quad (10.103a)$$

$$\sigma(z) = \int_z^{z_0} k_z(z') dz' \quad (10.103b)$$

$$k_z^2(z) = (|\mathbf{k}_h|/\omega)^2 (N_R^2(z) - \omega^2) \quad (10.103c)$$

$$\omega^2 = N_R^2(z) \cos^2 \gamma(z) = N_R^2(z) |\mathbf{k}_h|^2 / [|\mathbf{k}_h|^2 + k_z^2(z)] \quad (10.103d)$$

$$C [\equiv] L^{1/2} T^{-1}. \quad (10.103e)$$

In equation (10.103c) we exposed its  $z$  dependence in the vertical wavenumber,  $k_z(z)$ , along with that for the buoyancy frequency,  $N_R(z)$ . Importantly, recall that the angular frequency,  $\omega$ , and the horizontal wavevector,  $\mathbf{k}_h$ , are both spatially independent. Consequently, equation (10.103d) means that the angle made by the phase velocity relative to the horizontal,  $\gamma(z)$ , must change as the buoyancy frequency changes so that  $\omega^2 = N_R^2(z) \cos^2 \gamma(z)$  constant. Furthermore, note that the vertical wavenumber becomes imaginary in regions where  $\omega^2 > N_R^2(z)$ . So if a propagating wave enters a region of relatively low vertical stratification, then the wave becomes *evanescent*, which means that the wave exponentially decays when moving into the region. This behavior is exemplified by the wave guide discussed in Section 10.4.6. Conversely, if a wave enters a region of relatively large vertical stratification, the vertical wavenumber increases according to equation (10.103c). Correspondingly, the wave amplitude decreases according to equation (10.98).

#### 10.4.5 Structure of an internal gravity wave

The WKB expressions (10.103a)-(10.103e) are sufficient to determine the structure of an internal gravity wave moving through a region with non-constant background stratification. To do so requires the linear relations from Section 9.4, similar to what we did in Section 9.5 for plane waves with  $N_R$  constant. We start with equation (9.17c) that relates buoyancy and vertical velocity via

$$\partial_t b' = -w' N_R^2. \quad (10.104)$$

Time integrating the vertical velocity in equation (10.103a) leads to the buoyancy field

$$b' = i C N_R^2 / (\omega |k_z|^{1/2}) e^{i(k_h \cdot \mathbf{x} + \sigma(z) - \omega t)}, \quad (10.105)$$

where

$$\mathbf{k}_h \cdot \mathbf{x} = k_x x + k_y y. \quad (10.106)$$

For the horizontal velocity we use the non-divergence condition,  $\nabla \cdot \mathbf{v}' = 0$  (Section 9.5.2), as well as the vanishing vertical component to the relative vorticity,  $\partial_x v' = \partial_y u'$  (Section 9.5.3),

to find

$$\mathbf{u}' = -\frac{C \mathbf{k}_h}{|2 k_z|^{1/2} |\mathbf{k}_h|^2} [(k_z^{-1} \partial_z k_z) \sin(\mathbf{k}_h \cdot \mathbf{x} + \sigma(z) - \omega t) + 2 k_z \cos(\mathbf{k}_h \cdot \mathbf{x} + \sigma(z) - \omega t)]. \quad (10.107)$$

#### 10.4.6 Wave packets within a wave guide

We here highlight a particularly special behavior of internal gravity waves moving in a region with vertically varying stratification. This behavior results from having the angular frequency remain constant within the dispersion relation (10.103d). Again, the constancy of the angular frequency follows from our study of ray theory in Section 2.3, with the angular frequency of a wave a space-time constant if the base state is static.

##### **Wavenumber changes when moving through regions with changing $N_R$**

Consider a wave packet that moves vertically into a region where  $N_R(z)$  is decreasing. To keep the angular frequency fixed,  $\cos \gamma(z)$  must increase, which then requires the magnitude of  $\gamma(z)$  to decrease. A decreasing  $\gamma$  means that the phase velocity becomes more horizontal through decreasing the vertical wavenumber, which in turn means that the group velocity (equation (9.71)),

$$\mathbf{c}_g = \frac{N_R}{|\mathbf{k}|^3 |\mathbf{k}_h|} (k_z^2 \mathbf{k}_h - k_z |\mathbf{k}_h|^2 \hat{\mathbf{z}}), \quad (10.108)$$

becomes more vertical. The opposite happens for a packet moving into a region of enhanced vertical stratification, in which the magnitude of  $\gamma(z)$  increases toward  $\pi/2$  so that the group velocity becomes more horizontal. So in summary, the trajectory of a gravity wave packet becomes more horizontal as the packet moves into a region of stronger vertical stratification, whereas the trajectory becomes more vertical as the packet moves into a region of low stratification. These changes to the propagation angle are essential features of the wave packets move through a wave guide.

##### **The cusp-like trajectory of a packet entering a region where $\omega^2 > N_R^2$**

As wave packet becomes evanescent as it enters a region with  $\omega^2 > N_R^2(z)$ . At the level where  $\omega^2 = N_R^2(z)$ , the angle,  $\gamma$ , vanishes, so that the group velocity (10.108) also vanishes since  $k_z = 0$ . What does the ray path look like local to this region? To anticipate the answer, note that as a packet moves to a region of low vertical stratification, in which  $|k_z| \rightarrow 0$ , the horizontal component to the group velocity vanishes before the vertical component. Hence, as the packet approaches the region of weak stratification, its ray forms a vertical cusp as the packet stalls. Since the packet cannot propagate into a region with  $\omega^2 > N_R^2$ , then the wave reflects back into the region where  $\omega^2 < N_R^2$ . If the propagation region with  $\omega^2 < N_R^2$  sits between two regions with weak stratification with  $\omega^2 > N_R^2$ , then the bounded relatively high stratification region forms a *wave guide* for the internal gravity waves. We depict an idealized example of a wave guide in Figure 10.5.

We can compute an analytic expression for the wave packet ray as it approaches a *turning level*, at  $z = z_t$ , which is defined as the level where

$$N_R(z = z_t) = \omega \implies \text{turning level.} \quad (10.109)$$

The ray equations are given by

$$\frac{dx}{dt} = \mathbf{c}_g \cdot \hat{\mathbf{x}} = \frac{N_R k_z^2 k_x}{|\mathbf{k}|^3 |\mathbf{k}_h|} \quad (10.110a)$$

$$\frac{dz}{dt} = \mathbf{c}_g \cdot \hat{\mathbf{z}} = -\frac{N_R k_z |\mathbf{k}_h|^2}{|\mathbf{k}|^3 |\mathbf{k}_h|}. \quad (10.110b)$$

Assuming the packet is moving in the  $x$ - $z$  plane with  $k_x > 0$  leads to the trajectory equation

$$\frac{dz}{dx} = -\frac{|\mathbf{k}_h|^2}{k_x k_z} = -\frac{k_x}{k_z} = -\frac{\omega}{\sqrt{N_R^2 - \omega^2}}, \quad (10.111)$$

where the second equality made use of the dispersion relation (10.103c). So as already determined, the ray becomes vertical as the vertical wavenumber vanishes.

We now solve this trajectory equation near the turning level for a packet that is moving upward towards the turning level. For this purpose, take a Taylor series for the squared buoyancy frequency for points close to the turning level

$$N_R^2(z) \approx N_R^2(z_t) + (dN_R/dz)(z - z_t) = \omega^2 - |dN_R/dz|(z - z_t) = \omega^2 + |dN_R/dz|(z_t - z), \quad (10.112)$$

where  $N_R^2$  is a decreasing function as  $z$  increases towards  $z_t$ , which means that  $dN_R^2/dz = -|dN_R^2/dz|$ . Making use of this expression for  $N_R^2$  within the trajectory equation (10.111) leads to

$$\frac{dz}{dx} = -\frac{\omega}{\sqrt{|dN_R^2/dz|(z_t - z)}}. \quad (10.113)$$

We integrate along the ray as the packet moves from a point  $(x, z)$  to the turning level at  $(x_t, z_t)$  which yields the expression for the ray trajectory

$$(z_t - z)^{3/2} = \frac{(3\omega/2)(x_t - x)}{\sqrt{|dN_R^2/dz|}} \implies z_t - z = \left[ \frac{(3\omega/2)(x_t - x)}{\sqrt{|dN_R^2/dz|}} \right]^{2/3}. \quad (10.114)$$

We depict this trajectory in Figure 10.5 for a prototypical wave guide in which internal gravity waves propagate between two weakly stratified regions.

#### 10.4.7 Stretched vertical coordinate for hydrostatic gravity waves

Recall that for internal gravity waves, the velocity of a fluid particle is aligned with the group velocity (9.71). The group velocity reveals that waves maintaining the wavevector anisotropy

$$k_z^2 \gg |\mathbf{k}_h|^2, \quad (10.115)$$

have horizontal fluid particle displacements that are far larger than their vertical displacements, which means that such motions satisfy hydrostatic scalings (Sections ?? and ??). Consequently, such wavevectors characterize hydrostatic internal gravity waves, and the dispersion relation for these waves takes on the hydrostatic form

$$(|\mathbf{k}_h|^2 + k_z^2) \omega^2 = |\mathbf{k}_h|^2 N_R^2 \xrightarrow{k_z^2 \gg |\mathbf{k}_h|^2} k_z^2 \omega^2 \approx |\mathbf{k}_h|^2 N_R^2. \quad (10.116)$$

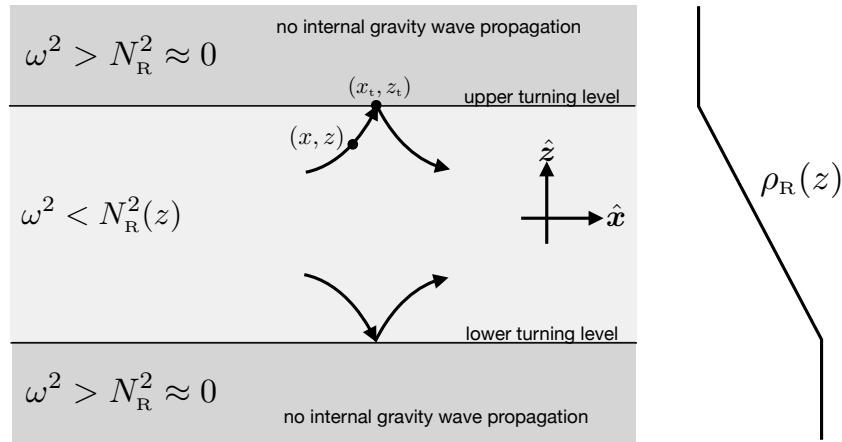


FIGURE 10.5: A waveguide for internal gravity waves, here depicted by two low stratification regions ( $N_R^2 \approx 0$ ) bounding a higher stratification region ( $N_R^2 > 0$ ). The density profile on the right provides an idealized realization of such a guide. The levels where  $N_R = \omega$  are referred to as *turning levels*. We depict sample rays approaching and leaving the turning levels, with the rays computed from the group velocity and dispersion relation. For example, a ray approaching the upper turning is computed from  $(x, t)$  to  $(x_t, z_t)$ , with the cusp-like trajectory given by equation (10.114).

Notably, with  $\mathbf{k}_h$  and  $\omega$  independent of depth, then the vertical component to the wavevector in a hydrostatic internal gravity wave equals to the buoyancy frequency times a constant,

$$k_z^2(z) = N_R^2(z) (|\mathbf{k}_h|^2 / \omega^2). \quad (10.117)$$

The hydrostatic vertical wavenumber (10.117) brings the vertical phase function (10.103b) to the form

$$\sigma(z) = \int_z^{z_0} k_z(z') dz' = \pm(|\mathbf{k}_h|/\omega) \int_z^{z_0} N_R(z') dz' \equiv \pm(|\mathbf{k}_h| N_R^*/\omega) z^* \equiv k_z^* z^*. \quad (10.118)$$

We here introduced the *WKB stretched vertical coordinate* and corresponding vertical wavenumber

$$z^* \equiv \frac{1}{N_R^*} \int_z^{z_0} N_R(z') dz' \quad \text{and} \quad k_z^* = \pm(|\mathbf{k}_h|/\omega) N_R^* \implies k_z = k_z^* N_R/N_R^*, \quad (10.119)$$

with  $N_R^*$  a constant buoyancy frequency that is commonly taken as the maximum value over the domain. This coordinate transformation is one-to-one since we assume that  $N_R(z) > 0$ , and it brings the vertical velocity (10.103a) to the form

$$w'(x, t) = C \sqrt{\frac{N_R^*}{N_R |k_z^*|}} \cos(k_x x + k_y y + k_z^* z^* - \omega t), \quad (10.120)$$

where the phase accords with waves moving through a reference state with constant buoyancy frequency. As noted on page 299 of [Gill \(1982\)](#), the stretched vertical coordinate is quite useful in the analysis of gravity wave spectra, for example, with the stretching allowing one to use insights from constant  $N_R$  to non-uniform  $N_R(z)$ .

#### 10.4.8 Comments and further study

See Section 8.12.1 of [Gill \(1982\)](#), Lecture 9 of [Pedlosky \(2003\)](#), Sections 3.5 and 3.6 of [Sutherland \(2010\)](#), and Section 7.5 of [Vallis \(2017\)](#) for similar presentations.

## 10.5 Vertically standing waves in a bounded domain

We now specialize the study from Section 10.4 to examine gravity waves in a domain that is bounded above and below. The lower boundary is assumed to be flat and rigid at  $z = -H$ . The upper boundary at  $z = \eta$  is generally free, thus supporting motions corresponding to barotropic waves and/or surface waves. However, we are concerned with the slower baroclinic (internal) modes, whose motion leaves a relatively tiny signature on free surface undulations. Hence, to focus on the baroclinic modes it is suitable to take the rigid lid boundary condition imposed at  $z = 0$ .

### 10.5.1 Formulation of the boundary value problem

We are interested in free waves in the horizontal, while the vertical is bounded above and below thus motivating the ansatz

$$w' = W(z) e^{i(k_x x + k_y y - \omega t)}. \quad (10.121)$$

Plugging this ansatz into the vertical velocity wave equation (9.40) leads to the boundary value problem

$$\frac{d^2W}{dz^2} + k_z^2 W = 0 \quad -H < z < 0 \quad (10.122a)$$

$$W = 0 \quad \text{at } z = -H \quad (10.122b)$$

$$\omega^2 dW/dz = g |\mathbf{k}_h|^2 W \quad \text{at } z = 0. \quad (10.122c)$$

In this equation we wrote the vertical wavenumber as

$$k_z^2(z) = \frac{|\mathbf{k}_h|^2 (N_R^2(z) - \omega^2)}{\omega^2}, \quad (10.123)$$

with oscillatory behavior for  $k_z^2 > 0$  and exponential behavior for  $k_z^2 < 0$ . We used the flat bottom kinematic boundary condition at  $z = -H$ . For the free surface we used the upper ocean boundary condition (9.53), thus rendering the boundary condition (10.122c). The vertical wavenumber has a depth dependence that arises from depth dependence of the buoyancy frequency. In contrast, the angular frequency remains depth independent, which follows since the reference state is time independent and so the angular frequency is a constant. We already encountered this property of the angular frequency in Section 10.4, as follows from the ray theory in Section 2.3.

### 10.5.2 Rigid lid boundary condition

The baroclinic motions of interest here have angular frequencies that are far smaller than  $N_R$ , in which case we can make use of the rigid lid approximation,  $W = 0$  at  $z = 0$ . Consequently, we are led to examine the following boundary value problem with a homogeneous Dirichlet boundary condition at both the surface ( $z = 0$ ) and bottom ( $z = -H$ )

$$\frac{d^2W}{dz^2} + k_z^2 W = 0 \quad -H < z < 0 \quad (10.124a)$$

$$W = 0 \quad \text{at } z = -H, 0. \quad (10.124b)$$

We offer comments on the free surface boundary condition in Section 10.5.7, which proves to be far more subtle mathematically than the rigid lid case.

### 10.5.3 Distinction between $\omega^2 < N_R^2$ and $\omega^2 > N_R^2$

The vertical wavenumber is real,  $k_z^2 > 0$ , in regions where the angular frequency is smaller than the buoyancy frequency,  $\omega^2 < N_R^2$ . In such regions, the solution to equation (10.124a) is oscillatory and, with specific discrete values of  $k_z$ , can satisfy the two boundary conditions. In contrast, the vertical wavenumber is imaginary,  $k_z^2 < 0$ , in regions where the angular frequency is larger than the buoyancy frequency,  $\omega^2 > N_R^2$ . In such regions, the solution to equation (10.124a) takes on a real exponential behavior. For example, if the buoyancy frequency is constant, then satisfying the  $z = -H$  boundary condition with  $\omega^2 > N_R^2$  requires  $W = A \sinh[k_z(z + H)]$ . Yet to also satisfy the  $z = 0$  boundary condition requires  $A \sinh(k_z H) = 0$ , which only has the trivial solution,  $A = 0$ .

To determine a general condition for a nontrivial solution, multiply the differential equation (10.124a) by  $W$  and vertically integrate over the domain. Use of the Dirichlet boundary conditions (10.124b) readily leads to the integral condition

$$\frac{|\mathbf{k}_h|^2}{\omega_m^2} \int_{-H}^0 W^2 (N_R^2 - \omega^2) dz = \int_{-H}^0 (dW/dz)^2 dz. \quad (10.125)$$

Assuming  $W$  is real, this equation has no solution if  $\omega^2 > N_R^2$  everywhere in the domain. However, it is possible for  $\omega^2 > N_R^2$  to occur in just a subset of the domain, in which case there is tunneling, refraction, and reflection across the turning levels where  $\omega^2 = N_R^2$ .

### 10.5.4 Sturm-Liouville eigenvalue problem

For the case of flat and rigid top and bottom boundaries with  $\omega^2 < N_R^2$  throughout the domain, the vertical boundary value problem (10.124a)-(10.124b) constitutes a *Sturm-Liouville* eigenvalue problem on the bounded domain  $-H \leq z \leq 0$ . The eigenfunction solutions form a countably infinite set of eigenmodes that provide a complete basis for square integrable functions that satisfy the homogeneous Dirichlet boundary conditions.<sup>6</sup> Consequently, any such function can be synthesized (i.e., represented) as a (generally infinite) sum of discrete eigenmodes. As seen below, the special case of constant  $N_R$  has sinusoidal eigenmodes, so that the sum constitutes the Fourier sine series studied in Section ??.

The quantized vertical eigenvalues lead to quantized angular eigenfrequencies, which we write as

$$\omega_m^2 = \frac{N_R^2 |\mathbf{k}_h|^2}{|\mathbf{k}_h|^2 + k_{zm}^2}, \quad (10.126)$$

where  $m$  is an integer that labels the discrete frequencies and vertical wavenumbers. To understand the basic mathematical properties of these eigenfrequencies, and associated eigenmodes, it is useful to consider a constant  $N_R$  with  $\omega^2 < N_R^2$ . For this case we have the discrete eigenmodes and corresponding eigenvalues

$$W_m = A \sin(k_{zm} z) \quad \text{with } k_{zm} = m \pi / H \text{ and } m = 1, 2, \dots \quad (10.127)$$

<sup>6</sup>The case of a free surface boundary is more mathematically subtle since it does *not* constitute a Sturm-Liouville system. We provide further comment on this case in Section (10.5.7).

The quantized vertical wavenumbers lead to the quantized angular eigenfrequencies

$$\omega_m^2 = \frac{N_R^2 |\mathbf{k}_h|^2}{|\mathbf{k}_h|^2 + (m\pi/H)^2}. \quad (10.128)$$

We do not consider  $m = 0$  since this value leads to the trivial eigenmode,  $W_{m=0} = 0$ . Additionally,  $W_{-m} = -W_m$ , so that the negative eigenmodes are not independent of the positive eigenmodes.

### 10.5.5 Orthogonality of the eigenmodes

To prove that the eigenmodes are orthogonal, consider the differential equation (10.124a) for two eigenfrequencies,  $\omega_m$  and  $\omega_n$ , with corresponding eigenmodes,  $W_m$  and  $W_n$ ,

$$\frac{d^2 W_m}{dz^2} + \frac{|\mathbf{k}_h|^2 W_m (N_R^2 - \omega_m^2)}{\omega_m^2} = 0 \quad \text{with} \quad W_m = 0 \quad \text{at} \quad z = 0, -H \quad (10.129a)$$

$$\frac{d^2 W_n}{dz^2} + \frac{|\mathbf{k}_h|^2 W_n (N_R^2 - \omega_n^2)}{\omega_n^2} = 0 \quad \text{with} \quad W_n = 0 \quad \text{at} \quad z = 0, -H. \quad (10.129b)$$

Multiply the  $m$  equation by  $W_n$  and the  $n$  equation by  $W_m$ , then subtract to find

$$\frac{d}{dz} \left[ W_n \frac{dW_m}{dz} - W_m \frac{dW_n}{dz} \right] + W_m W_n |\mathbf{k}_h|^2 N_R^2 (1/\omega_m^2 - 1/\omega_n^2) = 0. \quad (10.130)$$

Vertical integration with use of the homogeneous Dirichlet boundary conditions eliminates the derivative term, thus leaving

$$(1/\omega_m^2 - 1/\omega_n^2) \int_{-H}^0 W_m W_n N_R^2 dz = 0, \quad (10.131)$$

where we noted that the horizontal wavenumber,  $\mathbf{k}_h$ , and angular eigenfrequencies,  $\omega_m$  and  $\omega_n$ , are independent of  $z$  and so can be pulled outside of the integral. We thus find that the eigenmodes with  $m \neq n$  are orthogonal with a weighting function given by the squared buoyancy frequency

$$\int_{-H}^0 W_m W_n N_R^2 dz = 0 \quad m \neq n. \quad (10.132)$$

As shown in Figure 10.6, the presence of  $N_R^2$  as part of the normalization acts to scale the magnitude of the normalized eigenmodes.

### 10.5.6 Non-uniform vertical stratification

We already noted the sinusoidal eigenmodes (10.127) for the case of constant stratification. The case of non-constant vertical stratification is more common in realistic geophysical fluids. For general stratification profiles it is necessary employ numerical methods to determine the eigenfunctions and eigenfrequencies. However, so long as the stratification is gently varying, we can make use of the WKB results from Section 10.4.4 to write analytical expressions for the approximate eigenfunctions with  $N_R(z)$

$$W_m(z) = C |k_{zm}|^{-1/2} e^{i\sigma_m(z)} \quad (10.133a)$$

$$\sigma_m(z) = -\pi/2 + \int_z^0 k_{zm}(z') dz' \equiv -\pi/2 + \gamma_m(z) \quad (10.133b)$$

$$k_{zm}(z) = (|\mathbf{k}_h|/\omega_m) \sqrt{N_R^2(z) - \omega_m^2} \quad (10.133c)$$

$$\omega_m = N_R(z) |\mathbf{k}_h|^2 / [|\mathbf{k}_h|^2 + k_{zm}^2(z)]. \quad (10.133d)$$

We set the phase function,  $\sigma_m$ , to bring the real part of the vertical mode to the form

$$W_m(z) = C |k_{zm}|^{-1/2} \sin[\gamma_m(z)]. \quad (10.134)$$

Setting the upper bound to the integral at  $z_0 = 0$  ensures that  $W_m(z=0) = 0$ . To satisfy the lower boundary condition requires

$$W_m(-H) = 0 \implies \gamma_m(-H) = \int_{-H}^0 k_{zm}(z') dz' = m \pi. \quad (10.135)$$

For the special case of hydrostatic gravity waves, we follow the discussion in Section 10.4.7 by introducing the WKB stretched vertical coordinate (10.119),

$$z^*(z) = \frac{1}{N_R^*} \int_z^0 N_R(z') dz', \quad (10.136)$$

so that the phase function is

$$\gamma_m(z) = \int_z^0 k_{zm}(z') dz' = \frac{|\mathbf{k}_h|}{\omega_m} \int_z^0 N_R(z') dz' = k_{zm}^* z^*(z). \quad (10.137)$$

To satisfy the boundary conditions we must quantize the vertical wavenumber as

$$k_{zm}^* = m \pi / |z^*(-H)|, \quad (10.138)$$

in which case the complete set of vertical hydrostatic eigenmodes (within the WKB approximation) is given by

$$W_m(z) = C |k_{zm}(z)|^{-1/2} \sin \left[ m \pi \frac{\int_z^0 N_R(z') dz'}{\int_{-H}^0 N_R(z') dz'} \right] \quad (10.139a)$$

$$k_{zm}(z) = m \pi \frac{N_R(z)}{\int_{-H}^0 N_R(z') dz'}. \quad (10.139b)$$

In Figure 10.6 we show two sets of eigenmodes, one for a constant buoyancy frequency and one for an exponential profile.

### 10.5.7 A complete set of eigenmodes with a free surface boundary

*Boyce and DiPrima* (2009), and many other mathematical physics texts, present the theory of linear self-adjoint operators that form a Sturm-Liouville system. We made use of these results for the rigid lid problem, in particular by noting that the countably infinite set of eigenmodes forms a complete basis for functions satisfying the homogeneous Dirichlet conditions at the top and bottom boundaries. However, the free surface case (10.122a)-(10.122c) does not form a

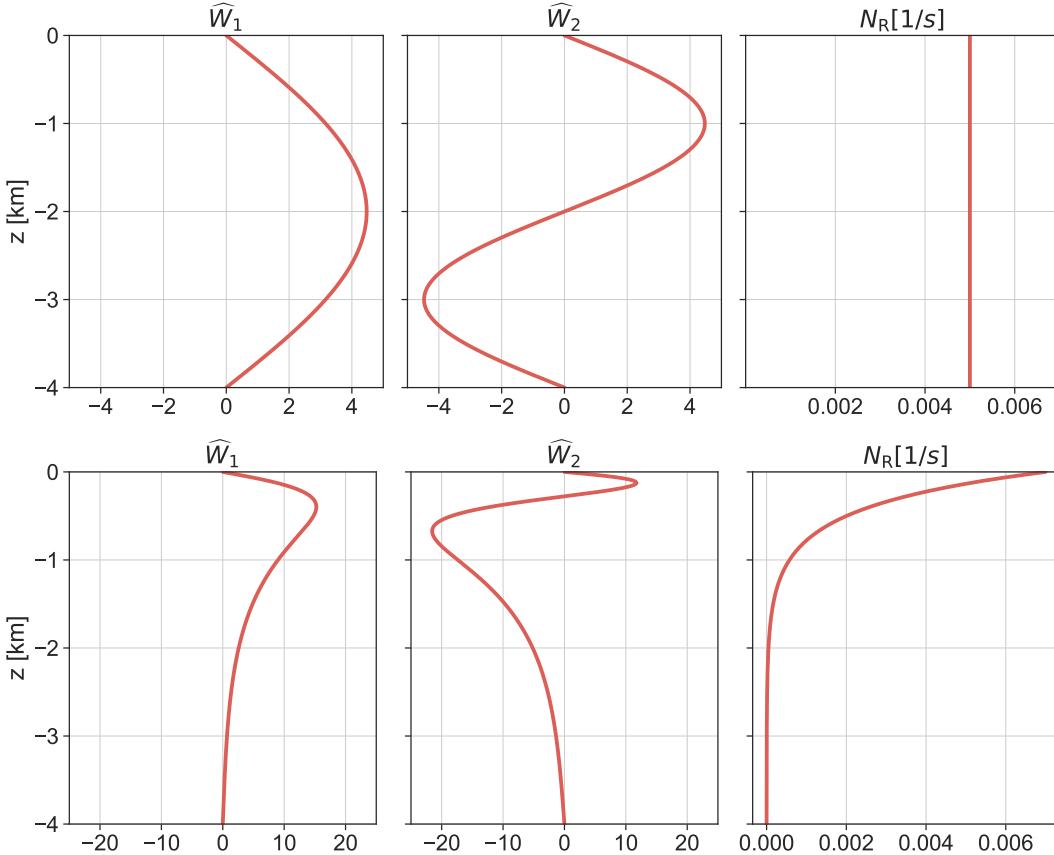


FIGURE 10.6: Two sets of normalized gravity wave eigenmodes,  $\widehat{W}_m$ , determined by normalizing the WKB vertical velocity eigenmodes,  $W_m$ , in equation (10.139a) according to equation (10.132),  $\int_{-H}^0 N_R^2 W_m^2 dz = 1$ . We plot the  $m = 1$  and  $m = 2$  modes along with the corresponding buoyancy frequency. The first set of eigenmodes arises from a constant buoyancy frequency, in which case the eigenmodes are the sinusoidal functions (10.127). The second set is for an exponential buoyancy frequency,  $N_R = (g \Delta) / (\delta H) e^{(z / (\delta H))}$ , with  $\delta = 0.05$ ,  $\Delta = 0.001$ , and  $H = 4000$  m. Note how the eigenmodes for the exponential buoyancy frequency are preferentially weighted in the upper ocean, where the buoyancy frequency is larger. Also note how the buoyancy frequency acts to scale the relative magnitude of the two sets of eigenmodes in order to ensure normalization holds with  $N_R^2$  as part of the normalization integral.

Sturm-Liouville system. To see why, it is useful to introduce the shorthand

$$\lambda = |\mathbf{k}_h|^2 / \omega^2, \quad (10.140)$$

in which the boundary value problem (10.122a)-(10.122c) takes on the equivalent form

$$-\frac{d^2 W}{dz^2} + |\mathbf{k}_h|^2 W = \lambda N_R^2 W \quad -H < z < 0 \quad (10.141a)$$

$$W = 0 \quad \text{at } z = -H \quad (10.141b)$$

$$\frac{dW}{dz} = g \lambda W \quad \text{at } z = 0. \quad (10.141c)$$

Notice how the eigenvalue,  $\lambda$ , appears in both the differential equation (10.141a) and the surface boundary condition (10.141c). Sturm-Liouville boundary value problems only have the eigenvalue appear in the differential equation. So the presence of the eigenvalue in the boundary condition means the boundary value problem (10.141a)-(10.141c) is *not* of the Sturm-

Liouville form. Remarkably, it was only the relatively recent work of [Yassin \(2021\)](#) who provided a mathematical theory for such non Sturm-Liouville eigenvalue problems associated with dynamical boundaries. Namely, [Yassin \(2021\)](#) generalized the Sturm-Liouville theory to enable construction of a complete set of orthogonal eigenmodes that includes both the interior baroclinic modes and the surface mode associated with boundary dynamics. [Yassin \(2021\)](#) and [Yassin and Griffies \(2022\)](#) also applied the mathematical theory to the coupling of *surface quasi-geostrophic* modes to interior baroclinic quasi-geostrophic modes (see Section ?? for an introduction to surface quasi-geostrophy).

### 10.5.8 Further study

Our presentation is consistent with Section 6.10 of [Gill \(1982\)](#), who also discusses the case of single lower boundary and no boundary at the top, which is relevant to the atmosphere. We also followed Lecture 9 of [Pedlosky \(2003\)](#), who also discusses the case of a free surface though in a heuristic manner that does not prove the completeness theorems of [Yassin \(2021\)](#).



## 10.6 Exercises

### EXERCISE 10.1: POLARIZATION RELATIONS FOR STATIONARY INERTIA-GRAVITY WAVES ON A MEAN FLOW

Following the discussion in Section 10.2.6, determine the polarization relations for stationary inertia-gravity waves in the presence of a constant reference flow,  $\mathbf{u}_R = U \hat{x}$  and with  $\omega_R = -\mathbf{u}_R \cdot \mathbf{k}_b > 0$ . Assume  $f^2 < \omega_R^2 < N_R^2$  so that the waves extend throughout the vertical. Write expressions for  $u'$ ,  $v'$ ,  $w'$ ,  $b'$ , and  $\varphi'$  in terms of the vertical velocity amplitude,  $\tilde{w}$ , as given by equation (10.39). Confirm that setting  $f = 0$  reduces the inertia-gravity results to the gravity wave case derived in Section 10.2.7.

### EXERCISE 10.2: MOUNTAIN GRAVITY WAVE FROUDE NUMBER BASED ON FREQUENCIES

Consider the non-dimensional *Froude number* for forced mountain gravity waves ( $f = 0$ )

$$\text{Fr} = \frac{|\mathbf{u}_R \cdot \mathbf{k}_b|}{N_R} = \frac{\omega_R}{N_R}, \quad (10.142)$$

which is the ratio of the forcing frequency (from flow over the monochromatic mountains) to the buoyancy frequency.

- Write the squared vertical wavenumber (10.42) for mountain gravity waves in terms of Fr.
- Write Fr as the ratio of two time scales. Provide an interpretation for why mountain waves are vertically trapped when one of those time scales is greater than the other, and vertically propagating for the alternative.

Hint: see Section 5.4.2 of [Sutherland \(2010\)](#) for a discussion of this non-dimensional number. The second part of this exercise is mostly meant to be answered in words.

### EXERCISE 10.3: MECHANICAL ENERGY OF A TRAPPED MOUNTAIN GRAVITY WAVE

We studied the energetics of propagating mountain gravity waves in Section 10.2.7. Here we consider the energetics of trapped mountain waves.

- Compute the mechanical energy for the trapped mountain gravity waves from Section 10.2.8.

- (b) Compute the phase averaged mechanical energy.
- (c) Evaluate the phase averaged mechanical energy at the ground (approximated by  $z = 0$  in the linear theory), and compare this energy to the phase averaged mechanical energy in a propagating mountain wave.

## EXERCISE 10.4: GENERATION OF NEAR-INERTIAL WAVES

Consider a linear Boussinesq ocean in a rotating reference frame with constant Coriolis parameter. Ignore the role of pressure gradients, internal friction, and vertical motion so that the equations of motion in an unstratified mixing layer are given by

$$\partial_t \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{u} = \mathbf{F}, \quad (10.143)$$

where we introduced acceleration from the wind stress,

$$\mathbf{F} = \boldsymbol{\tau}_{\text{wind}} / \rho_0. \quad (10.144)$$

The steady Ekman flow is given by

$$f \hat{\mathbf{z}} \times \mathbf{u}_e = \mathbf{F}, \quad (10.145)$$

which was studied in Chapter ???. Here we are interested in the transient solution that arises in the presence of a constant wind stress.

- (a) Write the solution to equation (10.143) with a space-time constant  $\tau_{\text{wind}}$ . Hint: first solve for a complex horizontal velocity (not to be confused with the vertical velocity)

$$\underline{\mathbf{u}} = u + iv. \quad (10.146)$$

Discuss the physics of the velocity field.

- (b) Discuss the case with wind forcing that has a particular horizontal scale,  $L_w$ , set by the atmospheric forcing, and an ocean with vertical stratification given by the buoyancy frequency,  $N$ . In particular, what is the typical angular frequency of the linear gravity wave motions? Hint: read Section 9.8.2.





## **Part II**

# **Flow instabilities**

Part I of this book is concerned with small amplitude wave fluctuations that move on the background of a prescribed and unchanging exact solution to the fluid equations. For example, in studying acoustic waves in Chapter 3, the background state is static and uniform, whereas when studying internal gravity waves in Chapter 9, the background is static yet stably stratified according to gravity. In each of these cases, the background state is stable to small perturbations so that the wave fluctuations do not lead to waves of growing amplitude. In Part II of the book we examine conditions required to ensure stability of background flow states, or conversely, what properties of the background state are necessary (and sometimes sufficient) to support instabilities.

### What enables an instability?

As noted in Section 2 of [Drazin and Reid \(2004\)](#), flow instabilities occur when a disturbance leads to the disequilibrium of forces within the fluid, and with this disequilibrium enabling the disturbance to grow. We explore numerous mechanisms for the growth of disturbances that manifest positive feedbacks energized by the background state. Examples include the kinetic energy that supports shear instability, the potential energy that supports baroclinic instability, and the angular momentum that supports centrifugal instability. A central goal of instability theory is to develop a physical understanding and mathematical recipe to determine what constitutes an unstable flow state. Furthermore, a flow state might be unstable, but not all disturbances initiate the instability. This situation motivates the study of conditions required of the disturbance for it to manifest the instability.

We focus on mechanisms that lead to a variety of geophysical flow instabilities, and in so doing develop associated mathematical analysis methods. Instabilities are the typical situation for geophysical flows rather than the exception, and they play a fundamental role in establishing properties of planetary circulations in both the atmosphere and ocean. Furthermore, many instabilities occurring in geophysical fluids are primarily understood in the absence of viscous dissipation, thus prompting us to focus exclusively on the inviscid equations.

### Distinguishing two classes of flow instabilities

We distinguish two general classes of fluid flow instabilities: *local* (or *parcel*) instabilities and *global* (or *wave*) instabilities. Local/parcel instabilities afford a local necessary and sufficient condition to determine whether the background flow state is unstable. For example, in the study of gravitational instability (in the absence of surface tension), we can measure the local squared buoyancy frequency: if  $N^2 < 0$  then the fluid is gravitationally unstable at the location where  $N^2 < 0$ . The study of symmetric instabilities in Chapter 11 explores a broader class of local instabilities, in which angular momentum and potential vorticity are central to determining stability conditions. Stability conditions are generally derived by considering an imagined displacement of a fluid parcel within the environment of the prescribed background state. If the forces acting on the parcel lead to furthering the displacement, then that signals a positive feedback indicating that the background state is unstable. The result of such local instabilities is a catastrophic breakdown of the background state without concern for any special space and/or time scale for the breakdown.

Global/wave instabilities arise from the interference of waves that lead to mutually constructive positive feedbacks, thus constituting the *wave resonance* interpretation of such instabilities. The stability analysis is generally referred to as the *method of normal modes*. We do not seek the most general solution to the linearized equations. Instead we ask if a single wave can

go unstable, and if so then what properties of such waves enable the instability to manifest? We are generally concerned with instabilities growing in time, with exponential growth of an unstable mode revealed by a complex wave angular frequency. Determining properties of the unstable waves (e.g., most unstable wavelength, wave velocity, and growth rate) requires solving an eigenvalue problem where boundary conditions often play a fundamental role. Shear instabilities of Chapter 13 and baroclinic instabilities of Chapter 14 are canonical examples of this sort of instability. Operationally, the means to reveal a global instability is similar to that used to study linear waves in Part I of this book. Namely, we determine a dispersion relation for small amplitude fluctuations, only now allowing the possibility of the wave frequency or wavevector to be complex. Temporally unstable waves occur with a complex wave frequency and real wavevector, whereas spatial instabilities arise with real wave frequency and complex wavevector. Finally, observe that local instabilities can be considered wave instabilities in which all waves are unstable.

### **Flow instabilities lead to a stable end state**

Flow instabilities are energized by the background state and, through the process of creating the instability and allowing it to grow to a nonlinear stage, act to eliminate the source for the instability. For example, the catastrophic vertical motion of a gravitationally unstable fluid column serves to remove the potential energy source of the instability, with the resulting equilibrium state gravitationally stable. In geophysical fluids, there are external forces (ultimately arising from solar heating, geothermal heating, or astronomical tides) that return the background state to an unstable condition, thus allowing for the fluid to undergo a multitude of successive instabilities. Through such successive instabilities, or further secondary instabilities that feed off the primary instability, the fluid generally evolves into an extremely complex state of linear and nonlinear waves, instabilities, and turbulence. The resulting fluid flow is comprised of a wide suite of space and time scales whose mathematical description requires methods from statistics and stochastics. We do not pursue the study of wave turbulence or fully developed turbulence in this book, but do appreciate that instabilities are the central means for geophysical fluids to transition into a turbulent state. Hence, understanding the mechanisms for flow instabilities offers insights into the nature of the associated turbulence.

### **We do not examine the effects of instabilities on the background state**

Throughout this study, we do not examine evolution of the prescribed background state. Ignoring changes to the background state is a sensible assumption when studying the motion of stable linear waves, whose amplitudes remain bounded and so offer only a tiny perturbation to the background state. However, for unstable flows, the assumption of a fixed background proves untenable when fluctuations become large enough to engage flow nonlinearities and thus to nontrivially impact the background state. It is for this reason that our analysis focuses exclusively on the early stages of an instability, whereby small amplitude assumptions allow for a focus on linear mechanisms leading to growth of a disturbance, rather than enable the rich complexities of interactions (some stable and some unstable) between the disturbance and an evolving background state.

### **Seek out visualizations of flow instabilities to help develop understanding**

This video from Prof. Mollo-Christensen provides an insightful introduction to the topic of fluid mechanical instabilities. This and many other videos from laboratory studies and numerical

simulations offer useful, if not essential, visual impressions to complement the somewhat intricate mathematical analysis in this part of the book.

#### MATHEMATICS IN THIS PART

The mathematics in this chapter rely mostly on the Cartesian tensor analysis and vector calculus from VOLUME 1, along with an appreciation for the methods of linear partial differential equations.

# Chapter 11

## SYMMETRIC FLOWS

In this chapter we study the stability of fluid flows exhibiting either axial symmetry (e.g., rotating fluid column) or spatial symmetry in one horizontal direction (e.g.,  $f$ -plane geostrophic fronts). This analysis supports the study of frontal features in geophysical fluids, where a front is a region of strong buoyancy gradients that leads to a jet-like geostrophic flow. Because of the assumed symmetry in the base flow state and the perturbation, the instabilities in this chapter are generically referred to as *symmetric instabilities*. However, the terminology is not universal, with flavors of symmetric instability often referred to as *centrifugal*, *inertial*, and *symmetric*.

We make use of three methods for stability analysis in this chapter. One consists of an energetic approach due to Rayleigh (*energetic stability analysis*); one is Lagrangian and considers force balances acting on a test fluid parcel (*parcel stability analysis*); and one is Eulerian and considers wave perturbations (*modal stability analysis*). For the instabilities considered in this chapter, each method leads to the same stability condition, which is here measured by local properties of the flow so that we refer to the method as *local stability analysis*. As part of the analysis we make use of material invariants, such as angular momentum, potential momentum, buoyancy, and potential vorticity. We restrict attention to the  $f$ -plane since that allows us to make use of potential momentum as a material invariant. It is notable that the flows considered in this chapter generally exhibit secondary overturning circulations, with such circulations studied here using the rudiments of *semi-geostrophy*.

Energetic and parcel methods probe the base state flow by perturbing test fluid parcels. A test parcel is assumed to have no impact on the flow state itself, meaning that the pressure field remains unchanged even as the parcel is moved. In effect, the test fluid parcel stability analysis makes the tacit assumption that pressure responses to a perturbation can be neglected for the purpose of detecting an unstable flow state. Ignoring the impacts on pressure is consistent with the notion of a test fluid element introduced in Section ??.<sup>1</sup> However, we questioned that approach when studying buoyancy in Section ??, where we computed the pressure response to the perturbation and found that the response has an important impact on the effective buoyancy felt by a finite fluid region. Acknowledging this limitation motivates us to complement the parcel approach with an Eulerian linear stability analysis using plane waves. This approach is dynamically consistent, and yet it works within the limitations of the linearized equations.

### READER'S GUIDE FOR THIS CHAPTER

We make use of the shallow water system from Chapters ?? and ??, in particular the study of angular momentum in Section ?? (see also Section ??). We also assume an understanding

<sup>1</sup>Since the concern here is with perfect fluids, the test fluid element from Section ?? is the same as a test fluid parcel.

---

of geostrophic flow as studied in Chapter ??, as well as the continuously stratified Boussinesq ocean from Chapter ??.

Gravitational instability is the canonical local instability and it is signaled by  $N^2 < 0$ . However, we do not consider gravitational instability in this chapter, with this instability introduced in Chapter ?? when studying buoyancy, and Chapter 12 when studying the Rayleigh-Taylor instability.

<b>11.1</b>	<b>Loose threads</b>	<b>329</b>
<b>11.2</b>	<b>Instabilities in this chapter</b>	<b>329</b>
11.2.1	Summary of the instabilities studied in this chapter	329
11.2.2	Nature of the base state and the perturbations	330
11.2.3	Comments	331
<b>11.3</b>	<b>Centrifugal instability of cyclostrophic flow</b>	<b>331</b>
11.3.1	Equations of motion	332
11.3.2	Elements of angular momentum	333
11.3.3	Energetic stability analysis	334
11.3.4	Instability condition in terms of absolute vorticity	336
11.3.5	Parcel stability analysis	336
11.3.6	Stability condition in terms of the surface height	338
11.3.7	Comments and further study	339
<b>11.4</b>	<b>Potential momentum on the <math>f</math>-plane</b>	<b>339</b>
11.4.1	Linear momentum and potential momentum	340
11.4.2	Zonal potential momentum on a $\beta$ -plane	341
11.4.3	Further reading	342
<b>11.5</b>	<b>Horizontal inertial instability of geostrophic fronts</b>	<b>342</b>
11.5.1	Equations of motion relative to the geostrophic equilibrium state	343
11.5.2	Stability analysis based on energetic arguments	343
11.5.3	Stability analysis based on parcel arguments	345
<b>11.6</b>	<b>Symmetric instability and the <math>f Q &lt; 0</math> criteria</b>	<b>349</b>
11.6.1	Equations using buoyancy as the vertical coordinate	349
11.6.2	Parcel stability analysis	350
11.6.3	Geometric perspective on the instability condition	351
<b>11.7</b>	<b>Symmetric instability and the wedge of instability</b>	<b>353</b>
11.7.1	Formulation of the basic equations	353
11.7.2	Assessing stability of the perturbation	354
11.7.3	Symmetrically unstable while inertially and gravitationally stable	355
<b>11.8</b>	<b>Symmetric instability and linear modal analysis</b>	<b>356</b>
11.8.1	Geostrophic base state and the perturbation equations	356
11.8.2	Inertia-vorticity oscillator equations	357
11.8.3	Ageostrophic overturning circulation streamfunction	358
11.8.4	Dispersion relation for meridional-vertical plane waves	359
11.8.5	Stability conditions	359
11.8.6	Ertel potential vorticity and local stability	361
11.8.7	Comments and further study	362
<b>11.9</b>	<b>Semi-geostrophy and ageostrophic overturning</b>	<b>362</b>
11.9.1	Hydrostatic and Boussinesq ocean on an $f$ -plane	362
11.9.2	Scaling for frontal flows	363
11.9.3	Geostrophic momentum approximation	364
11.9.4	Geostrophic evolution of the buoyancy gradient	365
11.9.5	Secondary ageostrophic circulation	366
11.9.6	Ageostrophic overturning circulation for a symmetric front	367

11.9.7	Connection to potential vorticity and symmetric instability . . . . .	368
11.9.8	Further study . . . . .	369
<b>11.10 Exercises</b>	. . . . .	<b>369</b>

## 11.1 Loose threads

- Work through symmetric instability in two shallow water layers. Does it work? Do we need to have mixing of the layers?
- Write the full oscillator equation for (11.75), including  $\Delta x$  and  $\Delta z$  motion.
- Energy analysis for each instability as per Chapter 8 of *Smyth and Carpenter* (2019), who use a modal approach and linearize the equations. Include energetics of gravitational instability.
- Maximum growth rate for symmetric instability in chapter 8 of *Smyth and Carpenter* (2019).
- Include more in Section 11.9.7 concerning stable and unstable symmetric fronts.
- More closely connect the ageostrophic overturning from Section 11.9 to symmetric instability. Make the case that the overturning is either stable or unstable, with the unstable form leading to symmetric instability.
- More about the growth of the squared buoyancy gradient in equation (11.170). Can that equation be written in a more revealing manner?
- Pursue the coordinate transformation from *Hoskins* (1975), in which case the discussion of semi-geostrophy could form a separate chapter.
- What is the potential vorticity for the linearized perturbations in Section 11.8?
- Sawyer-Eliasson front equation for gradient wind balanced cylindrical flow, as I Hunter Camp's work. Perhaps make into an exercise.

## 11.2 Instabilities in this chapter

The instabilities studied in this chapter are termed *local* since they are detected through a local condition that yields a necessary and sufficient condition for instability. The physical features of the flow instabilities considered here are quite similar, thus making it convenient to study them together. For centrifugal and horizontal inertial instabilities, we develop the stability conditions using both energetic and parcel analysis methods. For symmetric instability we use the parcel and modal methods.

### 11.2.1 Summary of the instabilities studied in this chapter

Before diving into the details, we here offer a summary of the instabilities studied in this chapter.

### Centrifugal instability of cyclostrophic balanced flow

Consider an equilibrium flow state under inviscid cyclostrophic balance. As studied in Section ??, cyclostrophic balance arises when pressure and centrifugal accelerations are balanced, with centrifugal accelerations arising from curvature in the fluid particle trajectory. The angular momentum is materially invariant when the equilibrium state is rotationally symmetric, as in an ideal circular vortex or a rotating circular tank (Section ??). Flow stability can be probed by horizontally displacing a rotationally symmetric circular ring of fluid parcels, with each parcel retaining its original angular momentum. If the parcels are displaced to a position where pressure and centrifugal accelerations further support the displacement, then the base state is unstable to centrifugal instability.

### Horizontal inertial instability of $f$ -plane geostrophic flow

Consider an equilibrium flow state in a barotropic fluid under inviscid geostrophic balance on an  $f$ -plane, whereby pressure and Coriolis accelerations are balanced (Chapter ??). The potential momentum is materially invariant when the equilibrium state is symmetric in a horizontal direction, as in a zonally or meridionally symmetric  $f$ -plane front. Flow stability is probed by horizontally displacing a symmetric line of fluid parcels, with each parcel retaining its original potential momentum. If the parcels are displaced to a position where pressure and Coriolis accelerations further support the displacement, then the base state is unstable to horizontal inertial instability.

### Isentropic inertial instability (symmetric instability) for $f$ -plane flow

Consider an equilibrium flow state under inviscid  $f$ -plane geostrophic balance in a baroclinic fluid. Potential momentum in the direction of the front as well as buoyancy are materially invariant when the base state is symmetric in a horizontal direction. Flow stability of the base flow is probed by isentropically displacing a symmetric line of fluid parcels, with each parcel retaining its original potential momentum and buoyancy. If the parcels are displaced to a position where pressure and Coriolis accelerations further support the displacement, then the base state is unstable to isentropic inertial instability, which we call *symmetric* instability.

#### 11.2.2 Nature of the base state and the perturbations

For the study of centrifugal instability, we make use of the shallow water model from Chapters ?? and ??, whereas we consider the continuously stratified Boussinesq ocean (Chapter ??) for the horizontal and isentropic inertial instabilities. When the fluid is continuously stratified and inviscid, all motion occurs along isentropes. However, when probing for centrifugal or horizontal inertial instabilities using parcel arguments, we examine stability to perturbations along geopotential surfaces. Such horizontal displacements generally cross isentropic surfaces in a baroclinic fluid. The isentropic inertial instability analysis in Section 11.6 maintains the adiabatic nature of displacements when probing for instabilities. Even so, these perturbations do not maintain a materially invariant potential vorticity. So what is it about these perturbations that makes them relevant to stability analysis?

The stability thought experiment using test fluid parcels assumes the parcels maintain their materially invariant property (e.g., buoyancy, angular momentum, potential momentum) as they probe stability of the surrounding fluid flow. That is the nature of the test parcels, as they do not directly interact with nor alter the surrounding fluid environment. Local flow stability

is examined by having test parcels cross surfaces of constant materially invariant properties. For example, gravitational stability is probed by test parcels moving across buoyancy surfaces (Section ??). Likewise, entering the *wedge of instability* for symmetrically unstable flow requires a test parcel to leave its constant buoyancy and constant potential momentum surfaces (Section 11.6).

### 11.2.3 Comments

Section 7.1 of [Holton and Hakim \(2013\)](#) distinguishes between *parcel* and *wave* instabilities. Analogously, [Cushman-Roisin and Beckers \(2011\)](#) use the terms *local* and *global*. Canonical examples of global instabilities are Kelvin-Helmholtz and baroclinic. We are not concerned with those instabilities in this chapter. Rather, the three instabilities considered here are examples of parcel or local instabilities.

Because of the rather close similarities between centrifugal and inertial instability, the oceanographic literature often uses the term centrifugal instability when referring to the inertial instability considered here (e.g., see [Thomas et al. \(2013\)](#) and [McWilliams \(2016\)](#)). However, we do not follow that usage since the inertial instability in this chapter is *not* associated with centrifugal accelerations. Rather, inertial instability is associated with Coriolis accelerations. We choose to follow the terminology of the atmospheric literature as detailed in the texts by [Holton and Hakim \(2013\)](#) and [Markowski and Richardson \(2010\)](#), which also follows the fluid mechanics terminology used by [Drazin and Reid \(1981\)](#). So in brief, we use the term *centrifugal instability* for an axisymmetric base state in cyclostrophic balance, and *inertial instability* for a two-dimensional base state in geostrophic balance.

## 11.3 Centrifugal instability of cyclostrophic flow

Consider flow of a single shallow water fluid layer in a rotating cylindrical tank with rotation about its vertical axis.<sup>2</sup> Throughout this analysis we assume all flow features maintain rotational symmetry (Figure 11.1). Hence, all dynamical fields are a spatial function only of the radial distance from the rotational axis (*axisymmetric*). We are interested in questions concerning flow stability as a function of the radial distribution of the angular momentum per mass,  $l^z(r)$ . In particular, we examine stability of *cyclostrophically balanced* flow, defined by flow whose radial acceleration vanishes so that the radial pressure gradient balances the centrifugal acceleration (Section ??). Furthermore, we examine stability under perturbations that also maintain axial symmetry. We find that such cyclostrophic flow is stable to rotationally symmetric perturbations so long as the squared angular momentum increases radially outward. This system provides a pedagogical introduction to the stability of rotating vortices in the ocean and atmosphere, and it establishes analysis methods we later use for inertial and symmetric instabilities arising in two-dimensional frontal regions. The stability condition for angular momentum is reflected in observed stable vortices in both the atmosphere and ocean.

The centrifugal instability described here does *not* rely on baroclinic structure. Rather, it arises from the imbalances between centrifugal and pressure forces along a geopotential surface.<sup>3</sup>

<sup>2</sup>The same ideas can be formulated for rotating Couette flow, in which fluid is placed between two rotating cylinders. We prefer the shallow water tank since it is a system considered elsewhere in this book as a laboratory model for rotating fluid mechanics.

<sup>3</sup>One may conceive of centrifugal instability in a baroclinic flow where parcel displacements maintain their angular momentum and buoyancy. That analysis would lead to isentropic centrifugal instability, which is directly analogous to the isentropic inertial instability discussed in Section 11.6. See [Buckingham et al. \(2021\)](#) for a

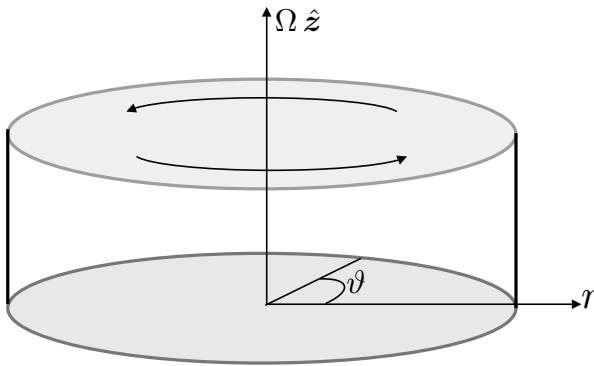


FIGURE 11.1: Rotating tank of shallow water fluid used to discuss centrifugal instability. We make use of cylindrical-polar coordinates from Section ?? to describe the flow. All flow features are assumed to be rotationally symmetric, including both the cyclostrophic balanced state and perturbations relative to the balanced state.

By using the shallow water fluid to study this instability, we directly connect to Section ??, where we studied angular momentum in a rotating tank of shallow water fluid.

### 11.3.1 Equations of motion

We studied the angular momentum of this rotating shallow water system in Section ??, where we made use of a rotating reference frame and polar coordinates  $(r, \vartheta)$  (Section ??) measured in the rotating frame. Here,  $r$  is the radial position from the rotational axis and  $\vartheta$  is the azimuthal angle measured counter-clockwise from the rotating  $x$ -axis. We furthermore derived the acceleration in cylindrical-polar coordinates in Exercise ???. Making use of those results allows us to write down the horizontal components to the inviscid velocity equation as well as the thickness equation

$$Dv^r/Dt = -g \partial_r \eta + r^{-3} (l^z)^2 \quad (11.1a)$$

$$Dl^z/Dt = -g \partial_\vartheta \eta \quad (11.1b)$$

$$r \partial_t \eta = -\partial_r (h r v^r) - \partial_\vartheta (h v^\vartheta). \quad (11.1c)$$

In these equations,  $h$  is the layer thickness and  $\eta$  is the free surface height for the shallow water layer (see Figure ??). The radial and azimuthal velocity components are given by

$$v^r = Dr/Dt \quad \text{and} \quad v^\vartheta = r D\vartheta/Dt, \quad (11.2)$$

and the angular momentum per mass computed about the rotation axis (the  $z$ -axis) is

$$l^z = \hat{\mathbf{z}} \cdot [\mathbf{r} \times (\mathbf{u} + \mathbf{U}_{\text{rigid}})] = r (v^\vartheta + r \Omega). \quad (11.3)$$

Material evolution of the radial velocity (equation (11.1a)) is affected by the radial pressure gradient plus centrifugal acceleration, whereas the material evolution of angular momentum (equation (11.1b)) is affected by angular gradients in the pressure field as realized by angular gradients in the free surface.

We consider equilibrium states where the radial acceleration vanishes. Such states are said to be in *cyclostrophic balance*, whereby the radial pressure gradient balances the centrifugal discussion of this system.

acceleration

$$Dv^r/Dt = 0 \implies g \partial_r \eta = (v^\vartheta + r \Omega)^2/r = r^{-3} (l^z)^2 \quad \text{cyclostrophic balance.} \quad (11.4)$$

Is cyclostrophically balanced flow stable? To answer this question we examine the more restricted problem of stability of rotational symmetric flow, and with perturbations also assumed to be rotationally symmetric. Rotational symmetry also means that all flow fields are a function only of the radial direction. As such, angular momentum is materially constant

$$\partial_\vartheta \eta = 0 \implies Dl^z/Dt = 0. \quad (11.5)$$

This constraint on the flow plays a fundamental role in the stability analysis.

### 11.3.2 Elements of angular momentum

Consider a spinning top in a gravitational field, in which conservation of angular momentum (holding in the absence of friction) keeps the top upright. Yet in the presence of friction, the angular momentum is dissipated so that the top eventually falls. There is an analog with a rotating fluid. Namely, we find that a state of zero angular momentum leads to centrifugal instability. To develop this result we study basic properties for angular momentum in the rotating tank.

#### Rigid-body motion

A fluid in rigid-body motion has  $v^\vartheta = 0$  and angular momentum per mass

$$l_{\text{rigid-body}}^z = r^2 \Omega. \quad (11.6)$$

Evidently, the magnitude of the angular momentum increases as the square of the radial distance. In the following we find it more convenient to use the square of the angular momentum (as it appears in the radial velocity equation (11.1a)), which also increases radially for the rigid-body motion

$$\frac{d[l^z(r)]^2}{dr} = 4r^3 \Omega^2 > 0. \quad (11.7)$$

We will find that flow is centrifugally unstable if the square of its angular momentum is a decreasing function of its radial distance. Such unstable configurations adjust through *centrifugal instability* towards a configuration with squared angular momentum that increases radially. The instability is termed “centrifugal” since it is the centrifugal acceleration that “throws outward” the fluid if its squared angular momentum decreases radially, thus bringing the fluid back into a stable state. More precisely, a flow where the angular momentum decreases outward has pressure gradients that cannot balance the centrifugal acceleration from the outward movement of angular momentum conserving fluid parcels.

#### Zero angular momentum flow

Consider a non-rotating tank with zero flow, so that the angular momentum is zero. In a perfect fluid the angular momentum of each fluid parcel remains zero, even if we start the tank rotating. That is, for a perfect fluid with zero angular momentum, the tank simply rotates but the fluid remains at rest in the inertial reference frame. To generate non-zero angular momentum for the fluid requires friction between the rotating tank and fluid. After sufficient time, friction

transfers angular momentum from the tank walls throughout the fluid, thus leading to a steady state flow in rigid-body motion. Upon reaching the steady rigid-body flow, there are no strains in the fluid and thus no viscous stresses to impart friction (Section ??).

In a flow with zero angular momentum, the relative angular velocity is given by

$$l^z = 0 \implies v^\vartheta = -r\Omega, \quad (11.8)$$

so that the flow is anti-cyclonic (i.e., directed counter to the tank's rotation). Furthermore, the zero angular momentum flow has the relative vorticity

$$l^z = 0 \implies \zeta = \frac{1}{r} \frac{d(rv^\vartheta)}{dr} = -2\Omega. \quad (11.9)$$

That is, the relative vorticity is anti-cyclonic with a Rossby number

$$l^z = 0 \implies \text{Ro} = \zeta/(2\Omega) = -1. \quad (11.10)$$

As we show in the following, the  $l^z = 0$  flow is the onset point for centrifugal instability. Again, we conceive of this flow as analogous to a rotating top that falls over (goes unstable) when its angular momentum vanishes.

### Establishing zero angular momentum flow

It is relatively simple to establish rigid-body motion of a fluid in a tank; all it takes is sufficient time for transients to relax and friction to fully transfer momentum from the tank walls throughout the fluid. For the complement task, consider a rigid-body flow that is subjected to an irreversible force that brings the fluid to zero angular momentum. The question we ask in our stability analysis is not concerned with details of how such forces arise. Instead, we are interested in what happens to the flow when it reaches zero angular momentum.

#### 11.3.3 Energetic stability analysis

As first introduced by Rayleigh, we consider a thought experiment in which two adjacent equal mass circular fluid rings are swapped, one originating from radial position  $r = r_1$  and the other at  $r = r_2 = r_1 + \Delta r$  (see Figure 11.2). Furthermore, assume that the radial velocity vanishes so that the kinetic energy of the rings is due only to their rotational motion. If swapping the rings decreases the kinetic energy in the base state, then the released kinetic energy can be used to fuel an instability.<sup>4</sup> In this case we say that the flow is *centrifugally unstable*, with this name used since it is the centrifugal acceleration from the circular parcel trajectory that promotes the instability. In general, any curved flow will be exposed to centrifugal instability if swapping parcels reduces the kinetic energy of the base state while maintaining constant angular momentum.

With no radial flow, kinetic energy only arises from angular motion so that a ring of mass  $\delta m = \rho \delta V$  and radius  $r_1$  has kinetic energy

$$E(r_1) = (\delta m/2) [v^\vartheta(r_1) + r_1 \Omega]^2 = (\delta m/2) [l^z(r_1)/r_1]^2. \quad (11.11)$$

---

<sup>4</sup>Gravitational potential energy plays no role here, as we are swapping fluid rings at the same vertical position.

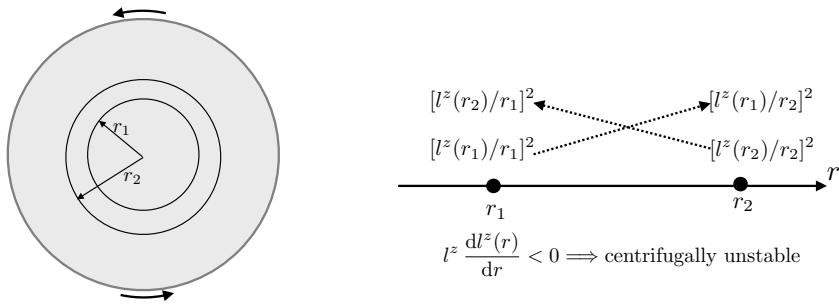


FIGURE 11.2: Illustrating Rayleigh’s energetic method for centrifugal instability. The initial configuration defines two rings of fluid at radii  $r_1$  and  $r_2 = r_1 + \Delta r$ , as shown in the left panel, with the two rings having a combined kinetic energy per mass  $2E_{\text{init}} = [l^z(r_1)/r_1]^2 + [l^z(r_2)/r_2]^2$ . Swapping the two rings while maintaining a constant angular momentum (and assuming nothing else changes) leads to the swapped kinetic energy per mass,  $2E_{\text{swap}} = [l^z(r_1)/r_2]^2 + [l^z(r_2)/r_1]^2$ . If  $E_{\text{swap}} - E_{\text{init}} < 0$  then energy is released upon the swap and the rotationally symmetric flow is centrifugally unstable. Energy is released upon swapping fluid rings if the angular momentum condition (11.15) is satisfied:  $d[l^z(r)]^2/dr < 0$ .

The initial kinetic energy for the two rings is thus given by the sum

$$E_{\text{init}} = (\delta m/2) ([l^z(r_1)/r_1]^2 + [l^z(r_2)/r_2]^2). \quad (11.12)$$

The equilibrium state and the perturbation each maintain rotational symmetry. Hence, when swapping their radial positions, the rings each maintain their respective angular momentum. But by changing radial positions their kinetic energy changes, thus leading to the kinetic energy of the swapped state

$$E_{\text{swap}} = (\delta m/2) ([l^z(r_1)/r_2]^2 + [l^z(r_2)/r_1]^2). \quad (11.13)$$

The difference in energy is given by

$$E_{\text{swap}} - E_{\text{init}} = (\delta m/2) ([l^z(r_2)]^2 - [l^z(r_1)]^2)(r_1^{-2} - r_2^{-2}). \quad (11.14)$$

Since  $r_2 = r_1 + \Delta r > r_1$ , we have a release of kinetic energy ( $E_{\text{swap}} - E_{\text{init}} < 0$ ) if the squared angular momentum decreases upon moving outward

$$\frac{d[l^z(r)]^2}{dr} = 2l^z(r) \frac{dl^z(r)}{dr} < 0 \implies \text{necessary + sufficient for centrifugally unstable flow.} \quad (11.15)$$

Recall from equation (11.7) that the angular momentum of rigid-body flow is an increasing function of radius. Hence, the instability condition (11.15) signals a distinctively non-rigid-body angular momentum configuration. We can consider a thought experiment in which a stable flow (one where  $d[l^z(r)]^2/dr > 0$ ) is somehow forced towards a centrifugally unstable state with  $d[l^z(r)]^2/dr < 0$ . A vanishing angular momentum at a particular radius,  $l^z = 0$ , is a sufficient condition to ensure  $d[l^z(r)]^2/dr = 0$  at that radius. Evidently, if the angular momentum of the cyclostrophic flow is reduced toward zero, the fluid will go centrifugally unstable upon reaching just below zero angular momentum.

### 11.3.4 Instability condition in terms of absolute vorticity

The angular momentum condition (11.15) is the traditional way to write the necessary and sufficient condition for centrifugal instability. However, to anticipate the role of vorticity found in the case of inertial and symmetric instability, we rewrite the stability condition (11.15) in terms of absolute vorticity. For this purpose, introduce the vertical component of the relative vorticity,

$$\zeta = \frac{1}{r} \frac{d(r v^\vartheta)}{dr}, \quad (11.16)$$

and thus write the angular velocity as the radial integral of the radius weighted vorticity

$$v^\vartheta = \frac{1}{r} \int_0^r r' \zeta(r') dr'. \quad (11.17)$$

This result allows us to write the angular momentum (11.3) in terms of the absolute vorticity

$$l^z = \Omega r^2 + \int_0^r r' \zeta(r') dr' = \int_0^r r' (\zeta(r') + 2\Omega) dr' = \int_0^r r' \zeta_a(r') dr'. \quad (11.18)$$

This equation allows us to write the stability condition (11.15) in the form

$$\frac{d[l^z(r)]^2}{dr} = 2 l^z \frac{dl^z(r)}{dr} = 2 r \zeta_a(r) \int_0^r r' \zeta_a(r') dr' < 0 \implies \text{centrifugally unstable}. \quad (11.19)$$

For  $d[l^z(r)]^2/dr < 0$  at a particular radius,  $r$ , and thus for the cyclostrophic flow to be unstable, requires the absolute vorticity at that radius to have the opposite sign to the integral of the radius weighted absolute vorticity from the origin to  $r$ . Since  $r \geq 0$ , a necessary condition for the instability is for the absolute vorticity to have a sign change somewhere within the region  $0 \leq r$ .

### 11.3.5 Parcel stability analysis

As a complement to the energetic stability analysis of Section 11.3.3, we here study the force balance in the radial momentum equation. This analysis shows that an unstable angular momentum profile corresponds to a situation where the outward centrifugal acceleration is stronger than the radially inward pressure gradient acceleration, thus allowing for the fluid parcel to be “thrown” outward.

We again assume rotational symmetry so that the angular momentum is a material invariant, and assume the base state is in cyclostrophic balance. As such, the radial momentum equation (11.1a) leads to

$$0 = -g \partial_r \bar{\eta} + r^{-3} (\bar{l}^z)^2, \quad (11.20)$$

with the overline denoting the cyclostrophic base state. Subtracting this equilibrium flow from the full momentum equation (11.1a) leads to an equation for radial acceleration of perturbations about the equilibrium state

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} + r^{-3} [(l^z)^2 - (\bar{l}^z)^2], \quad (11.21)$$

where

$$\eta' = \eta - \bar{\eta} \quad (11.22)$$

is the perturbation surface height. We focus now on the difference in squared angular momentum, with questions about the perturbation pressure gradient,  $-g \partial\eta'/\partial r$ , falling outside the parcel method of stability analysis.

### Probing stability by perturbing the radius of a circular fluid ring

Consider a perturbation realized by moving a constant mass circular fluid ring outward from its initial equilibrium state at radius  $r$  to a radius  $r + \Delta r$ . During expansion of the ring, its angular momentum remains constant due to the rotational symmetry, so that

$$(l^z)^2(r + \Delta r) = (l^z)^2(r) = (\bar{l}^z)^2(r), \quad (11.23)$$

where the second equality holds since we are starting the ring from its cyclostrophic initial condition (11.20). To determine the radial acceleration at  $r + \Delta r$  appearing on the right hand side of the radial velocity equation (11.21), we compute

$$(l^z)^2(r + \Delta r) - (\bar{l}^z)^2(r + \Delta r) = (\bar{l}^z)^2(r) - (\bar{l}^z)^2(r + \Delta r) \approx -\Delta r \frac{d(\bar{l}^z)^2(r)}{dr}, \quad (11.24)$$

so that the equation (11.21) evaluated at  $r + \Delta r$  is

$$\frac{Dv^r}{Dt} = -g \frac{\partial\eta'}{\partial r} - \frac{\Delta r}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr}. \quad (11.25)$$

We thus see that if the squared angular momentum decreases upon moving the ring to a larger radius, then the second right hand side term in equation (11.25) provides a positive radial acceleration, thus supporting the initial outward perturbation. Ignoring the possibility for the perturbation pressure gradient,  $-g\partial\eta'/\partial r$ , to counter-act the acceleration, we are left with the same instability condition (11.15) derived using energetic arguments.

### Describing the instability mechanism

When the fluid is in cyclostrophic balance, the radial pressure gradient acceleration (which is pointing radially inward) balances the centrifugal acceleration (which is pointing radially outward) as per equation (11.20). When a constant mass fluid ring is perturbed outward, from  $r$  to  $r + \Delta r$ , it carries its angular momentum,  $\bar{l}^z(r)$ , to the new location. This displaced ring is generally not in cyclostrophic balance with the pressure gradient at the new position. That is, the centrifugal acceleration of the displaced ring does not equal the pressure gradient of the new position

$$\underbrace{(r + \Delta r)^{-3} [\bar{l}^z(r)]^2}_{\text{centrifugal of displaced ring}} \neq \underbrace{g \frac{\partial\bar{\eta}(r + \Delta r)}{\partial r}}_{\text{pressure gradient at } r + \Delta r} \quad (11.26)$$

If the centrifugal acceleration of the displaced ring is greater than the local pressure gradient acceleration, then the ring will continue to move outward. That is, the outward centrifugal acceleration at the new location is greater than the inward pressure gradient so that the fluid ring is flung further outward. This process characterizes a centrifugally unstable state, and it occurs if the squared angular momentum decreases when moving radially outward, as per equation (11.15). Conversely, if the local pressure gradient is greater than the centrifugal acceleration of the displaced ring, then the ring returns to its original radius and exhibits stable centrifugal oscillations.

### Centrifugal oscillations

Ignoring the perturbation pressure gradient, and introducing a squared centrifugal angular frequency,

$$\sigma^2(r) \equiv \frac{1}{(r + \Delta r)^3} \frac{d(\bar{l}^z)^2(r)}{dr}, \quad (11.27)$$

leads to the free oscillator equation for the fluctuation of a fluid ring from its equilibrium radial position

$$\frac{D^2 \Delta r}{Dt^2} + \sigma^2 \Delta r = 0, \quad (11.28)$$

where  $u^r = D(\Delta r)/Dt$  is the radial velocity of a fluid parcel relative to its equilibrium radial position. For stable cases with  $\sigma^2 > 0$ , the ring exhibits harmonic centrifugal oscillations around the equilibrium radius with period  $2\pi/\sigma$ . In contrast, for the unstable case with  $\sigma^2 < 0$ , then  $\Delta r$  grows exponentially.

#### 11.3.6 Stability condition in terms of the surface height

The equilibrium angular momentum and free surface height are related by the cyclostrophic balance (11.20). So rather than focusing on the angular momentum, we can develop an equivalent stability condition in terms of radial derivatives of the free surface height. For this purpose, write the difference in angular momentum

$$(\bar{l}^z)^2(r) - (\bar{l}^z)^2(r + \Delta r) = g [r^3 d\bar{\eta}(r)/dr - (r + \Delta r)^3 d\bar{\eta}(r + \Delta r)/dr], \quad (11.29)$$

and then perform a Taylor series

$$F(r + \Delta r) \equiv (r + \Delta r)^3 d\bar{\eta}(r + \Delta r)/dr \quad (11.30a)$$

$$\approx F(r) + \Delta r dF(r)/dr \quad (11.30b)$$

$$= r^3 d\bar{\eta}(r)/dr + \Delta r d[r^3 d\bar{\eta}/dr]/dr, \quad (11.30c)$$

which renders

$$(\bar{l}^z)^2(r) - (\bar{l}^z)^2(r + \Delta r) \approx -g \Delta r \frac{d}{dr} \left[ r^3 \frac{d\bar{\eta}}{dr} \right]. \quad (11.31)$$

The radial momentum equation (11.25) thus takes the form

$$\frac{Dv^r}{Dt} = -g \frac{\partial \eta'}{\partial r} - \frac{g \Delta r}{(r + \Delta r)^3} \frac{d}{dr} \left[ r^3 \frac{d\bar{\eta}}{dr} \right]. \quad (11.32)$$

We are thus led to the instability condition written in terms of the surface height in the cyclostrophically balanced base state

$$\frac{d}{dr} \left[ r^3 \frac{d\bar{\eta}}{dr} \right] < 0 \implies \text{centrifugally unstable}. \quad (11.33)$$

A particular example of a centrifugally unstable free surface configuration is the decreasing function

$$\bar{\eta} = \eta_0 - a r^b \implies \left[ r^3 \frac{d\bar{\eta}}{dr} \right] = -a b (2 + b) r^{b+1}, \quad (11.34)$$

with  $\eta_0$  the free surface height at the origin,  $b$  a dimensionless constant, and  $a$  having dimensions so that  $a r^b$  has dimensions of length. Furthermore, we require with  $a b > 0$ , and  $b > -2$ . Recall

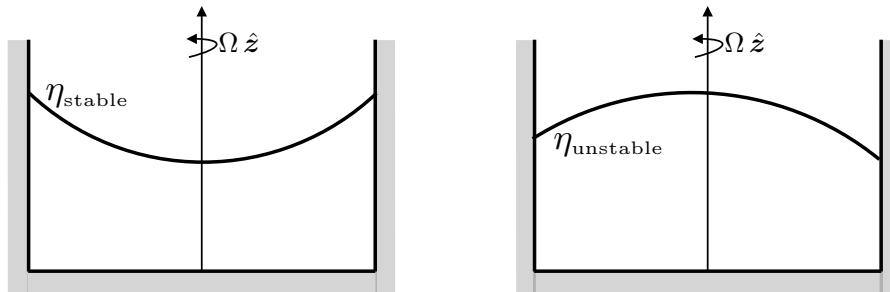


FIGURE 11.3: Example free surface heights for a shallow water layer in a rotating tank. The left panel exhibits a free surface height that is stable to centrifugal instability, whereas the right is unstable, with stability determined by the condition (11.33). The stable configuration approximates the parabolic profile that is realized for a rigid-body flow as in equation (11.35). The unstable profile has its layer thickness decrease with increasing radius, so that the radial pressure gradient acceleration cannot balance the centrifugal acceleration.

that the free surface is parabolic when the fluid is in rigid-body motion, as given by equation (??)

$$\bar{\eta} = \eta_0 + \Omega^2 r^2 / (2 g_e) \implies \left[ r^3 \frac{d\bar{\eta}}{dr} \right] = 4 \Omega^2 r^3 / g_e > 0, \quad (11.35)$$

which is stable.

### 11.3.7 Comments and further study

Chapter 3 of [Drazin and Reid \(1981\)](#) is the canonical reference for centrifugal instability, where they provide a stability analysis including both axisymmetric and non-axisymmetric perturbations. In our treatment, we also made use of the parcel arguments from Section 3.2 of [Markowski and Richardson \(2010\)](#). Furthermore, [Markowski and Richardson \(2010\)](#) comment on the perturbation pressure gradient in equation (11.25), emphasizing that parcel stability arguments generally ignore changes to the pressure gradient. Stated otherwise, a parcel analysis concerns the equilibrium angular momentum profile and its contribution to movement away from equilibrium. It is not concerned with back reaction from pressure perturbations associated with movement of parcels or fluid rings. A fuller treatment generally requires analysis beyond the parcel framework. Section 11.6 of [Kundu et al. \(2016\)](#) presents centrifugal instability in the context of viscous Couette flow, which is the flow of fluid between two rotating co-axial cylinders.

For a laboratory realization of centrifugal instability, see [this video](#) from the UCLA spin lab. The laboratory technique spins up a fluid to rigid-body motion in a tank, and then quickly reduces the rotation rate to induce  $d[l^z(r)]^2/dr < 0$ . The instability manifests as roll vortices aligned along the tank wall, with the character of the vortices a function of  $d[l^z(r)]^2/dr$  and other parameters.

## 11.4 Potential momentum on the *f*-plane

A front is a region of enhanced lateral gradients in the buoyancy field (baroclinic front) or sea level (shallow water front). These fronts generally have corresponding currents (jets) arising from geostrophic balance (when off-equator). Figure 11.4 illustrates a baroclinic front that is symmetric in the zonal direction so that the buoyancy field is only a function of latitude, depth, and time,  $b = b(y, z, t)$ . We likewise assume that all other fields possess zonal symmetry,

including pressure and velocity. Fronts can generally be oriented in any direction. Furthermore, there is rotational symmetry on the  $f$ -plane so that we can orient the horizontal coordinate axes as desired. In later sections we study the stability of fronts. Here, we establish some basic properties of *potential momentum*, which proves central to the stability analysis.

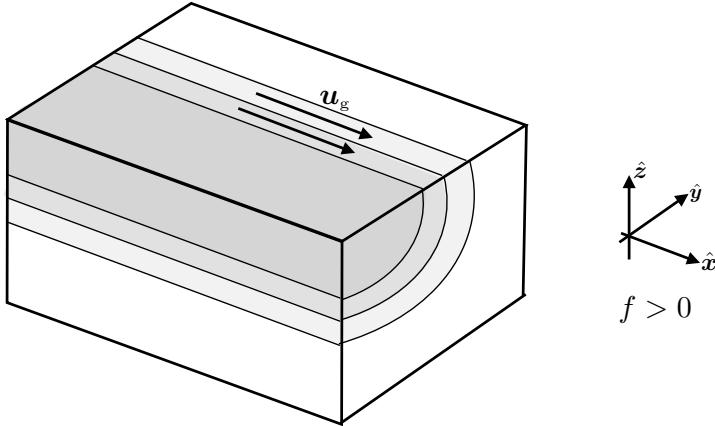


FIGURE 11.4: Example of a zonally symmetric baroclinic frontal region, showing iso-buoyancy surfaces with  $b = b(y, z, t)$ . With  $\partial b / \partial y < 0$  as drawn, the corresponding zonal thermal wind shear,  $f \partial u_g / \partial z = -\partial b / \partial y > 0$ , is eastward for a northern hemisphere front; i.e., stronger eastward flow with increasing height (see Section ??).

### 11.4.1 Linear momentum and potential momentum

The horizontal linear momentum per mass for a perfect Boussinesq ocean on an  $f$ -plane evolves according to (see Section ??)

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi. \quad (11.36)$$

Since  $f$  is a constant, this equation can be written

$$\frac{DM}{Dt} = -\nabla_h \varphi, \quad (11.37)$$

where we introduced the *potential momentum* per mass

$$\mathbf{M} \equiv \mathbf{u} + f \hat{\mathbf{z}} \times \mathbf{x} = (u - f y) \hat{\mathbf{x}} + (v + f x) \hat{\mathbf{y}}, \quad (11.38)$$

and noted that  $\mathbf{v} = D\mathbf{x}/Dt$ . Notably, the potential momentum remains materially invariant in directions where the horizontal pressure gradient vanishes. We described the potential momentum for a point particle in Section ??, and here make use of it to study the stability of certain  $f$ -plane fluid flows.

#### Materially constant zonal potential momentum

For the zonally aligned buoyancy front illustrated in Figure 11.4, we assume all fields are independent of the zonal direction so that

$$\partial_x \varphi = 0. \quad (11.39)$$

A vanishing zonal pressure gradient means that the zonal potential momentum per mass is a material invariant

$$DM^x/Dt = 0 \quad \text{where } M^x \equiv u - fy. \quad (11.40)$$

This material invariance greatly constrains the flow. For example, consider a fluid parcel at an initial latitude  $y_0$  with zonal velocity  $u_0$ . Movement of the parcel to a new latitude leads to the equality

$$u_0 - fy_0 = u_1 - fy_1, \quad (11.41)$$

so that the zonal velocity at the new latitude is given by

$$u_1 = u_0 - f(y_0 - y_1). \quad (11.42)$$

### Motivating the name potential momentum

Although *absolute momentum* or *geostrophic momentum* are more commonly used in the literature, we prefer the term *potential momentum* as motivated by the same reasoning used for potential temperature (Section ??), potential density (Section ??), and potential vorticity (Section ??). Namely, the zonal potential momentum identifies that amount of zonal linear momentum per mass (i.e., the velocity) that a parcel would acquire if moved from an arbitrary latitude  $y_1$ , to a reference latitude,  $y_0$ . More specifically, inverting equation (11.42) we see that

$$u_0 = M^x(y_1) + fy_0. \quad (11.43)$$

Hence, the quantity  $M^x(y_1)$  is the extra zonal momentum per mass available at the reference latitude,  $y_0$ , upon moving a parcel from  $y_1$  to  $y_0$ . We thus see that the potential momentum is a material invariant in the way that potential temperature is for a perfect fluid. Furthermore, as seen in Section 11.5, meridional gradients of  $M^x$  measure the *inertial* stability of a flow configuration in a manner directly analogous to how vertical gradients of potential temperature (or buoyancy) measure gravitational stability.

### Meridional potential momentum

There are occasions when a front exhibits meridional symmetry, in which case the perfect Boussinesq equations take the form

$$Du/Dt = fv - \partial_x \varphi \quad (11.44a)$$

$$Dv/Dt = -fu. \quad (11.44b)$$

In this case the meridional potential momentum is materially invariant

$$DM^y/Dt = 0 \quad \text{where } M^y \equiv v + fx. \quad (11.45)$$

#### 11.4.2 Zonal potential momentum on a $\beta$ -plane

The *f*-plane is rotationally invariant about the rotational axis. Correspondingly, we can write the momentum equation in the form (11.37), thus exposing the potential momentum. Material invariance for potential momentum holds along the symmetry direction of an arbitrarily oriented symmetric front.

The  $\beta$ -plane is not rotationally invariant. Rather, it only maintains symmetry along lines of constant latitude (zonal directions). Consequently, only zonally oriented symmetric fronts

maintain material invariance of zonal potential momentum. To see this property, write the zonal momentum equation in the form

$$\partial_t u + v \partial_y u + w \partial_z u - f v = 0, \quad (11.46)$$

where we assumed zonal symmetry ( $\partial_x = 0$ ). Now write the Coriolis parameter in the form

$$\Gamma = f_0 y + \beta y^2 / 2 \implies f = d\Gamma/dy, \quad (11.47)$$

so that the zonal momentum equation takes the form

$$\partial_t(u - \Gamma) + v \partial_y(u - \Gamma) + w \partial_z(u - \Gamma) = 0. \quad (11.48)$$

Evidently,  $M^x = u - \Gamma$  is materially invariant for this zonally symmetric front.

### 11.4.3 Further reading

See Section ?? for more discussion of potential momentum as it concerns a point particle. The term *absolute momentum* follows the arguments given on page 51 of [Markowski and Richardson \(2010\)](#).

## 11.5 Horizontal inertial instability of geostrophic fronts

We now examine stability of an  $f$ -plane geostrophically balanced front for an inviscid Boussinesq fluid. The analysis of centrifugal instability in Section 11.3 is closely followed, with rotational symmetry replaced by along-front symmetry and angular momentum replaced by potential momentum. We consider both the Rayleigh energetic stability analysis and the parcel analysis. Furthermore, the perturbations maintain symmetry in the along-front direction, so perturbations consist of a horizontally displaced row of parcels oriented along the front. Stability to more general perturbations, such as those that are not symmetric along the front, is not addressed here.

The results of our analysis are rotationally invariant since the  $f$ -plane is rotationally invariant. Hence, we choose to orient the coordinate system based on convenience whereby the  $x$ -axis is the along front axis and the  $y$ -axis is across the front (as in Figure 11.4). Furthermore, their relative orientation is chosen in the usual righthand sense, so that  $\hat{x} \times \hat{y} = \hat{z}$ , where  $\hat{z}$  is anti-parallel to gravity.

Although we make use of a continuously stratified Boussinesq fluid, the inertial instability considered in this section is not associated with baroclinicity. Rather, as for the centrifugal case in Section 11.3, it is associated with stability of an equilibrium state to horizontal displacements along geopotential surfaces. In a continuously stratified adiabatic fluid, horizontal displacements generally cross isentropic surfaces and so comprise irreversible perturbations. So long as the associated mixing of momentum is negligible, we can still make use of material invariance of potential momentum. We return to this point when studying isentropic inertial (symmetric) instability in Section 11.6, in which perturbations are along sloped buoyancy surfaces rather than geopotential surfaces.

### 11.5.1 Equations of motion relative to the geostrophic equilibrium state

The horizontal momentum equation for an inviscid Boussinesq fluid on an  $f$ -plane is given by (see Section ??)

$$\mathrm{D}\mathbf{u}/\mathrm{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi. \quad (11.49)$$

In the presence of along-front symmetry, an exact solution to the horizontal momentum equation is given by along-front geostrophic flow and zero flow across the front

$$u_g = -f^{-1} \partial_y \varphi \quad \text{and} \quad v = 0 \quad \text{and} \quad \partial_x \varphi = 0. \quad (11.50)$$

We examine the stability of this exact equilibrium base state<sup>5</sup> to perturbations aligned with the front. To study evolution of the perturbations relative to the base state, subtract the base state solution from the full momentum equation (11.49) to render

$$\mathrm{D}u/\mathrm{Dt} = fv \quad (11.51a)$$

$$\mathrm{D}v/\mathrm{Dt} = f(u_g - u). \quad (11.51b)$$

We continue to assume along-front symmetry in the perturbation so that there is no along-front pressure gradient ( $\partial_x \varphi = 0$ ) in equation (11.51a). Equation (11.51b) says that across-front accelerations are determined by deviations from geostrophy of the along-front velocity, and equation (11.51a) says that along-front accelerations are determined by the Coriolis acceleration arising from a non-zero across-front velocity.

Following the treatment of potential momentum in Section ??, we write the along-front momentum equation (11.51a) as the material time derivative of the along-front potential momentum per mass

$$M^x = u - fy, \quad (11.52)$$

bringing the suite of perturbation equations to

$$\mathrm{D}M^x/\mathrm{Dt} = 0 \quad (11.53a)$$

$$\mathrm{D}v/\mathrm{Dt} = f(u_g - u). \quad (11.53b)$$

Material invariance of the along-front potential momentum plays a fundamental role in the stability analysis, analogous to the role of angular momentum for centrifugal instability in Section 11.3. Finally, we can write the equations in terms of just  $M^x$  and  $v$  via

$$\mathrm{D}M^x/\mathrm{Dt} = 0 \quad (11.54a)$$

$$\mathrm{D}v/\mathrm{Dt} = f(M_g^x - M^x). \quad (11.54b)$$

### 11.5.2 Stability analysis based on energetic arguments

We follow the energetic arguments given in Section 11.3.3 for centrifugal instability of cyclostrophic flow. Here, we ask whether the along-front geostrophic flow is stable to a swap of two along-front oriented rows that have the same geopotential. If the swap releases kinetic energy from the base state, then the base state flow is *inertially unstable* to an along-front symmetric perturbation. In that case, perturbations spontaneously initiate inertial instability to affect a return to an inertially stable state.

---

<sup>5</sup>It is useful to confirm that  $u_g = -f^{-1} \partial_y \varphi$  and  $v = 0$  are indeed exact solutions to equation (11.49).

**Instability condition**

The kinetic energy per mass for the along-front geostrophic flow is given by

$$E = u_g^2/2 = (M_g^x + fy)^2/2, \quad (11.55)$$

where we replaced the geostrophic velocity with the geostrophic potential momentum through equation (11.52)

$$M_g^x = u_g - f y. \quad (11.56)$$

The kinetic energy per mass contained in two equal mass parcels at distinct meridional cross-front positions,  $y = y_1$ , and

$$y_2 = y_1 + (y_2 - y_1) = y_1 + \Delta y, \quad (11.57)$$

is given by

$$E_{\text{init}} = [M_g^x(y_1) + fy_1]^2/2 + [M_g^x(y_2) + fy_2]^2/2. \quad (11.58)$$

Swapping the parcels and making use of the invariance of potential momentum leads to the kinetic energy in the swapped state

$$E_{\text{swap}} = [M_g^x(y_1) + fy_2]^2/2 + [M_g^x(y_2) + fy_1]^2/2. \quad (11.59)$$

A bit of algebra leads to the difference in kinetic energy between the two states

$$E_{\text{swap}} - E_{\text{init}} = -f(y_2 - y_1)[M_g^x(y_2) - M_g^x(y_1)] = -f\Delta y\Delta M_g^x. \quad (11.60)$$

Now compute a Taylor series of the potential momentum,  $M_g^x(y_2)$ , relative to the across-front position,  $y = y_1$ ,

$$M_g^x(y_2) \approx M_g^x(y_1) + (y_2 - y_1) \frac{dM_g^x}{dy} \Big|_{y=y_1}, \quad (11.61)$$

which then leads to the energy difference

$$E_{\text{swap}} - E_{\text{init}} = -f(\Delta y)^2 \frac{dM_g^x}{dy} \quad (11.62a)$$

$$= -f(\Delta y)^2 (\partial_y u_g - f) \quad (11.62b)$$

$$= f(\Delta y)^2 (\zeta_g + f), \quad (11.62c)$$

where

$$\zeta_g = -\partial_y u_g \quad (11.63)$$

is the vertical component to the geostrophic relative vorticity for the symmetric base state. Evidently, kinetic energy is released upon swapping the two rows if the following condition is satisfied

$$-f\partial_y M_g^x = f(\zeta_g + f) < 0 \implies \text{inertial instability}. \quad (11.64)$$

The second instability condition says that the base state is unstable if the absolute geostrophic vorticity,  $\zeta_g + f$ , has an opposite sign to the planetary vorticity. This stability condition has a natural generalization to the case of symmetric instability considered in Sections 11.6 and 11.8.

### Interpreting the instability condition

The first instability condition in equation (11.64) arises if the cross-front gradient of the geostrophic potential momentum has the same sign as the Coriolis parameter. To help understand this condition we examine the inertial stability of a base state with zero flow. Zero flow in a rotating reference frame corresponds to rigid-body motion with potential momentum  $M_g^x = -fy$ . For this base state, the inertial stability condition (11.64) is given by

$$-f(\partial M_g^x / \partial y) = f^2 > 0, \quad (11.65)$$

thus signaling the rigid-body flow is inertially stable. By contrast, we infer that an inertially unstable base state is rendered unstable by having an along-front flow that overcomes the stabilizing contribution to potential momentum from planetary rotation. This situation is directly analogous to the centrifugal instability studied in Section 11.3.

### Summarizing the conditions for inertial instability

The instability condition (11.64) takes the following form for the northern and southern hemispheres. Again, the  $x$ -axis is oriented along the front and  $y$ -axis is across the front with  $\hat{x} \times \hat{y} = \hat{z}$  and  $M_g^x = u_g - fy$

$$\text{northern hemisphere } (f > 0): \quad \partial_y M_g^x > 0 \quad \zeta_g < -|f| \quad \partial_y u_g > +|f| \quad (11.66a)$$

$$\text{southern hemisphere } (f < 0): \quad \partial_y M_g^x < 0 \quad \zeta_g > +|f| \quad \partial_y u_g < -|f|. \quad (11.66b)$$

In both hemispheres, instability arises when the relative geostrophic vorticity is anti-cyclonic and larger in magnitude than the cyclonic vorticity from planetary rotation. Under such conditions, inertial instability allows the flow to readjust toward a state of less extreme relative vorticity, thus returning the flow to a state with absolute vorticity dominated by planetary vorticity. Equivalently, inertial instability arises for flows where  $\partial M_g^x / \partial y > 0$  in the northern hemisphere and  $\partial M_g^x / \partial y < 0$  in the southern hemisphere, so that potential momentum of the geostrophic base state increases poleward.

#### 11.5.3 Stability analysis based on parcel arguments

Following the analysis for centrifugal instability in Section 11.3.5, we here consider a parcel stability analysis to provide a force-balance interpretation of the inertial instability. For this purpose, return to the perturbation equations (11.53a) and (11.53b), again with the equilibrium base state of along-front geostrophic balance with zero motion in the across-front direction

$$fu_g = -\partial_y \varphi. \quad (11.67)$$

We examine the stability of this geostrophic base state with respect to along-front perturbations of fluid parcels. For this purpose, imagine moving a row of fluid parcels from position  $y$  to a new position,  $y + \Delta y$ . As per the usual assumption of a parcel analysis, this move is assumed to have no impact on the surrounding fluid environment, including the pressure field.<sup>6</sup> In general, the displaced row of fluid parcels will not be in geostrophic balance at the new position, thus providing for a non-zero acceleration in the across-front direction. If the acceleration is

---

<sup>6</sup>With this assumption, the pressure field remains in geostrophic balance with the unperturbed background geostrophic flow.

directed back to the original position, then the base state is stable and displaced parcels exhibit inertial oscillations in the  $x$ - $y$  plane. In contrast, the base state is inertially unstable if the net acceleration acts to further the initial displacement. Note that while the fluid parcels are moved meridionally, they will be displaced zonally due to the Coriolis acceleration. However, we focus attention on the meridional motion to determine whether the parcel accelerates further away from its initial latitude or returns. We illustrate facets of the stable and unstable situations in Figures 11.5 and 11.6, with details of these figures explained in the remainder of this section.

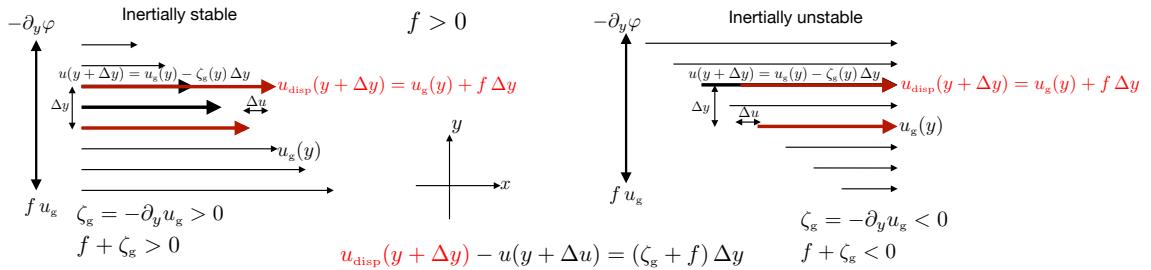


FIGURE 11.5: Schematics of an inertially stable (left panel) and inertially unstable (right panel) geostrophic base flow in the northern hemisphere. One of the red vectors is for the geostrophic velocity,  $u_g(y)$ , and the second for the velocity at a displaced position  $u_{\text{disp}}(y + \Delta y)$ . The displaced velocity is determined by conservation of zonal potential momentum,  $u_{\text{disp}}(y + \Delta y) = u_g(y) + f \Delta y$ , from equation (11.71). In the unperturbed flow, the eastward geostrophic velocity is generated by a northward pressure gradient acceleration that balances a southward Coriolis acceleration. The base state pressure gradient remains unchanged even as the fluid parcels are perturbed, with this assumption basic to the parcel method of stability analysis. At the displaced position we also show the base state velocity at the displaced position,  $u(y + \Delta y) = u_g(y) + \partial_y u(y) \Delta y = u_g(y) - \zeta_g(y) \Delta y$ . The difference between the displaced velocity and the base state velocity is  $u_{\text{disp}}(y + \Delta y) - u(y + \Delta y) = (f + \zeta_g) \Delta y$ . For the left panels, the flow has positive relative vorticity so that  $\zeta_g + f > 0$  and the flow is inertially stable. In this case, the displaced parcel has a southward Coriolis acceleration larger than the local Coriolis, thus returning the row of fluid parcels back towards its initial latitude. For the right panel, the flow has negative relative geostrophic vorticity, and that vorticity is strong enough so that the absolute geostrophic vorticity is negative,  $\zeta_g + f < 0$ , in which case the flow is inertially unstable. In this case, the displaced parcel has a southward Coriolis acceleration smaller than the local Coriolis. Hence, the northward pressure gradient is strong enough to keep the parcel moving northward, away from its initial position, thus signaling a base state that is inertially unstable.

### Mathematical formulation

Equation (11.53b) for the across-front motion is given by

$$Dv/Dt = f(u_g - u). \quad (11.68)$$

This equation holds everywhere, in particular at the displaced cross-front position,  $y + \Delta y$ . At this position, the right hand side has  $u(y + \Delta y)$ , which is the along-front velocity of the displaced parcel at the new position. Likewise,  $u_g(y + \Delta y)$  is the geostrophic velocity at  $y + \Delta y$  of the prescribed background flow. According to equation (11.67), the geostrophic velocity,  $u_g(y + \Delta y)$ , determines a Coriolis acceleration at the displaced position,  $y + \Delta y$ , with this Coriolis acceleration balanced by the cross-front pressure gradient, also evaluated at  $y + \Delta y$ .

To determine the sign of the cross-front acceleration in equation (11.68) acting on the displaced parcel at  $y + \Delta y$ , we make use of the material invariance of along-front potential momentum. This invariance means that each fluid parcel carries its potential momentum from the original position.<sup>7</sup> For example, the potential momentum of a fluid parcel initially at

<sup>7</sup>We offer a cautionary remark about notation. Namely,  $f$  is a constant so that  $f y$  is  $f$  times the latitude position  $y$ , and likewise for  $f(y + \Delta y)$ . In contrast, the zonal velocity,  $u$ , is a function of  $y$ , so that  $u(y)$  and

position  $y$  has the value given by the geostrophic base state at that latitude

$$M^x(y) = M_g^x(y) = u_g(y) - f y. \quad (11.69)$$

Material invariance means that the parcel retains this potential momentum as it moves to the new position,  $y + \Delta y$ . In turn, invariance of along-front potential momentum allows us to determine the along-front velocity of the parcel at the displaced position in terms of  $u_g(y)$  (recall the procedure leading to equation (11.42))

$$M^x(y + \Delta y) = u(y + \Delta y) - f(y + \Delta y) \quad (11.70a)$$

$$= M^x(y) \quad (11.70b)$$

$$= u_g(y) - f y, \quad (11.70c)$$

which leads to

$$u(y + \Delta y) = u_g(y) + f \Delta y. \quad (11.71)$$

The cross-front acceleration (11.68) at the new latitude position thus takes the form

$$\frac{Dv(y + \Delta y)}{Dt} = f [u_g(y + \Delta y) - u(y + \Delta y)] \quad (11.72a)$$

$$= f[u_g(y + \Delta y) - u_g(y) - f \Delta y] \quad (11.72b)$$

$$\approx f \Delta y \left[ \frac{\partial u_g}{\partial y} - f \right] \quad (11.72c)$$

$$= -f \Delta y (\zeta_g + f). \quad (11.72d)$$

The initial displacement grows,

$$\frac{1}{\Delta y} \frac{Dv(y + \Delta y)}{Dt} > 0, \quad (11.73)$$

if the following instability condition holds

$$f(\zeta_g + f) < 0 \implies \text{inertially unstable}, \quad (11.74)$$

which is the same instability condition as found via the energetic arguments in Section 11.5.2.

### Oscillations or exponential growth?

Let  $y$  be the arbitrary initial meridional position for a fluid parcel, and  $\Delta y$  the meridional position relative to  $y$ . The equation of motion (11.72d) can be written in terms of the relative position so that

$$\frac{d^2 \Delta y}{dt^2} + f(\zeta_g + f) \Delta y = 0. \quad (11.75)$$

If  $f(\zeta_g + f) < 0$  then  $\Delta y$  has an exponentially growing (and decaying) solution, thus indicating instability. For the alternative case with  $f(\zeta_g + f) > 0$  we expect to find inertial-like oscillations.

### A quick derivation of the instability condition

The displaced parcel at  $y + \Delta y$  has zonal velocity  $u(y + \Delta y) = u_g(y) + f \Delta y$  according to equation (11.71), which follows from the material conservation of zonal potential momentum. The velocity

---

$u(y + \Delta y)$  represent the zonal velocity evaluated at the latitude positions  $y$  and  $y + \Delta y$ , respectively.

of the background geostrophic flow at  $y + \Delta y$  is given by the Taylor series approximation relative to the flow at  $y$  (keeping terms up to first order):

$$u_g(y + \Delta y) \approx u_g(y) + \Delta y \partial_y u_g(y) = u_g(y) - \Delta y \zeta_g(g). \quad (11.76)$$

The difference between these two zonal velocities is

$$u_g(y + \Delta y) - u(y + \Delta y) = -(f + \zeta_g) \Delta y, \quad (11.77)$$

which recovers the inertial stability condition derived above.

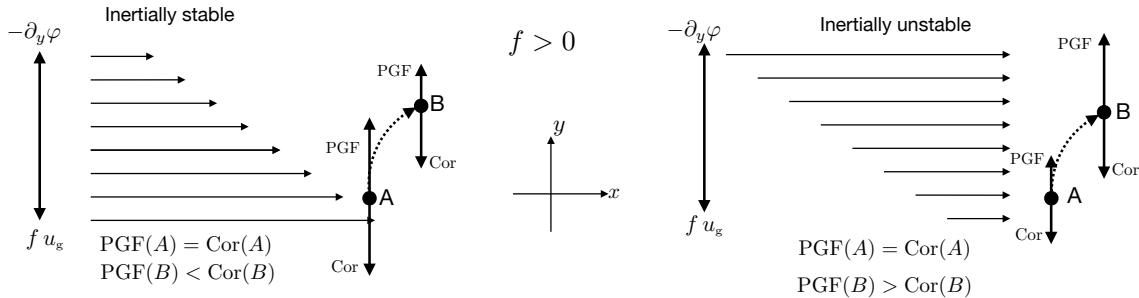


FIGURE 11.6: Further depiction of inertially stable geostrophic flow (left panel) and unstable geostrophic flow (right panel), complementing features presented in Figure 11.5. Here we show a fluid parcel starting at point  $A$  and displaced to point  $B$ , with the curved motion due to the zonal Coriolis acceleration that acts when the parcel is displaced in the meridional direction ( $f > 0$ ). For the inertially stable case, the pressure gradient acceleration acting on the displaced parcel at point  $B$  is weaker than the parcel's Coriolis acceleration, so that the parcel exhibits stable inertial oscillations in the  $x$ - $y$  plane. For the inertially unstable case, the pressure gradient acceleration acting on the parcel at point  $B$  is stronger than the parcel's Coriolis acceleration, so that the parcel continues to move further away and thus exhibits inertial instability.

### Summary of the parcel argument

As summarized in Figures 11.5 and 11.6, at the initial location in the base state, a parcel under geostrophic balance has its Coriolis acceleration balanced by its pressure acceleration. However, the displaced parcel generally will not be in geostrophic balance at the new location, in which case its Coriolis acceleration does not balance the local pressure gradient. The parcel approach determines a stability condition based on whether the imbalance leads to an acceleration back towards its initial position (stable inertial oscillations) or further away (inertial instability). For the unstable case, the Coriolis acceleration of the displaced parcel at its new location cannot match the local pressure gradient, thus causing the parcel to be thrust away from its initial latitude.

Along-front symmetry of both the base state and the perturbation ensures material invariance of along-front potential momentum. This invariance provides an explicit expression for the acceleration felt by the displaced parcel, thus determining a condition on stability of the base state to such symmetric perturbations. The method of analysis is directly analogous to that applied to the rotating tank of fluid in Section 11.3 for centrifugal instability, as well as for a vertical column of fluid in Section ?? for gravitational stability. We follow this procedure again in Section 11.6 for studying symmetric instability.

## 11.6 Symmetric instability and the $f Q < 0$ criteria

In this section we examine stability of a geostrophically balanced baroclinic front on an  $f$ -plane. We continue to assume the front is gravitationally stable ( $N^2 > 0$ ) and that it exhibits along-front symmetry so that the along-front potential momentum is a material invariant (Figure 11.4). Different from the inertial instability case in Section 11.5, we here assume the fluid to be adiabatic so that Archimedean buoyancy is a second materially invariant property. We investigate stability of a geostrophically balanced along-front flow using the parcel method and make use of buoyancy as the vertical coordinate. We follow up this analysis in Section 11.8 with an Eulerian modal analysis. Both approaches reveal the fundamental importance of potential vorticity as a signature of symmetric instability. More precisely, the product,  $f Q$ , if negative, indicates a symmetrically unstable front.

### 11.6.1 Equations using buoyancy as the vertical coordinate

Given the adiabatic nature of the fluid, and the role of baroclinicity, we are motivated to use buoyancy as the vertical coordinate for a Boussinesq fluid as detailed in Section ???. That is, we consider symmetric displacements of parcels along a constant buoyancy surface. By construction, this displacement is neutral to gravitational instability since it occurs along a constant buoyancy surface. However, a displaced parcel could still find itself in an unstable position depending on the potential momentum of the local environment relative to the buoyancy surfaces. The analysis proceeds just like for the horizontal inertial instability of Section 11.5.3, yet with the displacements here being isentropic (constant buoyancy) rather than horizontal (constant geopotential). We are thus motivated to refer to the ensuing instability as *isentropic inertial instability*, though note that it is more commonly referred to as *symmetric instability*.

With buoyancy as the vertical coordinate, the horizontal momentum equation is<sup>8</sup>

$$D\mathbf{u}/Dt + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_{hb} M \quad (11.78)$$

where  $\nabla_{hb}$  is a horizontal gradient computed along constant buoyancy surfaces. Furthermore,

$$M = \varphi - b z \quad (11.79)$$

is the Montgomery potential that contributes an acceleration via<sup>9</sup>

$$-\nabla_{hb} M = -\nabla_{hb} \varphi + b \nabla_{hb} z. \quad (11.80)$$

The first term arises from pressure gradients and the second from geopotential gradients, both computed along constant buoyancy surfaces. In the presence of along-front symmetry, an exact solution to the horizontal momentum equation is given by along-front geostrophic flow and zero cross-front flow

$$u_g = -\frac{1}{f \rho_0} \left[ \frac{\partial M}{\partial y} \right]_b \quad \text{and} \quad v = 0 \quad \text{and} \quad \left[ \frac{\partial M}{\partial x} \right]_b = 0. \quad (11.81)$$

<sup>8</sup>In Part ?? of this book, we develop the mathematical and physical details for generalized vertical coordinates, such as buoyancy or isopycnal coordinates. In particular, the essential features of the  $\nabla_{hb}$  gradient operator are provided in Figure ???. For purposes of the present chapter, one merely needs to know that  $\nabla_{hb}$  provides a measure of the property gradients computed along constant buoyancy surfaces, and yet with  $\nabla_{hb}$  still having only horizontal components:  $\nabla_{hb} = \hat{\mathbf{x}} (\partial/\partial x)_b + \hat{\mathbf{y}} (\partial/\partial y)_b$ , where the  $b$  subscript means that the derivative is computed while holding  $b$  constant.

<sup>9</sup>Be careful to distinguish the Montgomery potential,  $M$ , in equation (11.79) from the zonal potential momentum,  $M^x$ , in equation (11.40).

We examine stability of this base state to perturbations that are symmetric in the along-front direction. Subtracting the exact equilibrium solution from the full momentum equation (11.49) leads to

$$\frac{Du}{Dt} = fv \quad \text{and} \quad \frac{Dv}{Dt} = f(u_g - u), \quad (11.82)$$

where we assume along-front symmetry thus allowing us to drop the along-front gradient of the Montgomery potential. Following the treatment in Sections ?? and 11.5.1, we write the along-front momentum equation as the material time derivative of the along-front potential momentum per mass,  $M^x = u - fy$  (equation (11.52)), thus bringing the perturbation equations to

$$\frac{DM^x}{Dt} = 0 \quad \text{and} \quad \frac{Dv}{Dt} = f(M_g^x - M). \quad (11.83)$$

### 11.6.2 Parcel stability analysis

We follow the parcel analysis used for inertial instability in Section 11.5.3, starting with an equilibrium base state of along-front geostrophic balance with zero meridional motion and then examine the stability of this state to symmetric perturbations of fluid parcels along a constant buoyancy surface. For this purpose, imagine moving a row of fluid parcels from cross-front position  $y$  to position  $y + \Delta y$  while maintaining a fixed buoyancy *and* fixed potential momentum. Furthermore, assume the displacement has no impact on the base state, which is the typical assumption of the parcel approach to stability analysis. In general, the displaced row of parcels will not be in geostrophic balance at the new position, thus providing for a non-zero cross-front acceleration at that displaced position. The sign of that acceleration determines stability of the flow to the symmetric perturbation.

At the new cross-front position, the meridional acceleration acting on the displaced parcels is given by<sup>10</sup>

$$\frac{Dv(y + \Delta y)}{Dt} = f [M_g^x(y + \Delta y) - M^x(y + \Delta y)], \quad (11.84)$$

where  $M^x(y + \Delta y)$  is the potential momentum at  $y + \Delta y$ , and  $M_g^x(y + \Delta y)$  is the geostrophic potential momentum at  $y + \Delta y$ . To determine the sign of the acceleration acting on the displaced parcel, we make use of the material invariance of along-front potential momentum,

$$M^x(y + \Delta y) = M^x(y) = M_g^x(y), \quad (11.85)$$

so that the meridional acceleration is

$$\frac{Dv(y + \Delta y)}{Dt} = f \left[ \frac{\partial M_g^x}{\partial y} \right]_b \Delta y. \quad (11.86)$$

Evidently, if  $\Delta y > 0$  then the displacement is unstable if

$$f \left[ \frac{\partial M_g^x}{\partial y} \right]_b > 0 \implies \text{symmetrically unstable geostrophic state.} \quad (11.87)$$

We can write this condition in terms of the absolute geostrophic vorticity by noting that

$$\left[ \frac{\partial M_g^x}{\partial y} \right]_b = \left[ \frac{\partial u_g}{\partial y} \right]_b - f = -(\tilde{\zeta}_g + f), \quad (11.88)$$

---

<sup>10</sup>We only expose the meridional coordinate dependence in equation (11.84), but note that the velocity and potential momentum are also a function of the vertical position and time.

where  $\tilde{\zeta}_g$  is the relative vorticity of the geostrophic flow computed on buoyancy surfaces (see Section ??)

$$\tilde{\zeta}_g = \left[ \frac{\partial v_g}{\partial x} \right]_b - \left[ \frac{\partial u_g}{\partial y} \right]_b, \quad (11.89)$$

with  $v_g = 0$  for this zonal geostrophic front. Evidently, we have arrived at the alternative expression of a symmetrically unstable geostrophic base flow

$$f(f + \tilde{\zeta}_g) < 0 \implies \text{symmetrically unstable geostrophic state.} \quad (11.90)$$

Finally, we can introduce the Boussinesq Ertel potential vorticity in the form<sup>11</sup>

$$Q = (\boldsymbol{\omega} + f \hat{\mathbf{z}}) \cdot \nabla b = (\tilde{\zeta} + f) N^2, \quad (11.91)$$

with  $N^2 = \partial_z b > 0$  the squared buoyancy frequency for gravitationally stable flows. Bringing these results together leads to the equivalent expressions of the necessary and sufficient conditions for a geostrophic base flow to be symmetrically unstable

$$f \left[ \frac{\partial M_g^x}{\partial y} \right]_b > 0 \quad (11.92a)$$

$$f(\tilde{\zeta}_g + f) < 0 \quad (11.92b)$$

$$f Q_g < 0. \quad (11.92c)$$

### 11.6.3 Geometric perspective on the instability condition

The instability condition (11.92b) is a direct translation of the inertial instability condition (11.74), swapping horizontal displacements with displacements along buoyancy surfaces. Likewise, the instability condition (11.92a) is directly analogous to the inertial instability condition (11.64), only here with displacement along a buoyancy surface rather than a geopotential. In the northern hemisphere, if one moves in the  $+\hat{\mathbf{y}}$  direction on a constant buoyancy surface and encounters increasing values for the potential momentum, then the base state flow is symmetrically unstable. Conversely in the southern hemisphere, if one moves in  $-\hat{\mathbf{y}}$  direction on a constant buoyancy surface and encounters increasing values for the potential momentum, then the base state flow is symmetrically unstable. We illustrate this situation in Figure 11.7.

As another way to write the instability condition (11.92a), make use of the expression (??) to transform the derivative on buoyancy surfaces back to geopotential coordinates

$$\left[ \frac{\partial M_g^x}{\partial y} \right]_b = \left[ \frac{\partial M_g^x}{\partial y} \right]_z - \frac{\partial b / \partial y}{\partial b / \partial z} \frac{\partial M_g^x}{\partial z} \quad (11.93a)$$

$$= (\partial b / \partial z)^{-1} \left[ \frac{\partial M_g^x}{\partial y} \frac{\partial b}{\partial z} - \frac{\partial M_g^x}{\partial z} \frac{\partial b}{\partial y} \right] \quad (11.93b)$$

$$= -N^{-2} \hat{\mathbf{x}} \cdot (\nabla b \times \nabla M_g^x). \quad (11.93c)$$

Similarly, we have

$$\left[ \frac{\partial b}{\partial y} \right]_{M_g^x} = (\partial M_g^x / \partial z)^{-1} \left[ \frac{\partial b}{\partial y} \frac{\partial M_g^x}{\partial z} - \frac{\partial b}{\partial z} \frac{\partial M_g^x}{\partial y} \right] \quad (11.94a)$$

---

<sup>11</sup>We derive equation (11.91) when studying potential vorticity in the Boussinesq ocean using buoyancy coordinates; see equation (??).

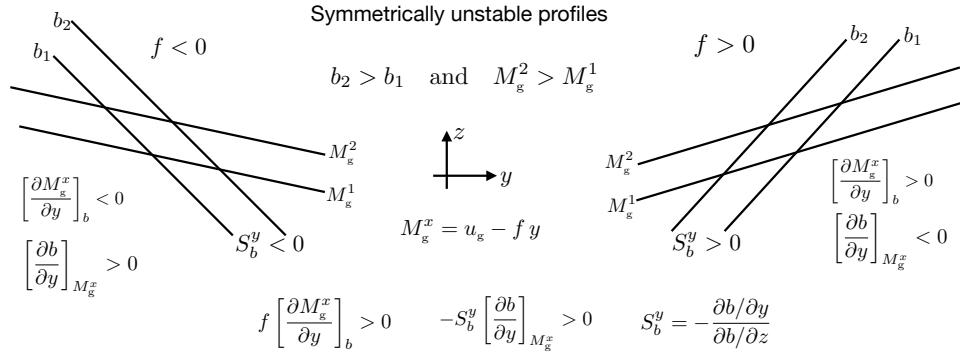


FIGURE 11.7: Example geostrophically balanced flow configurations that are symmetrically unstable; i.e., inertially unstable to a perturbation with symmetry along the front. We show example buoyancy surfaces and potential momentum surfaces for the southern hemisphere (left) and northern hemisphere (right). The instability conditions (11.99a) and (11.99b) are indicated on the respective panels. In both cases, surfaces of constant buoyancy are more steeply sloped than constant potential momentum surfaces. The  $x$  coordinate measures distance in the along-front direction and  $y$  measures distance in the cross-front direction, oriented so that  $\hat{x} \times \hat{y} = \hat{z}$  where  $\hat{z}$  is anti-parallel to gravity ( $\hat{x}$  is out of the page). A means to quickly judge whether a flow is symmetrically unstable is to note that the wedge of instability region between buoyancy and potential momentum surfaces provides a source of available potential energy. Symmetric instability can feed off the potential energy only when buoyancy surfaces are more steeply sloped than potential momentum surfaces.

$$= (\partial M_g^x / \partial z)^{-1} \hat{x} \cdot (\nabla b \times \nabla M_g^x), \quad (11.94b)$$

which then leads to

$$\frac{(\partial b / \partial y)_{M_g^x}}{\partial b / \partial z} = - \frac{(\partial M_g^x / \partial y)_b}{\partial M_g^x / \partial z}. \quad (11.95)$$

Furthermore, with  $M_g^x = u_g - fy$  we have

$$\partial_z M_g^x = \partial_z u_g = -f^{-1} (\partial b / \partial y)_z, \quad (11.96)$$

so that

$$S_b^y (\partial b / \partial y)_{M_g^x} = -f (\partial M_g^x / \partial y)_b, \quad (11.97)$$

where we introduced the slope of the meridional buoyancy surface

$$S_b^y \equiv - \frac{\partial b / \partial y}{\partial b / \partial z}. \quad (11.98)$$

Use of the identity (11.97) in the instability condition (11.92a) leads to the equivalent expressions of a symmetrically unstable geostrophically balanced state

$$f \left[ \frac{\partial M_g^x}{\partial y} \right]_b = -S_b^y \left[ \frac{\partial b}{\partial y} \right]_{M_g^x} > 0 \implies \text{symmetrically unstable} \quad (11.99a)$$

$$f N^2 \left[ \frac{\partial M_g^x}{\partial y} \right]_b = \left[ \frac{\partial b}{\partial y} \right]_z \left[ \frac{\partial b}{\partial y} \right]_{M_g^x} > 0 \implies \text{symmetrically unstable}. \quad (11.99b)$$

Consider a configuration with buoyancy slopes rising to the north, so that the meridional buoyancy slope is positive,  $S_b^y > 0$ . Equation (11.99a) says that if the buoyancy decreases moving in the  $+\hat{y}$  direction while moving along constant potential momentum surfaces, then the flow is symmetrically unstable. Equivalently, equation (11.99b) says that if buoyancy has

a meridional derivative of the same sign on both a constant  $z$  surface and a constant  $M_g^x$  surface, then the flow is symmetrically unstable. In Figure 11.7 we depict various properties of symmetrically unstable configurations.

## 11.7 Symmetric instability and the wedge of instability

We here formulate the stability of a symmetric front using geopotential vertical coordinate equations, thus serving as a complement to the analysis in Section 11.6 based on buoyancy as the vertical coordinate. Furthermore, we here focus on perturbations that lead to an instability if presented with a base flow state that is unstable. More precisely, we consider zonally symmetric perturbations (i.e., a row of fluid parcels) that carry buoyancy and zonal potential momentum from the initial position into a new position. The question is whether the displaced row of fluid parcels feels a net force that sends it back to where it came (stable perturbation), or if the force sends it further away (unstable perturbation). Through this analysis we show that flow configurations with buoyancy and potential momentum surfaces oriented as in Figure 11.7 admit unstable perturbations that probe the *wedge of instability* (see Figure 11.8). The existence of such perturbations is a signature of a symmetrically unstable background flow state.

### 11.7.1 Formulation of the basic equations

Take the inviscid Boussinesq equation (??) as a starting point, here written as

$$DM^x/Dt = 0 \quad \text{and} \quad Dv/Dt = f(M_g^x - M^x) \quad \text{and} \quad Dw/Dt = -\partial_z \varphi + b. \quad (11.100)$$

The base state is assumed to be in geostrophic and hydrostatic balance, in which

$$\text{base flow state} \iff Dv/Dt = 0 \quad \text{and} \quad M^x = M_g^x \quad \text{and} \quad \partial_z \varphi = bg. \quad (11.101)$$

To examine stability of this base state, displace a row of fluid parcels from its original  $y$ - $z$  position,  $(y_A, z_A)$ , to a new position,

$$y_B = y_A + \Delta y \quad \text{and} \quad z_B = z_A + \Delta z, \quad (11.102)$$

such as depicted in Figure 11.8. Notably, the buoyancy and potential momentum at point  $(y_B, z_B)$  are distinct from those at  $(y_A, z_A)$ . Importantly, the displacement materially preserves the buoyancy and potential momentum of the displaced row of parcels, so that

$$b_{\text{parcel}} = b(A) \quad \text{and} \quad M_{\text{parcel}}^x = M^x(A) = M_g^x(A). \quad (11.103)$$

At the displaced position, the row of parcels finds itself in a local environment where it feels a nonzero acceleration. Does this acceleration act to further the displacement (unstable perturbation) or return the row of parcels to its original position (stable perturbation)?

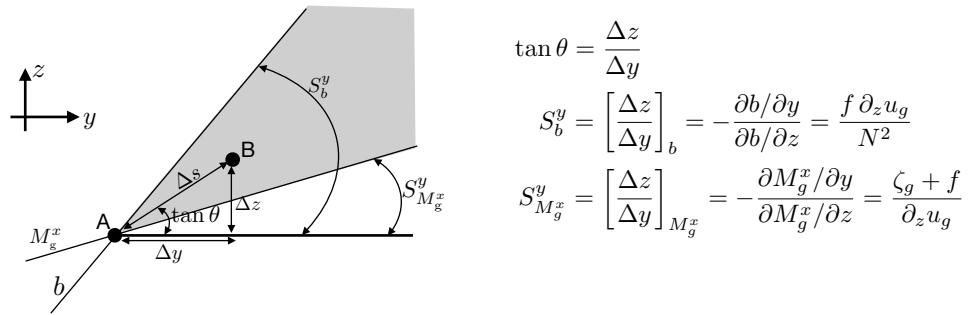


FIGURE 11.8: Illustrating the parcel argument for symmetric instability, whereby we displace a zonally ( $\hat{x}$  points out of the page) symmetric row of parcels from  $A$  to  $B$  while retaining their buoyancy,  $b(A)$ , and potential momentum,  $M^x(A) = M_g^x(A)$ . We also depict the slope of the displacement,  $\tan \theta = \Delta z / \Delta y$ , the slope of the buoyancy surface,  $S_b^y = (\Delta z / \Delta y)_b = f \partial_z u_g / N^2$ , and slope of the potential momentum surface,  $S_{M_g^x}^y = (\Delta y / \Delta z)_{M_g^x} = (\zeta_g + f) / \partial_z u_g$ . When the buoyancy surface is sloped more steeply than the potential momentum surface, then that base state flow is subject to symmetrically unstable perturbations, with the unstable perturbations extending into the gray wedge of instability.

### 11.7.2 Assessing stability of the perturbation

To answer the above stability question, write the acceleration for the parcel displaced from point  $A$  to point  $B$  in the form

$$\frac{D\mathbf{v}_{\text{parcel}}}{Dt} = f [M_g^x(B) - M_g^x(\text{parcel})] \hat{\mathbf{y}} + [b_{\text{parcel}} - \partial_z \varphi(B)] \hat{\mathbf{z}} \quad (11.104a)$$

$$= f [M_g^x(B) - M_g^x(\text{parcel})] \hat{\mathbf{y}} + [b_{\text{parcel}} - b(B)] \hat{\mathbf{z}}, \quad (11.104b)$$

where  $M_g^x$  and  $\partial_z \varphi$  are taken from the base state at point  $B$ , and for the second equality we set the base state vertical pressure gradient equal to the buoyancy according to the hydrostatic balance,  $\partial_z \varphi(B) = b(B)$ . We are interested in whether this displacement leads to an acceleration of the parcel further in the direction of the displacement, beyond point  $B$ , in which case the base state is symmetrically unstable, otherwise it is stable.

To determine stability conditions, use material invariance of the potential momentum and buoyancy in the form of equation (11.103), thus leading to the parcel acceleration

$$\frac{D\mathbf{v}_{\text{parcel}}}{Dt} = f [M_g^x(B) - M_g^x(A)] \hat{\mathbf{y}} + [b(A) - b(B)] \hat{\mathbf{z}}. \quad (11.105)$$

For orientation, let the unit vector in the direction of the displacement be written

$$\hat{\mathbf{s}} = \hat{\mathbf{y}} \cos \theta + \hat{\mathbf{z}} \sin \theta = \hat{\mathbf{y}} (\Delta y / \Delta s) + \hat{\mathbf{z}} (\Delta z / \Delta s), \quad (11.106)$$

where  $\theta$  is the angle of the displacement with respect to the horizontal plane, and

$$\Delta s = \sqrt{(\Delta y)^2 + (\Delta z)^2}, \quad (11.107)$$

is the distance of the displacement (see Figure 11.8). Taking the  $\hat{\mathbf{s}}$  projection of the parcel acceleration (11.105) leads to

$$\hat{\mathbf{s}} \cdot \frac{D\mathbf{v}_{\text{parcel}}}{Dt} = f [M_g^x(B) - M_g^x(A)] \cos \theta - [b(B) - b(A)] \sin \theta. \quad (11.108)$$

Now expand the terms on the right hand side in a Taylor series, expressing values at point  $B$  in terms of those at point  $A$ , so that (keeping terms up to first order)

$$b(B) - b(A) = \partial_y b \Delta y + \partial_z b \Delta z = \Delta y (\partial_y b + \partial_z b \tan \theta) \quad (11.109a)$$

$$M_g^x(B) - M_g^x(A) = \partial_y M_g^x \Delta y + \partial_z M_g^x \Delta z = \Delta y (\partial_y M_g^x + \partial_z M_g^x \tan \theta), \quad (11.109b)$$

with terms on the right hand side evaluated at point  $A$ . These expansions bring the equation of motion (11.108) to

$$\hat{s} \cdot \frac{D\mathbf{v}_{\text{parcel}}}{Dt} = f \Delta y (\partial_y M_g^x + \partial_z M_g^x \tan \theta) \cos \theta - \Delta y (\partial_y b + \partial_z b \tan \theta) \sin \theta. \quad (11.110)$$

For a northward and upward perturbation as depicted in Figure 11.8, the symmetric perturbation is unstable if the right hand side of equation (11.110) is positive, which requires

$$f (\partial_y M_g^x + \partial_z M_g^x \tan \theta) \cos \theta > (\partial_y b + N^2 \tan \theta) \sin \theta, \quad (11.111)$$

which can be written in terms of the slope of the buoyancy surface,  $S_b^y$ , the slope of the potential momentum surface,  $S_{M_g^x}^y$ , and the slope of the perturbation,  $\tan \theta$

$$f \partial_z M_g^x (\tan \theta - S_{M_g^x}^y) > \partial_z b (\tan \theta - S_b^y). \quad (11.112)$$

Finally, we make use of thermal wind

$$f \partial_z M_g^x = f \partial_z u_g = -\partial_y b, \quad (11.113)$$

so that the inequality (11.112) is

$$S_b^y (\tan \theta - S_{M_g^x}^y) > \tan \theta - S_b^y. \quad (11.114)$$

We now assume the following properties of the base state

$$S_b^y > 0 \quad \text{and} \quad S_{M_g^x}^y > 0 \quad \text{and} \quad \tan \theta > 0, \quad (11.115)$$

which correspond to the base state configuration in Figure 11.8. The inequality (11.114) yields the following condition for symmetric instability (as signaled by an acceleration in the direction of the displacement)

$$S_{M_g^x}^y < \tan \theta < S_b^y. \quad (11.116)$$

This slope condition means that a symmetric perturbation that falls within the wedge of instability in Figure 11.8 initiates a symmetric instability. Conversely, perturbations that fall outside of the wedge do not initiate a symmetric instability.

### 11.7.3 Symmetrically unstable while inertially and gravitationally stable

An alternative means to write the parcel acceleration equation (11.110) is given by

$$\hat{s} \cdot \frac{D\mathbf{v}_{\text{parcel}}}{Dt} = \Delta y \cos \theta [-f(f + \zeta_g) - N^2 \tan^2 \theta + 2 f \partial_z u_g \tan \theta], \quad (11.117)$$

where we used  $\partial_y M_g^x = -(f + \zeta_g)$  and  $\partial_z b = N^2$ . The assumptions (11.115) are equivalent to (it is useful to prove this)

$$f(f + \zeta_g) > 0 \quad \text{and} \quad N^2 > 0, \quad (11.118)$$

which means the base state flow is stable to horizontal inertial instability and to gravitational instability. We thus find from equation (11.117) that the perturbation is unstable if

$$2f\partial_z u_g \tan \theta > f(f + \zeta_g) + N^2 \tan^2 \theta. \quad (11.119)$$

This instability condition is equivalent to the purely geometric form given by equation (11.116).

## 11.8 Symmetric instability and linear modal analysis

We now consider symmetric instability following the linear stability analysis of [Hoskins \(1974\)](#). This approach recovers the same stability condition as for the parcel method in Section 11.6. However, linear stability analysis offers added insights concerning the central role of Ertel potential vorticity and it provides further details of the instability. The mathematical analysis here also offers useful practice for the study of baroclinic instability in Chapter 14.

### 11.8.1 Geostrophic base state and the perturbation equations

Our goal is to examine the stability of a zonally symmetric geostrophic front to zonally symmetric perturbations. Whereas the base state for the study of gravity waves has zero baroclinicity, here we consider a geostrophic front with meridional (across-front) baroclinicity. Rather than the buoyancy coordinates used for the parcel stability analysis in Section 11.6.1, we here make use of geopotential vertical coordinates and linearize the Boussinesq ocean equations. The starting point for this development is the  $f$ -plane version of the perfect Boussinesq ocean equations from Section ??

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + f \hat{\mathbf{z}} \times \mathbf{v} = -\nabla \varphi + b \hat{\mathbf{z}} \quad (11.120a)$$

$$\partial_t b + \mathbf{v} \cdot \nabla b = 0. \quad (11.120b)$$

#### Equations for the geostrophic base state

The equations describing the zonal geostrophic and hydrostatic base state are given by

$$f u_g = -\partial_y \varphi_g \quad \text{zonal flow in geostrophic balance} \quad (11.121a)$$

$$\partial_z \varphi_g = b_g \quad \text{hydrostatic balance} \quad (11.121b)$$

$$f \partial_z u_g = -\partial_y b_g \quad \text{thermal wind of zonal flow} \quad (11.121c)$$

$$\zeta_g = -\partial_y u_g \quad \text{relative vorticity.} \quad (11.121d)$$

The pressure field,  $\varphi_g(y, z)$ , has a meridional derivative that supports the zonal geostrophic flow,  $u_g(y, z)$ , and which is in hydrostatic balance with the base state buoyancy,  $b_g(y, z)$ . Combining geostrophy with the hydrostatic balance leads to the thermal wind shear in equation (11.121c). Due to zonal symmetry, the only contribution to the relative vorticity comes from the meridional shear of the zonal geostrophic flow,  $\zeta_g = -\partial_y u_g$ . Ertel potential vorticity (Chapter ??) of the base state proves to be central to the stability conditions

$$Q_g = (\omega_g + f \hat{\mathbf{z}}) \cdot \nabla b_g = N^2 (f + \zeta_g) - f (\partial_z u_g)^2, \quad (11.122)$$

where we made use of the thermal wind equation (11.121c) as well as the relative vorticity for the zonal geostrophic background flow

$$\omega_g = \nabla \times \hat{x} u_g = \hat{y} \partial_z u_g - \hat{z} \partial_y u_g = -\hat{y} f^{-1} \partial_y b_g + \hat{z} \zeta_g. \quad (11.123)$$

### Linearized perturbation equations

To develop perturbation equations, we decompose the flow into the geostrophic base state plus a perturbation. The perturbation is time dependent and ageostrophic, and, importantly, is assumed to retain the zonal symmetry of the base state

$$\mathbf{v}(y, z, t) = \mathbf{v}'(y, z, t) + \hat{x} u_g(y, z) \quad (11.124a)$$

$$\varphi(y, z, t) = \varphi'(y, z, t) + \varphi_g(y, z) \quad (11.124b)$$

$$b(y, z, t) = b'(y, z, t) + b_g(y, z). \quad (11.124c)$$

The linearized version of the velocity equation (11.120a) and buoyancy equation (11.120b) take the form

$$\partial_t u' = v' (f - \partial_y u_g) - w' \partial_z u_g \quad \text{zonal velocity} \quad (11.125a)$$

$$\partial_t v' = -f u' - \partial_y \varphi' \quad \text{meridional velocity} \quad (11.125b)$$

$$\partial_t w' = -\partial_z \varphi' + b' \quad \text{vertical velocity} \quad (11.125c)$$

$$\partial_t b' = -(v' \partial_y b_g + w' \partial_z b_g) \quad \text{buoyancy} \quad (11.125d)$$

$$\partial_y v' + \partial_z w' = 0 \quad \text{continuity.} \quad (11.125e)$$

Observe that the zonal velocity evolves according to the Coriolis acceleration plus the convergence of the advective flux of the base state geostrophic velocity

$$-v' \partial_y u_g - w' \partial_z u_g = -\nabla \cdot (\mathbf{v}' u_g). \quad (11.126)$$

Likewise, the buoyancy evolves according to the convergence of the advective flux of the base state buoyancy

$$-(v' \partial_y b_g + w' \partial_z b_g) = -\nabla \cdot (\mathbf{v}' b_g). \quad (11.127)$$

In both cases, the advection velocity arises from the non-divergent anomalous flow,  $\mathbf{v}'$ , with zonal symmetry meaning there is no contribution to the flux convergences from the anomalous zonal velocity,  $u'$ .

### 11.8.2 Inertia-vorticity oscillator equations

As in our study of buoyancy oscillations in Section 9.3 and inertia-gravity oscillations in Section 9.7, we here develop forced oscillator equations for the zonal and meridional velocity components. To do so, take a time derivative of the zonal velocity equation (11.125a) and then substitute the meridional velocity equation (11.125b) to find the zonal velocity oscillator equation

$$[\partial_{tt} + f(f + \zeta_g)] u' = -(f + \zeta_g) \partial_y \varphi' + \partial_z u_g (\partial_z \varphi' - b'). \quad (11.128)$$

Similar manipulations lead to the forced oscillator equation for the meridional velocity

$$[\partial_{tt} + f(f + \zeta_g)] v' = f w' \partial_z u_g - \partial_{ty} \varphi'. \quad (11.129)$$

The forcing functions on the right hand side of both equations (11.128) and (11.129) are generally nonzero. However, in their absence both the zonal and meridional velocity components satisfy a free oscillator equation with squared natural angular frequency,  $f(f + \zeta_g)$ . We refer to these as *inertia-vorticity* oscillations given the dual role of the Coriolis and vorticity for determining the oscillations. The motion is oscillatory if  $f(f + \zeta_g) > 0$ , whereas it is exponentially growing if  $f(f + \zeta_g) < 0$ . This stability condition for the oscillators is consistent with that found for the parcel stability analysis leading to equation (11.92a). To develop more insight into the nature of the instability for  $f(f + \zeta_g) < 0$ , we next develop an equation for the overturning streamfunction in the meridional-vertical plane.

### 11.8.3 Ageostrophic overturning circulation streamfunction

The continuity equation (11.125e) means that the meridional-vertical circulation is non-divergent so that we can introduce a streamfunction for the perturbation flow

$$\mathbf{v}' = -\hat{\mathbf{x}} \times \nabla \psi \implies v' = \partial_z \psi \quad \text{and} \quad w' = -\partial_y \psi. \quad (11.130)$$

The perturbed fields are ageostrophic and are sometimes referred to as the *ageostrophic secondary circulation*, with this circulation described by the ageostrophic overturning circulation streamfunction,  $\psi$ . Based on the oscillator equations (11.128) and (11.129), we anticipate that the overturning circulation is unstable for  $f(f + \zeta_g) < 0$ .

#### Zonal vorticity equation

Taking the  $z$  derivative of the meridional velocity equation (11.125b) and the  $y$  derivative of the vertical velocity equation (11.125c), and then subtracting, leads to the equation for the zonal component of the vorticity<sup>12</sup>

$$\partial_t(\partial_y w' - \partial_z v') - f \partial_z u' = \partial_y b', \quad (11.131)$$

which takes on the form using the overturning streamfunction

$$\partial_t(\partial_{yy} + \partial_{zz}) \psi = -f \partial_z u' - \partial_y b'. \quad (11.132)$$

#### Further assumptions for the base state flow

To enable the next steps of the derivation, which makes use of a plane wave ansatz in Section 11.8.4, we make use of the following assumptions about the base state:

$$\zeta_g = -\partial_y u_g \quad \text{constant relative vorticity} \quad (11.133a)$$

$$f \partial_z u_g = -\partial_y b_g \quad \text{constant thermal wind shear} \quad (11.133b)$$

$$N^2 = \partial_z b_g \quad \text{constant buoyancy frequency.} \quad (11.133c)$$

These assumptions are restrictive, and yet the resulting instability calculation provides great insights into base state flow configurations that are more general.

---

<sup>12</sup>The zonal component of the perturbation vorticity is given by  $\hat{\mathbf{x}} \cdot (\nabla \times \mathbf{v}') = \partial_y w' - \partial_z v' = -(\partial_{yy} + \partial_{zz})\psi$ .

### Eliminating the zonal velocity and the buoyancy

We can eliminate the perturbation zonal velocity,  $u'$ , and the perturbation buoyancy,  $b'$ , from the vorticity equation (11.132) by taking another time derivative. Upon doing so we make use of the zonal velocity equation (11.125a) and the buoyancy equation (11.125d), and make use of the assumptions (11.133a)-(11.133c) for the base state:

$$-\partial_{tt}(\partial_{yy} + \partial_{zz})\psi = f\partial_{zt}u' + \partial_{yt}b' \quad (11.134a)$$

$$= f\partial_z[v'(f + \zeta_g) - w'\partial_zu_g] - \partial_y(v'\partial_yb_g + w'\partial_zb_g) \quad (11.134b)$$

$$= f\partial_{zz}\psi(f + \zeta_g) - 2\partial_{yz}\psi\partial_yb_g + \partial_{yy}\psi\partial_zb_g \quad (11.134c)$$

$$= [f(f + \zeta_g)\partial_{zz} + 2f\partial_zu_g\partial_{yz} + N^2\partial_{yy}]\psi. \quad (11.134d)$$

Rearrangement leads to the equation for the ageostrophic overturning circulation streamfunction

$$[\partial_{tt}(\partial_{yy} + \partial_{zz}) + f(f + \zeta_g)\partial_{zz} + 2f\partial_zu_g\partial_{yz} + N^2\partial_{yy}]\psi = 0. \quad (11.135)$$

#### 11.8.4 Dispersion relation for meridional-vertical plane waves

Equation (11.135) is a constant coefficient partial differential equation for the overturning streamfunction. We examine its properties by considering a plane wave ansatz just as for our study of linear waves in Part I of this book. Since the flow is non-divergent, the plane waves are transverse. Furthermore, we continue to assume zonal symmetry with the perturbations, so that the waves have no  $x$  dependence. We thus consider waves propagating in the meridional-vertical plane, in which case we take the plane wave ansatz

$$\psi(y, z, t) = \tilde{\psi}e^{i(k_y y + k_z z - \omega t)}, \quad (11.136)$$

with Figure 11.9 depicting the wave geometry. Plugging this ansatz into the streamfunction equation (11.135) leads to the dispersion relation

$$\omega^2 = \frac{f(f + \zeta_g)k_z^2 + 2f\partial_zu_gk_yk_z + N^2k_y^2}{k_y^2 + k_z^2} \quad (11.137a)$$

$$= f(f + \zeta_g)\cos^2\alpha + 2f\partial_zu_g\sin\alpha\cos\alpha + N^2\sin^2\alpha \quad (11.137b)$$

$$= \cos^2\alpha[f(f + \zeta_g) + 2f\partial_zu_g\tan\alpha + N^2\tan^2\alpha]. \quad (11.137c)$$

Note that the convention for symmetric instability is to use the angle,  $\alpha$ , that the particle trajectories make with the vertical, whereas for our study of internal gravity waves (Section 9.5) and inertia-gravity waves (Section 9.8), we used the angle  $\gamma = \pi/2 - \alpha$  that the phase velocity makes with the horizontal. Either way, we see that the dispersion relation (11.137c) depends only on the wave direction along with the prescribed rotation and properties of the background flow. This “orientation-only” character of the dispersion relation is also shared by the dispersion relation for internal gravity waves (equation (9.65)) and inertia-gravity waves (equation (9.145)). Indeed, for the case with zero baroclinicity, in which  $u_g = 0$ , then the dispersion relation (11.137c) reduces to that for inertia-gravity waves.

#### 11.8.5 Stability conditions

The dispersion relation (11.137c) determines the squared angular frequency,  $\omega^2$ , for a plane wave ansatz. If  $\omega^2 > 0$  then there are propagating waves. However, if  $\omega^2 < 0$  then the waves

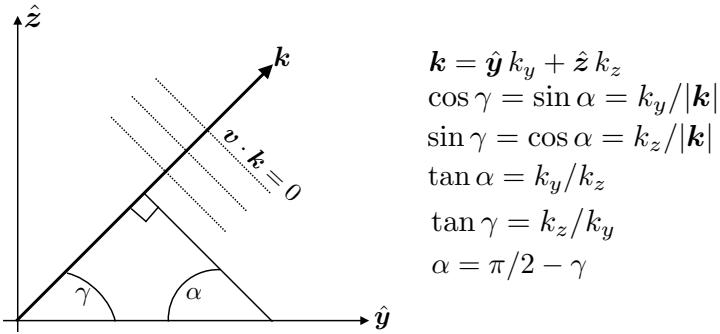


FIGURE 11.9: Transverse plane wave used to probe the stability of a baroclinic front. The phase vector,  $\mathbf{k}$ , makes an angle  $\gamma$  with the horizontal and  $\alpha$  with the vertical.

can be exponentially growing, in which case we say the flow is unstable. To determine the condition for stability, write the dispersion relation in the form

$$\omega^2 = F(\tan \alpha) \cos^2 \alpha, \quad (11.138)$$

where we introduced the quadratic stability function

$$F(\tau) = f(f + \zeta_g) + (2f \partial_z u_g) \tau + N^2 \tau^2. \quad (11.139)$$

The question of stability has been reduced to the question of whether there are wave perturbations that realize  $F(\tan \alpha) < 0$ , in which case  $\omega^2 < 0$ . To answer this question we write

$$F(\tau) = (\tau - R_+) (\tau - R_-), \quad (11.140)$$

with the two roots given by the quadratic formula

$$N^2 R_{\pm} = -f \partial_z u_g \pm \sqrt{(f \partial_z u_g)^2 - N^2 f(f + \zeta_g)} = -f \partial_z u_g \pm \sqrt{-f Q_g}, \quad (11.141)$$

where the final equality introduced the base state potential vorticity,  $Q_g$ , from equation (11.122). Remarkably, local stability has boiled down to the sign of  $f Q_g$  at any point within the fluid.

### Stable case

If  $f Q_g^b > 0$ , then the two roots,  $R_{\pm}$ , are complex conjugates. Yet  $\tan \alpha$  is real for all waves with real wavevectors. Hence, if  $f Q_g > 0$  then there are no waves for which  $F(\tan \alpha)$  vanishes, meaning that the stability function has one sign for all waves. To establish the sign, note that for very weak fronts in stably stratified fluids,

$$F(\tau) \approx f^2 + N^2 \tau^2 > 0, \quad (11.142)$$

which is the case for linear, and stable, inertia-gravity waves. Ramping up the strength of the geostrophic front, while maintaining  $f Q_g > 0$ , ensures that such geostrophic fronts are locally stable.

For the typical case in stratified geophysical fluids, we have  $N^2 > f^2$ , so that the maximum

angular frequency (11.137c) occurs with horizontally propagating waves where  $\alpha = \pi/2$  so that

$$\omega_{\max}^2 = N^2, \quad (11.143)$$

in which case fluid particles exhibit buoyancy oscillations in the vertical. At the other end of the spectrum, the minimum angular frequency occurs when  $\alpha = 0$  so that the wavevector is vertical and

$$\omega_{\min}^2 = f(f + \zeta_g), \quad (11.144)$$

which are inertial oscillations in the presence of a nonzero relative vorticity.

### Unstable case

If  $f Q_g < 0$  then the two roots,  $R_{\pm}$ , are real

$$R_{\pm} = \frac{-f \partial_z u_g \pm \sqrt{|f Q_g|}}{N^2}, \quad (11.145)$$

thus ensuring that there are waves that render  $F(\tan \alpha) = 0$ . Hence, the stability function changes sign so that waves are unstable if  $\omega^2 < 0$ . In the unstable case,  $\omega$  is not an angular frequency but instead it measures the exponential *growth rate* for the unstable wave.

### 11.8.6 Ertel potential vorticity and local stability

In developing the stability conditions in Section 11.8.5, appearance of the Ertel potential vorticity might not have seemed so remarkable. Namely, it seemingly just provided a useful shorthand for the discriminant of the square root appearing the quadratic formula. In fact, potential vorticity is central to the power of the stability analysis since it provides a direct connection to physical processes required to initiate symmetric instability. Namely, potential vorticity is materially invariant in a perfect fluid (Chapters ?? and ??). Hence, if a fluid starts in a stable state with  $f Q > 0$ , say it starts with zero baroclinicity and with stable stratification so that  $f Q = f^2 N^2 > 0$ , then it has very restrictive means to evolve into an unstable state with  $f Q < 0$ . One means for inducing  $f Q < 0$  is to introduce friction, diffusion, boundary processes, and/or other irreversible processes that materially alter  $Q$  in such a manner that brings  $f Q < 0$ . Another means is for a perfect fluid parcel to cross the equator so that  $f$  changes sign. If the parcel starts from a side of the equator with  $f Q > 0$ , then on the other side it has  $f Q < 0$  and so it is symmetrically unstable.

As noted by [Thomas et al. \(2013\)](#), the instability criteria,  $f Q < 0$ , embodies three instabilities. Namely, for a barotropic fluid, whereby  $\nabla_b b = 0$ , then the potential vorticity is given by

$$f Q = f(f + \zeta) N^2. \quad (11.146)$$

The  $f Q > 0$  stability criteria is satisfied with  $N^2 > 0$  (gravitational stability; Section ??) and  $f(f + \zeta) > 0$  (horizontal inertial stability; Section 11.5). In the presence of baroclinicity with a symmetric front, the  $f Q > 0$  stability criteria then reflects symmetric stability of the present section. [Buckingham et al. \(2021\)](#) offer a generalization of the stability criteria by allowing for curvature of the front, thus providing a criteria to determine stability with regard to gravitational, centrifugal, horizontal inertial, and symmetric instabilities.

### 11.8.7 Comments and further study

The bulk of this section is an elaboration of [Hoskins \(1974\)](#), which exposed the connection between symmetric instability and Ertel potential vorticity. The study from [Thomas et al. \(2013\)](#) provides an example of these ideas for the ocean, offering practical details to map out stability regimes across the spectrum of gravitational, inertial, and symmetric instabilities. [Buckingham et al. \(2021\)](#) generalized [Hoskins \(1974\)](#) to the allow for a curved front.

## 11.9 Semi-geostrophy and ageostrophic overturning

We here study the overturning circulation that arises in a symmetric front on an  $f$ -plane. This overturning circulation is ageostrophic, and it is sometimes referred to as the *ageostrophic secondary circulation*. [Hoskins \(1975\)](#) advanced the study of frontal secondary circulations by rationalizing the equations of *semi-geostrophy*, making use of the *geostrophic momentum approximation*. Semi-geostrophy is a balanced theory (i.e., gravity waves are filtered), just like in the quasi-geostrophic theory of Chapter ???. However, the semi-geostrophic equations allow us to study flows associated with fronts and where the cross-front flows can have relatively small length scales sufficient to reach order unity Rossby numbers. Hence, semi-geostrophic equations are quite useful for studying submesoscale flows surrounding geostrophic fronts and eddies (e.g., see the book chapter by [Thomas et al. \(2008\)](#)), as well as the study of atmospheric fronts associated with synoptic scale weather patterns (which motivated [Hoskins \(1975\)](#); see also Chapter 9 of [Holton and Hakim \(2013\)](#)). In this section we provide just a taste of semi-geostrophy sufficient to derive a diagnostic equation for the ageostrophic secondary overturning circulation.

### 11.9.1 Hydrostatic and Boussinesq ocean on an $f$ -plane

We develop the semi-geostrophic equations within the perfect fluid hydrostatic Boussinesq ocean equations (see Section ??) on an  $f$ -plane

$$\frac{D\mathbf{u}}{Dt} + f \hat{\mathbf{z}} \times \mathbf{u} = -\nabla_h \varphi \quad (11.147a)$$

$$\partial_z \varphi = b \quad (11.147b)$$

$$\frac{Db}{Dt} = 0 \quad (11.147c)$$

$$\nabla \cdot \mathbf{v} = \nabla_h \cdot \mathbf{u} + \partial_z w = 0. \quad (11.147d)$$

The geostrophic velocity is written in terms of the pressure field

$$\mathbf{u}_g = f^{-1} \hat{\mathbf{z}} \times \nabla_h \varphi \implies u_g = -f^{-1} \partial_y \varphi \quad \text{and} \quad v_g = f^{-1} \partial_x \varphi, \quad (11.148)$$

which is horizontally non-divergent on the  $f$ -plane. It is notable that this geostrophic velocity is written in terms of the pressure field,  $\varphi$ , just like in the planetary geostrophic theory of Chapter ???. It also accords with the gauge choice discussed in Section ?? for quasi-geostrophy.

Inserting the geostrophic velocity (11.148) into the horizontal momentum equation (11.147a) yields the identity

$$\mathbf{u} - \mathbf{u}_g = \hat{\mathbf{z}} \times \mathcal{D}\mathbf{u} \implies u - u_g = -\mathcal{D}v \quad \text{and} \quad v - v_g = \mathcal{D}u, \quad (11.149)$$

where we introduced the dimensionless material time derivative operator

$$\mathcal{D} = \frac{1}{f} \frac{\mathrm{D}}{\mathrm{D}t}. \quad (11.150)$$

The horizontal ageostrophic flow can thus be written

$$\mathbf{u}_{\text{ag}} = \mathbf{u} - \mathbf{u}_g = \hat{\mathbf{z}} \times \mathcal{D}\mathbf{u}, \quad (11.151)$$

with one iteration of this equation leading to

$$\mathbf{u}_{\text{ag}} = \hat{\mathbf{z}} \times \mathcal{D}\mathbf{u}_g - \mathcal{D}^2\mathbf{u} \implies u_{\text{ag}} = -\mathcal{D}(v_g + \mathcal{D}u) \quad \text{and} \quad v_{\text{ag}} = \mathcal{D}(u_g - \mathcal{D}v). \quad (11.152)$$

This identity motivates the *geostrophic momentum approximation* in Section 11.9.3 that ignores the  $\mathcal{D}^2\mathbf{u}$  term.

## 11.9.2 Scaling for frontal flows

For  $f$ -plane flow, we lose no generality by orienting the front parallel to the  $\hat{\mathbf{x}}$  direction, so that the across-front flow is in the  $\hat{\mathbf{y}}$  direction. We write  $U$  for the velocity scale of the along-front flow and  $L_x$  for the corresponding length scale.<sup>13</sup> Likewise, write  $V$  and  $L_y$  for the across-front scales, with our interest in flows satisfying the inequalities

$$L_x \gg L_y \quad \text{and} \quad U \gg V \quad \text{with} \quad L_x/L_y = U/V. \quad (11.153)$$

Inserting these scales into horizontal momentum equation (11.147a) yields

$$\frac{U}{T} + \frac{UU}{L_x} + \frac{UV}{L_y} - fV = -\Phi/L_x \quad (11.154a)$$

$$\frac{V}{T} + \frac{VU}{L_x} + \frac{VV}{L_y} + fU = -\Phi/L_y, \quad (11.154b)$$

where  $\Phi$  is the pressure scale. Now introduce two Rossby numbers according to

$$\text{Ro}_x = \frac{U}{fL_x} \quad \text{and} \quad \text{Ro}_y = \frac{V}{fL_y}, \quad (11.155)$$

and then divide equation (11.154a) by  $fV$  and equation (11.154b) by  $fU$ , thus yielding

$$\text{Ro}_x \frac{L_x}{TV} + \text{Ro}_x \frac{U}{V} + \text{Ro}_x \frac{L_x}{L_y} - 1 = -\frac{\Phi}{L_x fV} \quad (11.156a)$$

$$\text{Ro}_y \frac{L_y}{TU} + \text{Ro}_y \frac{V}{U} + \text{Ro}_y \frac{L_y}{L_x} + 1 = -\frac{\Phi}{L_y fU}. \quad (11.156b)$$

Now assume the time scale is advective so that

$$T = L_x/U = L_y/V, \quad (11.157)$$

---

<sup>13</sup>The radius of curvature for the front provides a suitable along-front length scale. Recall we introduced the radius of curvature in decomposing horizontal flows in Chapter ???. Also, the mathematics of the radius of curvature was introduced in Chapter ??.

in which case equations (11.156a) and (11.156b) become

$$\text{Ro}_x L_x / L_y + \text{Ro}_x U/V + \text{Ro}_x L_x / L_y - 1 = -\Phi / (L_x f V) \quad (11.158a)$$

$$\text{Ro}_y L_y / L_x + \text{Ro}_y V/U + \text{Ro}_y L_y / L_x + 1 = -\Phi / (L_y f U). \quad (11.158b)$$

Consider the across-front momentum equation (11.158b). Since  $L_y / L_x = V/U \ll 1$ , and assuming  $\text{Ro}_y$  is at most order unity, terms in the across-front momentum equation (11.158b) balance only if the along-front flow is geostrophic

$$f U = -\Phi / L_y. \quad (11.159)$$

In turn, the along-front flow has a small Rossby number,  $\text{Ro}_x \ll 1$ . We emphasize that the along-front flow is geostrophic even if  $\text{Ro}_y \sim 1$ . In contrast, even with  $\text{Ro}_x \ll 1$ , each term in equation (11.158a) can be order unity since  $U/V = L_x / L_y \gg 1$ , so that the across-front flow is not geostrophic. That is,  $\text{Ro}_y \sim 1$ , so that the across-front flow is ageostrophic.

Semi-geostrophy is designed to study flow along and across geostrophic fronts, with the along front flow geostrophic and across-front flow ageostrophic. Equivalently, semi-geostrophy is designed to study flows with relatively large cross-front shear induced relative vorticity (and small curvature induced relative vorticity), with correspondingly large vertical velocity magnitudes, both of which are signatures of  $\text{Ro}_y \sim 1$  flow (see Section 3 of [Hoskins \(1975\)](#)).<sup>14</sup>

### 11.9.3 Geostrophic momentum approximation

The *geostrophic momentum approximation* assumes the horizontal ageostrophic velocity is given by

$$\mathbf{u}_{\text{ag}} \approx \hat{\mathbf{z}} \times \mathcal{D}\mathbf{u}_{\text{g}}, \quad (11.160)$$

which results from truncating equation (11.152) at the leading order. From equation (11.152) we see that the geostrophic momentum approximation holds so long as the horizontal velocity components satisfy the following respective inequalities

$$|u| \gg |\mathcal{D}^2 u| \implies f^2 \gg \left| \frac{1}{u} \frac{\text{D}^2 u}{\text{D}t^2} \right| \quad (11.161a)$$

$$|v| \gg |\mathcal{D}^2 v| \implies f^2 \gg \left| \frac{1}{v} \frac{\text{D}^2 v}{\text{D}t^2} \right|. \quad (11.161b)$$

These inequalities are satisfied if the Lagrangian timescale of variability for the flow (both its magnitude and direction; see Section 3 of [Hoskins \(1975\)](#)) is much longer than an inertial period. Rearranging the geostrophic momentum approximation (11.160) leads to

$$\mathcal{D}\mathbf{u}_{\text{g}} + \hat{\mathbf{z}} \times \mathbf{u}_{\text{ag}} = 0, \quad (11.162)$$

which is equivalent to

$$\frac{\text{D}\mathbf{u}_{\text{g}}}{\text{D}t} + f \hat{\mathbf{z}} \times \mathbf{u}_{\text{ag}} = 0, \quad (11.163)$$

Hence, for the geostrophic momentum approximation, the material time evolution of the geostrophic velocity is forced by the Coriolis acceleration arising from the horizontal ageostrophic

---

<sup>14</sup>In Section ?? we decomposed relative vorticity into from cross-flow shear and along-flow curvature.

velocity. Furthermore, the material time derivative for the semi-geostrophic system is given by

$$\frac{D}{Dt} = \partial_t + \mathbf{v} \cdot \nabla = \partial_t + (\mathbf{u}_g + \mathbf{u}_{ag}) \cdot \nabla_h + w_{ag} \partial_z, \quad (11.164)$$

with

$$\mathbf{v} = (\mathbf{u}_g + \mathbf{u}_{ag}) + \hat{\mathbf{z}} w_{ag}. \quad (11.165)$$

Note that any vertical flow is ageostrophic, so that it is not really necessary to place the “ag” subscript on  $w_{ag}$ . Even so, we do so to remind us that it lives at the same order as the horizontal ageostrophic flow,  $\mathbf{u}_{ag}$ . Furthermore, note that the key distinction between semi-geostrophy and quasi-geostrophy is the presence of  $w_{ag}$  in the material time derivative operator for semi-geostrophy, whereas quasi-geostrophy only makes use of advection by the geostrophic flow (see Section ??).

The geostrophic momentum approximation is consistent with the frontal scaling given in Section 11.9.2, and so it offers a suitable starting point for the study of frontal dynamics. Even so, to make use of the geostrophic momentum approximation to evolve the geostrophic flow requires further work since the ageostrophic velocity appears as part of the material time derivative operator. This situation is analogous to quasi-geostrophy, whereby evolution of the geostrophic flow arises from ageostrophic processes. To derive a closed equation set, [Hoskins \(1975\)](#) developed the semi-geostrophic momentum equations via a coordinate transformation to *geostrophic coordinates*, making use also of material constancy of the buoyancy and potential vorticity.<sup>15</sup> We do not present that formulation here. Instead, we focus on deriving a diagnostic equation for the secondary ageostrophic overturning circulation, with the manipulations reminiscent of those used in Section 11.8 for symmetric instability of a baroclinic front.

#### 11.9.4 Geostrophic evolution of the buoyancy gradient

As a preface to the diagnostic equations for the ageostrophic flow, consider the evolution of the horizontal buoyancy gradient, considering just that evolution arising from geostrophic flow. Recall that for the Boussinesq ocean, the horizontal gradient of buoyancy provides a measure of baroclinicity (Section ??). Flow processes that increase the magnitude of the buoyancy gradient lead to growth of the thermal wind flow and thus the frontal strength.

Start by considering the buoyancy equation for adiabatic and geostrophic flow

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) b = 0. \quad (11.166)$$

Separately taking the zonal and meridional derivatives and rearranging leads to

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h)(\partial_x b) = -\partial_x \mathbf{u}_g \cdot \nabla_h b \quad \text{and} \quad (\partial_t + \mathbf{u}_g \cdot \nabla_h)(\partial_y b) = -\partial_y \mathbf{u}_g \cdot \nabla_h b, \quad (11.167)$$

which can be combined to

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h)(\nabla_h b) = \mathbf{Q}, \quad (11.168)$$

where we introduced the horizontal vector

$$\mathbf{Q} = -(\partial_x \mathbf{u}_g \cdot \nabla_h b) \hat{\mathbf{x}} - (\partial_y \mathbf{u}_g \cdot \nabla_h b) \hat{\mathbf{y}}. \quad (11.169)$$

---

<sup>15</sup>Some authors reserve the name *semi-geostrophy* for the transformed equations derived by [Hoskins \(1975\)](#).

Correspondingly, geostrophic evolution of the squared buoyancy gradient is

$$(\partial_t + \mathbf{u}_g \cdot \nabla_h) |\nabla_h b|^2 = 2 \mathbf{Q} \cdot \nabla_h b. \quad (11.170)$$

Evidently, the buoyancy gradient grows in magnitude if  $\mathbf{Q}$  has a positive projection onto the horizontal buoyancy gradient.

### 11.9.5 Secondary ageostrophic circulation

In Sections 11.5 and 11.6 we focused on the stability of a geostrophically balanced equilibrium with flow along a symmetric front. In addition to the geostrophic flow along the front, there is generally an ageostrophic circulation that circulates in the plane orthogonal to the front. We here derive an equation describing this overturning circulation, and then specialize that equation in Section 11.9.6 to a zonally symmetric front. For this purpose, start from the zonal momentum equation, buoyancy equation, and continuity equation within the semi-geostrophic system

$$\partial_t u_g + (\mathbf{u}_g \cdot \nabla_h) u_g + (\mathbf{u}_{ag} \cdot \nabla_h) u_g + w_{ag} \partial_z u_g - f v_{ag} = 0 \quad (11.171a)$$

$$\partial_t b + \mathbf{u}_g \cdot \nabla_h b + \mathbf{u}_{ag} \cdot \nabla_h b + w_{ag} N^2(z) = 0 \quad (11.171b)$$

$$\partial_x u_{ag} + \partial_y v_{ag} + \partial_z w_{ag} = 0. \quad (11.171c)$$

The vertical derivative of the zonal momentum equation (11.171a) leads to

$$\begin{aligned} \partial_t \partial_z u_g + (\partial_z \mathbf{u}_g \cdot \nabla_h) u_g + (\mathbf{u}_g \cdot \nabla_h) \partial_z u_g + (\partial_z \mathbf{u}_{ag} \cdot \nabla_h) u_g + (\mathbf{u}_{ag} \cdot \nabla_h) \partial_z u_g \\ + \partial_z (w_{ag} \partial_z u_g) - f \partial_z v_{ag} = 0, \end{aligned} \quad (11.172)$$

and the meridional derivative of the buoyancy equation (11.171b) yields

$$\partial_t \partial_y b + (\partial_y \mathbf{u}_g \cdot \nabla_h) b + (\mathbf{u}_g \cdot \nabla_h) \partial_y b + (\partial_y \mathbf{u}_{ag} \cdot \nabla_h) b + (\mathbf{u}_{ag} \cdot \nabla_h) \partial_y b + \partial_y w_{ag} N^2 = 0. \quad (11.173)$$

The thermal wind relation,

$$f \partial_z \mathbf{u}_g = \hat{\mathbf{z}} \times \nabla_h b \implies \partial_z u_g = -f^{-1} \partial_y b \quad \text{and} \quad \partial_z v_g = f^{-1} \partial_x b, \quad (11.174)$$

brings equation (11.172) to the form

$$\begin{aligned} \partial_t (\partial_y b) - f (\partial_z \mathbf{u}_g \cdot \nabla_h) u_g + (\mathbf{u}_g \cdot \nabla_h) \partial_y b - f (\partial_z \mathbf{u}_{ag} \cdot \nabla_h) u_g + (\mathbf{u}_{ag} \cdot \nabla_h) \partial_y b \\ + \partial_z (w_{ag} \partial_y b) + f^2 \partial_z v_{ag} = 0, \end{aligned} \quad (11.175)$$

and equation (11.173) becomes

$$\partial_t (\partial_y b) + (\partial_y \mathbf{u}_g \cdot \nabla_h) b + (\mathbf{u}_g \cdot \nabla_h) \partial_y b + (\partial_y \mathbf{u}_{ag} \cdot \nabla_h) b + (\mathbf{u}_{ag} \cdot \nabla_h) \partial_y b + \partial_y w_{ag} N^2 = 0. \quad (11.176)$$

Subtracting equation (11.176) from equation (11.175) eliminates the time derivative thus revealing the diagnostic relation

$$\begin{aligned} -f (\partial_z \mathbf{u}_g \cdot \nabla_h) u_g - f (\partial_z \mathbf{u}_{ag} \cdot \nabla_h) u_g + \partial_z (w_{ag} \partial_y b) + f^2 \partial_z v_{ag} - (\partial_y \mathbf{u}_g \cdot \nabla_h) b \\ - (\partial_y \mathbf{u}_{ag} \cdot \nabla_h) b - \partial_y w_{ag} N^2 = 0. \end{aligned} \quad (11.177)$$

Making use of thermal wind and horizontal non-divergence for the geostrophic velocity allows us to write

$$f(\partial_z \mathbf{u}_g \cdot \nabla_h) u_g = (\partial_y \mathbf{u}_g \cdot \nabla_h) b = -Q^y, \quad (11.178)$$

where  $Q^y$  is the meridional component of the geostrophic  $\mathbf{Q}$ -vector introduced in Section ?? (see equation (??)).<sup>16</sup>

$$\mathbf{Q} \equiv -(\partial_x \mathbf{u}_g \cdot \nabla_h) \hat{\mathbf{x}} - (\partial_y \mathbf{u}_g \cdot \nabla_h) \hat{\mathbf{y}}. \quad (11.179)$$

Introduction of  $Q^y$  into equation (11.177) yields

$$-f(\partial_z \mathbf{u}_{ag} \cdot \nabla_h) u_g + \partial_z w_{ag} \partial_y b + w_{ag} \partial_y z b + f^2 \partial_z v_{ag} - (\partial_y \mathbf{u}_{ag} \cdot \nabla_h) b - \partial_y w_{ag} N^2 = -2Q^y. \quad (11.180)$$

Again making use of thermal wind and  $\nabla_h \cdot \mathbf{u}_g = 0$  renders

$$-f(\partial_z \mathbf{u}_{ag} \cdot \nabla_h) u_g - (\partial_y \mathbf{u}_{ag} \cdot \nabla_h) b = f \partial_z u_{ag} \partial_y v_g - f \partial_z v_{ag} \partial_y u_g + f \partial_y u_{ag} \partial_z v_g + f \partial_y v_{ag} \partial_z u_g. \quad (11.181)$$

The mixed partial derivative of the buoyancy vanishes

$$\frac{\partial^2 b}{\partial z \partial y} = \frac{\partial}{\partial y} \frac{\partial b}{\partial z} = \frac{\partial N^2(z)}{\partial y} = 0, \quad (11.182)$$

which follows since the background vertical stratification is assumed to be independent of the horizontal direction. Bringing these results together into equation (11.180) leads to

$$\begin{aligned} & f \partial_z u_{ag} \partial_y v_g + f \partial_y u_{ag} \partial_z v_g + f \partial_y v_{ag} \partial_z u_g + \partial_z w_{ag} \partial_y b + f \partial_z v_{ag} (f - \partial_y u_g) \\ & \quad - \partial_y w_{ag} N^2 = -2Q^y, \end{aligned} \quad (11.183)$$

with another use of thermal wind yielding

$$f \partial_z u_{ag} \partial_y v_g + f \partial_y u_{ag} \partial_z v_g + f \partial_z u_g (\partial_y v_{ag} - \partial_z w_{ag}) + f \partial_z v_{ag} (f - \partial_y u_g) - \partial_y w_{ag} N^2 = -2Q^y. \quad (11.184)$$

This equation provides a relation for the ageostrophic cross-flow and vertical circulation,  $(v_{ag}, w_{ag})$ , written in terms of the buoyancy field and the geostrophic flow. We next consider flow surrounding a symmetric front, in which case equation (11.184) becomes a diagnostic equation for the ageostrophic overturning streamfunction.

### 11.9.6 Ageostrophic overturning circulation for a symmetric front

Equation ((11.184) is now specialized by assuming the zonal velocity is purely geostrophic (as in a zonal geostrophic front) so that

$$u_{ag} = 0. \quad (11.185)$$

For this flow, the ageostrophic flow in the cross-flow/depth plane is non-divergent

$$\partial_y v_{ag} + \partial_z w_{ag} = 0. \quad (11.186)$$

The diagnostic equation (11.184) now takes on the specialized form for a symmetric front

$$-2f \partial_z u_g \partial_z w_{ag} + f(f + \zeta_g) \partial_z v_{ag} - \partial_y w_{ag} N^2 = -2Q^y, \quad (11.187)$$

---

<sup>16</sup>Be careful to distinguish the geostrophic  $\mathbf{Q}$ -vector from the potential vorticity,  $Q$ . In particular, note the upright  $\mathbf{Q}$  versus the slanted  $Q$ .

where

$$\zeta_g = -\partial_y u_g \quad (11.188)$$

is the vertical component of the geostrophic relative vorticity. Introducing an overturning streamfunction for the cross-flow/vertical ageostrophic circulation

$$\mathbf{u}_{ag} = -\hat{\mathbf{x}} \times \nabla \psi \implies v_{ag} = \partial_z \psi \quad \text{and} \quad w_{ag} = -\partial_y \psi, \quad (11.189)$$

and using thermal wind brings equation (11.187) into the form

$$[N^2 \partial_{yy} - 2 \partial_y b \partial_{yz} + f(f + \zeta_g) \partial_{zz}] \psi = -2Q^y. \quad (11.190)$$

Equation (11.190) is useful for the study of ageostrophic ( $\text{Ro} \sim 1$ ) dynamics along a front in which there is an ageostrophic overturning circulation in response to geostrophic forcing from  $Q^y$ .

### 11.9.7 Connection to potential vorticity and symmetric instability

The partial differential equation (11.190) can be written

$$\mathcal{K}\psi = -2Q^y, \quad (11.191)$$

where

$$\mathcal{K} = N^2 \partial_{yy} - 2 \partial_y b \partial_{zy} + f(f + \zeta_g) \partial_{zz} \quad (11.192)$$

is a linear partial differential operator that is a function of the geostrophic flow and the buoyancy. Following the considerations in Section ??, we know that this operator is elliptic if the following inequality holds

$$(\partial_y b)^2 - N^2 f(f + \zeta_g) < 0. \quad (11.193)$$

We can relate the ellipticity condition (11.193) to the Ertel potential vorticity for the Boussinesq geostrophic flow. For this purpose, write the geostrophic vorticity as

$$\boldsymbol{\omega}_g = \nabla \times \mathbf{u}_g \quad (11.194a)$$

$$= -\hat{\mathbf{x}} \partial_z v_g + \hat{\mathbf{y}} \partial_z u_g + \hat{\mathbf{z}} \partial_x v_g - \partial_y u_g \quad (11.194b)$$

$$= -f^{-1} (\hat{\mathbf{x}} \partial_x b + \hat{\mathbf{y}} \partial_y b) + \hat{\mathbf{z}} (\partial_x v_g - \partial_y u_g). \quad (11.194c)$$

If we assume the front is zonally symmetric, then the relative vorticity in the geostrophic flow takes the form

$$\boldsymbol{\omega}_g^{2d} = -f^{-1} \partial_y b \hat{\mathbf{y}} - \hat{\mathbf{z}} \partial_y u_g, \quad (11.195)$$

in which case the Ertel potential vorticity (for the Boussinesq geostrophic flow) takes the form

$$Q_g^{2d} = \nabla b \cdot (\boldsymbol{\omega}_g + f \hat{\mathbf{z}}) \quad (11.196a)$$

$$= -f^{-1} (\partial_y b)^2 + N^2 (f - \partial_y u_g). \quad (11.196b)$$

$$= -f^{-1} (\partial_y b)^2 + N^2 (f + \zeta_g). \quad (11.196c)$$

Ellipticity of the PDE (11.190) is thus assured so long as

$$f Q_g^{2d} = -(\partial_y b)^2 + N^2 f(f + \zeta_g) > 0. \quad (11.197)$$

The PDE (11.6) transitions to a hyperbolic system when  $f Q_g^{2d} < 0$ , which is the condition for symmetric instability detailed in Section 11.6. Hence, when  $f Q_g^{2d} > 0$  the front is stable to symmetric instability, with the ageostrophic circulation acting to relax the front. In contrast, when  $f Q_g^{2d} > 0$  the front is symmetrically unstable.

### 11.9.8 Further study

Chapter 9 of *Holton and Hakim* (2013) provides a pedagogical discussion of semi-geostrophy and the study of atmospheric fronts.



## 11.10 Exercises

### EXERCISE 11.1: INERTIAL INSTABILITY IN A SHALLOW WATER LAYER

In Section 11.5 we developed the physics of horizontal inertial instability within the continuous Boussinesq equations. Consider the analysis instead within the context of a single layer of shallow water fluid. What are the conditions for inertial instability within the shallow water layer? Are they the same as for the continuous Boussinesq equations? Why?

### EXERCISE 11.2: SYMMETRIC INSTABILITY CRITERIA IN TERMS OF BALANCED RICHARDSON NUMBER

Consider the balanced Richardson number, defined for the geostrophic and hydrostatic balanced flow, along with a corresponding angle

$$Ri_b = \frac{N^2}{(\partial_z u_g)^2} \quad \text{and} \quad \phi_{Ri_b} \equiv \tan^{-1}(-1/Ri_b). \quad (11.198)$$

Show that the instability criteria,  $f Q < 0$ , from Section 11.8 can be written in the equivalent form

$$\phi_{Ri_b} < \phi_c \equiv \tan^{-1}(-(f + \zeta_g)/f). \quad (11.199)$$

As shown by *Thomas et al.* (2013), this alternative form of the instability criteria allows for a very effective means to characterize flow regimes conducive to the suite of local instabilities: gravitational, inertial, and symmetric.

### EXERCISE 11.3: GROUP VELOCITY FOR SYMMETRIC MERIDIONAL-VERTICAL WAVES

In this exercise we derive some properties for the group velocity of the stable meridional-vertical plane waves discussed in Section 11.8.4.

- (a) Derive the group velocity,  $\mathbf{c}_g$ , for the stable meridional-vertical waves whose dispersion relation is given by equation (11.137c).
- (b) Compute  $\mathbf{c}_g \cdot \mathbf{k}$ , where the wavevector is  $\mathbf{k} = k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}$ .





## Chapter 12

### STABILITY OF FLUID INTERFACES

In this chapter we study the *Rayleigh-Taylor* and *Kelvin-Helmholtz* instabilities as realized along an infinitesimal material fluid interface that separates two homogeneous (constant density) fluid layers of differing densities, and with the two fluids immiscible and inviscid. The gravitational body force is active throughout the fluid layers, and the surface tension is active at the interface between the layers. We do not consider Coriolis or centrifugal acceleration in this chapter (non-rotating reference frame), and the relative vorticity vanishes in both layers. The *normal mode method*, based on Fourier modal analysis, is suited to developing necessary and sufficient conditions for instability.

The physical system is highly idealized in its assumption that the interface between the fluid layers is infinitesimal, and furthermore that the fluids are inviscid and immiscible. Both assumptions accord with our treatment of surface gravity and capillary waves in Chapter 4. Even though quite idealized, the methods and concepts encountered in studying the instabilities are useful in more realistic cases. Furthermore, experimental results support the relevance of the instability conditions found in this chapter.

#### READER'S GUIDE FOR THIS CHAPTER

This chapter is an extension of the material in Chapter 4, which is concerned with stable linear surface waves on an interface arising from gravity and/or surface tension. The methods of modal instability analysis used in this chapter are also used for a variety of other instabilities elsewhere in this book.

<b>12.1</b>	<b>Governing equations</b>	<b>372</b>
12.1.1	Equations from potential theory	372
12.1.2	Kinematic boundary condition at the interface	373
12.1.3	Dynamic boundary condition at the interface	374
12.1.4	Pressures within the two layers	374
12.1.5	Linearized equations	375
<b>12.2</b>	<b>Rayleigh-Taylor instability</b>	<b>376</b>
12.2.1	Boundary value problem and dispersion relation	376
12.2.2	Stable traveling plane waves	377
12.2.3	Unstable exponentially growing plane waves	378
12.2.4	Further study	380
<b>12.3</b>	<b>Kelvin-Helmholtz instability</b>	<b>380</b>
12.3.1	Velocity potential	380
12.3.2	Dispersion relation from the interface conditions	380
12.3.3	Analysis of the stability condition	381

---

12.3.4 Insights from vorticity . . . . .	383
12.3.5 Insights from Bernoulli's theorem . . . . .	383
12.3.6 An energetic perspective on Kelvin-Helmholtz induced mixing . .	384
12.3.7 Further study . . . . .	387

---

## 12.1 Governing equations

Formulation of the stability problem in this chapter closely follows that pursued in Chapter 4 for surface gravity waves and capillary waves. In that study we assumed the waves occur on the interface between two regions of constant density, yet with the upper region having zero density, as motivated by waves on the free surface of the ocean underneath the much less massive atmosphere. Here, we allow for the upper region to have a nonzero density. Under certain circumstances, fluctuations on the interface lead to stable linear waves, and we develop the associated dispersion relations. However, a Rayleigh-Taylor instability occurs when the smaller density layer sits below a heavier layer, and with a base state that is at rest. More precisely, an instability occurs when the destabilizing effects from gravity overcome the stabilizing effects from surface tension. In the absence of surface tension, there is always a Rayleigh-Taylor instability when heavy fluid sits above light fluid.

For the Kelvin-Helmholtz instability, we consider the heavy fluid below the light fluid, yet with a base state velocity that differs between the two layers so that there is a velocity shear (more precisely, a velocity jump) at the interface and an associated vorticity singularity. An instability occurs when the destabilizing effects from the velocity shear are sufficiently strong so as to overcome the stabilizing effects from gravity and surface tension. In the absence of both gravity and surface tension, any velocity shear renders an instability of the interface which, as noted on page 485 of [Chandrasekhar \(1961\)](#), leads one to conclude that *the Kelvin-Helmholtz instability arises by the crinkling of the interface by the shear that is present; and this crinkling occurs even for the smallest differences in the velocities of the two fluids*. We find in Chapter 13 a modification to this result when considering a finite shear layer rather than the infinitesimal interface considered in the current chapter.

In this section we develop the equations describing motion of an interface separating two homogeneous, inviscid, and immiscible fluids in a non-rotating reference frame and with densities,  $\rho_1$  and  $\rho_2$ , and constant zonal velocities,  $\hat{\mathbf{x}} U_1$  and  $\hat{\mathbf{x}} U_2$ . (Figure 12.1). We ignore side boundaries by assuming the domain to be horizontally infinite. This is the idealized physical system that is used to study the Rayleigh-Taylor instability and the Kelvin-Helmholtz instability.

### 12.1.1 Equations from potential theory

The fluid in each layer has a constant density and so is incompressible, so that the velocity in each layer is non-divergent,  $\nabla \cdot \mathbf{v} = 0$ . Furthermore, we assume the flow is irrotational ( $\nabla \times \mathbf{v} = 0$ ) and so can make use of the potential theory detailed in Section 4.2. In particular, the velocity potential has a time tendency given by Bernoulli's equation of motion (4.16)

$$\partial_t \Psi = \Phi + \mathcal{K} + p/\rho = \mathcal{K} + p_d/\rho, \quad (12.1)$$

where

$$\mathbf{v} = -\nabla\Psi \quad \text{velocity and velocity potential} \quad (12.2a)$$

$$\Phi = g z \quad \text{geopotential} \quad (12.2b)$$

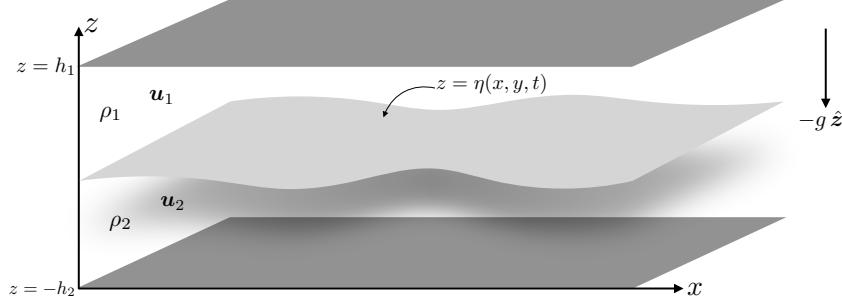


FIGURE 12.1: An infinite box filled with two homogeneous and immiscible fluids with densities,  $\rho_1$  and  $\rho_2$ , and horizontal velocities,  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . The material interface between the layers is located at  $z = \eta(x, y, t)$ , with  $\eta = 0$  when the interface is flat. The fluid region is denoted  $\mathcal{R}$ , which is infinite in the horizontal directions. For study of the Rayleigh-Taylor instability we assume rigid and flat plates located at  $z = -h_2$  and  $z = h_1$ , whereas for Kelvin-Helmholtz instability we let  $h_1$  and  $h_2$  go to infinity. The interface generally experiences a surface tension due to the density jump.

$$\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2 \quad \text{kinetic energy per mass} \quad (12.2c)$$

$$p = -\rho \Phi + p_d \quad p = \text{pressure and } p_d = \text{dynamic pressure.} \quad (12.2d)$$

The velocity potential,  $\Psi$ , satisfies Laplace's equation in the fluid interior and the no-normal flow kinematic boundary condition (Neumann boundary condition) at the rigid top and rigid bottom

$$\nabla^2 \Psi = 0 \quad \mathbf{x} \in \mathcal{R} \quad (12.3a)$$

$$\hat{\mathbf{n}} \cdot \nabla \Psi = 0 \quad \mathbf{x} \in \partial \mathcal{R}. \quad (12.3b)$$

When considering a background zonal velocity, we extract the velocity potential for the static background potential,  $\Psi_n^\circ = -U_n x$ , with  $n = 1, 2$  the layer index and with the remaining portion of the velocity potential capturing the perturbation relative to the background. The background velocity potential trivially satisfies Laplace's equation and the no-normal flow boundary condition at the flat and rigid top and bottom boundaries.

### 12.1.2 Kinematic boundary condition at the interface

Motion of the interface,  $z = \eta(x, y, t)$ , is determined by boundary conditions evaluated at the interface. The interface is material since the fluid layers are assumed to be immiscible. Focusing first on the lower side, in region 2, the kinematic boundary condition from Section ?? states that

$$(\mathbf{v}_2 - \mathbf{v}_\eta) \cdot \hat{\mathbf{n}}_2 = 0 \quad \text{at } z = \eta, \quad (12.4)$$

where  $\mathbf{v}_2$  is the fluid velocity in the lower region, and

$$\hat{\mathbf{n}}_2 = \hat{\mathbf{n}} = \frac{\nabla(z - \eta)}{|\nabla(z - \eta)|} \quad \text{at } z = \eta, \quad (12.5)$$

is the outward normal direction pointing into the upper layer. The analysis for layer-2 holds equivalently for layer-1, only with the normal direction pointing from layer-1 down to layer-2. We thus have

$$(\mathbf{v}_2 - \mathbf{v}_\eta) \cdot \hat{\mathbf{n}} = 0 \quad \text{and} \quad (\mathbf{v}_\eta - \mathbf{v}_1) \cdot \hat{\mathbf{n}} = 0 \implies (\mathbf{v}_2 - \mathbf{v}_1) \cdot \hat{\mathbf{n}} = 0. \quad (12.6)$$

That is, the normal components to the layer velocities match at the interface. Consequently, the normal derivative of the velocity potential also matches at the interface

$$\hat{\mathbf{n}} \cdot (\nabla \Psi_2 - \nabla \Psi_1) = 0 \quad \text{at } z = \eta. \quad (12.7)$$

The velocity of the interface,  $\mathbf{v}_\eta$ , has a normal component given by

$$|\nabla(z - \eta)| \mathbf{v}_\eta \cdot \hat{\mathbf{n}} = \partial_t \eta \quad \text{at } z = \eta, \quad (12.8)$$

so that the kinematic boundary condition (12.4) can be written in the equivalent form

$$(\partial_t + \mathbf{u}_1 \cdot \nabla) \eta = w_1 \quad \text{and} \quad (\partial_t + \mathbf{u}_2 \cdot \nabla) \eta = w_2 \quad \text{at } z = \eta. \quad (12.9)$$

Inserting the background velocity and the velocity potential leads to the kinematic boundary conditions at the fluid interface

$$(\partial_t + U_1 \partial_x - \nabla \Psi_1 \cdot \nabla) \eta = -\partial_z \Psi_1 \quad \text{and} \quad (\partial_t + U_2 \partial_x - \nabla \Psi_2 \cdot \nabla) \eta = -\partial_z \Psi_2 \quad \text{at } z = \eta. \quad (12.10)$$

### 12.1.3 Dynamic boundary condition at the interface

The dynamic boundary condition from Section ?? says that the pressure is continuous across the interface so long as we ignore surface tension. For the case with surface tension, the discussion in Section ?? reveals a pressure jump across the interface, which here takes the form given by the Young-Laplace formula (??),<sup>1</sup>

$$p_2 - p_1 = -\gamma \nabla_h^2 \eta \quad \text{at } z = \eta. \quad (12.11)$$

In this equation,  $\gamma > 0$  is the surface tension (dimensions of force per length = M T<sup>-2</sup>). Evidently, pressure on the concave side of the interface is higher than on the convex side. For example, if the interface extends upward then  $p_2 - p_1 > 0$  since the layer-2 fluid is on the concave side and so it has the higher interface pressure. This result also follows since  $\nabla_h^2 \eta < 0$  for an upward extension of the interface, which leads to a local free surface maximum. The treatment here follows our approach for surface capillary-gravity waves in Section 4.10.1.

### 12.1.4 Pressures within the two layers

We find it convenient to isolate the hydrostatic pressure within the two layers. For the upper layer we have

$$p_1^h = g \rho_1 (h_1 - z) \quad \text{for } \eta \leq z \leq h_1, \quad (12.12)$$

where we assumed pressure at  $z = h_1$  is zero. Similarly, the hydrostatic pressure in layer two is

$$p_2^h = p_1^h(\eta) + g \rho_2 (\eta - z) = g \rho_1 (h_1 - \eta) + g \rho_2 (\eta - z) \quad \text{for } -h_2 \leq z \leq 0. \quad (12.13)$$

Note that these hydrostatic pressures match at the interface

$$p_2^h(x, y, z = \eta, t) = p_1^h(x, y, z = \eta, t). \quad (12.14)$$

---

<sup>1</sup>We write  $\nabla_h^2 \eta$  to emphasize that the Laplacian is only acting in the horizontal directions. This notation is not needed when the Laplacian acts on the interface,  $\eta$ , since this field is just a spatial function of the horizontal positions,  $x, y$ . However, in equation (12.24) we replace  $\partial_t \eta(x, y, t)$  with  $-\partial_z \Psi(x, y, z = 0, t)$  as per the linearized kinematic boundary condition (12.20c) with  $u_1 = u_2 = 0$ . It is this replacement that makes it important to note that the Laplacian is acting just in the horizontal.

If there is fluid motion within the layers, then the layer interface is not flat and the pressure is not equal to its resting pressure. In this case we write the pressure as

$$p_1(x, y, z, t) = p_1^h(x, y, z, t) + \delta p_1(x, y, z, t) \quad (12.15a)$$

$$p_2(x, y, z, t) = p_2^h(x, y, z, t) + \delta p_2(x, y, z, t) - \gamma \nabla_h^2 \eta(x, y, t), \quad (12.15b)$$

where  $\delta p_1$  and  $\delta p_2$  are pressures associated with the fluid motion and that vanish when the fluid is at rest. Furthermore, these pressures are continuous at the interface just like the hydrostatic pressures

$$\delta p_1 = \delta p_2 \quad \text{at } z = \eta, \quad (12.16)$$

so that there remains a pressure jump (12.11) in the presence of surface tension

$$p_2 - p_1 = -\gamma \nabla_h^2 \eta \quad \text{at } z = \eta. \quad (12.17)$$

### 12.1.5 Linearized equations

As for the study of surface gravity waves and capillary waves in Chapter 4, linearization of this system is based on an assumed small slope for the interfaces. Hence, we follow the study of surface waves in Section 4.3 to linearize the Bernoulli equation of motion (12.1). Correspondingly, we linearize the kinematic boundary condition (12.9) and dynamic boundary condition (12.17). In particular, the linearized boundary conditions are evaluated at  $z = 0$  rather than at  $z = \eta$ , in which case the linearized version of  $p_2^h$  is a function just of  $z$  alone.

A new feature here beyond the case of surface waves concerns the presence of a background flow, in which case the kinetic energy contributes at linear order. In particular, the layer-1 Bernoulli equation yields

$$\rho_1 \partial_t \Psi_1 = g z \rho_1 + p_1 + \rho_1 [U_1 u_1 + (U_1)^2/2]. \quad (12.18)$$

The constant,  $\rho_1 (U_1)^2/2$ , can be eliminated by taking a gauge transformation, as detailed in Section 4.2.3 when studying surface waves. Hence, we drop this term in the following. Writing  $u_1 = -\partial_x \Psi_1$  then brings the linearized Bernoulli equation to the form

$$\rho_1 (\partial_t + U_1 \partial_x) \Psi_1 = g z \rho_1 + p_1, \quad (12.19)$$

so that the background zonal flow provides a constant advection of the velocity potential. We are thus led to the following linearized layer equations

$$\rho_1 (\partial_t + U_1 \partial_x) \Psi_1 = g z \rho_1 + p_1 \quad \text{linearized Bernoulli equation for layer 1} \quad (12.20a)$$

$$\rho_2 (\partial_t + U_2 \partial_x) \Psi_2 = g z \rho_2 + p_2 \quad \text{linearized Bernoulli equation for layer 2} \quad (12.20b)$$

$$(\partial_t + U_1 \partial_x) \eta = -\partial_z \Psi_1 \quad \text{linearized kinematic b.c. at } z = 0 \quad (12.20c)$$

$$(\partial_t + U_2 \partial_x) \eta = -\partial_z \Psi_2 \quad \text{linearized kinematic b.c. at } z = 0 \quad (12.20d)$$

$$\delta p_2 - \delta p_1 = 0 \quad \text{dynamic b.c. at } z = 0 \quad (12.20e)$$

$$p_2 - p_1 = -\gamma \nabla_h^2 \eta \quad \text{dynamic jump b.c. at } z = 0. \quad (12.20f)$$

Taking the difference between the Bernoulli equations in the two layers gives

$$\rho_2 (\partial_t + U_2 \partial_x) \Psi_2 - \rho_1 (\partial_t + U_1 \partial_x) \Psi_1 = g (\rho_2 z_2 - \rho_1 z_1) + p_2 - p_1, \quad (12.21)$$

and then evaluating this difference on the interface ( $z_1 = z_2 = \eta$ ) leads to

$$\rho_2 (\partial_t + U_2 \partial_x) \Psi_2 - \rho_1 (\partial_t + U_1 \partial_x) \Psi_1 = (g \delta\rho - \gamma \nabla_h^2) \eta \quad (12.22)$$

where the density difference is written

$$\delta\rho = \rho_2 - \rho_1. \quad (12.23)$$

Note that when multiplied by gravity, we evaluate the interface position at  $z = \eta$ , whereas other terms in the linear theory are evaluated at  $z = 0$ .<sup>2</sup>

## 12.2 Rayleigh-Taylor instability

For the Rayleigh-Taylor instability analysis, we assume the background flow is at rest so that  $U_1 = U_2 = 0$ . Hence, we are here examining stability of the rest state to small perturbations of the layer interface.

### 12.2.1 Boundary value problem and dispersion relation

With zero background flow it is a simple matter to eliminate the free surface from the interface condition (12.22). We do so by taking a time derivative and using the linearized kinematic boundary condition (12.20c) (or equivalently equation (12.20d))

$$\partial_{tt}(\rho_2 \Psi_2 - \rho_1 \Psi_1) = -[g \delta\rho - \gamma \nabla_h^2] \partial_z \Psi \quad \text{at } z = 0, \quad (12.24)$$

where  $w_1 = w_2 = -\partial_z \Psi$  at  $z = 0$ . We are thus led to the boundary value problem

$$\partial_z \Psi_1 = 0 \quad z = h_1 \quad (12.25a)$$

$$\nabla^2 \Psi_1 = 0 \quad 0 < z < h_1 \quad (12.25b)$$

$$\partial_{tt}(\rho_2 \Psi_2 - \rho_1 \Psi_1) = -[g \delta\rho - \gamma \nabla_h^2] \partial_z \Psi \quad z = 0 \quad (12.25c)$$

$$\nabla^2 \Psi_2 = 0 \quad -h_2 < z < 0 \quad (12.25d)$$

$$\partial_z \Psi_2 = 0 \quad z = -h_2. \quad (12.25e)$$

Following the approach for surface waves in Section 4.5, we seek a traveling plane wave solution with horizontal wavevector,

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} \quad \text{and} \quad \hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|, \quad (12.26)$$

and a wave ansatz in the form of a cosine modulated by a vertical structure function

$$\Psi(x, y, z, t) = A \Gamma(z) \cos(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad (12.27)$$

where  $A$  is a real amplitude. The solution to Laplace's equation with Neumann boundary

---

<sup>2</sup>We detailed this treatment of the boundary position when deriving the linear equations for surface waves in Section 4.3.4. The same considerations hold here.

conditions in the two half-domains is given by<sup>3</sup>

$$\Psi_1 = A \frac{\cosh[|\mathbf{k}|(z - h_1)]}{\sinh[-|\mathbf{k}|h_1]} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad 0 \leq z \leq h_1 \quad (12.28a)$$

$$\Psi_2 = A \frac{\cosh[|\mathbf{k}|(z + h_2)]}{\sinh[|\mathbf{k}|h_2]} \cos(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad -h_2 \leq z \leq 0. \quad (12.28b)$$

The dispersion relation is obtained by plugging equations (12.28a)-(12.28b) into the interface condition (12.25e), with the following pieces needed (recall each term is evaluated at  $z = 0$ )

$$\partial_{tt}(\rho_2 \Psi_2 - \rho_1 \Psi_1) = -\omega^2 A [\rho_2 \coth(|\mathbf{k}|h_2) - \rho_1 \coth(-|\mathbf{k}|h_1)] \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (12.29a)$$

$$\partial_z \Psi_1 = \partial_z \Psi_2 = A |\mathbf{k}| \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) \quad (12.29b)$$

$$(\partial_{xx} + \partial_{yy})\Psi = -|\mathbf{k}|^2 \Psi, \quad (12.29c)$$

thus leading to the dispersion relation

$$\omega^2 = \frac{|\mathbf{k}|(g \delta \rho + \gamma |\mathbf{k}|^2)}{\rho_1 \coth(|\mathbf{k}|h_1) + \rho_2 \coth(|\mathbf{k}|h_2)}. \quad (12.30)$$

As for the surface waves in Chapter 4, the horizontal wavenumber,  $|\mathbf{k}|$ , determines the vertical scale of the wave. We now examine various cases for stable and unstable waves.

## 12.2.2 Stable traveling plane waves

When the squared angular frequency (12.30) is positive,  $\omega^2 > 0$ , then the waves are interface waves related to those studied in Chapter 4.<sup>4</sup> For example, when the waves are so short that they do not feel the rigid boundaries at  $z = h_1$  and  $z = -h_2$ , then we can set  $\coth(|\mathbf{k}|h_2) \approx 1$  and  $\coth(|\mathbf{k}|h_1) \approx 1$ , in which case the dispersion relation takes on the approximate form

$$\omega^2 \approx \frac{|\mathbf{k}|(g \delta \rho + \gamma |\mathbf{k}|^2)}{\rho_1 + \rho_2} \quad \text{shortwave limit with } |\mathbf{k}|h_1 \gg 1 \text{ and } |\mathbf{k}|h_2 \gg 1. \quad (12.31)$$

This limit (when surface tension is set to zero) corresponds to the deep water waves from Section 4.5.5. For the longwave limit, in which case the waves feel the top and bottom boundaries, we set  $\coth(|\mathbf{k}|h_1) \approx 1/(|\mathbf{k}|h_1)$  and  $\coth(|\mathbf{k}|h_2) \approx 1/(|\mathbf{k}|h_2)$  so that

$$\omega^2 \approx \frac{|\mathbf{k}|^2 g \delta \rho}{\rho_1/h_1 + \rho_2/h_2} \quad \text{longwave limit with } |\mathbf{k}|h_1 \ll 1 \text{ and } |\mathbf{k}|h_2 \ll 1. \quad (12.32)$$

Note that we dropped the surface tension term since  $|\mathbf{k}|$  is very small in the longwave limit. Evidently, since  $\omega^2/|\mathbf{k}|^2$  is independent of  $\mathbf{k}$ , the longwaves are non-dispersive gravity waves and are thus the analog of non-dispersive shallow water gravity waves studied in Section 7.5. The waves here are affected by a modified value for the gravitational acceleration, which we

<sup>3</sup>As noted in Section 4.2.2, solutions to Laplace's equation do not support spatial oscillations in all three directions since the sum of the curvature in each direction (i.e., second partial derivatives) must vanish. Correspondingly, the velocity potential supports traveling waves in the horizontal and exponential behavior in the vertical.

<sup>4</sup>More precisely, if we set  $\rho_1 = 0$  as for a vacuum, then the waves are identical to the surface waves from Chapter 4.

write as<sup>5</sup>

$$g' H \equiv \frac{g \delta \rho}{\rho_1/h_1 + \rho_2/h_2} \quad \text{with} \quad H = h_1 + h_2. \quad (12.33)$$

By introducing this *reduced gravity*,  $g'$ , the two-layer shallow water dispersion relation (12.32) takes the form

$$\omega^2 = (g' H) |\mathbf{k}|^2. \quad (12.34)$$

Recall that the dispersion relation for a single shallow water layer (Section 7.5) is given by  $\omega^2 = (g H) |\mathbf{k}|^2$ . We thus see that long gravity waves on an interface between two fluid layers feels a reduced version of the gravitational acceleration, with  $g' \ll g$  when there is a small density difference. In fact, even the case of a single shallow water can be formulated as a two-layer system, with the upper layer having zero density so that the reduced gravity equals to  $g$ .

### 12.2.3 Unstable exponentially growing plane waves

The case with  $\omega^2 < 0$  leads to unstable wave growth. This case is rendered by  $\delta\rho = \rho_2 - \rho_1 < 0$ . That is, if there is heavy fluid above light fluid then the system can be unstable, depending on whether the gravitationally unstable stratification can overcome the stabilizing effects from surface tension. We now examine various cases to explore this unstable case, known as the *Rayleigh-Taylor instability*.

#### The growth rate

To express the temporal structure of the unstable wave, introduce the growth rate

$$\omega^2 = -\sigma^2 \implies \omega = \pm i\sigma, \quad (12.35)$$

where  $\sigma > 0$  is given by

$$\sigma = \left[ \frac{|\mathbf{k}|(g|\delta\rho| - \gamma|\mathbf{k}|^2)}{\rho_2 \coth(|\mathbf{k}|h_2) + \rho_1 \coth(|\mathbf{k}|h_1)} \right]^{1/2} > 0. \quad (12.36)$$

We furthermore express the velocity potential (12.28a) and (12.28b) as the real part of complex exponentials. In particular, write for the top layer

$$\Psi_1 = A\Gamma(z) \operatorname{Re}[e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}] = A\Gamma(z) e^{\pm\sigma t} \operatorname{Re}[e^{i\mathbf{k}\cdot\mathbf{x}}] = A\Gamma(z) e^{\pm\sigma t} \cos(\mathbf{k}\cdot\mathbf{x}). \quad (12.37)$$

The solution with the time behavior,  $e^{\sigma t}$ , is exponentially growing and this is the unstable wave.

#### All waves are unstable in the absence of surface tension

In the absence of surface tension, the growth rate is given by

$$\sigma = \left[ \frac{g|\mathbf{k}||\delta\rho|}{\rho_2 \coth(|\mathbf{k}|h_2) + \rho_1 \coth(|\mathbf{k}|h_1)} \right]^{1/2} \quad \text{if } \gamma = 0. \quad (12.38)$$

---

<sup>5</sup>The reduced gravity in a shallow water model is defined (e.g., see equation (??)) without the layer thicknesses introduced in equation (12.33). A key difference is that in the present section we start with the non-hydrostatic equations and then take the longwave limit, whereas in Section ?? we only work with the hydrostatic shallow water equations.

Evidently, all waves are unstable, with the smallest waves having the largest growth rate given approximately by

$$\sigma \approx \sqrt{g |\mathbf{k}| |\delta\rho| / (\rho_1 + \rho_2)} \quad \text{with } |\mathbf{k}| h_2 \gg 1 \text{ and } |\mathbf{k}| h_1 \gg 1. \quad (12.39)$$

We thus expect to find the smallest scales rapidly going unstable, with the instability halted only after all of the denser fluid occupies the lower layer.

### The stabilizing role of surface tension

The growth rate (12.36) vanishes at the critical wavenumber

$$|\mathbf{k}|_c^2 = g |\delta\rho| / \gamma. \quad (12.40)$$

All waves with wavenumbers larger than  $|\mathbf{k}|_c$  are stabilized by surface tension, in which case the stable linear waves are capillary-gravity waves. We studied the physics of surface tension in Section ??, where we noted that it can counteract the effects from gravity when the radius of curvature is sufficiently small. To get a sense for the size of these stable waves, recall our discussion of capillary-gravity waves in Section 4.10. For an air-water interface the surface tension is approximately  $\gamma = 0.072 \text{ N m}^{-1} = 0.072 \text{ kg s}^{-2}$ , along with the air density  $\rho_1 = 1 \text{ kg m}^{-3}$ , water density of  $\rho_2 = 1020 \text{ kg m}^{-3}$ . With these physical constants the critical wavenumber is given by

$$|\mathbf{k}|_c \approx 372 \text{ m}^{-1} \implies \Lambda_c = 2\pi / |\mathbf{k}|_c \approx 1.7 \text{ cm}. \quad (12.41)$$

Again, waves of wavelength smaller than  $\Lambda_c$  are stable.

### Maximum growth rate with both surface tension and gravity

We saw above that all waves are unstable without surface tension, with the growth rate increasing as  $|\mathbf{k}|^{1/2}$  as per equation (12.39). However, surface tension introduces a high wavenumber cutoff so that all waves with  $|\mathbf{k}| > |\mathbf{k}|_c^2$  are stable. Again, such waves are stabilized since surface tension dominates over gravity when the radius of curvature is small enough. What is the most unstable wave when there is both surface tension and gravity? One might expect that in the presence of random forcing, this wavenumber will be the one most visibly growing in any particular situation.

To simplify the algebra for computing the most unstable wave, we consider the growth rate (12.36) in the limit that the two rigid boundaries separate to infinity. This limit is not overly constraining since the waves exponentially decay away from the interface, and we expect that the most unstable wavenumber is within an order of magnitude of  $|\mathbf{k}|_c$ . With  $h_1, h_2$  set to infinity the dispersion relation is given by

$$\sigma^2 = \frac{g |\delta\rho| |\mathbf{k}| - \gamma |\mathbf{k}|^3}{\rho_2 + \rho_1}. \quad (12.42)$$

The wavenumber leading to the maximum growth rate is found by setting  $\partial\sigma^2/\partial|\mathbf{k}| = 0$ , in which case

$$|\mathbf{k}|_{\max}^2 = g |\delta\rho| / (3\gamma) = |\mathbf{k}|_c^2 / 3. \quad (12.43)$$

Using the numbers above for an air-water interface, we see that the wavelength for the most unstable wave is roughly  $1.7 \text{ cm} * \sqrt{3} \approx 3 \text{ cm}$ .

### 12.2.4 Further study

The current section follows the approach from section 2 from [Fetter and Walecka \(2003\)](#), whereas chapter X of [Chandrasekhar \(1961\)](#) and chapter 2 of [Sutherland \(2010\)](#) provide more detailed presentations. [The 18 minute mark of this video from Prof. Mollo-Christensen](#) provides a laboratory example of Rayleigh-Taylor instability.

## 12.3 Kelvin-Helmholtz instability

Kelvin-Helmholtz instability arises when the two fluid layers in Figure 12.1 are moving horizontally relative to each other. In the following we assume the velocities are zonal and written  $\hat{\mathbf{x}} U_1$  and  $\hat{\mathbf{x}} U_2$ . Furthermore, we assume the fluids are stably stratified so that  $\delta\rho > 0$ . Finally, to simplify the analysis, assume the rigid boundaries are moved to infinity so that the only boundary of concern is at the fluid interface.

### 12.3.1 Velocity potential

The velocity potential satisfying Laplace's equation in the two half spaces is given by

$$\Psi_1 = -U_1 x + A_1 e^{-|\mathbf{k}|z} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \quad 0 \leq z < \infty \quad (12.44a)$$

$$\Psi_2 = -U_2 x + A_2 e^{|\mathbf{k}|z} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \quad -\infty < z \leq 0. \quad (12.44b)$$

In contrast to the Rayleigh-Taylor instability, as given by equations (12.28a) and (12.28b), we here use complex exponentials, with the real part of each expression assumed. We also introduced two real amplitudes,  $A_1$  and  $A_2$ . Both of these features prove of use in the following analysis. Furthermore, the nonzero background flow makes it less convenient to eliminate the interface height, so that we explicitly consider its wave ansatz in the form

$$\eta = \eta_o e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}. \quad (12.45)$$

To develop a dispersion relation, we make use of the difference in the Bernoulli equation of motion when evaluated at the interface (equation (12.22)), as well as the kinematic and dynamic boundary conditions (12.20c)-(12.20f).

### 12.3.2 Dispersion relation from the interface conditions

Evaluating equation (12.22) at  $z = 0$  for the waves (12.44a), (12.44b), and (12.45), leads to the following relation between the wave amplitudes

$$i[\rho_2(-\omega + k_x U_2)A_2 - \rho_1(-\omega + k_x U_1)A_1] = (g\delta\rho + \gamma|\mathbf{k}|^2)\eta_o \quad (12.46)$$

Likewise, the linearized kinematic boundary conditions (12.20c) and (12.20d) render the relations

$$i(-\omega + k_x U_1)\eta_o = |\mathbf{k}|A_1 \quad (12.47a)$$

$$i(-\omega + k_x U_2)\eta_o = -|\mathbf{k}|A_2. \quad (12.47b)$$

Use of equations (12.47a) and (12.47b) in equation (12.46) leads to the dispersion relation

$$\rho_1(\omega - k_x U_1)^2 + \rho_2(\omega - k_x U_2)^2 = |\mathbf{k}|(g\delta\rho + \gamma|\mathbf{k}|^2). \quad (12.48)$$

We can readily check that this result agrees with that found for the Rayleigh-Taylor instability in equation (12.30). Expanding equation (12.48) and solving the quadratic expression leads to the more conventional form of the dispersion relation

$$\omega = \frac{k_x (U_1 \rho_1 + U_2 \rho_2)}{\rho_1 + \rho_2} \pm \sqrt{\frac{|\mathbf{k}| (g \delta\rho + \gamma |\mathbf{k}|^2)}{\rho_1 + \rho_2} - \frac{k_x^2 \rho_1 \rho_2 (U_1 - U_2)^2}{(\rho_1 + \rho_2)^2}}. \quad (12.49)$$

We consider facets of this dispersion relation in the following.

### 12.3.3 Analysis of the stability condition

The angular frequency is a real number, and the flow is stable, so long as the discriminant in equation (12.49) is positive,

$$\frac{\rho_1 \rho_2 (U_1 - U_2)^2}{(\rho_1 + \rho_2)} < \frac{|\mathbf{k}| (g \delta\rho + \gamma |\mathbf{k}|^2)}{k_x^2} \implies \text{stable state}, \quad (12.50)$$

with the perturbations organizing into stable linear capillary-gravity waves modified by the background velocity. To help understand this stability condition, we find it useful to consider a few special cases. To reduce algebra, assume the wavevector is aligned in the  $\hat{x}$  direction so that  $|\mathbf{k}|^2 = k_x^2$ . Now write the stability condition (12.50) as a condition on the squared velocity difference, in which case

$$(U_1 - U_2)^2 < \frac{(\rho_1 + \rho_2)}{\rho_1 \rho_2} (g \delta\rho / |\mathbf{k}| + \gamma |\mathbf{k}|) \implies \text{stable state}. \quad (12.51)$$

This equation says that the waves are stable so long as the squared velocity difference is insufficient to counteract the stabilizing effects from surface tension and gravity. Notice that the gravitational effects are weighted by the density difference,  $\delta\rho = \rho_2 - \rho_1$ , so that gravity plays no role when the densities are the same. In general, as the wave number increases (wavelength decreases), the effects from gravity acting to stabilize the waves become less important than those from surface tension, whereas the converse situation holds for low wavenumber waves (long wavelength). We illustrate these properties in Figure 12.2.

#### Most easily growing wave

The stability condition (12.51) indicates that that waves that are either small enough or large enough are stable, whereas waves of intermediate length are unstable in the presence of sufficient velocity difference (shear) across the interface. To find the wavenumber of the unstable wave that appears with the least amount of shear (e.g., minimum of the green curve in Figure 12.2), we find where the derivative,  $\partial/\partial|\mathbf{k}|$ , of the right hand of equation (12.51) vanishes, which leads to the critical squared wavenumber

$$|\mathbf{k}|_c^2 = g \delta\rho / \gamma, \quad (12.52)$$

which is the same as found for the Rayleigh-Taylor instability in Section 12.2.3. Evidently, when the wavenumber equals to  $|\mathbf{k}|_c$ , then the stability condition is most easily violated by the smallest squared shear, thus leading to exponential growth for this wave. Plugging in  $|\mathbf{k}|_c$  to the stability condition (12.51) leads to

$$(U_1 - U_2)_{\text{crit}}^4 < \frac{4 \gamma g (\rho_1 + \rho_2)^2 (\rho_2 - \rho_1)}{\rho_1^2 \rho_2^2} \implies \text{stable state}. \quad (12.53)$$

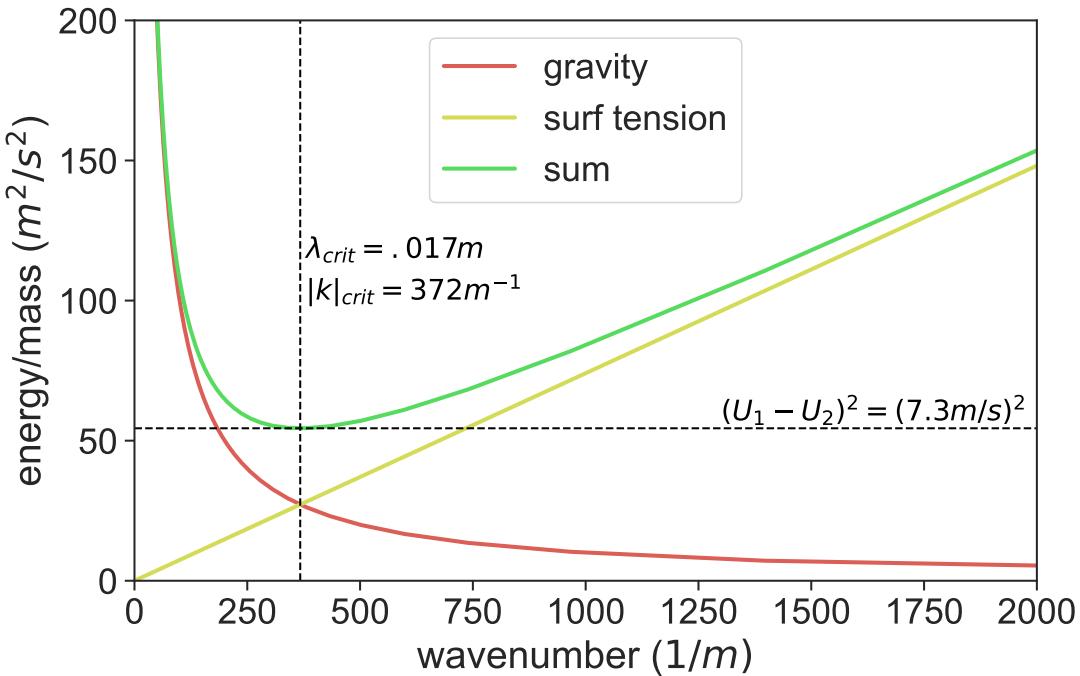


FIGURE 12.2: Stability diagram for Kelvin-Helmholtz instability of the interface between two constant density layers in a gravity field and with surface tension at the interface. The red curve arises from the gravity contribution to equation (12.51); the yellow curve is for the surface tension; and the green curve is the sum. If the squared velocity difference,  $(U_1 - U_2)^2$ , is above the green line then there are unstable waves whose wavenumbers are between the low and high wavenumber bounds defined by the green line. This figure is generated using the following numbers appropriate for the interface between the atmosphere (region 1) and ocean (region 2):  $\rho_1 = 1 \text{ kg m}^{-3}$ ,  $\rho_2 = 1020 \text{ kg m}^{-3}$ ,  $g = 9.8 \text{ m s}^{-2}$ ,  $\gamma = 0.072 \text{ kg s}^{-2}$ . The critical wavenumber from equation (12.52) equals to  $|\mathbf{k}|_c = \sqrt{g \delta \rho / \gamma} = 372 \text{ m}^{-1}$ , which corresponds to a critical wavelength  $\lambda_{crit} = 2\pi / |\mathbf{k}|_c = 0.017 \text{ m}$ . As the velocity difference increases from zero, this is the first wave that goes unstable when the velocity difference reaches  $U_1 - U_2 = 7.3 \text{ m s}^{-1} = 26 \text{ km hr}^{-1}$ .

This expression shows how surface tension and gravity act together to help maintain stability in the face of a velocity difference. Yet if the velocity difference grows, eventually the inequality swaps sign so that the flow becomes unstable. Furthermore, the wavenumber  $|\mathbf{k}|_c$  is the first wave to exponentially grow.

### The case of vanishing surface tension

If the surface tension vanishes then equation (12.53) says that there are always unstable high wavenumber waves for an arbitrarily small velocity difference. More precisely, we return to the general condition (12.51) with  $\gamma = 0$  to find

$$(U_1 - U_2)^2 < \frac{g(\rho_1 + \rho_2) \delta \rho}{|\mathbf{k}| \rho_1 \rho_2} \implies \text{stable state.} \quad (12.54)$$

Evidently, no matter how small the velocity difference, there are waves with high enough wavenumber that violate this inequality and thus lead to an instability. In Figure 12.2, the case with zero surface tension means that the green and red curves are identical, so that without the effects from surface tension, there is no high wavenumber cutoff for the instability.

### The case of vanishing gravity

If we align the layers horizontally rather than vertically, then gravity is no longer able to enhance stability in the face of the velocity difference. Just like in the case with zero surface tension, there are always waves that go unstable in this case. However, the unstable waves here have arbitrarily low wavenumber, so that in the absence of gravity there is no low wavenumber cutoff.

**Stability is enhanced when  $\rho_1/\rho_2 \ll 1$**

If the upper layer has a vanishingly small density relative to the lower layer, then the right hand side of the stability condition (12.51) becomes large. We say that this case is strongly stable since it takes a large velocity difference to produce an instability. A geophysically relevant example is air blowing over water with  $\gamma = 0.072 \text{ N m}^{-1} = 0.072 \text{ kg s}^{-2}$ ,  $\rho_2 = 1020 \text{ kg m}^{-3}$  and air density  $\rho_1 = 1 \text{ kg m}^{-3}$ . Equation (12.53) says that the most unstable wave is stimulated with an air-sea velocity difference

$$U_1 - U_2 = 7.3 \text{ m s}^{-1} = 26 \text{ km hr}^{-1}. \quad (12.55)$$

These numbers are used to generate Figure 12.2.

### 12.3.4 Insights from vorticity

The vorticity vanishes everywhere in the fluid, except at the interface. At the interface the velocity jump leads to a Dirac delta *vortex sheet*. We compute the vorticity of the base state by writing the velocity in terms of Heaviside step functions (equation (??))

$$\mathbf{v} = \hat{\mathbf{x}} [U_1 \mathcal{H}(z) + U_2 \mathcal{H}(-z)] \quad (12.56)$$

so that the vorticity is given by

$$\hat{\mathbf{y}} \cdot \nabla \times \mathbf{v} = \partial_z [U_1 \mathcal{H}(z) + U_2 \mathcal{H}(-z)] = (U_1 - U_2) \delta(z), \quad (12.57)$$

with Figure 12.3 providing an illustration. The vortex sheet is a stationary equilibrium state since the flow felt by adjacent vortices exactly cancels. However, as shown in Figure 12.4, the vortex sheet is unstable to small perturbations. Indeed, in the absence of gravity or surface tension then any perturbation is unstable, which is the Kelvin-Helmholtz instability.

### 12.3.5 Insights from Bernoulli's theorem

Figure 12.5 provides a schematic of the pressure forces active next to the interface in the presence of a wavelike perturbation. A wavelike perturbation along the interface gives rise, through Bernoulli's theorem, to pressure anomalies of opposite sign in the regions near to the interface.<sup>6</sup> In particular, consider the case of an interface that enters one of the two regions and so reduces the cross-sectional area for the fluid flow. The flow is incompressible, and so the flow speed increases in this region, with increased speed associated with an anomalously low pressure. The opposite occurs in the other region, where there is an anomalously high pressure. Evidently, the pressure dipoles increase the amplitude of the wavelike perturbation. If the pressure

<sup>6</sup>At the interface, the pressure has a jump given by the Young-Laplace equation (12.11). We are here interested in the region local to the boundary.

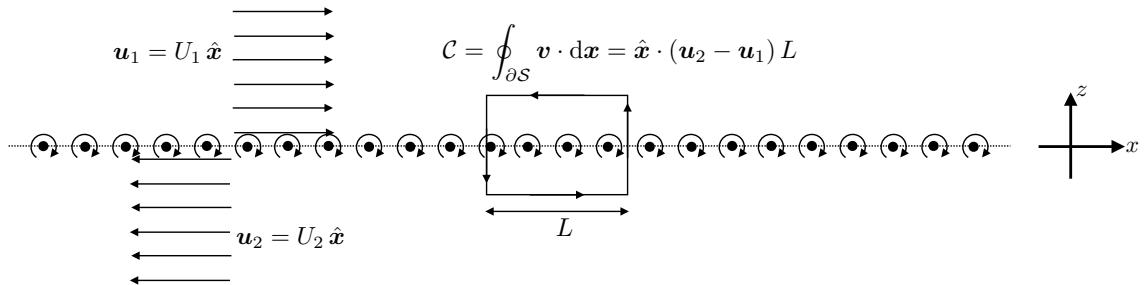


FIGURE 12.3: A vorticity perspective on the Kelvin-Helmholtz instability, whereby the velocity jump leads to a Dirac delta vorticity at the interface and a corresponding circulation for loops that enclose a portion of the sheet. Each circular arrow surrounding a black dot represents a point vortex induced by the velocity jump, with the  $z = 0$  plane filled with a continuum of such point vortices. The circulation around the sheet is  $\mathcal{C} = \hat{x} \cdot (\mathbf{u}_2 - \mathbf{u}_1) L$ , where  $L$  is the length of the side parallel to the sheet. This configuration is a stationary equilibrium since the flow felt by adjacent vortices exactly cancels so that they remain fixed. However, it is generally unstable, as shown in Figure 12.4.

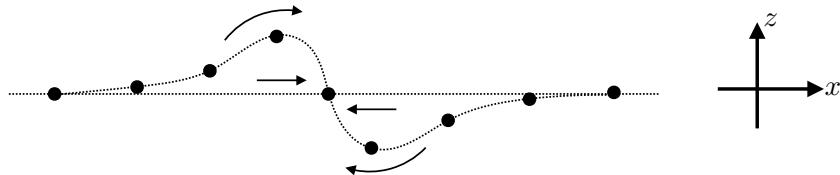


FIGURE 12.4: The equilibrium configuration from Figure 12.3 is unstable to small perturbations, and in the case of zero surface tension and zero gravity any perturbation is unstable. We understand the cause of the instability by observing that if a vortex is displaced away from the sheet, the flow from adjacent vortices and the background flow causes the sheet to roll-up on itself, which is the Kelvin-Helmholtz instability. This figure is a variant of Figure 7.1.3 of [Batchelor \(1967\)](#) (who provides full details of the vorticity interpretation of the Kelvin-Helmholtz instability), Figure 1.3 of [Drazin and Reid \(1981\)](#), and Figure 3.9 of [McWilliams \(2006\)](#).

perturbation is sufficiently strong to overcome the stabilizing effects from gravity and surface tension, then the perturbation grows and becomes nonlinear, which is the Kelvin-Helmholtz instability.

### 12.3.6 An energetic perspective on Kelvin-Helmholtz induced mixing

Kinetic energy of the background flow provides the energy source for the Kelvin-Helmholtz instability. Once the instability fully acts, it produces a well mixed state whereby the density and velocity are mixed within a region local to the initial interface. The kinetic energy of the final mixed state is less than the initial state, which we can infer since mixing removes the velocity jump across the interface; i.e., mixing smooths the velocity profile. Conversely, the gravitational potential energy is increased since some of the light fluid from the upper region is mixed with the heavy fluid from the lower region, and vice versa, thus raising the center of mass of the fluid column.<sup>7</sup>

A deductive analysis of the energetics of mixing is outside our scope. Indeed, without information about the pressure forces causing the base flow, we do not have sufficient information to perform an energy budget. However, we can study the energetic effects from mixing by making reasonable assumptions about the final flow profile, and in so doing we can quantitatively support the above inferences about the effects of mixing on kinetic energy and gravitational

<sup>7</sup>We studied this effect of mixing on gravitational potential energy in Section ??.

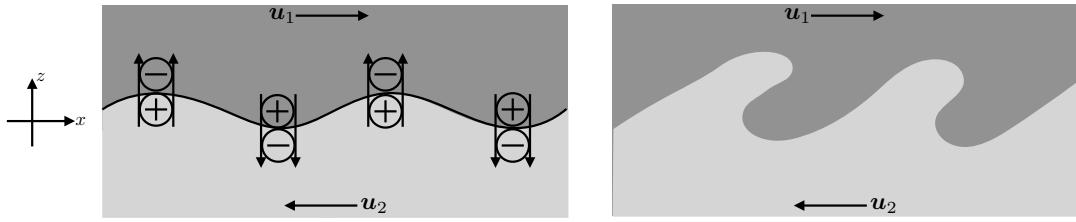


FIGURE 12.5: Schematic of a region near the interface between two uniform density and immiscible fluid regions with nonzero relative motion between the regions. The left panel shows a small wavelike perturbation along the interface that gives rise, through Bernoulli's theorem, to pressure anomalies of opposite signs in the regions on either side of the interface. Evidently, the pressure dipoles increase the amplitude of the wavelike perturbation, thus leading to a positive feedback. If the feedback is sufficiently strong to overcome the stabilizing effects from gravity and surface tension, then it will lead to the Kelvin-Helmholtz instability. The right panel shows the nonlinear stage in which the waves are growing and eventually break.

potential energy. For analytical tractability we assume the mixing region extends over the symmetric range,  $-H \leq z \leq H$ , with negligible signature of mixing outside of this range. Furthermore, we make use of the Boussinesq ocean from Chapter ?? with density a linear function of temperature and with a reference density  $\rho_0$ .

To estimate the thickness,  $H$ , of the mixing region, recall the expression (12.54) allows us to compute the low wavenumber cutoff for the case of Kelvin-Helmholtz instability in the absence of surface tension

$$|\mathbf{k}|_{\text{low}} = \frac{g(\rho_1 + \rho_2)\delta\rho}{\rho_1\rho_2(U_1 - U_2)^2}. \quad (12.58)$$

The unstable waves riding on the interface are exponentially decaying in the direction away from the interface, with their decay scale given by  $|\mathbf{k}|$ . Evidently, unstable Kelvin-Helmholtz waves extend a distance  $\sim 1/|\mathbf{k}|_{\text{low}}$  away from the interface, thus suggesting that a scale for the associated mixing is given by

$$H \sim \frac{\rho_1\rho_2(U_1 - U_2)^2}{g(\rho_1 + \rho_2)\delta\rho} \approx \frac{\rho_0(U_1 - U_2)^2}{2g\delta\rho}, \quad (12.59)$$

where the approximate expression made use of the oceanic Boussinesq approximation.

The initial density and velocity profiles are given by the jump across the interface at  $z = 0$ , with the initial velocity and density having a depth integral of

$$\int_{-H}^H u dz = H(U_1 + U_2) \quad \text{and} \quad \int_{-H}^H \rho dz = H(\rho_1 + \rho_2). \quad (12.60)$$

We assume the final density and velocity profiles have a depth integral equal to those in the initial state, thus ensuring that zonal momentum and heat are conserved by the mixing. For simplicity, assume the final state profiles are linear, in which case (see Figure 12.6)

$$\bar{\rho}(z) = \rho_2 - (\rho_2 - \rho_1)(1 + z/H)/2 = (\rho_1 + \rho_2)/2 - z\delta\rho/(2H) \quad \text{for } -H \leq z \leq H \quad (12.61a)$$

$$\bar{u}(z) = U_2 - (U_2 - U_1)(1 + z/H)/2 = (U_1 + U_2)/2 - z\delta U/(2H) \quad \text{for } -H \leq z \leq H. \quad (12.61b)$$

The depth integrated kinetic energy and gravitational potential energy per horizontal area

$$K = \frac{\rho_0}{2} \int_{-H}^H u^2 dz \quad \text{and} \quad P = g \int_{-H}^H z \rho dz, \quad (12.62)$$

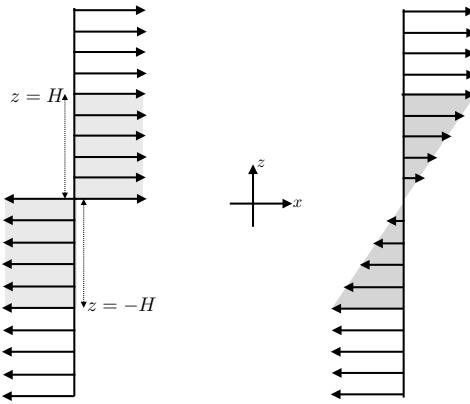


FIGURE 12.6: The initial velocity profile (left panel) for the study of Kelvin-Helmholtz instability, with  $U_1 = U_2$ . This profile is assumed to evolve to the linear profile (right panel) after mixing. The kinetic energy in the mixed profile is less than the kinetic energy in the initial profile. In turn, mixing raises the center of mass of the fluid so that it increases the gravitational potential energy. It is notable that the final state velocity profile shown here could be unstable to shear instability, as discussed in Chapter 13, in which case the shear instability will further act to homogenize the velocity and density.

take on the following initial values

$$K_{\text{init}} = \frac{\rho_0}{2} \int_{-H}^H u^2 dz = H \rho_0 (U_1^2 + U_2^2)/2 \quad (12.63a)$$

$$P_{\text{init}} = g \rho_2 \int_{-H}^0 z dz + g \rho_1 \int_0^H z dz = -g H^2 \delta \rho/2. \quad (12.63b)$$

Use of the linear profiles (12.61a) and (12.61b) render the final mixed state energies

$$K_{\text{final}} = \frac{\rho_0}{2} \int_{-H}^H u^2 dz = H \rho_0 [U_1^2 + U_2^2 + U_1 U_2] / 3 \quad (12.64a)$$

$$P_{\text{final}} = g \int_{-H}^H \bar{\rho} z dz = -g H^2 \delta \rho / 6. \quad (12.64b)$$

As anticipated, we find that the kinetic energy is reduced upon mixing whereas the gravitational potential energy increases

$$K_{\text{final}} - K_{\text{init}} = -H \rho_0 (U_1 - U_2)^2 / 6 < 0 \quad \text{and} \quad P_{\text{final}} - P_{\text{init}} = g H^2 \delta \rho / 3 > 0. \quad (12.65)$$

The ratio of the change in potential energy to the change in kinetic energy is

$$\frac{P_{\text{final}} - P_{\text{init}}}{K_{\text{final}} - K_{\text{init}}} = -\frac{2g}{\rho_0} \frac{\delta \rho / H}{[(U_1 - U_2) / H]^2}, \quad (12.66)$$

which is the ratio of the density stratification to the squared vertical shear. This ratio is a discrete version of the *gradient Richardson number* as discussed in Section 13.7.5. Indeed, in Section 13.7.5 we consider an energetic argument similar to that given here as applied to the final state linear sheared profile in Figure 12.6, which can be unstable to stratified shear instability if the Richardson number is less than 1/4.

### 12.3.7 Further study

Chapter XI of [Chandrasekhar \(1961\)](#) provides a detailed study of Kelvin-Helmholtz instability, with particular comments in Section 101b on the geophysical relevance of the critical shear. Section 2 of [Fetter and Walecka \(2003\)](#) also provides a presentation consistent with that given here. [The first half of this video from Prof. Mollo-Christensen](#) provides laboratory examples of Kelvin-Helmholtz instabilities. This [video from Prof. Worster's fluids lab](#) provides a vivid illustration of Kelvin-Helmholtz instability in a two-layer fluid.





## Chapter 13

### SHEAR INSTABILITY

In this chapter we study *shear instability* as realized in two canonical cases. The first concerns a meridionally sheared zonal flow in a horizontally non-divergent barotropic fluid, with this instability sometimes called *barotropic instability*. We derive integral stability conditions for arbitrary zonal flow profiles, and then study a special flow profile that admits an analytic solution. The analytic study exposes the underlying *wave resonance* mechanism active in shear instability, with this mechanism involving the interaction of *edge waves* that live on the background vorticity field (Section 6.5). We thus interpret shear instability as the constructive interference of two vorticity (edge) waves, where interference supports the mutual exponential growth of both waves that is characteristic of a modal instability.

The second kind of shear instability concerns vertically sheared flows in a gravitationally stratified fluid ( $N^2 > 0$ ) in the absence of planetary rotation ( $f = 0$ ). The stable vertical stratification creates a potential energy barrier that stabilizes the vertically sheared flow relative to the horizontally sheared case without gravity. If the kinetic energy of the vertically sheared flow is large enough, then the potential energy barrier can be overcome to produce a shear instability. Our study of *stratified shear instability* makes use of normal mode stability analysis just like for the barotropic shear flow. The *gradient Richardson number* provides a non-dimensional measure of the potential energy relative to the kinetic energy, with a normal mode instability occurring if the Richardson number is below a critical value. We offer a derivation following [Miles \(1961\)](#) revealing that the critical Richardson number is 1/4.

#### READER'S GUIDE FOR THIS CHAPTER

Surface tension is ignored throughout this chapter, so that we focus on sheared flows of scales larger than the  $\approx 10^{-3}$  m characteristic of capillary waves. To study horizontal shear instability, we assume familiarity with the horizontally non-divergent barotropic model from Chapter ?? and the associated wave mechanics in Sections 6.2 and 6.3. We make particular use of edge waves studied in Section 6.5 as part of our wave resonance interpretation of shear instability. For our study of stratified shear instability, we make use of the perfect Boussinesq fluid from Chapter ???. In linearizing the equations we follow many of the same steps used for the study of internal gravity waves in Chapter 9. The linear partial differential equation appearing in the stability analysis is known as the *Taylor-Goldstein* equation, which is very similar to the *Rayleigh* equation encountered in the barotropic shear instability. Methods and concepts from this chapter are very useful in the study of baroclinic instability in Chapter 14.

Kelvin and Helmholtz studied the interfacial instability described in Chapter 12. Even so, the continuous shear layer instability of the present chapter is also, sometimes, referred

to as Kelvin-Helmholtz instability.

---

<b>13.1</b>	<b>Loose threads</b>	<b>390</b>
<b>13.2</b>	<b>Global versus local instabilities</b>	<b>391</b>
<b>13.3</b>	<b>Governing barotropic equations</b>	<b>391</b>
13.3.1	Velocity equation	392
13.3.2	Eddy kinetic energy	392
13.3.3	Pressure equation	393
13.3.4	Meridional velocity equation	393
<b>13.4</b>	<b>Barotropic flow with a modulated wave ansatz</b>	<b>394</b>
13.4.1	Accounting for the edge wave direction	394
13.4.2	Polarization relations	394
13.4.3	Complex phase velocity	395
13.4.4	Rayleigh-Kuo equation for the streamfunction	396
13.4.5	Rayleigh-Kuo equation for the meridional displacement	396
13.4.6	Phase and zonal averaged eddy kinetic energy equation	397
13.4.7	Phase lines of unstable waves tilt into the shear	397
<b>13.5</b>	<b>Integral conditions necessary for shear instability</b>	<b>399</b>
13.5.1	Rayleigh-Kuo inflection point theorem	399
13.5.2	Fjørtoft's theorem	400
13.5.3	Critical latitude theorem	401
13.5.4	Stability conditions for sample profiles	402
<b>13.6</b>	<b>Interacting edge waves and shear instability</b>	<b>404</b>
13.6.1	Phase locked streamfunction and Rayleigh equation	406
13.6.2	Kinematic boundary condition at $y = \pm L$	406
13.6.3	Dynamic boundary condition at $y = L$	407
13.6.4	Dynamic boundary condition at $y = -L$	408
13.6.5	Phase velocity for phase locked edge waves	409
13.6.6	Dispersion relation and its interpretation	410
13.6.7	Plotting the $\tilde{\psi}(y)$ streamfunctions	413
13.6.8	Lack of mutual wave growth for stable flows	415
13.6.9	Further study	415
<b>13.7</b>	<b>Integral conditions for stratified shear stability/instability</b>	<b>416</b>
13.7.1	Governing equations	416
13.7.2	The linear vorticity equation	417
13.7.3	Taylor-Goldstein equation	418
13.7.4	Richardson number and the stability conditions	419
13.7.5	Richardson number and mixing energetics	421
13.7.6	Constraining the phase velocity of unstable waves	422
13.7.7	Further study	425
<b>13.8</b>	<b>A vertically sheared homogeneous fluid with a free surface</b>	<b>426</b>
13.8.1	Linearized governing equations	426
13.8.2	Hydrostatic fluctuations are stable	427
13.8.3	Taylor-Goldstein equation	428
13.8.4	Necessary conditions for instability	428
<b>13.9</b>	<b>Exercises</b>	<b>429</b>

---

## 13.1 Loose threads

- Develop Kelvin's cat's eye streamfunction math and illustrate it.

- Discuss Couette flow stability in Section 13.5.1 even though it satisfies the Rayleigh inflection point theorem.
- Look at the pressure field in the waves to offer a force balance interpretation of the instabilities. Also, to help understand why unstable waves tilt into the shear.
- Discuss pseudo-momentum argument as a more general stability argument than provided by Rayleigh criteria.

## 13.2 Global versus local instabilities

The wave/modal interpretation of shear instability that we pursue in this chapter contrasts to the variety of symmetric instabilities studied in Chapter 11. In particular, symmetric stability of a given flow can be deduced by the local necessary and sufficient condition,  $f Q < 0$ , with  $Q$  the Ertel potential vorticity. A corresponding mechanistic interpretation follows from parcel arguments. As emphasized by [Cushman-Roisin and Beckers \(2011\)](#) (see their Chapter 17), wave instabilities, such as shear instability of this chapter, are not characterized by a local flow property. The reason is that a wave instability arises from the constructive interaction between coherent wave motion, with that interaction a function of boundary conditions and phase relations. Hence, a quantitative understanding of shear instability requires the solution of an eigenvalue problem to determine properties of the interacting waves.

*Squire's theorem* ([Squire, 1933](#)) states that for every three-dimensional perturbation to a plane shear flow, there exists a more unstable two-dimensional perturbation.<sup>1</sup> Hence, to characterize the most unstable perturbations, it is sufficient to study shear instability in a two-dimensional flow. Even so, as presented in Chapter 3 of [Smyth and Carpenter \(2019\)](#), it can be pedagogically useful to start from the more general three-dimensional case and then show the validity of Squire's theorem.

## 13.3 Governing barotropic equations

As in the study of edge waves in Section 6.5, we here consider flow of a horizontally non-divergent barotropic fluid in the presence of a prescribed background zonal flow that is a function of latitude

$$\mathbf{u}_b = u_b(y) \hat{x}. \quad (13.1)$$

This background velocity has zero material acceleration

$$(\partial_t + \mathbf{u}_b \cdot \nabla) \mathbf{u}_b = 0, \quad (13.2)$$

so that it is in exact geostrophic balance with a background pressure gradient

$$f \hat{z} \times \mathbf{u}_b = -\nabla \varphi_b \implies f u_b = -\partial_y \varphi_b. \quad (13.3)$$

Much of this chapter (e.g., the linearized equations with a modulated wave ansatz in Section 13.4 and the interacting edge waves in Section 13.6) considers the case of a non-rotating reference frame ( $f = 0$ ), in which case the background flow is generated by an unspecified pressure gradient. In other sections we retain planetary rotation in the form of the  $\beta$  plane, such as for

---

<sup>1</sup>See [Drazin and Reid \(2004\)](#) or Section 11.8 of [Kundu et al. \(2016\)](#) for more details of Squire's theorem. We illustrate Squire's theorem in Section 14.7.3 when studying the growth rate of unstable Eady edge waves.

the integral stability conditions of Section 13.5. In the remainder of this section we develop the equations describing the velocity and the kinetic energy of fluid flow in the presence of the prescribed zonal background flow, and with details for generation of the background flow left unspecified.

### 13.3.1 Velocity equation

Writing the velocity and pressure as the sum of a background plus a fluctuation

$$\mathbf{u} = \mathbf{u}_b + \mathbf{u}' \quad \text{and} \quad \varphi = \varphi_b + \varphi', \quad (13.4)$$

leads to the material acceleration

$$\frac{D\mathbf{u}}{Dt} = (\partial_t + \mathbf{u} \cdot \nabla) \mathbf{u} = (\partial_t + \mathbf{u}' \cdot \nabla) \mathbf{u}' + (\mathbf{u}' \cdot \nabla) \mathbf{u}_b + (\mathbf{u}_b \cdot \nabla) \mathbf{u}', \quad (13.5)$$

and the corresponding equation of motion

$$(\partial_t + \mathbf{u}' \cdot \nabla) \mathbf{u}' + (\mathbf{u}' \cdot \nabla) \mathbf{u}_b + (\mathbf{u}_b \cdot \nabla) \mathbf{u}' + f \hat{\mathbf{z}} \times \mathbf{u}' = -\nabla \varphi', \quad (13.6)$$

which takes on the linearized form with the given background flow (13.1)

$$(\partial_t + u_b \partial_x) \mathbf{u}' + \hat{\mathbf{x}} v' \partial_y u_b + f \hat{\mathbf{z}} \times \mathbf{u}' = -\nabla \varphi'. \quad (13.7)$$

### 13.3.2 Eddy kinetic energy

As detailed in Section ??, energetic transfers in the horizontally non-divergent barotropic fluid only involve the kinetic energy since the gravitational potential energy is a constant. Working with the linearized equation of motion (13.7), we find that the kinetic energy per mass contained in the fluctuating flow,  $\mathbf{u}' \cdot \mathbf{u}' / 2$ , satisfies

$$[\partial_t + u_b \partial_x] (\mathbf{u}' \cdot \mathbf{u}') / 2 = -u' v' \partial_y u_b - \nabla \cdot (\mathbf{u}' \varphi). \quad (13.8)$$

We commonly refer to  $\mathbf{u}' \cdot \mathbf{u}' / 2$  as the *eddy kinetic energy*. The first term on the right hand side of equation (13.8) arises from fluctuations of the zonal and meridional velocity weighted by the meridional derivative of the background zonal flow. This term is associated with the transfer of kinetic energy from the background flow to the fluctuating flow.<sup>2</sup> The second term is the convergence of the pressure flux that is determined by the fluctuating flow. Taking a zonal average over the domain, and assuming all zonal boundary contributions vanish, leads to

$$\partial_t \bar{\mathcal{K}} = -\overline{u' v'} \partial_y u_b - \partial_y (\overline{v' \varphi'}), \quad (13.9)$$

where we defined the zonal averaged eddy kinetic energy in the fluctuating fields

$$\bar{\mathcal{K}} = \overline{\mathbf{u}' \cdot \mathbf{u}'} / 2. \quad (13.10)$$

---

<sup>2</sup>Note that we do not have access to the energy equation for the background flow since it is prescribed and remains static.

Evidently, equation (13.9) says that the zonal averaged kinetic energy of the fluctuating fields has an Eulerian time derivative determined by the following two terms:

$$-\partial_y(\overline{v' \varphi'}) = \text{meridional convergence of pressure flux} \quad (13.11a)$$

$$-\overline{u' v'} \partial_y u_b = \text{shear production.} \quad (13.11b)$$

We assume that the pressure convergence term vanishes when integrated over the meridional extent of the domain (e.g.,  $v' = 0$  along the meridional boundaries), in which case it represents a redistribution or transport that moves eddy kinetic energy around but does not alter its domain integrated value. In contrast, the shear production term is a source/sink that measures the rate that eddy kinetic energy is modified via the zonal correlations between  $u'$  and  $v'$  and as modulated by the meridional gradient of the background flow. As noted above, this term provides a transfer of kinetic energy between the mean flow and the fluctuating flow. It is expected that this term provides the source of the growing kinetic energy of an unstable wave. Note that its name arises since it is a nonzero production of kinetic energy in the presence of a background shear.

We can interpret the shear production term as the meridional flux of zonal momentum contained in the fluctuating field. Evidently, kinetic energy in the fluctuating fields increases if this flux is down the gradient of the background meridional shear

$$\overline{v' u'} \partial_y u_b < 0 \quad \text{growing kinetic energy of fluctuations.} \quad (13.12)$$

Such downgradient transport by the growing fluctuations acts in a direction that smooths the meridional shear of the background zonal flow.

### 13.3.3 Pressure equation

Taking the divergence of the linearized velocity equation (13.7), and noting that  $\nabla \cdot \mathbf{u}' = 0$ , leads to the Poisson equation for pressure

$$-\nabla^2 \varphi' = 2(\partial_y u_b)(\partial_x v') + \beta u' - f \zeta', \quad (13.13)$$

where we introduced the relative vorticity contained in the fluctuating field (i.e., the eddy vorticity)

$$\zeta' = \partial_x v' - \partial_y u'. \quad (13.14)$$

The analysis of interacting edge waves in Section 13.6 is formulated in a non-rotating reference frame, in which the pressure equation simplifies to

$$-\nabla^2 \varphi' = 2(\partial_y u_b)(\partial_x v'). \quad (13.15)$$

### 13.3.4 Meridional velocity equation

The meridional component of the linearized velocity equation (13.7) is given by

$$\partial_t v' + u_b \partial_x v' + f u' = -\partial_y \varphi'. \quad (13.16)$$

Setting  $f = 0$  and taking the Laplacian leads to

$$\partial_t(\nabla^2 v') + \nabla^2(u_b \partial_x v') = -\partial_y(\nabla^2 \varphi'). \quad (13.17)$$

Use of the pressure equation (13.15) then provides an equation for the meridional velocity

$$(\partial_t + u_b \partial_x)(\nabla^2 v') = (\partial_x v') \partial_{yy} u_b. \quad (13.18)$$

Evidently, the background zonal flow provides an advection of  $\nabla^2 v'$  as well as an interaction term on the right hand side.

## 13.4 Barotropic flow with a modulated wave ansatz

We here study a modulated wave ansatz for a non-rotating reference frame using the linearized flow in the presence of a background state with a meridionally sheared zonal flow. Following the development in Section 6.5, we expect plane traveling waves in the zonal direction, with a meridionally dependent modulation function. This expectation leads to the ansatz for the streamfunction

$$\psi(x, y, t) = \tilde{\psi}(y) e^{i(kx - \omega t)}. \quad (13.19)$$

### 13.4.1 Accounting for the edge wave direction

An edge wave can move in either the positive or negative  $\hat{x}$  direction, so that the wavevector is

$$\mathbf{k} = |\mathbf{k}| \hat{\mathbf{k}} = |k| \hat{\mathbf{k}} = \pm |k| \hat{x}. \quad (13.20)$$

A convenient means to incorporate the two directions of motion is to write the phase as

$$k x - \omega t = k [x - (\omega/k) t] = k (x - c t), \quad (13.21)$$

where the phase velocity is<sup>3</sup>

$$\mathbf{c}_p = (\omega/|k|) \hat{\mathbf{k}} = (\omega/k) \hat{x} = c \hat{x}. \quad (13.22)$$

For the geometry of this problem, an edge wave only moves along one-dimension, in which case it is sufficient to refer to  $c = \omega/k$  as the phase velocity and  $|c| = \omega/|k|$  the phase speed.<sup>4</sup> A real and positive wave velocity,  $c > 0$ , has  $\mathbf{k} = k \hat{x}$  with  $k > 0$ , thus indicating a stable edge wave with phase moving in the  $+\hat{x}$  direction. Conversely,  $c < 0$  has  $\mathbf{k} = k \hat{x} = -|k| \hat{x}$  with  $k < 0$ , thus indicating a stable edge wave with phase moving in the  $-\hat{x}$  direction.

### 13.4.2 Polarization relations

The horizontal velocity components are computed from the streamfunction via

$$\psi = \tilde{\psi} e^{i k (x - c t)} \quad (13.23a)$$

$$\mathbf{u}' = \hat{z} \times \nabla \psi = e^{i k (x - c t)} (-\hat{x} \partial_y + \hat{y} i k) \tilde{\psi} \quad (13.23b)$$

$$u' = (i/k) \partial_y v' = -e^{i k (x - c t)} \partial_y \tilde{\psi} \quad (13.23c)$$

$$v' = i k \psi = i k e^{i k (x - c t)} \tilde{\psi}. \quad (13.23d)$$

<sup>3</sup>Remember that for a stable wave,  $\omega \geq 0$  in this book (see Section 1.2.3).

<sup>4</sup>Many books refer to  $c$  as the phase speed, even though  $c$  can be positive or negative. However, as emphasized in Section 1.5.2, the phase speed is the non-negative magnitude of the phase velocity.

As a check we confirm that the horizontal velocity is non-divergent

$$\partial_x u' + \partial_y v' = i k \partial_y \tilde{\psi} e^{ik(x-ct)} (-1 + 1) = 0. \quad (13.24)$$

To get the pressure perturbation, return to the zonal component of the velocity equation (13.7) (with  $f = 0$ )

$$(\partial_t + u_b \partial_x) u' + v' \partial_y u_b = -\partial_x \varphi', \quad (13.25)$$

and make the ansatz

$$\varphi' = \tilde{\varphi}(y) e^{ik(x-ct)}. \quad (13.26)$$

Use of this pressure ansatz along with the horizontal velocity equations (13.23c) and (13.23d), yields the pressure amplitude in terms of the streamfunction amplitude

$$\tilde{\varphi} = [(u_b - c) \partial_y - \partial_y u_b] \tilde{\psi}. \quad (13.27)$$

In summary, the amplitude equations (polarization relations) for the horizontal velocity components and the pressure are given, in terms of the streamfunction, by

$$\psi = \tilde{\psi}(y) e^{ik(x-ct)} \quad (13.28a)$$

$$u' = \tilde{u}(y) e^{ik(x-ct)} \quad \text{and} \quad v' = \tilde{v}(y) e^{ik(x-ct)} \quad \text{and} \quad \varphi' = \tilde{\varphi}(y) e^{ik(x-ct)} \quad (13.28b)$$

$$\tilde{u} = -\partial_y \tilde{\psi} \quad \text{and} \quad \tilde{v} = i k \tilde{\psi} \quad \text{and} \quad \tilde{\varphi} = [(u_b - c) \partial_y - \partial_y u_b] \tilde{\psi}. \quad (13.28c)$$

### 13.4.3 Complex phase velocity

To investigate shear instability, we are interested in flow properties that lead to the phase velocity,  $c$ , having an imaginary part, in which case

$$c = c_r + i c_i = c_r + i \sigma/k. \quad (13.29)$$

In this case, it is the real part,  $c_r$ , that is the phase velocity whereas  $c_i$  measures the decay rate or growth rate for the wave amplitude

$$\psi(x, y, t) = \tilde{\psi}(y) e^{ik(x-c_r t)} e^{k c_i t} = \tilde{\psi}(y) e^{ik(x-c_r t)} e^{\sigma t}. \quad (13.30)$$

If  $c_i > 0$  then the wave is exponentially unstable with *growth rate*,

$$\sigma = k c_i, \quad (13.31)$$

whereas if  $\sigma < 0$  then the wave exponentially decays.

As we see in Section 13.4.4, if  $c$  is complex then it appears along with its complex conjugate, so that the exponentially growing mode and the decaying mode appear as a pair. Furthermore, note that the complex conjugate of the streamfunction (13.30) is given by

$$\psi^* = \tilde{\psi}^*(y) e^{-ik(x-c_r t)} e^{\sigma t}, \quad (13.32)$$

so that

$$|\psi|^2 = \psi^* \psi = |\tilde{\psi}|^2 e^{2\sigma t}, \quad (13.33)$$

along with the analogs for the horizontal velocity components

$$|u'|^2 = u' (u')^* = |\partial_y \tilde{\psi}|^2 e^{2\sigma t} \quad \text{and} \quad |v'|^2 = v' (v')^* = k^2 |\tilde{\psi}|^2 e^{2\sigma t}. \quad (13.34)$$

These identities are used when developing the phase averaged kinetic energy budget in Section 13.4.6.

#### 13.4.4 Rayleigh-Kuo equation for the streamfunction

Inserting the modulated wave ansatz (13.19) into the linearized vorticity equation (6.82) leads to the *Rayleigh equation*

$$(u_b - c) (\partial_{yy} - k^2) \tilde{\psi} + \partial_y \zeta_b \tilde{\psi} = 0. \quad (13.35)$$

Or, in the presence of  $\beta \neq 0$  we find the Rayleigh-Kuo equation (6.84)

$$(u_b - c) (\partial_{yy} - k^2) \tilde{\psi} + (\beta + \partial_y \zeta_b) \tilde{\psi} = 0. \quad (13.36)$$

For a study of instabilities, the phase velocity,  $c = \omega/k$ , and the streamfunction,  $\tilde{\psi}$ , are generally complex, whereas all other terms are real. Hence, the complex conjugate of the Rayleigh-Kuo equation (13.36) is given by

$$(u_b - c^*) (\partial_{yy} - k^2) \tilde{\psi}^* + (\beta + \partial_y \zeta_b) \tilde{\psi}^* = 0. \quad (13.37)$$

Evidently, if  $c$  satisfies the Rayleigh-Kuo equation (13.36) with streamfunction  $\tilde{\psi}$ , then  $c^*$  satisfies the complex conjugate equation (13.37) with streamfunction  $\tilde{\psi}^*$ . Hence, the phase velocities come in complex conjugate pairs.

#### 13.4.5 Rayleigh-Kuo equation for the meridional displacement

Let  $\xi(x, y, t)$  be the meridional component of the fluid particle displacement from its equilibrium position. This particle displacement satisfies the evolution equation

$$v' = (\partial_t + u \partial_x) \xi = [\partial_t + (u' + u_b) \partial_x] \xi. \quad (13.38)$$

Linearizing for small amplitude displacements leads to

$$v' = (\partial_t + u_b \partial_x) \xi, \quad (13.39)$$

with the usual ansatz,

$$\xi = \tilde{\xi}(y) e^{ik(x-ct)}, \quad (13.40)$$

yielding

$$\tilde{\psi} = (u_b - c) \tilde{\xi}. \quad (13.41)$$

Evidently, for small amplitude motion, the streamfunction amplitude,  $\tilde{\psi}$ , equals to the meridional particle excursion amplitude,  $\tilde{\xi}$ , multiplied by the difference between the phase velocity and the background flow,  $u_b - c$ .

With the relation (13.41), we can convert the Rayleigh equation (13.35) for the streamfunction into an equation for the meridional particle excursion. For this purpose make use of the derivative

$$\partial_{yy} \tilde{\psi} = \tilde{\xi} \partial_{yy} u_b + 2 \partial_y \tilde{\xi} \partial_y u_b + (u_b - c) \partial_{yy} \tilde{\xi}, \quad (13.42)$$

along with a few lines of algebra to derive the Rayleigh-Kuo equation in terms of the meridional excursion

$$\partial_y[(u_b - c)^2 \partial_y \tilde{\xi}] = (u_b - c) [-\beta + k^2 (u_b - c)] \tilde{\xi}. \quad (13.43)$$

We make use of equation (13.43) in Section 13.5.3 to derive a condition required for a modal perturbation to initiate an instability.

### 13.4.6 Phase and zonal averaged eddy kinetic energy equation

In Section 13.3.2 we developed the equation for the kinetic energy of the fluctuating field and took its zonal average to find

$$\partial_t(\overline{u' u'} + v' v')/2 = -\overline{u' v'} \partial_y u_b - \partial_y(\overline{v' \varphi'}). \quad (13.44)$$

Following the methods of complex variables and phase averaging as detailed in Section ??, use of the polarization relations from Section 13.4.2, and recognition that the phase velocity is generally complex according to equations (13.33) and (13.34), lead to<sup>5</sup>

$$2 \langle \overline{u' u'} \rangle = |\partial_y \tilde{\psi}|^2 e^{2\sigma t} \quad (13.45a)$$

$$2 \langle \overline{v' v'} \rangle = k^2 |\tilde{\psi}|^2 e^{2\sigma t} \quad (13.45b)$$

$$2 \langle \overline{u' v'} \rangle = k \operatorname{Re}[i \tilde{\psi}^* \partial_y \tilde{\psi}] e^{2\sigma t} = -k \operatorname{Im}[\tilde{\psi}^* \partial_y \tilde{\psi}] e^{2\sigma t} = k \operatorname{Im}[\tilde{\psi} \partial_y \tilde{\psi}^*] e^{2\sigma t}, \quad (13.45c)$$

along with the phase average of the meridional velocity and pressure

$$2 e^{-2\sigma t} \langle \overline{v' \varphi'} \rangle = k \operatorname{Re}[i \tilde{\psi} \tilde{\varphi}^*] \quad (13.46a)$$

$$= -k \operatorname{Im}[\tilde{\psi} \tilde{\varphi}^*] \quad (13.46b)$$

$$= -k \operatorname{Im}[(u_b - c) \tilde{\psi} \partial_y \tilde{\psi}^* - \partial_y u_b |\tilde{\psi}|^2] \quad (13.46c)$$

$$= -k \operatorname{Im}[(u_b - c) \tilde{\psi} \partial_y \tilde{\psi}^*], \quad (13.46d)$$

where we used  $\operatorname{Im}[\partial_y u_b |\tilde{\psi}|^2] = 0$  for the final equality. Bringing terms together leads to the phase and zonal averaged eddy kinetic energy budget

$$2\sigma \left[ |\partial_y \tilde{\psi}|^2 + k^2 |\tilde{\psi}|^2 \right] / 2 = \underbrace{-k \partial_y u_b \operatorname{Im} [\tilde{\psi} \partial_y \tilde{\psi}^*] / 2}_{\text{shear production}} + \underbrace{k \partial_y \operatorname{Im} [(u_b - c) \tilde{\psi} \partial_y \tilde{\psi}^*] / 2}_{\text{KE flux convergence}}, \quad (13.47)$$

where we recognize  $\sigma/k = c_i$  according to equation (13.29).

### 13.4.7 Phase lines of unstable waves tilt into the shear

As noted in Section 13.3.2, the only way for the globally integrated kinetic energy to grow is through the shear production term, since the globally integrated pressure flux convergence vanishes. A sufficient, though not necessary, condition for global kinetic energy growth is that the shear production is positive at each point within the fluid domain. From equation (13.47) we thus have

$$k \partial_y u_b \operatorname{Im} [\tilde{\psi}^* \partial_y \tilde{\psi}] > 0 \implies \text{sufficient condition for kinetic energy growth.} \quad (13.48)$$

<sup>5</sup>Exercise 13.1 derives the identity  $\operatorname{Im} [\tilde{\psi}^* \partial_y \tilde{\psi}] = -\operatorname{Im} [\tilde{\psi} \partial_y \tilde{\psi}^*]$  used in equation (13.45c).

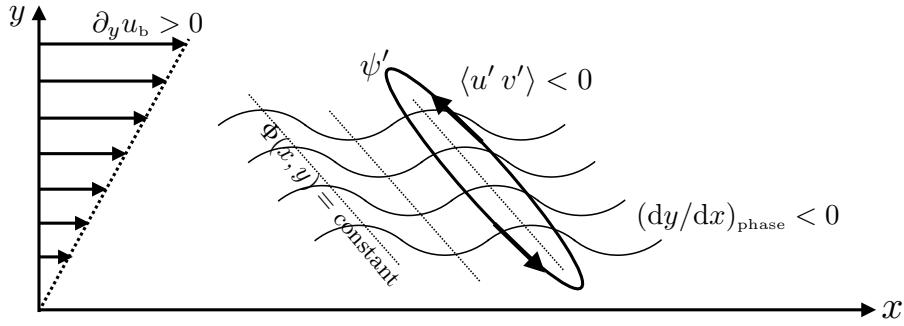


FIGURE 13.1: The lines of constant phase,  $\Phi(x, y) = kx - \alpha(y)$ , for the streamfunction (or the meridional velocity) in an unstable barotropic wave. The wave depicted here has an increasing kinetic energy since the phase lines are tilted into the shear. That is, this orientation of the wave ensures that the barotropic shear production increases the kinetic energy of the wave as per equation (13.56). Namely, since the waves are transverse, fluid particles move on constant phase lines, with the tilt shown here ensuring that  $\langle u' v' \rangle < 0$ . That is, a meridionally positive particle motion corresponds to a zonally negative motion, and vice versa. This behavior is characteristic of unstable waves whose kinetic energy grows by feeding off the unstable background shear state. Observe that if one placed a passive tracer in the flow, or an array of fluid particles, then they would be stretched to align with the shear rather than against the shear. The growing wave, however, is an active flow feature, with the present analysis indicating that energy growth for this feature requires phase lines to tilt into the shear as depicted here.

If the kinetic energy growth arises from shear instability, then equation (13.48) offers a sufficient condition for shear instability.

To develop a geometric view for the energy growth condition (13.48), introduce the phase,  $\alpha(y)$ , of the streamfunction via

$$\tilde{\psi}(y) = |\tilde{\psi}(y)| e^{i\alpha(y)}, \quad (13.49)$$

so that

$$\text{Im} [\tilde{\psi}^* \partial_y \tilde{\psi}] = \text{Im} [|\tilde{\psi}| e^{-i\alpha(y)} (\partial_y |\tilde{\psi}| + i |\tilde{\psi}| \partial_y \alpha) e^{i\alpha(y)}] = |\tilde{\psi}|^2 \partial_y \alpha, \quad (13.50)$$

thus bringing the sufficient condition (13.48) to the form

$$k \partial_y u_b |\tilde{\psi}|^2 \partial_y \alpha > 0 \implies \text{sufficient condition for energy growth.} \quad (13.51)$$

Now write the full streamfunction as

$$\psi = \tilde{\psi}(y) e^{ik(x-ct)} = |\tilde{\psi}(y)| e^{i\alpha(y)} e^{i(kx-ct)} = |\tilde{\psi}(y)| e^{i(\Phi(x,y)-ct)}, \quad (13.52)$$

where we defined the spatial phase function

$$\Phi(x, y) \equiv kx + \alpha(y). \quad (13.53)$$

Lines of constant phase are defined by

$$d\Phi = 0 = k dx + (\partial_y \alpha) dy \implies (dy/dx)_{\text{phase}} = -k/\partial_y \alpha, \quad (13.54)$$

so that the instability condition (13.48) can be written

$$k \partial_y u_b |\tilde{\psi}|^2 \partial_y \alpha = -k^2 |\tilde{\psi}|^2 \frac{\partial_y u_b}{(dy/dx)_{\text{phase}}} > 0 \implies \text{sufficient condition for energy growth.} \quad (13.55)$$

Simplifying this equation leads to the condition for the ratio of the shear and the phase slope

$$\frac{\partial_y u_b}{(dy/dx)_{\text{phase}}} < 0 \implies \text{sufficient condition for energy growth.} \quad (13.56)$$

This inequality says that kinetic energy of the wave grows when the wave's phase lines tilt into the background flow shear, such as depicted in Figure 13.1. The tilted phase lines reflect the ability of the wave to extract kinetic energy from the background state. This geometric property offers a visual indicator that the wave is acting on an unstable shear state, thus providing a valuable diagnostic tool for identifying shear instabilities as they are happening.

## 13.5 Integral conditions necessary for shear instability

In this section we develop integral conditions for stability of an inviscid horizontally sheared fluid on a  $\beta$  plane, with the conditions arrived at by forming spatial integrals of the Rayleigh equation (13.35) and the Rayleigh-Kuo equation (13.36). These integral conditions allow us to determine stability properties even without explicitly solving the detailed instability problem for a particular flow (see Section 13.6 for a solution example). We do so by deriving geometric conditions that are necessary for the flow to be unstable, or conversely that are sufficient to ensure the flow is stable. We also consider conditions needed for a particular perturbation to initiate an instability. Generally, the conditions we derive are necessary though not sufficient to ensure instability. Evidently, even if a flow satisfies the necessary conditions for instability, the flow can still be stable. This situation for the modal instabilities of sheared flows contrasts to the parcel instabilities from Chapter 11, where the instability conditions are both necessary and sufficient.

### 13.5.1 Rayleigh-Kuo inflection point theorem

Here we establish the *Rayleigh instability criteria*, also known as the *Rayleigh inflection-point criteria*. In the presence of  $\beta$ , it is known as the *Rayleigh-Kuo condition*. To proceed, start from the Rayleigh-Kuo equation (13.36) written as

$$(\partial_{yy} - k^2) \tilde{\psi} + \frac{(\beta - \partial_{yy} u_b) \tilde{\psi}}{u_b - c} = 0. \quad (13.57)$$

Multiplying by  $\tilde{\psi}^*$  (complex conjugate of  $\tilde{\psi}$ ) and integrating over the meridional extent of the domain,  $\mathcal{R}$ , leads to

$$\int_{\mathcal{R}} [\partial_y (\tilde{\psi}^* \partial_y \tilde{\psi}) - |\partial_y \tilde{\psi}|^2 - k^2 |\tilde{\psi}|^2] dy = - \int_{\mathcal{R}} \frac{(\beta - \partial_{yy} u_b) |\tilde{\psi}|^2}{u_b - c} dy. \quad (13.58)$$

If the streamfunction or its derivatives vanish on the meridional boundaries (or vanish when bounded away from the region of interest), then the left hand side is a negative real number. Now all terms on the right hand side are real except, possibly, the phase velocity,  $c$ . Hence, this equation is self-consistent only if the imaginary part of the right hand side vanishes

$$c_i \int_{\mathcal{R}} \frac{(\beta - \partial_{yy} u_b) |\tilde{\psi}|^2}{|u_b - c|^2} dy = 0. \quad (13.59)$$

This condition can be satisfied in two ways. The first way is if the phase velocity is real, so that  $c_i = 0$  and hence all waves are stable.<sup>6</sup> The second way is if the integral vanishes. For the integral to vanish requires  $\beta - \partial_{yy}u_b$  to change sign somewhere in the domain, since all the other terms in the integral are positive. That is, somewhere in the domain there must be an extrema of the base state's absolute vorticity,

$$\beta - \partial_{yy}u_b = \partial_y(f + \zeta_b). \quad (13.60)$$

There are many qualifiers to this result. In particular, for  $\beta - \partial_{yy}u_b$  to change sign in the domain represents a necessary condition for a shear instability, and yet it is not a sufficient condition for instability. Indeed, there are flow profiles that satisfy the inflection point criteria and yet there are still no growing wave modes. Turning the condition around we find that a sufficient condition for stability is that there are no sign changes for  $\beta - \partial_{yy}u_b$ . We summarize the result by stating the following theorem.

**RAYLEIGH-KUO INFLECTION POINT THEOREM:** Consider an inviscid and homogeneous (constant density) fluid, with flow in an inertial reference frame (no Coriolis) and with a base state of zonal flow with meridional shear. A necessary condition for shear instability is that there exists an inflection point in the base state zonal flow somewhere in the domain; i.e., where  $\partial_{yy}u_b = 0$  and so where the relative vorticity has an extrema,  $\partial_y\zeta_b = -\partial_{yy}u_b = 0$ . For flow on the  $\beta$ -plane, this criteria is generalized to  $\partial_y(f + \zeta_b) = \beta - \partial_{yy}u_b = 0$ , in which case the absolute vorticity must have an extrema in the domain in order to admit an instability. If there is no inflection point, then its absence is sufficient to conclude that the flow is stable to shear instability.

It is notable that  $\beta > 0$  always acts to stabilize the flow since its contribution requires a stronger background flow curvature to realize an inflection point. So if  $\beta$  is large enough then it can eliminate the inflection point,  $\partial_y(f + \zeta_b) = \beta - \partial_{yy}u_b = 0$ , from the domain, and in so doing it can stabilize the flow according to the Rayleigh-Kuo theorem. We can understand this stabilizing effect by noting that  $\beta$  supports planetary Rossby waves (Section 6.3), with such waves offering an alternative means to discharge the kinetic energy carried by shear in the base state flow.

### 13.5.2 Fjørtoft's theorem

Roughly 70 years after [Rayleigh \(1880\)](#) introduced the inflection point theorem, and a year after [Kuo \(1949\)](#) extended the inflection point theorem to the  $\beta$ -plane, [Fjørtoft \(1950\)](#) established another necessary condition for instability that is somewhat more constraining than the Rayleigh-Kuo theorem from Section 13.5.1.

To derive Fjørtoft's theorem, return to equation (13.58). Rather than focus on the imaginary part as done for the Rayleigh criteria, consider the real part

$$\int_{\mathcal{R}} \frac{(u_b - c_r)(\beta - \partial_{yy}u_b)|\tilde{\psi}|^2}{|u_b - c|^2} dy = \int_{\mathcal{R}} [|\partial_y\tilde{\psi}|^2 + k^2|\tilde{\psi}|^2] dy > 0. \quad (13.61)$$

---

<sup>6</sup>Recall from Section 13.6.1 that if  $c$  solves the Rayleigh equation with streamfunction  $\tilde{\psi}$ , then  $c^*$  also satisfies the equation with streamfunction  $\psi^*$ . Hence, for each decaying mode there is a growing mode. So a sufficient condition for instability is to find a wave in which  $c_i \neq 0$ .

We are interested in profiles that satisfy the Rayleigh criteria for instability, so that the integral (13.59) vanishes, in which case we have

$$\int_{\mathcal{R}} \frac{c_r (\beta - \partial_{yy} u_b) |\tilde{\psi}|^2}{|u_b - c|^2} dy = 0, \quad (13.62)$$

since  $c_r$  is a constant. Hence, equation (13.61) is trivially satisfied with any constant,  $U_s$ , inserted into the integral

$$\int_{\mathcal{R}} \frac{(u_b - U_s) (\beta - \partial_{yy} u_b) |\tilde{\psi}|^2}{|u_b - c|^2} dy > 0. \quad (13.63)$$

A particularly useful constant is the value of the zonal velocity at the inflection point,  $y = y_s$ , where the absolute vorticity has an extrema  $\partial_y(f + \zeta_b) = \beta - \partial_{yy} u_b = 0$ . We are thus led to the following theorem.

**FJØRTOFT'S THEOREM:** Under the same assumptions as the Rayleigh-Kuo theorem (Section 13.5.1), a necessary condition for shear instability is that  $(u_b - U_s) (\beta - \partial_{yy} u_b) > 0$  occurs somewhere in the domain in order to satisfy the condition (13.63). Here, the inflection point is determined by  $\partial_y(f + \zeta_b) = \beta - \partial_{yy} u_b(y_s) = 0$ , with  $U_s = u_b(y_s)$  the velocity at the inflection point.

Fjørtoft's theorem is rather subtle in its meaning, in particular it implies that an instability can occur only if the absolute vorticity has its maximum magnitude within the domain interior rather than at the domain boundary. One means to support this conclusion is by considering example zonal velocity profiles in Section 13.5.4.

### 13.5.3 Critical latitude theorem

The Rayleigh-Kuo condition and Fjørtoft's condition are statements about the geometry of the base flow state. Here we derive a condition necessary for a wave perturbation to support an instability. To do so, consider the Rayleigh-Kuo equation (13.43) written in terms of the meridional displacement

$$\partial_y [(u_b - c)^2 \partial_y \tilde{\xi}] = (u_b - c) [-\beta + k^2 (u_b - c)] \tilde{\xi}. \quad (13.64)$$

Following the approach taken for the Rayleigh-Kuo theorem in Section 13.5.1, multiply by  $\tilde{\xi}^*$ , integrate over the domain, and assume boundary contributions are zero so that

$$\int_{\mathcal{R}} (|\partial_y \xi|^2 + k^2 |\xi|^2) (u_b - c)^2 dy = \beta \int_{\mathcal{R}} (u_b - c) |\xi|^2 dy. \quad (13.65)$$

Writing

$$(u_b - c)^2 = (u_b - c_r)^2 - (c_i)^2 - 2 i c_i (u_b - c_r), \quad (13.66)$$

leads to the two conditions taken from the real and imaginary parts of equation (13.65)

$$\int_{\mathcal{R}} (|\partial_y \xi|^2 + k^2 |\xi|^2) [(u_b - c_r)^2 - (c_i)^2] dy = \beta \int_{\mathcal{R}} (u_b - c_r) |\xi|^2 dy \quad (13.67a)$$

$$2 c_i \int_{\mathcal{R}} (|\partial_y \xi|^2 + k^2 |\xi|^2) (u_b - c_r) dy = c_i \beta \int_{\mathcal{R}} |\xi|^2 dy. \quad (13.67b)$$

We are interested in unstable flows, in which case  $c_i \neq 0$ . So the second condition provides a statement about the value of the real phase velocity relative to the background flow

$$2 \int_{\mathcal{R}} (|\partial_y \xi|^2 + k^2 |\xi|^2) (u_b - c_r) dy = \beta \int_{\mathcal{R}} |\xi|^2 dy \geq 0. \quad (13.68)$$

Since the first portion of the left hand side integral is positive, the integral condition (13.68) can only be met if  $u_b - c_r$  is predominantly positive throughout the domain. For the special case of  $\beta = 0$ , we find the more definitive statement

$$\int_{\mathcal{R}} (|\partial_y \xi|^2 + k^2 |\xi|^2) (u_b - c_r) dy = 0 \quad \text{if } \beta = 0. \quad (13.69)$$

Evidently, when  $\beta = 0$  then  $u_b - c_r$  must change sign within the domain in order to support an instability. That is, for an instability to exist we must have the real phase velocity of a modal perturbation equal the background velocity at some latitude within the domain. For a wave perturbation to grow requires it to travel with the flow at least at one latitude, where the wave perturbation is stationary relative to the background flow and can thus extract kinetic energy to feed the growing wave. We are thus led to the *critical latitude theorem*.

**CRITICAL LATITUDE THEOREM:** Under the same assumptions as the Rayleigh-Kuo theorem (Section 13.5.1), a modal perturbation to a sheared flow with  $\beta = 0$  is able to initiate an instability if its real phase velocity equals to the background flow velocity somewhere in the domain, thus ensuring that the real phase velocity is within the range of the background velocity.

We return to this theorem in Section 13.7.6, where we find that this result also holds for stratified shear instability as represented by Howard's semi-circle theorem.

### 13.5.4 Stability conditions for sample profiles

We here consider a suite of example velocity profiles and discuss their stability properties as per the Rayleigh inflection point theorem and Fjørtoft's theorem, with consideration given only to the case of  $\beta = 0$ . We leave the velocity,  $U$ , and length  $L$ , scales arbitrary, noting that the stability theorems of Rayleigh and Fjørtoft are statements about the flow geometry rather than the scale of the flow.

#### Parabolic profile is stable

In Figure 13.2 we display a parabolic velocity profile, sometimes referred to as *Poiseuille flow*, which can be considered a highly smoothed version of the point jet studied in Section 6.5. Recall the point jet supports stable edge waves. In the absence of multiple point jets that can interact with one another, we expect the flow to be stable. Hence, extrapolating from the point jet motivates us to guess that the parabolic jet in Figure 13.2 is stable. Indeed, it is stable according to the Rayleigh inflection point theorem simply because there is no inflection point given that the derivative of the vorticity is constant.

#### Sinusoidal profile is unstable according Rayleigh and Fjørtoft

Figure 13.3 shows a sinusoidal velocity profile. This flow has an inflection point at  $y = 0$ , and so satisfies the Rayleigh inflection point condition for instability. It also satisfies the Fjørtoft

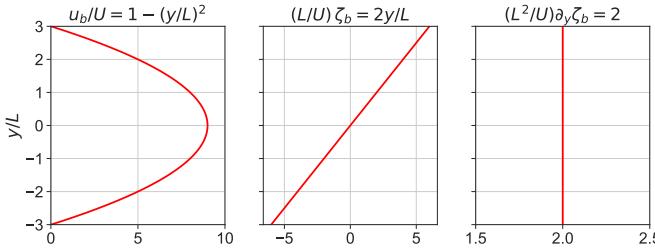


FIGURE 13.2: A parabolic velocity profile (Poiseuille flow) and its derivatives. Left panel:  $u_b/U = (y/L)^2$ . Middle panel: vorticity,  $(L/U) \zeta_b = -(L/U) \partial_y u_b = -2(y/L)$ . Right panel: derivative of the vorticity,  $(L^2/U) \partial_y \zeta_b = -(L^2/U) \partial_{yy} u_b = -2$ . Rayleigh's inflection point theorem says that this profile is stable since there is no inflection point.

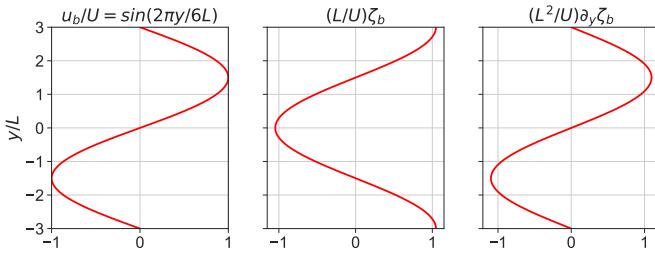


FIGURE 13.3: A sinusoidal velocity profile and its derivatives. Left panel:  $u_b/U = \sin(2\pi y/6L)$ . Middle panel: vorticity,  $(L/U) \zeta_b = -(L/U) \partial_y u_b = -(2\pi/6) \cos(2\pi y/6L)$ . Right panel: derivative of the vorticity,  $(L^2/U) \partial_y \zeta_b = -(2\pi/6)^2 u_b/U$ . This velocity profile satisfies both the Rayleigh and Fjørtoft criteria for instability.

condition for instability, which we see with  $U_s = 0$  at the  $y = 0$  inflection point so that

$$(u_b - U_s)(\beta - \partial_{yy} u_b) = (u_b - 0)(0 - \partial_{yy} u_b) = (2\pi y/6L)^2 (u_b)^2 > 0. \quad (13.70)$$

### Gaussian jet profile is unstable according to Rayleigh and Fjørtoft

Figure 13.4 shows a Gaussian jet profile. There are two inflection points,  $y = \pm L$ , so that this profile is unstable according to the Rayleigh inflection point theorem. It is also unstable due to Fjørtoft's theorem since the vorticity extrema are within the domain.

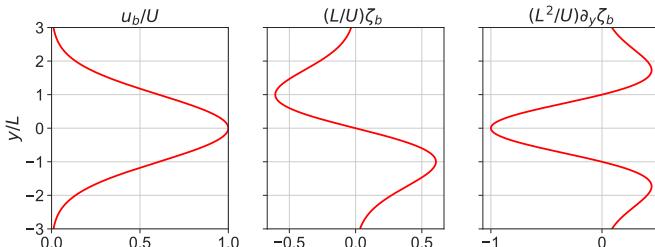


FIGURE 13.4: A Gaussian velocity profile and its derivatives. Left panel:  $u_b/U = e^{-(y/L)^2/2}$ . Middle panel: vorticity,  $(L/U) \zeta_b = -(L/U) \partial_y u_b = -(y/L) u_b/U$ . Right panel: derivative of the vorticity,  $(L^2/U) \partial_y \zeta_b = -(L^2/U) \partial_{yy} u_b = [(y/L)^2 - 1] (u_b/U)$ . This velocity profile is unstable according to the Rayleigh inflection point theorem, with the two inflection points at  $y = \pm L$ .

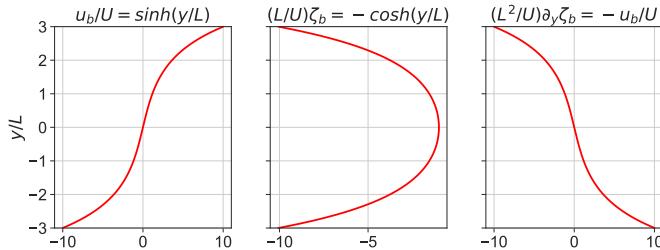


FIGURE 13.5: A sinh velocity profile and its derivatives. Left panel:  $u_b/U = \sinh(y/L)$ . Middle panel: vorticity,  $(L/U)\zeta_b = -(L/U)\partial_y u_b = -\cosh(y/L)$ . Right panel: derivative of the vorticity,  $(L^2/U)\partial_y \zeta_b = -(L^2/U)\partial_{yy}u_b = -\sinh(y/L)$ . This velocity profile satisfies the Rayleigh inflection point theorem, suggesting that it is unstable to shear instability. However, Fjørtoft's theorem says that it is stable.

### Sinh profile is unstable according to Rayleigh but stable according to Fjørtoft

Consider the sinh velocity profile,

$$u_b = U \sinh(y/L) \quad (13.71a)$$

$$\zeta_b = -\partial_y u_b = -(U/L) \cosh(y/L) \quad (13.71b)$$

$$\partial_y \zeta_b = -\partial_{yy} u_b = -(U/L^2) \sinh(y/L), \quad (13.71c)$$

as shown in Figure 13.5. The velocity vanishes at the  $y = 0$  inflection point (where  $\partial_y \zeta_b = -\partial_{yy} u_b = 0$ ), so that  $U_s = 0$ . Consequently, the velocity profile satisfies Rayleigh's necessary condition for instability. However, for Fjørtoft's theorem we note that

$$(u_b - U_s)(\beta - \partial_{yy} u_b) = (u_b - 0)(0 - \partial_{yy} u_b) = -(U/L)^2 \sinh^2(y/L) < 0. \quad (13.72)$$

Evidently, the flow is stable via the Fjørtoft theorem since  $(u_b - U_s)(\beta - \partial_{yy} u_b)$  is never positive in the domain. In particular, the vorticity extrema are at the latitudinal bounds of the domain (see middle panel of Figure 13.5), with the extrema increasing in magnitude as the domain is expanded. In Exercise 13.2 we consider  $\beta > 0$ , where we find that the sinh profile is also stable according to Fjørtoft.

## 13.6 Interacting edge waves and shear instability

So far in this chapter we developed sufficient conditions for sheared flows to be stable. Conversely, we developed necessary conditions for the sheared flow to be unstable. In this section we study a particular flow configuration and solve the Rayleigh equation to determine the dispersion relation and corresponding instability conditions. This case study exposes the interacting wave mechanism for shear instability. For this purpose, recall from Section 6.5 the theory for edge waves riding on a single jump in the base state vorticity, with salient features given by the following.

- Edge waves propagate along the interface where vorticity experiences a jump, with the waves trapped to this interface due to the exponential decay of the wave in the direction orthogonal to the jump surface.
- The edge wave phase velocity,  $c_p$ , is built from two contributions: one due to a Doppler shift from the background flow, and the other from the vorticity jump. The vorticity jump contribution is referred to as the intrinsic phase velocity,  $c_{\text{phase}}^{\text{intrinsic}}$ .

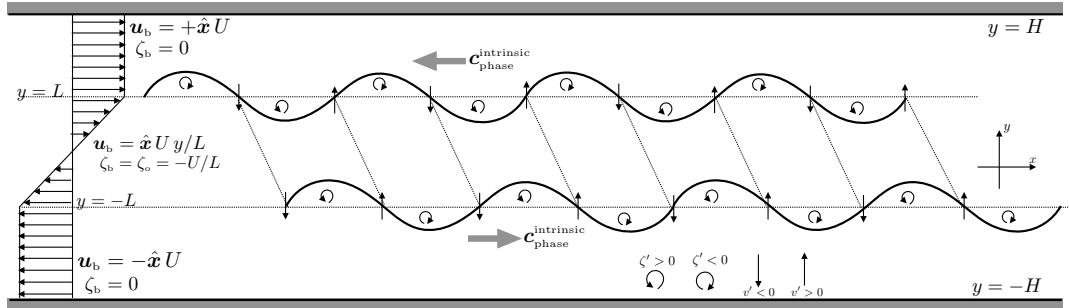


FIGURE 13.6: Depicting the interaction between two edge waves generated by a base flow that is unstable to shear instability. On the left we show the zonal base flow, with velocity kinks at  $y = \pm L$ . At  $y = -L$ , the vorticity jumps from zero to the south to a negative value to the north, so that the higher vorticity is to the south. This vorticity jump supports an edge wave with eastward intrinsic phase velocity,  $c_{\text{phase}}^{\text{intrinsic}}$ . For the  $y = L$  kink, the vorticity jumps from negative to zero when moving north. This vorticity configuration has higher vorticity to the north, thus supporting an edge wave with westward intrinsic phase velocity. These phase directions, both of which are contrary to the background flow, allow for Doppler shifting to make each of the edge waves stationary (i.e., zero real phase velocity). The oriented circles denote the anomalous vorticity associated with a particular fluctuation of the respective interfaces. For example, on the southern edge wave, a northward fluctuation brings fluid with higher relative vorticity northward, thus denoted by  $\zeta' > 0$  and a counter-clockwise oriented circle. The secondary circulations induced by the anomalous vorticity induce anomalous meridional flows, depicted by the meridional arrows. These meridional flows have maximum amplitude at the wave nodes, pointing upward in front of a wave peak and downward on the back side. Interactions between the waves are mutually constructive with a phase shift whereby the southern wave is shifted to the east of the northern wave (see Figure 13.1); i.e., with the  $y = L$  wave shifted ahead of the  $y = -L$  wave so that lines of constant phase are slanted into the background shear. With this phase shift, a northward anomalous flow from the southern wave enhances the peak of the northern wave, thus supporting its further growth. A similar enhancement occurs for the opposite wave, thus resulting in mutual amplification of each wave's amplitude. To optimize the growth requires the waves to be *phase locked*, whereby they stay stationary relative to one another, thus allowing the positive feedback to create the exponential wave growth.

- The intrinsic phase velocity is directed so that the higher relative vorticity is to the right when facing in the direction of the phase velocity. Equivalently, the phase velocity is directed towards the concave portion of the velocity profile. One can understand this orientation via the conservation of relative vorticity (holding for an inviscid flow in an inertial reference frame), as explained for the point jet in Figure 6.8.

We focus in this section on interactions between two edge waves, with the aim to characterize conditions that support shear instability. For analytical tractability, we focus on the case with zero planetary beta ( $\beta = 0$ ). This assumption serves to clearly expose the underlying *wave resonance* instability mechanism that arises from mutually reinforcing edge wave interactions. Additionally, it is convenient to set  $H \rightarrow \infty$  so that there are no solid boundaries, thus studying stability of a *free shear layer*.

Figure 13.6 illustrates the wave resonance mechanism, with the figure caption detailing the basic ingredients. In brief, the two edge waves, with oppositely directed intrinsic phase velocities, are *phase locked* due to Doppler shifting by the background flow. Phase locking means that the interacting waves move with the same phase velocity. Indeed, symmetry of the setup in Figure 13.6 leads to phase locked waves that are stationary, meaning that the real part of the phase velocity vanishes. When the waves are phase locked with a phase shift so that lines of constant phase are slanted into the background shear, then that orientation supports a mutual reinforcement of the opposing wave's amplitude, which is a signature of a modal instability. The goal of this section is to expose mathematical details supporting this figure.

### 13.6.1 Phase locked streamfunction and Rayleigh equation

Our mathematical task is to derive the streamfunction corresponding to the interacting edge waves supported by the velocity kinks at  $y = L$  and  $y = -L$  shown in Figure 13.6. Ingredients for this derivation were developed in Section 6.5 when studying a single edge wave along a single velocity jump. As for that analysis, start with the wave ansatz given by equation (13.30), written here again

$$\psi(x, y, t) = \tilde{\psi}(y) e^{ik(x - c_r t)} e^{kc_i t} = \tilde{\psi}(y) e^{ik(x - c_r t)} e^{\sigma t}. \quad (13.73)$$

As noted at the start of this section, writing a wave function with a single phase velocity is consistent with a phase locked system of two edge waves. With this wave ansatz we then solve the Rayleigh equation (13.35) (i.e., the vorticity equation) in the fluid region between the kinks, and use kinematic and dynamic boundary conditions to match the streamfunction across the kinks.

With the velocity profile in Figure 13.6, the meridional derivative of the background vorticity,  $\partial_y \zeta_b$ , has a Dirac delta at each of the vorticity jumps so that the Rayleigh equation (13.35) is

$$(u_b - c) (\partial_{yy} - k^2) \tilde{\psi} + \zeta_o [\delta(y + L) - \delta(y - L)] \tilde{\psi} = 0 \quad \text{with} \quad \zeta_o = -U/L. \quad (13.74)$$

The solution to this equation outside of the singular vorticity interfaces (at  $y = \pm L$ ) is given by

$$\tilde{\psi}(y) = \begin{cases} A_1 e^{-|k|(y-L)} & y \geq L \\ A_2 e^{|k|(y-L)} + A_3 e^{-|k|(y+L)} & -L \leq y \leq L \\ A_4 e^{|k|(y+L)} & y \leq -L, \end{cases} \quad (13.75)$$

where  $A_{1,2,3,4}$  are constants determined by the kinematic and dynamic boundary conditions applied (in the linear theory) at  $y = \pm L$ .

### 13.6.2 Kinematic boundary condition at $y = \pm L$

We derived the kinematic boundary condition in Section 6.5.4, which arises from the material nature of the interface. The velocity is continuous at  $y = \pm L$ , so that the kinematic boundary condition means that the streamfunction is also continuous at the kinks. Evaluating equation (13.75) at  $y = \pm L$  leads to the relations

$$A_1 = A_2 + A_3 e^{-2|k|L} \quad \text{and} \quad A_4 = A_3 + A_2 e^{-2|k|L}, \quad (13.76)$$

so that

$$\tilde{\psi}(y) = \begin{cases} A_{\text{up}} e^{-|k|(y-L)} + A_{\text{lo}} e^{-|k|(y+L)} & y \geq L \\ A_{\text{up}} e^{|k|(y-L)} + A_{\text{lo}} e^{-|k|(y+L)} & -L \leq y \leq L \\ A_{\text{up}} e^{|k|(y-L)} + A_{\text{lo}} e^{|k|(y+L)} & y \leq -L, \end{cases} \quad (13.77)$$

which can be written in the more succinct form

$$\tilde{\psi}(y) = A_{\text{up}} e^{-|k||y-L|} + A_{\text{lo}} e^{-|k||y+L|}, \quad (13.78)$$

where we wrote

$$A_{\text{up}} = A_2 \quad \text{and} \quad A_{\text{lo}} = A_3. \quad (13.79)$$

The notation (13.79) for the amplitudes correspond to the latitude ( $y = L$  or  $y = -L$ ) where their respective exponentials are maximized. Equation (13.78) reveals that the streamfunction

is the sum of two functions, each peaked at one of the kinks in the background flow.

The coefficients,  $A_{\text{up}}$  and  $A_{\text{lo}}$ , are generally complex. However, symmetry of the setup for when the edge waves are phase locked requires their magnitudes to be equal, in which case

$$A_{\text{up}} = \Gamma e^{i\theta_{\text{up}}} \quad \text{and} \quad A_{\text{lo}} = \Gamma e^{i\theta_{\text{lo}}} \quad \text{with} \quad \Gamma > 0, \quad (13.80)$$

and the corresponding streamfunction

$$\tilde{\psi}(y) = \Gamma \left[ e^{-|k||y-L|+i\theta_{\text{up}}} + e^{-|k||y+L|+i\theta_{\text{lo}}} \right]. \quad (13.81)$$

### 13.6.3 Dynamic boundary condition at $y = L$

As detailed in Section 6.5.5, the dynamic boundary condition is based on continuity of pressure across an interface (ignoring surface tension). Equivalently, it arises from integrating the Rayleigh equation (13.74) across each interface, with the Dirac delta rendering a jump condition. For this purpose, it is convenient to write the Rayleigh equation in the form of equation (6.85)

$$\partial_y[(u_b - c) \partial_y \tilde{\psi} - \tilde{\psi} \partial_y u_b] + [-k^2 (u_b - c)] \tilde{\psi} = 0. \quad (13.82)$$

Since  $u_b$  and  $\tilde{\psi}$  are everywhere continuous, integration across  $y = L$  yields

$$\lim_{\epsilon \rightarrow 0} \int_{L-\epsilon}^{L+\epsilon} \partial_y[(u_b - c) \partial_y \tilde{\psi} - \tilde{\psi} \partial_y u_b] dy = 0, \quad (13.83)$$

which leads to the jump condition

$$[(U - c) \partial_y \tilde{\psi} + \tilde{\psi} \zeta_o]_{y=L-\epsilon} = [(U - c) \partial_y \tilde{\psi}]_{y=L+\epsilon}, \quad (13.84)$$

where we noted that there is zero vorticity for  $y > L$ , and the zonal velocity equals to  $+U$  at  $y = L$ . Note that  $\epsilon$  is set to zero once evaluating the expressions. Making use of the streamfunction (13.77) and its derivative,

$$\partial_y \tilde{\psi} = |k| \begin{cases} -A_{\text{up}} e^{-|k|(y-L)} - A_{\text{lo}} e^{-|k|(y+L)} & y \geq L \\ A_{\text{up}} e^{|k|(y-L)} - A_{\text{lo}} e^{-|k|(y+L)} & -L \leq y \leq L \\ A_{\text{up}} e^{|k|(y-L)} + A_{\text{lo}} e^{|k|(y+L)} & y \leq -L, \end{cases} \quad (13.85)$$

brings the  $y = L$  jump boundary condition to

$$A_{\text{up}} [2(U - c)|k| + \zeta_o] + A_{\text{lo}} \zeta_o e^{-2|k|L} = 0, \quad (13.86)$$

which, with  $\zeta_o = -U/L$  (equation (13.74)), can be written as

$$A_{\text{up}} [2(1 - c/U)|k|L - 1] = A_{\text{lo}} e^{-2|k|L}. \quad (13.87)$$

Solving for the dimensionless phase velocity leads to

$$c/U = 1 - \frac{1}{2|k|L} \left[ 1 + \frac{A_{\text{lo}}}{A_{\text{up}}} e^{-2|k|L} \right] = 1 - \frac{1}{2|k|L} \left[ 1 + e^{-2|k|L-i\Delta\theta} \right], \quad (13.88)$$

where we introduced the phase for the amplitudes according to equation (13.80) and wrote the phase difference as

$$\Delta\theta = \theta_{\text{up}} - \theta_{\text{lo}}. \quad (13.89)$$

The phase velocity (13.88) is comprised of three terms

$$c = \underbrace{\frac{U}{2|k|L}}_{\text{Doppler}} - \underbrace{\frac{U}{2|k|L}}_{\text{free wave}} - \underbrace{\frac{U}{2|k|L} e^{-2|k|L-i\Delta\theta}}_{\text{interaction}}. \quad (13.90)$$

The first term arises from the Doppler shift via the background flow at the upper interface that is directed to the east,

$$\mathbf{u}_{\text{b}}(y = L) = U \hat{\mathbf{x}}. \quad (13.91)$$

The second term corresponds to the phase velocity of a free edge wave at the upper interface, with the westward intrinsic phase velocity

$$c_{\text{phase}}^{\text{intrinsic}} = -U/(2|k|L) \hat{\mathbf{x}} = \zeta_{\text{o}}/(2|k|) \hat{\mathbf{x}}. \quad (13.92)$$

For high wavenumbers (short wavelengths), the first and second terms dominate, with the phase velocity positive (eastward) as it is dominated by the Doppler term.

The final term in the phase velocity (13.90) arises from interactions between the upper and lower edge waves. The interaction decays both exponentially and algebraically according to  $|k|L$ , meaning that the edge waves have vanishingly small interactions for short wavelength zonal waves. Furthermore, the interaction term is generally a complex number that can be written

$$-\frac{U}{2|k|L} e^{-2|k|L-i\Delta\theta} = \frac{U e^{-2|k|L}}{2|k|L} [-\cos(\Delta\theta) + i \sin(\Delta\theta)] = c_r + i c_i. \quad (13.93)$$

Evidently, the growth rate,  $\sigma = c_i k$ , is positive (meaning the wave grows) if

$$k \sin(\Delta\theta) > 0 \implies \sigma = c_i k > 0 \implies \text{growing waves}. \quad (13.94)$$

For  $k > 0$  this condition means that the lines of constant phase are slanted into the background shear, just as anticipated by Figures 13.1 and 13.6.

#### 13.6.4 Dynamic boundary condition at $y = -L$

The interface at  $y = -L$  has the jump condition

$$[(-U - c) \partial_y \tilde{\psi}]_{y=-L-\epsilon} = [(-U - c) \partial_y \tilde{\psi} + \tilde{\psi} \zeta_{\text{o}}]_{y=-L+\epsilon}, \quad (13.95)$$

where the vorticity is zero for  $y < -L$  and the zonal velocity is  $-U$ . Making use of the streamfunction (13.77) and the derivative (13.85) brings the  $y = -L$  jump boundary condition to

$$A_{\text{up}} \zeta_{\text{o}} e^{-2|k|L} + A_{\text{lo}} [2(U + c)|k| + \zeta_{\text{o}}] = 0, \quad (13.96)$$

which can be written as

$$A_{\text{lo}} [2(1 + c/U)|k|L - 1] = A_{\text{up}} e^{-2|k|L}, \quad (13.97)$$

thus leading to the dimensionless phase velocity

$$c/U = -1 + \frac{1}{2|k|L} \left[ 1 + \frac{A_{\text{up}}}{A_{\text{lo}}} e^{-2|k|L} \right] = -1 + \frac{1}{2|k|L} \left[ 1 + e^{-2|k|L+i\Delta\theta} \right], \quad (13.98)$$

where we introduced the phase for the amplitudes according to equation (13.80). Just like for the  $y = L$  interface, we have the following interpretation for the phase velocity

$$c = \underbrace{-U}_{\text{Doppler}} + \underbrace{\frac{U}{2|k|L}}_{\text{free wave}} + \underbrace{\frac{U}{2|k|L} e^{-2|k|L+i\Delta\theta}}_{\text{interaction}} = c_r + i c_i. \quad (13.99)$$

At the  $y = -L$  interface the background flow is

$$\mathbf{u}_b(y = -L) = -U \hat{\mathbf{x}}, \quad (13.100)$$

so that the Doppler shift is in the opposite direction to that at the upper interface, as is the free edge wave intrinsic phase velocity

$$\mathbf{c}_{\text{phase}}^{\text{intrinsic}} = U/(2|k|L) \hat{\mathbf{x}} = -\zeta_0/(2|k|) \hat{\mathbf{x}}. \quad (13.101)$$

The interaction term in equation (13.99) shares much with that at the upper interface given by equation (13.90). Namely, both decay exponentially and algebraically according to  $|k|L$ . Furthermore, the interaction term is generally a complex number that can be written

$$\frac{U}{2|k|L} e^{-2|k|L+i\Delta\theta} = \frac{U e^{-2|k|L}}{2|k|L} [\cos(\Delta\theta) + i \sin(\Delta\theta)] = c_r + i c_i. \quad (13.102)$$

Just like for the upper interface, the growth rate,  $\sigma = c_i k$ , is positive (meaning the wave grows) if

$$k \Delta\theta > 0 \implies \sigma = c_i k > 0 \implies \text{growing waves.} \quad (13.103)$$

### 13.6.5 Phase velocity for phase locked edge waves

Application of the dynamical boundary condition at  $y = L$  and  $y = -L$  resulted in two expressions for the phase velocity as given by equations (13.88) and (13.98). The two phase velocities must be identical for the edge waves to be phase locked, with inspection of the two expressions leading to

$$c_r/U = 0 = \pm \left[ -1 + \frac{1}{2|k|L} + \frac{e^{-2|k|L} \cos \Delta\theta}{2|k|L} \right] \quad (13.104a)$$

$$c_i/U = \frac{\sin(\Delta\theta) e^{-2|k|L}}{2|k|L}. \quad (13.104b)$$

That is, the phase velocity for the phase locked edge waves is purely imaginary, so that the edge waves are stationary and have a streamfunction given by

$$\psi(x, y, t) = \tilde{\psi}(y) e^{ikx + k c_i t} = \tilde{\psi}(y) e^{ikx + \sigma t}. \quad (13.105)$$

### Growth rate

The phase velocity for the phase locked edge waves has a nonzero imaginary component when there is a nonzero phase shift between the two edge waves. Growth occurs when  $k c_i = \sigma > 0$ , which leads to the sufficient condition for instability

$$k \Delta\theta > 0 \implies \sigma > 0 \implies \text{unstable waves.} \quad (13.106)$$

### Vanishing real phase velocity

The vanishing real phase velocity puts a constraint on the zonal wavenumber, in which case

$$2|k|L = 1 + e^{-2|k|L} \cos \Delta\theta. \quad (13.107)$$

The maximum wavenumber that can satisfy this constraint is found when there is no phase shift ( $\Delta\theta = 0$ ), in which we define a critical wavenumber

$$(1 - 2|k|L) = e^{-2|k|L} \implies (|k|L)_{\text{crit}} = 0.6329. \quad (13.108)$$

This is a high wavenumber cutoff, meaning that for wavenumbers larger than  $(|k|L)_{\text{crit}}$  (i.e., relatively short waves), then there is a nonzero real component to the phase velocity. The nonzero real component to the phase velocity means that the two edge waves are in fact not phase locked, and as a result their amplitudes do not grow. Hence, the dimensionless wavenumber (13.108) separates the stable regime (high wavenumbers,  $|k|L > (|k|L)_{\text{crit}}$ ) from the unstable regime (low wavenumbers,  $|k|L < (|k|L)_{\text{crit}}$ ).

## 13.6.6 Dispersion relation and its interpretation

Here we derive the dispersion relation, which captures much of the information already gleaned by studying the phase velocity as well as some complementary perspectives.

### Deriving the dispersion relation

Combine the two jump conditions (13.86) and (13.96) into a matrix-vector equation

$$\begin{bmatrix} [2(U - c)|k| + \zeta_o] & \zeta_o e^{-2|k|L} \\ \zeta_o e^{-2|k|L} & [2(U + c)|k| + \zeta_o] \end{bmatrix} \begin{bmatrix} A_{\text{up}} \\ A_{\text{lo}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (13.109)$$

Nontrivial solutions to this equation exist if the determinant of the  $2 \times 2$  matrix vanishes, which then leads to the dispersion relation

$$(c/U)^2 = \frac{1}{(2|k|L)^2} [(1 - 2|k|L)^2 - e^{-4|k|L}] = \left[1 - \frac{1}{2|k|L}\right]^2 - \frac{e^{-4|k|L}}{(2|k|L)^2}. \quad (13.110)$$

### Interpreting the two roots to the dispersion relation

The dispersion relation (13.110) leads to two roots for the phase velocity,  $c = \omega/k$ , with roots either both real or both imaginary. We already saw in Section 13.6.5 that the real roots must vanish for phase locked edge waves, with a vanishing real phase velocity present when the horizontal wavenumber satisfies  $|k|L < (|k|L)_{\text{crit}}$  from equation (13.108). We interpret the nonzero real roots that arise with  $|k|L > (|k|L)_{\text{crit}}$  as stable and non-phase locked edge waves

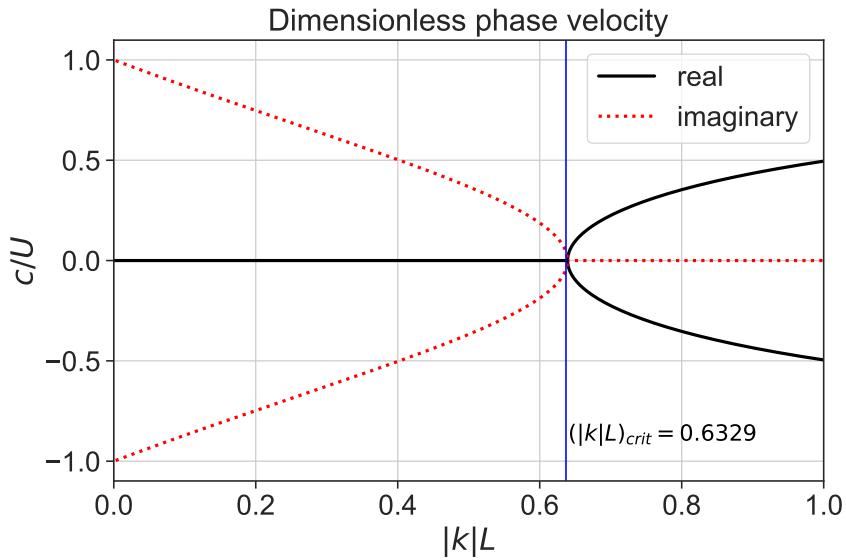


FIGURE 13.7: Dimensionless phase velocity,  $c/U = \omega/(kU)$ , from equation (13.110) derived for the piecewise linear velocity profile shown in Figure 13.6. A nonzero real phase velocity corresponds to stable edge waves that are not phase locked. If there are two edge waves, one on each of the velocity jumps and moving in the opposite directions, then they can interact. If the interaction is weak then it does not support a growing instability. Imaginary phase velocities correspond to phase locked edge waves with zero real phase velocity, so that the waves are stationary and have both an exponentially growing and decaying mode. The dimensionless wavenumber,  $|k|_{\text{crit}} L = 0.6329$ , separates the stable (high wavenumber) regime from the unstable (low wavenumber) regime.

moving in opposite directions along the velocity kinks at  $y = L$  and  $y = -L$ . For example, in the high wavenumber limit the roots are real and have the approximate value

$$c \approx \pm U \left[ 1 - \frac{1}{2|k|L} \right]. \quad (13.111)$$

In this limit, the two edge waves correspond to the isolated edge waves studied in Section 6.5. Evidently, the wavenumber is so large (small wavelength) that the meridional structure of the edge waves rapidly decays in the direction away from their respective interfaces. As a result the two edge waves are effectively free waves since their interaction is negligible and they are not phase locked. As  $|k|L$  gets smaller, the wavelength gets longer and the meridional extent of the two edge waves broadens. Interactions between the waves thus occurs, with these interactions possible even when the waves are stable. However, phase locking occurs for  $|k|L < (|k|L)_{\text{crit}}$ , with such waves unstable.

### Critical wavenumber separating the stable and unstable regimes

As seen in equation (13.108), the real part of the phase velocity vanishes when the dimensionless wavenumber satisfies

$$(1 - 2|k|L)^2 = e^{-4|k|L} \implies (|k|L)_{\text{crit}} = 0.6329. \quad (13.112)$$

That is, for lower dimensionless wavenumbers, the squared phase speed (13.110) becomes negative

$$(c/U)^2 < 0 \implies c = i c_i, \quad (13.113)$$

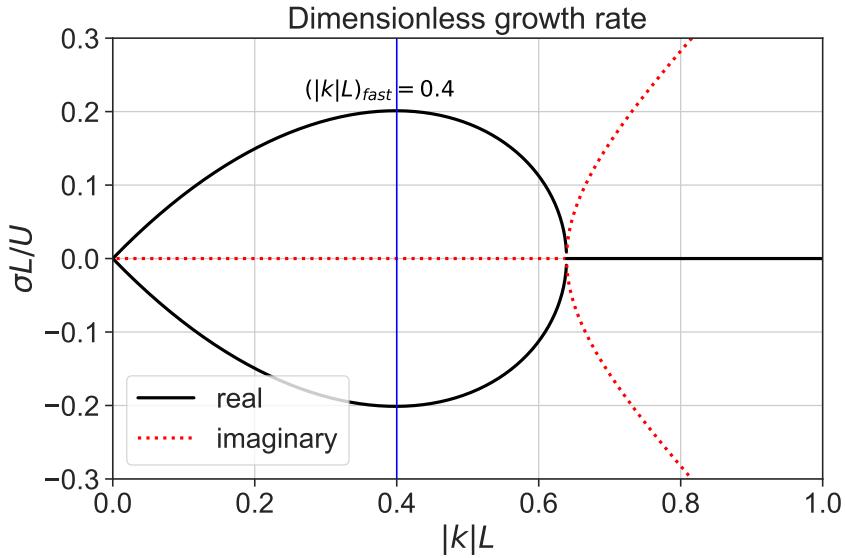


FIGURE 13.8: Dimensionless growth rate,  $\sigma L/U$ , from equation (13.114) derived for the piecewise linear velocity profile shown in Figure 13.6. A real growth rate corresponds to imaginary phase velocity from Figure 13.7, and represents exponentially growing ( $\sigma > 0$ ) and decaying ( $\sigma < 0$ ) modes. The dimensionless wavenumber,  $|k|_{\text{crit}} L = 0.6329$ , separates the stable (higher wavenumber) from unstable (lower wavenumber) regimes.

in which case the two real roots vanish and, for smaller wavenumbers, become two imaginary roots. One of the imaginary roots is an exponentially decaying mode and the other is the exponentially growing mode. Since the real part of the phase velocity vanishes, the modes are stationary relative to the shear layer, and thus they are either decaying or growing in place.

### Fastest growth rate

The growth rate,  $\sigma = c_i k$ , is given by

$$(\sigma L/U) = \pm \left[ e^{-4|k|L}/4 - (|k|L - 1/2)^2 \right]^{1/2}, \quad (13.114)$$

which we plot in Figure 13.8. The fastest growth rate is given by

$$\frac{d(\sigma L/U)^2}{d(|k|L)} = 0 \implies (|k|L)_{\text{fastest}} \approx 0.4 \implies \Lambda_{\text{fastest}} \approx 2\pi L/0.4 = 5\pi L. \quad (13.115)$$

Evidently, the most unstable mode has a zonal wavelength that is roughly 16 times the meridional width of the shear zone (see Figure 13.6). For this mode the positive dimensionless growth rate is

$$(\sigma L/U) \approx 0.2 \implies \sigma = 0.2 U/L. \quad (13.116)$$

The growth rate increases as the horizontal shear,  $U/L$ , increases, either by increasing the background flow speed,  $U$ , or by decreasing the width of the shear zone.

### Phase shift for the fastest growing unstable mode

Making use of equation (13.104b) renders the expression for the phase shift present for phase locked unstable edge waves

$$\sin(\Delta\theta) = \frac{2|k|L\sigma}{kU} e^{-2|k|L}. \quad (13.117)$$

Equation (13.116) says that the fastest growing mode has a growth rate of  $\sigma L/U \approx 0.2$  at the non-dimensional wavenumber  $|k|L \approx 0.4$  (equation (13.115)), so that

$$\sin(\Delta\theta) = \pm 2(\sigma L/U) e^{2|k|L} \approx \pm 0.4 e^{0.8}. \quad (13.118)$$

For an eastward intrinsic phase velocity, so that  $k > 0$ , the constraint (13.94) says that the growing wave has  $k \sin(\Delta\theta) > 0$ , with  $\sin(\Delta\theta) = 0.4 e^{0.8}$  yielding a phase shift of

$$\Delta\theta \approx 1.3\pi/2. \quad (13.119)$$

From Figure 13.6 we infer that the optimal alignment for growth of the two edge waves occurs with a  $\pi/2$  phase shift. However, this phase shift requires a larger  $|k|L$ , and a larger  $|k|L$  reduces the exponential appearing in the interaction between the two edge waves and thus slows the growth rate. Hence, the most unstable wave results from the dual need to optimize the relative phase of the two waves as well as the amplitude of the interaction.

### 13.6.7 Plotting the $\tilde{\psi}(y)$ streamfunctions

In Figure 13.9 we plot the streamfunction based on equation (13.78)

$$\tilde{\psi}(y) = A_{\text{up}} e^{-L|k||y/L-1|} + A_{\text{lo}} e^{-L|k||y/L+1|}, \quad (13.120)$$

for stable waves with  $kL = 1$  and  $kL = -1$ , and for an unstable wave with  $kL = 0.4$ . There are some subtleties with producing these plots, thus motivating the following presentation of details.

#### The stable streamfunction maximized at $y = L$

The stable streamfunctions have zero phase shifts (since they are stable waves) and they have unequal amplitudes,  $A_{\text{up}} \neq A_{\text{lo}}$ . The  $kL = -1$  wave is maximized on the  $y = L$  velocity kink since this is where the edge wave has a westward intrinsic phase velocity. The dimensionless phase velocity is given by the dispersion relation (13.110) with  $kL = -1$ , which yields (see also equation (13.90))

$$c/U \approx 0.495, \quad (13.121)$$

with the eastward phase velocity arising from dominance of the eastward background flow at  $y = L$ . That is, the stable edge wave maximized at  $y = L$  is swept eastward by the background flow, thus precluding it from being phase locked with the edge wave centered on  $y = -L$ . The amplitude ratio is computed from equation (13.87)

$$A_{\text{lo}}/A_{\text{up}} = e^2 [2(1 - c/U) - 1] \approx 0.068, \quad (13.122)$$

with the resulting streamfunction (13.120)

$$\tilde{\psi}^{(\text{L})}(y) = A_{\text{up}} \left[ e^{-|y/L-1|} + (A_{\text{lo}}/A_{\text{up}}) e^{-|y/L+1|} \right]. \quad (13.123)$$

We specify the amplitude,  $A_{\text{up}}$ , so that the maximum of  $\tilde{\psi}^{(\text{L})}(y)$  is unity as plotted in Figure 13.9.

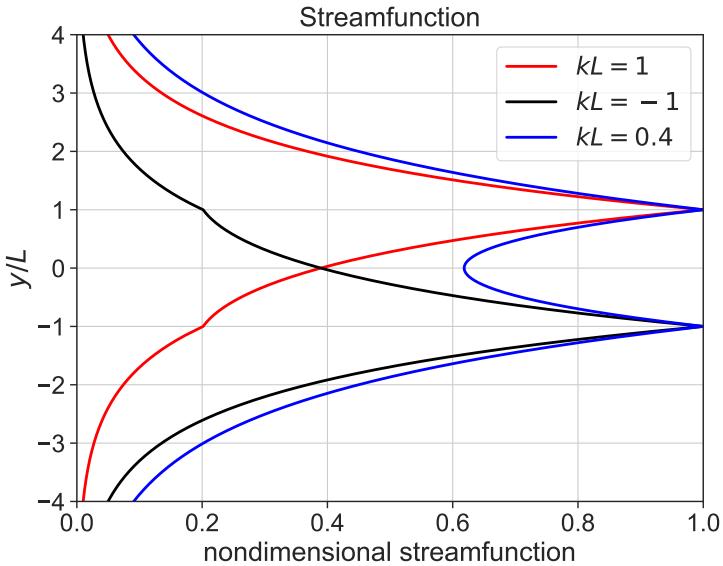


FIGURE 13.9: Streamfunction,  $\tilde{\psi}(y)$ , for stable and unstable edge waves on the free shear layer shown in Figure 13.6. Each streamfunction is non-dimensionalized with an amplitude set so that the maximum streamfunction is unity. The black and red streamfunctions are for stable and non-phase locked edge waves maximized at the upper velocity kink (red at  $y = L$ ) and lower kink (black at  $y = -L$ ). The two stable streamfunctions have zero phase shift ( $\theta_{\text{lo}} = \theta_{\text{up}} = 0$ ). The blue streamfunction is for the unstable mode with  $|k|L = 0.4$  and with the phase shift  $\Delta\theta = 1.3\pi/2$ , with the streamfunction given by equation (13.127). Details for how to compute these streamfunctions are given in Section 13.6.7.

### The stable streamfunction maximized at $y = -L$

The  $kL = 1$  wave is symmetric with respect to the  $kL = -1$  wave. Namely, the  $kL = 1$  wave is maximized on the  $y = -L$  velocity kink since this is where the edge wave has a westward intrinsic phase velocity. The dimensionless phase velocity is given by the dispersion relation (13.110) with  $kL = 1$ , which yields (see also equation (13.99))

$$c/U \approx -0.495, \quad (13.124)$$

with the westward phase velocity due to dominance of the westward background flow at  $y = -L$ . That is, the stable edge wave maximized at  $y = -L$  is swept westward by the background flow, thus precluding it from being phase locked with the edge wave centered on  $y = L$ . The amplitude ratio is computed from equation (13.97)

$$A_{\text{up}}/A_{\text{lo}} = e^2 [2(1 + c/U) - 1] \approx 0.068, \quad (13.125)$$

which is numerically equal to the ratio  $A_{\text{lo}}/A_{\text{up}}$  in equation (13.122) for the wave maximized at  $y = L$ . The resulting streamfunction (13.120) is

$$\tilde{\psi}^{(-L)}(y) = A_{\text{lo}} \left[ (A_{\text{up}}/A_{\text{lo}}) e^{-|y/L-1|} + e^{-|y/L+1|} \right]. \quad (13.126)$$

We specify the amplitude,  $A_{\text{lo}}$ , so that the maximum of  $\tilde{\psi}^{(-L)}(y)$  is unity as plotted in Figure 13.9.

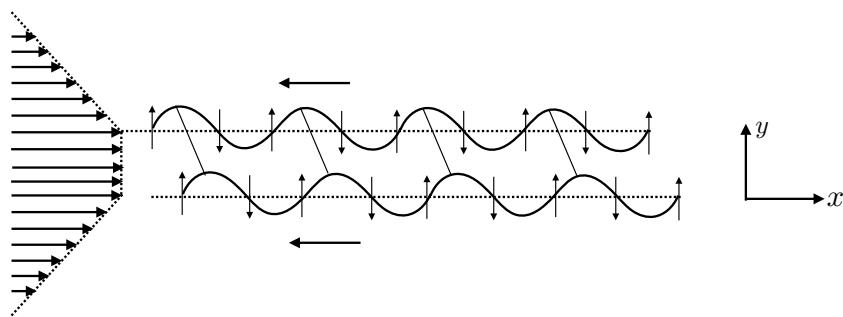


FIGURE 13.10: Two edge waves riding on their respective velocity jumps. There is no inflection point for this velocity profile, so that the velocity profile is stable according to the Rayleigh inflection point theorem from Section 13.5.1. We can mechanistically understand this result by noting that the two edge waves have a phase velocity in the same direction. Consequently, the waves cannot enter into the mutual resonance condition needed to create exponential growth. In this sketch, the secondary circulation from the lower wave enhances the amplitude of the upper wave. However, the upper wave diminishes the amplitude of the lower wave. Any other phase arrangement results in a similar situation whereby the two waves cannot mutually enhance each other's amplitude.

### The unstable streamfunction

When the zonal wavenumber gets smaller than  $(|k| L)_{\text{crit}} = 0.6329$ , then the two edge waves become phase locked and stationary (zero real phase velocity), with a phase shift allowing for mutual growth to manifest shear instability. In Figure 13.9 we plot the magnitude of the streamfunction for  $|k| L = 0.4$  according to equation (13.81)

$$\tilde{\psi}(y) = \Gamma e^{i\theta_{\text{up}}} \left[ e^{-L|k||y/L-1|} + e^{-L|k||y/L+1|+i\Delta\theta} \right], \quad (13.127)$$

where the phase shift is given by  $\Delta\theta = 1.3\pi/2$  as per equation (13.119), and the magnitude is computed by

$$|\tilde{\psi}| = \sqrt{\tilde{\psi}\tilde{\psi}^*}. \quad (13.128)$$

### 13.6.8 Lack of mutual wave growth for stable flows

The wave mechanism for shear instability offers a mechanistic understanding of the integral stability theorems from Section 13.5. For example, the case of Figure 13.6 illustrates how waves can mutually reinforce each other's amplitudes if there is an inflection point in the background flow, thus satisfying the necessary condition for Rayleigh's inflection point theorem. We sketch a profile in Figure 13.10 that has no inflection point, and for which the edge waves have their phase velocities in the same direction. If their relative phases are oriented so that the lower wave enhances the amplitude of the upper wave, as in Figure 13.10, then the upper wave in turn diminishes the amplitude of the lower. This result holds regardless the value for the relative phase. Whereas one wave's amplitude growth is supported by the other wave, that growth comes at the cost of diminishing the amplitude of its partner. We conclude that no matter what the relative phase relations, no mutual wave resonance can occur for the background flow profile of Figure 13.10 so there is no instability.

### 13.6.9 Further study

The linear sheared velocity profile in an unbounded domain was first analyzed by [Rayleigh \(1894\)](#) (volume II, page 393), and it forms the basis for most subsequent treatments, such as

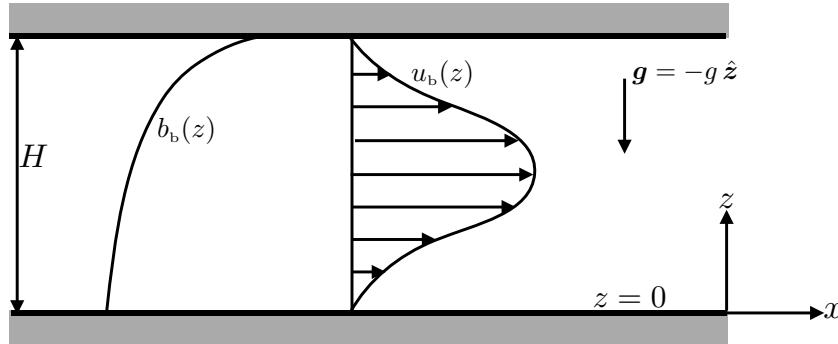


FIGURE 13.11: Schematic of the flow configuration used to examine stratified shear instability. The fluid is contained within a vertically bounded yet horizontally unbounded domain, with no-normal flow at the top and bottom meaning that the vertical flow vanishes at these boundaries,  $w(0) = w(H) = 0$ . The background flow is zonal with a vertical shear,  $u_b(z) \hat{x}$ . The background buoyancy,  $b_b(z)$ , is stably stratified with  $db_b/dz = N^2(z) > 0$ . The analysis in Section 13.7 reveals that the flow is stable if the gradient Richardson number,  $Ri = N^2/(du_b/dz)^2$ , is everywhere less than the critical value of 1/4. In contrast, the flow satisfies a necessary (though not sufficient) condition for shear instability if the Richardson number somewhere drops below 1/4.

[Chandrasekhar \(1961\)](#), [Drazin and Reid \(2004\)](#), [Vallis \(2017\)](#), and [Smyth and Carpenter \(2019\)](#). In particular, the treatment of [Smyth and Carpenter \(2019\)](#) offers many insights into the nature of the wave phases and phase speeds, with the presentation in this section emulating much from theirs. [This video from Prof. Mollo-Christensen](#) provides a pedagogical introduction to a variety of instabilities.

## 13.7 Integral conditions for stratified shear stability/instability

In this section we examine the stability of a vertically sheared zonal flow in the presence of gravity and with a gravitationally stable vertical buoyancy stratification. The fluid is contained within a vertically bounded domain yet with no boundaries in the horizontal, with Figure 13.11 providing an illustration of the configuration. Much in this discussion reflects that given for the meridionally sheared zonal flow in Section 13.5, here with the added feature of buoyancy stratification along with a slight increase in algebraic complexity.

### 13.7.1 Governing equations

We formulate the problem using the inviscid and adiabatic Boussinesq ocean equations in a rotating reference frame as developed in Section ??, with the governing equations given by

$$\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} + f \hat{z} \times \mathbf{v} = -\nabla \varphi + b \hat{z} \quad (13.129a)$$

$$(\partial_t + \mathbf{v} \cdot \nabla) b = 0 \quad (13.129b)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (13.129c)$$

where  $b$  is the buoyancy,  $\varphi$  is the pressure divided by the Boussinesq reference density, and  $\mathbf{v}$  is the non-divergent velocity.

The Coriolis acceleration generally plays a minor role in the mechanics of stratified shear instabilities, since the space-time scales are far too small for the Coriolis acceleration to play a noticeable role in the dynamics. We thus set  $f = 0$  in this section. In this case,

$$\mathbf{v} = u_b(z) \hat{x} + v_b(z) \hat{y}, \quad (13.130)$$

is an exact solution to the inviscid equations (13.129a)-(13.129c) if the background pressure has a zero horizontal gradient.<sup>7</sup> Evidently, the static and depth dependent horizontal flow (13.130) is an exact solution to the inviscid equations of motion in the absence of a Coriolis acceleration. It represents an unforced (free) mode for inviscid hydrostatic flow with pressure,  $\varphi(z)$ , and buoyancy  $b(z)$ , in exact hydrostatic balance.

With zero Coriolis acceleration, it is sufficient to orient the background flow along the  $\hat{x}$  axis so that we set  $v_b = 0$  in the following. We can further make use of Squire's theorem (Section 13.2) to note that the most unstable perturbation is two-dimensional and in the  $x$ - $z$  plane (the plane of the background flow). We thus focus on stability of two-dimensional flows in the form

$$\mathbf{v} = [u_b(z) + u'(x, z, t)] \hat{x} + w'(x, z, t) \hat{z}, \quad (13.131)$$

where  $u' \hat{x} + w' \hat{z}$  is the non-divergent perturbation flow in the  $x$ - $z$  plane

$$\partial_x u' + \partial_z w' = 0. \quad (13.132)$$

Given the above flow perturbation, we decompose the buoyancy and pressure fields according to

$$\varphi = \varphi_b(z) + \varphi'(x, z, t) \quad \text{and} \quad b = b_b(z) + b'(x, z, t), \quad (13.133)$$

where

$$\frac{d\varphi_b}{dz} = b_b \quad \text{and} \quad \frac{db_b}{dz} = N^2 > 0. \quad (13.134)$$

In this manner, the governing equations (13.129a)-(13.129c) take the form

$$\partial_t u' + (u_b + u') \partial_x u' + w' \partial_z (u_b + u') = -\partial_x \varphi' \quad (13.135a)$$

$$\partial_t w' + (u_b + u') \partial_x w' + w' \partial_z w' = -\partial_z \varphi' + b' \quad (13.135b)$$

$$\partial_t b' + (u_b + u') \partial_x b' + w' \partial_z b' = -w' N^2 \quad (13.135c)$$

$$\partial_x u' + \partial_z w' = 0. \quad (13.135d)$$

### 13.7.2 The linear vorticity equation

To develop an integral condition for stability, linearize the governing equations (13.135a)-(13.135d) (i.e., drop all terms with products of primed fields) to have

$$\partial_t u' + u_b \partial_x u' + w' \partial_z u_b = -\partial_x \varphi' \quad (13.136a)$$

$$\partial_t w' + u_b \partial_x w' = -\partial_z \varphi' + b' \quad (13.136b)$$

$$(\partial_t + u_b \partial_x) b' = -w' N^2 \quad (13.136c)$$

$$\partial_x u' + \partial_z w' = 0. \quad (13.136d)$$

Taking the  $z$ -derivative of the zonal equation (13.136a) and the  $x$ -derivative of the vertical equation (13.136b), and then subtracting, eliminates the pressure gradient to render

$$\partial_t (\partial_z u' - \partial_x w') + \partial_z (u_b \partial_x u') - u_b \partial_{xx} w' + \partial_z (w' \partial_z u_b) = -\partial_x b'. \quad (13.137)$$

A bit of rearrangement, and use of the continuity equation (13.136d), yields the linearized vorticity equation

$$(\partial_t + u_b \partial_x) (\partial_z u' - \partial_x w') + w' \partial_{zz} u_b = -\partial_x b', \quad (13.138)$$

<sup>7</sup>This point is noted in Section 21 of [Drazin and Reid \(2004\)](#).

with

$$\partial_z u' - \partial_x w' = \hat{\mathbf{y}} \cdot (\nabla \times \mathbf{v}') \quad (13.139)$$

the meridional component to the vorticity of the fluctuating flow.

Since the  $x$ - $z$  flow is non-divergent (equation (13.136d)), it is convenient to introduce a zonal-depth streamfunction

$$u' \hat{\mathbf{x}} + w' \hat{\mathbf{z}} = \hat{\mathbf{y}} \times \nabla \psi = \partial_z \psi \hat{\mathbf{x}} - \partial_x \psi \hat{\mathbf{z}}, \quad (13.140)$$

in which case the linear vorticity equation (13.138) becomes

$$(\partial_t + u_b \partial_x)(\partial_{xx} + \partial_{zz})\psi - \partial_x \psi \partial_{zz} u_b = -\partial_x b', \quad (13.141)$$

where the meridional component to the vorticity is given by the Laplacian of the streamfunction

$$\hat{\mathbf{y}} \cdot (\nabla \times \mathbf{v}') = (\partial_{xx} + \partial_{zz})\psi. \quad (13.142)$$

Likewise, the linear buoyancy equation (13.136c) becomes

$$(\partial_t + u_b \partial_x)b' = \partial_x \psi N^2. \quad (13.143)$$

Since the fluid is contained by a top and bottom boundary as per Figure 13.11, the no-normal flow condition for the vertical velocity,

$$w' = 0 \quad \text{at } z = 0, H, \quad (13.144)$$

means that the streamfunction is a spatial constant along the top and bottom boundaries

$$\partial_x \psi = 0 \quad \text{at } z = 0, H. \quad (13.145)$$

Without loss of generality we take these constant streamfunction values to be zero, thus rendering the homogeneous Dirichlet boundary conditions

$$\psi = 0 \quad \text{at } z = 0, H. \quad (13.146)$$

### 13.7.3 Taylor-Goldstein equation

The vorticity equation (13.141) and buoyancy equation (13.143) are two coupled linear partial differential equations with independent variables,  $(x, z, t)$ , and with  $z$ -dependent coefficients via  $d^2 u_b(z)/dz^2$  and  $N^2(z)$ . To study the stability of this flow, we pursue a modal analysis based on the wave ansatz

$$\psi(x, z, t) = \tilde{\psi}(z) e^{ik(x-ct)} \quad \text{and} \quad b'(x, z, t) = \tilde{b}(z) e^{ik(x-ct)}, \quad (13.147)$$

which accords with the wave ansatz (13.147) used to study stability of meridionally sheared flows in Section 13.4. In particular, the phase velocity,

$$\mathbf{c}_p = c \hat{\mathbf{x}} = (\omega/k) \hat{\mathbf{x}}, \quad (13.148)$$

follows the conventions in Section 13.4.1. Use of the ansatz (13.147) brings the vorticity equation (13.141) and buoyancy equation (13.143) to

$$(-\omega + u_b k)(-k^2 + \partial_{zz})\tilde{\psi} - k \tilde{\psi} \partial_{zz} u_b = -k \tilde{b} \quad (13.149a)$$

$$(-\omega + u_b k)\tilde{b} = -k \tilde{\psi} N^2. \quad (13.149b)$$

Substituting the buoyancy equation (13.149b) into the vorticity equation (13.149a) leads to the *Taylor-Goldstein equation* for the streamfunction

$$(u_b - c) \left[ \frac{d^2}{dz^2} - k^2 \right] \tilde{\psi} + \left[ \frac{N^2}{u_b - c} - \frac{d^2 u_b}{dz^2} \right] \tilde{\psi} = 0. \quad (13.150)$$

In the special case of  $N^2 = 0$  and with  $z$  swapped to  $y$ , the Taylor-Goldstein equation reduces to the Rayleigh equation (13.35).

### 13.7.4 Richardson number and the stability conditions

#### An inspired transformation of the Taylor-Goldstein equation

To develop conditions for stability/instability, we make the inspired transformation of the streamfunction to

$$\tilde{\psi} = \tilde{\phi} \sqrt{u_b - c} \quad \text{and} \quad \tilde{\phi}(0) = \tilde{\phi}(H) = 0. \quad (13.151)$$

A bit of algebra yields the second derivative

$$(u_b - c) \frac{d^2 \tilde{\psi}}{dz^2} = \left[ -\frac{1}{4(u_b - c)^{1/2}} \left( \frac{du_b}{dz} \right)^2 + \frac{(u_b - c)^{1/2}}{2} \frac{d^2 u_b}{dz^2} \right] \tilde{\phi} + (u_b - c)^{1/2} \frac{d}{dz} \left[ (u_b - c) \frac{d \tilde{\phi}}{dz} \right], \quad (13.152)$$

and the consequent transformation of the Taylor-Goldstein equation (13.150)

$$\left[ \frac{1}{u_b - c} \left[ N^2 - \frac{1}{4} \left( \frac{du_b}{dz} \right)^2 \right] - (u_b - c) k^2 - \frac{1}{2} \frac{d^2 u_b}{dz^2} \right] \tilde{\phi} + \frac{d}{dz} \left[ (u_b - c) \frac{d \tilde{\phi}}{dz} \right] = 0. \quad (13.153)$$

This form of the Taylor-Goldstein equation is actually a bit less compact than the original form (13.150). Even so, as we now show, it offers an elegant stability condition in terms of the gradient Richardson number.

#### Sufficient condition for stability and necessary condition for instability

To develop an integral stability theorem, multiply equation (13.153) by the complex conjugate,  $\tilde{\phi}^*$ , and integrate over the depth range  $0 \leq z \leq H$ . Performing this integral on the derivative term in equation (13.153) leads to

$$\int_0^H \tilde{\phi}^* \frac{d}{dz} \left[ (u_b - c) \frac{d \tilde{\phi}}{dz} \right] dz = \int_0^H \frac{d}{dz} \left[ \tilde{\phi}^* (u_b - c) \frac{d \tilde{\phi}}{dz} \right] dz - \int_0^H (u_b - c) \left| \frac{d \tilde{\phi}}{dz} \right|^2 dz, \quad (13.154)$$

with the total derivative term vanishing through use of the homogeneous Dirichlet boundary conditions in equation (13.151). Rearrangement thus renders

$$\int_0^H \frac{|\tilde{\phi}|^2}{u_b - c} \left[ N^2 - \frac{1}{4} \left( \frac{du_b}{dz} \right)^2 \right] dz = \int_0^H (u_b - c) \left[ k^2 |\tilde{\phi}|^2 + \left| \frac{d\tilde{\phi}}{dz} \right|^2 \right] dz + \frac{1}{2} \int_0^H \frac{d^2 u_b}{dz^2} |\tilde{\phi}|^2 dz. \quad (13.155)$$

This integral condition provides the basis for developing a sufficient condition for stability and, conversely, a necessary condition for instability.

The final term on the right hand side of equation (13.155) is real, so that the imaginary part of this equation is given by

$$c_i \int_0^H \frac{|\tilde{\phi}|^2 (du_b/dz)^2 (Ri - 1/4)}{|u_b - c|^2} dz = -c_i \int_0^H \left[ k^2 |\tilde{\phi}|^2 + \left| \frac{d\tilde{\phi}}{dz} \right|^2 \right] dz, \quad (13.156)$$

where we used the identity

$$\frac{1}{u_b - c} = \frac{u_b - c^*}{(u_b - c)(u_b - c^*)} = \frac{u_b - c_r + i c_i}{|u_b - c|^2}, \quad (13.157)$$

and introduced the *gradient Richardson number* (assuming nonzero vertical shear)

$$Ri = \frac{N^2}{(du_b/dz)^2}. \quad (13.158)$$

Evidently, if the Richardson number is greater than the critical value,

$$Ri_{crit} = 1/4, \quad (13.159)$$

throughout the vertical column, then the only way to satisfy equation (13.156) is for  $c_i = 0$ , which establishes a sufficient condition for stratified shear stability

$$Ri(z) > Ri_{crit} = 1/4 \quad \forall z \in [0, H] \implies \text{sufficient condition for stability.} \quad (13.160)$$

Conversely, a necessary condition for stratified shear instability is for the Richardson number to be less than 1/4 somewhere in the vertical column

$$Ri(z) < Ri_{crit} = 1/4 \quad \text{for some } z \in [0, H] \implies \text{necessary condition for instability.} \quad (13.161)$$

We emphasize that a Richardson number less than the critical value is necessary but not sufficient for stratified shear instability. Although for many purposes it is sufficient, there are examples where it is not. For example, flows near boundaries generally require a smaller Richardson number to go unstable. Even so, for most geophysical applications,  $Ri_{crit} = 1/4$  is a very good indicator for stratified shear instability.

### Comparison to barotropic shear instability

The gradient Richardson number (13.158) is the ratio of the buoyancy stratification to the vertical shear of the horizontal flow. It is notable that the necessary condition for instability,  $Ri < 1/4$ , does not depend on the curvature of the background flow, which contrasts to the case of barotropic shear instability studied in Section 13.5. For stratified shear instability, there is a direct struggle by the vertical shear to overcome the stabilizing effects from the background buoyancy stratification.

### 13.7.5 Richardson number and mixing energetics

The gradient Richardson number (13.158) provides a measure of the struggle between stabilizing effects from vertical stratification to the destabilizing effects from vertical shear. Here we provide an interpretation of the Richardson number in terms of mixing energetics, with mixing induced by the stratified shear instability. For this purpose we examine the thought experiment illustrated in Figure 13.12, with the analysis emulating that in Section 12.3.6 for mixing induced by Kelvin-Helmholtz instability.

#### Initial and final states of the thought experiment

Consider a Boussinesq ocean whose initial flow has a linear shear and a linear density profile

$$u_b(z) = U_0 + (z/H) \delta U \quad \text{and} \quad \rho_b(z) = \rho_0 - (z/H) \delta \rho, \quad (13.162)$$

where  $U_0$  is the velocity at  $z = 0$ ,  $\rho_0$  is the Boussinesq reference density,  $\delta \rho > 0$  is a constant that sets the strength of the vertical stratification, and  $\delta U$  is a constant that sets the strength of the vertical shear. The initial Richardson number is assumed to be less than the critical value of  $1/4$ ,

$$Ri = \frac{N^2}{(\partial_z u_b)^2} = \frac{g}{\rho_0} \frac{(\delta \rho / H)}{(\delta U / H)^2} < 1/4, \quad (13.163)$$

thus satisfying the necessary condition for stratified shear instability. For the final state, assume the fluid completely mixes to produce a uniform density and uniform velocity, which are taken to be the vertical average of the initial values

$$\bar{\rho} = \rho_0 - \delta \rho / 2 \quad \text{and} \quad \bar{U} = U_0 + \delta U / 2. \quad (13.164)$$

We might think of this configuration as a tiny region where the shear and density are well approximated with a linear vertical profile, and where shear induced mixing homogenizes the density and velocity.

#### Change in the potential and kinetic energies per area

The difference in the potential energy (per horizontal area) between the final and initial states is given by

$$P_{\text{final}} - P_{\text{init}} = g \int_0^H z [\bar{\rho} - \rho_b(z)] dz = g \delta \rho H^2 / 12. \quad (13.165)$$

An increase in potential energy upon mixing is anticipated by the study in Section 12.3.6 of Kelvin-Helmholtz induced mixing, and from the more general examination of potential energy in Section ???. The basic idea is that vertical mixing raises the center of mass of the fluid, thus increasing the gravitational potential energy.

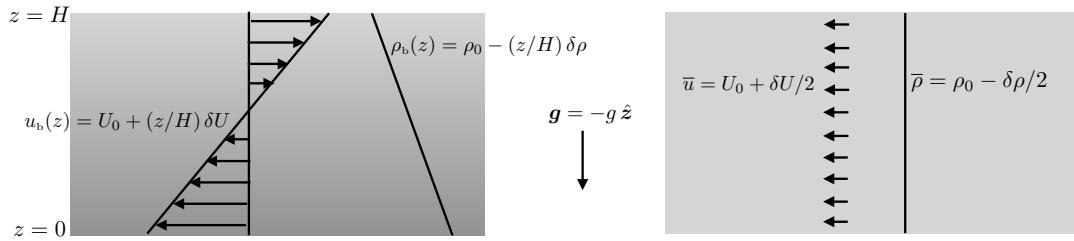


FIGURE 13.12: Left panel: initial conditions for a linearly stratified flow specified by a constant density parameter,  $\delta\rho > 0$ , that sets the strength of the linear stratification, and a constant shear parameter,  $\delta U$ , that sets the strength of the linear vertical shear. Right panel: a homogenized final state in which the density equals to the average of the initial density,  $\bar{\rho} = \rho_0 - \delta\rho/2$ , and the velocity equals to the average of the initial velocity,  $\bar{u} = U_0 + \delta U/2$ . The ratio of the potential energy change to the kinetic energy change is proportional to the gradient Richardson number as given by equation (13.167).

Where does the energy come from to raise the center of mass? For this thought experiment it must come from the kinetic energy, as that is the only other energy source. We thus anticipate that the kinetic energy decreases upon mixing. Indeed, the change in the kinetic energy (per horizontal area) is given by

$$K_{\text{final}} - K_{\text{init}} = \frac{\rho_0}{2} \int_0^H (\bar{u}^2 - u_b^2) dz = -\frac{\rho_0}{2} \int_0^H (u_b - \bar{u})^2 dz = -\rho_0 H (\delta U)^2 / 24. \quad (13.166)$$

A decrease in kinetic energy follows since the square of a spatially variable velocity is always greater than the square of its spatial average.

Taking the absolute ratio of the change in potential energy to the change in kinetic energy renders

$$\left| \frac{P_{\text{final}} - P_{\text{init}}}{K_{\text{final}} - K_{\text{init}}} \right| = \frac{2g}{\rho_0} \frac{\delta\rho/H}{(\delta U/H)^2} = 2 \text{Ri}, \quad (13.167)$$

where the second equality introduced the Richardson number (13.163) of the initial state. Evidently, the Richardson number in this thought experiment is directly proportional to the absolute ratio of the potential energy change upon mixing to the kinetic energy change upon mixing. In Exercise 13.3 we find a similar result for a two-layer system, yet with a different proportionality constant. These two thought experiments support the general connection between the Richardson number and the energy ratio. Namely, the numerator is proportional to the potential energy increase due to mixing (presenting a barrier to mixing), and the denominator is proportional to the kinetic energy available from the shear to overcome the potential energy barrier.

### 13.7.6 Constraining the phase velocity of unstable waves

The inspired transformation (13.151) is not the only one that can extract useful information about stability from the Taylor-Goldstein equation (13.150). Another transformation, following [Howard \(1961\)](#), is motivated by considering the vertical displacement of a fluid particle under small amplitude linear waves, with this approach following that given in Section 13.4.5, where we considered meridional displacements of fluid particles. The particle following approach motivates a transformation of the Taylor-Goldstein equation that then leads to constraints on the real and imaginary part of the phase velocity for unstable waves. That is, we derive conditions that must be satisfied for a wave to grow when placed in an unstable background flow. The development leads to the *critical height theorem* as well as *Howard's semi-circle*

theorem, which are generalizations of the critical latitude theorem from Section 13.5.3.

### Taylor-Goldstein equation in terms of a fluid particle's vertical displacement

Introduce the field,  $\eta(x, z, t)$ , which measures the vertical displacement of a fluid particle that has its rest vertical position at  $z$ . As such, this field satisfies

$$w' = [\partial_t + (u' + u_b) \partial_x] \eta \approx (\partial_t + u_b \partial_x) \eta, \quad (13.168)$$

with the linearization step following from the assumed small displacements. As defined,  $\eta(x, z, t)$  is the vertical displacement of a fluid particle that has its rest position at  $z$ . Now assume the displacements are generated by small amplitude waves of the form

$$\eta(x, z, t) = \tilde{\eta}(z) e^{ik(x-ct)}. \quad (13.169)$$

The corresponding wave ansatz for the vertical velocity,

$$w' = -\partial_x \psi = -i k \tilde{\psi}(z) e^{ik(x-ct)}, \quad (13.170)$$

leads to the relation between the streamfunction amplitude and the particle displacement amplitude

$$\tilde{\psi} = -(u_b - c)\tilde{\eta}. \quad (13.171)$$

The corresponding relation between second derivatives

$$-(u_b - c) \frac{d^2 \tilde{\psi}}{dz^2} = \frac{d}{dz} \left[ (u_b - c)^2 \frac{d\tilde{\eta}}{dz} \right] + \tilde{\eta} (u_b - c) \frac{d^2 u_b}{dz^2} \quad (13.172)$$

leads to the Taylor-Goldstein equation (13.150) written in terms of the vertical particle displacement

$$\frac{d}{dz} \left[ (u_b - c)^2 \frac{d\tilde{\eta}}{dz} \right] = [-N^2 + k^2 (u_b - c)^2] \tilde{\eta} \quad \text{with } \tilde{\eta}(0) = \tilde{\eta}(H) = 0. \quad (13.173)$$

Multiplying by  $\tilde{\eta}^*$ , then integrating over the vertical extent of the domain and using the homogeneous Dirichlet boundary conditions,  $\tilde{\eta}(0) = \tilde{\eta}(H) = 0$ , leads to the identity

$$\int_0^H N^2 |\tilde{\eta}|^2 dz = \int_0^H \left[ k^2 |\tilde{\eta}|^2 + \left| \frac{d\tilde{\eta}}{dz} \right|^2 \right] (u_b - c)^2 dz. \quad (13.174)$$

### The critical height theorem

The imaginary part of the integral condition (13.174) leads to the constraint

$$2c_i \int_0^H \left[ k^2 |\tilde{\eta}|^2 + \left| \frac{d\tilde{\eta}}{dz} \right|^2 \right] (u_b - c_r) dz = 0, \quad (13.175)$$

where we used

$$(u_b - c)^2 = (u_b - c_r)^2 - c_i^2 - 2ic_i(u_b - c_r). \quad (13.176)$$

The constraint (13.175) holds in the presence of an instability ( $c_i \neq 0$ ) only if the real part of the phase velocity lives within the bounds of the background flow

$$u_b^{\min} < c_r < u_b^{\max} \implies u_b(z_{\text{crit}}) = c_r. \quad (13.177)$$

That is, for a wave to grow in the presence of an unstable background flow, its phase velocity must have a real part within the bounds of the background flow. It follows that there is a vertical position,  $z = z_{\text{crit}}$ , where the background flow and the real part of the phase velocity are equal,

$$u_b(z_{\text{crit}}) = c_r. \quad (13.178)$$

This *critical height theorem* is the vertical analog to the critical latitude theorem from Section 13.5.3. Evidently, for a wave to grow, the real part of its phase velocity must match that of the background flow at no less than one vertical position. Otherwise, if  $c_r$  is outside of the bound (13.177), then the wave simply moves too fast for it to extract energy from the background flow.

### Howard's semi-circle theorem

The real part of the integral condition (13.174) leads to the constraint

$$\int_0^H \mathcal{Q} [(u_b - c_r)^2 - c_i^2] dz \geq 0, \quad (13.179)$$

where we introduced the shorthand for the non-negative quantity

$$\mathcal{Q} = k^2 |\tilde{\eta}|^2 + \left| \frac{d\tilde{\eta}}{dz} \right|^2 \geq 0. \quad (13.180)$$

For the case with  $c_i \neq 0$ , which means there is an unstable wave, then equation (13.175) implies

$$2 \int_0^H \mathcal{Q} (u_b - c_r) dz = 0. \quad (13.181)$$

We can thus add this term to the constraint (13.179) to render

$$\int_0^H \mathcal{Q} [(u_b - c_r)^2 - c_i^2] dz = \int_0^H \mathcal{Q} [(u_b - c_r)^2 - c_i^2 + 2 c_r (u_b - c_r)] dz \geq 0, \quad (13.182)$$

which then leads to

$$\int_0^H \mathcal{Q} u_b^2 dz \geq \int_0^H \mathcal{Q} (c_r^2 + c_i^2) dz \implies (c_r^2 + c_i^2) \leq \frac{\int_0^H \mathcal{Q} u_b^2 dz}{\int_0^H \mathcal{Q} dz}. \quad (13.183)$$

This inequality places a constraint on the real and imaginary part of the phase velocity. However, it is not so practical since we need to specify  $\mathcal{Q}$  from equation (13.180), which requires information about the wavenumber and particle displacements.

A more useful practical constraint can be derived by starting from the inequality

$$\int_0^H \mathcal{Q} (u_b - u_b^{\min}) (u_b - u_b^{\max}) dz = \int_0^H \mathcal{Q} [u_b^2 - u_b u_b^{\min} - u_b u_b^{\max} + u_b^{\min} u_b^{\max}] dz \leq 0, \quad (13.184)$$

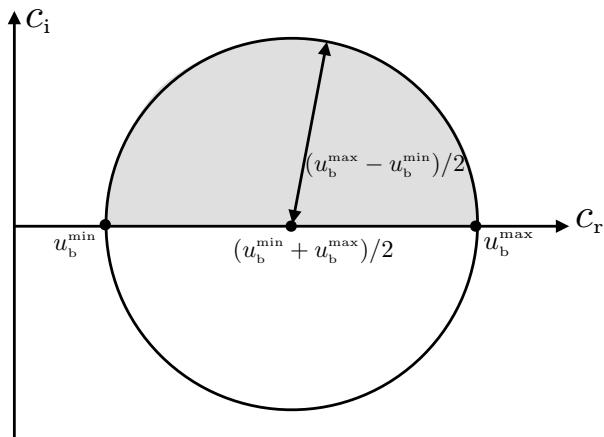


FIGURE 13.13: Illustrating Howard’s semi-circle theorem (13.188) in the complex plane defined by the phase velocity. The semi-circle theorem says that growing waves (with  $c_i > 0$ ) in a sheared flow have their phase velocity within the gray region, with radius  $(u_b^{\max} - u_b^{\min})/2$  and center along the real axis at  $c_r = (u_b^{\min} + u_b^{\max})/2$ .

which follows from the definition of the  $u_b^{\min}$  and  $u_b^{\max}$ , and recalling that  $\mathcal{Q} \geq 0$ . Making use of equation (13.183) allows us to replace  $u_b^2$  with  $c_r^2 + c_i^2$  and still maintain the inequality, so that

$$\int_0^H \mathcal{Q} [c_r^2 + c_i^2 - u_b (u_b^{\min} + u_b^{\max}) + u_b^{\min} u_b^{\max}] dz \leq 0, \quad (13.185)$$

and use of the constraint (13.181) then allows us to replace  $u_b$  with  $c_r$  to have

$$\int_0^H \mathcal{Q} [c_r^2 + c_i^2 - c_r (u_b^{\min} + u_b^{\max}) + u_b^{\min} u_b^{\max}] dz \leq 0. \quad (13.186)$$

This constraint is more practical since it only involves the phase velocity and the maximum and minimum values of the background velocity, each of which are independent of  $z$ . Evidently, since  $\mathcal{Q}$  is positive, we must have

$$c_r^2 + c_i^2 - c_r (u_b^{\min} + u_b^{\max}) + u_b^{\min} u_b^{\max} \leq 0, \quad (13.187)$$

which can be rearranged to

$$[c_r - (u_b^{\min} + u_b^{\max})/2]^2 + c_i^2 \leq [(u_b^{\max} - u_b^{\min})/2]^2. \quad (13.188)$$

This constraint is depicted in Figure 13.13, where we see that unstable waves (with  $c_i > 0$ ) have phase velocities that sit within the semi-circle in the upper half of the complex plane whose center is along the real axis with  $c_r = (u_b^{\min} + u_b^{\max})/2$  and whose radius is  $(u_b^{\max} - u_b^{\min})/2$ .

The semi-circle theorem is particularly useful when designing numerical algorithms to find unstable waves. Also note that we derived this theorem starting from the case of a stratified shear layer. However, the same result holds for unstable waves in the presence of a meridionally sheared barotropic flow discussed as earlier in this chapter.

### 13.7.7 Further study

The  $\text{Ri} = 1/4$  stability argument was first presented by [Miles \(1961\)](#) and soon thereafter it was extended by [Howard \(1961\)](#). Our derivation of Howard’s semi-circle theorem (13.188) follows Section 11.7 of [Kundu et al. \(2016\)](#).

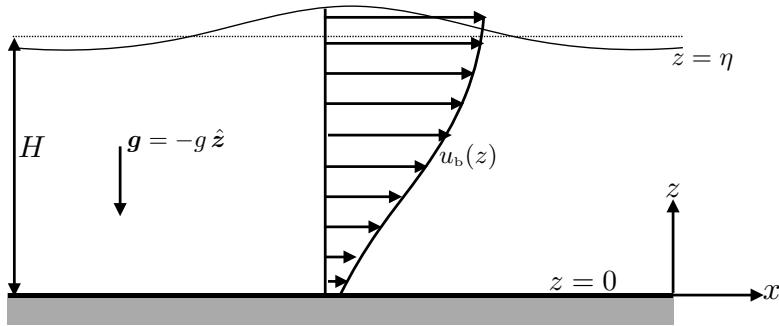


FIGURE 13.14: Schematic of the flow configuration used to examine vertical shear instability in a homogeneous inviscid flow with a free upper surface. The background flow is zonal with a vertical shear,  $u_b(z) \hat{x}$ . The fluid domain has a rigid lower bound at  $z = 0$ , where the bottom kinematic boundary condition means that  $w(0) = 0$ . The upper surface is free at  $z = \eta = H + \eta'(x, t)$ , so that the linearized upper kinematic boundary condition is  $(\partial_t + u_b \partial_x) \eta' = w'$ , and this boundary condition is evaluated at  $z = H$  rather than  $z = \eta$ .

## 13.8 A vertically sheared homogeneous fluid with a free surface

In this chapter we studied horizontal sheared flow in a homogeneous fluid, and then rotated the geometry to consider vertically sheared flow in a gravitationally stratified fluid. Here we consider a vertically sheared homogeneous fluid moving over a flat surface at  $z = 0$  and with a free surface at  $z = \eta$ , as depicted in Figure 13.14. Directly translating the analysis of stratified shear instability from Section 13.7 to the present case, we might presume that since the Richardson number vanishes here, then any flow satisfies the necessary condition for vertical shear instability. However, besides having a vanishing buoyancy (since the fluid is homogeneous), the free upper surface further distinguishes this configuration from the vertically stratified shear instability studied in Section 13.7 (where we assumed the flow to occur between two rigid boundaries). The presence of a free surface removes the ability to make general statements about the stability, thus illustrating the importance of boundary conditions when studying modal instabilities. This section also serves to highlight the role of non-hydrostatic pressure for enabling vertical shear instability.

### 13.8.1 Linearized governing equations

The linearized governing equations are given by equations (13.136a)-(13.136d) with identically zero buoyancy

$$\partial_t u' + u_b \partial_x u' + w' \partial_z u_b = -\partial_x \varphi' \quad (13.189a)$$

$$\partial_t w' + u_b \partial_x w' = -\partial_z \varphi' \quad (13.189b)$$

$$\partial_x u' + \partial_z w' = 0, \quad (13.189c)$$

along with the bottom kinematic boundary condition and the linearized kinematic free surface boundary condition

$$w' = 0 \quad \text{at } z = 0 \quad (13.190a)$$

$$(\partial_t + u_b \partial_x) \eta' = w' \quad \text{at } z = H. \quad (13.190b)$$

Since the  $x$ - $z$  flow is non-divergent we can introduce a streamfunction as in equation (13.140)

$$u' \hat{x} + w' \hat{z} = \hat{\mathbf{y}} \times \nabla \psi = \partial_z \psi \hat{x} - \partial_x \psi \hat{z}, \quad (13.191)$$

which yields the governing equations

$$(\partial_t + u_b \partial_x) \partial_z \psi - \partial_x \psi \partial_z u_b = -\partial_x \varphi' \quad \text{for } 0 < z < H \quad (13.192a)$$

$$(\partial_t + u_b \partial_x) \partial_x \psi = \partial_z \varphi' \quad \text{for } 0 < z < H \quad (13.192b)$$

$$\psi = 0 \quad \text{at } z = 0 \quad (13.192c)$$

$$(\partial_t + u_b \partial_x) \eta' = -\partial_x \psi \quad \text{at } z = H. \quad (13.192d)$$

### 13.8.2 Hydrostatic fluctuations are stable

If we assume the fluctuations are hydrostatic then the zonal pressure gradient is given by the zonal derivative of the free surface (as in the case of a hydrostatic shallow water fluid),

$$\partial_x \varphi' = g \partial_x \eta', \quad (13.193)$$

so that the zonal velocity equation (13.192a) becomes

$$(\partial_t + u_b \partial_x) \partial_z \psi - \partial_x \psi \partial_z u_b = -g \partial_x \eta'. \quad (13.194)$$

Now assume the small amplitude fluctuations of the streamfunction and free surface take the propagating plane wave form

$$\psi(x, z, t) = \tilde{\psi}(z) e^{ik(x-ct)} \quad \text{and} \quad \eta'(x, t) = \tilde{\eta} e^{ik(x-ct)}. \quad (13.195)$$

For these fluctuations the kinematic boundary conditions (13.192c) and (13.192d) become

$$\tilde{\psi} = 0 \quad \text{at } z = 0 \quad (13.196a)$$

$$(u_b - c)\tilde{\eta} = -\tilde{\psi} \quad \text{at } z = H. \quad (13.196b)$$

Similarly, the zonal velocity equation (13.194) renders the relations between the amplitude functions

$$(u_b - c)\partial_z \tilde{\psi} - \tilde{\psi} \partial_z u_b = -g \tilde{\eta} \quad \text{for } 0 < z < H. \quad (13.197)$$

Dividing by  $(u_b - c)^2$  brings about

$$\frac{d}{dz} \left[ \frac{\tilde{\psi}}{u_b - c} \right] = -\frac{g \tilde{\eta}}{(u_b - c)^2}, \quad (13.198)$$

whose vertical integral renders

$$\left. \frac{\tilde{\psi}}{u_b - c} \right|_{z=H} - \left. \frac{\tilde{\psi}}{u_b - c} \right|_{z=0} = -g \tilde{\eta} \int_0^H \frac{dz}{(u_b - c)^2}, \quad (13.199)$$

where we evaluated the upper integral limit at  $z = H$  as per the linear theory. Making use of the kinematic boundary conditions (13.196a) and (13.196b) brings this equation into the form

$$g \int_0^H \frac{dz}{(u_b - c)^2} = 1. \quad (13.200)$$

This identity can only be satisfied if the phase velocity,  $c$ , has zero imaginary part, in which case we conclude that the fluctuations are stable to vertical shear instability. In particular, for

$u_b$  a constant we recover the case considered in Exercise 7.4 for gravity waves moving on a constant background flow, in which the phase velocity is  $c = u_b \pm \sqrt{g H} = u_b \pm c_{\text{grav}}$ .

We understand why the hydrostatic fluctuations are stable by appealing to the discussion of shallow water dynamics in Section ???. In that section we observed that a homogeneous fluid layer with a hydrostatic pressure (i.e., a shallow water fluid layer) renders a horizontal pressure gradient that is depth independent, which in turn leads to a depth independent horizontal velocity. That is, the homogeneous hydrostatic fluid cannot support a vertical shear in the horizontal flow, so that there can be no vertical shear instability. That is, when studying stability of a vertically sheared homogeneous fluid layer, hydrostatic fluctuations are depth independent and so they cannot lead to a vertical shear instability.

### 13.8.3 Taylor-Goldstein equation

Accepting that unstable fluctuations must involve a non-hydrostatic pressure, we make use of the analysis in Section 13.7.3 for stratified shear instability, in which we formed the linearized vorticity equation and introduced the streamfunction to render the Taylor-Goldstein equation (13.150). Specializing Taylor-Goldstein to the case of zero buoyancy ( $N^2 = 0$ ) yields

$$(u_b - c) \left[ \frac{d^2}{dz^2} - k^2 \right] \tilde{\psi} = \frac{d^2 u_b}{dz^2} \tilde{\psi} \quad (13.201a)$$

$$\tilde{\psi} = 0 \quad \text{at } z = 0 \quad (13.201b)$$

$$(u_b - c)\tilde{\eta} = -\tilde{\psi} \quad \text{at } z = H. \quad (13.201c)$$

In deriving the necessary condition for vertical shear instability in Section 13.7.4, we made use of the assumed rigid top and bottom boundaries to eliminate the two boundary contributions appearing in equation (13.154). Yet with a free surface the  $z = H$  boundary contribution no longer vanishes, instead it satisfies the kinematic condition (13.201c).

### 13.8.4 Necessary conditions for instability

To develop an integral condition, rather than introduce the ansatz (13.151), we work directly with the Taylor-Goldstein equation (13.201a). Since  $N^2 = 0$ , the manipulations are quite similar to those used to derive the Rayleigh-Kuo inflection point theorem in Section 13.5.1. For this purpose, multiply the Taylor-Goldstein equation (13.151) by  $\tilde{\psi}^*$  and rearrange to yield

$$\frac{d}{dz} \left[ \tilde{\psi}^* \frac{d\tilde{\psi}}{dz} \right] = \left[ k^2 + \frac{1}{u_b - c} \frac{d^2 u_b}{dz^2} \right] |\tilde{\psi}|^2 + \left| \frac{d\tilde{\psi}}{dz} \right|^2 \quad (13.202)$$

Vertical integration and use of the kinematic boundary conditions leads to

$$\tilde{\psi}^* \frac{d\tilde{\psi}}{dz} \Big|_{z=H} = \int_0^H \left[ k^2 |\tilde{\psi}|^2 + \left| \frac{d\tilde{\psi}}{dz} \right|^2 \right] dz + \int_0^H \frac{d^2 u_b}{dz^2} \frac{|\tilde{\psi}|^2}{u_b - c} dz. \quad (13.203)$$

Taking the imaginary part of this equation and exposing the amplitude and phase for  $\tilde{\psi}$ ,

$$\tilde{\psi} = |\tilde{\psi}(z)| e^{i\alpha(z)}, \quad (13.204)$$

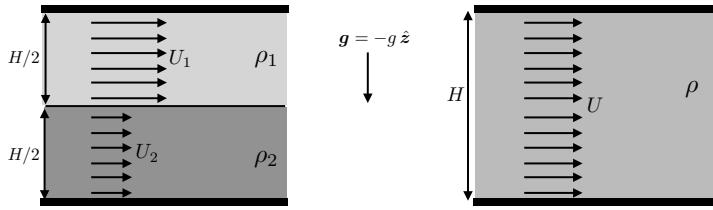


FIGURE 13.15: Initial (left) and final (right) conditions for a two-layer stratified flow that completely mixes its density and velocity. In Exercise 13.3 we work through the energetics of mixing and show that the ratio of the potential energy change to the kinetic energy change is proportional to a discrete version of the gradient Richardson number.

leads to

$$\left[ |\tilde{\psi}|^2 \frac{d\alpha}{dz} \right]_{z=H} = c_i \int_0^H \frac{d^2 u_b}{dz^2} \frac{|\tilde{\psi}|^2}{|u_b - c|^2} dz. \quad (13.205)$$

The left hand side vanishes for a rigid upper boundary, in which case we recover the Rayleigh-Kuo inflection point theorem (13.59). But with a free surface there are other situations that support instability since the left hand side is no longer zero. That is, allowing for a free surface opens up further avenues for instabilities. When developing necessary conditions for baroclinic instability in Section 14.8, we also find a fundamental role for boundary processes.



## 13.9 Exercises

### EXERCISE 13.1: COMPLEX CONJUGATE FORMULA

In equations (13.45c) and (13.48) we made use of the identity

$$\text{Im} [\tilde{\psi}^* \partial_y \tilde{\psi}] = -\text{Im} [\tilde{\psi} \partial_y \tilde{\psi}^*]. \quad (13.206)$$

Derive this identity. Hint: make use of equation (13.49) for the streamfunction.

### EXERCISE 13.2: FJØRTOFT'S THEOREM WITH SINH PROFILE AND $\beta \neq 0$

Consider the sinh velocity profile from Section 13.5.4. Show that with  $\beta \neq 0$  the flow remains stable according to Fjørtoft's theorem.

### EXERCISE 13.3: RICHARDSON NUMBER AND ENERGETICS OF MIXING

Rework the energetics from Section 13.7.5, only now with the two-layer stratified shear profile shown in Figure 13.15. In particular, compute the ratio of the potential energy increase to the kinetic energy decrease and show that this ratio is proportional to a discrete version of the gradient Richardson number

$$\frac{P_{\text{final}} - P_{\text{init}}}{K_{\text{init}} - K_{\text{final}}} \propto \text{Ri}. \quad (13.207)$$

Assume a Boussinesq ocean for purposes of computing the kinetic energy. Hint: the solution to this exercise is detailed in Section 14.1 of [Cushman-Roisin and Beckers \(2011\)](#).





## Chapter 14

# QUASI-GEOSTROPHIC WAVES AND BAROCLINIC INSTABILITY

In this chapter we study waves and *baroclinic instability* in a continuously stratified quasi-geostrophic flow within a vertically bounded domain. The vertically bounded domain is more relevant to the ocean, with the atmosphere top boundary more suitably assumed to be at infinity. However, beyond the ocean case, we are motivated to consider the vertically bounded domain since it is assumed for the *Eady model* of baroclinic instability ([Eady, 1949](#)), which is a particularly elegant example of baroclinic instability. We furthermore find it is convenient to study waves in combination with baroclinic instability since baroclinic instability follows a *wave resonance* mechanism, much like that considered for shear instability in Chapter 13.

The waves we encounter in this chapter include *planetary Rossby waves*, *topographic Rossby waves*, and *Eady edge waves*. Recall that we encountered Rossby waves and edge waves in the horizontally non-divergent barotropic model of Chapter 6, as well as Rossby waves and topographic waves for the shallow water model in Chapter 7. In the present chapter, we work fully within the quasi-geostrophic theory and allow for continuous stratification. In addition to Rossby waves, we here consider a thermal wind background state, in which the quasi-geostrophic flow supports Eady edge waves. As shown here, Eady edge waves rely on the presence of a buoyancy gradient along the top and/or bottom boundaries of the domain. Eady waves are the primary actors in the Eady model of baroclinic instability.

The study of baroclinic instability is motivated by asking whether a thermal wind flow is stable to small amplitude geostrophic perturbations. Under certain circumstances, there are unstable wave modes whose energy grows by feeding off the potential energy of the thermal wind state. The Eady model considers the constructive interference of edge waves that leads to wave resonance. This resonance then drives the mutual growth of the edge waves, which constitutes baroclinic instability. Baroclinic instability dominates the fluctuations of the large-scale (order deformation radius) flows in the middle latitude atmosphere (*synoptic scale*) and the *mesoscale* ocean.

### READER'S GUIDE FOR THIS CHAPTER

In this chapter we make use of the governing equations of continuously stratified quasi-geostrophic flow from Chapter ??, as well as the Rossby wave discussions in Chapter 6 (horizontally non-divergent barotropic model) and Section 7.9 (shallow water layer). The wave resonance interpretation of Eady's model of baroclinic instability closely follows that given for shear instability in Chapter 13.

<b>14.1</b>	<b>Loose threads</b>	<b>432</b>
<b>14.2</b>	<b>Equations of quasi-geostrophy</b>	<b>433</b>
14.2.1	Equations in the fluid interior	433
14.2.2	Boundary conditions for vertically bounded domain	434
<b>14.3</b>	<b>Linear fluctuations on a zonal geostrophic background flow</b>	<b>435</b>
14.3.1	Zonal geostrophic background flow	436
14.3.2	Background state is an exact quasi-geostrophic solution	436
14.3.3	Fluctuating streamfunction, potential vorticity, and buoyancy	437
14.3.4	Linearized upper boundary condition	438
14.3.5	Linearized bottom boundary condition	438
<b>14.4</b>	<b>Vertically bounded planetary Rossby waves</b>	<b>439</b>
14.4.1	Governing linear equations	439
14.4.2	Barotropic mode and baroclinic modes	440
14.4.3	Planetary Rossby waves with constant $N^2$	441
<b>14.5</b>	<b>Topographic Rossby waves</b>	<b>442</b>
14.5.1	Eigenvalue problem for $\tilde{\psi}$	443
14.5.2	Topographic Rossby waves with constant $N^2$	444
<b>14.6</b>	<b>Non-interacting Eady waves</b>	<b>445</b>
14.6.1	Assumptions for Eady waves	445
14.6.2	Streamfunction equation	446
14.6.3	Bottom trapped Eady waves	447
14.6.4	Upper surface trapped Eady waves	448
14.6.5	Comparing the top and bottom dispersion relations	448
14.6.6	Meridional and vertical motion within an Eady wave	449
14.6.7	Further study	451
<b>14.7</b>	<b>Interacting Eady waves and baroclinic instability</b>	<b>451</b>
14.7.1	Streamfunction solution	451
14.7.2	Dispersion relation	452
14.7.3	Growth rate	453
14.7.4	Further study	454
<b>14.8</b>	<b>Necessary conditions for instability</b>	<b>454</b>
14.8.1	Summary of the governing linear equations	455
14.8.2	Steps for deriving the necessary conditions	455
14.8.3	Necessary conditions based on the imaginary part	456
14.8.4	Necessary conditions based on the real part	457
14.8.5	Necessary condition for instability of the Eady model	457
14.8.6	Effects from adding a bottom slope to the Eady model	458
14.8.7	Flat bottom with constant buoyancy along the two boundaries	458
<b>14.9</b>	<b>Energetics of small amplitude fluctuations</b>	<b>459</b>
14.9.1	Time derivative terms	459
14.9.2	Advection by the background zonal flow	460
14.9.3	Summary of the energy equation	461
14.9.4	Horizontal and thermal wind shear production	461
14.9.5	Meridional and vertical eddy buoyancy fluxes	462
14.9.6	Tilting phase lines of unstable baroclinic waves	463
14.9.7	Caveats for extending the wedge of instability to parcels	465
14.9.8	Further reading	466

---

## 14.1 Loose threads

- Split this chapter into one with the linear waves and placed into Part I, and another chapter focusing on baroclinic instability.

- Vertically propagating Rossby waves with an infinite top
- Continuous modes
- Show some vertical baroclinic modes in Section 14.4.1 for constant  $N^2$  and for an exponential thermocline.
- Polarization relations for Eady waves.
- Group velocity for Eady waves
- Energetics for planetary waves, topographic waves, and Eady waves. Prove that phase averaged energy flux equals to the group velocity times the energy.
- Discuss the neutral wave case with  $u_b = c$  and further explore Footnote 10 in Chapter 9 of [Vallis \(2017\)](#).
- Offer further elaborations from [Vallis \(2017\)](#) and [Cushman-Roisin and Beckers \(2011\)](#).
- Eliassen-Palm fluxes and potential vorticity fluxes in Section 14.9.
- Discuss Figures 13.1 and 13.4 of [Gill \(1982\)](#).
- Charney model

## 14.2 Equations of quasi-geostrophy

As developed in Chapter ??, quasi-geostrophy is concerned with the evolution of hydrostatic and nearly geostrophic flow in the presence of a prescribed and gravitationally stable background state that is itself in hydrostatic and geostrophic balance. In this section we summarize salient features of continuously stratified quasi-geostrophy that are useful in the study of quasi-geostrophic waves and baroclinic instability.

### 14.2.1 Equations in the fluid interior

Within the interior of the fluid domain, the quasi-geostrophic buoyancy equation (??), relative vorticity equation (??), and potential vorticity equation (??), are given by

$$(\partial_t + \mathbf{u} \cdot \nabla) b = -w N^2 \quad \text{buoyancy equation} \quad (14.1a)$$

$$(\partial_t + \mathbf{u} \cdot \nabla) \zeta = -\beta v + f_0 \partial_z w \quad \text{relative vorticity equation} \quad (14.1b)$$

$$(\partial_t + \mathbf{u} \cdot \nabla) q = 0, \quad \text{quasi-geostrophic potential vorticity equation,} \quad (14.1c)$$

in which

$$\psi = p/(\rho_0 f_0) \quad \text{quasi-geostrophic streamfunction (}p=\text{pressure)} \quad (14.2a)$$

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi \quad \text{non-divergent geostrophic velocity} \quad (14.2b)$$

$$b = f_0 \partial_z \psi \quad \text{buoyancy} \quad (14.2c)$$

$$\zeta = \hat{\mathbf{z}} \cdot (\nabla \times \mathbf{u}) = \nabla_h^2 \psi \quad \text{geostrophic relative vorticity} \quad (14.2d)$$

$$q = f_0 + \beta y + \zeta + f_0 \partial_z (b/N^2) \quad \text{quasi-geostrophic potential vorticity} \quad (14.2e)$$

$$q = f_0 + \beta y + \nabla_h^2 \psi + f_0^2 \partial_z (\partial_z \psi / N^2) \quad \text{potential vorticity in terms of streamfunction.} \quad (14.2f)$$

The prescribed time-independent parameters are given by

$$N^2(z) > 0 \quad \text{background stratification} \quad (14.3a)$$

$$f(y) = f_o + \beta y \quad \beta\text{-plane Coriolis frequency.} \quad (14.3b)$$

Quasi-geostrophy is based on the following scale assumptions

$$\text{Ro} = U/f_o L \ll 1 \quad \text{small Rossby number} \quad (14.4a)$$

$$\beta L/f_o \ll 1 \quad \beta\text{-plane approximation} \quad (14.4b)$$

$$\text{Bu}(z) = [L_d(z)/L]^2 = [N(z) H/f_o]^2 L^{-2} \sim 1 \quad \text{order unity Burger number} \quad (14.4c)$$

$$\text{Ri}(z) = \text{Ro}^{-2} \text{Bu}(z) \gg 1 \quad \text{very large Richardson number.} \quad (14.4d)$$

The small Rossby number,  $\text{Ro}$ , means that the flow is strongly affected by the Coriolis acceleration, which furthermore means it is in near geostrophic balance. This scaling holds for both the background flow, which is prescribed, and perturbations to the background. The order unity Burger number says that the horizontal length scale of the flow,  $L$ , is comparable to the deformation radius,  $L_d$ , with the deformation radius a function of the vertical scale of motion,  $H$  (also the vertical size of the Eady model domain), the  $f$ -plane Coriolis parameter,  $f_o$ , and the prescribed background buoyancy frequency,  $N(z)$ . The large Richardson number,  $\text{Ri}$ , means that the flow is stable with respect to stratified shear instability (Chapter 13). Note that when writing the potential vorticity, the constant  $f_o$  can be dropped since it plays no role in the potential vorticity equation.

The evolution of buoyancy (equation (14.1a)) and relative vorticity (equation (14.1b)) are both impacted by the ageostrophic vertical velocity,  $w$ . However, we do not need to explicitly compute  $w$  to evolve the flow within the domain interior. The reason is that we can instead evolve the quasi-geostrophic potential vorticity through equation (14.1c). Thereafter, we solve the Poisson equation (with boundary conditions) for the geostrophic streamfunction

$$\nabla_h^2 \psi + f_o^2 \partial_z (\partial_z \psi / N^2) = q - (f_o + \beta y). \quad (14.5)$$

Upon updating the streamfunction we then update the velocity, buoyancy, and relative vorticity. Even though  $w$  is unnecessary for updating the flow state, it can be useful for a variety of diagnostic purposes. In Section ?? we derive a diagnostic equation for this ageostrophic velocity component.

### 14.2.2 Boundary conditions for vertically bounded domain

For the lateral boundaries, we assume either an infinite horizontal domain with all fields assumed to be finite or vanishing at spatial infinity, or assume doubly periodic domains. In this manner, the lateral boundaries play no fundamental role in the dynamics of concern in this chapter. In contrast, we are concerned with vertically bounded domains in which the top and bottom boundaries are central to the dynamics. To establish the corresponding boundary conditions, we make use of the buoyancy equation (14.1a), given the presence of the vertical velocity. Furthermore, we are concerned with perfect fluid quasi-geostrophy, so consider just the kinematic boundary conditions. The dynamic boundary conditions involve frictional stresses and are not considered here.

The top and bottom boundary conditions for quasi-geostrophy were studied in Sections ?? and ???. For the upper surface (top) boundary, the vertical velocity is vanishingly small

relative to interior vertical motion, thus prompting the rigid lid approximation. Evaluating the buoyancy equation (14.1a) at the rigid lid top boundary means that the boundary buoyancy is materially invariant

$$(\partial_t + \mathbf{u} \cdot \nabla_h) b = 0 \implies (\partial_t + \mathbf{u} \cdot \nabla_h) \partial_z \psi = 0 \quad \text{at } z = \bar{\eta}. \quad (14.6)$$

To within the accuracy of quasi-geostrophy, this boundary condition is evaluated at the resting position of the top boundary,  $z = \bar{\eta}$ , which is typically taken as  $\bar{\eta} = 0$ .

The analogous boundary condition at the domain bottom,  $z = \eta_b(x, y)$ , is given by

$$(\partial_t + \mathbf{u} \cdot \nabla_h) b = -N^2 \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \bar{\eta}_b. \quad (14.7)$$

which takes on the following form with the streamfunction

$$(\partial_t + \mathbf{u} \cdot \nabla_h) (f_0 \partial_z \psi) = -N^2 \mathbf{u} \cdot \nabla \eta_b \quad \text{at } z = \bar{\eta}_b. \quad (14.8)$$

The bottom boundary condition is evaluated at the horizontally averaged position,  $z = \bar{\eta}_b$ , since the more precise boundary location,  $z = \eta_b(x, y)$ , is one order higher in Rossby number and so is dropped from quasi-geostrophic theory. Correspondingly, quasi-geostrophic theory is formally valid only for very gently sloping bottom boundaries. Note that when  $\eta_b$  is a constant, then  $\nabla_h \eta_b = 0$ , so that the bottom boundary condition reduces to the material invariance of the boundary buoyancy, just like the top boundary condition.

Quasi-geostrophic theory based on studies of just the top and/or bottom boundary conditions is known as *surface quasi-geostrophy* (e.g., [Held et al. \(1995\)](#), [Yassin and Griffies \(2022\)](#)), where the interior potential vorticity is assumed to be a constant, which can be set to zero without loss of generality. Surface quasi-geostrophy shares mathematically elements with the study of surface gravity waves in Chapter 4. Namely, surface quasi-geostrophy supports edge waves that are exponentially trapped at the boundaries and with a vertical length scale inversely proportional to the horizontal wavelength (i.e., shorter edge waves are more trapped next to the boundary than longer edge waves). As we see later in this chapter, edge waves are central to baroclinic instability as realized in the Eady model.

## 14.3 Linear fluctuations on a zonal geostrophic background flow

We here formulate equations for a zonal background flow state and the linear fluctuations relative to that flow. The static and prescribed background state is assumed to be in thermal wind balance, and all perturbations to that background state satisfy the scalings of quasi-geostrophy. All background fields have a “b” subscript to remind us that these fields are prescribed.

The linear analysis in this section forms the baseline for subsequent sections that specialize these results. In particular, in Section 14.4 we specialize to the case of zero thermal wind flow with planetary beta, thus considering planetary Rossby waves. In Section 14.5, we set planetary beta to zero but allow for a topographic slope, thus considering topographic Rossby waves. In Section 14.6, we maintain the thermal wind state but assume zero planetary beta and zero topographic beta, thus focusing on the mechanics of *Eady edge waves*, whose existence along either the top or bottom boundary relies on the presence of a horizontal buoyancy gradient along that boundary. In Section 14.7 we study the unstable Eady edge waves arising in the Eady model, thus forming the wave mechanism for baroclinic instability.

### 14.3.1 Zonal geostrophic background flow

Consider a geostrophic background state described by a streamfunction,  $\Psi_b(y, z)$ , with a corresponding zonal flow and thermal wind shear

$$u_b(y, z) = -\partial_y \Psi_b \quad \text{and} \quad \partial_z u_b = -\partial_{yz} \Psi_b = -f_o^{-1} \partial_y b_b, \quad (14.9)$$

where  $b_b(y, z)$  is the prescribed background buoyancy field supporting the background geostrophic flow.

A particular example of buoyancy supporting a geostrophic flow with a constant thermal wind shear is given by the sloped and planar buoyancy surfaces

$$b_b = N^2 z + (\partial_y b_b) y = N^2 z - f_o (\partial_z u_b) y, \quad \text{with } N^2, \partial_y b_b, \text{ and } \partial_z u_b \text{ constants.} \quad (14.10)$$

We consider this special case when studying Eady waves in Section 14.6 and the Eady model of baroclinic instability in Section 14.7. Although very idealized, it provides the core features of Eady waves and baroclinic instability, and it does so in an analytically tractable manner.

The buoyancy (14.10) can be written in a geometric form by introducing the meridional slope of the background buoyancy surfaces. To derive an expression for the slope, note that constant buoyancy surfaces are defined by<sup>1</sup>

$$b_b = \text{constant} \implies db_b = 0 = (\partial b_b / \partial y) dy + (\partial b_b / \partial z) dz, \quad (14.11)$$

which means that the slope of constant buoyancy surfaces is given by

$$\left[ \frac{dz}{dy} \right]_{b_b} = -\frac{\partial_y b_b}{\partial_z b_b} = -\frac{\partial_y b_b}{N^2} = \frac{f_o \partial_z u_b}{N^2}, \quad (14.12)$$

so that the buoyancy (14.10) is

$$b_b = N^2 [z - y (dz/dy)_{b_b}] \quad \text{with } (dz/dy)_{b_b} \text{ a constant slope.} \quad (14.13)$$

Figure 14.1 illustrates such a buoyancy field with a constant and positive meridional slope,  $(dz/dy)_{b_b} > 0$ .

It is common to forget the minus signs appearing in the first and second expressions for the slope in equation (14.12). Examination of Figure 14.1 quickly remedies this mistake. Namely, this figure depicts a buoyancy field that has a positive slope towards the north,  $(dz/dy)_{b_b} > 0$ . This slope arises due to the reduction in buoyancy moving north ( $\partial_y b_b < 0$ ) in the presence of a vertically stable stratification ( $\partial_z b_b > 0$ ).

### 14.3.2 Background state is an exact quasi-geostrophic solution

We here show that the thermal wind flow (14.9) identically satisfies the quasi-geostrophic equations, thus constituting an exact solution to quasi-geostrophy. For this purpose, introduce the quasi-geostrophic potential vorticity for the background flow

$$q_b = f_o + \beta y + \zeta_b + f_o^2 \partial_z (N^{-2} \partial_z \Psi_b) = f_o + \beta y - \partial_y u_b + f_o^2 \partial_z (N^{-2} \partial_z \Psi_b), \quad (14.14)$$

---

<sup>1</sup>See Section ?? for mathematical details on the treatment of generalized vertical coordinates.

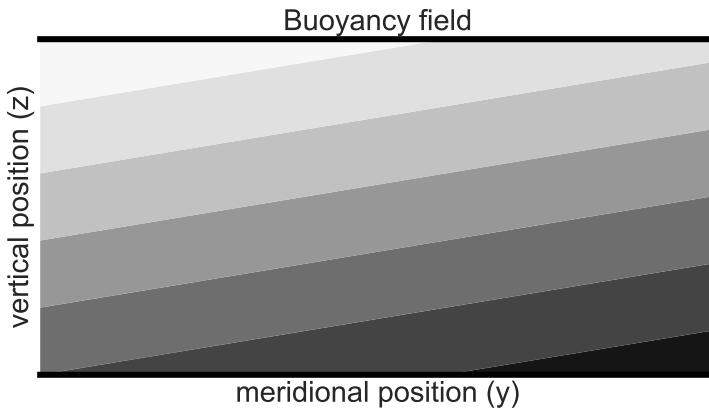


FIGURE 14.1: Example background buoyancy field written as the linear function,  $b_b = N^2 [z - y (\partial_z b_b)]$ , as per equation (14.13). The squared buoyancy frequency is positive,  $N^2 = \partial_z b_b > 0$ , whereas the northward buoyancy gradient is negative,  $\partial_y b_b < 0$ . Correspondingly, the buoyancy slope,  $(\partial_z b_b) = -\partial_y b_b / \partial_z b_b = f_o \partial_z u_b / N^2 > 0$ , is positive to the north (north is to the right). There are no lateral boundaries, so that the fluid is assumed to be unbounded in the horizontal directions. However, the top and bottom are rigid.

where we introduced the relative vorticity of the background zonal geostrophic flow

$$\zeta_b = -\partial_y u_b. \quad (14.15)$$

Since  $q_b$  is static and a spatial function just of  $(y, z)$ , it trivially satisfies the potential vorticity equation

$$(\partial_t + u_b \hat{x} \cdot \nabla_h) q_b = u_b \partial_x q_b = 0. \quad (14.16)$$

Likewise, the background buoyancy field,  $b_b(y, z)$ , trivially satisfies the top boundary condition (14.6) since

$$(\partial_t + u_b \hat{x} \cdot \nabla_h) b_b = u_b \partial_x b_b = 0 \quad \text{at } z = \bar{\eta}. \quad (14.17)$$

For the bottom boundary condition (14.8) we have

$$(\partial_t + u_b \partial_x) b_b + u_b \partial_x (N^2 \eta_b) = 0 \quad \text{at } z = \bar{\eta}_b, \quad (14.18)$$

which is satisfied if the topography is a function only of latitude,  $\eta_b = \eta_b(y)$ . Hence, the background thermal wind flow,  $u_b(y, z)$ , and corresponding buoyancy,  $b_b(y, z)$ , are exact solutions to quasi-geostrophy if the bottom is either flat or has a meridional slope.

### 14.3.3 Fluctuating streamfunction, potential vorticity, and buoyancy

Now consider fluctuations relative to the zonal geostrophic flow, with streamfunction and buoyancy decomposed as

$$\psi(x, y, z, t) = \Psi_b(y, z) + \psi'(x, y, z, t) \quad \text{and} \quad b(x, y, z, t) = b_b(y, z) + b'(x, y, z, t). \quad (14.19)$$

The corresponding quasi-geostrophic potential vorticity is decomposed according to

$$q(x, y, z, t) = q_b(y, z) + q'(x, y, z, t), \quad (14.20)$$

with  $q_b$  given by equation (14.14). The fluctuating potential vorticity arises from relative vorticity and stretching

$$q' = \nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi'). \quad (14.21)$$

Substituting these expressions into the quasi-geostrophic potential vorticity equation (14.1c) leads to

$$[\partial_t + (\mathbf{u}' + \mathbf{u}_b) \cdot \nabla] (q_b + q') = [\partial_t + (\mathbf{u}' + \mathbf{u}_b) \cdot \nabla] q' + \mathbf{u}' \cdot \nabla q_b = 0, \quad (14.22)$$

where we set

$$\partial_t q_b + \mathbf{u}_b \cdot \nabla q_b = 0, \quad (14.23)$$

since  $q_b$  is an exact solution to quasi-geostrophy (Section 14.3.2). Rearranging equation (14.22), and setting  $\mathbf{u}_b = u_b \hat{\mathbf{x}}$ , leads to

$$[\partial_t + (u_b \hat{\mathbf{x}} + \mathbf{u}') \cdot \nabla_h] q' + v' \partial_y q_b = 0. \quad (14.24)$$

We are concerned in this chapter with linear theory, with the linearized equation for the perturbation potential vorticity given by

$$(\partial_t + u_b \partial_x) q' = -v' \partial_y q_b \implies (\partial_t + u_b \partial_x) [\nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi')] = -\partial_x \psi' \partial_y q_b. \quad (14.25)$$

Evidently, in the linear theory we find that  $q'$  is advected by the zonal background flow,  $u_b$ , and it has a source determined by the anomalous meridional advection of the background quasi-geostrophic potential vorticity,  $q_b$ .

For the buoyancy,  $b(x, y, z, t) = b_b(y, z) + b'(x, y, z, t)$ , we linearize the buoyancy equation (14.1a) to find

$$(\partial_t + u_b \partial_x) b' = -v' \partial_y b_b - w' N^2. \quad (14.26)$$

The right hand side source terms can be written in terms of the slope of the buoyancy surfaces (14.12)

$$(\partial_t + u_b \partial_x) b' = -N^2 [w' - v' (dz/dy)_{b_b}]. \quad (14.27)$$

#### 14.3.4 Linearized upper boundary condition

The upper surface boundary condition (14.6) is

$$\partial_t b' + (u_b \partial_x + \mathbf{u}' \cdot \nabla_h) (b_b + b') = 0, \quad (14.28)$$

which linearizes to

$$(\partial_t + u_b \partial_x) b' = -v' \partial_y b_b \implies (\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' \partial_z u_b. \quad (14.29)$$

As we see in Section 14.6, advection by the zonal geostrophic flow provides a frequency shift to the linear waves, whereas the source,  $\partial_x \psi' \partial_z u_b$ , supports the propagation of surface trapped edge waves relative to the flow.

#### 14.3.5 Linearized bottom boundary condition

The bottom boundary condition (14.8), with  $\eta_b = \eta_b(y)$ , is given by

$$f_o (\partial_t + u_b \partial_x + \mathbf{u}' \cdot \nabla_h) \partial_z \psi' = v' (f_o \partial_z u_b - N^2 \partial_y \eta_b), \quad (14.30)$$

which linearizes to

$$f_o (\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' (f_o \partial_z u_b - N^2 \partial_y \eta_b). \quad (14.31)$$

The right hand side can be written in a geometric manner by introducing the slope of the background buoyancy as per equation (14.12), so that

$$(dz/dy)_{b_b} = -\partial_y b_b / \partial_z b_b = f_0 \partial_z u_b / N^2. \quad (14.32)$$

The bottom boundary condition (14.31) can thus be written as

$$f_0 (\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' N^2 [(dz/dy)_{b_b} - \partial_y \eta_b], \quad (14.33)$$

in which the forcing on the right hand side is proportional to the difference between the buoyancy slope and bottom topography slope. When the bottom topography is flat then the bottom boundary condition reduces to the same condition as the top

$$(\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' \partial_z u_b. \quad (14.34)$$

## 14.4 Vertically bounded planetary Rossby waves

In this section we simplify the background state by assuming zero background flow with flat top and bottom boundaries, yet affected by a nonzero planetary beta. That is, we here study continuously stratified planetary Rossby waves in a vertically bounded domain. We already encountered planetary Rossby waves in Section 6.3 for the horizontally non-divergent barotropic model, and in Section 7.9 for the shallow water model. The goal here is to extend those earlier discussions to the case of continuous stratification.

### 14.4.1 Governing linear equations

Again, we assume the background state has a flat bottom with zero background thermal wind flow, so that the background consists solely of planetary beta. Hence, the background potential vorticity equals to the planetary vorticity

$$q_b = f_0 + \beta y \implies \nabla q_b = \beta \hat{\mathbf{y}}. \quad (14.35)$$

The linearized potential vorticity equation (14.25) and linearized boundary conditions (14.29) and (14.31), each with zero background flow, are given by

$$\partial_t q' + \beta v' = 0 \implies \partial_t [\nabla_h^2 \psi' + f_0^2 \partial_z (N^{-2} \partial_z \psi')] + \beta \partial_x \psi' = 0 \quad (14.36a)$$

$$\partial_{tz} \psi' = 0 \quad \text{at } z = \bar{\eta} \text{ and } z = \bar{\eta}_b. \quad (14.36b)$$

In the presence of a horizontally homogeneous vertical stratification,  $N^2(z) > 0$ , we can introduce a wave ansatz consisting of horizontally traveling free plane waves that are modulated by a vertically dependent (generally complex) amplitude

$$\psi'(x, y, z, t) = \tilde{\psi}(z) e^{i(k_x x + k_y y - \omega t)} \implies \mathbf{u} = (\hat{\mathbf{z}} \times i \mathbf{k}) \tilde{\psi}(z) e^{i(k_x x + k_y y - \omega t)}. \quad (14.37)$$

As we will see when analytically calculating the vertical structure for constant  $N^2$  in Section 14.4.3, the vertical structure takes the form of vertically standing waves, which accords with our prior experience of waves in a bounded domain (e.g., acoustic waves in a rectangular cavity in Exercise 3.1; surface gravity waves in Section 4.8). Plugging the ansatz (14.37) into the streamfunction equation (14.36a) and boundary condition equation (14.36b) leads to the

*Sturm-Liouville* eigenvalue problem<sup>2</sup>

$$\frac{d}{dz} \left[ \frac{f_o^2}{N^2} \frac{d\tilde{\psi}}{dz} \right] = -\lambda \tilde{\psi} \quad (14.38a)$$

$$\frac{d\tilde{\psi}}{dz} = 0 \quad \text{at } z = \bar{\eta}, \bar{\eta}_b \quad (14.38b)$$

$$\lambda = -(|\mathbf{k}|^2 + \beta k_x/\omega). \quad (14.38c)$$

We refer to solutions  $\tilde{\psi}$  as *eigenmodes* with  $\lambda$  the corresponding *eigenvalues*. From the Sturm-Liouville theory we know there are a countably infinite number of eigenmodes, with the higher eigenvalues corresponding to modes with more zero crossings.

#### 14.4.2 Barotropic mode and baroclinic modes

Multiplying the eigenvalue equation (14.38a) by  $\tilde{\psi}^*$ , integrating over the depth of the fluid, and making use of the Neumann boundary conditions (14.38b), leads to the expression for the eigenvalue in terms of the eigenmodes

$$\lambda = -(|\mathbf{k}|^2 + \beta k_x/\omega) = \frac{\int_{\bar{\eta}_b}^{\bar{\eta}} |(f_o/N) d\tilde{\psi}/dz|^2 dz}{\int_{\bar{\eta}_b}^{\bar{\eta}} |\tilde{\psi}|^2 dz}, \quad (14.39)$$

which proves that the eigenvalues are non-negative

$$\lambda \geq 0. \quad (14.40)$$

##### Barotropic mode

The case with zero eigenvalue,  $\lambda = 0$ , has a dispersion relation given by

$$\omega = -k_x \beta / |\mathbf{k}|^2 \quad \text{barotropic mode}, \quad (14.41)$$

and a corresponding depth independent eigenmode ( $d\tilde{\psi}/dz = 0$ ), with this mode referred to as the *barotropic mode*. Notice that the barotropic mode's frequency is unbounded as the wavenumber gets smaller (longer waves). That is, the longer wavelength modes have higher frequency. Also note that the phase velocity (equation (1.26))

$$\mathbf{c}_p = (\omega/|\mathbf{k}|) \hat{\mathbf{k}} = (\omega/|\mathbf{k}|^2) \mathbf{k}, \quad (14.42)$$

has a westward component

$$(\mathbf{c}_p \cdot \hat{\mathbf{x}})^{\text{barotropic}} = (\omega/|\mathbf{k}|^2) \mathbf{k} \cdot \hat{\mathbf{x}} = -k_x^2 \beta / |\mathbf{k}|^2 < 0. \quad (14.43)$$

This westward phase velocity accords with our earlier studies of Rossby waves. Even though this barotropic mode appears in a stratified fluid, its properties are equivalent to that of the Rossby waves appearing in a horizontally non-divergent barotropic model as studied in Section 6.3, further supporting this mode being referred to it as the barotropic mode.

---

<sup>2</sup>Sturm-Liouville eigenvalue problems have a well established theory in differential equations and mathematical physics. Chapter 11 of *Boyce and DiPrima* (2009) is a classic reference, now in its 9th edition!

### Baroclinic modes

Equation (14.39) shows that all nonzero eigenvalues are positive

$$\lambda = -(|\mathbf{k}|^2 + \beta k_x/\omega) > 0. \quad (14.44)$$

Furthermore, upon vertically integrating the eigenvalue equation (14.38a) and using the boundary conditions (14.38b) we find that all higher modes have zero depth integral

$$\int_{\bar{\eta}_b}^{\bar{\eta}} \tilde{\psi} dz = 0. \quad (14.45)$$

We refer to these as *baroclinic modes* since they are depth dependent, with the baroclinic modes having eigenvalues that form a monotonically increasing and countably infinite sequence. Each successive baroclinic mode has one more zero crossing, and thus more vertical structure. We display this behavior in Section 14.4.3 when analytically determining the eigenmodes with a constant  $N^2$ .

#### 14.4.3 Planetary Rossby waves with constant $N^2$

An analytic solution to the eigenvalue problem (14.38a)-(14.38b) for vertically bounded Rossby waves can be found for the case of constant  $N^2$ , whereby the eigenmode is given by the cosine function

$$\tilde{\psi} = \tilde{\psi}_0 \cos[n \pi (z - \bar{\eta}_b)/H] \quad \text{with } H = \bar{\eta} - \bar{\eta}_b, \quad (14.46)$$

with  $\tilde{\psi}_0$  a constant real amplitude. The eigenvalues can be found through equation (14.39),

$$\lambda = (n \pi / L_d)^2 \quad \text{with } n = 0, 1, 2, 3, \dots \text{ and } L_d = N H / f_0, \quad (14.47)$$

which then lead to the dispersion relation for the baroclinic modes

$$\omega = -\frac{\beta k_x}{(n \pi / L_d)^2 + k_x^2 + k_y^2}. \quad (14.48)$$

Since the angular frequency is non-negative (Section 1.2.3), all planetary Rossby wave modes have a westward,  $k_x < 0$ , phase propagation, just like we saw for the barotropic planetary wave with  $n = 0$ . We display this property by writing the zonal component of the phase velocity

$$\mathbf{c}_p \cdot \hat{\mathbf{x}} = (\omega / |\mathbf{k}|^2) k_x = -\frac{\beta k_x^2 / |\mathbf{k}|^2}{(n \pi / L_d)^2 + k_x^2 + k_y^2} < 0. \quad (14.49)$$

Observe that a useful way to non-dimensionalize the dispersion relation (14.48) is to write

$$\omega / (\beta L_d) = -\frac{k_x L_d}{(n \pi)^2 + (k_x L_d)^2 + (k_y L_d)^2}, \quad (14.50)$$

with this expression plotted in Figure 14.2 for the  $n = 0, 1, 2$  Rossby wave modes with  $k_y = 0$ .

A packet of Rossby waves with carrier wavevector  $\mathbf{k}$  moves with the group velocity

$$\mathbf{c}_g = \nabla_{\mathbf{k}} \omega = \frac{\beta [\hat{\mathbf{x}} (k_x^2 - k_y^2 - (n \pi / L_d)^2) + 2 k_x k_y \hat{\mathbf{y}}]}{[(n \pi / L_d)^2 + k_x^2 + k_y^2]^2}. \quad (14.51)$$

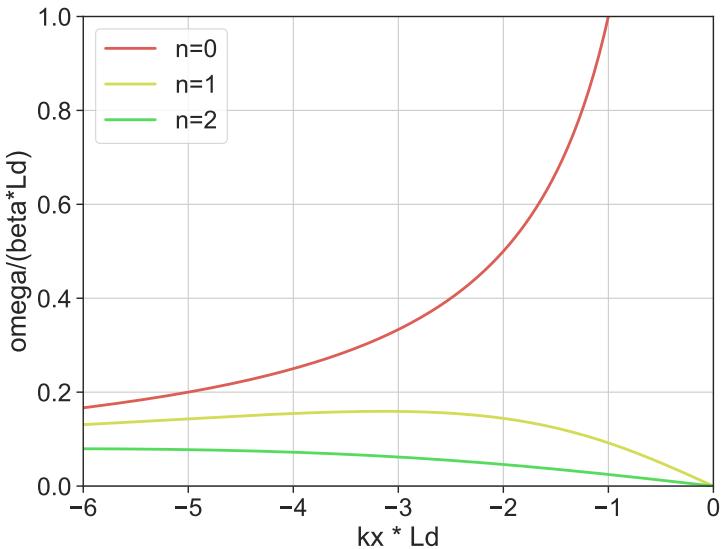


FIGURE 14.2: Dispersion relation for planetary Rossby waves according to equation (14.50),  $\omega / (\beta L_d) = -k_x L_d / [(n \pi)^2 + (k_x L_d)^2 + (k_y L_d)^2]$ , showing the  $n = 0$  barotropic mode and  $n = 1, 2$  baroclinic modes, each with  $k_y = 0$ . A positive angular frequency is realized with  $k_x < 0$ , so that the phase propagates to the west. The frequency for the barotropic mode is unbounded approaching the origin, so that long barotropic Rossby waves are high frequency waves. Both the barotropic and baroclinic wave modes have a vanishing frequency as the waves become shorter ( $|\mathbf{k}| \rightarrow \infty$ ).

In Section 6.3 we studied Rossby waves in the horizontally non-divergent barotropic model, and then in Section 7.9 we studied Rossby waves in a single shallow water layer. Both of those discussions share much with the present case, in particular the shallow water Rossby waves since they have a finite deformation radius whereas the deformation radius is formally infinite in the barotropic model. Hence, each baroclinic mode found in the continuously stratified case can be understood as a Rossby wave in a single shallow water layer whose deformation radius,

$$L_d = c_{\text{grav}} / f_o = \sqrt{g H} / f, \quad (14.52)$$

is chosen to fit that of the baroclinic mode. By extension, we can understand the geometry of baroclinic Rossby wave packets, including their reflection from surfaces, by referring to the earlier shallow water discussion.

## 14.5 Topographic Rossby waves

Now set  $\beta = 0$ , to eliminate planetary Rossby waves, but maintain a background potential vorticity gradient by allowing the bottom to have a nonzero slope. In so doing we extend the study of topographic Rossby waves from the shallow water in Section 7.4.6 to the continuously stratified case. The resulting linearized potential vorticity equation and boundary conditions are

$$\partial_t q' = 0 \implies \partial_t [\nabla_h^2 \psi' + f_0^2 \partial_z (N^{-2} \partial_z \psi')] = 0 \quad (14.53a)$$

$$\partial_{tz} \psi' = 0 \quad \text{at } z = \bar{\eta}_b. \quad (14.53b)$$

$$f_o \partial_{tz} \psi' + N^2 \hat{\mathbf{z}} \cdot (\nabla \psi' \times \nabla \eta_b) = 0 \quad \text{at } z = \bar{\eta}_b. \quad (14.53c)$$

To study plane waves in the horizontal, assume the special case of a gentle and constant linear slope,  $\partial_y \eta_b$ , in the meridional direction so that

$$\eta_b = \eta_{\text{const}} + y \partial_y \eta_b \quad \text{with } |\partial_y \eta_b| \ll 1, \quad (14.54)$$

in which case the bottom boundary condition (14.53c) becomes

$$f_o \partial_{tz} \psi' + N^2 \partial_y \eta_b \partial_x \psi' = 0 \quad \text{at } z = \bar{\eta}_b. \quad (14.55)$$

### 14.5.1 Eigenvalue problem for $\tilde{\psi}$

Taking the linear bottom slope (14.54) allows us to consider the same wave ansatz (14.37) as used for planetary Rossby waves in Section 14.4.1, which then converts equations (14.53a), (14.53b), and (14.55) into the eigenvalue problem

$$\omega \frac{d}{dz} \left[ \frac{f_o^2}{N^2} \frac{d\tilde{\psi}}{dz} \right] = \omega |\mathbf{k}|^2 \tilde{\psi} \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.56a)$$

$$\omega \frac{d\tilde{\psi}}{dz} = 0 \quad \text{at } z = \bar{\eta}, \quad (14.56b)$$

$$f_o \omega \frac{d\tilde{\psi}}{dz} = N^2 \partial_y \eta_b k_x \tilde{\psi} \quad \text{at } z = \bar{\eta}_b. \quad (14.56c)$$

The differential equation (14.56a) and surface boundary condition (14.56b) can be satisfied with a zero frequency wave,  $\omega = 0$ , in which case there is no propagating wave. However, the bottom boundary condition (14.56c) cannot be satisfied with  $\omega = 0$  in the presence of a nonzero bottom slope and nonzero Coriolis frequency. Evidently, the Coriolis frequency plays a fundamental role in supporting a propagating topographic Rossby waves in the presence of a sloping bottom.

Cancelling the angular frequency in equation (14.56a) (since  $\omega \neq 0$  for topographic waves), then multiplying by  $\tilde{\psi}^*$  and integrating over the depth of the domain leads to

$$\int_{\bar{\eta}_b}^{\bar{\eta}} \frac{d}{dz} \left[ \tilde{\psi}^* \frac{f_o^2}{N^2} \frac{d\tilde{\psi}}{dz} \right] dz = \int_{\bar{\eta}_b}^{\bar{\eta}} \left[ |\mathbf{k}|^2 |\tilde{\psi}|^2 + \frac{f_o^2}{N^2} \left| \frac{d\tilde{\psi}}{dz} \right|^2 \right] dz. \quad (14.57)$$

Use of the surface boundary condition (14.56b) and bottom boundary condition (14.56c) render

$$f_o \partial_y \eta_b k_x = -\frac{\omega}{|\tilde{\psi}(\bar{\eta}_b)|^2} \int_{\bar{\eta}_b}^{\bar{\eta}} \left[ |\mathbf{k}|^2 |\tilde{\psi}|^2 + \frac{f_o^2}{N^2} \left| \frac{d\tilde{\psi}}{dz} \right|^2 \right] dz. \quad (14.58)$$

The right hand side is a negative number (recall  $\omega > 0$ ), which then orients the zonal component to the phase velocity according to the sign of  $f_o \partial_y \eta_b$ . For example, in the northern hemisphere with a bottom slope rising to the north, so that  $f_o \partial_y \eta_b > 0$ , then  $k_x < 0$ , which means that the zonal phase velocity is to the west. Likewise, for the southern hemisphere, a bottom slope that is rising to the south has  $f_o \partial_y \eta_b > 0$ , which also yields  $k_x < 0$ .

### 14.5.2 Topographic Rossby waves with constant $N^2$

Following our approach for planetary Rossby waves in Section 14.4.3, assume the buoyancy frequency is constant so that the eigenvalue problem (14.56a)-(14.56c) reduces to

$$\frac{d^2\tilde{\psi}}{dz^2} = k_R^2 \tilde{\psi} \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.59a)$$

$$f_o \omega \frac{d\tilde{\psi}}{dz} = N^2 \partial_y \eta_b k_x \tilde{\psi} \quad \text{at } z = \bar{\eta}_b \quad (14.59b)$$

$$\omega \frac{d\tilde{\psi}}{dz} = 0 \quad \text{at } z = \bar{\eta}, \quad (14.59c)$$

where we introduced the inverse *Rossby height*

$$k_R = |\mathbf{k}| N / f_o = |\mathbf{k}| L_d / H \quad \text{with } L_d = N H / f_o. \quad (14.60)$$

As we see below,  $k_R^{-1}$  defines an exponential scale height over which the wave decays moving away from the bottom boundary.

#### Bottom trapped streamfunction

To further simplify the analysis, assume the upper boundary, at  $z = \bar{\eta}$ , is far enough away that it can be ignored. In this case the streamfunction takes on the bottom trapped form

$$\tilde{\psi} = \tilde{\psi}_o e^{-k_R(z - \bar{\eta}_b)}. \quad (14.61)$$

To determine what is “far enough away”, evaluate the streamfunction at  $z = \bar{\eta}$ , whereby

$$\tilde{\psi}(z = \bar{\eta}) = \tilde{\psi}_o e^{-k_R H} = \tilde{\psi}_o e^{-|\mathbf{k}| L_d}. \quad (14.62)$$

This streamfunction is exponentially small for wavenumbers satisfying

$$|\mathbf{k}| \gg L_d^{-1}. \quad (14.63)$$

That is, the topographic Rossby waves do not feel the upper boundary if their horizontal wavelength is small compared to the deformation radius

$$\Lambda \ll 2\pi L_d. \quad (14.64)$$

#### Dispersion relation

Inserting the streamfunction (14.61) into the bottom boundary condition (14.59b) leads to the dispersion relation

$$\omega = -N \partial_y \eta_b k_x / |\mathbf{k}|. \quad (14.65)$$

Since the angular frequency is positive, the wave vector is constrained so that  $\partial_y \eta_b k_x < 0$ . That is, if the bottom slope is rising to the north ( $\partial_y \eta_b > 0$ ) then the phase velocity has a westward component ( $k_x < 0$ ), whereas a slope rising to the south ( $\partial_y \eta_b < 0$ ) has an eastward phase velocity ( $k_x > 0$ ). We can also see this orientation by looking at the zonal component to the phase velocity

$$\mathbf{c}_p \cdot \hat{x} = (\omega / |\mathbf{k}|^2) k_x = -N \partial_y \eta_b k_x^2 / |\mathbf{k}|^3, \quad (14.66)$$

with the sign of  $\partial_y \eta_b$  determining the orientation of the zonal phase velocity.

Evidently, the topographic slope acts as a background potential vorticity gradient just like planetary beta. To further this correspondence, define *topographic beta*

$$\beta_{\text{topo}} = f_0 \partial_y \eta_b / H, \quad (14.67)$$

so that the dispersion relation (14.65) takes on the form

$$\omega = -\frac{\beta_{\text{topo}} k_x L_d}{|\mathbf{k}|}, \quad (14.68)$$

which shares features with the planetary Rossby wave from Section 14.4.1. However, in contrast to planetary beta, the topographic beta can be either sign.

## 14.6 Non-interacting Eady waves

The geometric expression for the bottom boundary condition (14.33) suggests that sloping buoyancy surfaces support waves in a manner akin to sloping bottom topography. Whereas a sloping bottom in the presence of flat buoyancy surfaces supports topographic Rossby waves, a sloping buoyancy surface in the presence of a flat bottom or rigid lid top supports *Eady waves*. More precisely, a nonzero gradient of boundary buoyancy supports Eady edge waves.

### 14.6.1 Assumptions for Eady waves

We make the following assumptions to support an analytic derivation of the dispersion relation for Eady waves.

- Disable planetary Rossby waves by setting  $\beta = 0$ .
- Disable topographic Rossby waves by setting  $\nabla \eta_b = 0$ .
- The flow occurs between a flat bottom at  $z = \bar{\eta}_b = 0$  and flat top at  $z = \bar{\eta} = H$ .
- A linear thermal wind front is supported by the buoyancy in equation (14.10), in which  $N^2$ ,  $\partial_y b_b$ , and  $\partial_z u_b$  are constants, thus implying that the buoyancy slope,  $(dz/dy)_b$ , is also a constant;
- The background zonal geostrophic flow is a linear function of vertical position so that

$$u_b = u_b(z) = U_0 + \partial_z u_b (z - \bar{\eta}_b), \quad (14.69)$$

where

$$\partial_z u_b = \text{constant}. \quad (14.70)$$

In turn, the corresponding geostrophic streamfunction is

$$\Psi_b = -y [U_0 + \partial_z u_b (z - \bar{\eta}_b)]. \quad (14.71)$$

- With these assumptions the background potential vorticity (14.14) is a constant throughout the fluid interior

$$q_b = f_0 \implies \nabla q_b = 0, \quad (14.72)$$

and thus it plays no dynamical role.

Each of the above assumptions is rather restrictive. Indeed, the use of a top boundary is not very relevant to the atmosphere. However, these assumptions offer a streamlined means to analytically reveal the core physics of Eady waves, and in turn the interaction of such waves to produce baroclinic instability (Section 14.7). It is for this reason that the Eady model has proven so compelling pedagogically.

### 14.6.2 Streamfunction equation

The linearized potential vorticity equation (14.25) and corresponding boundary conditions (14.29) and (14.33) take on the form

$$[\partial_t + u_b(z) \partial_x] q' = 0 \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.73a)$$

$$[\partial_t + u_b(\bar{\eta}) \partial_x] \partial_z \psi' = \partial_x \psi' \partial_z u_b \quad z = \bar{\eta} \quad (14.73b)$$

$$(\partial_t + u_b(\bar{\eta}_b) \partial_x) \partial_z \psi' = \partial_x \psi' \partial_z u_b \quad z = \bar{\eta}_b, \quad (14.73c)$$

with the bottom boundary condition resulting from the meridional gradient of the background buoyancy, or equivalently the zonal thermal wind shear

$$N^2 (dz/dy)_b = -\partial_y b_b = f_0 \partial_z u_b. \quad (14.74)$$

The simplifying assumptions from Section 14.6.1 have led to a linear boundary value problem in which the only spatial dependence is in the vertical. Hence, we can consider the familiar wave ansatz (14.37) for the streamfunction, in which fluctuations are organized into horizontally propagating free plane waves that are modulated by a vertically dependent amplitude function

$$\psi'(x, y, z, t) = \tilde{\psi}(z) e^{i(k_x x + k_y y - \omega t)}. \quad (14.75)$$

With this ansatz we find the fluctuating potential vorticity (14.21) is given by

$$q' = (f_0/N)^2 (\partial_{zz} - k_R^2) \psi', \quad (14.76)$$

and the linear potential vorticity equation (14.73a) is

$$(-\omega + u_b k_x) (\partial_{zz} - k_R^2) \tilde{\psi} = 0, \quad (14.77)$$

where

$$k_R = |\mathbf{k}| N/f_0 = |\mathbf{k}| L_d/H \quad (14.78)$$

is the inverse Rossby height originally introduced for topographic waves in equation (14.60).

Throughout the analysis we assume the angular frequency satisfies

$$\omega \neq u_b k_x, \quad (14.79)$$

which then means that the streamfunction equation (14.77) reduces to

$$\frac{d^2 \tilde{\psi}}{dz^2} = k_R^2 \tilde{\psi}, \quad (14.80)$$

which is the same equation as satisfied by the topographic Rossby wave (14.59a). As seen below, the angular frequency assumption (14.79) indeed holds for Eady waves.

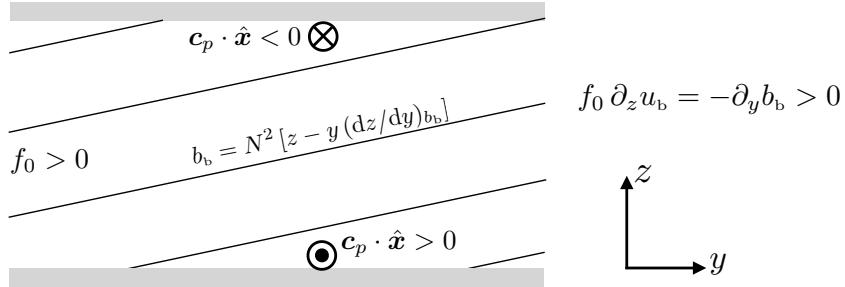


FIGURE 14.3: Depicting the orientation for the intrinsic phase velocity of Eady edge waves for a northern hemisphere thermal wind balanced base state with buoyancy decreasing northward. Relative to the local background flow, the upper boundary has westward propagating Eady waves whereas the Eady waves are eastward at the lower boundary (see Section 14.6.5).

### 14.6.3 Bottom trapped Eady waves

As in our discussion of topographic Rossby waves in Section 14.5, assume here that the upper boundary is far from the lower boundary, with “far” determined by a horizontal wavenumber satisfying equation (14.63), i.e.,

$$|\mathbf{k}| \gg L_d^{-1}. \quad (14.81)$$

In this case, the bottom trapped streamfunction solution to equation (14.80) is given by

$$\tilde{\psi} = \tilde{\psi}_{\text{bot}} e^{-k_R(z - \bar{\eta}_b)}, \quad (14.82)$$

which is, as expected, the same as for the topographic Rossby wave (14.61). Even though they have the same form for their streamfunctions, the Eady wave dispersion relation differs from that of the topographic Rossby wave. Here, we derive the dispersion relation through use of the bottom boundary condition (14.73c), which takes the form

$$(-\omega + u_b(\bar{\eta}_b) k_x)(-k_R) = k_x \partial_z u_b, \quad (14.83)$$

thus leading to the dispersion relation

$$\omega_{\text{bot}} = k_x u_b(\bar{\eta}_b) + (k_x/k_R) \partial_z u_b. \quad (14.84)$$

The vertical shear portion of the right hand side can be written in the following equivalent forms (again, each evaluated at  $z = \bar{\eta}_b$ )

$$\frac{f_0 \partial_z u_b}{N} = \frac{H}{L_d} \frac{\partial u_b}{\partial z} = -N^{-1} \partial_y b_b = N (\mathrm{d}z/\mathrm{d}y)_{b_b}, \quad (14.85)$$

where the final equality introduced the meridional buoyancy slope (14.12). We are thus led to the dispersion relation for the bottom trapped edge waves

$$\omega_{\text{bot}} = k_x [u_b + k_R^{-1} \partial_z u_b]_{z=\bar{\eta}_b} = k_x \left[ u_b + \frac{N}{|\mathbf{k}|} \left( \frac{\mathrm{d}z}{\mathrm{d}y} \right)_{b_b} \right]_{z=\bar{\eta}_b}. \quad (14.86)$$

The  $k_x u_b(\bar{\eta}_b)$  term provides a Doppler shifted frequency due to motion of the background zonal flow at the bottom. The next term arises from thermal wind shear in the presence of rotation and stratification, with the zonal thermal wind shear reliant on the meridionally sloped

buoyancy surfaces that intersect the bottom.

#### 14.6.4 Upper surface trapped Eady waves

Proceeding just like for the bottom, we now focus on the upper (top) boundary and assume the bottom boundary is far away. In this case the upper surface trapped streamfunction is given by

$$\tilde{\psi} = \psi_{\text{top}} e^{k_R(z - \bar{\eta})}, \quad (14.87)$$

which, when used in the upper boundary condition (14.29), leads to the dispersion relation

$$\omega_{\text{top}} - k_x u_b(\bar{\eta}) = -(k_x/k_R) \partial_z u_b. \quad (14.88)$$

This relation can be written just like equation (14.86) for the bottom boundary condition, only with a swapped sign on the buoyancy slope term

$$\omega_{\text{top}} = k_x [u_b - k_R^{-1} \partial_z u_b]_{z=\bar{\eta}} = k_x \left[ u_b - \frac{N}{|\mathbf{k}|} \left( \frac{dz}{dy} \right)_{b_b} \right]_{z=\bar{\eta}}. \quad (14.89)$$

As for the bottom trapped Eady waves, the term  $k_x u_b(\bar{\eta})$  is a Doppler shift in the frequency arising from the zonal velocity at the upper boundary that couples to the zonal component of the phase velocity. The other term arises from the thermal wind shear coupled to rotation and stratification, which itself relies on the slope of the buoyancy surfaces that intersect the upper boundary.

#### 14.6.5 Comparing the top and bottom dispersion relations

It is useful to compare the dispersion relations (14.86) and (14.89), rewritten here as the  $\hat{x}$  component of the phase velocity

$$\mathbf{c}_p \cdot \hat{x} = (\omega/|\mathbf{k}|^2) k_x, \quad (14.90)$$

which are given by

$$(\mathbf{c}_p \cdot \hat{x})_{\text{bot}} = \frac{k_x^2}{|\mathbf{k}|^2} \left[ u_b + \frac{H \partial_z u_b}{L_d |\mathbf{k}|} \right]_{z=\bar{\eta}_b} = \frac{k_x^2}{|\mathbf{k}|^2} \left[ u_b + \frac{N}{|\mathbf{k}|} \left( \frac{dz}{dy} \right)_{b_b} \right]_{z=\bar{\eta}_b} \quad (14.91a)$$

$$(\mathbf{c}_p \cdot \hat{x})_{\text{top}} = \frac{k_x^2}{|\mathbf{k}|^2} \left[ u_b - \frac{H \partial_z u_b}{L_d |\mathbf{k}|} \right]_{z=\bar{\eta}} = \frac{k_x^2}{|\mathbf{k}|^2} \left[ u_b - \frac{N}{|\mathbf{k}|} \left( \frac{dz}{dy} \right)_{b_b} \right]_{z=\bar{\eta}}. \quad (14.91b)$$

The first difference between these expressions arises from the differences in the background zonal flow at  $z = \bar{\eta}_b$  versus  $z = \bar{\eta}$ . Assuming  $\partial_z u_b > 0$ , the background flow at the top has a larger eastward value than the background flow at the bottom, as given by

$$u_b(\bar{\eta}) = u_b(\bar{\eta}_b) + (\bar{\eta} - \bar{\eta}_b) \partial_z u_b = u_b(\bar{\eta}_b) + H \partial_z u_b. \quad (14.92)$$

The second difference between the bottom dispersion relation (14.91a) and top dispersion relation (14.91b) arises from the swapped signs for the buoyancy slope term. This term represents the *intrinsic frequency* for the two edge waves (i.e., the frequency seen by an observer

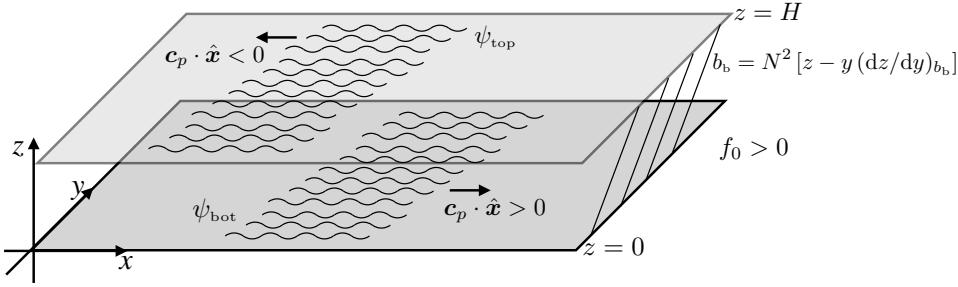


FIGURE 14.4: A perspective view of the intrinsic phase velocity of Eady edge waves for a northern hemisphere thermal wind balanced base state with buoyancy increasing northward. The upper boundary has westward propagating Eady waves whereas the Eady waves are eastward at the lower boundary. The top and bottom boundaries are flat, and so we write  $z = \bar{\eta} = H$  for the upper boundary and  $z = \bar{\eta}_b = 0$  for the lower boundary.

moving with the background flow)

$$\omega_{\text{bot}}^{\text{intrinsic}} = \frac{N k_x^{\text{bot}}}{|\mathbf{k}|} \left( \frac{dz}{dy} \right)_{b_b} \quad \text{and} \quad \omega_{\text{top}}^{\text{intrinsic}} = -\frac{N k_x^{\text{top}}}{|\mathbf{k}|} \left( \frac{dz}{dy} \right)_{b_b}. \quad (14.93)$$

For surfaces of constant buoyancy that slope upward to the north,  $(dz/dy)_{b_b} > 0$ , the intrinsic frequency for the bottom wave is positive if the phase velocity is eastward relative to the background flow ( $k_x^{\text{bot}} > 0$ ), whereas a positive frequency for the top wave requires a westward phase velocity ( $k_x^{\text{top}} < 0$ ). We also see this orientation by writing the intrinsic components to the zonal phase velocities

$$(\mathbf{c}_p \cdot \hat{x})_{\text{bot}}^{\text{intrinsic}} = \frac{N k_x^2}{|\mathbf{k}|^2} \left( \frac{dz}{dy} \right)_{b_b} > 0 \quad \text{and} \quad (\mathbf{c}_p \cdot \hat{x})_{\text{top}}^{\text{intrinsic}} = -\frac{N k_x^2}{|\mathbf{k}|^2} \left( \frac{dz}{dy} \right)_{b_b} < 0. \quad (14.94)$$

The zonal component to the phase velocity in equations (14.91a) and (14.91b) are equal for that wavevector whose magnitude satisfies

$$L_d |\mathbf{k}| = 2, \quad (14.95)$$

in which case

$$(\mathbf{c}_p \cdot \hat{x})_{\text{top}} = (k_x^2/|\mathbf{k}|^2) [u_b(\bar{\eta}) - (H/2) \partial_z u_b] = (k_x^2/|\mathbf{k}|^2) [u_b(\bar{\eta}_b) + (H/2) \partial_z u_b] = (\mathbf{c}_p \cdot \hat{x})_{\text{bot}}. \quad (14.96)$$

Evidently, waves with  $L_d |\mathbf{k}| = 2$  have the opportunity to phase lock and thus to interact. This possibility motivates the work in Section 14.7 whereby we include interactions in the formulation.

#### 14.6.6 Meridional and vertical motion within an Eady wave

To help understand fluid particle motion in a stable Eady wave, we compute the meridional velocity,  $v'$ , and vertical velocity,  $w'$ , for a wave moving along the bottom boundary, with a similar calculation holding along the top. Recall that since the flow is quasi-geostrophic that the vertical velocity is much smaller than the horizontal. Even so, it is nonzero and we can compute it for the Eady wave, thus revealing motion in the meridional-depth plane. For this calculation we use the streamfunction for a bottom trapped Eady wave

$$\psi'(x, y, z, t) = \tilde{\psi}_o e^{-k_R(z - \bar{\eta}_b)} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t)}, \quad (14.97)$$

with  $\omega_{\text{bot}}$  given by the dispersion relation (14.86).

The meridional velocity of fluid particles within the bottom Eady wave is given by

$$v' = \partial_x \psi' = i k_x \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} e^{i(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t)}, \quad (14.98)$$

whose real part is

$$v' = -k_x \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t). \quad (14.99)$$

A similar calculation leads to the zonal velocity within the wave

$$u' = k_y \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t). \quad (14.100)$$

Calculating the vertical velocity takes a bit more work, for which we make use of the linearized buoyancy equation (14.26) so that

$$w' N^2 = -(\partial_t + u_b \partial_x) b' - v' \partial_y b_b. \quad (14.101)$$

The first term is given by

$$-(\partial_t + u_b \partial_x) b' = i k_R f_o (-\omega_{\text{bot}} + k_x u_b) \psi', \quad (14.102)$$

whose real part is

$$-(\partial_t + u_b \partial_x) b' = k_R f_o (\omega_{\text{bot}} - k_x u_b) \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t), \quad (14.103)$$

and the meridional velocity term is

$$-v' \partial_y b_b = k_x \partial_y b_b \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t), \quad (14.104)$$

so that

$$w' N^2 = \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t) [k_x \partial_y b_b + k_R f_o (\omega_{\text{bot}} - k_x u_b)]. \quad (14.105)$$

Making use of the dispersion relation (14.86) for the bottom trapped Eady edge waves leads to

$$k_x \partial_y b_b + k_R f_o (\omega_{\text{bot}} - k_x u_b) = k_x \partial_y b_b + k_R f_o k_x (k_R^{-1} - z) \partial_z u_b \quad (14.106a)$$

$$= k_R k_x z \partial_y b_b, \quad (14.106b)$$

so that the vertical velocity component is

$$w' = -k_R k_x z (dz/dy)_b \tilde{\psi}_o e^{-k_R(z-\bar{\eta}_b)} \sin(\mathbf{k} \cdot \mathbf{x} - \omega_{\text{bot}} t) \quad (14.107)$$

and its ratio with the meridional velocity is

$$w'/v' = z k_R (dz/dy)_b = z |\mathbf{k}| (N/f_o) (dz/dy)_b. \quad (14.108)$$

For vertical positions less than the Rossby height,  $k_R^{-1}$ , the fluid particle motion is more horizontal than the buoyancy slope, whereas the motion is more vertical than the slope for positions higher than  $k_R^{-1}$ . Also note that the amplitude of the motion exponentially decays moving away from the bottom boundary with an efolding height  $k_R^{-1}$ . Figure 14.5 provides a schematic of this motion.

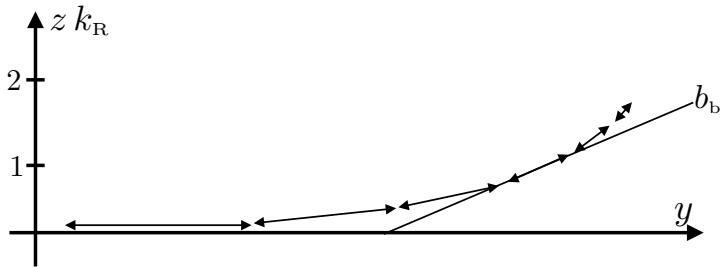


FIGURE 14.5: A meridional-vertical view of fluid particle velocity within a stable bottom trapped Eady wave, as given by equation (14.108). The particle velocity is parallel to the bottom when at  $z = \bar{\eta}_b = 0$ , and parallel to the buoyancy surface at vertical position  $z k_R = 1$ , and further steepens when moving higher in the column. The amplitude of the oscillatory motion exponentially decays away from the bottom with an e-folding scale given by the Rossby height,  $k_R^{-1}$ . This figure is adapted from Figure 13.1(e) in [Gill \(1982\)](#).

### 14.6.7 Further study

The treatment in this section is compatible with Section 13.2 of [Gill \(1982\)](#), Chapter 19 of [Pedlosky \(2003\)](#) and Section 8.7 of [Smyth and Carpenter \(2019\)](#).

## 14.7 Interacting Eady waves and baroclinic instability

The analysis from Section 14.6 revealed the presence of Eady edge waves that are exponentially trapped next to the rigid top and bottom boundaries, with the wave supported by the thermal wind shear intersecting the boundaries. Furthermore, the zonal component to the phase velocity of the two waves is equal (equation (14.6.5)) for a horizontal wavenumber

$$\omega_{\text{top}} = \omega_{\text{bot}} \implies |\mathbf{k}| = 2 L_d^{-1}. \quad (14.109)$$

For this wave, and for longer waves (smaller wavenumbers), we cannot ignore wave interactions. Under specific circumstances detailed in this section, the waves constructively interact to support mutual exponential growth, thus signaling *baroclinic instability*. The treatment is directly analogous to our study in Section 13.6 of interacting vortex edge waves in a horizontally sheared fluid.

### 14.7.1 Streamfunction solution

As before, we take the streamfunction ansatz (14.75), which builds in an assumption that the Eady waves at the two boundaries move with the same phase velocity to thus allow them to interact. We are thus led to the differential equation (14.80) and boundary conditions (14.73c) and (14.83)

$$\frac{d^2 \tilde{\psi}}{dz^2} = k_R^2 \tilde{\psi} \quad 0 < z < H \quad (14.110a)$$

$$[-\omega + u_b(H) k_x] \partial_z \tilde{\psi} = k_x \partial_z u_b \tilde{\psi} \quad z = H \quad (14.110b)$$

$$[-\omega + u_b(0) k_x] \partial_z \tilde{\psi} = k_x \partial_z u_b \tilde{\psi} \quad z = 0, \quad (14.110c)$$

where we made use of the inverse Rossby height (14.60)

$$k_R = |\mathbf{k}| N / f_o = |\mathbf{k}| L_d / H \quad \text{with } L_d = N H / f_o. \quad (14.111)$$

Additionally, we set

$$\bar{\eta} = H \quad \text{and} \quad \bar{\eta}_b = 0, \quad (14.112)$$

to help reduce notational clutter (see Figure 14.4).

### Solution to the boundary value problem

The solution to the differential equation (14.110a) is given by the hyperbolic functions

$$\tilde{\psi} = C_1 \cosh(k_R z) + C_2 \sinh(k_R z), \quad (14.113)$$

where the coefficients  $C_1$  and  $C_2$ , are specified from the boundary conditions. Making use of the upper surface boundary condition (14.110b) and bottom boundary condition (14.110c) leads to the coupled linear equations

$$(-\tilde{c} + H \partial_z u_b) k_R [C_1 \sinh(k_R H) + C_2 \cosh(k_R H)] = \partial_z u_b [C_1 \cosh(k_R H) + C_2 \sinh(k_R H)] \quad (14.114a)$$

$$-\tilde{c} k_R C_2 = C_1 \partial_z u_b, \quad (14.114b)$$

where we introduced the shorthand

$$\tilde{c} \equiv \omega/k_x - u_b(0), \quad (14.115)$$

and noted that the constant thermal wind shear means that

$$u_b(H) = u_b(0) + H \partial_z u_b. \quad (14.116)$$

#### 14.7.2 Dispersion relation

Equations (14.114a) and (14.114b) are two homogeneous linear equations for the two unknowns,  $C_1$  and  $C_2$ . A nontrivial solution exists if the determinant of the  $2 \times 2$  coefficient matrix vanishes. This condition leads to the quadratic equation for  $\tilde{c}$

$$\tilde{c}^2 - \tilde{c} H \partial_z u_b + (\partial_z u_b/k_R)^2 (H k_R \coth(k_R H) - 1) = 0, \quad (14.117)$$

which has two roots given by the quadratic formula

$$\tilde{c} = \frac{H \partial_z u_b}{2} \pm \frac{\partial_z u_b}{k_R} \left[ \left( \frac{k_R H}{4} \right)^2 - k_R H \coth(k_R H) + 1 \right]^{1/2}. \quad (14.118)$$

A further simplification arises from use of the identity

$$2 \coth x = \tanh(x/2) + \coth(x/2), \quad (14.119)$$

in which case we have the roots<sup>3</sup>

$$\omega/k_x = u_b(z=0) + \frac{H \partial_z u_b}{2} \pm \frac{\partial_z u_b}{k_R} \left[ \left( \frac{k_R H}{2} - \coth(k_R H/2) \right) \left( \frac{k_R H}{2} - \tanh(k_R H/2) \right) \right]^{1/2}. \quad (14.120)$$

---

<sup>3</sup>The function  $L(x) = \coth x - 1/x$  is known as the Langevin function. It appears in the study of the statistical mechanics of paramagnetism; e.g., see Section 3.8 of *Pethria and Beale* (2021).

Since

$$x \geq \tanh x, \quad (14.121)$$

the only way to realize a negative discriminant is for

$$k_R H \leq 2 \coth(k_R H/2). \quad (14.122)$$

A graphical solution to the equality finds the critical nondimensional wavenumber

$$k_R H = |\mathbf{k}| L_d \approx 2.399. \quad (14.123)$$

Longer waves that satisfy the inequality are baroclinically unstable

$$k_R H = |\mathbf{k}| L_d < 2.399 \implies \text{unstable} \quad (14.124)$$

Note how  $|\mathbf{k}| L_d \approx 2.399$  is rather close to the  $|\mathbf{k}| L_d = 2$  value from equation (14.109), computed by setting the frequencies equal for two non-interacting Eady edge waves.

### 14.7.3 Growth rate

For unstable waves, the growth rate is given by the imaginary part of the angular frequency

$$\sigma = \frac{|k_x \partial_z u_b|}{k_R} \left[ \left| \frac{k_R H}{2} - \coth(k_R H/2) \right| \left| \frac{k_R H}{2} - \tanh(k_R H/2) \right| \right]^{1/2}. \quad (14.125)$$

Due to the  $k_x$  factor in the front, we see that for any given horizontal wavenumber,

$$|\mathbf{k}| = \sqrt{k_x^2 + k_y^2}, \quad (14.126)$$

the growth rate is maximized for waves with a zonal phase velocity, in which case

$$|\mathbf{k}| = |k_x|. \quad (14.127)$$

That is, the most unstable waves whose phase velocity is aligned with the background flow<sup>4</sup>

$$\sigma(\mathbf{k} = \hat{\mathbf{x}} k_x) = \frac{H |\partial_z u_b|}{L_d} \left[ \left| \frac{|\mathbf{k}| L_d}{2} - \coth(|\mathbf{k}| L_d/2) \right| \left( \frac{|\mathbf{k}| L_d}{2} - \tanh(|\mathbf{k}| L_d/2) \right) \right]^{1/2}. \quad (14.128)$$

Since the horizontal flow is non-divergent, Eady waves are horizontally transverse so that fluid particle movement is perpendicular to the phase velocity. With a zonal phase velocity, fluid particles move in the meridional direction.

#### Maximum growth rate

To determine the maximum growth rate we compute the wavenumber,  $|\mathbf{k}| = |k_x|$ , that satisfies

$$\frac{d\sigma^2}{dk_x} = 0, \quad (14.129)$$

which yields

$$|\mathbf{k}|_{\max} = |k_x|_{\max} \approx 1.6/L_d \implies \Lambda_{\max} \approx (2\pi/1.6) L_d \approx 3.9 L_d. \quad (14.130)$$

---

<sup>4</sup>This is a version of Squires theorem described at the start of Chapter 13.

Evidently, the fastest growing Eady waves are purely zonal and have wavelength about four times the deformation radius. It is this connection to the deformation radius that directly connects these unstable Eady waves to synoptic eddies in the atmosphere and mesoscale eddies in the ocean.

### Eady growth rate

The growth rate (14.128) is the product of an inverse time scale,  $T_{\text{Eady}}^{-1}$ , and a non-dimensional function, where the Eady time scale is given by

$$T_{\text{Eady}} = \frac{L_d}{H \partial_z u_b} = \frac{1}{f_o} \frac{N}{\partial_z u_b} = \frac{\sqrt{\text{Ri}}}{f_o} = [N |(\text{d}z/\text{d}y)_{b_b}|]^{-1}. \quad (14.131)$$

We here introduced the *balanced Richardson number* from Section ??

$$\text{Ri} = N^2 / (\partial_z u_b)^2 = f_o^2 [N (\text{d}z/\text{d}y)_{b_b}]^{-2}. \quad (14.132)$$

For quasi-geostrophic flows, the balanced Richardson number is normally quite large. Evaluating the growth rate expression (14.128) with the fastest growing wave (14.130) renders

$$\sigma_{\max} = 0.31 T_{\text{Eady}}^{-1} = 0.31 |\partial_z u_b (f_o/N)| = 0.31 \frac{|f_o|}{\sqrt{\text{Ri}}}. \quad (14.133)$$

This maximum growth rate is sometimes generically called *the Eady growth rate*. For a background state with  $\text{Ri} = 100$  and  $f_o = 10^{-4} \text{ s}^{-1}$ , we find the growth rate of the most unstable Eady wave

$$\sigma_{\max} \approx (3.7 \text{ days})^{-1}, \quad (14.134)$$

which accords with the growth rate of middle latitude atmospheric cyclones. For the ocean, the Eady growth rate is roughly ten times slower than the atmosphere (e.g., see Figure 1 in [Treguier et al. \(1997\)](#)).

#### 14.7.4 Further study

Our presentation of the Eady model is consistent with that found in [Pedlosky \(2003\)](#) and [Vallis \(2017\)](#). Although maths and pictures are revealing, it is also useful to observe laboratory rotating tank experiments to further ones understanding of the Eady model and baroclinic instability.

## 14.8 Necessary conditions for instability

In Section 13.5 we established the Rayleigh-Kuo inflection point theorem as well as Fjørtoft's theorem, each establishing necessary conditions for shear instability. Here we pursue a similar approach for baroclinic instability of a zonal background flow

$$u_b = u_b(y, z), \quad (14.135)$$

and maintain nonzero planetary  $\beta$  so that

$$f = f_o + \beta y. \quad (14.136)$$

The analysis of non-zonal flows greatly adds to the technical overhead, largely since the flow is no longer perpendicular to the planetary vorticity gradient. We thus restrict attention to zonal background flows.

As for the shear instability case, the necessary conditions for instability identifies cases whereby instabilities are possible. However, the integral conditions are not sufficient conditions. Hence, detailed calculations are required to determine if the flow is indeed unstable even if it satisfies the necessary conditions. Both the derivation of the necessary conditions, and their form, offer insights into the mechanics of baroclinic instability.

### 14.8.1 Summary of the governing linear equations

From Section 14.3.3, we have the governing linear equation (14.25) for the geostrophic streamfunction and the boundary conditions (14.29) and (14.33)

$$(\partial_t + u_b \partial_x) [\nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi')] = -\partial_x \psi' \partial_y q_b \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.137a)$$

$$(\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' \partial_z u_b \quad z = \bar{\eta} \quad (14.137b)$$

$$f_o (\partial_t + u_b \partial_x) \partial_z \psi' = \partial_x \psi' N^2 [(\mathrm{d}z/\mathrm{d}y)_b - \partial_y \eta_b] \quad z = \bar{\eta}_b, \quad (14.137c)$$

where the meridional derivative of the background potential vorticity is given by equation (14.14)

$$\partial_y q_b = \partial_y [f_o + \beta y - \partial_y u_b + f_o^2 \partial_z (N^{-2} \partial_z \Psi_b)] = \beta - \partial_{yy} u_b + f_o^2 \partial_{yz} (N^{-2} \partial_z \Psi_b). \quad (14.138)$$

The background flow is a function of  $(y, z)$ , which means it only supports freely propagating plane waves in the zonal direction. We thus consider the following wave ansatz for the streamfunction

$$\psi'(x, y, z, t) = \tilde{\psi}(y, z) e^{i(k_x x - \omega t)}. \quad (14.139)$$

Use of this ansatz in the boundary value problem (14.137a)-(14.137c) yields

$$(u_b - c) [(\partial_{yy} - k_x^2) \tilde{\psi} + f_o^2 \partial_z (N^{-2} \partial_z \tilde{\psi})] = -\tilde{\psi} \partial_y q_b \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.140a)$$

$$(u_b - c) \partial_z \tilde{\psi} = \tilde{\psi} \partial_z u_b \quad z = \bar{\eta} \quad (14.140b)$$

$$f_o (u_b - c) \partial_z \tilde{\psi} = N^2 \tilde{\psi} [(\mathrm{d}z/\mathrm{d}y)_b - \partial_y \eta_b] \quad z = \bar{\eta}_b, \quad (14.140c)$$

where the zonal phase velocity is written

$$\mathbf{c}_p = (\omega/|\mathbf{k}|) \hat{\mathbf{k}} = (\omega/k_x) \hat{\mathbf{x}} = c \hat{\mathbf{x}}. \quad (14.141)$$

Note that the phase velocity,  $c$ , is generally complex, which means that the streamfunction,  $\tilde{\psi}$ , is also complex. Just like for the horizontal shear case in Section 13.6, if  $c$  and  $\tilde{\psi}$  satisfy the boundary value problem (14.140a)-(14.140c), then so do their complex conjugates,  $c^*$  and  $\tilde{\psi}^*$ .

### 14.8.2 Steps for deriving the necessary conditions

To develop a necessary condition for instability, multiply the differential equation (14.140a) by  $\tilde{\psi}^*$  and integrate over the full domain

$$\int \left[ \tilde{\psi}^* (\partial_{yy} - k_x^2) \tilde{\psi} + \tilde{\psi}^* f_o^2 \partial_z (N^{-2} \partial_z \tilde{\psi}) + \frac{|\tilde{\psi}|^2 \partial_y q_b}{U - c} \right] \mathrm{d}y \mathrm{d}z = 0. \quad (14.142)$$

**Meridional derivative term**

Consider the meridional derivative term, in which we have

$$\int \tilde{\psi}^* \partial_{yy} \tilde{\psi} dy = \int [\partial_y (\tilde{\psi}^* \partial_y \tilde{\psi}) - |\partial_y \tilde{\psi}|^2] dy. \quad (14.143)$$

Assuming either meridionally periodic conditions, or fields that decay at infinity, allows us to drop the total derivative term to have

$$\int \tilde{\psi}^* \partial_{yy} \tilde{\psi} dy = - \int |\partial_y \tilde{\psi}|^2 dy. \quad (14.144)$$

**Vertical derivative term**

For the vertical derivative term we have

$$\int \tilde{\psi}^* \partial_z (N^{-2} \partial_z \tilde{\psi}) dz = \int [\partial_z (\tilde{\psi}^* N^{-2} \partial_z \tilde{\psi}) - N^{-2} |\partial_z \tilde{\psi}|^2] dz. \quad (14.145)$$

The boundary conditions (14.140b) and (14.140c) lead to

$$[\tilde{\psi}^* N^{-2} \partial_z \tilde{\psi}]_{z=\bar{\eta}} = \left[ \frac{|\tilde{\psi}|^2 \partial_z u_b}{N^2 (u_b - c)} \right]_{z=\bar{\eta}} \quad (14.146a)$$

$$[\tilde{\psi}^* N^{-2} \partial_z \tilde{\psi}]_{z=\bar{\eta}_b} = \left[ \frac{|\tilde{\psi}|^2 [(dz/dy)_b - \partial_y \eta_b]}{f_o (u_b - c)} \right]_{z=\bar{\eta}_b} \quad (14.146b)$$

It is important to note that these boundary terms are generally nonzero, and so they play a role in determining the necessary conditions for instability.

### 14.8.3 Necessary conditions based on the imaginary part

The various pieces bring the integral (14.142) into the form

$$\begin{aligned} & \int [k_x^2 |\tilde{\psi}|^2 + |\partial_y \tilde{\psi}|^2 + (f_o/N)^2 |\partial_z \tilde{\psi}|^2] dy dz \\ &= \int \frac{|\tilde{\psi}|^2 \partial_y q_b}{u_b - c} dy dz + \int_{z=\bar{\eta}} \frac{f_o^2 |\tilde{\psi}|^2 \partial_z u_b}{N^2 (u_b - c)} dy - \int_{z=\bar{\eta}_b} \frac{f_o |\tilde{\psi}|^2 [(dz/dy)_b - \partial_y \eta_b]}{(u_b - c)} dy. \end{aligned} \quad (14.147)$$

The left hand side is a real and non-negative number. For consistency, the imaginary part of the right hand side must vanish. Making use of the identity

$$\frac{1}{u_b - c} = \frac{u_b - c_r - i c_i}{|u_b - c|^2} \quad (14.148)$$

leads to the constraint

$$c_i \left[ \int \frac{|\tilde{\psi}|^2 \partial_y q_b}{|u_b - c|^2} dy dz + \int_{z=\bar{\eta}} \frac{f_o^2 |\tilde{\psi}|^2 \partial_z u_b}{N^2 |u_b - c|^2} dy - \int_{z=\bar{\eta}_b} \frac{f_o |\tilde{\psi}|^2 [(dz/dy)_b - \partial_y \eta_b]}{|u_b - c|^2} dy \right] = 0. \quad (14.149)$$

An instability exists only if  $c_i \neq 0$ . We thus see that if the sum of the integrals does not vanish, then there can be no instability. That is, a sufficient condition for baroclinic stability of a zonal

geostrophic flow,  $u_b(y, z)$ , is that the sum of the integrals is nonzero. Conversely, a necessary condition for baroclinic instability of the zonal flow is that the sum of the three integrals in equation (14.149) vanishes. This necessary condition is known as the *Charney-Stern-Pedlosky* condition ([Charney and Stern, 1962](#); [Pedlosky, 1964](#)).

In summary, the necessary conditions for instability are that the following three functions must not have the same sign everywhere

$$\partial_y q_b \quad \text{must change sign within } \bar{\eta}_b < z < \bar{\eta} \quad (14.150a)$$

$$f_o \partial_y u_b = -\partial_y b_b \quad \text{must change sign along } z = \bar{\eta} \quad (14.150b)$$

$$(dz/dy)_{b_b} - \partial_y \eta_b \quad \text{must change sign along } z = \bar{\eta}_b. \quad (14.150c)$$

Conversely, if each of these conditions fails, then that is sufficient to conclude that the flow is stable to baroclinic instability.

#### 14.8.4 Necessary conditions based on the real part

Following from Fjørtoft's theorem for shear instability in Section 13.5.2, consider the real part of equation (14.147), which says

$$\begin{aligned} \Gamma = & \int \frac{(u_b - c_r) |\tilde{\psi}|^2 \partial_y q_b}{|u_b - c|^2} dy dz \\ & + \int_{z=\bar{\eta}} \frac{(u_b - c_r) f_o^2 |\tilde{\psi}|^2 \partial_z u_b}{N^2 |u_b - c|^2} dy - \int_{z=\bar{\eta}_b} \frac{(u_b - c_r) f_o |\tilde{\psi}|^2 [(dz/dy)_{b_b} - \partial_y \eta_b]}{|u_b - c|^2} dy, \end{aligned} \quad (14.151)$$

where we introduced the non-negative number

$$\Gamma = \int \left[ k_x^2 |\tilde{\psi}|^2 + |\partial_y \tilde{\psi}|^2 + (f_o/N)^2 |\partial_z \tilde{\psi}|^2 \right] dy dz \geq 0. \quad (14.152)$$

Given that the condition (14.149) is maintained (with  $c_i \neq 0$ ), then we can replace  $c_r$  in equation (14.151) with an arbitrary constant, referred to as  $U_s$  in Fjørtoft's theorem from Section 13.5.2. We conclude that a sufficient condition for stability is if there is any constant,  $U_s$ , whereby all of the following conditions hold

$$(u_b - c_r) \partial_y q_b < 0 \quad \bar{\eta}_b < z < \bar{\eta} \quad (14.153a)$$

$$(u_b - U_s) \partial_z u_b < 0 \quad z = \bar{\eta} \quad (14.153b)$$

$$-(u_b - U_s) [(dz/dy)_{b_b} - \partial_y \eta_b] < 0 \quad z = \bar{\eta}_b. \quad (14.153c)$$

#### 14.8.5 Necessary condition for instability of the Eady model

To help understand features of the necessary condition (14.149), and the summary statements (14.150a)–(14.150c), we consider some special cases, starting with the Eady model from Sections 14.6 and 14.7. In this case, there is a constant interior potential vorticity of the background state, so that  $\partial_y q_b = 0$ . Hence, the necessary condition for instability (14.149) reduces to a condition on the boundary integrals

$$\int_{z=\bar{\eta}} \frac{|\tilde{\psi}|^2 \partial_z u_b}{N^2 |u_b - c|^2} dy = \int_{z=\bar{\eta}_b} \frac{|\tilde{\psi}|^2 \partial_z u_b}{N^2 |u_b - c|^2} dy, \quad (14.154)$$

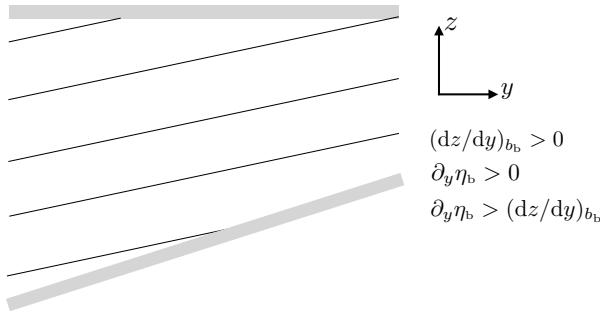


FIGURE 14.6: The Eady model with constant sloping buoyancy surfaces and with a bottom slope that is steeper than the buoyancy surfaces. This flow state is baroclinically stable according to equation (14.157) since the slope of the topography is steeper than the buoyancy surface.

where we set  $\partial_y \eta_b = 0$  as per the Eady model, and the buoyancy slope is

$$(dz/dy)_b = -\partial_y b_b / \partial_z b_b = f_o \partial_z u_b / N^2. \quad (14.155)$$

Furthermore,  $N^2$  and  $\partial_z u_b$  are constants in the Eady model, and the zonal flow is a function only of vertical,  $u_b(z)$ , so that the necessary condition for instability (14.154) reduces to

$$\left[ \frac{|\tilde{\psi}|^2}{|u_b - c|^2} \right]_{z=\bar{\eta}} = \left[ \frac{|\tilde{\psi}|^2}{|u_b - c|^2} \right]_{z=\bar{\eta}_b}. \quad (14.156)$$

For this condition to hold requires the streamfunction to be nonzero at both boundaries. We are ensured that the streamfunction is not exponentially small at the boundaries when the edge waves have small enough wavenumber that the waves can overlap and have a nontrivial presence at both boundaries

#### 14.8.6 Effects from adding a bottom slope to the Eady model

Consider the Eady model but with a sloping bottom, in which case the necessary condition for instability, equation (14.149), becomes

$$\int_{z=\bar{\eta}} \frac{|\tilde{\psi}|^2 (dz/dy)_b}{|u_b - c|^2} dy = \int_{z=\bar{\eta}_b} \frac{|\tilde{\psi}|^2 [(dz/dy)_b - \partial_y \eta_b]}{|u_b - c|^2} dy. \quad (14.157)$$

For the Eady model the buoyancy slope is constant, and let us assume it is positive. If the bottom topography slope has the same sign as the buoyancy slope, but is larger in magnitude, then the right hand side of equation (14.157) is negative whereas the left hand side is positive. We reach a contradiction, which means this necessary condition for instability cannot be satisfied. Evidently, the flow is baroclinically stable if the bottom slope is steeper than the buoyancy slope. We depict this situation in Figure 14.6.

#### 14.8.7 Flat bottom with constant buoyancy along the two boundaries

The Eady model has horizontally varying buoyancy along the top and bottom boundaries, and stability of the Eady waves is determined by the boundary buoyancy. But what if there is a constant buoyancy along the two boundaries (so  $(dz/dy)_b = 0$ ) and with the bottom boundary flat ( $\partial_y \eta_b = 0$ )? In this case there are no topographic waves nor Eady waves along

the boundaries, and both boundary integrals in the necessary condition (14.149) vanish. The necessary condition for instability (14.149) reduces to

$$\int \frac{|\tilde{\psi}|^2 \partial_y q_b}{|u_b - c|^2} dy dz = 0. \quad (14.158)$$

We thus find that in the absence of boundary contributions, the necessary condition for baroclinic instability holds so long as the meridional derivative of the background potential vorticity changes sign somewhere in the domain. This condition is reminiscent of the Rayleigh-Kuo inflection point theorem for horizontal shear instability (Section 13.5).

## 14.9 Energetics of small amplitude fluctuations

In this section we offer another means to help understand the mechanics of quasi-geostrophic waves and baroclinic instability, as well as conditions required for a fluctuation to grow. Elements of this section are shared with Section 14.8, yet here we focus on the mechanical energy budget of the fluctuating fields. As with Section 14.8, the results here are more general than those found for the Eady model in Section 14.7, since we allow for the general linear theory from Sections 14.3 and 14.8.

Recall the energetics of quasi-geostrophy as studied in Section ??, where we identify the mechanical energy per mass of the fluctuations

$$E = [\nabla_h \psi' \cdot \nabla_h \psi' + (f_o \partial_z \psi' / N)^2] / 2, \quad (14.159)$$

which is the sum of the kinetic energy per mass

$$\nabla_h \psi' \cdot \nabla_h \psi' / 2 = \mathbf{u}' \cdot \mathbf{u}' / 2, \quad (14.160)$$

plus the available potential energy per mass

$$(f_o \partial_z \psi' / N)^2 / 2 = (b' / N)^2 / 2. \quad (14.161)$$

To derive the energy budget equation, we follow the standard procedure in quasi-geostrophy by multiplying the potential vorticity equation (14.25) by  $\psi'$

$$\psi' (\partial_t + u_b \partial_x) [\nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi')] = -\psi' \partial_x \psi' \partial_y q_b. \quad (14.162)$$

We now move derivatives around, with the manipulations relatively straightforward with  $u_b = 0$ , but somewhat more tedious with  $u_b \neq 0$ .

### 14.9.1 Time derivative terms

The time derivative can be written

$$\psi' \partial_t (\nabla_h^2 \psi') = \psi' \nabla_h \cdot \nabla_h (\partial_t \psi') \quad (14.163a)$$

$$= \nabla_h \cdot [\psi' \nabla_h (\partial_t \psi')] - \nabla_h \psi' \cdot \nabla_h (\partial_t \psi') \quad (14.163b)$$

$$= \nabla_h \cdot [\psi' \nabla_h (\partial_t \psi')] - \partial_t (\nabla_h \psi' \cdot \nabla_h \psi') / 2, \quad (14.163c)$$

along with

$$\psi' \partial_t \partial_z [(f_o^2/N^2) \partial_z \psi'] = \psi' \partial_z [(f_o^2/N^2) \partial_{zt} \psi'] \quad (14.164a)$$

$$= \partial_z [\psi' (f_o^2/N^2) \partial_{zt} \psi'] - \partial_z \psi' (f_o^2/N^2) \partial_{zt} \psi' \quad (14.164b)$$

$$= \partial_z [\psi' (f_o^2/N^2) \partial_{zt} \psi'] - \partial_t (f_o \partial_z \psi' / N)^2 / 2, \quad (14.164c)$$

which yields

$$\psi' \partial_t [\nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi')] = -\partial_t E + \nabla_h \cdot [\psi' \nabla_h (\partial_t \psi')] + \partial_z [\psi' (f_o^2/N^2) \partial_{zt} \psi']. \quad (14.165)$$

### 14.9.2 Advection by the background zonal flow

To account for advection by the zonal background geostrophic flow, it is important to remember that  $u_b = u_b(y, z)$ , so that

$$\psi' u_b \partial_x (\nabla_h^2 \psi') = \psi' u_b \nabla_h \cdot \nabla_h (\partial_x \psi') \quad (14.166a)$$

$$= \nabla_h \cdot [\psi' u_b \nabla_h (\partial_x \psi')] - \nabla_h (\psi' u_b) \cdot \nabla_h (\partial_x \psi') \quad (14.166b)$$

$$= \nabla_h \cdot [\psi' u_b \nabla_h (\partial_x \psi')] - \psi' \nabla_h u_b \cdot \nabla_h (\partial_x \psi') - u_b \nabla_h \psi' \cdot \nabla_h (\partial_x \psi') \quad (14.166c)$$

$$= \nabla_h \cdot [\psi' u_b \nabla_h (\partial_x \psi')] - \psi' \partial_y u_b \partial_{xy} \psi' - u_b \partial_x (\nabla_h \psi' \cdot \nabla_h \psi') / 2. \quad (14.166d)$$

One further step is key, whereby we write

$$\psi' \partial_y u_b \partial_{xy} \psi' = \partial_x (\psi' \partial_y u_b \partial_y \psi') - \partial_x \psi' \partial_y u_b \partial_y \psi', \quad (14.167)$$

so that

$$\begin{aligned} \psi' u_b \partial_x (\nabla_h^2 \psi') &= \nabla_h \cdot [\psi' u_b \nabla_h (\partial_x \psi')] - \partial_x (\psi' \partial_y u_b \partial_y \psi') + \partial_x \psi' \partial_y u_b \partial_y \psi' \\ &\quad - u_b \partial_x (\nabla_h \psi' \cdot \nabla_h \psi') / 2. \end{aligned} \quad (14.168)$$

A similar set of steps for the stretching term (temporarily dropping the constant  $f_o^2$  for brevity) leads to

$$\psi' u_b \partial_x \partial_z (N^{-2} \partial_z \psi') = \psi' u_b \partial_z (N^{-2} \partial_{xz} \psi') \quad (14.169a)$$

$$= \partial_z [\psi' u_b N^{-2} \partial_{xz} \psi'] - \partial_z (\psi' u_b) N^{-2} \partial_{xz} \psi' \quad (14.169b)$$

$$= \partial_z [\psi' u_b N^{-2} \partial_{xz} \psi'] - \psi' \partial_z u_b N^{-2} \partial_{xz} \psi' - u_b \partial_z \psi' N^{-2} \partial_{xz} \psi' \quad (14.169c)$$

$$= \partial_z [\psi' u_b N^{-2} \partial_{xz} \psi'] - \psi' \partial_z u_b N^{-2} \partial_{xz} \psi' - u_b \partial_x (\partial_z \psi' / N)^2 / 2. \quad (14.169d)$$

As for the relative vorticity terms, write

$$\psi' \partial_z u_b N^{-2} \partial_{xz} \psi' = \partial_x (\psi' \partial_z u_b N^{-2} \partial_z \psi') - \partial_x \psi' \partial_z u_b N^{-2} \partial_z \psi', \quad (14.170)$$

which then leads to

$$\begin{aligned} \psi' u_b \partial_{xz} (f_o^2 N^{-2} \partial_z \psi') &= \partial_z (\psi' u_b f_o^2 N^{-2} \partial_{xz} \psi') - \partial_x (\psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi') \\ &\quad + \partial_x \psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi' - u_b \partial_x (\partial_z \psi' / N)^2 / 2. \end{aligned} \quad (14.171)$$

### 14.9.3 Summary of the energy equation

Bringing terms together gives

$$\begin{aligned} 0 &= \psi' (\partial_t + u_b \partial_x) [\nabla_h^2 \psi' + f_o^2 \partial_z (N^{-2} \partial_z \psi')] + \psi' \partial_x \psi' \partial_y q_b \\ &= -(\partial_t + u_b \partial_x) E - \nabla_h \cdot \mathbf{F}^h - \partial_z F^z + \partial_x \psi' \partial_y u_b \partial_y \psi' + \partial_x \psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi', \end{aligned} \quad (14.172)$$

where we introduced the horizontal and vertical components to the energy flux vector

$$\mathbf{F}^h = -\psi' (\partial_t + u_b \partial_x) \nabla_h \psi' + \hat{\mathbf{x}} \psi' (-\psi' \partial_y q_b / 2 + \partial_y u_b \partial_y \psi' + \partial_z u_b f_o^2 N^{-2} \partial_z \psi') \quad (14.173a)$$

$$F^z = -\psi' f_o^2 N^{-2} (\partial_t + u_b \partial_x) \partial_z \psi' = -\psi' f_o N^{-2} (\partial_t + u_b \partial_x) b', \quad (14.173b)$$

which gives the energy equation

$$(\partial_t + u_b \partial_x) E = -\nabla_h \cdot \mathbf{F}^h - \partial_z F^z + \partial_x \psi' \partial_y u_b \partial_y \psi' + \partial_x \psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi'. \quad (14.174)$$

Since  $u_b$  is independent of  $x$ , one may choose to place the zonal advection term inside of the horizontal flux by noting that

$$u_b \partial_x E = \partial_x (u_b E), \quad (14.175)$$

so that

$$\partial_t E = -\nabla_h \cdot (E u_b \hat{\mathbf{x}} + \mathbf{F}^h) - \partial_z F^z + \partial_x \psi' \partial_y u_b \partial_y \psi' + \partial_x \psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi'. \quad (14.176)$$

### 14.9.4 Horizontal and thermal wind shear production

Equation (14.176) indicates that time changes to the mechanical energy contained in the small amplitude fluctuations is driven by the convergence of an energy flux, plus two source terms arising from shears in the background zonal geostrophic flow. The flux convergence acts to move energy around, and its domain integral provides possible sources for energy at the boundaries. We do not study this term in this section since it is quite dependent on assumptions about the boundaries, though we comment on this limitation in the analysis in Section 14.9.7.

Instead, we focus here on the two source terms in equation (14.176). The source arising from horizontal shear can be written

$$\partial_x \psi' \partial_y u_b \partial_y \psi' = -v' u' \partial_y u_b, \quad (14.177)$$

and the source arising from vertical shear is

$$\partial_x \psi' \partial_z u_b f_o^2 N^{-2} \partial_z \psi' = v' b' (f_o \partial_z u_b / N^2) = v' b' (dz/dy)_{b_b}, \quad (14.178)$$

where the final step introduced the slope of the background buoyancy surface according to equation (14.12).

#### Horizontal shear production and thermal wind shear production

As in the study of horizontal shear instability in Section 13.4.6, we identify

$$-v' u' \partial_y u_b = \text{horizontal shear production}. \quad (14.179)$$

This shear production term is generally rather small for geostrophic flows in comparison to the ageostrophic flows considered in Chapter 13. In analogy, we introduce

$$v' b' (dz/dy)_{b_b} = -v' b' \partial_y b_b / N^2 = \text{thermal wind shear production.} \quad (14.180)$$

The thermal wind shear production is fundamentally distinct from horizontal shear production. The reason is that  $v' b' (dz/dy)_{b_b}$  arises from the potential energy in the background state that supports the thermal wind shear, rather than from the background kinetic energy that supports horizontal (or vertical) shear production described in Section 13.4.6. For quasi-geostrophic flows, kinetic energy sourced shear production is generally far smaller than potential energy sourced thermal wind shear production.

#### 14.9.5 Meridional and vertical eddy buoyancy fluxes

As shown in the following, fluctuations that increase their energy through making use of the prescribed background potential energy have a tendency to flux buoyancy down the meridional gradient and upward. These two effects act in a manner that increases the total quasi-geostrophic mechanical energy (sum of kinetic plus available potential) of the fluctuations, while feeding off the available potential energy of the background flow.

##### **Meridionally downgradient flux of buoyancy increases mechanical energy of the fluctuations**

Mechanical energy of the fluctuations increases through the thermal wind shear production (14.180) if the meridional *eddy buoyancy flux* is down the meridional gradient of the background buoyancy

$$v' b' \partial_y b_b = -v' b' f_o \partial_z u_b < 0 \quad \text{mechanical energy of fluctuations increases.} \quad (14.181)$$

In practice, this downgradient behavior occurs over a space and/or time and/or ensemble average, in which we use an overbar as a generic average so that

$$\overline{v' b'} \partial_y b_b = -\overline{v' b'} f_o \partial_z u_b < 0 \quad \text{averaged mechanical energy of fluctuations increases.} \quad (14.182)$$

This downgradient meridional buoyancy flux leads to a poleward heat transport for flows where temperature dominates buoyancy. Furthermore, the poleward heat transport extracts potential energy from the background flow since it acts in a direction that flattens the background buoyancy surfaces. Such fluctuations increase their energy at the expense of the available potential energy of the background state.

##### **Fluid particle motion leading to an increase in available potential energy of the fluctuations**

Available potential energy contained in the fluctuations evolves according to

$$(\partial_t + u_b \partial_x)[(f_o \partial_z \psi' / N)^2 / 2] = (\partial_t + u_b \partial_x)(b'/N)^2 / 2 \quad (14.183a)$$

$$= (b'/N^2)(\partial_t + u_b \partial_x)b' \quad (14.183b)$$

$$= -b' [w' - v' (dz/dy)_{b_b}] \quad (14.183c)$$

$$= -b' (w' + v' \partial_y b_b / N^2), \quad (14.183d)$$

where we used the linear buoyancy equation (14.27) for the penultimate step. The available potential energy of the fluctuations increases if

$$b' [w' - v' (dz/dy)_{b_b}] = b' (w' + v' \partial_y b_b / N^2) < 0 \quad \text{APE of the fluctuations increases.} \quad (14.184)$$

An increasing available potential energy in the fluctuations is expected for a growing disturbance, in which fluctuating buoyancy surfaces have growing amplitudes. The condition (14.184) takes on a geometric expression by cancelling the buoyancy fluctuation and writing

$$w'/v' < (dz/dy)_{b_b} \implies \text{APE of the fluctuations increases,} \quad (14.185)$$

whereas the fluctuations maintain a fixed available potential energy if the meridional and vertical velocity components align with the background buoyancy surfaces

$$w'/v' = (dz/dy)_{b_b} \implies \text{APE of the fluctuations remains constant.} \quad (14.186)$$

We have more to say on these geometric conditions in Section 14.9.7.

### Kinetic energy of the fluctuations

Substituting the available potential energy equation (14.183d) into the mechanical energy equation (14.174) leads to the equation for the kinetic energy of the fluctuations

$$(\partial_t + u_b \partial_x)(\mathbf{u}' \cdot \mathbf{u}'/2) = -\nabla_h \cdot \mathbf{F}^h - \partial_z F^z - v' u' \partial_y u_b + w' b'. \quad (14.187)$$

Growth in the kinetic energy of the fluctuations arises if positive buoyancy anomalies are fluxed upward,

$$w' b' > 0 \quad \text{kinetic energy of the fluctuations increases.} \quad (14.188)$$

This flux lowers the center of mass for the fluid and so decreases the potential energy.

### Summary of the inequalities

We summarize the considerations thus far by noting that the mechanical energy, kinetic energy, and available potential energy of the fluctuating field increases for fluid particle displacements that are sloped between the horizontal plane and the constant buoyancy surface passing through the origin as in the *wedge of instability* in Figure 14.7

$$w'/v' \leq (dz/dy)_{b_b} \implies N^2 w' b' < -v' b' \partial_y b_b. \quad (14.189)$$

Recall that  $w' b' > 0$  means that kinetic energy in the fluctuating fields increases (equation (14.188)), and  $v' b' \partial_y b_b < 0$  means that the mechanical energy increases (equation (14.181)). Inequality (14.189) thus ensures that available potential energy for the fluctuations also increases.

### 14.9.6 Tilting phase lines of unstable baroclinic waves

In Section 13.4.7 we showed that phase lines of unstable barotropic waves tilt into the meridional shear of the zonal flow, with this orientation allowing the wave perturbations to grow by extracting kinetic energy from the background shear. Here we pursue a similar argument to reveal that unstable baroclinic waves tilt into the vertical shear of the zonal flow, with this

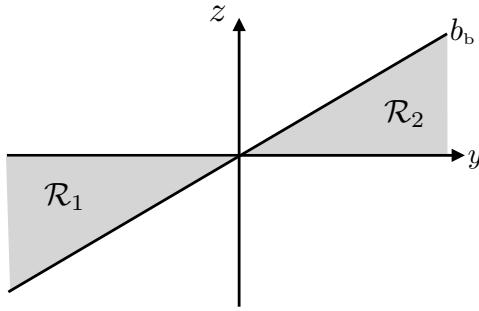


FIGURE 14.7: Depicting the wedge of instability for baroclinic instability. The background buoyancy is oriented with  $\partial_y b_b < 0$ , so that less buoyant fluid sits to the north (to the right), as well as  $\partial_z b_b > 0$ , so that the fluid is stably stratified in the vertical. A buoyancy fluctuation,  $b'$ , in region  $\mathcal{R}_1$  is more buoyant than  $b_b$ , so that  $b' > 0$ , whereas in region  $\mathcal{R}_2$  we find  $b' < 0$  relative to  $b_b$ . Hence, motion of fluid within a fluctuation that moves from region  $\mathcal{R}_1$  to region  $\mathcal{R}_2$  represents motion of relatively buoyant fluid ( $b' > 0$ ) upward ( $w' > 0$ ). This motion thus has  $w' b' > 0$ , which leads to an increase in kinetic energy of the fluctuation (equation (14.188)). This same motion also carries  $v' b' > 0$  so that  $v' b' \partial_z b_b < 0$ , meaning that the mechanical energy of the fluctuation increases (equation (14.181)). Finally, this motion ensures that  $w' b' \leq v' b' (dz/dy)_{b_b}$ , so that the available potential energy of the fluctuation increases (equation (14.189)). Motion from region  $\mathcal{R}_2$  to region  $\mathcal{R}_1$  reverses all signs of the perturbations, so that their products remain the same, thus ensuring that energies increase for fluctuations moving from  $\mathcal{R}_2$  to region  $\mathcal{R}_1$ .

orientation allowing the waves to extract potential energy from the background thermal wind flow.

Start with the phase average of the inequality (14.182) that provides a sufficient condition for wave fluctuations to increase their mechanical energy

$$\langle v' b' \rangle f_o \partial_z u_b > 0 \quad \text{mechanical energy of fluctuations increases.} \quad (14.190)$$

Writing the streamfunction in the form

$$\psi'(x, y, z, t) = \tilde{\psi}(z) e^{i(k_x x + k_y y - \omega_r t) + \omega_i t} = |\tilde{\psi}(z)| e^{i(k_x x + k_y y + \alpha(z) - \omega_r t) + \omega_i t}, \quad (14.191)$$

leads to

$$v' = \partial_x \psi' = i k_x |\tilde{\psi}| e^{i(k_x x + k_y y + \alpha - \omega_r t) + \omega_i t} \quad (14.192a)$$

$$b' = f_o \partial_z \psi' = f_o [\partial_z |\tilde{\psi}| + i |\tilde{\psi}| \partial_z \alpha] e^{i(k_x x + k_y y + \alpha - \omega_r t) + \omega_i t}, \quad (14.192b)$$

and with the phase averaging identity (??) yielding

$$\langle v' b' \rangle = (1/2) \operatorname{Re}[v' b'^*] = (1/2) k_x f_o |\tilde{\psi}|^2 e^{2\omega_i t} \partial_z \alpha. \quad (14.193)$$

We focus on phase lines in the  $x$ - $z$  plane (so that  $k_y = 0$ ), as that is the plane of the zonal flow with a vertical shear and so it is the plane of the most unstable wave as per *Squire's theorem*. We thus consider the spatial phase function

$$\Phi(x, z) \equiv k_x x + \alpha(z), \quad (14.194)$$

whose constant surfaces are defined by

$$d\Phi = 0 = k_x dx + (\partial_z \alpha) dz \implies (dz/dx)_{\text{phase}} = -k_x / \partial_z \alpha, \quad (14.195)$$

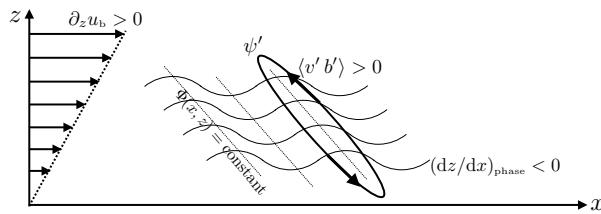


FIGURE 14.8: The lines of constant phase,  $\Phi(x, z) = kx - \alpha(z)$ , for the streamfunction (or the meridional velocity) in a northern hemisphere ( $f_o > 0$ ) baroclinic wave. The wave depicted here has an increasing mechanical energy since the phase lines are tilted into the vertical shear of the zonal background flow, thus allowing for thermal wind shear production to increase the mechanical energy as per equation (14.197). This figure is directly analogous to the case of barotropic shear production as given by Figure 13.1.

so that the instability condition (14.190) takes the form

$$(1/2) k_x^2 f_o^2 |\tilde{\psi}|^2 e^{2\omega_l t} \frac{\partial_z u_b}{(dz/dx)_{\text{phase}}} < 0 \implies \text{sufficient condition for energy growth.} \quad (14.196)$$

Simplifying this equation leads to the condition for the ratio of the vertical shear to the phase slope

$$\frac{\partial_z u_b}{(dz/dx)_{\text{phase}}} < 0 \implies \text{sufficient condition for energy growth.} \quad (14.197)$$

This inequality says that mechanical energy of the wave grows when the wave's phase lines tilt into the background vertical shear, such as depicted in Figure 14.8. This condition is directly analogous to equation (13.56) and Figure 13.1, which hold for unstable barotropic waves on a meridional sheared zonal flow. Here, the tilt of the phase lines reflects the ability of the wave to extract potential energy from the background state, thus leading to growth of energy for the unstable wave. The geometric property (14.197) offers a visual indicator that the wave is acting on a baroclinically unstable background state, thus providing a valuable diagnostic tool for identifying when baroclinic instability is happening.

### 14.9.7 Caveats for extending the wedge of instability to parcels

Inequality (14.189) refers to the orientation of fluid particles affected by small amplitude wave fluctuations that lead to baroclinic instability, as illustrated in Figure 14.7. Even so, this figure is the basis for a parcel argument that goes beyond that of a small amplitude wave argument, with the parcel argument presented in many texts, such as Section 7.6 of [Pedlosky \(1987\)](#), Lecture 20 of [Pedlosky \(2003\)](#), Section 13.15 of [Kundu et al. \(2016\)](#), and Section 9.4.1 of [Vallis \(2017\)](#), and originating from [Eady \(1949\)](#). Here we offer some caveats about this argument.

#### Baroclinic instability versus symmetric instability

The parcel arguments associated with the wedge of instability do not represent necessary conditions for baroclinic instability, and so they are no substitute for the integral statements in Section 14.8. This situation contrasts to parcel arguments considered for symmetric instability, where the necessary and sufficient condition for symmetric instability is summarized by its wedge of instability shown in Figure 11.8.

Furthermore, the parcel argument used for symmetric instability is based on a force balance, and so it incorporates accelerations from both buoyancy and Coriolis to account for the vertical and horizontal motion of the parcels. In contrast, the wedge of instability in Figure 14.7 is

based on energetic arguments for baroclinic instability, with Coriolis acceleration absent from such arguments.

### Baroclinic instability is a wave instability not a local instability

As explored in our study of the Eady model in Section 14.7, baroclinic instability mechanistically occurs through wave resonance. The realization of wave resonance relies on boundary conditions and phase locking of waves so that they can mutually reinforce one another. This mechanism is directly analogous to that considered for horizontal shear instability in Chapter 13.

In contrast, parcel arguments, such as used for symmetric instability, do not know about boundary conditions or phase locking. Instead, they are only concerned with local environmental conditions. Consider the following case in point for limitations of parcel arguments for baroclinic instability. Namely, observe that the wedge of instability does not distinguish between flat bottom boundary or sloping bottom boundary. However, in Section 14.8.6 we found that the Eady model is stable in the presence of a bottom slope that is steeper than the buoyancy surfaces (Figure 14.6).

### Additional points

Focusing specifically on the arguments leading up to inequality (14.189), note that these arguments have ignored all contributions from the convergence of energy fluxes that appear in the energy equation (14.174). These fluxes, particularly at domain boundaries, provide further influences on the domain integrated energetics and hence on stability of the flow (e.g., boundary terms are critical for the necessary instability conditions in Section 14.8). Additional caveats are raised by [Heifetz et al. \(1998\)](#) related to the problems with ignoring pressure fluctuations acting on the fluid particles. These caveats are analogous to those raised in Section ?? when studying effective buoyancy.

### 14.9.8 Further reading

Chapter 17 of [Cushman-Roisin and Beckers \(2011\)](#) and Section 3 of [Held \(2019\)](#) offer qualifiers similar to those raised in Section 14.9.7 regarding the conceptual limitations of parcel arguments for describing the mechanism of instabilities occurring via wave resonance.



## **Part III**

### **End matter**



---

## Appendix A

### GLOSSARY OF CONCEPTS AND TERMS

**action** The action is the time integral of the Lagrangian density. Hamilton’s principle states that the functional variation of the action is stationary for physically realized time evolution.

[xi](#)

**advection equation** The advection equation,  $(\partial_t + \mathbf{v} \cdot \nabla) C = 0$ , is the canonical transport equation that arises from an Eulerian perspective on the evolution of a scalar field in the presence of fluid flow. We encounter the advection equation throughout this book.

**angular frequency** The angular frequency,  $\omega$ , is related to the wave period via  $2\pi/\omega$ , which is a non-negative number. It measures the number of times the wave phase changes by  $2\pi$  during a single wave period. In this book, we always assume  $\omega \geq 0$  since the wave period is non-negative. For unstable waves, the angular frequency is complex, with the real part related to the period and imaginary part determining the exponential growth and/or decay of the wave.

**ansatz** Ansatz is a German word meaning educated guess, trial form, or assumed function form. We use the term when making a guess for the mathematical structure for the solution to a differential equation. Most notably, we make the plane wave ansatz when seeking the dispersion relation connecting the space and time patterns within a wave.

**barystatic sealevel** When freshwater enters the ocean, such as from melting continental ice sheets, it adds to the ocean mass and in turn increases global mean sea level. This change is referred to as barystatic sealevel change according to the sea level terminology paper from [Gregory et al. \(2019\)](#). Although ocean salinity changes upon changing the freshwater content, the net effect on global mean sea level is almost entirely barystatic since the global halosteric effect is negligible. We can understand why the global halosteric effect is so tiny by recognizing that freshwater entering the ocean sees its salinity increase whilst the ambient seawater is itself freshened. These compensating salinity changes (which are often mistakenly ignored) have corresponding compensating sea level changes, thus bringing the global halosteric effect to near zero.

**body forces** A body force acts throughout the extent of a fluid element, and it is synonymous with external force. Examples include the gravitational force, as well as the Coriolis and centrifugal forces arising from the rotating planetary reference frame. These forces are also known as long range forces.

**budget equations** Much of this book is concerned with deriving and understanding equations that describe the evolution of fluid properties, with such equations (differential or integral)

derived from physical principles such as Newton's laws of motion, Hamilton's principle of stationary action, Noether's theorem, thermodynamic laws, mass conservation, and vorticity mechanics. These are the budget equations that form the theoretical foundation of continuum mechanics. [ix](#)

**carrier wave** Carrier waves are the plane wave modes that comprise a wave packet. From an engineering perspective, the carrier wave carries the most useful information contained in the wave packet signal, hence the name. [17](#), [31](#)

**causal free space Green's function** The causal free space Green's function arises in evolutionary partial differential equations, such as the diffusion equation and wave equation. It refers to the solution in the absence of spatial boundary conditions (other than decay at spatial infinity) and with zero values for times before the initial time. [22](#)

**conjugate symmetry** Conjugate symmetry refers to the identity holding for Fourier transform amplitude functions,  $\mathcal{A}(\mathbf{k}) = \mathcal{A}^*(-\mathbf{k})$ , whereas conjugate anti-symmetry holds when  $\mathcal{A}(\mathbf{k}) = -\mathcal{A}^*(-\mathbf{k})$ . [21](#)

**contact force** A contact force acts on the boundary of a fluid element, with examples including stresses from pressure and from friction. Contact forces are local forces. Contact forces are sometimes referred to as internal forces, since they arise from local interactions internal to the fluid, as distinct from body forces that arise from long range external forces that act throughout the body of a fluid element. Contact forces are also called tractions in some areas of continuum mechanics. Contact forces are molecular in origin, though we are unconcerned in this book with details of the molecular dynamics leading to these forces. Contact forces act on a region of a continuous media through the area integrated stresses acting on the boundary enclosing the region. [xix](#)

**diagnostic equation** A diagnostic equation determines the value of a field at a particular time instance. An example is the non-divergence condition,  $\nabla \cdot \mathbf{v} = 0$ , satisfied by velocity in a Boussinesq ocean. There are generally no time derivatives appearing in diagnostic equations, though this property is generally a function of the chosen coordinate system. [xi](#)

**Dirac delta** The Dirac delta provides an idealization of a point source,  $\delta(\mathbf{x})$ , and it is formally infinite when evaluated at  $\mathbf{x} = 0$  whereas it vanishes at all other points. The Dirac delta plays a central role in the theory of Green's functions. When multiplied by mass, it provides the mass density for a point particle. In mathematics, the Dirac delta is known as a generalized function or a distribution. [32](#), [41](#)

**dispersion** Dispersion of waves refers to the presence of distinct wave speeds as a function of wavenumber, so that a packet of waves spread and change shape as it propagates. Mathematically, wave dispersion is reflected in a nonzero  $\partial_{k_j} \mathbf{c}_g$ . Dispersion is a property inherent in the waves and has nothing to do with dissipation. [24](#), [25](#)

**dispersion relation** Forces acting in a fluid determine the dispersion relation satisfied by linear waves, with the dispersion relation connecting the wave angular frequency,  $\omega$ , to the wavevector,  $\mathbf{k}$ . Hence, the dispersion relation couples the space and time structure of a wave so that once a wavevector is chosen then the angular frequency is specified. [2](#), [5](#)

**dispersion tensor** The dispersion tensor is determined by the second derivative of the dispersion relation evaluated at the carrier wavevector,  $\mathbf{k}_0$ , so that  $\mathbb{K}^{ab} = \partial^2 \varpi / \partial k_a \partial k_b |_{\mathbf{k}=\mathbf{k}_0}$ . The dispersion tensor is symmetric, as appropriate for a diffusion tensor. However, there is no guarantee that it is positive-definite. Hence, dispersion can both broaden (as for diffusion) or sharpen (as for anti-diffusion) the wave packet. 25

**dispersive waves** Dispersive waves are characterized by phases whose speeds are a function of the wavenumber, so that the phase velocity is distinct from the group velocity. As a result, a localized disturbance that initially has many wavelengths will spread out or change shape as it propagates. Most waves we study in this book are dispersive, such as surface gravity waves, internal gravity waves, inertia waves, and Rossby waves. 10, 34

**domain of influence** For a wave disturbance acting at a point in space-time, the domain of influence is the region of space-time that can be affected by the disturbance, given the propagation speed of the wave. 23

**elements pillar** The elements pillar of geophysical fluid mechanics comprises the physical and mathematical formulation of conceptual models used to garner insight into rotating and stratified fluid motion. This pillar is concerned with setting the stage by deductively and descriptively exposing how physical concepts are mathematically expressed to describe geophysical fluid flows. ix

**emergent phenomena pillar** The emergent phenomena pillar of geophysical fluid mechanics studies solutions to equations that describe phenomena, such as waves, instabilities, turbulence, and general circulation, all of which emerge from the fundamental equations based on first principles. These phenomena can emerge in manners that are far from simple to understand deductively, particularly when considering nonlinear behavior such as turbulence. ix

**emergent scale** There are two general types of dimensional scales that we use to non-dimensionalize a mathematical physics equation. One is the emergent scale, which emerges from the flow itself. Emergent scales, such as the length scale and velocity scale of the flow, are specified by the subjective interest of the theorist though these scales are not under direct control. That is, we choose to focus on flows with a particular scale for purposes of examining the corresponding equations that describe that flow regime. A key example concerns our study of planetary geostrophy and quasi-geostrophy, where we choose to focus on flows of a particular scale where the Coriolis acceleration is of leading order importance. xii

**Euler identity** The Euler identity is  $e^{-i\omega t} = \cos(\omega t) - i \sin(\omega t)$ . Within this book, the Euler identity is used throughout Fourier analysis and wave mechanics. 11

**evolution equation** An evolution equation determines the time tendency (Eulerian evolution) of a quantity such as the temperature or velocity. Terms in the prognostic equation are referred to as time tendencies. Evolution equations are also referred to as prognostic equations. x

**external scale** There are two general types of dimensional scales that we use to non-dimensionalize a mathematical physics equation. One is the external scale, with examples in this book being the gravitational acceleration, Coriolis parameter, and specified background or reference state. External scales are set by the geophysical parameter regime in which the

flow occurs, and as such they are under direct control of the theorist or experimentalist. The other scale is emergent, and is a property of the flow. [xii](#)

**Fourier analysis** Fourier analysis is the representation of a field or signal as a superposition of sinusoidal basis functions (sines, cosines, or complex exponentials), each with a definite wavenumber (or frequency) and amplitude. We make use of Fourier analysis in our study of wave mechanics throughout this book. [16](#)

**Fresnel integral** Fresnel integrals are oscillatory integrals with a quadratic phase. They describe how waves with curved phase fronts interfere, and they are central to near-field diffraction, wave propagation, and quantum free-particle evolution. [40](#)

**generalized vertical coordinate** A generalized vertical coordinate,  $\sigma$ , has a one-to-one invertible relation with the geopotential vertical coordinate,  $z$ , so that  $\sigma = \sigma(x, y, z, t)$ , and yet this coordinate is typically not orthogonal to the horizontal Cartesian coordinates. Generalized vertical coordinates are commonly used as the basis for numerical ocean and atmosphere models, and frequently used for theoretical formulations. [15](#)

**geophysical fluid dynamics** Geophysical fluid dynamics is a branch of fluid mechanics concerned with natural fluid motion on a rotating and gravitating body such as a planet or star. [ix](#)

**geophysical fluid mechanics** Geophysical fluid mechanics is a branch of theoretical physics concerned with natural fluid motion on a rotating and gravitating body such as a planet or star, making use of concepts and methods from classical continuum mechanics and thermodynamics. [vii](#)

**geostrophic balance** The geostrophic balance is a diagnostic balance between the pressure gradient acceleration and the Coriolis acceleration. It is well maintained for the large-scale and low frequency middle to high latitude motions in the atmosphere and ocean. Geostrophic balance does not hold near the equator, since the Coriolis parameter vanishes there. [xi](#)

**global instabilities** Global fluid instabilities arise from the constructive interference of waves and so involve the solution of an eigenvalue problem to determine properties of unstable waves. At most, a necessary condition can be derived to determine whether a global instability exists. Global instabilities are also referred to as wave instabilities. [xxi](#)

**Green's function** The Green's function is the formal inverse of a linear differential operator. Knowledge of the Green's function allows one to write the solution to a linear differential equation as an integral, with the Green's function acting as a kernel. [22, 132](#)

**group velocity** The group velocity is the velocity of the energy within the wave disturbance. It is distinct from the wave phase velocity for dispersive waves, but equal to the phase velocity for non-dispersive waves. [18, 34](#)

**harmonic waves** Harmonic waves have a regular sinusoidal-like spatial pattern. [7](#)

**Heaviside step function** The Heaviside step function,  $\mathcal{H}(x)$ , also known as the unit step function, is a discontinuous non-dimensional mathematical function that outputs 0 for negative input values and unity for positive input values. It takes values  $\mathcal{H}(\tau) = 0$  for  $\tau < 0$

and  $\mathcal{H}(\tau) = 1$  for  $\tau > 0$ . The derivative of a Heaviside step function is the Dirac delta:  $d\mathcal{H}/dx = \delta(x)$ . Note that we do not define the Heaviside at  $x = 0$ , though some authors give it a value of  $\mathcal{H}(0) = 1/2$ . For our purposes, the properties of the Heaviside step function remain unchanged whether it is defined at  $x = 0$  or not. See footnote on page 20 of *Stakgold* (2000a) for more details. 22

**Heisenberg uncertainty principle** The Heisenberg uncertainty principle is the quantum wave mechanical realization of the uncertainty relation holding in Fourier analysis. That is, the Heisenberg uncertainty principle is the Fourier-based statement that non-commuting observables cannot be sharply known at the same time. The uncertainty reflects physics, not measurement flaws. 31

**hydrodynamics** A branch of fluid mechanics concerned with the flow of a homogeneous (constant density) incompressible fluid. vii

**hydrostatic balance** The hydrostatic balance is a diagnostic balance between the vertical pressure gradient force and the weight of fluid. The exact hydrostatic balance holds for a static fluid in a gravity field. The approximate hydrostatic balance holds quite well for moving fluids with scales of motion such that the vertical scales are far smaller than the horizontal scales. xi

**hyperbolic partial differential equation** A hyperbolic partial differential equation a second order PDE whose solutions have finite propagation speed, can exhibit oscillations, and typically require initial conditions in time. Disturbances propagate with finite speed along characteristic curves (or surfaces). 9, 32

**internal gravity waves** An internal gravity wave is a transverse wave that is comprised of fluid particles undergoing a simple harmonic oscillation within a continuously and stably stratified buoyancy field. The angular frequency of the oscillation is determined by the buoyancy stratification and the sine of the angle the wave's group velocity makes with respect to the vertical (equivalently, the cosine of the angle the wave's phase velocity makes with horizontal). 249

**irreversible process** A physical process that results in the increase of entropy. Processes that increase the entropy of a fluid particle include the mixing of momentum such as through viscous friction; the mixing of matter such as through the diffusion of constituents in a multi-component fluid; and the mixing of enthalpy (diffusion of heat) in a fluid with variable temperature. vii

**local instabilities** Local fluid instabilities are afforded a local necessary and sufficient condition to determine whether the fluid base state is unstable to perturbations. Gravitational instability provides the canonical example, along with centrifugal and symmetric instabilities. Local instabilities are also referred to as parcel instabilities. xxi

**material time derivative** The material time derivative measures the evolution of a fluid property along a trajectory defined by a fluid particle. This time derivative is also referred to as the Lagrangian or substantive time derivative. 14

**method of stationary phase** The method of stationary phase is an asymptotic technique that identifies which Fourier components dominate a rapidly oscillatory integral. Contributions come mainly from points where the phase is stationary, leading directly to ray paths and group-velocity propagation in wave physics. 38

**modulation function** A modulation function is used to localize a wave train into a wave packet. 17

**momentum based viewpoint** Determining the forces, either directly or indirectly, provides physical insight into the cause of fluid flow and its changes. This approach is referred to a momentum based viewpoint since it is based on working directly with the momentum equation (i.e., Newton's second law of motion). This viewpoint is distinct from a vorticity viewpoint whereby the primary concern is with terms contributing to the evolution of vorticity. x

**monochromatic pattern** A monochromatic pattern is a feature characterized by having all points maintain the same time periodic motion with single angular frequency,  $\omega$ . 11

**monochromatic wave** A monochromatic wave is characterized by having all points maintain the same time periodic motion with single angular frequency,  $\omega$ . 11

**narrow band** A narrow band wave packet is one that has a relatively narrow spread in  $k$ -space. 29

**near inertial waves** Inertial waves with  $\omega \approx |f|$  are referred to as near inertial waves, which refers to their angular frequency being close to the Coriolis frequency. Since near inertial waves have their wavevector oriented close to the rotation axis, they also have a vanishingly small group velocity. 158

**non-dimensionalization** Non-dimensionalization is the process of removing all physical dimensions from an equation of motion, and in turn to identify a set of non-dimensional numbers that characterize a particular flow regime. xii

**non-dispersive waves** Non-dispersive waves are have wave phases whose speeds are independent of the wavenumber, so that the phase velocity is identical to the group velocity. Examples include acoustic waves and shallow water gravity waves in a non-rotating reference frame. 9

**parcel instabilities** Parcel instabilities are afforded a local necessary and sufficient condition to determine whether the fluid base state is unstable to perturbations. Gravitational instability provides the cancial example, along with centrifugal and symmetric instabilities. Parcel instabilities are also referred to as local instabilities. xxii

**Parseval's identity** Parseval's identity states that the total energy (or variance) of a wave field is the same whether computed it in  $x$ -space or  $k$ -space. It is also referred to as the Bessel-Parseval identity for discrete Fourier transforms and the Parseval-Plancherel formula for continuous Fourier transforms. 41

**perfect fluid** A fluid that flows in the absence of irreversible processes so that the motion is reversible and the specific entropy remains constant following a fluid particle. A

**perfectfluid** is a continuum of infinitesimal material fluid parcels. Some authors use the term *ideal fluid*, but we eschew that term to avoid confusion with *ideal gas*. [vii](#)

**phase** The phase of a wave refers to the argument of the exponential or the trigonometric portion of a linear wave. A phase change by  $2\pi$  radians returns to the same wave pattern. [12](#)

**phase average** A phase average refers to an average of a field computed over the phase of the field. This choice for an averaging operation is particularly relevant when the fluctuating field involves quasi-linear waves. [2](#)

**phase speed** The phase speed,  $C_p$ , is a non-negative number. It is not a vector so that we do not consider components to the phase speed. Rather, we consider components to the phase velocity vector, which is a vector whose direction is given by the wavevector,  $\mathbf{k}$ , and whose non-negative magnitude, along with the angular frequency, determine the phase speed. [9](#), [14](#)

**phase velocity** The phase velocity,  $\mathbf{c}_p$ , is the velocity of a point fixed on a phase surface and traveling in the direction of the wavevector. Its magnitude is the phase speed,  $C_p$ , which is the speed that the phase moves in the direction of the wavevector. [10](#), [14](#)

**plane waves** Plane waves are characterized by a single wavevector and single angular frequency, with the wave exhibiting symmetry in directions perpendicular to the propagation direction (hence the “plane” in its name). Any linear wave, be it acoustic, gravity, Rossby, etc., can be decomposed into a sum or integral of plane waves with a suite of frequencies and wavevectors, and with modulation by an amplitude function. [12](#)

**planetary vorticity** Planetary vorticity refers to the vorticity imparted to every fluid due to its existence in a rotating planetary reference frame. [xix](#)

**potential vorticity** Potential vorticity is a strategically chosen component of the vorticity vector that melds mechanics (vorticity) to thermodynamics (stratification). Material conservation properties of potential vorticity render important constraints on fluid motion, thus promoting it as a primary field in the study of geophysical fluid mechanics. [xix](#)

**prognostic equation** A prognostic equation determines the time tendency (Eulerian evolution) of a quantity such as the temperature or velocity. Terms in the prognostic equation are referred to as time tendencies. A prognostic equation is also referred to as an evolution equation. [x](#), [xi](#)

**real fluid** A fluid whose flow is affected by irreversible processes arising from momentum mixing (nonzero viscous friction); enthalpy mixing (nonzero diffusivity for temperature); matter mixing (nonzero diffusivity of matter constituents); and through sources such as radiation and chemical reactions. The specific entropy increases following a fluid particle moving in a real fluid. [vii](#)

**reduced wavenumber** The reduced wavenumber,  $\bar{k}$ , is the wavenumber divided by  $2\pi$ , so that  $\bar{k} = |\mathbf{k}|/2\pi$ . The reduced wavenumber measures the number of waves per unit length. [16](#)

**Riemann-Legesque lemma** If a signal is integrable and not wildly singular, its Fourier transform vanishes at high frequency since fast oscillations cancel. We make use of this result when studying wave packets. [38](#)

**Rossby waves** Rossby waves are dispersive waves that arise in a rotating reference frame with differential rotation, as for a rotating planet or the  $\beta$ -plane. They are typically studied via the quasi-geostrophic equations, and are sometimes referred to as vorticity waves. Also, long Rossby waves are supported by planetary geostrophic flows, which are Rossby waves in the limit where the horizontal wavenumber vanishes. Rossby waves play an important role in large-scale geophysical flows. 10, 228

**specular reflection** Wave packets, in the geometrical optics approximation, exhibit specular reflection if the angle the incident wave packet makes with the reflecting surface is preserved upon reflection. Rossby waves, electromagnetic waves, and acoustic waves exhibit specular reflection, whereas internal gravity waves and inertial waves exhibit non-specular reflection. 273

**Stokes drift** The Stokes drift refers to the Stokes mean velocity, which is the difference between the Lagrangian mean velocity and the Eulerian mean velocity. 2

**superposition principle** The superposition principle states that the sum of any two solutions to the same linear equation is itself a solution. Correspondingly, the response of a physical system to two sources is the same as the sum of the response of the system to each source applied separately. The superposition principle is the basis for the Green's function method, and for Fourier analysis. 2

**Sverdrup balance** The Sverdrup balance is a diagnostic balance between vertically integrated meridional transport and the wind stress curl. In particular, a positive wind stress curl leads to northward vertically integrated flow. This balance helps to explain the steady equatorward ocean circulation appearing in the eastern portion of middle latitude gyres. xi

**time tendency** Those terms in a prognostic Eulerian equation that contribute to the time evolution are referred to as time tendencies. For prognostic equations, knowledge of the processes contributing to the net time tendency enables a prediction of flow properties. x

**uncertainty relation** The uncertainty relation states that it is not possible to simultaneously specify the position of a wave packet with arbitrary precision in both  $x$ -space and  $k$ -space. That is, precision in one space corresponds to imprecision in the other. In quantum wave mechanics, we refer to the Heisenberg uncertainty principle. 5, 31, 37

**vorticity** Vorticity is the curl of the velocity,  $\omega = \nabla \times \mathbf{v}$ . It plays a leading role in the study of geophysical fluid flows, where it is important to distinguish the relative vorticity,  $\omega = \nabla \times \mathbf{v}$ , from the planetary vorticity,  $2\Omega$ . xix

**vorticity based viewpoint** A variety of vorticity constraints offer the means to deduce flow properties without determining forces, thus prompting the **vorticity based viewpoint** that is distinct from the momentum-based approach, thus prompting the importance of vortex mechanics in geophysical fluid mechanics. x, 476

**wave function** The wave function plays a fundamental role in describing waves. Example wave functions include the velocity potential (acoustic and surface waves), the streamfunction (Rossby waves), and free surface height (shallow water waves). All other dynamical fields can be generated from the wave function, thus allowing us to focus on characterizing how the wave appears through study of the wave function. The wave function has both an  $x$ -space (geographic/height space) representation as well as a  $\mathbf{k}$ -space (wavevector space) representation. These two representations offer complementary characterizations of wave properties, with the transformation between these representations provided by Fourier's integral theorem. 5, 11

**wave instabilities** Wave instabilities arise from the constructive interference of waves and so involve the solution of an eigenvalue problem to determine properties of unstable waves. At most, a necessary condition can be derived to determine whether a wave instability exists. Wave instabilities are also referred to as global instabilities. xxi

**wave kinematics** Wave kinematics is concerned with wave properties arising from the existence of a dispersion relation yet is unconcerned with the forces that determine this relation. 5

**wave length** The wave length for a periodic wave is the distance between two identical features on the wave, such as the peak-to-peak distance. The wavelength equals to  $2\pi/|\mathbf{k}|$ , where  $\mathbf{k}$  is the wavevector. 14

**wave mechanics** Wave mechanics in this book refers to the study of the kinematics and dynamics of linear waves. 2

**wave packet** The modulation of waves serves to organize waves into  $x$ -space regions known as wave trains. A wave packet is a train of free waves that are organized into a localized (and moving) region of  $x$ -space. Localization is enabled by a modulation function and the plane wave modes that are organized within the packet are known as carrier waves. 17

**wave trains** The modulation of waves serves to organize waves into  $x$ -space regions known as wave trains. 17

**wavemaker** A wavemaker is the generator of wave packets or wave trains. 18

**wavenumber** The wavenumber is the magnitude of the wavevector,  $|\mathbf{k}|$ . The wavenumber measures  $2\pi$  times the number of waves per unit length, so that the wavenumber measures the spatial angular frequency of a wave. 8, 12, 15

**wavenumber vector** The wavenumber vector,  $\mathbf{k}$ , is another name for the wavevector. 12

**wavevector** The wavevector,  $\mathbf{k}$ , determines the direction of the wave phase propagation as well as the wavelength,  $2\pi/|\mathbf{k}|$ . The wavevector is related, through the dispersion relation, to the wave angular frequency. 5, 12



---

## Appendix B

### LIST OF ACRONYMS

**AI** artificial intelligence xvii

**GFD** geophysical fluid dynamics ix

**GFM** geophysical fluid mechanics vii

**GVC** generalized vertical coordinate xx



## Appendix C

### LIST OF SYMBOLS

Many symbols encountered in this book are defined local to their usage and are not used far outside of that location. Many other symbols appear in a variety of places and are included in the tables given below. Additionally, we generally aim to respect the following conventions.

- Many symbols are adorned with extra labels. One usage exposes tensor indices, with tensor indices written using the slanted math font, such as  $F^i$  for the component  $i$  of the vector  $\mathbf{F}$ . Another usage expresses part of the name for the symbol, with the label written with the upright sans serif. Examples include the “ $b$ ” in  $\eta_b$  for the position of the bottom solid boundary of a fluid domain, and the “ $h$ ” in  $\nabla_h$  for the horizontal gradient operator.
- We strive for unique symbols to represent distinct mathematical and/or physical objects. Yet that goal must confront the multitude of mathematical expressions appearing in this book. We have chosen, on rare occasions, to allow some symbols to carry multiple meanings. In such cases we emphasize the particular meaning of the symbol to help avoid confusion with its alternative meaning.

#### NON-DIMENSIONAL NUMBERS

SYMBOL	NAME	MEANING
Bu	Burger	$Bu = (\text{deformation radius}/\text{horizontal length scale of flow})^2 = (L_d/L)^2$
Db	Deborah	$Db = \text{relaxation time}/\text{observation time}$
Ek	Ekman	$Ek = \text{vertical frictional acceleration}/\text{planetary Coriolis acceleration}$
Fr	Froude	$Fr = \text{fluid particle speed}/\text{fluid wave speed} = U/c$
Ge	Geostrophic	$Ge = \text{horizontal accelerations from Coriolis}/\text{pressure acceleration} = f U L \rho_a/p$
Kn	Knudsen	$Kn = \text{molecular mean free path}/\text{macroscopic length scale}$
Ma	Mach	$Ma = \text{fluid particle speed}/\text{sound wave speed} = U/c_s$
Pr	Prandtl	$Pr = \text{viscosity}/\text{diffusivity} = \mu/\kappa$
Pe	Peclet	$Pe = \text{advective transport}/\text{diffusive transport} = U L/\kappa$
Re	Reynolds	$Re = \text{inertial acceleration}/\text{frictional acceleration} = U L/\nu$
Ri	Richardson	$Ri = \text{squared buoyancy frequency}/\text{squared vertical shear}$
Ro	Rossby	$Ro = \text{horizontal inertial acceleration}/\text{planetary Coriolis acceleration} = U/(f L)$

SYMBOL	MEANING
$\mathcal{A}$	wave action
$A^L(\mathbf{a}, T)$	Lagrangian representation of a fluid property as a function of material coordinates and time
$\mathbf{a}$	coordinate position for a fluid particle using arbitrary material/Lagrangian coordinates
$\mathbf{A}, \mathbf{A}$	second order skew symmetric tensor with elements satisfying $A^{mn} = -A^{nm}$
$A^\vee$	Avogadro's number: $A^\vee = 6.0222 \times 10^{23}$ mole $^{-1}$
$\mathbf{B}$	baroclinicity vector: $\mathbf{B} = \nabla\rho \times (-\rho^{-1} \nabla p) = (\nabla\rho \times \nabla p)/\rho^2$
$\mathcal{B}$	base (or reference) manifold for describing the space of continuum matter
$b$	Archimidean buoyancy with $b > 0$ for relatively light fluid: $b = -g(\rho - \rho_0)/\rho$
$C$	tracer concentration = mass of tracer per mass of fluid = tracer mass fraction
$C_d$	dimensionless bottom drag coefficient: $C_d > 0$
$\mathcal{C}$	circulation of velocity around the boundary of a surface $\mathcal{C} \equiv \oint_{\partial\mathcal{S}} \mathbf{v} \cdot d\mathbf{r}$
$c_{\text{grav}}$	shallow water gravity wave speed: $c_{\text{grav}} = \sqrt{g H}$
$\mathbf{c}_g$	wave group velocity, given by wavevector gradient of dispersion relation: $\mathbf{c}_g = \nabla_k \varpi(\mathbf{k})$
$c_p$	wave phase velocity: $\mathbf{c}_p = C_p \hat{\mathbf{k}}$
$C_p$	wave phase speed
$c_s$	sound speed: $c_s^{-2} = [\partial\rho/\partial p]_{\Theta,S}$
$c_p$	heat capacity at constant pressure: $c_p = [\partial\mathcal{H}/\partial T]_{p,C}$
$\mathbf{E}, \mathbf{E}$	second order eddy transport tensor for tracers, and with elements $E^{mn}$
$\mathbb{E}^1, \mathbb{E}^2, \mathbb{E}^3$	one (line), two (plane), and three dimensional Euclidean space
$\mathcal{E}$	total energy per mass of a fluid element = sum of internal plus mechanical energies
$e_a$	basis vectors for a chosen coordinate system, with index $a = 1, 2, 3$ for 3-dimensional space
$e^a$	basis one-forms for a chosen coordinate system, with index $a = 1, 2, 3$ for 3-dimensional space
$f$	Coriolis parameter, also the planetary vorticity: $f = 2\Omega \sin\phi$
$f_\circ$	Coriolis parameter at a particular latitude: $f_\circ = 2\Omega \sin\phi_0$
$\mathbf{F}$	frictional acceleration vector
$F^i{}_I$	deformation matrix, which transforms between $\mathbf{x}$ -space (Eulerian) and $\mathbf{a}$ -space (Lagrangian)
$G$	water mass transformation, with dimensions of mass per time
$G = G^{\text{grav}}$	Newton's gravitational constant: $G = 6.674 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2} = 6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$
$G(\mathbf{x} \mathbf{x}_0)$	Green's function with $\mathbf{x}$ the observation point (or field point) and $\mathbf{x}_0$ the source point
$\tilde{G}(\mathbf{x} \mathbf{x}_0)$	modified Green's function for Laplace's operator with Neumann boundary conditions
$G^\ddagger(\mathbf{x} \mathbf{x}_0)$	adjoint Green's function for non-self adjoint operators such as the diffusion operator
$\mathcal{G}(\mathbf{x} \mathbf{x}_0)$	free space Green's function; i.e., the Green's function without boundaries
$\mathbf{G}$	velocity gradient tensor with elements $G^i{}_j$
$\mathcal{G}$	Gibbs potential per mass of a fluid element
$g_e$	gravitational acceleration from central gravity due to just the mass of the planet
$g$	effective gravitational acceleration from central gravity + planetary centrifugal as evaluated at the Earth's surface: $g \approx 9.8 \text{ m s}^{-2}$

SYMBOL	MEANING
$g'$	reduced gravity defined between two shallow water layers: $g'_{k+1/2} = g(\rho_{k+1} - \rho_k)/\rho_{\text{ref}} \ll g$
$\mathfrak{g}$	metric tensor (symmetric positive definite second order tensor) with components $\mathfrak{g}_{ab}$
$\mathbf{g}$	square root of the metric tensor determinant: $\mathbf{g} = \sqrt{\det(\mathfrak{g}_{mn})}$
$\mathbf{g}^E$	square root of the metric tensor determinant using Eulerian coordinates: $\mathbf{g}^E = \sqrt{\det(\mathfrak{g}(\mathbf{x}))}$
$\mathbf{g}^L$	square root of the metric tensor determinant using Lagrangian coordinates: $\mathbf{g}^L = \sqrt{\det(\mathfrak{g}(\mathbf{a}, T))}$
$h_k$	layer thickness for a shallow water fluid: $h_k = \eta_{k-1/2} - \eta_{k+1/2} = \delta_k \eta_{k-1/2}$
$h$	layer thickness for a continuously stratified fluid: $h = \bar{h} \delta\sigma$
$\mathbf{h}$	specific thickness for a generalized vertical coordinate: $\mathbf{h} = \partial z / \partial \sigma = 1 / (\partial \sigma / \partial z)$
$\mathcal{H}(x)$	Heaviside step function: $\mathcal{H}(x) = 0$ for $x < 0$ whereas $\mathcal{H}(x) = 1$ for $x > 0$
$H$	vertical length scale of the flow under consideration
$H$	sometimes used as depth of the ocean bottom: $z = -H(x, y) = \eta_b(x, y)$
$H$	Hamiltonian energy function
$\mathcal{H}$	Hamiltonian density used in field theory; dimensions energy per volume (when in 3d space)
$\mathcal{H}$	enthalpy per mass of a fluid element
$\mathbf{I}$	unit tensor or Kronecker tensor: $\mathbf{I} = \delta^{ab} \mathbf{e}_a \otimes \mathbf{e}_b = \delta^a{}_b \mathbf{e}_a \otimes \mathbf{e}^b = \delta_a{}^b \mathbf{e}^a \otimes \mathbf{e}_b = \delta_{ab} \mathbf{e}^a \otimes \mathbf{e}^b$
$\mathfrak{J}$	internal energy per mass of a fluid element
$i$	$i = \sqrt{-1}$ used for imaginary numbers
$i, j, k$	tensor indices/labels for Eulerian coordinates
$I, J, K$	tensor indices/labels for Lagrangian coordinates
$\text{Im}[\cdot]$	imaginary part of a complex number; e.g., $\text{Im}[e^{-i\omega t}] = -\sin(\omega t)$
$\mathbf{J}$	tracer flux; for material tracers the dimensions are mass per time per area
$\mathbf{k}$	wavevector (dimensions inverse length) for a wave of wavelength $\Lambda = 2\pi/ \mathbf{k} $
$\hat{\mathbf{k}}$	unit vector in the direction of a wave: $\mathbf{k} = \hat{\mathbf{k}}  \mathbf{k} $ (as distinct from the vertical unit vector, $\hat{\mathbf{z}}$ )
$ \mathbf{k} $	wavenumber: $ \mathbf{k}  = 2\pi/\Lambda$
$K$	kinetic energy for a particle of mass $m$ : $K = m \mathbf{V} \cdot \mathbf{V}/2$
$K$	kinetic energy for a system of $N$ particles, $\sum_{n=1}^N m^n \mathbf{V}^n \cdot \mathbf{V}^n$
$\mathcal{K}$	kinetic energy per mass of a fluid element arising from macroscopic motion: $\mathcal{K} = \mathbf{v} \cdot \mathbf{v}/2$
$\mathcal{K}^{\text{hyd}}$	kinetic energy per mass for an approximate hydrostatic flow: $\mathcal{K}^{\text{hyd}} = \mathbf{u} \cdot \mathbf{u}/2$
$\mathcal{K}^{\text{sw}}$	kinetic energy per horizontal area for a shallow water layer: $\mathcal{K}^{\text{sw}} = \rho h \mathbf{u} \cdot \mathbf{u}/2$
$\mathbf{K}, \mathbf{K}'$	positive and symmetric second order tensor parameterizing diffusive mixing
$k$	integer index to label a layer in a shallow water model with $k = 1, N$ layers ( $k = 1$ is top layer)
$k_B$	Boltzmann constant: $k_B = 1.3806 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1} = R^* / A^*$
$k_R$	Rossby height/depth: $k_R =  \mathbf{k}  N / f_0$ with horizontal wavenumber $ \mathbf{k}  = \sqrt{k_x^2 + k_y^2}$
$L$	Lagrangian used in Lagrangian mechanics: kinetic minus potential energies: $L = K - P$
$L$	length scale for a particular physical feature and commonly used in scale analysis

LATIN SYMBOLS AND THEIR MEANING

SYMBOL	MEANING
$\mathcal{L}$	Lagrangian density used in field theory; dimensions energy per volume (when in 3d space)
$L_d$	deformation radius: (a) shallow water $L_d = \sqrt{g H}/f$ ; (b) continuous internal $L_d = H N/f$
$\mathcal{M}$	mechanical energy per mass of a fluid element arising from macroscopic motion
$\mathcal{M}^{sw}$	mechanical energy per area of a shallow water fluid column: $\mathcal{M}^{sw} = \mathcal{K}^{sw} + \mathcal{P}^{sw}$
$\mathbf{M}$	moment of inertia tensor
$\mathbf{M}$	potential momentum vector: $\mathbf{M} = \mathbf{u} + 2\boldsymbol{\Omega} \times \mathbf{X}$
$M$	Montgomery potential for continuously stratified fluid $M = \varphi - bz$
$M$	mass, as in the mass of a fluid region, $M = \int_{\mathcal{R}} \rho dV$
$M_k^{dyn}$	Montgomery potential for a shallow water layer: $M_k^{dyn} = \sum_{j=0}^{k-1} g_{j+1/2}^r \eta_{j+1/2}$
$M^{air}$	mass per mole of air: $M^{air} = 28.8 \times 10^{-3} \text{ kg mole}^{-1}$
$N$	buoyancy frequency
$\mathcal{O}$	order of magnitude
$P$	potential energy of a physical system, with corresponding force $\mathbf{F} = -\nabla P$
$\mathcal{P}_k^{sw}$	potential energy per horizontal area for a shallow water fluid: $\mathcal{P}_k^{sw} = g \rho_k \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} z dz$
$\mathcal{P}$	phase of a wave
$\mathcal{P}_\sigma$	generalized momentum for discrete particle system: $\mathcal{P}_\sigma = \partial L / \partial \dot{\xi}^\sigma$
$\mathcal{P}$	generalized momentum density for continuous media: $\mathcal{P} = \partial \mathcal{L} / \partial (\partial_t \psi)$
$\mathbf{P}$	linear momentum of a physical system
$p$	pressure at a point in the fluid
$p_a$	pressure applied to the ocean surface from the atmosphere or cryosphere
$p_b$	pressure at the bottom of a fluid column, at the fluid-solid earth interface
$p_{slp}$	sea level pressure with an area average, $\langle p_{slp} \rangle = 101.325 \times 10^3 \text{ N m}^{-2}$
$p_{k-1/2}$	hydrostatic pressure at the layer interface with vertical position $z = \eta_{k-1/2}$
$p_k^{dyn}$	dynamic pressure in a shallow water layer: $p_k^{dyn} = \rho_{ref} \sum_{j=0}^{k-1} g_{j+1/2}^r \eta_{j+1/2}$
$P_k$	pressure integrated over a shallow water layer: $P_k \equiv \int_{\eta_{k+1/2}}^{\eta_{k-1/2}} p_k(z) dz = h_k (g \rho_k h_k / 2 + p_{k-1/2})$
$Q$	potential vorticity for continuously stratified (Ertel PV) or shallow water (Rossby PV)
$q$	quasi-geostrophic potential vorticity either for a continuous fluid or shallow water fluid
$Q_m$	mass flux (mass per horizontal area per time) across ocean surface: $Q_m > 0 \text{ enters ocean}$
$\mathcal{Q}_m$	mass flux (mass per surface area per time) across ocean surface: $\mathcal{Q}_m dS = Q_m dA$
$Q_c$	turbulent tracer flux (tracer per horiz area per time) across ocean surface: $Q_c > 0 \text{ enters ocean}$
$\mathcal{Q}_c$	turbulent tracer flux (tracer per surface area per time) across ocean surface: $QCcal dS = Q_c dA$
$r$	radial distance of a point relative to an origin
$\mathbf{R}$	rotation tensor: $2 R^m{}_n = \partial_n v^m - \partial^m v_n = -2 R_n{}^m$
$\mathbb{R}^1$	real number line
$\mathbb{R}^2$	two-dimensional space of real numbers
$\mathbb{R}^3$	three-dimensional space of real numbers
$R$	radius of a sphere

LATIN SYMBOLS AND THEIR MEANING

SYMBOL	MEANING
$R_e$	radius of sphere whose volume approximates that of the earth: $R_e = 6.371 \times 10^6 \text{ m}$
$R^g$	universal gas constant: $R^g = 8.314 \text{ J mole}^{-1} \text{ K}^{-1} = 8.314 \text{ kg m}^2 \text{ s}^{-2} \text{ mole}^{-1} \text{ K}^{-1}$
$R^{\text{air}}$	specific gas constant for air: $R^{\text{air}} = R^g/M^{\text{air}} = 2.938 \times 10^2 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}$
$\mathcal{R}$	arbitrary region or manifold
$\mathcal{R}^a{}_b$	orthogonal rotation matrix
$\text{Re}[\cdot]$	real part of a complex number; e.g., $\text{Re}[e^{-i\omega t}] = \cos(\omega t)$
$\mathcal{S}$	spatial manifold
$\mathcal{S}$	entropy per mass of a fluid element
$\mathcal{S} = \mathcal{S}^{\text{action}}$	action: time integral of the Lagrangian: $\mathcal{S} = \int_{t_A}^{t_B} L \, dt$
$\mathbf{S}$	strain rate tensor: $2\mathbf{S} = \nabla \mathbf{v} + (\nabla \mathbf{v})^T$
$\mathbf{S}^{\text{dev}}$	deviatoric strain rate tensor: $\mathbf{S}^{\text{dev}} = \mathbf{S} - S^q_q/3$
$S$	salt concentration = mass of salt in a fluid element per mass of seawater
$S$	Absolute Salinity, generically referred to as salinity: $S = 1000 \mathcal{S}$
$s$	expression for a generic surface: $s = s(x, y, z, t)$ .
$s$	arc-length along a curve $\mathbf{x}(s)$ with infinitesimal increment $ds = \sqrt{d\mathbf{x} \cdot d\mathbf{x}}$
$\hat{s}$	unit tangent to a curve, also written as $\hat{s} = \hat{\mathbf{t}}$ (see below)
$\text{sgn}$	sign function related to Heaviside step function via $\text{sgn}(x) = 2\mathcal{H}(x) - 1$
$T'$	absolute thermodynamic <i>in situ</i> temperature (Kelvin if in a thermodynamic equation)
$T$	time scale for a particular physical process and commonly used in scale analysis
$T$	time (universal Newtonian time) measured in the Lagrangian reference frame
$t$	time (universal Newtonian time) measured in the Eulerian reference frame
$\tau$	general symbol for time as considered in the tensor analysis chapters
$\mathbf{T}$	stress tensor with natural elements $T^m{}_n$
$\mathbb{T}^{\text{kinetic}}$	kinetic stress tensor: $\mathbb{T}^{\text{kinetic}} = -\rho \mathbf{v} \otimes \mathbf{v}$
$\mathbb{T}^{\text{sw kinetic}}$	kinetic stress tensor for shallow water fluid: $\mathbb{T}^{\text{sw kinetic}} = -\rho \mathbf{u} \otimes \mathbf{u}$
$\hat{\mathbf{t}}$	unit tangent to a curve: $\hat{\mathbf{t}} = d\mathbf{x}/ds$ , where $s$ is the arc-length so that $ds = \sqrt{d\mathbf{x} \cdot d\mathbf{x}}$
$\mathbf{u}$	horizontal velocity of a fluid particle, with Cartesian representation: $\mathbf{u} = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v$
$U$	horizontal velocity scale of the flow under consideration
$\mathbf{U}$	depth integrated horizontal velocity: $\mathbf{U} = \int_{\eta_b}^{\eta} \mathbf{u} \, dz$
$V$	volume, as in the volume of a fluid region, $V = \int_{\mathcal{R}} dV$
$\mathbf{v}$	velocity of a fluid particle: $\mathbf{v} = D\mathbf{x}/Dt$ , with Cartesian components $\mathbf{v} = \hat{\mathbf{x}} u + \hat{\mathbf{y}} v + \hat{\mathbf{z}} w$
$\mathbf{v}^*$	eddy-induced velocity
$\mathbf{v}^\dagger$	residual velocity of a fluid particle: $\mathbf{v}^\dagger = \mathbf{v} + \mathbf{v}^*$
$\mathbf{v}^{(b)}$	velocity of a point on a region boundary
$\mathbf{v}^L(\mathbf{a}, T)$	Lagrangian velocity of a fluid particle so that $\mathbf{v}^L(\mathbf{a}, T) = \mathbf{v}[\mathbf{x} = \boldsymbol{\varphi}(\mathbf{a}, T), t = T]$
$\mathbf{v}_I$	velocity of a fluid particle measured in the inertial/absolute reference frame: $\mathbf{v}_I = \mathbf{v} + \boldsymbol{\Omega} \times \mathbf{x}$
$W$	vertical velocity scale of the flow under consideration
$\mathcal{W}$	on-shell action
$w$	vertical component to the velocity: $w = Dz/Dt$
$w^{\text{dia}}$	dia-surface flux = volume per surface area per time crossing a $\sigma$ -surface: $w^{\text{dia}} = (1/ \nabla \sigma ) \dot{\sigma}$
$w^{(\dot{\sigma})}$	dia-surface velocity = volume per <i>horizontal</i> area per time crossing $\sigma$ -surface: $w^{(\dot{\sigma})} = \dot{\sigma} \partial z/\partial \sigma$

---

## LATIN SYMBOLS AND THEIR MEANING

SYMBOL	MEANING
$(x, y, z)$	triplet of Cartesian coordinates
$\mathbf{x}$	spatial position as a line segment with an arrow pointing from an origin to the position of a particle
$x$	spatial position represented by either general coordinates or Cartesian coordinates
$\hat{\mathbf{x}}$	initial position for a fluid particle using arbitrary coordinates
$(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$	triplet of Cartesian unit vectors oriented in a righthand sense
$\mathbf{X}(t)$	position for a point particle defining a trajectory through space-time
$\mathbf{X}(\mathbf{a}, T)$	position of a material fluid particle expressed using material coordinates
$z_\sigma$	specific thickness for a generalized vertical coordinate: $z_\sigma = \partial z / \partial \sigma = h$

## GREEK SYMBOLS AND THEIR MEANING

SYMBOL	MEANING
$\alpha$	thermal expansion: $\alpha = -\rho^{-1} \partial \rho / \partial \theta$ or $\alpha = -\rho^{-1} \partial \rho / \partial \Theta$ or $\alpha = -\rho^{-1} \partial \rho / \partial T$
$\alpha_T$	thermal expansion in terms of <i>in situ</i> temp: $\alpha = -\rho^{-1} \partial \rho / \partial T$
$\alpha^{(\Theta)}$	thermal expansion in terms of Conservative Temperature: $\alpha^{(\Theta)} = -\rho^{-1} \partial \rho / \partial \Theta$
$\alpha_{\text{aspect}}$	aspect ratio; ratio of vertical to horizontal scales of the flow: $\alpha_{\text{aspect}} = H/L$
$\beta, \beta^{(S)}$	haline (saline) contraction coefficient: $\beta = \beta^{(S)} = \rho^{-1} \partial \rho / \partial S$
$\beta$	meridional derivative of planetary vorticity: $\beta = \partial_y f$
$\hat{\gamma}$	dianeutral unit direction perpendicular to the neutral tangent plane
$\delta_{ab}$	Kronecker delta, which is the metric for Euclidean space with Cartesian coordinates
$\delta_b^a$	components to the Kronecker tensor in arbitrary coordinates
$\epsilon$	kinetic energy dissipation from viscosity (energy per time per mass)
$\epsilon_{ab}$	components to the permutation symbol in two space dimensions
$\epsilon_{abc}$	components to the permutation symbol in three space dimensions
$\varepsilon_{abc}$	components to the Levi-Civita symbol in three space dimensions: $\varepsilon_{abc} = \sqrt{\det(g_{ab})} \epsilon_{abc}$
$\zeta$	vertical component to the relative vorticity; e.g., $\zeta = \partial_x v - \partial_y u$
$\zeta_a$	vertical component to the absolute vorticity; e.g., $\zeta_a = f + \zeta$
$\eta$	vertical position of the free upper surface of a fluid domain: $z = \eta(x, y, t)$
$\eta$	vertical position of a generalized vertical coordinate surface: $z = \eta(x, y, \sigma, t)$ , with $\sigma$ the generalized vertical coordinate
$\eta_{k-1/2}$	vertical position of the top interface of the $k$ shallow water layer
$\eta_{k+1/2}$	vertical position of the lower interface of the $k$ shallow water layer
$\eta_b = -H$	vertical position of static solid-earth boundary: $z = \eta_b(x, y) = -H(x, y)$
$\theta$	potential temperature
$\Theta$	Conservative Temperature
$\kappa$	molecular kinematic diffusivity
$\kappa_T$	molecular diffusivity for <i>in situ</i> temperature in water: $\kappa_T = 1.4 \times 10^{-7} \text{ m}^2 \text{ s}^{-1}$
$\kappa_S$	molecular diffusivity for salt in water: $\kappa_S = 1.5 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$
$\kappa_{\text{eddy}}$	kinematic eddy diffusivity: $\kappa_{\text{eddy}} \gg \kappa$
$\Lambda$	wavelength of a wave: $\Lambda = 2\pi/ \mathbf{k} $ , where $\mathbf{k}$ is the wavevector and $ \mathbf{k} $ the wavenumber.
$\lambda$	reduced wavelength of a wave: $\lambda = \Lambda/(2\pi) = 1/ \mathbf{k} $ .
$\lambda$	longitude on the sphere: $0 \leq \lambda \leq 2\pi$
$\mu_n$	chemical potential for constituent $n$ within a fluid (energy per mass)
$\tilde{\mu}_n$	chemical potential for constituent $n$ within a fluid (energy per mole number)
$\mu$	relative chemical potential for a binary fluid
$\mu$	chemical potential for seawater: $\mu = \mu_{\text{salt}} - \mu_{\text{water}}$
$\mu_{\text{vsc}}$	dynamic viscosity = $\rho \nu$
$\nu_s$	specific volume: $\nu_s = \rho^{-1}$
$\nu$	molecular kinematic viscosity
$\nu_{\text{air}}$	molecular kinematic viscosity of air: $\nu_{\text{air}} \approx 1.3 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$
$\nu_{\text{water}}$	molecular kinematic viscosity of fresh water: $\nu_{\text{water}} \approx 10^{-6} \text{ m}^2 \text{ s}^{-1}$
$\nu_{\text{eddy}}$	eddy viscosity: $\nu_{\text{eddy}} \gg \nu$
$\xi^a$	a'th component to a generalized coordinate
$\Pi$	Exner function
$\Pi$	Boussinesq dynamic enthalpy

## GREEK SYMBOLS AND THEIR MEANING

SYMBOL	MEANING
$\rho$	Eulerian <i>in situ</i> density (mass per volume) of a fluid element: $\rho = \rho(\mathbf{x}, t)$
$\rho^L$	mass density following a fluid particle trajectory (Lagrangian mass density): $\rho^L = \rho^L(\mathbf{a}, T)$
$\dot{\rho}^L$	initial mass density in Lagrangian space-time: $\dot{\rho}^L = \rho^L(\mathbf{a}, T = t_0)$
$\rho_0$	constant reference density used for the Boussinesq ocean
$\rho_{\text{ref}}$	constant reference density used for the shallow water fluid
$\varrho$	potential density referenced to a specified pressure
$\sigma$	generalized vertical coordinate, $\sigma = \sigma(x, y, z, t)$
$\tau$	stress vector such as from winds or bottom stresses acting on the ocean
$\mathbb{T}$	frictional stress tensor
$\varphi$	pressure divided by the Boussinesq reference density: $\varphi = p/\rho_0$
$\varphi$	sometimes used as the variable for parameterizing a curve
$\phi$	latitude on the sphere: $-\pi/2 \leq \phi \leq \phi/2$
$\Phi_e$	gravitational potential from a spherical and homogeneous earth
$\Phi$	geopotential from central gravity plus planetary centrifugal; also, potential energy per mass
$\Phi$	inverse flow map that generates an inverse mapping of the fluid continuum: $\mathbf{a} = \Phi(\mathbf{x}, t)$
$\boldsymbol{\varphi}$	motion field that maps the fluid continuum as time evolves: $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{a}, T)$
$\psi$	streamfunction for two-dimensional non-divergent flow: $\mathbf{u} = \hat{\mathbf{z}} \times \nabla \psi$
$\Psi$	vector streamfunction for three-dimensional non-divergent flow: $\mathbf{v} = \nabla \times \Psi$
$\omega$	relative vorticity: $\boldsymbol{\omega} = \nabla \times \mathbf{v}$
$\omega$	angular frequency for a wave so that the wave period is $2\pi/ \omega $
$\varpi$	dispersion relation for linear waves, relating angular frequency to the wavevector: $\omega = \varpi(\mathbf{k})$
$\Omega$	angular velocity for a rotating reference frame
$\Omega$	earth's angular velocity oriented through the north pole: $ \Omega  = 7.2921 \times 10^{-5} \text{ s}^{-1}$

## MATHEMATICAL OPERATIONS AND RELATIONS

SYMBOL	MEANING
$[\equiv]$	"has dimensions" for use in referring to the physical dimensions
$\times$	vector cross product
$\nabla$	covariant derivative operator, which acts on a $(p, q)$ tensor to produce a $(p, q + 1)$ tensor.
$\nabla$	gradient operator
$\nabla_h$	horizontal gradient operator on constant $z$ surface: $\nabla_h = \hat{\mathbf{x}} (\partial/\partial x)_z + \hat{\mathbf{y}} (\partial/\partial y)_z = \hat{\mathbf{x}} \partial_x + \hat{\mathbf{y}} \partial_y$
$\nabla \cdot$	divergence operator that acts on a vector to produce a scalar
$\nabla \times$	curl operator
$\nabla_\sigma$	horizontal gradient on constant $\sigma$ -surface: $\nabla_\sigma = \hat{\mathbf{x}} (\partial/\partial x)_\sigma + \hat{\mathbf{y}} (\partial/\partial y)_\sigma$
$\partial/\partial\sigma$	vertical partial derivative with general vertical coordinate: $\partial_\sigma = \partial/\partial\sigma = \partial/\partial\sigma = (\partial z/\partial\sigma) \partial/\partial z$
$\partial/\partial t$	Eulerian time derivative acting at a fixed spatial position, $\mathbf{x}$ , also written as $\partial_t$
$[\partial/\partial t]_\sigma$	time derivative computed on constant $\sigma$ -surface
$D/Dt$	material, Lagrangian, or substantial time derivative following a fluid particle
$D_r/Dt$	time derivative following a ray (integral lines of the group velocity): $D_r/Dt = \partial/\partial t + \mathbf{c}_g \cdot \nabla$
$D_g/Dt$	time derivative following the horizontal geostrophic flow $D_g/Dt = \partial/\partial t + \mathbf{u}_g \cdot \nabla$
$\ddot{\phantom{x}}$	inexact differential operator commonly found in thermodynamics
$\delta$	virtual displacement (also the variation) for Lagrangian mechanics and Hamilton's principle
$\delta$	differential increment that signals an object following the fluid flow
$\delta(x)$	one-dimensional Dirac delta with dimensions of inverse length
$\delta^{(2)}(\mathbf{x})$	two-dimensional Dirac delta with dimensions of inverse area
$\delta(\mathbf{x})$	three-dimensional Dirac delta with dimensions of inverse volume
$\delta(t)$	temporal Dirac delta with dimensions of inverse time
$\Delta$	finite difference increment in space: $\Delta_x, \Delta_y, \Delta_z, \Delta_\sigma$
$dA$	infinitesimal horizontal area element: $dA = dx dy$
$d^3a$	infinitesimal region of material space: $d^3a = da db dc$
$d\mathcal{S}$	infinitesimal area element on a surface
$dV$	infinitesimal volume element, sometimes written $dV = d\mathbf{x}$
$d\mathbf{x}$	infinitesimal volume element, with Cartesian expression $d\mathbf{x} = dV = dx dy dz$
$\delta V$	infinitesimal volume for a region moving with the fluid (Lagrangian region)
$\int_{\mathcal{R}} dV$	volume integral over an arbitrary region, $\mathcal{R}$
$\int_{\mathcal{R}(\mathbf{v})} dV$	volume integral over a region following the fluid flow (Lagrangian integral)
$\int_{\mathcal{S}} d\mathcal{S}$	surface integral over an arbitrary surface $\mathcal{S}$
$\oint_{\partial\mathcal{R}} d\mathcal{S}$	surface integral over a closed surface $\partial\mathcal{R}$ that bounds the volume $\mathcal{R}$
$\oint d\ell$	closed line integral over a periodic domain
$\oint_{\partial\mathcal{S}} d\ell$	counter-clockwise closed line integral over the boundary of a surface, $\partial\mathcal{S}$
$\sim$	"similar to" or "scales as"
$\approx$	approximately equal to
$\dot{\Psi}$	time derivative following a trajectory; for fluid particle trajectories then, $\dot{\Psi} = D\Psi/Dt$



## BIBLIOGRAPHY

Acheson, D., *Elementary Fluid Dynamics*, Oxford Applied Mathematics and Computing Science Series, Oxford, Oxford, 1990. 5

Anderson, P. W., More is different, *Science*, 177, 393–396, doi:10.1126/science.177.4047.39, 1972. ix

Andrews, D. G., and M. E. McIntyre, On wave action and its relatives, *Journal of Fluid Mechanics*, 89, 647–664, doi:10.1017/S0022112078002785, 1978. 43

Badin, G., and F. Crisciani, *Variational formulation of fluid and geophysical fluid dynamics*, Springer International Publishing, Switzerland, 218 pages, 2018. xii

Batchelor, G. K., *An Introduction to Fluid Dynamics*, Cambridge University Press, Cambridge, England, 1967. 384

Bender, C., and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill Book Company, New York, 593 pp, 1978. 85

Beyer, W., *Standard Mathematical Tables*, 24th Edition, CRC Press, 1973. 30

Bohm, D., *Quantum Theory*, Dover Publications, Inc., New York, 646 pp, 1951. 31, 37

Boyce, W. E., and R. C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, 9th ed., Wiley Publications, 2009. 318, 440

Bretherton, F. P., The general linearized theory of wave propagation, in *Mathematical Problems in the Geophysical Sciences*, edited by W. H. Reid, pp. 61–102, American Mathematical Society, 1971. 5, 8, 43, 54, 56

Bretherton, F. P., and C. J. R. Garrett, Wavetrains in inhomogeneous moving media, *Proceedings of the Royal Society A*, 302, 529–554, doi:10.1098/rspa.1968.0034, 1969. 43, 54

Brown, E., *Waves, tides, and shallow-water processes*, 227 pp., The Open University, Milton Keys, UK, 1999. 130

Buckingham, C. E., J. Gula, and X. Carton, The role of curvature in modifying frontal instabilities. Part I: Review of theory and presentation of a nondimensional instability criterion, *Journal of Physical Oceanography*, 51, 299–315, doi:10.1175/JPO-D-19-0265.1, 2021. 331, 361, 362

- Böhler, O., *Waves and mean flows*, 2nd ed., Cambridge University Press, Cambridge, UK, doi:10.1017/CBO9781107478701, 2014. 52, 64, 136, 143
- Buijsman, M. C., B. K. A. an Samuel M. Kelly, and A. F. Waterhouse, Internal gravity waves, in *Encyclopedia of Ocean Sciences, Reference Module in Earth Systems and Environmental Sciences*, vol. 3, pp. 622–632, Elsevier, doi:10.1016/B978-0-12-409548-9.04160-9, 2019. 276, 291
- Butt, T., P. Russell, and R. Grigg, *Surf Science - An Introduction To Waves For Surfing*, University of Hawaii Press, 2004. 10
- Chandrasekhar, S., *Hydrodynamic and Hydromagnetic Stability*, Dover Publications, New York, 654 pp, 1961. 372, 380, 387, 416
- Charney, J. G., and M. E. Stern, On the stability of internal baroclinic jets in a rotating atmosphere, *Journal of the Atmospheric Sciences*, 19, 159–172, doi:10.1175/1520-0469(1962)019<0159:OTSOIB>2.0.CO;2, 1962. 457
- Cohen-Tannoudji, C., B. Diu, and F. Laloë, *Quantum Mechanics: Volumes One and Two*, John Wiley and Sons, New York, 1524 pp, 1977. 34, 35, 37, 41, 166
- Cushman-Roisin, B., and J.-M. Beckers, *Introduction to Geophysical Fluid Dynamics*, Academic Press, Amsterdam, 828, 2011. 130, 245, 249, 263, 291, 331, 391, 429, 433, 466
- Davidson, P., *Turbulence: an introduction for scientists and engineers*, Oxford University Press, Oxford, UK, 2015. 147
- Drazin, P., and R. Johnson, *Solitons: an introduction*, Cambridge University Press, Cambridge, UK, 226 pp, 1989. 125
- Drazin, P., and W. Reid, *Hydrodynamic stability*, Cambridge University Press, Cambridge, UK, 527 pp, 1981. 331, 339, 384
- Drazin, P., and W. Reid, *Hydrodynamic stability*, 2nd ed., Cambridge University Press, Cambridge, UK, 605 pp, 2004. 324, 391, 416, 417
- Eady, E. T., Long waves and cyclone waves, *Tellus*, 1, 33–52, doi:10.1111/j.2153-3490.1949.tb01265.x, 1949. 431, 465
- Elmore, W. C., and M. A. Heald, *Physics of Waves*, Dover Publications, New York, 1969. 11
- Fetter, A. L., and J. D. Walecka, *Theoretical Mechanics of Particles and Continua*, Dover Publications, Mineola, New York, 570 pp, 2003. 34, 35, 37, 41, 63, 81, 85, 93, 96, 109, 114, 125, 132, 135, 380, 387
- Feynman, R. P., R. B. Leighton, and M. L. Sands, *The Feynman lectures on physics*, Addison-Wesley Publishing Company, Reading, Mass, 1963. 68
- Fjørtoft, R., An application of integral theorems in deriving criteria for laminar flows and for the baroclinic circular vortex, *Geophysics Publications*, 17, 1–52, 1950. 400
- Gasiorowicz, S., *Quantum Physics*, Wiley International, Hoboken, NJ, USA, 1974. 31, 37

- Gill, A., *Atmosphere-Ocean Dynamics*, International Geophysics Series, vol. 30, Academic Press, London, 662 + xv pp, 1982. 249, 292, 314, 320, 433, 451
- Goldstein, H., *Classical Mechanics*, Addison-Wesley Publishing, Reading, MA, 1980. 49
- Gregory, J., S. M. Griffies, C. Hughes, J. Lowe, J. Church, I. Fukimori, N. Gomez, R. Kopp, F. Landerer, R. Ponte, D. Stammer, M. Tamisiea, and R. van den Wal, Concepts and terminology for sea level—mean, variability and change, both local and global, *Surveys in Geophysics*, 40, 1251–1289, doi:10.1007/s10712-019-09555-7, 2019. 469
- Haidvogel, D. B., and P. B. Rhines, Waves and circulation driven by oscillatory winds in an idealized ocean basin, *Geophysical and Astrophysical Fluid Dynamics*, 25, 1–63, doi:10.1080/03091928308221747, 1983. 166
- Heifetz, E., P. Alpert, and A. da Silva, On the parcel method and the baroclinic wedge of instability, *Journal of Atmospheric Sciences*, 55, 788–795, doi:10.1175/1520-0469(1998)055<0788:OTPMAT>2.0.CO;2, 1998. 466
- Held, I., R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson, Surface quasi-geostrophic dynamics, *Journal of Fluid Mechanics*, 282, 1–20, doi:10.1017/S0022112095000012, 1995. 114, 435
- Held, I. M., 100 years of progress in understanding the general circulation of the atmosphere, *Meteorological Monographs*, 59(6), 6.1–6.23, doi:10.1175/amsmonographs-d-18-0017.1, 2019. 466
- Holton, J. R., and G. J. Hakim, *An Introduction to Dynamic Meteorology*, 5th ed., Academic Press, Waltham, Mass, USA, 532 pp, 2013. 331, 362, 369
- Hoskins, B. J., The role of potential vorticity in symmetric stability and instability, *Quarterly Journal of the Royal Meteorological Society*, 100, 480–482, 1974. 356, 362
- Hoskins, B. J., The geostrophic momentum approximation and the semi-geostrophic equations, *Journal of Atmospheric Sciences*, 32, 233–242, 1975. 329, 362, 364, 365
- Howard, L. N., Note on a paper of John W. Miles, *Journal of Fluid Mechanics*, 10, 509–512, doi:10.1017/S0022112061000317, 1961. 422, 425
- Johnson, R. S., *A Modern Introduction to the Mathematical Theory of Water Waves*, Cambridge Texts in Applied Mathematics, 1997. 128
- José, J. V., and E. J. Saletan, *Classical Dynamics: A Contemporary Approach*, Cambridge Press, 670 pp., 1998. 44
- Kundu, P., I. Cohen, and D. Dowling, *Fluid Mechanics*, Academic Press, 921 + xxiv pp, 2016. 63, 76, 249, 291, 339, 391, 425, 465
- Kuo, H. L., Dynamic stability of two-dimensional nondivergent flow in a barotropic atmosphere, *Journal of the Meteorology*, 6, 105–122, doi:10.1175/1520-0469(1949)006<0105:DIOTDN>2.0. CO;2, 1949. 400
- Lamb, H., *Hydrodynamics*, 6th ed., Cambridge University Press, 1993. 107
- Landau, L. D., and E. M. Lifshitz, *Mechanics*, Pergamon Press, Oxford, UK, 170 pp, 1976. 44

- Landau, L. D., and E. M. Lifshitz, *Fluid Mechanics*, Pergamon Press, Oxford, UK, 539 pp, 1987. [63](#)
- Lighthill, J., *Waves in Fluids*, Cambridge University Press, Cambridge, UK, 504 pp, 1978. [41](#), [43](#), [68](#), [85](#), [123](#), [147](#), [249](#), [291](#)
- Longuet-Higgins, H. C., Planetary waves on a rotating sphere, *Proceedings of the Royal Society of London A*, **279**, 446–473, doi:10.1098/rspa.1964.0116, 1964. [178](#), [180](#), [184](#)
- Luke, J. C., A variational principle for a fluid with a free surface, *Journal of Fluid Mechanics*, **27**, 395–397, doi:10.1017/S0022112067000412, 1967. [109](#), [111](#), [112](#)
- MacKinnon, J., Louis St. Laurent, and A. C. Naveira Garabato, Diapycnal mixing processes in the ocean interior, in *Ocean Circulation and Climate, 2nd Edition: A 21st century perspective, International Geophysics Series*, vol. 103, edited by G. Siedler, S. M. Griffies, J. Gould, and J. Church, pp. 159–183, Academic Press, doi:10.1016/B978-0-12-391851-2.00007-6, 2013. [276](#)
- MacKinnon, J., Z. Zhao, C. Whalen, A. Waterhouse, D. Trossman, O. Sun, L. S. Laurent, H. Simmons, K. Polzin, R. Pinkel, A. Pickering, N. Norton, J. Nash, R. Musgrave, L. Merchant, A. Melet, B. Mater, S. Legg, W. Large, E. Kunze, J. Klymak, M. Jochum, S. Jayne, R. Hallberg, S. M. Griffies, P. Gent, S. Diggs, G. Danabasoglu, E. Chassignet, M. Buijsman, F. Bryan, B. Briegleb, A. Barna, B. Arbic, J. Ansong, and M. Alford, Climate process team on internal-wave driven ocean mixing, *Bulletin of the American Meteorological Society*, pp. 2429–2454, doi:10.1175/BAMS-D-16-0030.1, 2017. [276](#)
- Marion, J. B., and S. T. Thornton, *Classical Dynamics of Particles and Systems*, Harcourt Brace Jovanovich, San Diego, USA, 602 pp, 1988. [x](#), [49](#)
- Markowski, P., and Y. Richardson, *Mesoscale Meteorology in Midlatitudes*, Wiley-Blackwell Publishers, Oxford, UK, 2010. [331](#), [339](#), [342](#)
- McWilliams, J., *Fundamentals of Geophysical Fluid Dynamics*, Cambridge University Press, Cambridge, Cambridge, UK, 2006. [166](#), [384](#)
- McWilliams, J. C., Submesoscale currents in the ocean, *Proceedings of the Royal Society, A* **472**, doi:10.1098/rspa.2016.0117, 2016. [331](#)
- Mermin, N. D., What's wrong with these equations?, *Physics Today*, **42**, 9–11, doi:10.1063/1.2811173, 1989. [viii](#)
- Milder, D. M., A note regarding ‘on hamilton’s principle for surface waves’, *Journal of Fluid Mechanics*, **83**, 159–161, doi:10.1017/S0022112077001116, 1977. [97](#), [109](#), [110](#)
- Miles, J. W., On the stability of heterogeneous shear flows, *Journal of Fluid Mechanics*, **10**, 496–508, doi:10.1017/S0022112061000305, 1961. [389](#), [425](#)
- Miles, J. W., On hamilton’s principle for surface waves, *Journal of Fluid Mechanics*, **83**, 153–158, doi:10.1017/S0022112077001104, 1977. [97](#), [109](#)
- Mory, M., Inertial oscillations, in *Rotating Fluids in Geophysical and Industrial Applications*, International Centre for Mechanical Sciences (Courses and Lectures), Springer, Vienna, doi:10.1007/978-3-7091-2602-8\_8, 1992. [147](#)

- Neumann, G., and W. Pierson, *Principles of Physical Oceanography*, Prentice-Hall, Englewood Cliffs, USA, 545 pp, 1966. 130
- Olbers, D. J., J. Willebrand, and C. Eden, *Ocean Dynamics*, 1st ed., Springer, Berlin, Germany, 704 pages, 2012. xii, 5, 44, 49, 62
- Pathria, R. K., and P. D. Beale, *Statistical Mechanics*, 4 ed., Academic Press (Elsevier), Amsterdam, 2021. 452
- Pedlosky, J., The stability of currents in the atmosphere and ocean, Part I, *Journal of the Atmospheric Sciences*, 21, 201–219, doi:10.1175/1520-0469(1964)021<0201:TSOCIT>2.0.CO;2, 1964. 457
- Pedlosky, J., *Geophysical Fluid Dynamics*, 2nd ed., Springer-Verlag, Berlin Heidelberg New York, 710 + xv pp, 1987. 465
- Pedlosky, J., *Waves in the Ocean and Atmosphere: Introduction to Wave Dynamics*, Springer-Verlag, Berlin Heidelberg New York, 206 + viii pp, 2003. 5, 125, 177, 197, 249, 279, 291, 292, 314, 320, 451, 454, 465
- Pierce, A. D., Wave equation for sound in fluids with unsteady inhomogeneous flow, *Journal of the Acoustical Society of America*, 87, 2292–2299, doi:10.1121/1.399073, 1990. 70, 85, 93
- Pratt, L. J., and J. A. Whitehead, *Rotating Hydraulics: Nonlinear Topographic Effects in the Ocean and Atmosphere*, 589 pp., Springer Science + Business Media, New York, 2008. 209, 246
- Rayleigh, L., On the stability, or instability, of certain fluid motions, *Proceedings of the London Mathematical Society*, 9, 57–70, doi:10.1112/plms/s1-11.1.57, 1880. 400
- Rayleigh, L., *The Theory of Sound*, Macmillan, London, UK, 1894. 415
- Salmon, R., *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, Oxford, England, 378 + xiii pp., 1998. xii
- Seliger, R. L., and G. B. Whitham, Variational principles in continuum mechanics, *Proceedings of the Royal Society A*, 305, 1–25, doi:10.1098/rspa.1968.0103, 1968. 111
- Smyth, W. D., and J. R. Carpenter, *Instability in Geophysical Flows*, Cambridge University Press, Cambridge, UK, 327 pp, 2019. 189, 329, 391, 416, 451
- Squire, H. B., On the stability for three-dimensional disturbances of viscous fluid flow between parallel walls, *Proceedings of the Royal Society of London. Series A*, 142, 621–628, doi:10.1098/rspa.1933.0193, 1933. 391
- Stakgold, I., *Boundary value problems of mathematical physics, volume I*, SIAM, Philadelphia, 340 pp, 2000a. 473
- Stakgold, I., *Boundary value problems of mathematical physics, volume II*, SIAM, Philadelphia, 408 pp, 2000b. 22, 28, 37
- Stern, M., *Ocean circulation physics, International Geophysics Series*, vol. 19, Academic Press, New York, New York, 246 pp, 1975. 147, 161

- Sutherland, B. R., *Internal Gravity Waves*, Cambridge University Press, 2010. 5, 176, 184, 240, 249, 269, 291, 292, 314, 320, 380
- Symon, K., *Mechanics*, Addison-Wesley Publishing Co., Reading, MA, USA, 639 pp, 1971. x
- Thomas, L., A. Tandon, and A. Mahadevan, Submesoscale processes and dynamics, in *Eddy resolving ocean models*, edited by M. Hecht and H. Hasumi, Geophysical Monograph 177, pp. 17–38, American Geophysical Union, 2008. 362
- Thomas, L., J. R. Taylor, R. Ferrari, and T. Joyce, Symmetric instability in the Gulf Stream, *Deep Sea Research II*, 91, 96–110, doi:10.1016/j.dsr2.2013.02.025, 2013. 331, 361, 362, 369
- Thorne, K., and R. Blandford, *Modern Classical Physics*, Princeton University Press, Princeton, USA, 1511 + xl pp, 2017. 5, 48, 63, 93
- Towne, D. H., *Wave phenomena*, Dover Publications, New York, 482 pp, 1967. 68
- Tracy, E. R., A. J. Brizard, A. S. Richardson, and A. N. Kaufman, *Ray Tracing and Beyond: Phase Space Methods in Plasma Wave Theory*, Cambridge University Press, 2014. xv, 44, 52, 54
- Treguier, A. M., I. M. Held, and V. D. Larichev, On the parameterization of quasi-geostrophic eddies in primitive equation ocean models, *Journal of Physical Oceanography*, 27, 567–580, doi:10.1175/1520-0485(1997)027<0567:POQEIP>2.0.CO;2, 1997. 454
- Tromp, J., *Theoretical and Computational Seismology*, Princeton University Press, Princeton, USA, 2025. xv
- Tyler, R. H., and R. Käse, A string function for describing the propagation of large-scale energy anomalies in a rotating fluid, *Geophysical and Astrophysical Fluid Dynamics*, 92, 31–64, doi:10.1080/03091920008203710, 2000. 231
- Tyler, R. H., and R. Käse, A string function for describing the propagation of baroclinic anomalies in the ocean, *Journal of Physical Oceanography*, 31, 765–776, doi:10.1175/1520-0485(2001)031<0765:ASFFDT>2.0.CO;2, 2001. 231
- Vallis, G. K., *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-scale Circulation*, 2nd ed., Cambridge University Press, Cambridge, 946 + xxv pp, 2017. 5, 112, 125, 167, 184, 189, 192, 244, 246, 249, 289, 291, 292, 314, 416, 433, 454, 465
- Vallis, G. K., *Essentials of Atmospheric and Oceanic Dynamics*, 1st ed., Cambridge University Press, Cambridge, 2019. 246
- Whitham, G. B., *Linear and nonlinear waves*, John Wiley and Sons, New York, 636 + xvi pp, 1974. 9, 39, 41, 43, 52, 54, 96, 109, 111, 112
- Yassin, H., Normal modes with boundary dynamics in geophysical fluids, *Journal of Mathematical Physics*, 62, doi:10.1063/5.0048273, 2021. 114, 320
- Yassin, H., and S. M. Griffies, On the discrete normal modes of quasigeostrophic theory, *Journal of Physical Oceanography*, 52, 243–259, doi:10.1175/JPO-D-21-0199.1, 2022. 114, 320, 435
- Zeitlin, V., *Geophysical Fluid Dynamics: Understanding (almost) everything with rotating shallow water models*, Oxford University Press, 2018. xix
- Zinnser, W., *Writing to Learn*, Harper Perennial, 1993. xiv

# INDEX

- absolute momentum, 341
- acoustic wave
  - density, 72
  - pressure, 72
  - temperature, 73
  - velocity, 72
- acoustic waves, 63, 71
  - density, 74
  - energetics, 77
  - Eulerian, 69
  - Lagrangian, 65
  - piston wavemaker, 81
  - pressure, 74
  - radiation, 81
  - speed, 68, 71
  - velocity, 74
- action, xi, 52
  - phase averaged, 54
- adiabatic invariant, 54
- adiabatic lapse rate, 73
- advection equation, 9
- ageostrophic overturning, 362
- ageostrophic secondary circulation, 358, 362
- angular momentum, 333
- ansatz, 17
- association versus causation, xi
- available potential energy, 254, 271
- axisymmetric flow, 331
- balances, xi
- baroclinic instability, 431, 451
  - heat transport, 461
  - necessary condition, 456
- baroclinic mode, 211, 249, 439, 440
- baroclinic Rossby waves, 439
- baroclinic velocity, 215
- barotropic instability, 389
- barotropic mode, 211, 249, 439, 440
- barotropic model, 166
- barotropic Rossby waves, 439
- barotropic velocity, 215
- Bernoulli
  - equation of motion, 100
  - potential, 100
  - theorem, 383
- beta
  - effective, 204
- body forces, xix
- Boussinesq ocean, 65
- budget analysis, ix
- buoyancy oscillations, 255
- capillary
  - waves, 95, 132
- carrier wave, 17
- causal relations, ix
- causality, 235
- centrifugal instability, 327, 330, 331, 333
- centrifugal oscillations, 338
- classical field theory, 109
- compression, 70
- conjugate anti-symmetry, 20
- conjugate symmetry, 20, 130
- conservation of wave crests, 46
- contact forces, xix
- critical height theorem, 422, 423
- critical latitude theorem, 401
- critical levels, 296
- critical reflection, 276
- cyclostrophic balance, 330, 331
- D'Alembert

- wave solution, 238  
decibel scale, 75  
deformation radius, 244, 440  
shallow water, 202, 203, 220, 243  
diagnostic equation, xi  
dimensional analysis, xii  
dispersion, 23  
dispersion relation, 2, 5, 9  
acoustic, 73  
inertia-gravity waves, 220  
inertial waves, 153  
interacting Eady waves, 452  
shallow water gravity wave, 207  
variational principle, 57  
dispersion tensor, 24  
domain of influence, 22  
Doppler shift, 174, 237, 295  
Rossby waves, 174  
Doppler-shifted frequency, 175  
Eady edge waves, 431  
Eady growth rate, 454  
Eady model, 431  
Eady waves, 435, 445  
eddy kinetic energy, 392  
edge waves, 165, 174, 184, 389, 431, 435  
edge waves interactions, 404  
effective beta, 204  
effective buoyancy, 255  
eigenfunctions, 316  
eigenvalues, 316  
eikonal approximation, 43  
eikonal equation, 48  
eikonal wave ansatz, 43, 85, 88  
elliptic PDE  
geostrophic adjustment, 243  
emergent scales, xii  
energetic stability analysis, 327, 343  
energy equipartition  
acoustic, 78  
oscillator, 78  
Euler equation, 69  
Euler identity, 11  
Euler-Lagrange field equation, 52  
evanescence, 162, 311  
evanescent gravity waves, 301  
evanescent waves, 308  
external scales, xii  
extrinsic frequency, 175  
Fjørtoft's theorem, 400  
fluid  
dynamics, xviii  
kinematics, xviii  
Fourier analysis, 2, 16, 117  
free shear layer, 405  
frequency  
extrinsic, 175  
ground-based, 175  
intrinsic, 175  
Fresnel integral, 39  
frontal equations, 362  
Froude number, 208, 236, 297, 299  
Galilean  
transformation, 233, 295  
gauge  
function, 72, 99  
transformation, 100  
generalized momentum density, 53  
geometric optics, 43, 48  
geostrophic  
adjustment, 240, 243  
balance, xi  
geostrophic coordinates, 365  
geostrophic momentum approximation, 362  
global instability, xxi, 324, 330  
gradient Richardson number, 386, 389, 420, 421  
gravest mode, 129  
gravity waves, 198  
non-rotating, 242  
polarized, 223  
speed, 207  
two layers, 209  
Green's function  
for wave equation, 21  
group velocity, 48  
growth rate, 361, 378, 395  
Eady waves, 453  
Hamilton's equations for rays, 48, 51  
Hamilton's principle, xi, 52, 109  
Hamiltonian density, 53  
harmonic function, 99  
Heaviside step function, 241

- Heisenberg uncertainty principle, 30  
Helmholtz  
    equation, 81  
Howard's semi-circle theorem, 422, 424  
hydraulic control, 209, 237  
hydraulic jump, 209, 237  
hydrodynamics, vii  
hyperbolic PDE, 9  
    domain of influence, 22  
inertia-gravity waves, 199, 202, 218, 220, 232  
inertia-vorticity oscillations, 357  
inertial  
    near inertial waves, 158  
    waves dispersion relation, 152, 153  
    waves group velocity, 155  
inertial instability, 327, 330, 342  
inertial period, 244  
initial value problem, 241  
instabilities  
    normal modes, 324, 371  
instability  
    baroclinic, 431, 451  
    barotropic, 389  
    centrifugal, 327, 330, 331  
    critical height, 422, 423  
    critical latitude, 401  
    Fjørtoft's theorem, 400  
    global, xxi, 324, 330  
    inertial, 327, 330, 342  
    interfaces, 371  
    isentropic inertial, 349  
    Kelvin-Helmholtz, 371  
    local, 327, 329  
    necessary conditions, 399  
    phase tilt, 397  
    Rayleigh-Kuo theorem, 399  
    Rayleigh-Taylor, 371, 376  
    stratified shear, 389  
    sufficient conditions, 399  
    symmetric, 327, 330, 349, 353, 356  
    wave, xxi, 324, 330  
internal gravity waves, 249, 291  
    stationary, 293  
internal tides, 293  
intrinsic frequency, 175, 448  
inversion, 240  
irreversible process, vii  
irrotational, 99  
Kelvin waves, 216  
Kelvin's minimum kinetic energy, 107  
Kelvin-Helmholtz instability, 371  
kinematic boundary condition  
    surface waves, 113  
Lagrangian  
    phase averaged, 54  
Lagrangian acoustic wave equation, 69  
Lagrangian density, 52  
lee waves, 307  
local instability, 327, 329  
long waves, 191, 231  
longitudinal waves, 63, 64, 71, 74, 206  
Lorentzian, 306  
Luke's variational principle, 109  
Mach number, 63, 64, 68, 74  
mass conservation  
    Lagrangian, 66  
mean value theorem, 66  
mesoscale ocean, 431  
method of images, 83  
modal stability analysis, 327  
modulation function, 17  
momentum  
    absolute, 341  
    geostrophic, 341  
    potential, 330, 339  
momentum argument, 172  
momentum-based, x  
monochromatic patterns, 11  
monochromatic waves, 11  
mountain drag, 300  
mountain waves, 291, 293, 297  
natural boundary conditions, 109  
near inertial wave, 281  
near inertial waves, 158  
near-inertial waves, 281, 321  
Newton's  
    third law, 259  
non-dimensionalization, xii  
non-dispersive waves, 207  
non-divergent  
    barotropic model, 166  
normal mode method, 324, 371

- ocean submesoscale, 362  
overturning circulation, 358
- parcel method, 255  
parcel stability analysis, 327  
perfect fluid, vii  
phase averaged action, 54  
phase averaged Lagrangian, 54  
phase locked, 405  
phase speed, 47, 394  
phase velocity, 394  
plane wave, 12  
planetary  
— vorticity, xix  
planetary Rossby waves, 166, 202, 431, 439  
planetary waves  
— dispersion relation, 174  
Poincaré waves, 218  
point jet, 186  
Poiseuille flow, 402  
poleward heat transport, 461  
potential flow, 95, 98  
potential momentum, 330, 339  
potential vorticity  
— inertia-gravity waves, 279  
— inversion, 240  
— symmetric instability, 361  
pressure  
— Lagrange multiplier, x  
— non-hydrostatic, 148  
prognostic equation, x  
pseudo-westward phase, 177, 188
- quasi-geostrophy, 203
- rarefaction, 70  
ray equations, 51  
ray theory, 43  
Rayleigh equation, 389, 396, 406  
Rayleigh inflection-point theorem, 399  
Rayleigh waves, 184  
Rayleigh-Kuo equation, 185, 396  
Rayleigh-Kuo theorem, 399  
Rayleigh-Taylor instability, 371, 376, 378  
real fluid, vii  
reduced gravity, 377  
— model, 213  
reduced variational principle, 53
- reflection  
— non-specular, 273  
— specular, 273  
refraction, 49, 127  
resonance, 405  
Richardson number, 386, 389, 420, 421  
— balanced, 454  
Riemann-Legesque lemma, 38  
rigid-body motion, 333  
Rossby height, 444  
Rossby waves, 166, 199  
— baroclinic, 440  
— barotropic, 440  
— dispersion circle, 179, 227  
— dispersion relation, 174  
— group velocity, 178  
— stationary, 177  
— topographic, 442  
rotating hydraulics, 209  
rotation tensor, 69
- scalar  
— potential, 99  
— potential harmonic, 99  
scalar field theory, 97  
scale analysis, xii  
scales  
— emergent, xii  
— external, xii  
secondary circulation, 362  
seiche mode, 129  
self-adjoint, 316  
semi-geostrophy, 327, 362  
separation of variables, 117  
shallow water  
— deformation radius, 203  
— gravity wave dispersion, 220  
— gravity wave speed, 205  
— gravity waves, 205  
— inertia-gravity waves, 218  
— linearized equations, 194  
— quasi-geostrophy, 203  
— wave equation, 197  
— waves, 193  
shear instability, 389  
— stratified, 416  
shear production, 461  
sign-function, 241

- Snell's law, 127  
soliton, 125  
sound pressure level, 75  
sound speed, 68, 70  
sound waves, 63  
specular reflection, 183  
Squire's theorem, 391, 417, 464  
St. Andrew's cross, 269  
stability analysis  
    energetic, 327, 334, 343  
    modal, 327, 356  
    parcel, 327, 336, 345, 350  
stationary phase, 37  
stationary waves, 235, 297  
Stokes  
    correction, 136  
    drift, 122, 136  
    drift and surfing, 137  
    drift for surface prototypical waves, 139  
Stokes drift, 2  
strain rate tensor, 69  
stratified shear instability, 416  
stress-energy-momentum, 53  
    acoustic, 86  
stretching, 151  
Sturm-Liouville, 439  
Sturm-Liouville problem, 316  
sub-inertial wave, 175, 202  
super-inertial waves, 201  
superposition principle, 2, 12, 16, 74, 117  
surface  
    gravity waves, 95, 112  
surface gravity waves, 10  
    dispersion relation, 119  
    kinematic boundary condition, 113  
    longwave limit, 125  
    shortwave limit, 125  
    wave breaking, 126  
surface quasi-geostrophy, 318  
surface tension, 132, 374  
Sverdrup balance, xi  
symmetric instability, 327, 330, 349, 353, 356  
synoptic scale atmosphere, 431  
Taylor-Goldstein, 389, 428  
Taylor-Goldstein equation, 418  
Taylor-Proudman effect, 147, 160  
test  
    fluid element, 255, 266  
thermal wind shear production, 461  
tilting, 151  
time  
    tendency, x  
topographic beta, 445  
topographic form stress  
    gravity waves, 300  
topographic Rossby waves, 199, 202, 431, 442  
topographic waves, 199  
topography forcing, 292  
tracer  
    mechanics, xx  
transverse waves, 71, 152  
trapped gravity waves, 301  
traveling wave, 12  
turning level, 312  
uncertainty relation, 5, 29, 30, 36  
variation, 108  
    action, 109  
velocity potential, 72  
vertical  
    stiffness, 147  
virial theorem, 78  
vortex  
    sheet, 383  
vortical mode, 165  
vortical waves, 172  
vorticity  
    acoustic waves, 72  
    stretching, 151  
    tilting, 151  
wave  
    acoustic, 63, 71  
    action, 43, 54, 59, 91  
    amplitude, 12  
    angular frequency, 12  
    anisotropic, 177  
    baroclinic mode, 211, 249  
    barotropic mode, 211, 249  
    barotropic vorticity, 165  
    capillary, 95, 132  
    carrier, 17

- deep water waves, 120  
dispersion relation, 2, 5, 9, 17, 73  
dispersive packet, 34  
Eady, 435, 445  
edge, 165, 174, 184, 389, 404, 431, 435  
eikonal ansatz, 43  
energy, 43  
equation, 9  
evanescent, 311  
function, 11  
gravity, 202  
gravity wave critical reflection, 274, 276  
group velocity, 17  
guide, 312  
harmonic, 7, 8  
inertia-gravity, 281  
inertia-gravity dispersion, 281  
inertial, 147  
inertial polarization, 156  
inertial radial, 158  
interfacial, 95  
interference, 2  
internal gravity, 249  
internal gravity energetics, 271  
internal gravity polarization, 260  
internal gravity reflection, 273  
internal inertia-gravity, 278  
Kelvin, 216  
kinematics, 5  
length, 13  
longitudinal, 63, 64, 206  
maker, 17  
mathematics, 5  
mechanics, 2  
monochromatic, 7, 8, 11  
narrow band packet, 29  
near inertial, 281  
near-inertial, 281, 321  
non-dispersive, 74, 207  
non-dispersive packet, 32  
nonlinear, 125  
number, 5, 13  
packet, 16, 17  
period, 13  
phase, 5, 12, 44  
phase distance, 15  
phase speed, 14, 47  
phase velocity, 9, 14, 47  
plane, 12  
planetary Rossby, 431, 439  
polarization relation, 223  
pseudo-west phase, 177, 188  
ray, 48, 49  
reduced wavelength, 15  
reduced wavenumber, 15  
refraction, 127  
resonance, 324, 389, 405, 431  
Rossby, 165, 202  
Rossby phase velocity, 176  
Rossby reflection, 181  
Rossby shallow water, 226  
seiche, 129  
shallow water, 193  
shallow water gravity, 198  
shallow water inertia-gravity, 199  
shallow water Rossby, 199  
shallow water topographic, 199  
shallow water waves, 120  
sound, 63  
spatial frequency, 15  
standard phase, 13  
standing, 7, 8, 128  
standing packet, 27  
stationary, 7, 8, 293  
stationary phase, 37  
sub-inertial, 153, 202, 220  
super-inertial, 201, 220  
superposition, 16  
surface, 95  
topographic Rossby, 431, 442  
trains, 17  
transverse, 167  
traveling, 7, 8, 12, 117  
turning level, 312  
uncertainty, 29  
uncertainty relation, 30, 36  
vector, 5, 12  
vortical and divergent motions, 279  
vorticity, 165  
wave packet, 29, 130  
wavenumber, 12, 15  
wavevector, 12  
WKB ansatz, 43  
wave equation

- domain of influence, 22  
Lagrangian, 69  
wave instability, [xxi](#), 324, 330  
wave packet  
    surface gravity waves, 130  
waves  
    Rayleigh, 184  
    stationary, 162, 235, 297  
wedge of instability, 463  
symmetric, 330, 351, 353  
weir, 208  
Whitham's variational principle, 43, 53, 85, 92  
WKB approximation, 43, 309  
    gravity waves, 307  
WKB stretching, 314, 318  
WKB wave ansatz, 43, 309  
WKBJ asymptotic method, 85  
Young-Laplace formula, 132, 374